ChemistrySELECT

Supporting Information

Examination of the Effect of N-terminal Diproline and Charged Side Chains on the Stabilization of Helical Conformation in Alanine–based Short Peptides: A Molecular Dynamics Study

Bhupesh Goyal,* Kinshuk Raj Srivastava, and Susheel Durani*

Computational Section

Modeling of peptides

The peptides were modeled using *in-house* software package CAPM (Computer Aided Peptide Modeling),[1] capable of handling D-amino acids effectively. *In-house* program PDBmake was used to generate coordinates of CAPM modeled structure.

Molecular dynamics simulations and preparation of equilibrium ensembles

The molecular dynamics simulations were performed with Gromos96 43a1 force field in GROningen MAchine for Chemical Simulations (GROMACS) 3.3.3 in a periodic box with water as explicit-solvent.[2,3] The GROMOS96 force field has been widely used for the conformational analysis of peptides in a number of recent studies.[4] The simulations were performed under NVT condition, *viz.*, fixed number of particles, constant volume, and constant temperature. The non-bonded list cutoff was 1.4 nm with a shift at 0.8 nm. The integration step was 2 fs. Initial velocities were drawn from Maxwellian distribution. The temperature was coupled to an external bath with relaxation time constant of 0.1 ps. The bond lengths were constrained with SHAKE[5] to geometric accuracy 10^{-4}. The electrostatics was treated by the Particle Mesh Ewald (PME)[6,7] method implementing a Coulomb cutoff of 1.4 nm, a Fourier spacing of 0.12 nm, and an interpolation order of 4.

The peptides were modeled in PPII conformation with $\varphi^{L/D} = -/+75°$, $\psi^{L/D} = +/- 145°$. The modeled peptides constrained to centre of the periodic cubic box were soaked in SPC (Simple Point Charge) water model,[8] which was added to 1 atm density at 298 K. First the solute was energy minimized, then the solvent while restraining solute, and finally, both were energy minimized after removing restraint. The molecular dynamics simulations were initialized and

the initial 3 ns trajectory was exempted from the analysis as a pre-equilibration period. The total simulation time was 250 ns for all model peptides. The simulations were performed in multiple runs in parallel and have been merged together to generate the equilibrium in order to avoid the biasness for the starting conformer over the evolution of equilibria. The five different MD simulations of length 50 ns each have been merged together to avoid the biasness for the starting conformer. The trajectories were sampled at 10 ps interval for all model peptides.

Analysis and characterization of macrostate, polypeptide microstates

Conformational microstates were clustered in Cartesian space with root-mean-square deviation (RMSD) cutoff of 0.15 nm over backbone atoms (N, Cβ, Cα, C), giving microstates diminishing in population, viz., diminishing thermodynamic stability. The clustering was performed in GROMACS package according to Daura *et al.* algorithm.[9] This procedure is widely used for conformational clustering in a number of recent studies.[10] In this procedure, conformer with largest number of neighbors was defined as central member of the first cluster or the most-populated microstate. All members of this microstate were removed from the ensemble, and the procedure was iterated until all the remaining conformers in the ensemble were assigned to specific microstates, diminishing in population. We considered the most-populated first microstate as the ordered state and evaluated its stability with regard to remaining microstates considered as unordered state. The most-populated first microstate considered as the ordered state because it has maximum thermodynamic stability compared to other microstates. The radius-of-gyration (R_g) was computed using the g_gyrate utility in GROMACS. The percentage occupancy of the macrostate in α, β and PPII basins was evaluated computed using *in-house* program. The definition of φ, ψ basins in Ramachandran diagram that were adopted in the present study is as follows: α ($^{L/D}\varphi$ = –/+ 20 to –/+ 100, $^{L/D}\psi$ = –/+ 20 to –/+ 80), β ($^{L/D}\varphi$ = –/+ 90 to –/+ 170, $^{L/D}\psi$ = +/– 80 to +/– 180), and PPII ($^{L/D}\varphi$ = –

/+ 30 to –/+ 90, $^{L/D}\psi$ = +/– 80 to +/– 170). The percentage population of specific φ, ψ basins was evaluated using *in-house* written scripts. The hydrogen bonds were enumerated to 0.35 nm distance (N−O) and 30° angle (H−N−O) cutoff. The hydrogen bonds are defined as short-ranged (SR; i→i ± 2), medium-ranged (MR; i→i ± 3, i→i ± 4) and long-ranged (LR; i→i ± 5, i→i ± ≥ 6) according to sequence separation between donor and acceptor residue. The radial distribution functions of specific solvent atoms were calculated over the most-populated microstate in each ensemble using g_rdf utility in GROMACS.

References

[1] S. Durani, *Acc. Chem. Res.* **2008**, *41*, 1301–1308.

[2] E. Lindahl, B. Hess, D. van der Spoel, *J. Mol. Mod.* **2001**, *7*, 306–317.

[3] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual, User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, **1996**, pp. 1–1042.

[4] a) A. Kumar, S. Srivastava, S. Tripathi, S. K. Singh, S. Srikrishna, A. Sharma, *J. Biomol. Struct. Dyn.* **2016**, *34*, 1252–1263; b) L. Baweja, K. Balamurugan, V. Subramanian, A. Dhawan, *J. Mol. Graph. Model.* **2015**, *61*, 175–185; c) V. Minicozzi, R. Chiaraluce, V. Consalvi, C. Giordano, C. Narcisi, P. Punzi, G. C. Rossi, S. Morante, *J. Biol. Chem.* **2014**, *289*, 11242–11252; d) I. Autiero, E. Langella, M. Saviano, *Mol. BioSys.* **2013**, *9*, 2835–2841; e) S. T. Ngo, M. S. Li, *J. Phys. Chem. B* **2012**, *116*, 10165−10175; f) M. H. Viet, S. T. Ngo, N. S. Lam, M. S. Li, *J. Phys. Chem. B* **2011**, *115*, 7433–7446; g) C. Yang, J. Li, Y. Li, X. Zhu, *J. Mol. Struct. THEOCHEM* **2009**, *895*, 1–8; h) Y. Sun, Z. Qian, G. Wei, *Phys. Chem. Chem. Phys*. **2016**, ***18***, 12582–12591.

[5] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phy.* **1977**, *23*, 327–341.

[6] T. Darden, D. York, L. G. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[7] U. Essmann, M. L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.

[8] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, *Interaction models for water in relation to protein hydration. Intermolecular Forces*; (Ed.: B, Pullman), Reidel Publishing Company: Dordrecht, The Netherlands, **1981**, pp. 331–342.

[9] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem.* **1999**, *111*, 249–253; *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240.

[10] a) J. K. Marzinek, R. Lakshminarayanan, E. Goh, R. G. Huber, S. Panzade, C. Verma, P. J. Bond, *Sci. Rep.* **2016**, *6*, 19160; b) G. Rossetti, F. Musiani, E. Abad, D. Dibenedetto, H. Mouhib, C. O. Fernandez, P. Carloni, *Phys. Chem. Chem. Phys.* **2016**, *18*, 5702–5706; c) P. Gupta, B. Liu, D. Klepacki, V. Gupta, K. Schulten, A. S. Mankin, N. Vázquez-Laslop, *Nat. Chem. Biol.* **2016**, *12*, 153–158; d) G. O. Rutter, A. H. Brown, D. Quigley, T. R. Walsh, M. P. Allen, *Phys. Chem. Chem. Phys.* **2015**, *17*, 31741–31749; e) B. Goyal, K. Patel, K. R. Srivastava, S. Durani, *RSC Adv.* **2015**, *5*, 105400–105408; f) Y. Sun, Z. Qian, C. Guo, G. Wei, *Biomacromolecules* **2015**, *16*, 2940–2949; g) A. K. Somavarapu, K. P. Kepp, *ChemPhysChem* **2015**, *16*, 3278–3289; h) S. R. Gerben, J. A. Lemkul, A. M. Brown, D. R. Bevan, *J. Biomol. Struct. Dyn.* **2014**, *32*, 1817–1832; i) C. Merten, F. Li, K. Bravo-Rodriguez, E. Sanchez-Garcia, Y. Xu, W. Sander, *Phys. Chem. Chem. Phys.* **2014**, *16*, 5627–5633; j) F. Godschalk, S. Genheden, P. Söderhjelm, U. Ryde, *Phys. Chem. Chem. Phys.* **2013**, *15*, 7731–7739.