

Electronic Supplementary Information

A Combined NMR/MD/QM Approach for Structure and Dynamics Elucidations in the Solution State: Pilot Studies Using Tetrapeptides

Abil E. Aliev, Denis Courtier-Murias, Simrath Bhandal and Shen Zhou*

Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

E-mail: a.e.aliev@ucl.ac.uk

Tel: +44 (0) 20 7679 4616; Fax: +44 (0) 20 7679 7463

Contents

NMR Measurements and Calculations	S2
Quantum Mechanical Calculations	S2
MD Calculations	S2
Experimental and Calculated Data	S3
References	S7

NMR MEASUREMENTS AND CALCULATIONS

Gly-Pro-Gly-Gly (GPGG) and Val-Ala-Pro-Gly (VAPG) were purchased from Sigma-Aldrich and were used without further purification. Solution ^1H NMR spectra were recorded on Bruker NMR spectrometers AVANCE400, AVANCE500 and AVANCE600 equipped with a Bruker 5 mm cryoprobe (600 MHz) or room temperature probes (400 and 500 MHz). Data acquisition and processing were performed using standard Bruker XwinNMR (version 2.6) and TopSpin (version 2.1) software. ^1H and ^{13}C chemical shifts in D_2O were calibrated using sodium 3-(trimethylsilyl)propionate (TSP, ^1H 0 ppm) and dioxane (^1H 3.75 ppm, ^{13}C 67.19 ppm).

Two-dimensional NOESY spectra with varying mixing times were recorded in $\text{H}_2\text{O}:\text{D}_2\text{O}$ (20:1). Two different methods were used for estimating distances from each spectrum: the intensity ratio method using the ratio of cross peak and diagonal peak intensities,^[1,2] and a simplified version of the growth rates method,^[3] in which cross peak intensities measured at a single mixing time are used.

The IUPAC labelling of nuclei are used in this work.^[4]

QUANTUM MECHANICAL CALCULATIONS

All quantum mechanical calculations were carried out using *Gaussian 03*.^[5] Geometry optimisations were done using the HF/6-31G(d), HF/6-31+G(d) and MP2/cc-pVDZ levels of theory for tetrapeptides. Additional QM geometry optimisations were undertaken for the C^γ -endo and C^γ -exo conformations of *trans-N*-acetyl-L-proline at the CCSD/cc-pVDZ and MP2/cc-pVTZ levels of theory in order to follow the change of the H^α - $\text{H}^{\beta 3}$ distance, which was used as a reference for determining other interproton distances from NMR spectra and MD simulations. Additional frequency calculations were also undertaken in order to verify that the optimized geometries do indeed correspond to true minima. For the determination of the coefficients of the Karplus equation for the $^3J_{\text{CH}}$ couplings, we used a method similar to that described by Suardíaz *et al.*, with the optimized molecular geometries from the B3LYP 6-31G(d,p) calculations and the $^3J_{\text{CH}}$ couplings from the B3LYP/EPR-III/B3LYP/6-31G(d,p) calculations.^[6] Water solvent effects were introduced in all the quantum mechanical calculations reported in this work via self-consistent reaction field (SCRF) theory calculations using the IEFPCM method,^[7] as implemented in *Gaussian 03*.

MD CALCULATIONS

One molecule of zwitterionic GPGG or VAPG surrounded by >800 water molecules in a cubic box were used in MD simulations with periodic boundary conditions. No counterions were included in MD calculations in order to imitate the experimentally studied systems by NMR as closely as possible. The structures used were energy minimised and the water molecules were allowed to relax (with positionally restrained solute) prior to MD production steps. The OPLS-AA/L force field^[8] (with TIP4P water)^[9] was used.

In simulations using *GROMACS* (version 4.0.2),^[10] neighbour lists were updated every 5th step. A cubic box of $30 \times 30 \times 30 \text{ \AA}^3$ was used for GPGG. A slightly larger box of $35 \times 35 \times 35 \text{ \AA}^3$ was used for VAPG. An integration step of 2 fs was used. Typical cut-off distances were used for the OPLS-AA/L force field.^[8] The PME (particle mesh Ewald)^[11] correction was used for the long range electrostatic interactions. Langevin dynamics, with a reference temperature of 300 K and a weak frictional constant of 10 ps^{-1} was employed.^[12] A Parrinello-Rahman scheme was employed for pressure control at 1 bar using a coupling constant of 0.5 ps.^[13] Prior to production runs, the system was minimized using L-BFGS, steepest-descent and conjugate gradient algorithms. Minimization steps were followed by 3 steps of equilibration. The system was first equilibrated for 40 ps with the positionally restrained solute molecule to allow water molecules to equilibrate around it, followed by an *NVT* molecular dynamics for 10 ps without restraints and then by 100 ps of *NPT* dynamics with an isothermal compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ and a reference pressure of 1.0 bar. Production simulations were performed for 2000 ns (GPGG) and 1092 ns (VAPG) using *NPT* ensemble.

Distances from the MD simulations were calculated in a way similar to that used in NMR measurements described above: (i) internuclear distances (r_i) for pairs of hydrogen atoms were calculated in each MD frame i ; (ii) a quantity equal to r_i^{-6} were calculated as a measure of the expected NOE in each frame, η_i ; (iii) the sum of r_i^{-6} were used as a measure of the expected total NOE over the full length of the MD run (4×10^6 frames over 2 μs MD run in this case); (iv) using $r = 2.4 \text{ \AA}$ as the reference $\text{H}^\alpha\text{-H}^{\beta 3}$ distance, internuclear distances for other proton pairs were calculated using the $\eta \sim r^{-6}$ relationship.

EXPERIMENTAL AND CALCULATED DATA

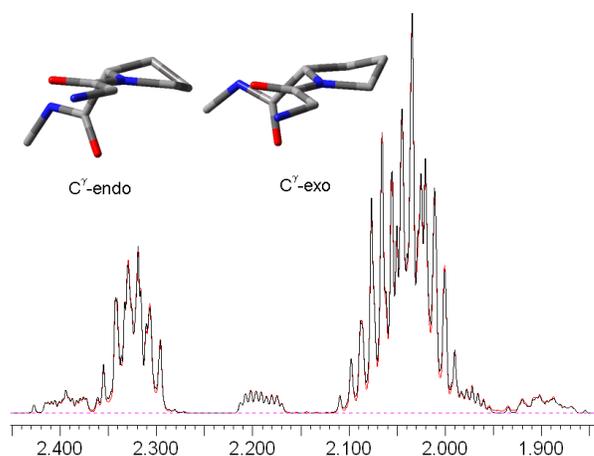


Figure S1. The overlapping view of experimental (red, 298 K, 600 MHz) and calculated (black) ¹H NMR spectra of Pro protons of *cis*- and *trans*-GPGG in D₂O. The low frequency areas of the spectra are shown together with the C^γ-endo and C^γ-exo geometries of the pyrrolidine ring for the major *trans*-rotamer.

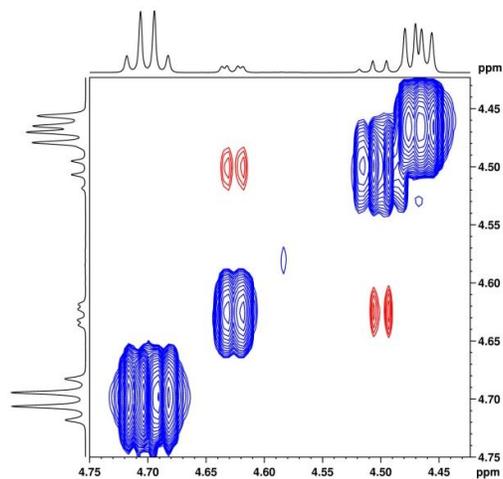
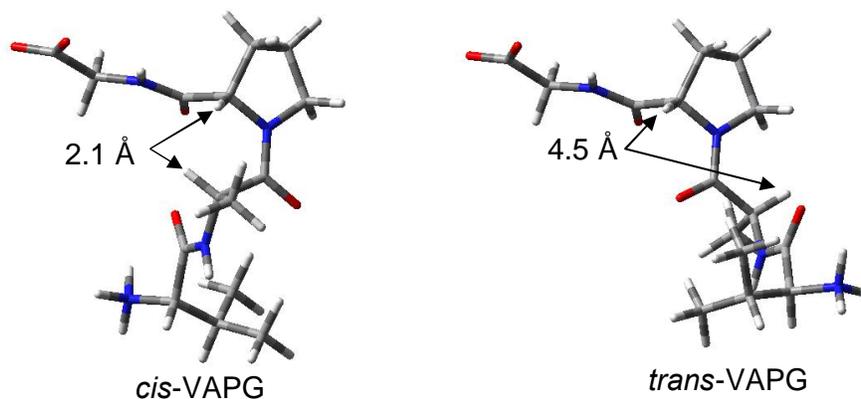


Figure S2. The 2D NOESY spectrum of VAPG in D₂O at 298 K (600 MHz, mixing time 600 ms) showing H^α protons of Pro and Ala. A NOE cross-peak is observed for the signals with lower intensity (corresponding to the *cis*-rotamer about the peptide bond preceding Pro), whereas no such cross peak is observed for the signals with higher intensity (corresponding to the *trans*-rotamer).

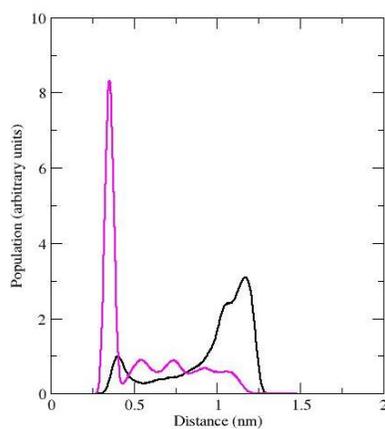


Figure S3. Distribution of conformers as a function of the distance d_{ter} from the MD simulations (OPLS-AA/TIP4P) of GPGG with charged δNH_3^+ and δCOO^- terminal groups (magenta) and with neutral δNH_2 and $-\text{COOH}$ terminal groups (black). A 125 ns long MD run was used in both cases.

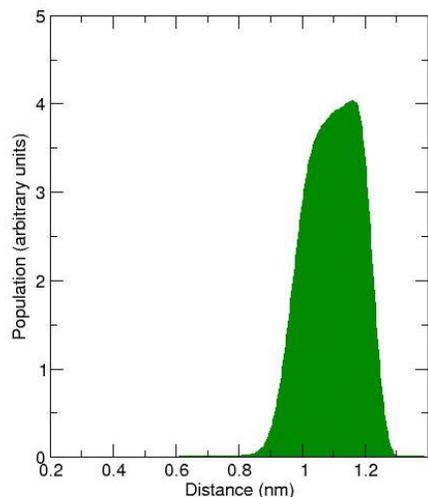


Figure S4. Distribution of conformers as a function of the distance d_{ter} between the terminal N atom of Val and the terminal C atom of Gly in VAPG with charged -NH_3^+ and -COO^- terminal groups from the MD simulations (OPLS-AA/L, 1092 ns).

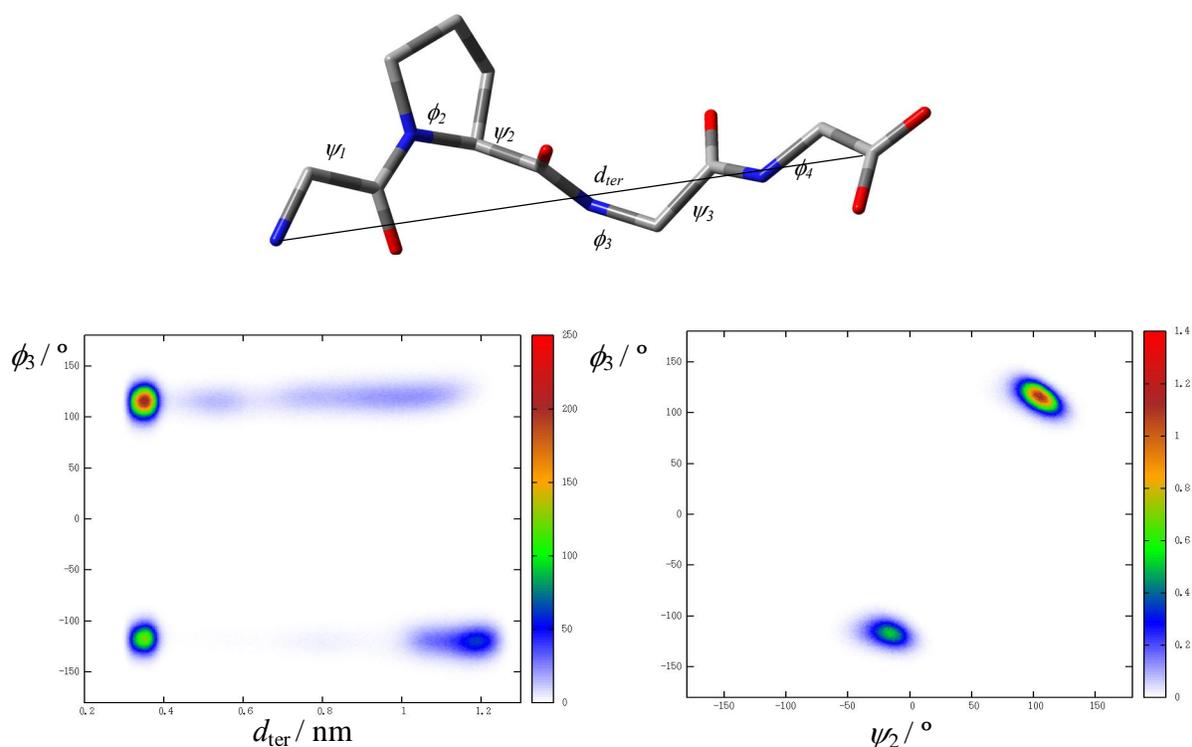
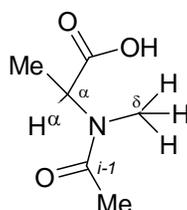


Figure S5. The OPLS-AA/L MD distributions of ϕ_3 vs d_{ter} (left) and ϕ_3 vs ψ_2 (right) for GPGG. Two folded conformations *a* and *b* were identified at $d_{\text{ter}} \sim 0.35$ nm with $\phi_3 \in 115^\circ / \psi_2 \in 115^\circ$ and $\phi_3 \in -115^\circ / \psi_2 \in -15^\circ$, respectively, for further QM calculations.

Table S1. Averaged $^3J_{CH}$ -couplings (in Hz) predicted by MD calculations for carbonyl carbons of GPPG. The RMS deviations from the experimental values are also included.

Pair of coupled nuclei	OPLS-AA/L TIP4P	Experimental values
H ^α (Pro2)δC(Gly1) ^a	1.52	1.3
H ^α (Gly3)δC(Pro2) ^b	3.91	4.1
H ^α (Gly4)δC(Gly3) ^c	3.53	3.7
H ^N (Gly4)δC(Gly4) ^d	0.88	1.1
RMS (Hz)	0.20	-

^a The following Karplus equation was used: $^3J(\text{H}^\alpha, \text{C}^{i-1}) = 4.24 \cos^2 \theta - 1.51 \cos \theta + 0.63$, where θ is the dihedral angle between two coupled nuclei. The coefficients of this Karplus equation were derived from B3LYP/EPR-III//B3LYP/6-31G(d,p) calculations using the following molecule:



^b The following Karplus equation was used: $^3J(\text{H}^\alpha, \text{C}^{i-1}) = -2.75 \cos^2 \phi + 1.08 \cos \phi + 5.06$, where ϕ is the backbone dihedral angle, defined as $\text{C}^{i-1}-\text{N}^i-\text{C}_\alpha^i-\text{C}^i$. The coefficients of this Karplus equation were derived from B3LYP/EPR-III//B3LYP/6-31G(d,p) calculations using triglycine.

^c The following Karplus equation was used: $^3J(\text{H}^\alpha, \text{C}^{i-1}) = -2.66 \cos^2 \phi + 1.24 \cos \phi + 5.10$, where ϕ is the backbone dihedral angle, defined as $\text{C}^{i-1}-\text{N}^i-\text{C}_\alpha^i-\text{C}^i$. The coefficients of this Karplus equation were derived from B3LYP/EPR-III//B3LYP/6-31G(d,p) calculations using diglycine.

^d The following Karplus equation was used: $^3J(\text{H}^N, \text{C}^i) = 5.81 \cos^2 \phi + 1.43 \cos \phi - 0.27$, where ϕ is the backbone dihedral angle, defined as $\text{C}^{i-1}-\text{N}^i-\text{C}_\alpha^i-\text{C}^i$. The coefficients of this Karplus equation were derived from B3LYP/EPR-III//B3LYP/6-31G(d,p) calculations using diglycine.

Table S2. Internuclear distances in VAPG obtained from NMR measurements in D₂O and from HF/6-31+G(d) and MP2/cc-pVDZ calculations in H₂O with the IEFPCM solvation model.

	Distances (Å) by NMR	Calculated Distances (Å) by HF/6-31+G(d)	Calculated Distances (Å) by MP2/cc-pVDZ
NH(A2)-H ^α (V1)	2.5	2.33	2.31
NH(G4)-H ^α (P3)	2.5	2.30	2.32
NH(G4)-H ^α (G4)	2.8	2.82	2.83
H ^α (A2)-H ^β (A2)	2.7	2.67	2.68
H ^α (A2)-H ^{δ3} (P3)	2.4	2.35	2.36
H ^α (A2)-H ^{δ2} (P3)	2.7	2.34	2.34
H ^α (V1)-H ^{γ1γ2} (V1)	3.4	3.54	3.55
RMS (Å)	-	0.16	0.16

References

- [1] M. Reggelin, H. Hoffman, M. Köck, D.F. Mierke, *J. Am. Chem. Soc.* **1992**, *114*, 3272.
- [2] D. Neuhaus, M.P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, New York, 2nd edn, 2000.
- [3] D. Huster, K. Arnold, K. Gawrisch, *J. Phys. Chem. B* **1999**, *103*, 243.
- [4] J.L. Markley, A. Bax, Y. Arata, C.W. Hilbers, R. Kaptein, B.D. Sykes, P.E. Wright, K. Wüthrich, *Pure & Appl. Chem.* **1998**, *70*, 117.
- [5] *Gaussian 03*, Revision D.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, **2004**.
- [6] R. Suardíaz, C. Pérez, J.M. García de la Vega, J. San Fabián, R.H. Contreras. *Chem. Phys. Lett.* **2007**, *442*, 119-123.
- [7] E. Cancès and B. Mennucci, *J. Math. Chem.* **1998**, *23*, 309.
- [8] a) W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236; b) G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, *J. Phys. Chem. B* **2001**, *105*, 6474.
- [9] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [10] a) B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435. b) D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701; c) E. Lindahl, B. Hess, D. van der Spoel, *J. Mol. Model.* **2001**, *7*, 306.
- [11] U. Essman, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577.
- [12] W.F. van Gunsteren, H.J.C. Berendsen, *Mol. Simul.* **1988**, *1*, 173.
- [13] M. Parrinello, A. Rahman. *J. Appl. Phys.* **1981**, *52*, 7182.