John von Neumann Institute for Computing

Computer Simulation of Biomolecular Systems: Where Do We Stand?

W. F. van Gunsteren, D. P. Geerke

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 49-56, 2008.

http://www.fz-juelich.de/nic-series/volume40

Computer Simulation of Biomolecular Systems: Where Do We Stand?

Wilfred F. van Gunsteren and Daan P. Geerke

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH,
8093 Zurich, Switzerland
E-mail: wfvgn@igc.phys.chem.ethz.ch

The four major aspects that determine the quality of the ensemble of molecular conformations as obtained from biomolecular simulation are reviewed and illustrated with examples.

1 Introduction

Over the past three decades, simulation of the dynamics of biomolecular systems at the atomic level has developed from short-time simulations of simple molecular models [1, 2] to orders of magnitude larger simulations based on detailed and much more accurate molecular models [3]. The improved accuracy has turned molecular dynamics (MD) simulation into a standard method for an atomic interpretation of experimental data on biomolecular systems [4]. Yet, much progress is still to be made in order to use MD simulation to accurately predict various properties of biomolecular systems [5].

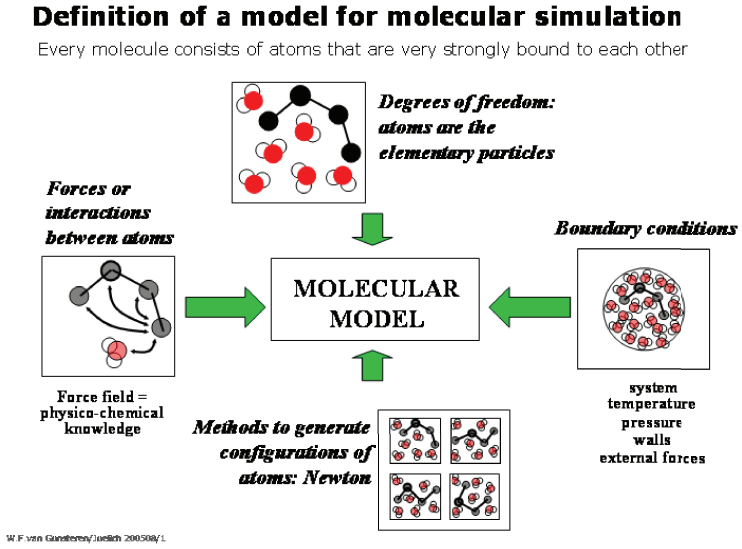

Figure 1.

Figure 1 illustrates the four choices to be made when defining a molecular model for molecular simulation: (1) which degrees of freedom are to be explicitly simulated; (2) how are the forces governing the motion along these degrees of freedom calculated in a necessarily approximative manner; (3) how is the motion of the system propagated in time such that the relevant configurational space is widely and sufficiently sampled; (4) how are the spatial and thermodynamic boundary conditions imposed upon the motion of the system. In this short paper we illustrate the state of the art with respect to these four aspects of modelling using examples from our own work.

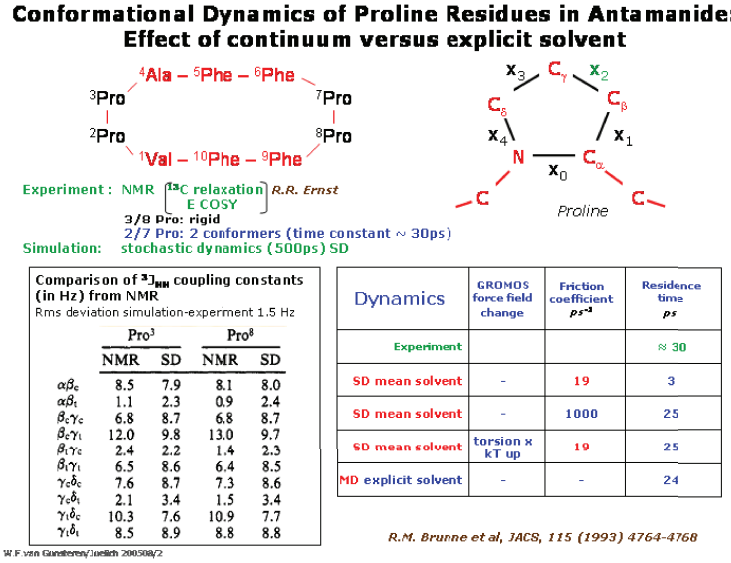

Figure 2.

2 Choice of Degrees of Freedom

In Figure 2, the importance of explicitly including solvent degrees of freedom is illustrated. It compares the puckering residence time of the χ_2 torsional angle of residues ^{2}Pro and ^{7}Pro in the cyclic polypeptide antamanide as obtained from experiment [6] with that as obtained from simulations using either a mean implicit solvent [7] or explicit water solvent molecules [8]. Using stochastic dynamics (SD) simulation with a friction coefficient of 19 ps^{-1}, typical for water at room temperature and pressure, the puckering rate is ten times too high compared to experiment. Using a mean solvation model it could only be reduced by either using an artificially high friction coefficient of 1000 ps^{-1} or by artificially increasing the torsional barrier by about 2.5 kJ mol^{-1}. These unjustified changes in the molecular model can, however, be avoided by explicitly simulating the water degrees of freedom, which results in the correct puckering rate (see Figure 2).

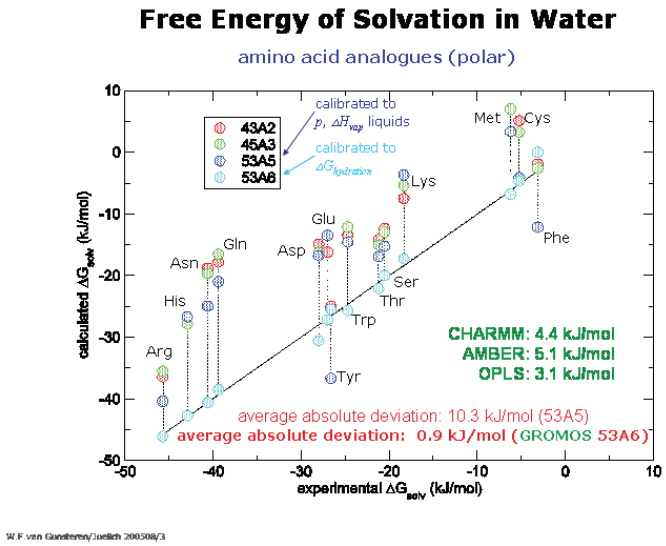

Figure 3.

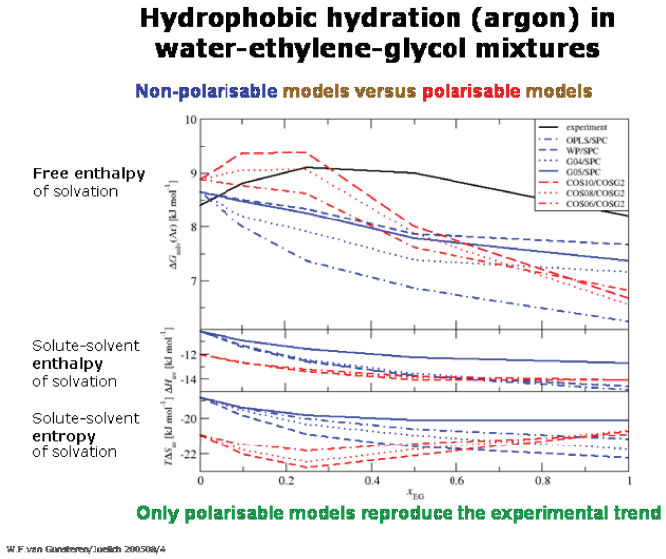

Figure 4.

3 Choice of Interatomic Interactions

In Figure 3, it is illustrated that the free energy of aqueous solvation of the side chains of polar amino acid residues is insufficiently negative for some widely used biomolecu-

lar force fields [9], which would lead to incorrect partitioning between polar and apolar solvents and would over-stabilize folded protein structure. Therefore, more recent force fields, such as the GROMOS 53A6 one [10], lead to a better description of the folding equilibrium [11].

Figure 4 illustrates that inclusion of atomic polarisability will be essential to reach an improved level of accuracy, i.e. beyond 1 kJ mol $^{-1}$. The experimentally observed non-linear behaviour of the free enthalpy of solvation of argon in water-ethylene glycol mixtures as function of the ethylene glycol mole fraction is only reproduced in MD simulations using polarisable molecular models (COS models), whereas a more or less linear behaviour is obtained when using non-polarisable models [12]. Calculation of the solute-solvent entropy of solvation shows that the non-linearity is an entropic effect, which cannot be modelled using a mean or continuum solvent model.

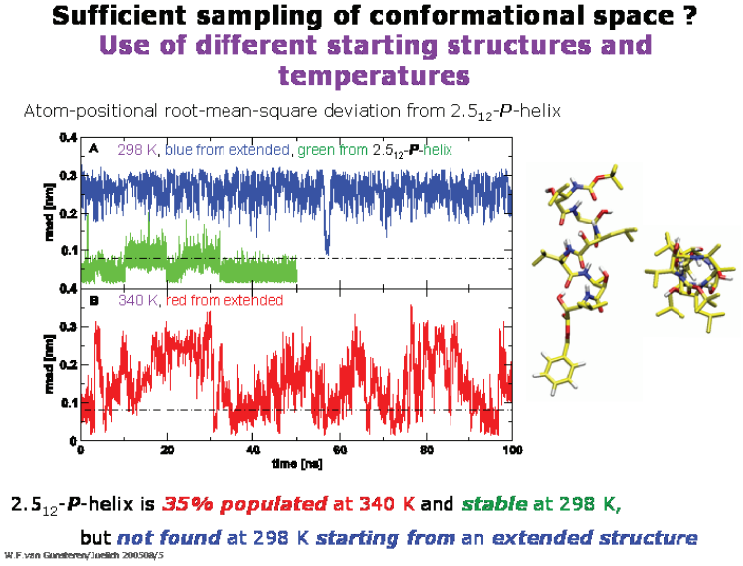

Figure 5.

4 Sufficient Searching and Sampling of Conformational Space

In Figure 5 it is illustrated that even 100 ns of MD simulation of an eight-residue β-peptide in methanol at room temperature may be insufficient to find the most stable 2.5$_{12}$-P-helical fold [13]. Starting the MD simulation at 298 K from this helical structure shows a low root-mean-square deviation (rmsd) from this fold, whereas starting at 298 K from an extended conformation the helical structure is not populated. At 340 K the sampling of conformational space is much widened: a large number of (un)folding events is observed and the helical conformation is present 35% of the time. This example illustrates the need

for search and sampling enhancement techniques, of which Figure 6 classifies the most important ones [14].

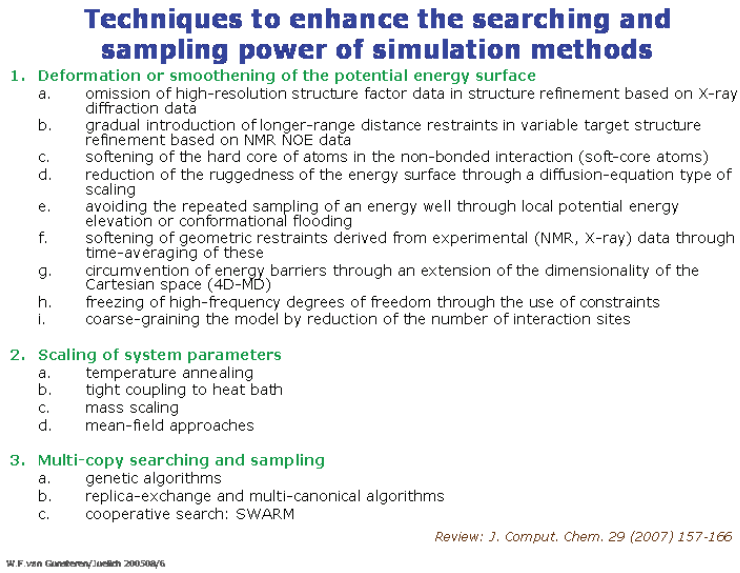

Figure 6.

5 Choice of the Appropriate Thermodynamic State Point

Figure 7 illustrates the effect of pH on the folding equilibrium of a seven residue β-peptide in methanol [15]. The only protonisable groups of this peptide with aliphatic side chains are the amino- and carboxy-termini. The backbone atom-positional root-mean-square deviation of the MD trajectory structures from the most stable (both computationally and experimentally) 3_{14}-helical fold shows that only if the protonation state corresponds to the experimental pH, the helical fold is the most populated one.

6 Conclusion

We have briefly illustrated that the following factors are essential to obtain a high quality ensemble of molecular conformations in a molecular simulation.

1. Inclusion of the relevant degrees of freedom: solvent and co-solvents.
2. Use of a thermodynamically calibrated force field, with a solvent model that is compatible with the solute one, and possibly inclusion of polarisability.
3. Sufficient and Boltzmann-weighted sampling of conformational space.

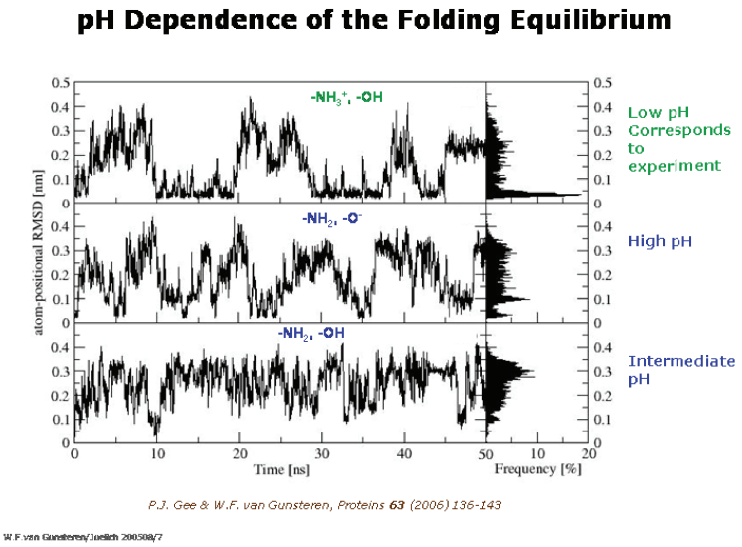

Figure 7.

4. Use of the appropriate (experimental) thermodynamic state point and spatial boundary conditions: temperature, pressure, pH, ionic strength, co-solvents, etc.

Acknowledgments

Financial support from the National Centre of Competence in Research (NCCR) in Structural Biology and grant number 200021-109227 of the Swiss National Science Foundation is gratefully acknowledged.

References

1. A. Warshel, M. Levitt, *Theoretical Studies of Enzymatic Reactions: Dielectric Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme*, J. Mol. Biol. **103**, 227–249, 1976.
2. J. A. McCammon, B. R. Gelin, M. Karplus, *Dynamics of Folded Proteins*, Nature **267**, 585–590, 1977.
3. W. F. van Gunsteren, J. Dolenc, *Biomolecular Simulation: Historical Picture and Future Perspectives*, Biochem. Soc. Trans. **36**, 11–15, 2008.
4. W. F. van Gunsteren, J. Dolenc, A. E. Mark, *Molecular Simulation as an Aid to Experimentalists*, Curr. Opin. Struct. Biology **18**, 149–153, 2008.
5. W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glaettli, P. H. Huenenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. van der Vegt, H. B. Yu, *Biomolec-*

ular Modelling: Goals, Problems, Perspectives, Angew. Chem. **118**, 4168–4198, 2006; Angew. Chem. Int. Ed. **45**, 4064–4092, 2006.
6. Z. L. Madi, C. Griesinger, R. R. Ernst, *Conformational Dynamics of Proline Residues in Antamanide. J Coupling Analysis of Strongly Coupled Spin Systems Based on E.COSY Spectra*, J. Am. Chem. Soc. **112**, 2908–2914, 1990.
7. R. M. Brunne, W. F. van Gunsteren, R. Brueschweiler, R. R. Ernst, *Molecular Dynamics Simulation of the Proline Conformational Equilibrium and Dynamics in Antamanide Using the GROMOS Force Field*, J. Am. Chem. Soc. **115**, 4764–4768, 1993.
8. J. W. Peng, C. A. Schiffer, P. Xu, W. F. van Gunsteren, R. R. Ernst, *Investigations of Peptide Hydration Using NMR and Molecular Dynamics Simulations: A study of Effects of Water on the Conformation and Dynamics of Antamanide*, J. Biomol. NMR **8**, 453–476, 1996.
9. M. R. Shirts, J. W. Pitera, W. C. Swope, V. S. Pande, *Extremely Precise Free Energy Calculations of Amino Acid Side Chain Analogs: Comparison of Common Molecular Mechanics Force Fields for Proteins*, J. Chem. Phys. **119**, 5740–5761, 2003.
10. C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6*, J. Comput. Chem. **25**, 1656-1676, 2004.
11. C. Oostenbrink, T. A. Soares, N. F. A. van der Vegt, W. F. van Gunsteren, *Validation of the 53A6 GROMOS force field*, Eur. Biophys J. **34**, 273–284, 2005.
12. D. P. Geerke, W. F. van Gunsteren, *The performance of non-polarizable and polarizable force-field parameter sets for ethylene glycol in molecular dynamics simulation of the pure liquid and its aqueous mixtures*, Mol. Phys. **105**, 1861–1881, 2007.
13. A. Glaettli, W. F. van Gunsteren, *Are NMR-derived model structures for peptides representative for the ensemble of structures adopted in solution? Probing the fourth helical secondary structure of β-peptides by molecular dynamics simulation*, Angew. Chem. **116**, 6472-6476, 2004; Angew. Chem. Int. Ed. Engl. **43**, 6312-6316, 2004.
14. M. Christen, W. F. van Gunsteren, *On searching in, sampling of, and dynamically moving through conformational space of biomolecular systems: a review*, J. Comput. Chem. **29**, 157–166, 2007.
15. P. J. Gee, W. F. van Gunsteren, *Terminal-group effects on the folding behaviour of selected beta-peptides*, Proteins **63**, 136–143, 2006.