VOLUME 35 NUMBER 6

ACCOUNTS of CHEMICAL RESEARCH®

JUNE 2002

GUEST EDITORIAL

Molecular Dynamics Simulations of Biomolecules

The long-range goal of molecular approaches to biology is to describe living systems in terms of chemistry and physics. Great progress has been made during the past 70 years in applying the quantum mechanical equations representing the underlying physical laws to chemical problems involving the structures and reactions of small molecules. (This work was recognized by the Nobel Prize in Chemistry awarded to Walter Kohn and John Pople in 1998.) Computational studies of thermodynamics and dynamics of mesoscopic systems of biological interest have been attempted only more recently. Classical mechanics is adequate for describing most of the properties of these systems, and the molecular dynamics simulation method is the theoretical approach which is best suited for such studies.

It is now 25 years since the first molecular dynamics simulation of a macromolecule of biological interest was published.[1] The simulation concerned the bovine pancreatic trypsin inhibitor (BPTI), which has served as the "hydrogen molecule" of protein dynamics because of its small size, high stability, and relatively accurate X-ray structure, available in 1975;[2] interestingly, its physiological functions remain unknown. Although this simulation was done in vacuum with a crude molecular mechanics potential and lasted for only 9.2 ps, the results were instrumental in replacing our view of proteins as relatively rigid structures (In 1981, Sir D. L. Phillips commented: "Brass models of DNA and a variety of proteins dominated the scene and much of the thinking".[3]) with the realization that they were dynamic systems, whose internal motions play a functional role. Of course, there were already experimental data (such as the pioneering hydrogen exchange studies of Linderstrom-Lang and his co-worker),[4,5] pointing in this direction. It is now recognized that the X-ray structure of a protein provides the average atomic positions, but the atoms exhibit fluid-like motions of sizable amplitudes about these averages. The new understanding of protein dynamics subsumed the static picture in that the average positions are still useful for the discussion of many aspects of biomolecular function in the language of structural chemistry. However, the recognition of the importance of fluctuations opened the way for more sophisticated and accurate interpretations of functional properties.

The conceptual changes resulting from the early studies make one marvel at how much of great interest could be learned with so little — such poor potentials, such small systems, so little computer time. This is, of course, one of the great benefits of taking the initial, somewhat faltering steps in a new field where the questions are qualitative rather than quantitative and any insights, even if crude, are better than none at all. Subsequent applications have been concerned with more detailed interpretations and predictions of phenomena that require the use of improved methods. Simulations based on ever more refined potentials and the longer runs required for improved statistics are becoming possible for increasingly complex systems as a result of the progress in the available computers, which seem to continue to increase in speed by a factor of 2 or so every 18 months, in accord with Moore's law.

This special issue of *Accounts of Chemical Research* reminds me that when we first began this work at Harvard, both chemists and biologists expressed their feeling that the calculations were a waste of time — my chemistry colleagues, who were rather a conservative lot, felt that detailed treatments of such complex systems were impossible, and biologists believed that even if they were possible, they would add little, if anything, of importance to our knowledge. Experience has proved the contrary. There has been a very rapid development in molecular dynamics simulations that is providing a basis for a more

10.1021/ar020082r CCC: $22.00
Published on Web 06/18/2002

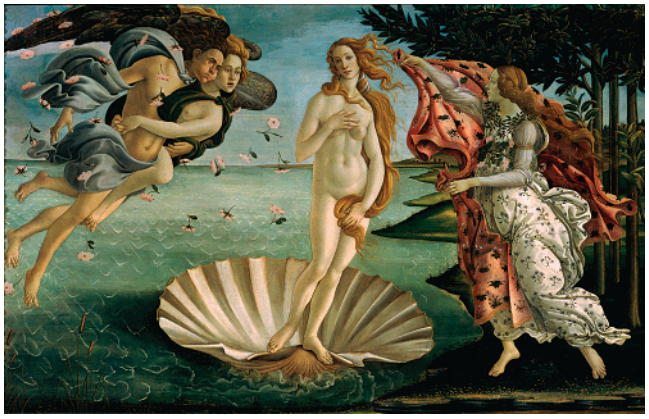

FIGURE 1. "The Birth of Venus", 1484, by Sandro Botticelli; Galleria degli Uffizi, Florence, Italy/SuperStock.

complete understanding of biologically important macromolecules and is aiding in the interpretation of experiments concerned with their properties. Although such studies do not, in themselves, constitute a complete theoretical approach to biology, they represent a fundamental contribution to our knowledge of the properties and functions of the essential components of living systems.

Venus rose from the sea fully formed in all her glory, as pictured in "The Birth of Venus" by Sandro Botticelli (Figure 1). Such creation rarely occurs in science, particularly today, and it is useful to recall the prior developments that provided a basis for our attempting the first molecular dynamics simulation of a protein.[1] Two essential elements were the existence of molecular dynamics simulation methods and the implementation of potential functions for macromolecular systems like proteins. Molecular dynamics had developed along two pathways which came together in the study of biomolecule dynamics. One of these, usually referred to as trajectory calculations, has an ancient history that goes back to two-body scattering problems for which analytic solutions can be achieved. However, even for only three particles with realistic interactions, difficulties arise. An example is provided by the simplest chemical reaction, $H + H_2 \rightarrow H_2 + H$, for which a prototype calculation was attempted by Hirschfelder, Eyring, and Topley in 1936.[6] They were able to calculate only a few steps along a single trajectory, and it was nearly 30 years later that the availability of computers made it possible for us to complete the calculation.[7] That my group had been studying the trajectories of small molecule reactions for a number of years set the stage for the molecular dynamics simulation of a protein.

The other pathway in molecular dynamics had been concerned with physical rather than chemical interactions and with the thermodynamics and dynamic properties of a large number of particles, rather than the calculations of specific trajectories of a few particles. Although the basic ideas go back to van der Waals and Boltzmann, the modern era began with the work of Alder and Wainright on hard-sphere liquids in the late 1950s.[8] The paper by Rahman[9] in 1964, on a molecular dynamics simulation of liquid argon with a soft sphere (Lennard-Jones) potential, represented an important next step. Simulations of so-called complex fluids followed; the now classic study of liquid water by Stillinger and Rahman was published in 1974.[10] Our contact with Anesur Rahman, who was a constant source of advice and encouragement, resulted in a collaboration on the simulation of the alanine dipeptide in water, which was undertaken before the BPTI simulation, but only published afterwards.[11]

The energy functions for obtaining the forces required for implementing such dynamical methods for macromolecules had to be empirical potentials of the molecular mechanics type. Even today semiempirical quantum mechanical potentials are too slow (and not necessarily better than empirical potentials) to do the many energy derivative calculations (on the order of 1000 per picosecond) required for the simulations (nanoseconds or longer) that are now state-of-the-art. Although many persons have contributed to the development of empirical potentials, our work at Harvard owed most to the research of two groups. One is that of Harold Scheraga, whose work on the parametrization of nonbonded interactions in proteins was particularly important.[12] The other is the group of Shneior Lifson; their consistent force field (CFF) for small molecules[13] was very useful to Bruce Gelin,[14] who wrote the major portion of the original macromolecular simulation program employed for the BPTI calculation. Because of the lack of large scale computing facilities in the United States, the initial calculations were performed in France at CECAM, an organization for scientific computing directed by Carl Moser at Orsay. A month-long workshop (one really worked in workshops in those days) held there in 1976 (it was organized by Herman Berendsen) permitted us to do simulations with the program that had been finished just in time. Moreover, this workshop brought together a number of the key people doing molecular dynamics simulations today and led to the rapid dissemination of our methodology.

Molecular dynamics simulations of proteins, as of many other systems (e.g., liquids), can, in principle, provide the ultimate details of motional phenomena. The primary limitation of simulation methods is that they are approximate. It is here that experiment plays an essential role in validating the simulation methods; that is, comparisons with experimental data serve to test the accuracy of the calculated results and provide criteria for improving the methodology. The experimental approaches to bimolecular dynamics are limited as to the information that can be obtained from them; e.g., if one is concerned with the time scale of motions, the frequency spectrum covered by experiments such as nuclear magnetic resonance (NMR) is incomplete,[15] so that motional models that are able to rationalize the data can be inaccurate. When experimental comparisons indicate that the simulations are meaningful, their capacity for providing detailed results often makes it possible to examine specific aspects of the atomic motions far more easily than by making measurements.

In the 25 years between 1977 and 2002, molecular dynamics simulations of biomolecules have undergone an explosive development and have been applied to a wide range of problems, some of which are reviewed in this issue of *Accounts*. It is useful to separate three types of applications of molecular dynamics simulation methods for the study of macromolecules of biological interest. The first uses the simulation simply as a means of sampling configuration space. This is involved in the utilization of molecular dynamics, often with simulated annealing protocols, to determine or refine structures with data obtained from experiments, such as X-ray diffraction or nuclear magnetic resonance. The second uses simulations to determine equilibrium averages, including structural and motional properties (e.g., atomic mean-square fluctuation amplitudes) and the thermodynamics of the system. For such applications, it is necessary that the simulations adequately sample configuration space, as in the first application, with the additional condition that each point along the trajectory be weighted by the appropriate Boltzmann factor. The third area employs simulations to examine the actual dynamics. Here, not only is adequate sampling of configuration space with appropriate Boltzmann weighting required, but it must be done so as to properly represent the time development of the system. All three types of applications of molecular dynamics simulations are illustrated in this issue of *Accounts*.

With science moving as rapidly as it does today, there are exciting new applications of molecular dynamics simulations which it was not possible to include in this issue. Examples are real-time molecular dynamics of water penetration through a membrane protein,[16] of the formation of lipid bilayers,[17,18] and of the folding and unfolding of three-stranded β-sheet peptides,[19] though not yet of actual proteins. Also, a recent simulation which addresses directly a medical problem (the destruction of fibrous caps and their relation to heart attacks) is indicative of what the future holds.[20] It is clear that molecular dynamics simulations are providing and will continue to provide insights into the richness of the internal motions of biomolecules, whose evolutionary selection may well have included not only the average structures but also their motional correlates, as encoded in the potential energy surface.

In closing this Editorial, I can do no better than quote from Chapter 3, "The Relation of Physics to Other Sciences", in the wonderful Feynman *Lectures on Physics*, published in 1963:[21]

> Certainly no subject or field is making more progress on so many fronts at the present moment, than biology, and if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that *all things are made of atoms* [italics in the original], and that *everything that living things do can be understood in terms of the jigglings and wigglings of atoms* [italics added].

I would like to thank Paul Barbara, who did all the hard work as the real Editor of this issue. My contribution was to suggest authors.

References

(1) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature* **1977**, *267*, 585−590.
(2) Deisenhofer, J.; Steigemann, W. Crystallographic refinement and the structure of the bovine pancreatic trypsin inhibitor at 1.5 Å resolution. *Acta Crystallogr.* **1975**, *B31*, 238−250.
(3) Phillips, D. C. In *Biomolecular Stereodynamics, II*; Sarma, R. H., Ed.; Adenine Press: Guilderland, NY, 1981; p 497.
(4) Linderstrom-Lang, K. Deuterium exchange between peptides and water. *Chem. Soc. (London)* **1955**, Spec. Publ. 2, 1−20.
(5) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. *Adv. Protein Chem.* **1989**, *21*, 287−386.
(6) Hirschfelder, J. A.; Eyring, H.; Topley, B. Reactions Involving Hydrogen Molecules and Atoms. *J. Chem. Phys.* **1936**, *4*, 170−177.
(7) Karplus, M.; Porter, R. N.; Sharma, R. D. Exchange Reactions with Activation Energy. I. Simple Barrier Potential for (H,H_2). *J. Chem. Phys.* **1965**, *43*, 3259−3287.
(8) Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. 1. General method. *J. Chem. Phys.* **1959**, *31*, 459−466.
(9) Rahman, A. Correlations in motion of atoms in liquid argon. *Phys. Rev.* **1964**, *136*, A405.
(10) Stillinger, F. H.; Rahman, A. Improved simulation of liquid water by molecular-dynamics. *J. Chem. Phys.* **1974**, *60*, 1545−1557.
(11) Rossky, P. J.; Karplus, M.; Rahman, A. A Model for the Simulation of an Aqueous Dipeptide Solution. *Biopolymers* **1979**, *18*, 825−854.
(12) Scheraga, H. A. Calculations of the conformations of small molecules. *Adv. Phys. Org. Chem.* **1968**, *6*, 103−184.
(13) Lifson, S.; Warshel, A. Consistent Force Field for Calculations of Conformations Vibrational Spectra and Enthalpies of Cycloalkanes and n-Alkane Molecules. *J. Chem. Phys.* **1969**, *49*, 5116−5129.
(14) Gelin, B. R. Application of Empirical Energy Functions to Conformational Problems in Biochemical Systems. Harvard University, April 1976.
(15) Wand, A. J. Dynamic activation of protein function: A view emerging from NMR spectroscopy. *Nat. Struct. Biol.* **2001**, *8*, 926−931.
(16) de Groot, B. L.; Grubmüller, H. Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science* **2001**, *294*, 2353−2357.
(17) Marrink, S. J.; Tieleman, D. P.; Mark, A. E. Molecular Dynamics Simulation of the Kinetics of Spontaneous Micelle Formation. *J. Phys. Chem. B* **2000**, *104*, 12165−12173.
(18) Bogusz, S.; Venable, R. M.; Pastor, R. W. Molecular Dynamics Simulations of Octyl Glucoside Micelles: Dynamic Properties. *J. Phys. Chem. B* **2001**, *105*, 8312−8321
(19) Cavalli, A.; Ferrara, P.; Caflisch, A. Weak temperature dependence of the free energy surface and folding pathways of structured peptides. *Proteins: Struct., Funct. Genet.* **2002**, *47*, 305−314.
(20) Stultz, C. Localized Unfolding of Collagen Explains Collagenase Cleavage Near Imino-Poor Sites. *J. Mol. Biol.* **2002**, in press.
(21) Feynman, R. P.; Leighton, R. B.; Sands, M. *The Feynman Lectures in Physics*; Addison-Wesley: Reading, MA, 1963; Vol. I, p 3.6.

Martin Karplus

Harvard University and Université Louis Pasteur

Guest Editor

AR020082R