Journal of Computer-Aided Molecular Design, 2 (1988) 217–221
ESCOM

J-CAMD 042

Dynamic simulation as an essential tool in molecular modeling

H.J.C. Berendsen

Laboratory of Physical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

Key words: Molecular dynamics; Dynamic simulation; Molecular modeling; Modeling requirements

INTRODUCTION

An attempt is made to categorize the purposes and methods for molecular modeling, followed by a short summary of methods for molecular dynamics simulation and its applications to complex molecular systems. The power of dynamic simulation methods to answer relevant questions concerning molecular behaviour is emphasized, including the determination of free energies in nonlinear systems.

A SURVEY OF MODELING PURPOSES

The requirements for computer modeling of molecular systems depend on the purpose that is to be served by the modeling exercise. In order of increasing predictive power, but also increasing complexity, the purpose could be*

(a) Visualisation of a known structure;

(b) Modification of a known structure by operator manipulation;

(c) Computation of properties of a known or modified structure;

(d) Generation of the nearest structure with local minimum energy;

(e) Generation of a (large, and preferably complete) set of structures with local minimum energies;

(f) Generation of the structure with global minimum energy;

*The term *structure* stands for the combination of *topology* (number and type of atoms and their covalent connectivity), *configuration* (stereospecific arrangement), *conformation* (relatively stable set of spatial arrangements) and *3D structure* (actual atomic coordinates). A conformation comprises an ensemble of 3D structures, representative for the conformation at a given temperature and in a defined environment.

(g) Computation of static mechanical properties of a given conformation in equilibrium under specified external conditions as temperature, pressure, solvent, etc.;

(h) As *(g)*, but including dynamical properties;

(i) As *(g)*, but including thermodynamic properties as free energy and entropy;

(k) Computation of thermodynamic properties of reversible processes, such as binding properties, equilibrium constants, partition coefficients, activities, etc.;

(l) Computation of properties in irreversible processes, excluding changes in covalent structure, such as transport properties and nonlinear effects; and/or

(m) Computation of reaction paths, involving changes in covalency.

The modeling methods to be used for the purposes listed here, also increase in complexity going down the list. Purposes *(a)* and *(b)* are purely technical and will not be considered here. Purposes *(c)* and *(d)* belong to the realm of *molecular mechanics*; they require a reliable description of the potential energy of a molecular system as a function of all atomic coordinates. Efficient methods are available to generate the nearest local energy minimum in a multidimensional space. All subsequent purposes have the same requirements with respect to an adequate description of the potential field; this is central to any type of molecular modeling beyond pure graphic visualisation!

Purposes *(e)* and *(f)* require, in addition, efficient multidimensional search and optimisation algorithms for strongly nonlinear functions: not a trivial matter! The general solution of purpose *(f)* may not even exist. Note that for diffusive systems, as liquids and solutions, it is not a relevant question to ask for a survey of local minima, while a global minimum does not exist in such cases.

Purposes *(g)*, *(h)*, *(i)* and *(k)* allow the determination of *all* properties of interest of equilibrium systems. They require the generation of at least an *ensemble* of 3D structures, representative for the given environmental conditions, i.e., occurring with their Boltzmann probability. Such ensembles can be generated either by *Monte Carlo* or *molecular dynamics* simulations. This suffices for purposes *(g)* and, if possible, *(i)*, and allows the determination of average structures, energies, etc., with their fluctuations, and, if possible, entropies and free energies. Ensembles for purpose *(h)* can only be generated by molecular dynamics simulations; they yield additional information about internal motions, stabilities of substructures, spectroscopic properties and transport properties as diffusion.

Purpose *(i)*, the computation of thermodynamic properties from an equilibrium ensemble, can only be realised in special cases, and then only with difficulty. Most practical questions regarding free energies, including the construction of phase diagrams, relate to purpose *(k)*. They require the simulation, either by Monte Carlo or molecular dynamics, of a series of ensembles of intermediate states along the reversible path of the process of interest. The recently developed method of integration over a continuous molecular dynamics simulation in which the reversible path is slowly realised, is very promising for practical applications.

Purpose *(l)* can only be accomplished by molecular dynamics simulations, for which non-equilibrium versions have been developed. This represents an, as yet, almost unexplored range of inter-

esting problems, including strongly nonlinear systems supporting stable dissipative structures and solitons.

Purpose *(m)* requires the intermixing of dynamic simulations and quantum mechanical treatments of a reactive subsystem. Methods for quantum dynamical simulations are available, but seem prohibitive in their computational requirement for practical applications in reactive systems. Simplifications are being developed, with static quantum mechanical computations incorporated into classical simulations. Slow events are approached by modifications of classical transition state theory in which molecular dynamics details are incorporated. Rapid developments are expected in this field in the near future.

APPLICATIONS OF MOLECULAR DYNAMICS TO COMPLEX SYSTEMS

As is apparent from the survey given above, for many practical applications it is necessary to generate a full dynamic trajectory that can be considered representative for the behaviour of the system of interest. Such molecular dynamics simulations [1–8] are quite straightforward (once a reliable potential function is available), but they are extremely computer-intensive.

Simulation of atomic motion in complex macromolecular systems by molecular dynamics is only 12 years old: the first simulation of a small protein (BPTI) was carried out by Andrew McCammon at a CECAM workshop in 1976 [9, 10]. The development of computer capabilities over the past few decennia has been such that every six years roughly a factor of ten increase in computer power has become available [11]. This trend is expected to continue for some time to come in view of the recent development of cheap and powerful parallel computer architectures. Together with methodological developments in the techniques of efficient simulation, the available computer power now allows simulations of complex chemical systems involving tens of thousands of degrees of freedom. Sophisticated applications, such as the prediction of binding properties of substrates or inhibitors to native or modified proteins, are now within reach, while predictions of tertiary structure from primary sequences and prediction of heat and solvent stability of modified macromolecules are within sight.

Methodological improvements have included the use of constraints for bond lengths and angles [12–16] enabling a four-fold increase of the allowed time step, stochastic dynamics [17–19] enabling the omission of solvent degrees of freedom in the dynamical behaviour of solutes, constant temperature and pressure algorithms [20, 21] enabling simulation under more natural and practical conditions, and the development of practical methods to compute entropies and free energies. Since the latter determine many properties of interest such as solubilities, partition coefficients, binding and association constants, conformational stabilities, phase boundaries, and reaction kinetics through the thermodynamic properties of activated complexes, they are of prime importance for the future application of modeling methods.

The problem with the computation of entropies and free energies is that these quantities are integrals over phase space rather than the much more accessible ensemble averages as energy, pressure and structure. Although for some special cases entropies can be computed from estimates of the distribution function in phase space [22–24], in general this is not possible. The most promising method is *thermodynamic integration* [25–39], providing a free energy difference between two states by computing the reversible work done in the process of transforming one state into the other. The transformation must be carried out over a reversible path, either by slowly changing

parameters, or by simulating a number of intermediate steps. Free energy differences between two states can often be obtained from the difference of two non-physical processes: For example, the binding constant of a substrate S to a modified enzyme E' can be found by considering the following two processes: *(a)* change E into E', *(b)* change ES into E'S. The difference of the free energies of these two alchemic processes yields the ratio of the required binding constant to the (known) binding constant of S to the native enzyme E. Computations of this kind can be of such importance in drug design and protein engineering that a revolutionary development can be expected.

Surveys describing recent work from our laboratory can be found in references [40–44]. These include applications of molecular dynamics simulations to proteins in water [40], to DNA [41], to bilayer membranes [42, 44], to prediction of binding constants [43] and to structure refinement using 2D NMR [44].

REFERENCES

1 Hockney, R.W. and Eastwood, J.W., Computer Simulation Using Particles, McGraw-Hill, New York, NY, 1981.
2 Van Gunsteren, W.F. and Berendsen, H.J.C., Biochem. Soc. Trans., 10 (1982) 301–305.
3 Berendsen, H.J.C. and van Gunsteren, W.F., In Barnes A.J., Orville-Thomas, W.J. and Yarwood, J. (Eds.) Molecular Liquids, Dynamics and Interactions, NATO/ASI Series, C135, Reidel, Dordrecht, 1984, pp. 475–500.
4 Hermans, J. (Ed.) Molecular Dynamics and Protein Structure, Polycrystal Book Service, P.O. Box 27, Western Springs, IL 60558, 1985.
5 Van Gunsteren, W.F. and Berendsen, H.J.C., In Hermans, J. (Ed.) Molecular Dynamics and Protein Structure, ibid., pp. 5–14.
6 Ciccotti, G. and Hoover, W.G. (Eds.) Molecular Dynamics Simulation of Statistical-Mechanical Systems, Proc. Int. Enrico Fermi School of Physics, Course 97, Elsevier, Amsterdam, 1986.
7 Ciccotti, G., Frenkel, D. and McDonald, J.R. (Eds.) Simulation of Liquids and Solids: Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics, Elsevier, Amsterdam, 1987.
8 Allen, M.P. and Tildesley, D.J., Computer Simulation of Liquids, Clarendon Press, Oxford, 1987.
9 McCammon, J.A., In Berendsen, H.J.C. (Ed.) Models for Protein Dynamics, Report CECAM Workshop May–July 1976, CECAM, Orsay, 1976.
10 McCammon, J.A., Gelin, B.R. and Karplus, M., Nature, 267 (1977) 585–590.
11 Berendsen, H.J.C., In Barnes, A.J. et al. (Eds.) Molecular Liquids, Dynamics and Interactions, NATO/ASI Series, C135 Reidel, Dordrecht, 1984, pp. 561–564.
12 Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C., J. Comput. Phys., 23 (1977) 327–341.
13 Van Gunsteren, W.F. and Berendsen, H.J.C., Mol. Phys., 34 (1977) 1311–1327.
14 Van Gunsteren, W.F. and Karplus, M., J. Comput. Chem., 1 (1980) 266–274.
15 Van Gunsteren, W.F. and Karplus, M., Biochemistry, 21 (1982) 2259–2274.
16 Ryckaert, J.P., Mol. Phys., 55 (1985) 549–556.
17 Ermak, D.L., and Buckholtz, H., J. Comput. Phys., 35 (1980) 169–182.
18 Van Gunsteren, W.F., Berendsen H.J.C. and Rullmann, J.A.C., Mol. Phys., 44 (1981) 69–95.
19 Van Gunsteren, W.F. and Berendsen, H.J.C., Mol. Simulation, 1 (1988) 173–185.
20 Nosé, S., Mol. Phys., 52 (1984) 255–268.
21 Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F., DiNola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684–3690.
22 Karplus, M. and Kushic, J.N., Macromolecules, 14 (1981) 325–332.
23 Edholm, O. and Berendsen, H.J.C., Mol. Phys., 51 (1984) 1011–1028.
24 DiNola, A., Berendsen, H.J.C. and Edholm, O., Macromolecules, 17 (1984) 2044–2050.
25 Kirkwood, J.G., J. Chem. Phys., 3 (1935) 300–313.
26 Postma, J.P.M., Berendsen, H.J.C. and Haak, J.R., J. Chem. Soc. Faraday Trans., 17 (1982) 55–67.

27 Postma, J.P.M., MD of H_2O, A Molecular Dynamics Study, Ph. D. Dissertation, University of Groningen, Groningen, The Netherlands, 1985.
28 Tembe, B.L. and McCammon, J.A., Comput. Chem., 8 (1984) 281–283.
29 Berendsen, H.J.C., Postma, J.P.M. and Van Gunsteren, W.F., In Hermans, J. (Ed.) Molecular Dynamics and Protein Structure, Polycrystal Book Service, P.O. Box 27, Western Springs, IL 60558, 1985, pp. 43–46.
30 Jorgensen, W.L., and Ravimohan, C., J. Chem. Phys., 83 (1985) 3050–3054.
31 Straatsma, T.P., Berendsen, H.J.C. and Postma, J.P.M., J. Chem. Phys., 85 (1986) 6720–6727.
32 Straatsma, T.P. and Berendsen, H.J.C., J. Chem. Phys., 1988, accepted for publication.
33 Straatsma, T.P., Free Energy Evalution by Molecular Dynamics Simulation, Ph.D. Dissertation, University of Groningen, Groningen, The Netherlands, 1987.
34 Van Gunsteren, W.F. and Berendsen, H.J.C., J. Comput.-Aided Mol. Design, 1 (1987) 171–176.
35 Mezei, M., J. Chem. Phys., 86 (1987) 7084–7088.
36 Singh, U.C., Brown, F.K., Bash, P.A. and Kollman, P.A., J. Am. Chem. Soc., 109 (1987) 1607–1614.
37 Bash, P.A., Singh, U.C., Langridge, R. and Kollman, P.A., Science, 236 (1987) 564–568.
38 Rao, S.N., Singh, U.C., Bash, P.A. and Kollman, P.A., Nature, 328 (1987) 551–554.
39 Van Gunsteren, W.F., In Van Gunsteren, W.F. and Weiner, P. (Eds.) Computer simulation of biomolecular systems: Theoretical and experimental applications, ESCOM, Leiden, in press.
40 Berendsen, H.J.C., van Gunsteren, W.F., Zwinderman, H.R.J. and Geurtsen, R.G., Ann. N.Y Acad. Sci., 482 (1986) 269–285.
41 Van Gunsteren, W.F., Berendsen, H.J.C., Geurtsen, R.G. and Zwinderman, H.R.J. Ann. N.Y. Acad. Sci., 482 (1986) 287–303.
42 Berendsen, H.J.C., In Ciccotti, G. and Hoover, W.G. (Eds.) Molecular Dynamics Simulation of Statistical Mechanical Systems, Proc. Int. Enrico Fermi School of Physics, Course 97, Elsevier, Amsterdam, 1986, pp. 496–519.
43 Berendsen, H.J.C., Comput. Phys. Comm., 44 (1987) 233–242.
44 Berendsen, H.J.C., van Gunsteren, W.F., Egberts, E. and de Vlieg, J., In Jensen, K.F. and Truhlar, D.G. (Eds.) Supercomputer Research in Chemistry and Chemical Engineering, ACS Symposium Series 353, American Chemical Society, Washington, DC, 1987, pp. 106–122.