Theor Chem Acc (2000) 103:332–334
DOI 10.1007/s002149900025

Theoretical
Chemistry Accounts


*Perspective*

# Perspective on "Dynamics of folded proteins"

## McCammon JA, Gelin BR, Karplus M (1977) Nature 267: 585–590

**David A. Case**

Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

Received: 24 February 1999 / Accepted: 12 March 1999 / Published online: 28 June 1999

**Abstract.** This paper provides an overview of the title paper, discussing its background and significance, some details of the results, and subsequent developments that were stimulated by the work.

**Key words:** Molecular dynamics – Proteins – Historical perspective

### 1 Background

One of the key advances in computational chemistry during the last quarter of the twentieth century was the development of simulation methods to study motions of atoms and molecules in condensed-phase environments. The paper reviewed here is noteworthy for its pioneering application of molecular dynamics techniques to proteins, and (along with the 1976 Warshel and Levitt paper on lysozyme, discussed elsewhere in this issue) can be considered to herald an increased interest among the computational chemistry community on the problems dealing with the structures and dynamics of biological macromolecules.

In molecular dynamics simulations, one uses numerical integration techniques to obtain approximate solutions to Newton's equations of motion:

$$ma \equiv m(\mathrm{d}^2x/\mathrm{d}t^2) = F = -(\partial V/\partial x) \ . \qquad (1)$$

Here $m$ is an atomic mass, $a$ is the acceleration (the second derivative of position $x$ with respect to time), and $F$ is the force (expressed as the derivative of a potential $V$ with respect to position). Under the assumptions usually used, the effect of electronic motion is folded into the potential-energy function $V(x)$, so that the dynamic variables become the position of each atomic nucleus as a function of time.

The use of classical trajectory techniques to study atomic and molecular collisions in the gas phase can be traced back to the 1930s, and became an established technique in the 1950s and 1960s. Some of this background is contained in the perspectives in this issue on the 1959 paper by Ford and Wheeler [1], the 1970 paper by Miller [2], and the 1965 paper by Karplus, Porter, and Sharma [3]. The first two of these deal in part with the approximation in Eq. (1) that uses classical ideas, rather than quantum dynamics, to describe nuclear motion; this greatly simplifies the computational analysis, and in most cases should lead to relatively small errors. The third paper contains a lucid description of the use of classical trajectories to analyze elementary gas-phase collisions; this background was very much "in the air" in the Karplus group, where similar calculations continued well into the 1970s [4, 5].

A second thread in computational chemistry that made it possible to carry out dynamical calculations on proteins was the development of suitable potential-energy functions (or "force fields") that describe energies as a function of configuration. The fundamental ideas trace back to ideas long used by vibrational spectroscopists to interpret infrared spectra [6]. The use of computers in conformational analysis of floppier molecules (i.e. to locate likely structures that are local minima on the potential-energy surface) had begun in earnest in the 1960s [7–12], and by the early 1970s many applications to peptide and protein systems had been reported. The work of the Lifson group in developing the consistent force field (CFF) was particularly influential to graduate student Bruce Gelin, whose principal dissertation project involved the development of a similar suite of codes to handle the "bookkeeping" involved in computing and using potential-energy functions in polypeptides, and the calibration of parameters for a force field with fewer terms than that of the CFF. In this model, the energy of a polypeptide is expressed in terms of the contributions of each local bond, angle and dihedral distortion terms (comprising a simplified version of traditional vibrational force fields), and longer-range, "nonbonded" contributions that represented steric and electrostatic interactions:

$$E_{\mathrm{MM}} = \sum_{\text{bonds}} K_r(r - r_{\mathrm{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\mathrm{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} + \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right] . \quad (2)$$

This description is in many ways a quite simple one, lacking local anharmonic terms and cross-terms connecting bonds and angles, for example. Still, even for a small protein in vacuum there are many thousands of terms to be evaluated at each step, and many force constants ($K_r$, $A_{ij}$, $q_i$, etc.) to be estimated. In much early work, the value of the dielectric parameter $\varepsilon$ was taken to be the interatomic distance $R$ in angstroms, both to partially mimic solvent screening and to avoid expensive square root calculations. Gelin and postdoctoral researcher Andy McCammon (who had interests in both atomic and hydrodynamic models of proteins) worked together to construct a workable computer program to perform protein trajectories and to analyze the results, primarily in terms of time-correlation functions.

A third background thread for the application of molecular dynamics techniques to proteins came from the computational physics community on simulations of simple liquids. Here the perspectives in this issue on the 1964 paper by Rahman (on liquid argon) and on the 1971 paper by Rahman and Stillinger (on liquid water) may be helpful. The results of simulations and of other statistical models for dense liquids had led to a dynamical picture of diffusional motion whose character is dominated by relatively "hard" collisions among atoms which is augmented, but not fundamentally changed, by much "softer" (more slowly varying) attractive or cohesive forces that serve to hold the liquid together. The reasonable success of very simple potentials (such as hard-sphere models) in explaining the dynamical and equilibrium structural properties of simple liquids [13, 14] fueled the hope that many interesting features of biomolecular dynamics could be determined even with quite crude potential-energy functions, as long as the size scale and general characteristics of the rapidly varying parts of the potential were approximately correct. This belief had to contend with a common feeling (expressed at Harvard and elsewhere) that the number of adjustable parameters in force fields representing molecules as complex as proteins was so great as to severely compromise the usefulness of such calculations.

## 2 Results

The simulations reported in the paper dealt with bovine pancreatic trypsin inhibitor (BPTI), a small (58 amino acid) protease inhibitor whose structure is stabilized by three disulfide crosslinks. The simulation began with a set of X-ray coordinates and zero velocities, and ran 100 steps of numerical integration of Newton's equations of motion, with each time step being about 1 fs. Since the X-ray coordinates did not correspond to a minimum of the potential-energy surface, during the dynamics some of the potential energy was converted to kinetic energy, so that after 100 steps the mean internal temperature was 140 K. At this point, velocities were scaled by a factor of 1.5 (corresponding to scaling the kinetic energy by a factor of 2.25) and 250 additional equilibration steps were carried out, leading to an internal temperature of 285 K. Following this, 9,000 additional "production" steps were computed and analyzed, during which the average internal temperature rose to about 295 K.

The analysis of the results concentrated on the time-averaged structure and on correlation functions that illustrate the time-dependence of internal motions. A key conclusion was that the internal motion was fluidlike at ordinary temperatures, i.e. that "the dynamics of atomic displacements are dominated by collisions with neighboring atoms, at least on the picosecond time scale." Hence, it is argued, "...many of the dynamical properties (though not necessarily the correct average structure) can be obtained from any potential function which includes the forces that depend strongly upon distance (covalent and hydrogen bonds, nonbonded repulsions) and provides sufficient attractive interactions to preserve the compact structure of the native system."

## 3 Discussion

A striking feature (for most readers at the end of the twentieth century) is the short time scale of the simulation. This reflects both the speed of contemporary computers, and of the state of conformational analysis at the time, which typically analyzed calculations consisting of only a few hundreds of steps of energy minimization. The authors recognized that velocity equilibration was not complete even after 9 ps, and it was clear that significant computer resources would be needed for systematic studies of protein dynamics. For this reason, also, this initial paper is no longer consulted for quantitative details about fluctuations, although most of the results presented stand up well compared to later studies. Within 2 years, results from a 100-ps simulation were available [15], and the time scale and size of simulations continued to grow. Advances made in both algorithms and computers over the next two decades may be highlighted by comparison to a recent molecular dynamics simulation [16] of a system with about 20 times as many atoms as in the initial BPTI simulation, for a time period of 1 μs, $10^5$ times longer than the initial 10-ps time period. Even today, however, the size and time scales of affordable simulations is an important limitation to obtaining reliable answers for many interesting questions.

The field of macromolecular simulations has grown so rapidly that one no longer looks to the old primary literature for useful overviews, but rather turns to textbooks and monographs [17–21]. For those interested in some historical perspective, though, a few additional papers that represent the thinking in the Karplus group at the time are worth consulting. These include a comparison of hard-collision models with vibrational theories in understanding protein internal dynamics [22]; the

inclusion of solvent effects into the simulations [23]; and analyses of oxygen binding to hemoglobin and myoglobin [24, 25]. A Scientific American article by Karplus and McCammon, although published much later, retains much of the "flavor" of the early investigations [26]. (It should be recognized, however, that the computer-generated graphics used to illustrate this article were not available in 1977; more typical of the time were line-printer plots that were later hand-traced for publication.)

The title *Nature* paper was accompanied in the same issue by a "News and Views" commentary by Barry Robson [27], who correctly noted that "the potential for future application of molecular dynamics and related techniques seems enormous", and that a key impact of molecular dynamics studies would be to counter an overly rigid view of protein structure arising from the fact that crystallographic results are typically presented as a single (average) structure. It is perhaps sobering, however, to consider Robson's suggested application for the new technique, that of using computations to address a controversy between the conventional double-helix model of DNA and a then-current "zipper" model. The difficulties to be encountered in moving from qualitative insights about the nature of protein fluctuations to reasonably reliable estimates of the relative energies of different conformations have been underestimated at one time or another by almost everyone in the field. Recent analyses of the conformational energetics of DNA, using computers many orders of magnitude more powerful than those available in 1977, illustrate problems with accuracy of potentials, treatment of electrostatics, and sampling problems that still frustrate straightforward efforts to "resolve" biochemical problems by recourse to computation [28, 29].

The introduction of molecular dynamics techniques into biomolecular simulation certainly qualifies as a key development in computational chemistry, one that required considerable foresight in tackling what were obviously very difficult problems. Martin Karplus [30] would later write, "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, of course, is one of the great benefits of taking the initial, somewhat faltering steps in a new field where the questions are qualitative rather than quantative and any insights, even if crude, are better than none at all." In the beginning, emphasis was on picosecond motions in proteins that could be directly simulated, but it soon became clear (at least in principle) how to use the ability of molecular dynamics simulations to sample a Boltzmann distribution to make connections both to equilibrium thermodynamics [31] and to "activated" kinetic events that take place on slower time scales [32]. A few years later it was also realized that dynamical simulated annealing (i.e. a high-temperature simulation followed by slow cooling to low temperatures) can be a robust method for searching for low-energy conformers that is far superior to energy minimization techniques. The use of molecular-dynamics-based simulated annealing as part of an optimization strategy is now nearly universal in crystallographic and NMR structure refinement [33]. Overall, biomolecular molecular dynamics simulation is now no longer a specialized task for computational chemists, but an increasingly reliable tool that is a (nearly) routine adjunct to experimental studies for those interested in the general field of structural biology.

*Acknowledgements*. It has been most enjoyable to remember the many wonderful interactions I have had, both personal and scientific, with the authors of the title paper.

## References

1. Ford KW, Wheeler JA (1959) Ann Phys (NY) 7: 287
2. Miller WH (1970) J Chem Phys 53: 1949
3. Karplus M, Porter RN, Sharma RD (1965) J Chem Phys 43: 3259
4. Brumer P, Karplus M (1973) Discuss Faraday Soc 55: 80
5. Wang ISY, Karplus M (1973) J Am Chem Soc 95: 8160
6. Wilson EB, Decius JC, Cross PC (1980) Molecular vibrations. Dover, New York, pp 312–346
7. Hendrickson JB (1961) J Am Chem Soc 83: 4537
8. Hendrickson JB (1964) J Am Chem Soc 86: 4854
9. Lifson S, Warshel A (1968) J Chem Phys 49: 5116
10. Levitt M, Lifson S (1969) J Mol Biol 46: 269
11. Scheraga HA (1971) Chem Rev 71: 195
12. Allinger NL (1976) Adv Phys Org Chem 13: 1
13. Weeks JD, Chandler D, Anderson HC (1971) J Chem Phys 54: 5237
14. Barker JA, Henderson D (1972) Annu Rev Phys Chem 23: 439
15. Karplus M, McCammon JA (1979) Nature 277: 578
16. Duan Y, Kollman PA (1998) Science 282: 740
17. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Clarendon, Oxford
18. Frenkel D, Smit B (1996) Understanding molecular simulation: from algorithms to applications. Academic Press, San Diego
19. Brooks CL III, Karplus M, Pettitt BM (1988) Adv Chem Phys 71: 1
20. van Gunsteren WF, Weiner PK, Wilkinson AJ (eds) (1993) Computer simulations of biomolecular systems, vol 2. ESCOM, Leiden
21. van Gunsteren WF, Weiner PK, Wilkinson AJ (eds) (1997) Computer simulations of biomolecular systems, vol 3. ESCOM, Leiden
22. McCammon JA, Wolynes PG, Karplus M (1979) Biochemistry 18: 928
23. Rossky PJ, Karplus M (1979) J Am Chem Soc 101: 1913
24. Gelin BR, Karplus M (1977) Proc Natl Acad Sci USA 74: 801
25. Case DA, Karplus M (1979) J Mol Biol 132: 343
26. Karplus M, McCammon JA (1986) Sci Am 254: 42
27. Robson B (1977) Nature 267: 577
28. Srinivasan J, Cheatham TE III, Kollman P, Case DA (1998) J Am Chem Soc 120: 9401
29. Jayaram B, Sprous D, Young MA, Beveridge D (1998) J Am Chem Soc 120: 10629
30. Karplus M (1993) In: van Gunsteren WF, Weiner PK, Wilkinson AJ (eds) Computer simulation of biomolecular systems: theoretical and experimental applications. vol 2. ESCOM, Leiden, pp v–xvii
31. Pearlman DA, Rao BG (1998) In: Schleyer P von R, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer HF III (eds) Encyclopedia of computational chemistry. Wiley, Chichester, pp 1036–1061
32. Northrup SH, Pear MR, Lee C-Y, McCammon JA, Karplus M (1982) Proc Natl Acad Sci USA 79: 4035
33. Brünger AT, Karplus M (1991) Acc Chem Res 24: 54