Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics

Supplementary Information

Rapid Calculation of Protein Chemical Shifts Using Bond Polarization Theory and its Application to Protein Structure Refinement

Igor Jakovkin, Marco Klipfel, Claudia Muhle-Goll, Anne S. Ulrich, Burkhard Luy and Ulrich Sternberg

Contents

1. Adjustment of ^{13}C reference for proteins

Table SI1: Comparison of the BPT calculated $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts (CS) in NH_3^+-$(Ala_4X)_4Ala_4$-COO^- (X=corresponding amino acid) peptides with the consensus α-helical geometry (φ = -57.80° and ψ = -47.00°) with the averaged experimental $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts in helical structures as defined by Zhang et al.[1] The differences ΔCS between BPT calculated and averaged experimental chemical shifts are comparable with the systematic deviations between BPT calculated and experimental chemical shifts in Ubiquitin (PDB ID 1D3Z). All calculations were performed using original BPT parameterization[2,3] for ^{13}C and additional reference value for ^{13}C CS calculation in proteins could be defined.

Amino acid	BPT CS* (in ppm)	Averaged CS (in ppm)	ΔCS (in ppm)	ΔCS in Ubiquitin (in ppm)
$^{13}C_\alpha$ nucleus				
GLY	40.33±0.19	46.91±1.10	6.58±1.29	4.59±2.11
ILE	57.90±0.47	64.57±1.74	6.67±2.21	5.29±2.31
LYS	55.77±0.29	58.93±1.44	3.16±1.73	2.69±2.22
SER	55.78±0.11	60.88±1.61	5.10±1.72	4.58±2.86
THR	55.29±0.05	65.61±2.39	10.32±2.44	5.77±1.98
VAL	55.70±0.15	66.16±1.55	10.46±1.70	5.97±4.24
Reference $^{13}C_\alpha$				4.8 ppm
PRO	50.73±0.21	65.49±1.08	14.76±1.29	16.24±0.16
Reference $^{13}C_\alpha$ PRO				16.2 ppm
$^{13}C_\beta$ nucleus				
ARG	37.22±0.22	30.14±1.14	-7.08±1.36	-7.77±0.39
GLN	35.26±0.21	28.51±0.92	-6.75±1.13	-8.02±0.95
GLU	35.79±0.38	29.37±0.99	-6.42±1.37	-8.39±1.86
ILE	46.15±0.12	37.60±1.15	-8.55±1.27	-9.91±2.95
VAL	45.46±0.19	31.49±0.72	-13.97±0.91	-11.13±2.43
Reference $^{13}C_\beta$				-9.0 ppm

*All ^{13}C chemical shifts are referenced to DSS.

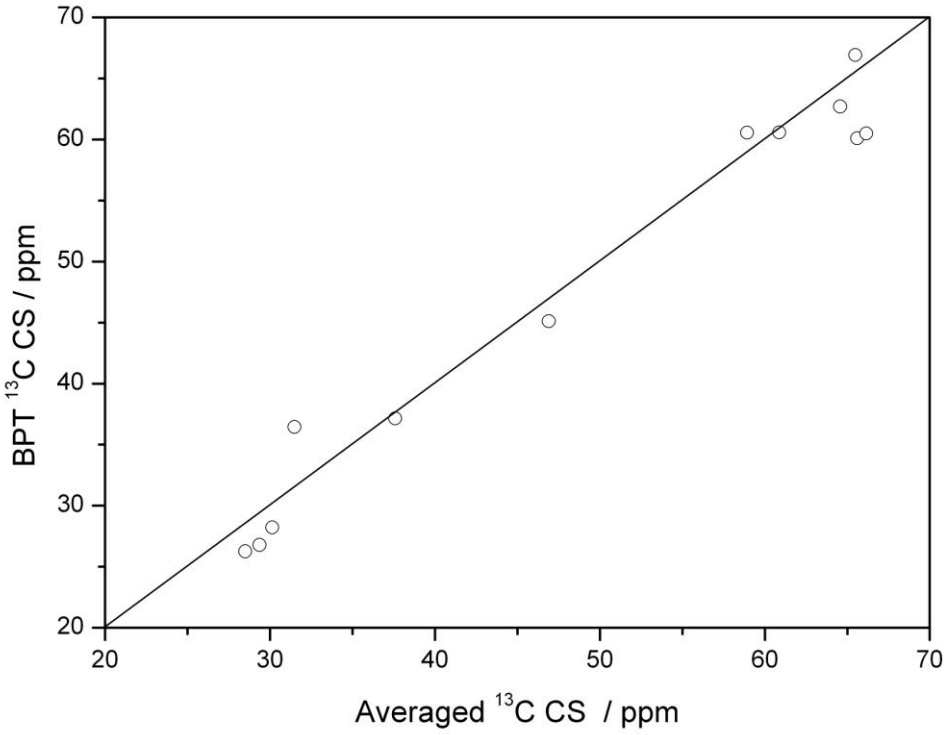

Figure SI1: Correlation between the BPT calculated $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts in peptides with consensus α-helical geometry and the averaged experimental $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts in helical structures as defined by Zhang et al.[1] (correlation coefficient R=0.983 and standard deviation SD=2.90 ppm; BPT parameterization includes additional references given in Table SI1).

2. ^{15}N chemical shift calculations in crystalline tripeptides

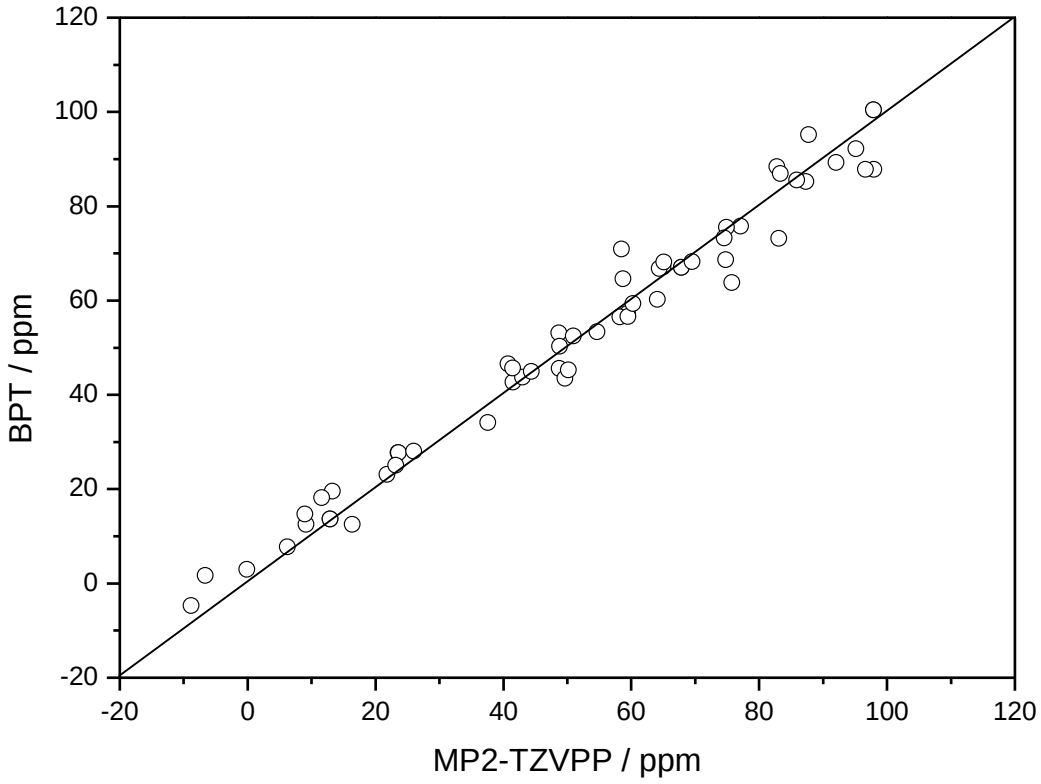

Figure SI2: Correlation of the isotropic ^{15}N nuclear shielding values from *ab initio* calculations at the MP2/TZVPP level with values obtained by BPT (correlation coefficient R=0.988, and standard deviation SD=4.7 ppm).

Table SI2: Comparison of the experimental[4] and BPT calculated ^{15}N chemical shifts in central residues of crystalline tripeptides. Note that peptide trimers (cf. Figure SI2) are not sufficient representation of crystalline environment neither in BPT nor in DFT[4] framework.

^{15}N Chemical shift in ppm					
Molecule	Experiment	BPT – Unit cell	BPT – Trimer	DFT – Trimer	BPT – Monomer
APG	132.4±0.5	136.1	124.4	141	156.7
AGG	104.8±0.5	104.0	108.5	113	135.6
GGV	112.8±0.5	114.7	110.8	122	121.4
MAE		2.1	4.6	8.7	21.2

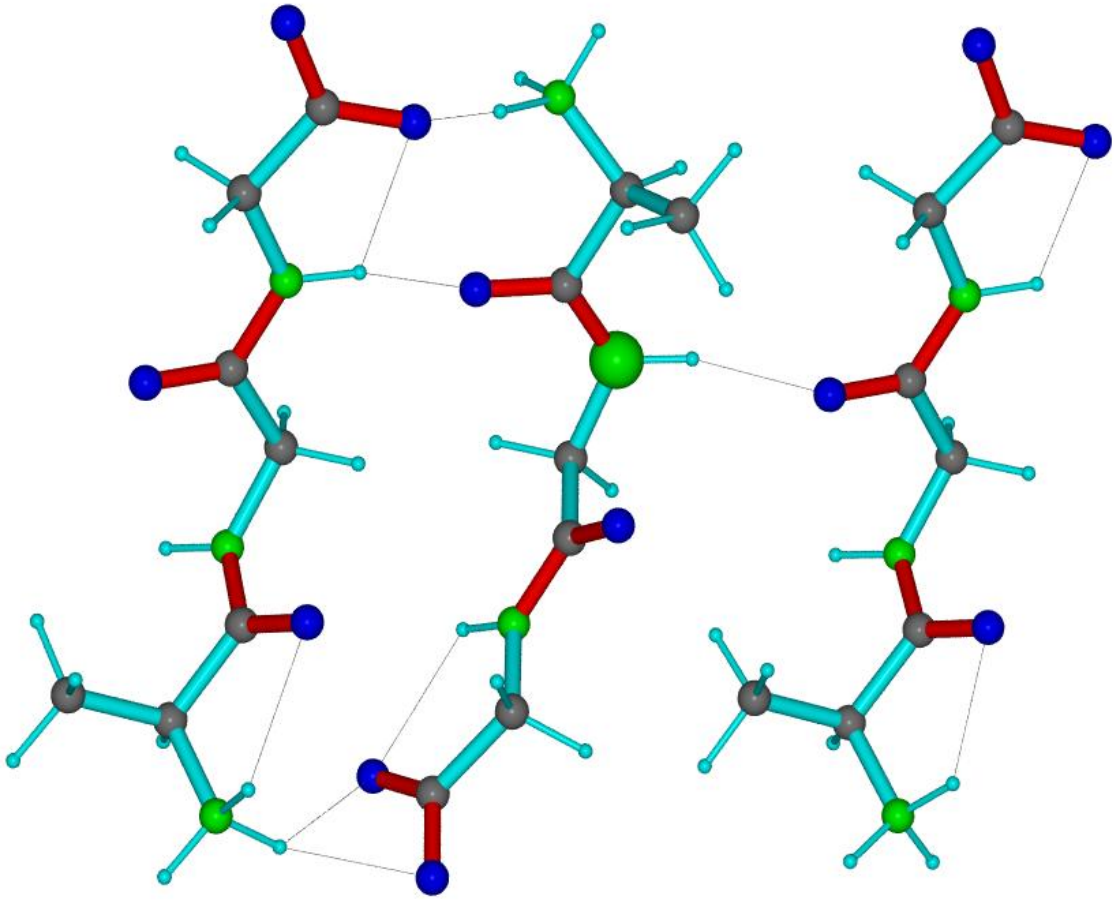

Figure SI3A: Structure of AGG trimer used in DFT and BPT calculations. Target nitrogen is emphasized and hydrogen bonds are shown in gray.

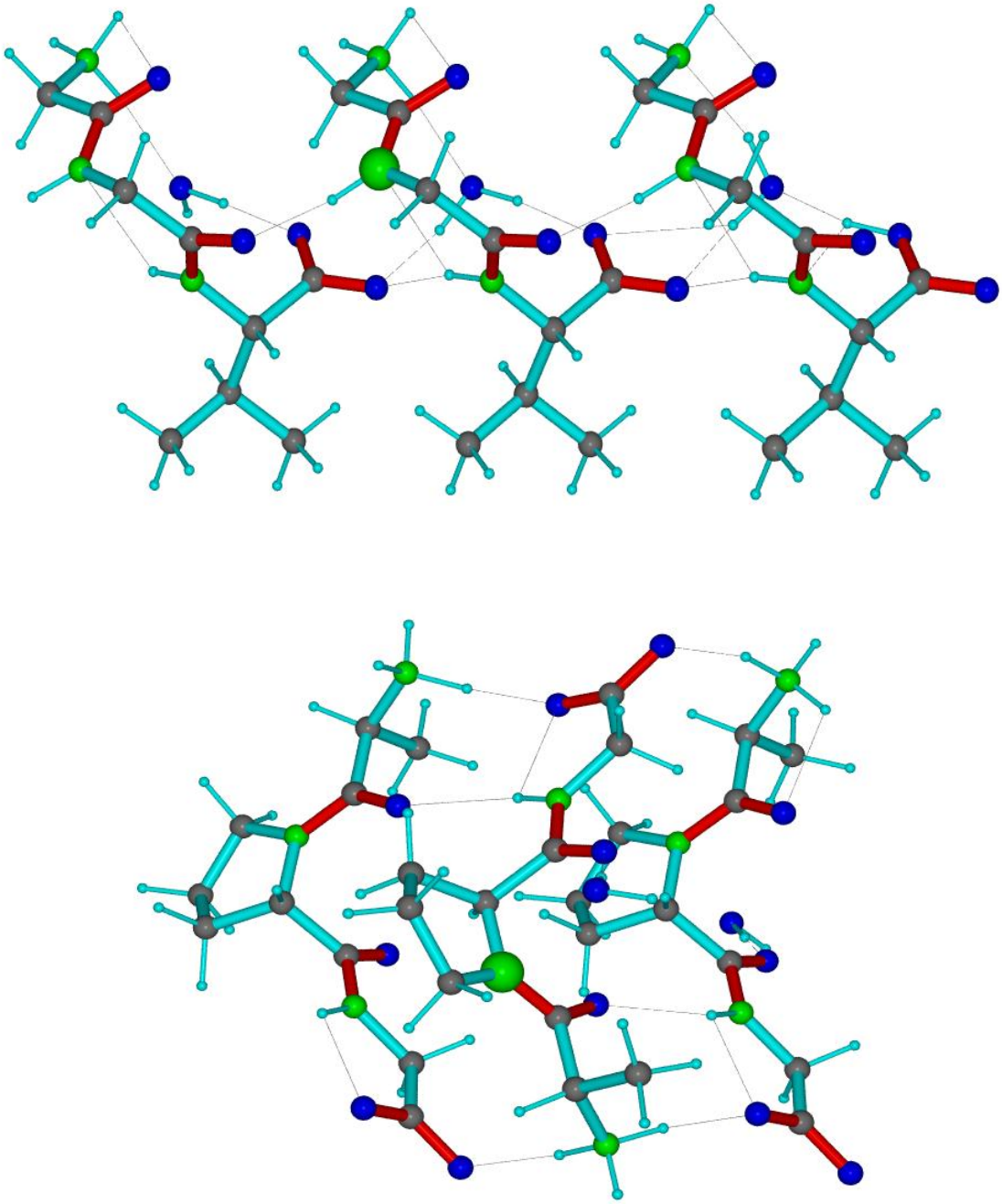

Figure SI3B: Structures of GGV (at the top) and APG (at the bottom) trimers used in DFT and BPT calculations. Target nitrogen is emphasized and hydrogen bonds are shown in gray.

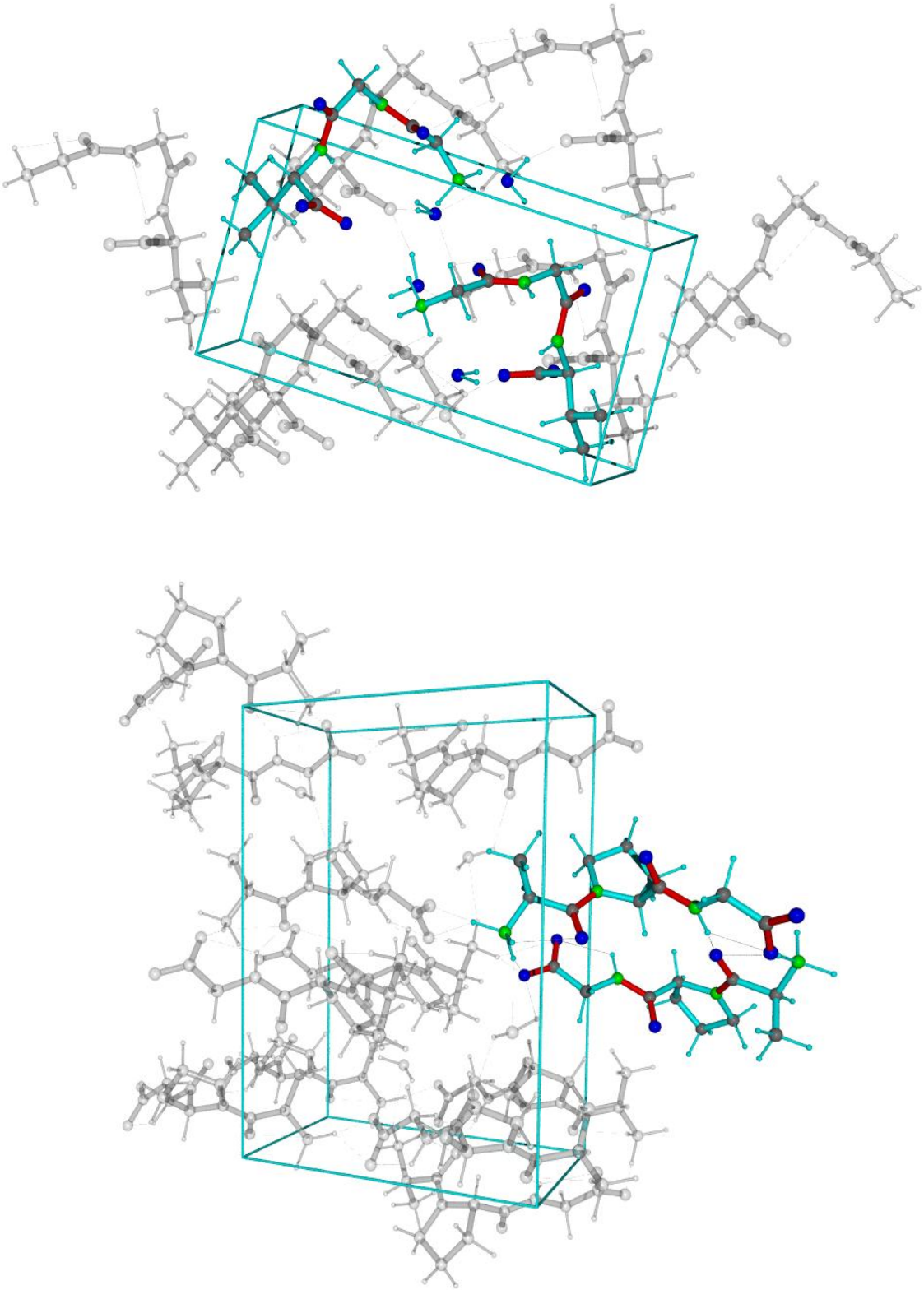

Figure SI4: Unit cells of GGV (at the top) and APG (at the bottom) tripeptides used in BPT calculations with periodic boundary conditions. Structures belonging to the asymmetric unit are shown in full colour.

Table SI3: Experimental[5] and calculated chemical shifts (in ppm) of N-formyl-L-Met-L-Leu-L-Phe-OH (all ^{13}C CS are referenced to DSS): MD – averaged CS values over 10 ps molecular dynamics trajectory; FSS – CS values obtained via full structure search of the conformational space[5]; ^{1}H opt. – CS values of the structure obtained via full structure search after hydrogen position optimization.

Residue	Experiment	BPT – MD	BPT – FSS	BPT – ^{1}H opt.	DFT* – ^{1}H opt.
^{15}N					
Met	125.5	124.3	125.2	127.1	146.6 (98.0)
Leu	116.2	117.5	119.2	124.6	127.2 (117.4)
Phe	107.6	109.3	113.6	101.4	109.6 (135.0)
MAE		1.4	3.1	5.4	11.4
$^{13}C_{\alpha}$					
Met	52.0	54.8	54.9	55.0	58.4 (127.5)
Leu	56.8	55.2	55.3	55.1	63.9 (122.0)
Phe	54.4	54.4	54.8	54.1	62.7 (123.2)
MAE		1.5	1.6	1.7	7.3
$^{13}C_{\beta}$					
Met	37.9	39.7	39.7	40.1	47.4 (138.5)
Leu	40.7	41.1	41.0	41.4	49.8 (136.1)
Phe	36.9	42.0	41.9	41.8	37.9 (148.0)
MAE		2.4	2.4	2.6	6.5

* Nuclear shieldings are given in brackets. ^{13}C nuclear shieldings were converted into chemical shifts using calculated reference value for TMS in Table SI4. 1.7 ppm were added to each calculated ^{13}C chemical shift for conversion from TMS to DSS chemical shift scale[6]. ^{15}N nuclear shieldings were converted into chemical shifts using the reference value of 244.6 ppm[4].

Table SI4: Experimental[7,8] and calculated (tight geometry optimization[9,10], DFT B3LYP with 6-311++G(d,p) basis set) geometry parameters and chemical shieldings of TMS. Calculated shielding was used for conversion of ^{13}C nuclear shieldings into chemical shifts given in Table SI3.

Parameters of the calculated and experimental TMS geometry and chemical shielding				
Molecule	Si–C in Å	C–H in Å	Si–C–H in degrees	σ in ppm
Experimental	1.877±0.004	1.110±0.003	111.0±0.2	188.1
DFT B3LYP	1.891	1.095	111.3	184.2

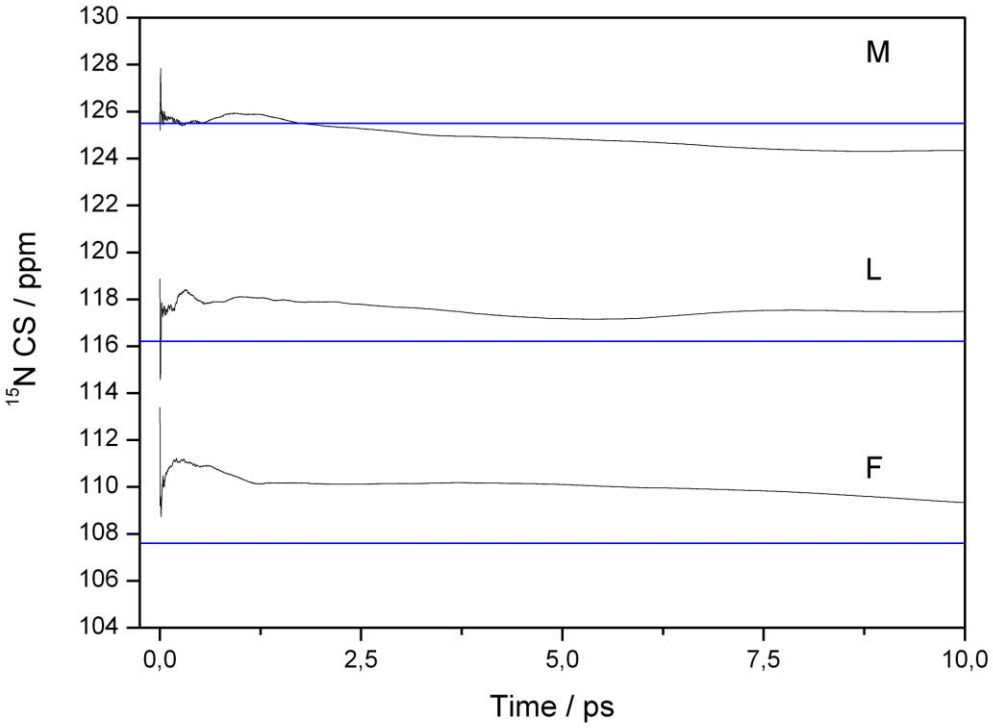

Figure SI4: Calculated averaged ^{15}N chemical shift in N-formyl-L-Met-L-Leu-L-Phe-OH during molecular dynamics (MD). To minimize the influence of the MD equilibration phase on the averaged chemical shift time averaging was performed using exponential memory decay function[11].

3. Definition of $^{13}C_\alpha$-RMSD

Definition of the conformationally averaged $^{13}C_\alpha$ chemical shift root mean square deviation between calculated and experimental chemical shift values ($^{13}C_\alpha$-RMSD) as introduced by Vila et al. for validation of protein structures[12] and comparison of computational methods[13]:

$$^{13}C_\alpha - RMSD = \left[(1/N) \sum_{i=1}^{N} (\delta_{\exp,i} - (1/\Omega) \sum_{j=1}^{\Omega} \delta_{calc,i,j})^2 \right]^{1/2} \qquad \text{(SI1)}$$

with N – number of amino acids in protein sequence, Ω – number of protein structures in ensemble and δ - chemical shift of $^{13}C_\alpha$.

4. References

[1] Zhang, H.; Neal, S.; Wishart, D. S. RefDB: A database of uniformly referenced protein chemical shifts; *J. Biomol. NMR* **2003**, *25*, 173-195

[2] Prieß, W.; Sternberg, U. Fast calculation of ^{13}C NMR chemical shift tensors using bond polarization model; *J. Mol. Struct. (Theochem)* **2001**, *544*, 181

[3] Sternberg, U.; Koch, F.-Th.; Prieß, W.; Witter, R. Crystal structure refinements of cellulose polymorphs using ^{13}C solid state chemical shifts; *Cellulose* **2003**, *10*, 189-199

[4] Waddell, K. W.; Chekmenev E. Y.;Wittebort R. J. Single-Crystal Studies on Peptide Prolyl and Glycyl ^{15}N Shielding Tensors; *J. Am. Chem. Soc.* **2005**, *127*, 9035

[5] Rienstra, C. M.; Tucker-Kellogg, L.; Jaroniec, C. P.; Hohwy, M.; Reif, B.; McMahon, M. T.; Tidor, B.; Lozano-Perez, T.; Griffin, R. G. *De novo* determination of peptide structure with solid-state magic-angle spinning NMR spectroscopy; *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10260

[6] Wishart, D. S,; Bigam, C. G,; Yao, J,; Abildgaard, F,; Dyson, H. J,; Oldfield, E,; Markley, J. L.; Sykes, B. D. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR; *J. Biomol. NMR* **1995**, *6*, 135–140

[7] Campanelli, A. R.; Ramondo, F.; Domenicano, A.; Hargittai, I. Toward a more accurate silicon stereochemistry: an electron diffraction study of the molecular structure of tetramethylsilane; *Struct. Chem.* **2000**, *11*, 155-160

[8] Jameson, A. K.; Jameson, C. J. Gas-phase ^{13}C chemical shifts in the zero-pressure limit: refinements to the absolute shielding scale for ^{13}C; *J. Chem. Phys. Lett.* **1997**, *134*, 461

[9] Csaszar, P. and Pulay, P. Geometry optimization by direct inversion in the iterative subspace; *J. Mol. Struct. (Theochem)* **1984**, *114*, 31-34

[10] Farkas, Ö. and Schlegel, H. B. Methods for optimizing large molecules. II. Quadratic search; *J. Chem. Phys.* **1999**, *111*, 10806-10814

[11] Sternberg, U.; Witter, R.; Ulrich, A. S. All-atom molecular dynamics simulations using orientational constraints; *J. Biomol. NMR* **2007**, *38*, 23-39

[12] Vila, J. A.; Villegas, M. E.; Baldoni, H. A.; Scheraga, H. A. Predicting $^{13}C^{\alpha}$ chemical shifts for validation of protein structures; *J. Biomol. NMR* **2007**, *38*, 221-235

[13] Vila, J. A.; Baldoni, H. A.; Scheraga, H. A. Performance of Density Functional Models to Reproduce Observed $^{13}C^{\alpha}$ Chemical Shifts of Proteins in Solution; *J. Comput. Chem.* **2009**, *30*, 884-892