

Research article

Open Access

# Identification of new, well-populated amino-acid sidechain rotamers involving hydroxyl-hydrogen atoms and sulfhydryl-hydrogen atoms

Bosco K Ho\* and David A Agard

Address: Department of Biochemistry, Howard Hughes Medical Institute, University of California, San Francisco, San Francisco CA 94158-2517

Email: Bosco K Ho\* - apposite@gmail.com; David A Agard - agard@msg.ucsf.edu

\* Corresponding author

Published: 7 October 2008

Received: 16 April 2008

BMC Structural Biology 2008, 8:41 doi:10.1186/1472-6807-8-41

Accepted: 7 October 2008

This article is available from: <http://www.biomedcentral.com/1472-6807/8/41>

© 2008 Ho and Agard; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

**Background:** An important element in homology modeling is the use of rotamers to parameterize the sidechain conformation. Despite the many libraries of sidechain rotamers that have been developed, a number of rotamers have been overlooked, due to the fact that they involve hydrogen atoms.

**Results:** We identify new, well-populated rotamers that involve the hydroxyl-hydrogen atoms of Ser, Thr and Tyr, and the sulfhydryl-hydrogen atom of Cys, using high-resolution crystal structures (<1.2 Å). Although there were refinement artifacts in these structures, comparison with the electron-density maps allowed the placement of hydrogen atoms involved in hydrogen bonds. The  $\chi_2$  rotamers in Ser, Thr and Cys are consistent with tetrahedral bonding, while the  $\chi_3$  rotamers in Tyr are consistent with trigonal-planar bonding. Similar rotamers are found in hydrogen atoms that were computationally placed with the Reduce program from the Richardson lab.

**Conclusion:** Knowledge of these new rotamers will improve the evaluation of hydrogen-bonding networks in protein structures.

## Background

One important conformational parameter of a protein structure is the sidechain  $\chi$  torsion angle [1]. In crystal structures, these torsion angles were found to be rotameric [2]: they cluster around specific values, values that can be explained in terms of relatively simple stereo-chemical considerations [3,4]. Consequently, libraries of sidechain rotamers have been compiled [5,6]. These libraries have proven useful in parameterizing sidechain conformations for homology modeling [7], monte-carlo simulations [8], and protein design [9]. Rotamer libraries are also used to build and verify crystallographic models [10]. Although the sidechain rotamers have been extensively studied,

there remain a number of rotamers involving hydrogen atoms that have been overlooked.

Due to the difficulty in placing hydrogen atoms in protein electron density maps, it has long been customary to omit hydrogen atoms in reporting the crystal structure of a protein. However, Richardson and co-workers showed that positions of hydrogen atoms in high-resolution crystal structures can be confidently projected from the topology of the heavy atoms [11]. The projected hydrogen atoms, in most cases, form better van-der-Waals contacts with the neighboring atoms than do the heavy atoms themselves. The heavy atoms accommodate the packing of the hydro-

gen atoms, even though the hydrogen atoms cannot be seen in the crystal structure.

One of the reasons why the projection of hydrogen atoms works so well is that the positions of most of the sidechain hydrogen atoms are stereo-chemically restricted. For example, the H $\beta$  atom of Val can only adopt one tetrahedral-bonding position off the C $\beta$  position given that 3 other C atoms are also bound to C $\beta$ . For other hydrogen atoms, symmetry between equivalent methyl-hydrogen atoms results in similar restrictions. For example, in Val, the three equivalent H $\gamma$ 1 atoms bound to C $\gamma$ 1 saturate the three available tetrahedral-bonding positions at C $\gamma$ 1. Nevertheless, there exist four types of sidechain hydrogen atoms in which there is ambiguity in projecting their positions. For instance, in Ser, there are three different ways to place the H $\gamma$  atom onto the tetrahedral-bonding positions of the O $\gamma$  atom. This freedom is also found in the hydroxyl-hydrogen atoms of Thr and Tyr, and in the sulfhydryl-hydrogen in Cys.

As the positions of most of the sidechain hydrogen atoms are so restricted, little attention has been paid to their conformation in crystal structures. However, given the growing number of structures containing hydrogen atoms in the data bank, it has become practical to revisit the question of sidechain hydrogen rotamers for the four classes of ambiguous sidechain hydrogen atoms. The positions of these sidechain hydrogen atoms should be parameterized by  $\chi$  torsion angles, and we would like to know if these angles display rotameric preferences. Here, we study the distributions of these  $\chi$  torsion angles in three data-sets: (1) high-resolution X-ray structures that contains explicit

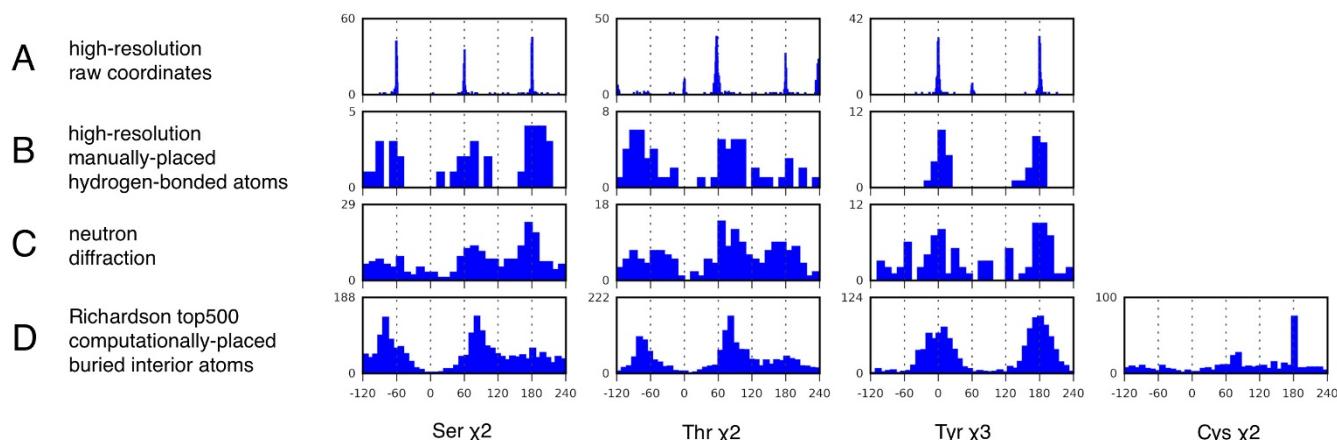
hydrogen atoms, (2) neutron diffraction structures and (3) structures with computationally-placed hydrogen positions.

## Results

### The rotamers in high-resolution X-ray structures with hydrogen atoms

For the first part of the analysis, we use high-resolution X-ray structures that have explicit hydrogen atoms in the Ser, Thr, Tyr and Cys residues. As such, the data-set consists of structures found in the RCSB.ORG website [12] with resolution < 1.2 Å, where hydrogen atoms are found in the structure. The hydrogen atoms are filtered for residues with no alternate conformations and where the neighboring heavy atom has a B-factor < 40. The structures were further selected depending on the availability of the electron density maps in the Electron Density Server [13]. This results in 27 structures: 1AHQ, 1DY5, 1JM1, 1M40, 1RW1, 2AXW, 2FDN, 1BXO, 1EUW, 1KQP, 1MUW, 1TT8, 2BF9, 2FFY, 1C75, 1F94, 1L9L, 1O7L, 1UICS, 2CAL, 2H5C, 1CEX, 1GOV, 1LS1, 1RB9, 2AWK, 2ERL.

We first generate the  $\chi$ -angle distributions of the hydrogen atom positions directly from the coordinates reported in the high-resolution structures (Figure 1A). Hydrogen atoms were found for 232 Ser, 356 Thr and 187 Tyr. The distributions show exceptionally sharp peaks, which correspond to rotameric preferences. In Ser, the hydrogen atom position is defined by  $\chi_2 = \text{Ca}-\text{C}\beta-\text{O}\gamma-\text{H}\gamma$ , where the  $\chi_2$  distribution shows three peaks, corresponding to tetrahedral bonding on the O $\gamma$  atom. In Thr, the hydrogen atom position is defined by  $\chi = \text{Ca}-\text{C}\beta-\text{O}\gamma_1-\text{H}\gamma_1$ , where the distribution shows 4 peaks. Two of the peaks are



**Figure 1**  
**Distributions of the hydrogen-based  $\chi$  angles of Ser, Thr, Tyr, and Cys.** (A) The distributions calculated directly from high-resolution ( $< 1.2 \text{ \AA}$ ) structures deposited in the PDB. (B) The distributions of the same structures where the hydrogen atoms have been manually placed to fit the electron-density of hydrogen-bonded hydrogen atoms. (C) The distributions from the structures determined by neutron diffraction. (D) The distributions from the hydrogen atoms that have been computationally-placed with Reduce from the top500 database of non-redundant structures from the Richardson lab.

expected from tetrahedral bonding but the peak  $\chi_2 = -60^\circ$  is missing. Instead there are two other peaks at  $\chi_2 = -120^\circ$  and  $\chi_2 = 0^\circ$ . In Tyr, the hydrogen position is defined by  $\chi_3 = C\delta 1-C\epsilon 1-O\zeta-H\zeta$  where the distribution has three peaks, of which 2 are expected from trigonal-planar bonding and there is an unexpected peak at  $\chi_3 = 60^\circ$ .

In many cases with high-resolution structures, the hydrogen atoms are not actually refined, but are instead placed automatically. As such their positions can be unreliable. To check this, we carefully inspected the  $2F_0-F_c$  maps for the hydrogen atoms of the 27 structures in this data-set and discovered that no electron density was observable for any of these hydrogen atoms at normal contour levels of the maps (Figure 2). Therefore, the reported positions of these hydrogen atoms must be an artifact of automatic placement. The sharp peaks in Figure 1 arise from pre-defined restrictions on the  $\chi$  angles, and not intrinsic structural propensities. Of the 27 structures, there was one notable exception in 2H5C which contains a large number of rotamers found far from the pre-defined rotamers ( $> 10^\circ$ ) where the hydrogen atom positions have been carefully refined against the chemistry of nearby atoms.

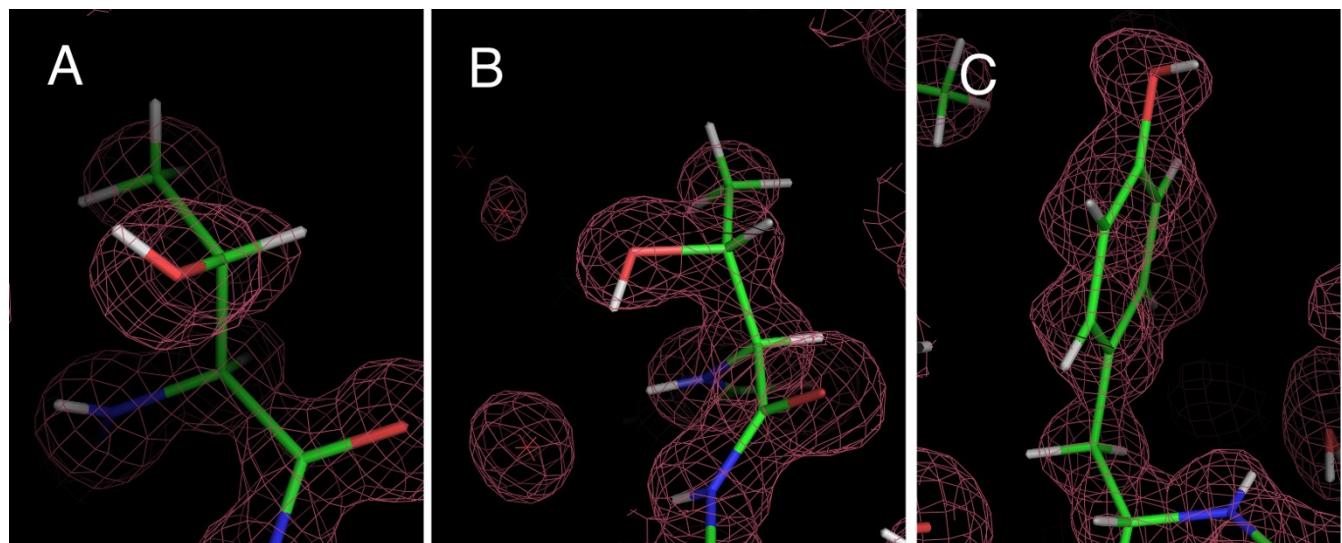
Given the absence of agreement with the other distributions (below), and incompatibilities with tetrahedral-bonding and trigonal-planar bonding, it must be concluded that in Ser, the two rotamers at  $\chi_2 = -120^\circ$  (Figure 2A) and  $\chi_2 = 0^\circ$  (Figure 2B) are artifacts especially since these rotamers pack the hydrogen atom against a carbon

atom. Similarly, in Tyr, the off-planar rotamer at  $\chi_3 = 60^\circ$  (Figure 2C) is an artifact.

#### **Manually-placed rotamers in X-ray structures using hydrogen-bonds**

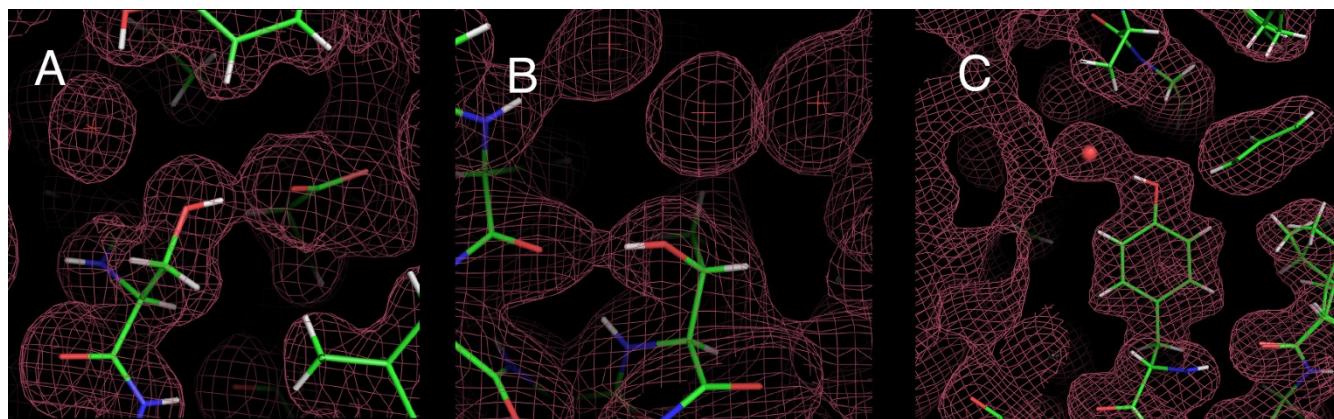
In general, the  $2F_0-F_c$  electron density for the hydroxyl-hydrogen atoms is too weak to be observed at contour levels that are typically used. However, in these high-resolution structures, even at a contour level of  $\sigma = 0$ , there are regions of the  $2F_0-F_c$  map that result in well-defined contours of the molecule. Given that heavy atoms can be clearly distinguished in these maps, in regions where the map is well defined, the position of hydrogen atoms can be determined from density between the donor and acceptor atoms of a hydrogen bond (Figure 3). Such hydrogen-bonded hydrogen atoms were identified for 37 Ser, 39 Tyr and 60 Thr residues, and the hydrogen atom positions were manually placed [see Additional File 1].

The  $\chi$ -angle distributions and rotamers were calculated from these manually-placed hydrogen atoms (Figure 1B and Table 1; also see Additional File 2). The Ser  $\chi_2$  distribution splits into three rotamers at  $-79^\circ$ ,  $69^\circ$  and  $189^\circ$  (Figure 1), corresponding to the three angles of tetrahedral-bonding of the Hy atom to the O $\gamma$  atom. The  $189^\circ$  rotamer is favored, presumably because it positions the Hy atom furthest away from the C $\beta$  atom. The Thr  $\chi_2$  distributions contain two well-populated rotamers at  $\chi_2 = -74^\circ$  and  $80^\circ$ . Surprisingly, there is no clear peak at the tetrahedral-bonding position that places the hydrogen atom away from the bulk of the sidechain. The Tyr  $\chi_3$  distribu-



**Figure 2**

**Examples of the poor fits of the hydrogen positions to the  $2F_0-F_c$  electron-density map (at  $0.5\sigma$  contour level) in high-resolution X-ray structures.** These conformations are refinement artifacts (A) Thr-A41 [1C75] at  $\chi_2 = -120^\circ$ , (B) Thr-A104 [1IM1] at  $\chi_2 = 0^\circ$ ; and (C) Tyr-A3 [1RW1] at  $\chi_3 = 60^\circ$ .

**Figure 3**

**Examples of hydrogen positions that were manually-placed to the  $2F_0-F_c$  electron-density of a hydrogen-bond (at the  $0\sigma$  contour level).** (A) Ser-A205 [2AWK] with a hydrogen bond to a Glu. (B) Thr-A30 [ILSI] with a hydrogen bond to a backbone carbonyl O atom. (C) Tyr-149 [LCFX] with a hydrogen bond to water.

tion splits into two rotamers at  $5^\circ$  and  $174^\circ$ , consistent with trigonal-planar bonding (Figure 1). However, given that the aromatic ring is symmetric, these two rotamers are equivalent.

#### Comparison with neutron structures

We can compare the high-resolution manually-placed distributions to distributions from neutron structures. While the neutron structures are generally of a lower resolution,

**Table 1: Rotamers from hydrogen-bonded hydrogen atoms fitted from electron density in high-resolution structures.**

Type of $\chi$ Angle	Rotamer	Population
Ser $\chi_2$ : 37 counts	$-79 \pm 18^\circ$	27%
	$69 \pm 24^\circ$	30%
	$189 \pm 16^\circ$	43%
Thr $\chi_2$ : 60 counts	$-74 \pm 26^\circ$	47%
	$80 \pm 18^\circ$	35%
	$176 \pm 35^\circ$	18%
Tyr $\chi_3$ : 39 counts	$5 \pm 9^\circ$	49%
	$174 \pm 12^\circ$	51%

hydrogen is a strong and negative neutron scatterer, which should allow reasonably accurate positioning of the hydrogen atoms. We found 16 neutron structures in the PDB with hydrogen atoms, containing a total of 215 Ser, 154 Thr and 85 Tyr residues. There were no Cys hydrogen atoms. In keeping with the lower resolution, the distributions in the neutron structures (Figure 1C) are more diffuse than those of the high-resolution X-ray structures (Figure 1B). In the neutron structures, the Ser distribution has a peak near  $180^\circ$ . In Thr, the distribution has one delineated peak at  $60^\circ$ . In Tyr, the distribution is diffuse with weak peaks at  $0^\circ$  and  $180^\circ$ . Qualitatively, there is agreement with distributions derived from the manually-placed hydrogen atoms of the high-resolution crystal structures (Figure 1B).

#### Comparison with computationally-placed hydrogen atoms

We can also compare the high-resolution manually-placed distributions to distributions derived from computationally-placed hydrogen atoms. This allows us to evaluate algorithms that project hydrogen positions from the coordinates of heavy atoms. We use a representative non-homologous set of high resolution structures ( $< 1.8 \text{ \AA}$ ), provided by the Richardson lab [6], where missing hydrogen atoms have been computationally-placed using the program Reduce [14]. These hydrogen atoms were computationally placed by optimizing hydrogen bonds and steric contacts with neighboring atoms.

In the structures with Reduce-placed hydrogen atoms, we found an artifact in the surface hydrogen atoms. As there are few neighboring contacts on the surface to help determine the position of hydrogen atoms, many surface hydrogen atoms remain at the default value, resulting in a pronounced peak at  $180^\circ$  (data not shown). This peak

can be removed if we eliminate surface residues. Furthermore, as the Reduce algorithm uses steric contacts to optimize hydrogen positions, we need to use well-packed hydrogen atoms. Consequently, we only consider buried interior hydrogen atoms, defined as atoms with > 8 neighboring atoms, where a neighboring atom is defined if it is within 3.5 Å of another atom. We also filter out residues with alternate conformations and atoms where the B-factor > 40. There were hydrogen atoms from 5768 Ser, 5932 Thr, 3645 Tyr and 660 Cys residues in 480 structures. The large size of this data set gives the most reliable statistics.

The  $\chi$ -angle distributions and rotamers were calculated from the Reduce-placed hydrogen atoms (Figure 1D and Table 2; also see Additional file 2). In the  $\chi_2$  distribution of Ser, there are only two peaks, as opposed to three peaks in the high-resolution crystal structures (Figure 1). Both the Thr and Tyr distributions qualitatively reproduced the manually-placed hydrogen atom distributions. Apart from the missing  $\chi_2 = 189^\circ$  rotamer in Ser, the Reduce-placed rotamers show remarkable agreement with the manually-placed rotamers. One possible reason for this

**Table 2: Rotamers from computationally-placed hydrogen atoms in the Richardson set of structures.**

Type of $\chi$ Angle	Rotamer	Population
Ser $\chi_2$ : 5768 counts	$-75 \pm 25^\circ$	30%
	$81 \pm 22^\circ$	30%
	$180 \pm 26^\circ$	40%
Thr $\chi_2$ : 5932 counts	$-70 \pm 24^\circ$	25%
	$80 \pm 20^\circ$	37%
	$178 \pm 24^\circ$	38%
Tyr $\chi_3$ : 3645 counts	$2 \pm 27^\circ$	41%
	$180 \pm 23^\circ$	59%
Cys $\chi_2$ : 660 counts	$-66 \pm 32^\circ$	20%
	$72 \pm 27^\circ$	38%
	$179 \pm 24^\circ$	42%

agreement is that Reduce algorithm makes extensive use hydrogen-bonding optimization, which reflects the choice of using hydrogen-bonding density to manually place the hydrogen atoms in the high-resolution structures.

Given the robust performance of Reduce, we investigated the position of sulphydryl-hydrogen atoms in the set of structures provided by the Richardson lab. The hydrogen atoms are considered only if there are no other Cys residues within 4.5 Å of the SG atom of the Cys in order to avoid disulfide-bonded Cystines. The position of the sulfhydryl-hydrogen in Cys is determined by the  $\chi_2 = \text{Ca}-\text{C}\beta-\text{S}\gamma-\text{Hy}$  angle. The Cys  $\chi_2$  distribution show a dominant rotamer at  $\chi_2 = 181^\circ$  (Figure 1), which places the hydrogen atom furthest away from the backbone, in between the two  $\text{H}\beta$  atoms.

## Methods

Data was collected from PDB structures using in-house Python scripts. The distributions in Figure 1 were binned using a rough guideline of ~4 bins for a peak. For the X-ray structures, 360 bins were used for the distribution from the raw coordinates, and 30 bins for the distributions from the manually-placed hydrogen atoms. For neutron structures, we used 30 bins. For the computationally-placed distributions, we used 50 bins. To calculate the means and distributions for the rotamers, we divided up the  $\chi$  range into 3 equal partitions in Ser, Thr and Cys, and 2 partitions for Tyr.

## Conclusion

Based on experimental data, we find that certain sidechain torsion angles involving hydrogen atoms have strongly preferred orientations and should thus be considered rotameric. Although there were serious artifacts found in the reported coordinates of high-resolution X-ray structures, reliable hydrogen atom positions could be directly derived from the electron-density maps of hydrogen-bonded hydrogen atoms. The  $\chi$ -angle distributions of these hydrogen-bonded hydrogen atoms match the distribution of hydrogen atoms that were computationally placed by the program Reduce [14].

## Authors' contributions

BKH carried out the study and drafted the manuscript. DAA provided support and guidance. This work was supported by the Howard Hughes Medical Institute.

## Additional material

### Additional file 1

*Manually-fitted hydrogen atoms in high resolution structures. In these high-resolution X-ray structures, the hydroxyl hydrogen atoms were removed if there is no corresponding electron density but if there is sufficient electron density due to hydrogen-bonding, the hydrogen atom position were fitted manually.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1472-6807-8-41-S1.zip>]

### Additional file 2

*List of hydroxyl-hydrogen  $\chi$  rotamers in high-resolution PDB files. List of the residues used to calculate the hydroxyl-hydrogen  $\chi$ -angle rotamers in Ser, Thr and Tyr. Data-sets provided for both the manually-placed hydrogen atoms in the high-resolution structures, and the Reduce-placed hydrogen atoms in the Richardson data-set.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1472-6807-8-41-S2.zip>]

## References

1. Ramachandran GN, Sasisekharan V: **Conformation of polypeptides and proteins.** *Adv Protein Chem* 1968, **23**:283-438.
2. Janin J, Wodak S: **Conformation of amino acid side-chains in proteins.** *J Mol Biol* 1978, **125**(3):357-386.
3. 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**(2):295-310.
4. Dunbrack RL Jr, Karplus M: **Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains.** *Nat Struct Biol* 1994, **1**(5):334-340.
5. 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**(4):775-791.
6. Lovell SC, Word JM, Richardson JS, Richardson DC: **The penultimate rotamer library.** *Proteins* 2000, **40**(3):389-408.
7. Bower MJ, Cohen FE, Dunbrack RL Jr: **Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool.** *J Mol Biol* 1997, **267**(5):1268-1282.
8. Lasters I, De Maeyer M, Desmet J: **Enhanced dead-end elimination in the search for the global minimum energy conformation of a collection of protein side chains.** *Protein Eng* 1995, **8**(8):815-822.
9. Dahiyat BI, Mayo SL: **De novo protein design: fully automated sequence selection.** *Science* 1997, **278**(5335):82-87.
10. Laskowski RA, Moss DS, Thornton JM: **Main-chain bond lengths and bond angles in protein structures.** *J Mol Biol* 1993, **231**(4):1049-1067.
11. Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC: **Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms.** *J Mol Biol* 1999, **285**(4):1711-1733.
12. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nucleic Acids Res* 2000, **28**:235-242.
13. Kleywegt GJ, Harris MR, Zou JY, Taylor TC, Wahlby A, Jones TA: **The Uppsala Electron-Density Server.** *Acta Crystallogr D* 2004, **60**:2240-2249.
14. Word JM, Lovell SC, Richardson JS, Richardson DC: **Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation.** *J Mol Biol* 1999, **285**(4):1735-1747.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

