Animation: a useful tool for protein molecular dynamicists, applied to hydrogen bonds in the active site of elastase

Stanley M. Swanson, Tomasz Wesolowski, Maciej Geller and Edgar F. Meyer

Biographics Laboratory, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX USA

Massive amounts of coordinate data result from molecular dynamics calculations. The animation program MDKINO is a simple but powerful tool for previewing or reviewing the results. In recent simulations of elastase, we have examined hydrogen bonding patterns, conformational changes involving shifts in ring positions and rotations of amino acid side chains, electric fields in interatomic space, and electric forces acting on chosen nuclei. Animation is also useful for checking on the stability of calculations in progress. Simple programming techniques achieve acceptable levels of animation with readily available hardware (PS330 or PS390 display with a serial interface to a laboratory VAX). In about half an hour, it is possible to make and watch a color stereo "movie" of a selected subsystem of a simulation (up to 1 000 frames of about 100 atoms each).

Keywords: computer animation, molecular dynamics, hydrogen bonding, elastase, EC 3.4.21.36

THE UTILITY OF GRAPHICS

Interactive computer graphics has already proven its worth in the study of complex molecules and their interactions.[1,2] Molecular dynamics calculations[3] extend the insights of static molecular modeling at the cost of multiplying the data analysis burden manyfold; more graphical tools are needed. Animation lets one examine the results qualitatively and focus on interesting events.

The eye (mind) is often able to pick out slow, large-scale qualitative changes in conformation, such as the shifting of functional groups with respect to each other. This chemically interesting motion and its topological consequences (e.g., the breaking and formation of hydrogen bonds) are not easy to find in the raw coordinates. Clues suggested by animation may indicate which collective motions or which observables should be subjected to more rigorous statistical analysis.

Molecular dynamics calculations, which are often run at remote supercomputer sites, may exhibit anomalous behavior, such as the escape of water molecules from solvation shells and problematic contacts between a few atoms. Current production versions of AMBER[4] do not adequately monitor for these events, so it is now our practice to transmit subsets of the saved coordinates over BITNET or INTERNET to check on the stability of calculations in progress.

In this paper we will first discuss the programs and their performance, then report on their application to a system of biochemical interest.

IMPLEMENTATION

Unfortunately, the equipment available to our laboratory and to many others still lacks the bandwidth and computational speed required to present thousands of frames of real-time animation of entire proteins with surrounding solvent. However, by selectively viewing the active site motions (say, 8 to 10 residues), it is possible to animate up to 1 000 frames, enough for a typical simulation.

This laboratory has an Evans and Sutherland (E&S) PS330 system with 2 megabytes of mass memory linked to a VAX 11/750 over a 19.2 kilobaud asynchronous terminal line. Although it requires about 15 minutes of download time to fill the graphics memory, initial frames can be seen in a few seconds. The frames in the E&S memory can be viewed in forward, reverse, and stop motion (stepping individual frames).

The set of programs in the implementation includes a topology generator that reads AMBER residue dictionaries to establish bonding in a selected subset of the entire protein-solvent-substrate complex; several programs to extract the selected coordinates from the simulation snapshots and a

Drs. Wesolowski and Geller are on leave from the University of Warsaw, Department of Biophysics, Warsaw, Poland

Address reprint requests to Dr. Swanson at the Biographics Laboratory, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843-2128, USA.
Received 10 March 1989; accepted 6 June 1989

downloading program (MDKINO) to send vector lists to the display. MDKINO allows one to modify the color choices for features in the images and determines hydrogen bonds in each frame by geometric criteria. MDKINO can be modified easily to emphasize other features of the data; another version, MDEKINO, shows electric fields in the neighborhood of chosen atoms, or indicates the electric forces acting on specified nuclei.

Although some existing software systems allow a limited amount of animation (e.g., HYDRA[5]), a simplified graphics interface was designed to maximize the amount of displayable information. Most of our 2 megabytes of mass storage is available for vector images, with space for about 200 000 endpoints. A minimal function net permits scaling, rotation and translation of the center of view. Stereo views can be achieved easily within the hierarchical structure of display objects in the E&S graphics language. A frame needs about 100 bytes for overhead and about 8 bytes for each endpoint, using vector normalization. (Block normalization saves one or two bytes per endpoint, but we have not exploited this.) A single line segment requires two endpoints, while n sequentially connected segments need only $n+1$ endpoints. We typically view about 100 bonds, in up to 1 000 frames of animation. More complicated pictures, such as electric fields, reduce the number of available frames.

The "level of detail" animation technique suggested by E&S[6] suffices for even 1 000 frames, although there is evidence that a single level of conditionals slows down the frame generation with many hundreds of frames. For a sample function net and a number of good ideas for information coding, see Weiner *et al.*,[7] who have previously applied this technique to molecular dynamics. We have reported another animation program, KINO,[8] which uses other methods.

THE ACTIVE SITE OF ELASTASE

Ever since the charge-relay[9] was proposed to explain the mechanism of the serine proteases, the spatial arrangement of active-site H-bond donors and acceptors has been a matter of ongoing interest. Crystallographic results are the prime source of experimental evidence for specific, functionally important H-bonding interactions in biological macromolecular complexes. The existence of a chain of H-bonds from the buried, catalytic Ser 195 O_γ through the catalytic tetrad and further to a remote site on the exterior of the enzyme has been observed.[10] (Residue numbers correspond to the chymotrypsinogen numbering system.) To begin a systematic study of the influence of this H-bond network on the catalytic mechanism of serine proteases, we performed a molecular dynamic simulation of native porcine pancreatic elastase (PPE, EC 3.4.21.36) with the active site surrounded by a dome of 630 water molecules. Our aim was to study whether there is any tendency to form an H-bond between the O_γ–H group of Ser 195 and N_ϵ of His 57. This bond is crucial for the first catalytic step; an ideal hydrogen bond would have an O_γ–N_ϵ distance of 2.85 Å and a C_β–O_γ–N_ϵ angle of 111°. PPE has a longer distance (3.2 Å) and a smaller angle (85°), similar to active sites of trypsin and trypsinogen that are exposed to water. In a survey[11] containing eight such structures, this distance varies from 2.8 to 3.1 Å, and the angle varies between 77° and 94°. The small angles indicate poor hydrogen bonding. The same survey also includes complexes, in which the active site is inaccessible to solvent, where this distance and angle are qualitatively different (2.6 Å, 108° average). We conclude that the solvent adversely influences the 195 O_γ–N_ϵ 57 interaction. Also arguing for solvent influence on SER 195 O_γ is its large thermal parameter ($B=22$ Å^2 in PPE[10]) when compared to nearby protein atoms ($B=4$ to 6 Å^2). Our calculations of the simulated active site have been fully described.[12]

EXAMPLES

We have studied hydrogen bonding patterns dynamically by drawing additional lines in individual frames to indicate which of the possible donor-acceptor pairs actually satisfies geometrical criteria (distance, angles) for good hydrogen bonding. By changing the picture to show electric fields in the neighborhood of a few atoms of interest, we studied the forces contributing to hydrogen bonding. These forces are usually called "electrostatic" because they are calculated from the "static" Coulomb potential; we call them simply electric, not because we include electrodynamic effects, but to remind us that their sources move at frequencies in the infrared.

We give several examples from a 100 ps molecular dynamics simulation of PPE.[12] Coordinate sampling at a 0.1 ps interval gives 1 000 frames. The active-site region of PPE showed a significant change between two families of conformational states of the side chain of Ser 195 during a 2 ps period of this simulation; the conformation of this side chain is crucial for catalysis. Of course, the best way to illustrate the utility of the program MDKINO is to show animation sequences; we can present only a few frames here.[13] Color Plate 1 (a,b,c) shows three frames (numbers 1, 638 and 771) that are typical of the initial conformational states, the transition state and the final conformational states, respectively. The catalytic triad (Asp 102, His 57 and Ser 195) is shown in red, water molecules are in green, and nearby residues are in blue. Yellow lines indicate hydrogen bonds.

The animation showed that during the whole simulation, the catalytically crucial hydrogen bond between the O_γ–H group (Ser 195) and the N_ϵ acceptor (His 57) was *never* formed in native PPE. Rather than with the O_γ–H group, this acceptor formed hydrogen bonds with water molecules. The water "O 3" was particularly important.

Another version of the program, MDEKINO, depicts the electric forces acting on selected nuclei in the same frames (Color Plate 2a,b,c; yellow bars). Electric fields and forces were calculated according to the model of electrostatic interactions employed by AMBER 3.0, with a distance-dependent dielectric "constant" ($1/r_{ij}$) that mimics the effect of shielding and with partial charges computed quantum mechanically. By the conventions of AMBER, these are net forces due to "nonbonded" interactions. A force vector begins on a nucleus and points in the direction of its free end. In the initial and transition states, the electric forces drawing the N_ϵ atom of His 57 and one hydrogen atom of the water molecule (O 3) together are much stronger and better directed than the net forces acting on the side chain atoms of Ser 195. Even in the final state, where the forces on N_ϵ and O_γ–H are comparable, O_γ forms hydrogen bonds

to the water "O 3" and another residue rather than to N_ϵ; the hydrogen bond to N_ϵ is not geometrically possible.

An option in MDEKINO can display electric fields in a small region of space outside the van der Waals radii of nearby atoms. Color Plate 2d shows the electric field in the space between the side chains of Ser 195 and His 57 for the transition state frame. The field vectors are bicolored, beginning on a point of a sampling grid in light blue and ending in dark blue. The extent and spacing of the grid is user definable. Knowledge of the dynamical behavior of this field may help to estimate its influence on proton transfer between these two groups, which is the next step of the catalytic process in elastase.

SUMMARY

These dynamics calculations and animation gave us some insights into the role of water in elastase catalysis that we could not have obtained so easily from any other technique. These insights include:

(1) The active site of PPE is not rigid, in contrast to older preconceptions based on time-averaged crystallographic coordinates.
(2) In the native, hydrated state, the O_γ–H group of Ser 195 is mainly located in one of two distinct regions without any possibility for the formation of the catalytically crucial H-bond with N_ϵ of His 57.
(3) A competition between molecules of water in the active-site region and the O_γ–H group for H-bond formation with the N_ϵ atom of His 57 favors the water.
(4) Even at the transition where the O_γ atom rotated past His 57, the calculated electric forces opposed the formation of this crucial H-bond. Especially here, the combination of molecular dynamics, electric field calculations, and sequential visualization via MDKINO presents a myriad of data in a comprehensible form.

The simulation reported here has now pointed us to the next step. The molecular dynamics simulation of the system now includes an oligopeptide substrate, a dome of hydration, and a shell of water to produce an active complex (of nearly 9000 atoms) with a buried catalytic tetrad (under study).

It is not sufficient to calculate molecular dynamics simulations. These simulations must be linked, wherever possible, with experimental (crystallographic, nmr and other spectroscopic) evidence. Moreover, our graphical resources are still imperfect. Thus, we continue to develop ad hoc methods to explore the implications of these structural and computational observations. A desire to understand the nature of interactions in structural biochemistry continues to stimulate the development of more comprehensive graphical tools for visualization and conceptualization of these complex phenomena.

ACKNOWLEDGEMENTS

Financial support has come from the National Science Foundation (DMB-8517286). Discussions with Rosemarie Swanson have greatly sharpened the exposition. T.W. and M.G. wish to acknowledge travel support from the Polish Academy of Sciences within the project C.P.B.R. 3.13 and C.P.B.R. 01.06. We also gratefully acknowledge the support of Dr. John Dinkel, Associate Provost for Computing, Texas A&M University, and of the Texas Agricultural Experiment Station.

REFERENCES

1 Collins, D.M., Cotton, F.A., Hazen, E.E., Meyer, E.F. and Morimoto, C.N. Protein crystal structures: quicker, cheaper approaches. *Science* 1975, **190**, 1047–1053
2 Jones, T.A. Interactive computer graphics: FRODO, in *Methods in Enzymology*. Academic Press, New York, 1985, **115**, 157–171
3 McCammon, J.A., and Harvey, S.C. *Dynamics of Proteins and Nucleic Acids*. Cambridge University Press, Cambridge, 1987
4 Weiner, P.K., and Kollman, P.A. *J. Comp. Chem.* 1981, **2**, 287–303; Singh, U.Ch., Weiner, P.K., Cadwell, L., and Kollman, P., AMBER: Assisted Model Building With Energy Refinement, Version 2, 1985; Case, D., Seibel, G., and Bush, P., Version 3.0, 1986 with update 1987, Department of Pharmaceutical Chemistry, University of California, San Francisco
5 Hubbard, R.E. in *Computer Aided Molecular Design*. Oyez Scientific, UK, 1985, 99–106
6 "Conditional Referencing," PS 300 Document Set, Volume 2b, Graphics Programming, Evans and Sutherland Computer Corporation, Salt Lake City, Utah, 1984, 12–16, 23–25
7 Weiner, P.K., Gallion, S.L., Westhof, E. and Levy, R.M. Visualization of energetics and conformations from molecular computer simulations. *J. Molecular Graphics* 1986, **4**, 203–207
8 Fujita, T., Swanson, S.M. and Meyer, E.F. Intermolecular enzyme-ligand animation in the active site of porcine pancreatic elastase with acetyl-alanine-proline-alanine by means of molecular dynamics calculation. *J. Molecular Graphics* 1986, **4**, 208–212
9 Blow, D.M. Structure and mechanism of chymotrypsin. *Acc. Chem. Res.* 1976, **9**, 145–152
10 Meyer, E., Cole, G., Radhakrishnan, R. and Epp, O. The crystal structure of native porcine pancreatic elastase at 1.65 Å Resolution. *Acta Cryst.* 1988, **B44**, 26–38
11 Marquart, M., Walter, J., Deisenhofer, J., Bode, W. and Huber, R. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Cryst.* 1983, **B39**, 480–490
12 Geller, M., Carlson-Golab, G., Lesyng, B., Swanson, S.M. and Meyer, E.F. How rigid is the active site of native porcine elastase? molecular dynamics simulations, *Biopolymers*, submitted
13 MDKINO is being made available through the Quantum Chemistry Exchange