# Supplementary Information

# Fast and accurate predictions of protein NMR chemical shifts from interatomic distances

Kai J. Kohlhoff, Paul Robustelli, Andrea Cavalli, Xavier Salvatella and Michele Vendruscolo

*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK*

## Hydrogen bonding term

The effects of hydrogen bonds originating on backbone $H_N$ and O atoms are implemented in CamShift by assuming a correlation between chemical shift and hydrogen bond energy using results that were recently reported by David Baker and coworkers[1]. They determined the orientation and distance dependence of hydrogen bond dimerization energies using both quantum mechanical calculations and structural analysis for hydrogen bonds between side chain and backbone atoms[1].

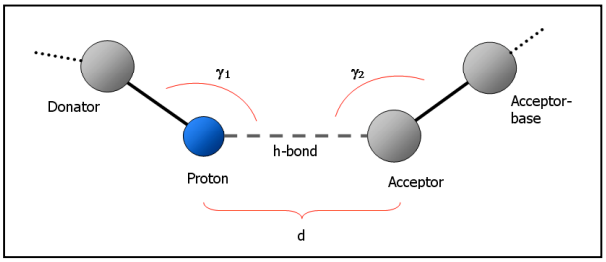


Figure S1. Geometric parameters used to model hydrogen bonding contributions to chemical shifts; hydrogen atoms (blue), heavy atoms (gray). The angles characterising the hydrogen bond geometry are denoted as $\gamma_1$ and $\gamma_2$ (corresponding to $\theta$ and $\psi$ in Ref[1]), and the hydrogen bond length is indicated as *d*.

The parameters that are used in the calculation of hydrogen bond energies are shown in Fig. S1. In the original study, data were calculated separately for different types of secondary structure elements[1]. For our purposes, we considered average values and fitted a single curve. We thus avoided having to subdivide the data used for training according to different secondary structure elements; a benefit in this procedure is that the chemical shift term remains differentiable and thus suitable for being used in restrained molecular dynamics simulations. Averages extracted from the data were fitted with a potential of the Lennard-Jones form

$$E(x) = p_1\left(\left(\frac{p_2}{x} + p_4\right)^r - \left(\frac{p_3}{x} + p_5\right)^s\right) \qquad \text{(S1)}$$

where $E(x)$ denotes the energy corresponding to the parameter $x$ (an angle or distance), and $p_1$ through $p_5$, $r$, and $s$ are parameters determined by a Marquardt-Levenberg nonlinear least-squares fit.

| | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | r | s |
|---|---|---|---|---|---|---|---|
| $\gamma_1$ | 15.0486 | -1.79469 | -1.79469 | 1.49398 | 1.49398 | 15.5 | 2.5 |
| $\gamma_2$ | 12.1514 | -1.12258 | -1.12258 | 1.30953 | 1.30953 | 15.5 | 3 |
| D | -0.403531 | 13.9148 | -0.421725 | -3.08016 | 0.987251 | 3 | -16.5 |

Table S1. Values of the three parameters describing the backbone-backbone hydrogen bonding term in Eq. (S1), which are averaged over different types of secondary structure elements.

The chemical shift term for each of the three parameters is therefore

$$\delta_{hbond,x} = \alpha_{hbond,x}\left(p_{1,x}\left(\left(\frac{p_{2,x}}{x}+p_{4,x}\right)^{r}-\left(\frac{p_{3,x}}{x}+p_{5,x}\right)^{s}\right)\right) \quad \text{(S2)}$$

where $x$ indicates $\gamma_1$, $\gamma_2$, or $d$. The three different contributions are included in the CamShift equation with individual coefficients. The full chemical shift term for hydrogen bonds is therefore given by

$$\delta_{hbond} = \delta_{hbond,\gamma_1} + \delta_{hbond,\gamma_2} + \delta_{hbond,d} \quad \text{(S3)}$$

**Ring current term**

CamShift implements the classic point-dipole method[2]

$$\delta_{ring} = \sum_i \alpha_{ring,i} \sum_{R(i)}\left[\frac{1-3\cos^2(\theta)}{r^3}\right] \quad \text{(S4)}$$

where θ is the angle between the vector normal to the ring plane and the vector connecting the ring centre and the target atom, $r$ is the distance between that target atom and the ring centre (Fig. S2), $i$ runs over the five different ring types (Phe, Tyr, His, $Trp_1$, and $Trp_2$, indicated by the corresponding amino acid three-letter code, with the latter two denoting the two different rings of tryptophan), and $R(i)$ the set of all rings in a protein that are of type $i$.

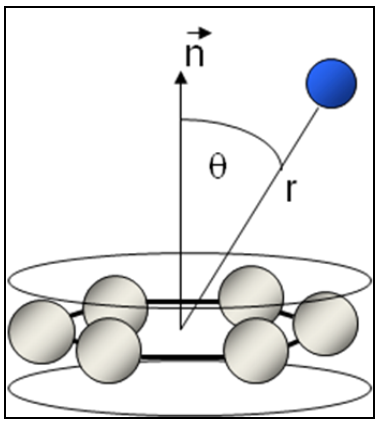


Figure S2: Illustration of the parameters used for the ring current term (Eq. S4). Atoms in the ring are shown as gray spheres and the query atom as a blue sphere at position $\boldsymbol{r}$; $\vec{n}$ is the normal vector to the ring plane and θ the angle between $\boldsymbol{r}$ and $\vec{n}$.

## Dihedral angle term

Dihedral angle contributions are fitted with a periodic function of the form

$$f(\theta) = p_1 \cos\left[3(\theta + p_4)\right] + p_2 \cos(\theta + p_5) + p_3 \qquad \text{(S5)}$$

where θ indicates one among the three backbone dihedral angles ϕ, ψ, and $\chi_1$. The value of the parameters are given in Table S2.

| | | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ |
|---|---|---|---|---|---|---|
| Hα | ϕ | 0.3 | 0.4 | 0.0 | 0.1 | 3.0 |
| | ψ | -0.170287 | 0.278618 | 0.0609298 | 2.48692 | -3.07025 |
| | $\chi_1$ | -0.0244394 | -0.137248 | 0.0560274 | -5.42115 | 2.27543 |
| Cα | ϕ | 1.5 | -3.0 | 0.0 | 2.45 | -2.7 |
| | ψ | -0.4 | -1.5 | 0.7 | 4.8 | -8.4 |
| | $\chi_1$ | 0.1 | 0.9 | 0.6 | -6.0 | 2.2 |
| $H_N$ | ϕ | -0.4 | -0.3 | 0.2 | 3.6 | -3.4 |
| | ψ | -0.104686 | 0.174676 | 0.0639679 | 3.51401 | -2.82811 |
| | $\chi_1$ | 0.0343048 | 0.0448763 | 0.0406634 | 3.42026 | 4.15534 |
| N | ϕ | -2.1 | -1.2 | 0.2 | 3.2 | -2.7 |
| | ψ | -1.4 | 2.2 | 0.3 | 4.2 | -2.5 |
| | $\chi_1$ | 0.5 | 1.2 | -0.1 | 2.2 | 2.9 |
| C' | ϕ | 0.9 | -2.9 | -1.2 | 2.4 | -2.5 |
| | ψ | -0.7 | -0.9 | 0.0 | 5.1 | -8.3 |
| | $\chi_1$ | 0.1 | -0.5 | -0.1 | 19.9 | 11 |
| Cβ | ϕ | 1.0 | 1.8 | 0.0 | 0.8 | 2.5 |
| | ψ | 0.5 | 2.0 | 0.8 | -0.6 | 2.6 |
| | $\chi_1$ | -0.4 | 0.5 | 1.1 | 8.9 | 11 |

Table S2. Parameters used for the dihedral angle term (Eq. S5)

## Fitting of the CamShift parameters

Functions used to describe the hydrogen bonding (Eq. S1) were fitted with a Marquardt-Levenberg nonlinear least-squares fit, while the dihedral angle (Eq. S5) contributions were fitted using a combination of steepest descent and conjugate gradient minimization. The CamShift equation, which is a linear combination of these functions and of the distance-based terms described by Eq. 1 in the main text, was optimized by using a least-squares fit to find the set of coefficients that minimizes the deviations between predicted and experimental chemical shift values.

Distances along the backbone, which represent the dominant contribution to the chemical shifts, and all other contributions that are common among amino acids were fitted across the complete trainings data set (without glycine and proline residues) in order to minimize overfitting effects, while distances to side-chain atoms serve as residue-specific additional degrees of freedom to reduce the remaining differences between experimental and predicted chemical shifts. Since the fit includes many different side-chain atoms, the number of coefficients reported in Table S6 is significantly reduced on a per-residue basis. For glycine and proline residues we present separate sets of coefficients, since we found that because of their particular structural properties these two types of residues both disturb and suffer from coefficients optimized across all residue types.

## Exclusion of distances with low variance-to-mean ratio

The inclusion of near-constant distances (in particular between covalently bonded atoms) can cause instabilities in the fitting of the parameters in Eq. 1 in the main text. An iterative procedure was therefore used to exclude distances with limited variance in order to avoid such a behaviour. We first ranked all distances in the dataset of structures used for the fitting of the parameters according to their variance-to-mean ratios. We then started to iterate the fitting procedure in the following way: at each iteration, the top (i.e. the less variable) 10 distances were removed from the list and the chemical shift RMSD for a test set was calculated. The iterative elimination was repeated until a significant deterioration of the test set RMSD was observed. This procedure resulted in the exclusion of a number of distances ranging from 40 (for N atoms) to 100 (for Hα atoms), and provided a stable predictor while keeping a high accuracy. Excluded distances are listed with a coefficient 0 in Table S6.

| CamShift | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| PDB | BMRB | Chain | Cα | Cβ | Hα | $H_N$ | N | C' |
| 1iar | 4094 | A | 1.01 | 1.11 | 0.21 | 0.39 | 2.40 | 0.87 |
| 1bdo | 4425 | A | 1.31 | 1.30 | 0.24 | 0.45 | 2.49 | |
| 1kmi | 4472 | Y | 1.10 | 1.21 | 0.23 | 0.62 | 2.87 | |
| 1kqr | 5275 | A | 1.18 | 1.28 | 0.32 | 0.70 | 3.23 | 1.10 |
| 1mms | 5513 | A | 1.46 | 1.26 | 0.30 | 0.64 | 3.13 | 1.18 |
| 1fh9 | 7264 | A | 1.21 | 1.37 | | 0.60 | 3.07 | 1.28 |
| 2ihb | 7272 | B | 1.19 | 0.81 | 0.21 | 0.49 | 2.28 | 1.10 |
| | | | | | | | | |
| SPARTA | | | | | | | | |
| PDB | BMRB | Chain | Cα | Cβ | Hα | $H_N$ | N | C' |
| 1iar | 4094 | A | 0.86 | 0.96 | 0.21 | 0.42 | 2.37 | 0.77 |
| 1bdo | 4425 | A | 0.99 | 1.08 | 0.25 | 0.41 | 2.48 | |
| 1kmi | 4472 | Y | 0.97 | 1.02 | 0.26 | 0.63 | 2.57 | |
| 1kqr | 5275 | A | 0.99 | 1.11 | 0.34 | 0.56 | 3.01 | 1.11 |
| 1mms | 5513 | A | 1.36 | 1.42 | 0.31 | 0.52 | 2.94 | 1.01 |
| 1fh9 | 7264 | A | 1.04 | 1.30 | | 0.63 | 2.75 | 1.25 |
| 2ihb | 7272 | B | 1.19 | 0.78 | 0.22 | 0.54 | 2.34 | 1.05 |
| | | | | | | | | |
| CamShift distances | | | | | | | | |
| PDB | BMRB | Chain | Cα | Cβ | Hα | $H_N$ | N | C' |
| 1iar | 4094 | A | 1.04 | 1.02 | 0.24 | 0.40 | 2.47 | 0.88 |
| 1bdo | 4425 | A | 1.38 | 1.30 | 0.25 | 0.50 | 2.54 | |
| 1kmi | 4472 | Y | 1.12 | 1.28 | 0.26 | 0.66 | 2.84 | |
| 1kqr | 5275 | A | 1.23 | 1.39 | 0.37 | 0.70 | 3.22 | 1.12 |
| 1mms | 5513 | A | 1.47 | 1.27 | 0.34 | 0.63 | 3.15 | 1.18 |
| 1fh9 | 7264 | A | 1.23 | 1.41 | | 0.63 | 3.08 | 1.32 |
| 2ihb | 7272 | B | 1.20 | 0.80 | 0.26 | 0.52 | 2.27 | 1.13 |
| | | | | | | | | |
| SHIFTX | | | | | | | | |
| PDB | BMRB | Chain | Cα | Cβ | Hα | $H_N$ | N | C' |
| 1iar | 4094 | A | 0.98 | 1.18 | 0.24 | 0.43 | 2.71 | 1.02 |
| 1bdo | 4425 | A | 1.03 | 1.35 | 0.25 | 0.49 | 1.96 | |
| 1kmi | 4472 | Y | 1.13 | 1.09 | 0.28 | 0.65 | 2.99 | |
| 1kqr | 5275 | A | 1.13 | 1.23 | 0.36 | 0.59 | 3.01 | 1.11 |
| 1mms | 5513 | A | 1.47 | 1.38 | 0.32 | 0.60 | 2.92 | 1.20 |
| 1fh9 | 7264 | A | 1.18 | 1.45 | | 0.61 | 2.79 | 1.35 |
| 2ihb | 7272 | B | 1.16 | 0.92 | 0.23 | 0.54 | 2.11 | 1.05 |

Table S3.
Comparison between the predictions provided by SPARTA[3], SHIFTX[4] and Camshift for a test set of seven proteins, which was previously used to compare SHIFTX and SPARTA. Chemical shift datasets were taken from the RefDB and re-referenced according to the method described by Zhang and coworkers[5]. All values are RMSDs in ppm.

| N | $H_N$ | Hα | Cα | Cβ | C' |
|---|---|---|---|---|---|
| 3.01 ± 0.18 | 0.55 ± 0.04 | 0.28 ± 0.03 | 1.26± 0.10 | 1.28± 0.27 | 1.44± 0.55 |

Table S4.
Averages and standard deviations in ppm for ten randomly selected subsets of seven proteins extracted from the 28-protein test set.

| | | N | $H_N$ | Hα | Cα | Cβ | C' |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| CamShift | All | 2.87 | 0.58 | 0.26 | 1.21 | 1.22 | 1.15 |
| | Helices | 2.26 | 0.49 | 0.21 | 1.05 | 0.94 | 1.10 |
| | Strands | 3.29 | 0.61 | 0.30 | 1.21 | 1.40 | 1.17 |
| | Coil | 2.95 | 0.59 | 0.27 | 1.31 | 1.38 | 1.09 |
| | Turns | 3.44 | 0.73 | 0.29 | 1.35 | 1.41 | 1.24 |
| | | | | | | | |
| SPARTA | All | 2.70 | 0.56 | 0.27 | 1.06 | 1.15 | 1.10 |
| | Helices | 2.18 | 0.51 | 0.21 | 1.01 | 0.88 | 1.04 |
| | Strands | 2.80 | 0.53 | 0.33 | 1.01 | 1.37 | 1.18 |
| | Coil | 3.22 | 0.62 | 0.28 | 1.09 | 1.41 | 1.03 |
| | Turns | 3.19 | 0.63 | 0.30 | 1.13 | 1.17 | 1.18 |
| | | | | | | | |
| SHIFTX | All | 2.73 | 0.57 | 0.29 | 1.17 | 1.28 | 1.20 |
| | Helices | 2.27 | 0.50 | 0.23 | 1.03 | 0.98 | 1.08 |
| | Strands | 2.81 | 0.58 | 0.32 | 1.15 | 1.48 | 1.18 |
| | Coil | 3.18 | 0.58 | 0.27 | 1.19 | 1.51 | 1.25 |
| | Turns | 3.20 | 0.69 | 0.35 | 1.37 | 1.41 | 1.40 |

Table S5.
Comparison of the predictions of chemical shifts for the 7-protein test set detailed in Table S3 divided for secondary structure elements; units are in ppm. For comparison, we also report the overall results of the chemical shift values predicted by PROSHIFT[6] (http://www.meilerlab.org/web/index.php), which is another commonly used prediction method: 3.19 ppm (N), 0.60 ppm ($H_N$), 0.33 ppm (Hα), 2.40 ppm (Cα), 2.62 ppm (Cβ) , 2.50 ppm (C´).

## Constants for all amino acids

Category 1: Random coil values (in ppm)

| Amino acid | Hα | Cα | $H_N$ | N | C' | Cβ | Ncorr[1] |
|---|---|---|---|---|---|---|---|
| ALA | 4.432847 | 52.26644 | 8.24 | 123.8 | 177.1 | 19 | -0.2 |
| ARG | 4.35015 | 56.24026 | 8.23 | 120.5 | 176.5 | 30.3 | 1.45 |
| ASN | 4.806587 | 53.30902 | 8.3 | 118.7 | 175.5 | 39 | 1.15 |
| ASP | 4.725179 | 54.158222 | 8.34 | 120.4 | 177.2 | 40.8 | 1.18 |
| CYS | 4.65 | 57.791772 | 8.32 | 118.6 | 175.1 | 41.8 | 2.43 |
| GLN | 4.405826 | 56.062813 | 8.32 | 119.8 | 176.3 | 30.1 | 1.57 |
| GLU | 4.364552 | 56.503798 | 8.42 | 120.2 | 176.1 | 29.7 | 1.8 |
| GLY | 3.97 | 45.486533 | 8.33 | 108.8 | 173.6 | 0 | 1.07 |
| HIS | 4.615947 | 55.660394 | 8.32 | 118.2 | 175.1 | 32 | 1.2 |
| ILE | 4.193333 | 61.206185 | 8.05 | 119.9 | 176.8 | 37.5 | 4.14 |
| LEU | 4.402746 | 54.957697 | 8.16 | 121.8 | 177.1 | 41.9 | 1.5 |
| LYS | 4.36807 | 56.243301 | 8.29 | 120.4 | 176.5 | 32.3 | 1.55 |
| MET | 4.461933 | 55.367 | 8.28 | 119.6 | 175.5 | 32.8 | 1.33 |
| PHE | 4.430244 | 57.64624 | 8.3 | 120.3 | 175.8 | 39.3 | 1.25 |
| PRO | 4.658447 | 63.066266 | 0 | 0 | 176 | 31.7 | 0.6 |
| SER | 4.504638 | 58.144025 | 8.31 | 115.7 | 173.7 | 62.7 | 2.59 |
| THR | 4.324869 | 62.54202 | 8.15 | 113.6 | 175.2 | 68.1 | 3.2 |
| TRP | 4.412784 | 56.208648 | 8.15 | 121.3 | 175.8 | 28.3 | 1.64 |
| TYR | 4.339301 | 57.33818 | 8.12 | 120.3 | 175.7 | 38.7 | 1.75 |
| VAL | 4.086605 | 62.512466 | 8.03 | 119.2 | 177.1 | 31.7 | 3.83 |

[1]Correction applied to the chemical shifts of the nitrogen atom of residue i+1 if residue i is of the type listed in the first column

## Categories, functional forms, and parameters for all amino acids except GLY and PRO

| | | Hα | Cα | $H_N$ | N | C' | Cβ |
|---|---|---|---|---|---|---|---|
| Category 2: Backbone distances | | $f(d) = c\,d^{\beta}, \beta = 1$ | | | | | |
| | | **c units: ppm/length** | | | | | |
| Atom | residue | | | | | | |
| N | i-1 | 0.1049538 | -0.5514491 | -0.7220064 | -0.9896088 | -0.0335257 | -0.2953833 |
| CA | i-1 | 0 | -0.3171652 | 0.7184281 | 0 | -0.3228718 | -0.1623656 |
| HA | i-1 | 0.0867546 | -2.2950958 | -1.2431488 | -3.3287317 | -0.3565204 | 0.4139583 |
| C | i-1 | 0 | 5.0389158 | 0 | 10.747981 | 0 | -1.2077866 |
| O | i-1 | -0.3532599 | -0.9486056 | 0 | -1.6351735 | -0.4007998 | 0.9900209 |
| N | i | 0 | 0 | 0 | -4.72E-10 | 0 | 3.6029795 |
| H | i | -0.2327538 | 3.9203989 | 1.40E-11 | -3.5391268 | -0.3422196 | -0.7672058 |
| CA | i | 0 | -1.95E-10 | 0 | -15.905456 | 0 | 0 |
| HA | i | -5.37E-13 | 7.88841 | 0.0647727 | -2.1882201 | 0 | 1.52206 |
| C | i | 0 | 6.2111825 | 0.356169 | 1.7859756 | -4.07E-11 | -5.7922812 |
| O | i | 0.1744758 | 3.6965816 | -0.5820469 | 4.2262422 | 0 | 1.2132408 |
| N | i+1 | 0 | 0.8900045 | -0.4391426 | 1.0503018 | 0 | 1.1459425 |
| H | i+1 | 0.1487117 | 1.0575798 | 0 | 4.1696025 | 0 | 0.9454372 |
| CA | i+1 | -0.0532967 | 0.8431783 | 0.3153333 | -3.5151185 | -1.4947525 | -2.4331324 |
| HA | i+1 | 0.0388204 | -0.0068419 | -0.0476424 | 0.7841273 | -0.1945234 | 0.1294023 |
| C | i+1 | 0.1117464 | 0.5577506 | -0.0367456 | 0.0492691 | -0.8447972 | 0.1112132 |
| Category 3: Side-chain distances | | $f(d) = c\,d^{\beta}, \beta = 1$ | | | | | |
| | | **c units: ppm/length** | | | | | |
| Amino acid | Atom | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ALA | CB | 0 | -3.4713212 | 0 | 0 | 0 | 1.06E-09 |
| ALA | 1HB | 0 | 0 | 0 | 0 | 0 | 0 |
| ALA | 2HB | 0 | 0 | 0 | 0 | 0 | 0 |
| ALA | 3HB | 0 | 0 | 0 | 0 | 0 | 0 |
| ARG | CB | 0 | 0 | 0 | 0 | 0 | -9.60E-12 |
| ARG | CG | 0 | 1.1144674 | 0 | 0.8725275 | 0 | 0 |
| ARG | CD | 0 | 0.8162443 | 0 | 0 | 0 | 0.2271032 |
| ARG | NE | -0.0430368 | -2.8863841 | 0 | 0.0024325 | -0.2415749 | -2.3789452 |
| ARG | CZ | 0 | 1.2386056 | 0 | 0 | 0 | -0.5131809 |
| ARG | NH1 | 0 | 0 | 0 | 0 | 0 | 0.688123 |
| ARG | NH2 | 0.0191117 | 0 | 0 | 0 | 0 | 0.688123 |
| ARG | 1HB | 0.0714608 | -1.9456232 | 0.0085882 | -1.6105041 | 0.0172898 | 1.1794897 |
| ARG | 2HB | 0.0714608 | -1.9456232 | 0.0085882 | -1.6105041 | 0.0172898 | 1.1794897 |
| ARG | 1HG | -0.0239778 | -0.2422652 | -0.0352422 | -0.8101128 | 0.0405182 | -0.1239578 |
| ARG | 2HG | -0.0239778 | -0.2422652 | -0.0352422 | -0.8101128 | 0.0405182 | -0.1239578 |
| ARG | 1HD | -0.0087136 | 0.3207482 | 0.0121353 | -0.3670633 | -0.0337978 | -0.2067592 |
| ARG | 2HD | -0.0087136 | 0.3207482 | 0.0121353 | -0.3670633 | -0.0337978 | -0.2067592 |
| ARG | HE | -0.0095285 | 1.2349324 | -0.0173074 | -0.4236829 | 0.0948721 | 0.886851 |
| ARG | 1HH1 | 0 | 0 | 0 | 0 | 0 | 0 |
| ARG | 2HH1 | 0 | 0 | 0 | 0 | 0 | 0 |
| ARG | 1HH2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ARG | 2HH2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ASN | CB | 0 | 0 | 0 | 0 | 0 | -4.15E-10 |
| ASN | CG | -0.1831709 | -13.440288 | 0 | -2.0759051 | 0 | -2.2470142 |
| ASN | OD1 | -0.0327841 | 5.0124168 | -0.0625137 | -0.4411607 | -0.3516236 | 0 |
| ASN | ND2 | 0 | 4.1041014 | 0 | -0.0950434 | 0 | 1.7681132 |
| ASN | 1HB | 0.100685 | 0.6762769 | 0.02455 | -1.2563696 | 0.138793 | 0 |
| ASN | 2HB | 0.100685 | 0.6762769 | 0.02455 | -1.2563696 | 0.138793 | 0 |
| ASN | 1HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ASN | 2HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ASP | CB | 0 | 6.5765072 | 0 | -3.8690467 | 0 | -2.41E-09 |
| ASP | CG | 0 | -5.8322385 | -0.0110424 | 4.5435659 | 0 | 0.4223025 |
| ASP | OD1 | -0.1025933 | 1.641523 | -0.0778158 | -2.8572188 | -0.1968005 | -0.5720464 |
| ASP | OD2 | -0.0972178 | 1.6768363 | 0.0083144 | -2.6882814 | -0.1607423 | 0.7996824 |
| ASP | 1HB | 0.1218923 | -1.778769 | 0.0294315 | -0.0839807 | 0.0088177 | 0 |
| ASP | 2HB | 0.1218923 | -1.778769 | 0.0294315 | -0.0839807 | 0.0088177 | 0 |
| CYS | CB | 0 | 0 | 0 | 0 | 0 | 1.10E-09 |
| CYS | SG | 0 | 0 | -0.0071872 | -0.9895489 | -0.1222304 | 0 |
| CYS | 1HB | 0 | 0 | -0.0085162 | -2.3032725 | -0.0245121 | 0 |
| CYS | 2HB | 0 | 0 | -0.0085162 | -2.3032725 | -0.0245121 | 0 |
| CYS | HG | 0 | 0 | 0 | 0 | 0 | -5.3074295 |
| GLN | CB | 0 | 0 | 0 | -3.7525202 | 0 | -7.19E-12 |
| GLN | CG | 0 | -0.4566143 | 0 | -4.0155098 | 0 | 17.326394 |
| GLN | CD | 0 | 3.5883887 | 0 | 6.9017111 | 0 | -13.849589 |
| GLN | OE1 | -0.0259782 | -1.0240332 | -0.0870417 | -2.8876885 | 0.0138941 | 4.5666691 |
| GLN | NE2 | 0.0092093 | -1.0116499 | 0 | -2.0799035 | -0.0306656 | 3.9474494 |
| GLN | 1HB | 0.0460137 | -2.1809066 | 0.0038255 | -0.538439 | -0.1263032 | 8.8285368 |
| GLN | 2HB | 0.0460137 | -2.1809066 | 0.0038255 | -0.538439 | -0.1263032 | 8.8285368 |
| GLN | 1HG | -0.0323311 | 0.7429182 | 0.0268177 | 0.4377085 | -0.0014186 | -8.8389784 |
| GLN | 2HG | -0.0323311 | 0.7429182 | 0.0268177 | 0.4377085 | -0.0014186 | -8.8389784 |
| GLN | 1HE2 | 0 | 0 | 0 | 0 | 0 | 0 |
| GLN | 2HE2 | 0 | 0 | 0 | 0 | 0 | 0 |
| GLU | CB | 0 | 0 | 0 | -0.3773162 | 0 | -6.36E-10 |
| GLU | CG | 0 | -2.8934267 | 0 | -4.9704113 | 0 | 0 |
| GLU | CD | 0 | 4.5000857 | 0 | 5.2939901 | 0 | -4.7840764 |
| GLU | OE1 | -0.0187235 | -1.3830948 | -0.1589678 | -2.1837438 | 0.0061113 | 1.3325745 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| GLU | OE2 | -0.0240431 | -1.3199339 | -0.1077505 | -1.8104882 | 0.1390527 | 1.1347852 |
| GLU | 1HB | 0.0534452 | -1.2034429 | 0.1043308 | -1.1019617 | -0.0804855 | 0 |
| GLU | 2HB | 0.0534452 | -1.2034429 | 0.1043308 | -1.1019617 | -0.0804855 | 0 |
| GLU | 1HG | -0.0084319 | 1.0118245 | 0.0687767 | 0.4988155 | -0.0505778 | 1.2801966 |
| GLU | 2HG | -0.0084319 | 1.0118245 | 0.0687767 | 0.4988155 | -0.0505778 | 1.2801966 |
| HIS | CB | 0 | 1.8283033 | 0 | -16.929154 | 0 | -3.88E-10 |
| HIS | CG | -0.1530754 | -2.6194459 | 0 | 2.4334007 | 0 | 0 |
| HIS | ND1 | 0 | -0.4199474 | 0 | -2.8866673 | 0 | 0 |
| HIS | CD2 | 0 | 0 | 0 | 0 | 0 | -3.6891183 |
| HIS | CE1 | 0 | 0 | 0 | 0 | 0 | 0 |
| HIS | NE2 | 0 | -4.9231018 | 0 | 0 | 0 | 0 |
| HIS | 1HB | 0.0667815 | -4.5033366 | 0.0224923 | 2.8658186 | 0.0057744 | 0 |
| HIS | 2HB | 0.0667815 | -4.5033366 | 0.0224923 | 2.8658186 | 0.0057744 | 0 |
| HIS | HD1 | 0 | 7.9189163 | -0.030249 | -0.7754716 | -0.0588564 | 2.6143485 |
| HIS | HD2 | 0 | 7.9189163 | -0.030249 | -0.7754716 | -0.0588564 | 2.6143485 |
| HIS | HE1 | 0 | -2.2028053 | -0.014255 | 3.2003381 | 0.0239427 | -1.5821813 |
| ILE | CB | 0 | 0 | 0 | 0 | 0 | -1.13E-09 |
| ILE | CG1 | 0 | -1.6648856 | 0 | 0 | 0 | 28.041084 |
| ILE | CG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ILE | CD1 | 0 | 0 | 0 | 0 | 0 | 30.373639 |
| ILE | HB | 0 | 1.2970047 | 0 | -4.7185183 | 0 | 22.654756 |
| ILE | 1HG1 | -0.0111565 | 0 | -0.0006385 | -0.5067059 | -0.1804216 | -33.38388 |
| ILE | 2HG1 | -0.0111565 | 0 | -0.0006385 | -0.5067059 | -0.1804216 | -33.38388 |
| ILE | 1HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ILE | 2HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ILE | 3HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| ILE | 1HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| ILE | 2HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| ILE | 3HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | CB | 0 | 3.517946 | 0 | 0 | 0 | 1.72E-10 |
| LEU | CG | 0 | 3.8861386 | -0.0712586 | -1.8316832 | 0 | 0 |
| LEU | CD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | CD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | 1HB | 0.0024544 | -3.5176338 | 0.0203021 | -2.0087278 | -0.2383368 | 0 |
| LEU | 2HB | 0.0024544 | -3.5176338 | 0.0203021 | -2.0087278 | -0.2383368 | 0 |
| LEU | HG | 0 | -0.7096024 | 0.002288 | 0.0301227 | 0.1279977 | 0.5441892 |
| LEU | 1HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | 2HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | 3HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | 1HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | 2HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEU | 3HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| LYS | CB | 0 | 0 | 0 | -6.3060844 | 0 | -3.19E-10 |
| LYS | CG | 0 | -0.3613907 | 0 | 1.6582529 | 0 | -11.609509 |
| LYS | CD | 0 | 1.4293193 | 0 | -0.5101093 | 0 | -0.1443795 |
| LYS | CE | 0 | -2.1458839 | 0 | 0 | 0 | -0.7009138 |
| LYS | NZ | 0 | 0.5528206 | 0 | 0 | -0.0459614 | 0.1855675 |
| LYS | 1HB | 0.0464037 | -1.3608475 | -0.0065542 | 0.0779636 | -0.053668 | 0 |
| LYS | 2HB | 0.0464037 | -1.3608475 | -0.0065542 | 0.0779636 | -0.053668 | 0 |
| LYS | 1HG | -0.0259734 | 0.2360132 | -0.0174149 | -0.7856875 | 0.0015136 | 4.4708471 |
| LYS | 2HG | -0.0259734 | 0.2360132 | -0.0174149 | -0.7856875 | 0.0015136 | 4.4708471 |
| LYS | 1HD | -0.0118798 | -0.0534137 | 0.0076768 | 0.0678926 | 0.0212682 | -0.0182633 |
| LYS | 2HD | -0.0118798 | -0.0534137 | 0.0076768 | 0.0678926 | 0.0212682 | -0.0182633 |
| LYS | 1HE | 0.0022253 | 0.5994697 | -0.0125536 | -0.0203758 | -0.0098667 | 0.2455037 |
| LYS | 2HE | 0.0022253 | 0.5994697 | -0.0125536 | -0.0203758 | -0.0098667 | 0.2455037 |
| LYS | 1HZ | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| LYS | 2HZ | 0 | 0 | 0 | 0 | 0 | 0 |
| LYS | 3HZ | 0 | 0 | 0 | 0 | 0 | 0 |
| MET | CB | 0 | 0 | 0 | 0 | 0 | 1.68E-09 |
| MET | CG | 0 | -0.6459234 | -0.6525533 | 0 | 0 | 0 |
| MET | SD | -0.1636483 | 2.8351113 | 0.1364456 | 1.2823161 | -0.0427728 | -6.3588798 |
| MET | CE | 0 | 0 | 0 | -0.5457662 | 0 | 0 |
| MET | 1HB | 0.1377561 | -3.7815143 | 0.0148497 | -2.247036 | -0.045965 | 0 |
| MET | 2HB | 0.1377561 | -3.7815143 | 0.0148497 | -2.247036 | -0.045965 | 0 |
| MET | 1HG | -0.0084985 | 0.9431252 | 0.1833973 | -0.991317 | 0.079936 | 4.3040213 |
| MET | 2HG | -0.0084985 | 0.9431252 | 0.1833973 | -0.991317 | 0.079936 | 4.3040213 |
| MET | 1HE | 0 | 0 | 0 | 0 | 0 | 0 |
| MET | 2HE | 0 | 0 | 0 | 0 | 0 | 0 |
| MET | 3HE | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | CB | 0 | 0 | -0.0108482 | -7.139512 | 1.6285523 | 7.67E-10 |
| PHE | CG | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | CD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | CD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | CE1 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | CE2 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | CZ | 0 | 0 | 0 | 0.9072064 | 0 | 0 |
| PHE | 1HB | 0 | -0.3469764 | -0.032161 | -0.8021542 | -0.8248405 | 0 |
| PHE | 2HB | 0 | -0.3469764 | -0.032161 | -0.8021542 | -0.8248405 | 0 |
| PHE | HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | HE1 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | HE2 | 0 | 0 | 0 | 0 | 0 | 0 |
| PHE | HZ | 0 | 0 | 0 | 0 | 0 | 0 |
| SER | CB | 0 | 1.7940727 | 0 | 0 | 0 | 4.04E-12 |
| SER | OG | -0.1130518 | 0.7939235 | -0.5580253 | -8.0106234 | -3.3810222 | 0 |
| SER | 1HB | 0.0652575 | -1.3559362 | -0.026532 | -1.3635124 | -0.07761 | -4.5539859 |
| SER | 2HB | 0.0652575 | -1.3559362 | -0.026532 | -1.3635124 | -0.07761 | -4.5539859 |
| SER | HG | -0.0087883 | 0 | 0.4586186 | 4.4828757 | 2.9825818 | 6.0879134 |
| THR | CB | 0 | 0 | 0 | 8.1381237 | 0 | -7.14E-12 |
| THR | OG1 | -0.0535248 | -2.0076927 | 0.028154 | -4.6797542 | -0.6556241 | 1.2314574 |
| THR | CG2 | 0 | 2.80742 | 0 | 0 | 0 | 0 |
| THR | HB | 0.0444565 | -1.8833967 | -0.0463287 | -6.9078573 | 0.4753306 | 0 |
| THR | HG1 | 0 | 0 | 0 | 0 | 0 | 0 |
| THR | 1HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| THR | 2HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| THR | 3HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| TRP | CB | 0 | 0 | 0 | 0 | 0 | -1.64E-12 |
| TRP | CG | 0 | 0 | 0 | 0 | 0 | 0 |
| TRP | CD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TRP | CD2 | 0 | 12.350985 | 0 | 0 | 0 | 0 |
| TRP | NE1 | 0 | 0 | 0 | 0 | 0 | 6.2259977 |
| TRP | CE2 | 0 | 0 | 0 | -7.9272257 | 0 | 0 |
| TRP | CE3 | 0 | 0 | 0 | 9.0887743 | 0 | 0 |
| TRP | CZ2 | 0 | 0 | 0 | -5.0425674 | 0 | -2.939783 |
| TRP | CZ3 | 0 | 0 | 0 | 0 | 0 | 0 |
| TRP | CH2 | 0 | 0 | 0 | 0 | 0 | 0 |
| TRP | 1HB | 0.0092938 | -0.4214938 | 0.1033322 | -1.8068207 | -0.2268331 | 0 |
| TRP | 2HB | 0.0092938 | -0.4214938 | 0.1033322 | -1.8068207 | -0.2268331 | 0 |
| TRP | HD1 | -0.2154728 | 12.502242 | -0.4010585 | -8.5114412 | 0.0953497 | 0.0297149 |
| TRP | HE1 | 0 | -11.943648 | 0.3700701 | 14.366154 | 0 | -5.3341363 |
| TRP | HE3 | -0.3873828 | 0 | -0.2975314 | -5.7956769 | 0 | -1.1800541 |
| TRP | HZ2 | 0 | -7.5751791 | 0 | 0 | 0 | 1.1527539 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| TRP | HZ3 | 0 | 0 | 0 | 0 | 0 | 2.5990143 |
| TRP | HH2 | 0.3277147 | 3.8883626 | 0 | 0 | 0.0876171 | 0 |
| TYR | CB | 0 | 0 | 0 | -10.697018 | 0 | 2.86E-13 |
| TYR | CG | 0 | 2.9675601 | 0 | 5.0955463 | 0 | 0 |
| TYR | CD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | CD2 | 0 | -1.6679126 | 0 | 0 | 0 | 0 |
| TYR | CE1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | CE2 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | CZ | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | OH | 0.0158655 | -3.6715087 | 0 | -1.4610716 | 0 | -1.4619849 |
| TYR | 1HB | 0 | 4.8082255 | -0.0116527 | 0.4579165 | -0.0937108 | -0.181595 |
| TYR | 2HB | 0 | 4.8082255 | -0.0116527 | 0.4579165 | -0.0937108 | -0.181595 |
| TYR | HD1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | HD2 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | HE1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | HE2 | 0 | 0 | 0 | 0 | 0 | 0 |
| TYR | HH | 0 | 0 | 0 | 0 | 0 | 1.4251078 |
| VAL | CB | 0 | 0 | 0 | -1.9792373 | 0 | -3.17E-12 |
| VAL | CG1 | 0 | 0 | 0 | 0 | 0 | 0.3495978 |
| VAL | CG2 | 0 | 0 | 0 | 0 | 0 | 0.3495978 |
| VAL | HB | 0 | -0.752457 | -0.0188545 | -4.2789195 | -0.5444149 | 0 |
| VAL | 1HG1 | 0 | 0 | 0 | 0 | 0 | 0 |
| VAL | 2HG1 | 0 | 0 | 0 | 0 | 0 | 0 |
| VAL | 3HG1 | 0 | 0 | 0 | 0 | 0 | 0 |
| VAL | 1HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| VAL | 2HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| VAL | 3HG2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Category 4: Nonbonded I | **$f(d) = c\ d^{\beta}$, $\beta = -3$**<br>**c units: ppm·length$^3$** | | | | | | |
| Atom | | | | | | | |
| Sub category 4a: sp3 hybridized | | | | | | | |
| C | | -2.9677854 | -10.347271 | -5.3813248 | -70.57317 | 2.7653856 | -14.770212 |
| H | | 0.6715652 | 3.4507071 | 1.4552245 | 34.697589 | 1.2320259 | 4.8983259 |
| N | | -1.6214566 | -12.598247 | 2.5053915 | -88.714197 | 31.323841 | -14.856998 |
| O | | 0.13518 | -3.566465 | 3.3033813 | 15.271538 | 31.172001 | -4.2014278 |
| S | | 6.2758 | -101.81223 | -0.8774168 | 89.182573 | 0.4631502 | 25.322355 |
| Sub category 4b: sp2 hybridized | | | | | | | |
| C | | -0.2035104 | -9.754433 | -14.43259 | -43.345602 | -27.91772 | -1.6525552 |
| N | | 1.2355083 | -9.6285342 | 17.498842 | 78.051259 | 40.257083 | 0.7468027 |
| O | | 8.1829427 | -8.6495093 | 9.1407986 | 53.274593 | 24.047757 | -9.2063201 |
| Category 5: Nonbonded II | **$f(d) = c\ d^{\beta}$, $\beta = 1$**<br>**c units: ppm/length** | | | | | | |
| Atom | | | | | | | |
| Sub category 5a: sp3 hybridized | | | | | | | |
| C | | 0.0026981 | 0.0340218 | -0.006955 | -0.0305132 | -0.0149591 | 0.0439096 |
| H | | -0.0035563 | -0.004969 | 0.0116062 | -0.1324914 | -0.0314517 | -0.0111027 |
| N | | 0.0093801 | 0.0474206 | -0.0631763 | 0.3450078 | -0.0069307 | 0.0436933 |
| O | | 0.0013892 | 0.0041549 | 0.0012358 | 0.085475 | -0.0624505 | -0.0125857 |
| S | | -0.0326842 | 0.359245 | 0.0129472 | -0.3312136 | -0.0197605 | -0.0796109 |
| Sub category 5b: sp2 hybridized | | | | | | | |
| C | | -0.0060841 | 0.0254618 | 0.03526 | 0.0402185 | 0.0658226 | 0.0122074 |
| N | | -0.0204519 | 0.0257897 | -0.0491208 | -0.255377 | -0.0809803 | -0.0353783 |
| O | | -0.0380543 | 0.0448192 | -0.0283907 | -0.184789 | -0.1061224 | 0.0308797 |
| Category 6: Aromatic rings | **$f(\theta, r) = c\ (1 - 3\cos^2(\theta))/r^3$, for parameters $\theta$ and r see Fig. S2**<br>**c units: ppm·length$^3$** | | | | | | |
| Ring type | | | | | | | |
| PHE | | 0.0197106 | 0.0107214 | 0.0248126 | -0.0121893 | 0.0379475 | 0.0228755 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| TYR | | | 0.0187376 | 0.0066462 | 0.0190884 | -0.0325441 | 0.0321684 | 0.0220022 |
| TRP2 | | | 0.0112872 | 0.0250235 | 0.0199792 | -0.0196035 | 0.0024794 | 0.0273681 |
| TRP1 | | | 0.0202472 | 0.0060401 | 0.0255547 | -0.0134924 | 0.0344568 | 0.0192496 |
| HIS | | | 0.011913 | 0.0036752 | 0.0130313 | -0.0110219 | 0.0176256 | 0.0131036 |
| Category 7: Dihedral angles | | | $f(\theta) = c\,(p_1 \cos(3(\theta + p_4)) + p_2 \cos(\theta + p_5) + p_3)$, where $\theta \in \{\phi, \varphi, \chi_1\}$. For parameters $p_1$ through $p_5$ see Table S2<br>c units: ppm; p units: dimensionless | | | | | |
| Angle | | | | | | | | |
| $\phi$ | | | 0.3960648 | 0.4778015 | 0.1404857 | 0.1306177 | -0.0663646 | 0.637021 |
| $\varphi$ | | | 0.176954 | 0.3692802 | 0.2091438 | 0.1913853 | 0.0596866 | 0.7236717 |
| $\chi_1$ | | | -0.0277769 | 0.1878352 | -0.2363657 | 0.2126343 | 0.2778522 | 0.1438652 |
| Category 8: Hydrogen bonds | | | $f(\theta) = c\,E(\theta)$, where $\theta \in \{d, \gamma_1, \gamma_2\}$. For the function E(x) see Eq. S1 and Table S1.<br>c units: ppm | | | | | |
| Origin | residue | Parameter | | | | | | |
| O | i-1 | d | -2.82E-08 | -5.15E-06 | -1.00E-07 | -7.527E-05 | 8.48E-06 | 3.302E-05 |
| O | i-1 | $\gamma_1$ | -0.0005796 | 0.0014532 | -0.0012312 | -0.0180411 | 0.0039887 | -0.0016318 |
| O | i-1 | $\gamma_2$ | 6.866E-05 | -0.0001394 | 0.0001251 | 0.0018755 | -0.0003917 | 0.0001679 |
| H | i | d | -0.001389 | 0.0030283 | -0.0142872 | -0.0134281 | -0.0043336 | -0.006617 |
| H | i | $\gamma_1$ | 0.0001071 | 0.0010661 | 0.0044821 | 0.0003982 | -0.0004837 | 0.0007823 |
| H | i | $\gamma_2$ | -5.55E-07 | -0.0001326 | -0.0005106 | -0.0001003 | 5.464E-05 | -3.278E-05 |
| O | i | d | 1.14E-06 | -2.569E-05 | -5.47E-06 | -8.44E-05 | -1.693E-05 | -1.792E-05 |
| O | i | $\gamma_1$ | 0.0003084 | -0.0002918 | 0.0002329 | 0.0065627 | -0.011352 | 0.001168 |
| O | i | $\gamma_2$ | -2.808E-05 | 3.277E-05 | -1.567E-05 | -0.0006374 | 0.0011653 | -0.0001242 |
| H | i+1 | d | -0.0017769 | -0.0052441 | 0.0011321 | -0.0155729 | -0.007736 | 0.001717 |
| H | i+1 | $\gamma_1$ | 0.0008555 | 0.0034481 | 0.000172 | 0.0117728 | 0.005713 | -0.0008433 |
| H | i+1 | $\gamma_2$ | -8.147E-05 | -0.0003803 | -9.97E-06 | -0.0012193 | -0.0005443 | 0.0001076 |
| Category 9: Extra distances | | | $f(d) = c\,d^{\beta}$, $\beta = 1$<br>c units: ppm/length | | | | | |
| Atom 1 | res 1 | atom 2 | res 2 | | | | | |
| H | i | HA | i | -0.2327538 | -0.7434611 | 0.0647727 | 15.413797 | 2.1192318 | -1.4816841 |
| H | i | C | i | 0.4706843 | -2.9678515 | 0.356169 | -0.5081131 | -0.3422196 | 1.2693928 |
| H | i | CB | i | 0.1105158 | 0.2527171 | 0.4513892 | 4.0841294 | -0.5979143 | -0.7672058 |
| C | i-1 | HA | i | -0.245715 | -0.5493254 | -0.1247354 | 5.3175934 | 1.4109851 | 0.3143816 |
| C | i-1 | C | i | 0.6633011 | -4.4043652 | -0.7657828 | -7.0572238 | -1.8892012 | 1.6552051 |
| C | i-1 | CB | i | 0.1545141 | -0.6886219 | -0.1308588 | -0.3235057 | 0.6930224 | -1.2077866 |
| O | i | HA | i | 0.1744758 | -7.050432 | 0.1035089 | -2.3941224 | -1.3898243 | 0.3139368 |
| O | i | N | i | -0.4090631 | 2.5129779 | 0.6346263 | 4.2262422 | 1.715943 | -0.3510702 |
| O | i | CB | i | 0.0299336 | 0.5613173 | 0.0991565 | 1.1221295 | 0.1724717 | 1.2132408 |
| N | i+1 | HA | i | 0.0455486 | -6.196119 | 0.1243992 | -5.4083632 | -0.2092066 | 0.6083014 |
| N | i+1 | N | i | -0.2246819 | 0.5916958 | -0.0851907 | 1.0503018 | 1.2968193 | -2.2086765 |
| N | i+1 | CB | i | -0.0349463 | 0.0887913 | -0.0992274 | -1.2712689 | -0.0724524 | 1.1459425 |
| O | i-1 | HA | i-1 | -0.1545062 | 0.9721613 | -0.4524309 | -1.1365341 | -0.5756206 | -0.5649802 |
| O | i-1 | N | i-1 | -0.2047421 | -0.0910651 | -0.4601405 | -1.2419222 | 0.2337042 | -0.2127165 |
| O | i-1 | CB | i-1 | -0.0351304 | 0.0416089 | 0.0069608 | -2.3622365 | 0.205132 | 0.0623775 |
| N | i | HA | i-1 | -0.1740476 | 3.0052241 | 1.3906446 | -3.3287317 | 0.0626519 | -1.2720737 |
| N | i | N | i-1 | -0.1117979 | 0.4661329 | 0.8949212 | -0.9896088 | 0.3919709 | 0.3149075 |
| N | i | CB | i-1 | 0.0189609 | 0.0051997 | -0.0271163 | 2.1847223 | -0.1618808 | -0.104322 |
| CG | i | HA | i | 0.1288138 | -0.8250044 | 0.027143 | 0.6921567 | 0.0426615 | -0.3198625 |
| CG | i | N | i | -0.0763197 | -0.0433575 | 0.1319916 | 2.8791465 | -0.1113439 | 0.2268996 |
| CG | i | C | i | -0.1053257 | 0.0469574 | -0.0870101 | 1.8044619 | 0.0823026 | -0.1846964 |
| CG | i | C | i-1 | 0.0026591 | -0.0793623 | -0.0520901 | 0.362815 | 0.0017187 | -0.0785277 |
| CG | i | N | i+1 | 0.0421798 | -0.1619192 | 0.0506248 | -0.2754379 | 0.0777308 | 0.0512366 |
| CG | i-1 | CA | i | 0.0042209 | -0.0016555 | 0.0111822 | -0.1476944 | -0.0103181 | 0.0103673 |
| CG | i+1 | CA | i | 0.0009935 | 0.0177661 | -0.0039802 | 0.0286774 | -0.0373196 | 0.0036293 |
| CA | i-1 | CA | i+1 | -0.0904907 | 0.6891161 | -0.0285188 | -1.2469408 | 0.5263286 | 1.0992811 |
| Category 10: Disulfide bond CYS | | | $f(d) = c\,d^{\beta}$, $\beta = 1$<br>c units: ppm/length | | | | | |
| Atom 1 | res 1 | atom 2 | res 2 | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| S | i | S | x | 0.0115552 | -2.8517004 | -0.1060554 | -1.4131442 | -0.3708814 | 11.361261 |

## Additional parameters for GLY

| | | Hα | Cα | $H_N$ | N | C' | Cβ | |
|---|---|---|---|---|---|---|---|---|
| Category 2: Backbone distances | | $f(d) = c\,d^{\beta}, \beta = 1$ **c units: ppm/length** | | | | | | |
| Atom | residue | | | | | | | |
| N | i-1 | 0.0663175 | -0.4879558 | -0.679174 | -0.7475809 | -0.0373979 | | 0 |
| CA | i-1 | 0 | -0.5317572 | 0.7014888 | 0 | -0.9441592 | | 0 |
| HA | i-1 | 0.0805372 | -1.7137742 | -1.0601918 | -3.7811287 | -0.3017085 | | 0 |
| C | i-1 | 0 | 3.8804893 | 0 | 11.661072 | 0 | | 0 |
| O | i-1 | -0.3009441 | -1.0573319 | 0 | -2.1254723 | -0.0635957 | | 0 |
| N | i | 0 | 0 | 0 | -4.02E-10 | 0 | | 0 |
| H | i | -0.1629661 | 2.8169776 | 3.68E-11 | 15.246462 | 0.3813989 | | 0 |
| CA | i | 0 | 1.91E-10 | 0 | -18.048166 | 0 | | 0 |
| HA | i | 3.39E-12 | 6.6313875 | 0.0135944 | 8.3760367 | 0 | | 0 |
| C | i | 0 | 6.7846053 | 0.5119419 | 12.00611 | 9.83E-13 | | 0 |
| O | i | 0.0989333 | 3.9973166 | -0.95712 | 3.4703765 | 0 | | 0 |
| N | i+1 | 0 | -1.5019093 | -0.9095434 | 1.1213514 | 0 | | 0 |
| H | i+1 | 0.1331114 | 0.7995692 | 0 | 3.9804866 | 0 | | 0 |
| CA | i+1 | 0.0646256 | 1.0488699 | 0.447443 | -2.8770606 | -1.215509 | | 0 |
| HA | i+1 | 0.0221346 | -0.0029701 | -0.0566733 | 0.6644562 | -0.2252219 | | 0 |
| C | i+1 | 0.1054101 | 0.5482872 | -0.0732623 | -0.070327 | -0.9356821 | | 0 |
| Category 3: Side chain distances | | $f(d) = c\,d^{\beta}, \beta = 1$ **c units: ppm/length** | | | | | | |
| Amino acid | Atom | | | | | | | |
| GLY | 2HA | 0.059929 | -3.5303299 | 0.030421 | 1.3803243 | 0.4653395 | | 0 |
| Category 4: Nonbonded I | | $f(d) = c\,d^{\beta}, \beta = -3$ **c units: ppm·length$^3$** | | | | | | |
| Atom | | | | | | | | |
| Sub category 4a: sp3 hybridized | | | | | | | | |
| C | | -3.0168132 | -8.5977612 | -5.8425338 | -71.591195 | 2.8434093 | | 0 |
| H | | 0.6934661 | 3.4918232 | 1.4585048 | 34.950005 | 1.1159714 | | 0 |
| N | | -1.5118058 | -11.682521 | 1.0705717 | -84.79476 | 31.208091 | | 0 |
| O | | -0.0116203 | -2.7975315 | 2.986752 | 15.415656 | 33.073499 | | 0 |
| S | | 6.0126235 | -90.045535 | -0.9345985 | 92.216868 | 2.2352785 | | 0 |
| Sub category 4b: sp2 hybridized | | | | | | | | |
| C | | 0.0261763 | -4.6399331 | -12.228723 | -35.763162 | -23.299868 | | 0 |
| N | | 1.7596832 | -24.154797 | 17.096711 | 93.833337 | 45.921117 | | 0 |
| O | | 7.515325 | -8.1137498 | 8.5922145 | 52.818653 | 24.239118 | | 0 |
| Category 5: Nonbonded II | | $f(d) = c\,d^{\beta}, \beta = 1$ **c units: ppm/length** | | | | | | |
| Atom | | | | | | | | |
| Sub category 5a: sp3 hybridized | | | | | | | | |
| C | | 0.0019265 | 0.0309129 | -0.0054204 | -0.0207984 | -0.0151925 | | 0 |
| H | | -0.0031884 | -0.0067515 | 0.0121696 | -0.1312018 | -0.0307875 | | 0 |
| N | | 0.0084885 | 0.0471441 | -0.0599493 | 0.3297395 | -0.0085599 | | 0 |
| O | | 0.0018812 | -0.0014212 | 0.0013845 | 0.0865619 | -0.0690416 | | 0 |
| S | | -0.032383 | 0.3254299 | 0.0090494 | -0.3311426 | -0.0239321 | | 0 |
| Sub category 5b: sp2 hybridized | | | | | | | | |
| C | | -0.0072228 | 0.0120399 | 0.0281911 | 0.0224532 | 0.052758 | | 0 |
| N | | -0.0251185 | 0.0727875 | -0.0591581 | -0.3550119 | -0.0906942 | | 0 |
| O | | -0.0346209 | 0.0412071 | -0.0269135 | -0.1633652 | -0.1020419 | | 0 |
| Category 6: Aromatic rings | | $f(\theta, r) = c\,(1 - 3\cos^2(\theta))/r^3$, **for parameters θ and r see Fig. S2** **c units: ppm/length** | | | | | | |
| Ring type | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| PHE | | | | 0.0187097 | 0.0109849 | 0.0235628 | -0.0095191 | 0.0367113 | 0 |
| TYR | | | | 0.0190572 | 0.006835 | 0.0152577 | -0.034505 | 0.0291119 | 0 |
| TRP2 | | | | 0.0116558 | 0.0225183 | 0.0224577 | -0.020596 | 0.0041195 | 0 |
| TRP1 | | | | 0.0198815 | 0.0079561 | 0.0213769 | -0.0133618 | 0.0330346 | 0 |
| HIS | | | | 0.0116936 | 0.002803 | 0.0128327 | -0.0118153 | 0.0169792 | 0 |
| Category 7: Dihedral angles | | | $f(\theta) = c\,(p_1 \cos(3(\theta + p_4)) + p_2 \cos(\theta + p_5) + p_3)$, where $\theta \in \{\phi, \varphi, \chi_1\}$. For parameters $p_1$ through $p_5$ see Table S2<br>c units: ppm; p units: dimensionless | | | | | | |
| Angle | | | | | | | | | |
| $\phi$ | | | | 0.4864764 | 0.4169099 | 0.2348974 | 0.2576444 | 0.1099509 | 0 |
| $\varphi$ | | | | 0.2353466 | 0.2842276 | -0.0211099 | 0.1762757 | 0.1410861 | 0 |
| $\chi_1$ | | | | 0.0417891 | 0.092503 | -0.2358528 | 0.1582806 | 0.4117375 | 0 |
| Category 8: Hydrogen bonds | | | $f(\theta) = c\,E(\theta)$, where $\theta \in \{d, \gamma_1, \gamma_2\}$. For function E(x) see Eq. S1 and Table S1<br>c units: ppm | | | | | | |
| Origin | residue | parameter | | | | | | | |
| O | i-1 | d | | -2.91E-07 | -7.02E-06 | -7.27E-07 | -8.416E-05 | 9.89E-06 | 0 |
| O | i-1 | $\gamma_1$ | | -0.000418 | 0.0006044 | -0.0012924 | -0.0165728 | 0.0046344 | 0 |
| O | i-1 | $\gamma_2$ | | 5.246E-05 | -4.566E-05 | 0.0001313 | 0.0017067 | -0.0004593 | 0 |
| H | i | d | | -0.001257 | 0.0030459 | -0.0148477 | -0.0126117 | -0.0049753 | 0 |
| H | i | $\gamma_1$ | | 0.0001398 | 0.0013573 | 0.0051242 | 0.0018565 | -0.0006142 | 0 |
| H | i | $\gamma_2$ | | -4.26E-06 | -0.0001636 | -0.0005739 | -0.0002636 | 6.704E-05 | 0 |
| O | i | d | | 1.47E-06 | -2.472E-05 | -6.01E-06 | -7.949E-05 | -1.65E-05 | 0 |
| O | i | $\gamma_1$ | | 0.0001921 | -0.0002637 | 2.224E-05 | 0.0055339 | -0.0107069 | 0 |
| O | i | $\gamma_2$ | | -1.612E-05 | 3.028E-05 | 6.29E-06 | -0.0005244 | 0.0010978 | 0 |
| H | i+1 | d | | -0.0018399 | -0.0050053 | 0.0009498 | -0.0132878 | -0.0099268 | 0 |
| H | i+1 | $\gamma_1$ | | 0.0009027 | 0.00308 | 9.249E-05 | 0.0128629 | 0.0060278 | 0 |
| H | i+1 | $\gamma_2$ | | -8.296E-05 | -0.0003454 | -8.88E-07 | -0.0013345 | -0.0005737 | 0 |
| Category 9: Extra distances | | | $f(d) = c\,d^{\beta}$, $\beta = 1$<br>c units: ppm/length | | | | | | |
| Atom 1 | residue 1 | atom 2 | residue 2 | | | | | | |
| H | i | HA | i | -0.1629661 | -0.1483798 | 0.0135944 | 8.766847 | 2.1013204 | 0 |
| H | i | C | i | 0.4122232 | -4.2093755 | 0.5119419 | -6.7520934 | 0.3813989 | 0 |
| H | i | CB | i | 0 | 0 | 0 | 0 | 0 | 0 |
| C | i-1 | HA | i | -0.0651481 | 0.2987554 | 0.0891024 | 5.6900611 | 0.7835696 | 0 |
| C | i-1 | C | i | 0.5298099 | -4.9560248 | -0.9512783 | -8.4671022 | -0.8618283 | 0 |
| C | i-1 | CB | i | 0 | 0 | 0 | 0 | 0 | 0 |
| O | i | HA | i | 0.0989333 | -7.3344972 | 0.0778165 | -4.6575886 | -0.8408604 | 0 |
| O | i | N | i | -0.337499 | 3.5153174 | 1.0217624 | 3.4703765 | 0.6362337 | 0 |
| O | i | CB | i | 0 | 0 | 0 | 0 | 0 | 0 |
| N | i+1 | HA | i | -0.0386912 | -5.2279072 | 0.0624355 | -5.4231033 | -0.1365515 | 0 |
| N | i+1 | N | i | -0.0449491 | 2.1436293 | 0.4497398 | 1.1213514 | -0.0124586 | 0 |
| N | i+1 | CB | i | 0 | 0 | 0 | 0 | 0 | 0 |
| O | i-1 | HA | i-1 | -0.1476902 | 1.2056681 | -0.340777 | -2.517914 | -0.363172 | 0 |
| O | i-1 | N | i-1 | -0.1780503 | -0.1273927 | -0.4324141 | -1.1224459 | 0.2225914 | 0 |
| O | i-1 | CB | i-1 | -0.0321439 | -0.0121417 | 0.0164976 | -2.2819998 | 0.1328632 | 0 |
| N | i | HA | i-1 | -0.1467708 | 2.8171414 | 1.1421495 | -3.7811287 | 0.279518 | 0 |
| N | i | N | i-1 | -0.0704203 | 0.4414006 | 0.8694495 | -0.7475809 | 0.4309026 | 0 |
| N | i | CB | i-1 | 0.0173538 | 0.0600871 | -0.0325651 | 2.1044602 | -0.0954739 | 0 |
| CG | i | HA | i | 0.1217622 | -0.9510271 | 0.028993 | 0.3537814 | 0.109656 | 0 |
| CG | i | N | i | -0.096412 | -0.011697 | 0.2195154 | 2.1215038 | -0.1739345 | 0 |
| CG | i | C | i | -0.0905658 | 0.0338832 | -0.0730184 | 1.9813139 | 0.1401974 | 0 |
| CG | i | C | i-1 | 0.0240178 | 0.0879133 | -0.1232099 | 1.0299104 | 0.052646 | 0 |
| CG | i | N | i+1 | 0.0367706 | -0.0177007 | 0.0122126 | -0.5722059 | -0.0112693 | 0 |
| CG | i-1 | CA | i | 0.0040347 | -0.0018201 | 0.0107078 | -0.1536396 | -0.0100844 | 0 |
| CG | i+1 | CA | i | 0.0013451 | 0.0164845 | -0.0026209 | 0.0327525 | -0.0375762 | 0 |
| CA | i-1 | CA | i+1 | -0.1203359 | 0.6912473 | -0.0638527 | -1.0492797 | 0.6711163 | 0 |

## Additional parameters for PRO

| | | Hα | Cα | $H_N$ | N | C' | Cβ |
|---|---|---|---|---|---|---|---|
| Category 2: Backbone distances | | $f(d) = c\ d^{\beta}, \beta = 1$ **c units: ppm/length** | | | | | |
| Atom | residue | | | | | | |
| N | i-1 | 0.0297782 | -0.3084801 | 0 | 0 | 0.0570203 | -0.2483017 |
| CA | i-1 | 0 | 0.8047905 | 0 | 0 | 0.4554952 | -1.6760849 |
| HA | i-1 | -0.0194261 | -2.0300362 | 0 | 0 | -0.3650702 | 0.2375751 |
| C | i-1 | 0 | 1.3222739 | 0 | 0 | 0 | -0.4814236 |
| O | i-1 | -0.2139526 | -0.7680869 | 0 | 0 | 0.0210265 | 1.1264962 |
| N | i | 0 | 0 | 0 | 0 | 0 | 1.6972503 |
| H | i | 0 | 0 | 0 | 0 | 0 | 0 |
| CA | i | 0 | 8.85E-11 | 0 | 0 | 0 | 0 |
| HA | i | -1.88E-12 | 7.724755 | 0 | 0 | 0 | 1.1828645 |
| C | i | 0 | 1.7933113 | 0 | 0 | -2.16E-11 | -4.7655492 |
| O | i | 0.0466054 | 7.9039908 | 0 | 0 | 0 | 0.9695745 |
| N | i+1 | 0 | 3.9079032 | 0 | 0 | 0 | 0.4914757 |
| H | i+1 | 0.2079951 | 1.1819019 | 0 | 0 | 0 | 1.2655731 |
| CA | i+1 | 0.0271823 | 0.7888304 | 0 | 0 | -1.1011533 | -2.2653164 |
| HA | i+1 | 0.0331915 | 0.0588155 | 0 | 0 | -0.1836946 | 0.1337957 |
| C | i+1 | 0.1143958 | 0.5778967 | 0 | 0 | -0.860993 | 0.0970634 |
| Category 3: Side chain distances | | $f(d) = c\ d^{\beta}, \beta = 1$ **c units: ppm/length** | | | | | |
| Amino acid | Atom | | | | | | |
| PRO | CB | 0.7663159 | -6.674831 | 0 | 0 | -6.5330139 | -1.16E-09 |
| PRO | CG | -1.0693319 | -4.4517713 | 0 | 0 | -9.0566327 | 37.196099 |
| PRO | CD | 0.5137461 | -1.5048913 | 0 | 0 | -3.7791219 | 3.9716195 |
| PRO | 1HB | -0.7856916 | 0.1989625 | 0 | 0 | 2.3072704 | 21.034246 |
| PRO | 2HB | -0.7856916 | 0.1989625 | 0 | 0 | 2.3072704 | 21.034246 |
| PRO | 1HG | 0.644203 | 2.0322733 | 0 | 0 | 3.8140635 | -21.435762 |
| PRO | 2HG | 0.644203 | 2.0322733 | 0 | 0 | 3.8140635 | -21.435762 |
| PRO | 1HD | -0.1377958 | 1.5457336 | 0 | 0 | 2.0392284 | -2.8604453 |
| PRO | 2HD | -0.1377958 | 1.5457336 | 0 | 0 | 2.0392284 | -2.8604453 |
| Category 4: Nonbonded I | | $f(d) = c\ d^{\beta}, \beta = -3$ **c units: ppm·length$^3$** | | | | | |
| Atom | | | | | | | |
| Sub category 4a: sp3 hybridized | | | | | | | |
| C | | -2.8890318 | -11.570118 | 0 | 0 | 0.7454209 | -15.431234 |
| H | | 0.678145 | 3.8625409 | 0 | 0 | 1.5414621 | 4.9580328 |
| N | | -1.6775879 | -14.305472 | 0 | 0 | 30.44214 | -13.276479 |
| O | | 0.1396715 | -3.2341673 | 0 | 0 | 31.004812 | -3.7642266 |
| S | | 6.1752519 | -97.777556 | 0 | 0 | -0.1268803 | 23.362399 |
| Sub category 4b: sp2 hybridized | | | | | | | |
| C | | -0.0160994 | -10.477399 | 0 | 0 | -29.372628 | -1.8666475 |
| N | | 1.0437019 | -5.6041119 | 0 | 0 | 29.917336 | -4.3766057 |
| O | | 8.0417551 | -8.575915 | 0 | 0 | 24.741825 | -8.6247771 |
| Category 5: Nonbonded II | | $f(d) = c\ d^{\beta}, \beta = 1$ **c units: ppm/length** | | | | | |
| Atom | | | | | | | |
| Sub category 5a: sp3 hybridized | | | | | | | |
| C | | 0.0023439 | 0.0353139 | 0 | 0 | -0.0095636 | 0.0451183 |
| H | | -0.0037569 | -0.0063923 | 0 | 0 | -0.0329768 | -0.0118314 |
| N | | 0.0097003 | 0.053891 | 0 | 0 | -0.0036915 | 0.037665 |
| O | | 0.0009191 | 0.0031997 | 0 | 0 | -0.0619811 | -0.0114114 |
| S | | -0.0327975 | 0.348281 | 0 | 0 | -0.0211074 | -0.0725498 |
| Sub category 5b: sp2 hybridized | | | | | | | |
| C | | -0.0070156 | 0.0257928 | 0 | 0 | 0.0698801 | 0.0109669 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N | | | -0.0190643 | 0.0141753 | 0 | 0 | -0.0497686 | -0.0161789 |
| O | | | -0.0374516 | 0.0469155 | 0 | 0 | -0.1080708 | 0.0300053 |
| Category 6: Aromatic rings | | | $f(\theta, r) = c\,(1 - 3\cos^2(\theta))/r^3$<br>**c units: ppm/length** | | | | | |
| Ring type | | | | | | | | |
| PHE | | | 0.0186948 | 0.0109772 | 0 | 0 | 0.0369795 | 0.0228389 |
| TYR | | | 0.0185362 | 0.0068703 | 0 | 0 | 0.0333031 | 0.0210853 |
| TRP2 | | | 0.0119457 | 0.0208549 | 0 | 0 | 0.0008892 | 0.0260824 |
| TRP1 | | | 0.0194632 | 0.0120662 | 0 | 0 | 0.0365632 | 0.0191159 |
| HIS | | | 0.0121215 | 0.0038377 | 0 | 0 | 0.0193006 | 0.0144385 |
| Category 7: Dihedral angles | | | $f(\theta) = c\,(p_1\cos(3(\theta + p_4)) + p_2\cos(\theta + p_5) + p_3)$, where $\theta \in \{\phi, \varphi, \chi_1\}$. For parameters $p_1$ through $p_5$ see Table S2<br>**c units: ppm; p units: dimensionless** | | | | | |
| Angle | | | | | | | | |
| $\phi$ | | | 0.4106709 | 0.481939 | 0 | 0 | -0.0141746 | 0.5496927 |
| $\varphi$ | | | 0.2242678 | 0.4381989 | 0 | 0 | 0.1017863 | 0.7894501 |
| $\chi_1$ | | | 0.0579212 | 0.0899445 | 0 | 0 | 0.433849 | 0.1204433 |
| Category 8: Hydrogen bonds | | | $f(\theta) = c\,E(\theta)$, where $\theta \in \{d, \gamma_1, \gamma_2\}$. For function E(x) see Eq. S1 and Table S1<br>**c units: ppm** | | | | | |
| Origin | residue | parameter | | | | | | |
| O | i-1 | d | 1.07E-08 | -6.69E-06 | 0 | 0 | 8.95E-06 | 3.469E-05 |
| O | i-1 | $\gamma_1$ | -0.0005608 | 0.0019834 | 0 | 0 | 0.0044442 | -0.0018979 |
| O | i-1 | $\gamma_2$ | 6.722E-05 | -0.0001937 | 0 | 0 | -0.0004373 | 0.0001956 |
| H | i | d | -0.0015129 | 0.0025176 | 0 | 0 | -0.0047888 | -0.0068513 |
| H | i | $\gamma_1$ | -2.591E-05 | 0.0012989 | 0 | 0 | -0.0003295 | 0.0003949 |
| H | i | $\gamma_2$ | 1.342E-05 | -0.0001582 | 0 | 0 | 3.708E-05 | 8.74E-06 |
| O | i | d | 1.89E-06 | -2.615E-05 | 0 | 0 | -1.792E-05 | -1.858E-05 |
| O | i | $\gamma_1$ | 0.000219 | -0.0001293 | 0 | 0 | -0.0111639 | 0.0010705 |
| O | i | $\gamma_2$ | -1.871E-05 | 1.616E-05 | 0 | 0 | 0.0011469 | -0.0001135 |
| H | i+1 | d | -0.0018291 | -0.0047666 | 0 | 0 | -0.0083926 | 0.0016062 |
| H | i+1 | $\gamma_1$ | 0.0007124 | 0.0034823 | 0 | 0 | 0.0057975 | -0.0005487 |
| H | i+1 | $\gamma_2$ | -6.613E-05 | -0.0003854 | 0 | 0 | -0.0005543 | 7.667E-05 |
| Category 9: Extra distances | | | $f(d) = c\,d^{\beta}$, $\beta = 1$<br>**c units: ppm/length** | | | | | |
| Atom 1 | residue 1 | atom 2 | residue 2 | | | | | |
| H | i | HA | i | 0 | 0 | 0 | 0 | 0 | 0 |
| H | i | C | i | 0 | 0 | 0 | 0 | 0 | 0 |
| H | i | CB | i | 0 | 0 | 0 | 0 | 0 | 0 |
| C | i-1 | HA | i | 0.0239546 | 0.0550619 | 0 | 0 | -0.1309367 | 0.6048279 |
| C | i-1 | C | i | 0.3639351 | -1.8245109 | 0 | 0 | -2.1589941 | 1.0121608 |
| C | i-1 | CB | i | 0.0601703 | -1.1517461 | 0 | 0 | 1.4440016 | -0.4814236 |
| O | i | HA | i | 0.0466054 | -7.9758753 | 0 | 0 | -0.4737369 | -0.4962168 |
| O | i | N | i | 0.0608192 | 0.0143416 | 0 | 0 | 0.4166167 | 1.236902 |
| O | i | CB | i | -0.0032265 | 0.177568 | 0 | 0 | 0.3098538 | 0.9695745 |
| N | i+1 | HA | i | -0.3380021 | -7.1442735 | 0 | 0 | 0.4825454 | -0.0616248 |
| N | i+1 | N | i | 0.2101381 | -1.4483086 | 0 | 0 | 0.2440395 | -0.098548 |
| N | i+1 | CB | i | -0.1088596 | -0.1018623 | 0 | 0 | 0.0799778 | 0.4914757 |
| O | i-1 | HA | i-1 | -0.1535562 | 0.9541831 | 0 | 0 | -0.3530593 | -0.1973697 |
| O | i-1 | N | i-1 | -0.1617153 | -0.0384509 | 0 | 0 | 0.2598648 | -0.2525967 |
| O | i-1 | CB | i-1 | -0.0505202 | 0.0836814 | 0 | 0 | 0.2533118 | -0.0658209 |
| N | i | HA | i-1 | -0.1062667 | 2.7549949 | 0 | 0 | 0.1927719 | -0.595766 |
| N | i | N | i-1 | -0.0340523 | 0.2772183 | 0 | 0 | 0.2516512 | 0.3693537 |
| N | i | CB | i-1 | 0.0358333 | -0.0347704 | 0 | 0 | -0.2073699 | 0.0314129 |
| CG | i | HA | i | 0.1186278 | -0.5677036 | 0 | 0 | 0.0794325 | -0.2580184 |
| CG | i | N | i | -0.0881744 | 0.4042956 | 0 | 0 | -0.117154 | 0.2480631 |
| CG | i | C | i | -0.0746942 | 0.3357447 | 0 | 0 | -0.0332108 | -0.0770119 |
| CG | i | C | i-1 | 0.0187258 | -0.1100216 | 0 | 0 | 0.0023606 | -0.1108395 |
| CG | i | N | i+1 | 0.0290884 | -0.105078 | 0 | 0 | 0.0711681 | -0.035068 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CG | i-1 | CA | i | 0.0042968 | -0.0012085 | 0 | 0 | -0.0091117 | 0.0070518 |
| CG | i+1 | CA | i | 0.001109 | 0.0159162 | 0 | 0 | -0.0382256 | 0.0025576 |
| CA | i-1 | CA | i+1 | -0.1514723 | 0.7027408 | 0 | 0 | 0.4796896 | 0.3274878 |

Table S6. Values of the parameters of Eq. 1 in the main text. The equations for the terms $f(x)$ describing the various contributions are also given. Chemical shifts δ are calculated as random coil values $c_0$ plus the sum over the $f(x)$ terms for all categories. Additional parameters for the hydrogen bonding and the dihedral angle terms can be found in Tables S1 and S2, respectively. For categories 2, 3, 4, and 5, distances are calculated from the query atom to each of the atoms listed. The units of the parameters c are also given.

**References**


(1) Morozov, A. V.; Kortemme, T.; Tsemekhman, K.; Baker, D. *Proc. Natl. Acad. Sci. USA* **2004**, 101, 6946-6951.
(2) Pople, J. A. *Mol. Phys.* **1958**, 1, 175-180.
(3) Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, 38, 289-302.
(4) Neal, S.; Nip, A. M.; Zhang, H. Y.; Wishart, D. S. *J. Biomol. NMR* **2003**, 26, 215-240.
(5) Zhang, H. Y.; Neal, S.; Wishart, D. S. *J. Biomol. NMR* **2003**, 25, 173-195.
(6) Meiler, J. *J. Biomol. NMR* **2003**, 26, 25-37.