ELSEVIER

Graphical visualization of mean hydration from molecular dynamics simulations

Jed Pitera* and Peter Kollman†

**Graduate Group in Biophysics and †Department of Pharmaceutical Chemistry, University of California, San Francisco, California, USA*

How does one characterize water solvating a complex solute? Specific hydration of proteins and nucleic acids plays a key role in many biological processes. However, traditional pairwise descriptions of solvent structure (radial distribution functions, etc.) are incapable of adequately describing the hydration of these complex solutes. We have developed methods to visualize the average three-dimensional water structure surrounding a solute, as seen in a molecular dynamics (MD) simulation. Applications to simple solutes [sodium ion, N*-methyl acetamide, 18-crown-6, (hydroxymethyl)phenols] are presented, and the extension of the method to larger molecules of biochemical interest is discussed. © 1998 by Elsevier Science Inc.*

Keywords: hydration visualization, molecular dynamics, AMBER, MidasPlus, sodium ion, N*-methyl acetamide, crown ether, (hydroxymethyl)phenol*

INTRODUCTION

Water is the solvent for virtually every biological process. Whether inside or outside a living cell, water surrounds the proteins, nucleic acids, and small molecules necessary for life. In many cases, water molecules make specific hydrogen bonds and electrostatic interactions with the surface atoms of these molecules. More generally, water is responsible for the hydrophobic effect that stabilizes the structure of proteins and drives many macromolecular interactions. Also, the water surrounding two molecules must be displaced if those molecules are to interact with one another. For example, both the enzyme active site and the substrate must be desolvated before they can interact. More practically, water molecules must be displaced from a drug-binding site on a protein before an inhibitor can bind. Clearly, it is important to describe and understand the hydration of biological molecules.

In this article, we describe the development of some software tools that can be used in conjunction with molecular dynamics to describe the average structure of water molecules surrounding a solute. The present study is limited to small molecules, but these methods have been applied to biomolecular systems, including proteins and nucleic acids. The data we have collected show the position and orientation of water molecules in the tightly bound first and second shells of hydration surrounding each molecule. Molecular dynamics is used as a tool to generate a large number of realistic solvent conformations, and the average properties of these conformations are extracted and visualized.

Traditionally, water structure has been described by radial distribution functions [$G(r)$], which show the probability distribution of distances between an atom of the solute and a type of solvent atom (methane carbon and water oxygens, for example). For solutes with a large number of atoms, such as a typical protein (2 000+ atoms), such an atom-based description is intractable. Instead, a description that is independent of the individual atoms of the solute is necessary. We have decided to describe the space surrounding the solute using a fixed Cartesian grid [$F(x, y, z)$]. This permits our data collection and visualization software to be used on a broad range of systems—from methane to DNA.

Two main types of data were collected in this study. The first is the Cartesian analog of the radial distribution function—$G(x, y, z)$ rather than $G(r)$—the water oxygen probability density. Properly normalized, this is the probability of finding a water oxygen in a particular volume of space during the molecular dynamics simulation. When converged, it gives a clear picture of the most favorable positions for water molecules near the solute, what is thought of as the first shell of hydration. This sort of technique has previously been applied to simulations of biological molecules by Lounnas et al.[1] Subramanian and Beveridge (e.g., see Ref. 2) have similarly used a

Color Plates for this article are on pages 386–388.

Address reprint requests to: P. Kollman, Department of Pharmaceutical Chemistry, University of California, San Francisco, UCSF Box 0446, San Francisco, California 94143, USA.

Received 28 October 1997; revised 30 January 1998; accepted 2 February 1998.

655 Avenue of the Americas, New York, NY 10010

1093-3263/97/$17.00
PII S1093-3263(98)00006-0

superposition of snapshots along a molecular dynamics trajectory to suggest the hydration structure around nucleic acids. The superposition approach is hampered, however, by limited sampling: the 20 or so structures used can only suggest the highest maxima of the water probability density, and contain little information about moderate- to low-probability regions. The continuous data collected in this study have been filtered and displayed on the basis of simple statistical analysis, providing information about the structure of the water in both high-probability (traditional hydration sites) and low-probability regions.

While these data are useful, they do not completely describe the solvent structure. There is no information about the orientation of the water molecules in the probability density. Consequently, we augmented the probability data with the mean dipolar vector field. These data show the mean orientation of water molecules in a particular sector of space, indicating hydrogen bond and electrostatic interactions with the solute.

It must also be noted that prior studies of the hydration of both small solutes[3] and macromolecules[2] have examined the pairwise water–solute interactions in substantially more detail, including careful calculation of energetics and orientational parameters. The intent of this work, in contrast, is to describe a simple computational framework for the qualitative analysis of solvent structure that makes use of readily available software tools and interactive three-dimensional modeling to provide a vivid picture of solvation.

METHODS

Simulation details

Molecular dynamics simulations were carried out for each solute solvated by a cube of water molecules. All simulations were run using the SANDER module of AMBER 4.1[4] and the Cornell et al. force field.[5] The TIP3P water model was used. This model has been shown to reproduce accurately the free energies of solvation and hydrogen bond geometries for small solutes when used in conjunction with electrostatic potential derived charges.[6] In addition, the TIP3P model adequately simulates the microscopic structure and bulk properties of water. Simulations were run with fully periodic boundary conditions. The systems were equilibrated at constant pressure, but data were collected from simulations at constant volume to facilitate the use of a fixed grid. Nonbonded interactions were truncated at 8 Å. The solute was held fixed in the center of the periodic box, but the solvent molecules were free to move and coupled to a temperature bath at 300 K.

Data collection

Modifications were made to the SANDER molecular dynamics software to permit real-time collection of information about the solvent structure. The simulation box was divided into cubic bins, 0.5 Å on a side. Data were collected every molecular dynamics time step (2 fs), to give sufficient statistics for the observed properties. For the probability densities, a three-dimensional array of integers is maintained, one per grid bin. At each time step, the grid position of every water oxygen is determined. One "count" is then added to the integer array element at the corresponding grid position. At the end of the simulation, the array is output in a format suitable for the display software. The procedure for collecting the mean dipolar vectors is similar. Three variables are maintained per grid point in addition to the integer array: real accumulators for the *x*, *y*, and *z* components of the dipolar vector. At each time step, the list of water molecules is again traversed. The position of each water oxygen is determined, and the integer counter incremented as above. In addition, Cartesian components of the dipolar vector (vector sum of the O $\rightarrow$ H1 and O $\rightarrow$ H2 vectors) are calculated and added to the appropriate *x*, *y*, and *z* accumulators. When the simulation is complete, the mean *x*, *y*, and *z* components of the dipolar vector are calculated for each grid bin. These are output as a list of grid coordinates and vector components.

Visualization

The MidasPlus software suite[7] was used to display and manipulate the data in this study, since it allows for real-time interactive manipulation of three-dimensional models as stereo images. All the data in this study are resolved in three spatial dimensions, but all the images included below are two-dimensional (2-D) views for ease of viewing. The probability densities were projected onto a model of the solute, using the MidasPlus program and its Density delegate, a facility for displaying electron densities (or other scalar fields) atop molecular models. The Density program allows interactive display, contouring, and coloring of scalar fields. For each probability density, the mean density and standard deviation were calculated. All contouring was done at some number of standard deviations above the mean density, to differentiate the tightly associated hydration shell from the bulk solvent. The mean dipolar vectors were similarly displayed using MidasPlus. The Discern delegate was modified to permit interactive display, coloring, and contouring of vector fields. Contours for the vector images were chosen by hand to show the significant features of the hydration shell, including hydrogen bond donors or acceptors, while minimizing clutter. Consequently, the vector images show data points only for high-probability regions of the solvent shell. The image of water residencies or lifetimes in the sodium ion hydration shell was also displayed using Discern. The wireframe images are direct screen captures from MidasPlus, using the snapshot utility. All of the solid rendered images were generated using the MidasPlus rendering tools Conic[8] and Ribbonjr. All visualization and data display was carried out on a Silicon Graphics IRIS Indigo2 (150-MHz R4400, Elan graphics) running IRIX 5.2.

RESULTS

The methods described above were applied to four different solutes: a sodium ion, *N*-methyl acetamide, the crown ether 18-crown-6, and various isomers of (hydroxymethyl)phenols. The results for each system are presented below.

Sodium ion

The hydration of a sodium ion was selected as an initial test case for our methods. The simple structure of the Na^+ hydration shell also permits a gradual introduction to our various graphical representations of hydration, from probability densities to mean dipolar vectors and lifetime data. Color Plate 1 shows these differing views of Na^+ hydration. Color Plate 1a

is a snapshot from a molecular dynamics trajectory, illustrating the difficulty in determining the structure of the solvent shell from instantaneous frames of molecular dynamics. Color Plate 1b, displaying the water oxygen probability density around the ion from the same simulation, gives a much clearer picture of the first and second hydration shells. Color Plate 1c adds orientational information to this view, showing the increased directional order of waters in the first hydration shell relative to the second. Our software also allows collection of dynamic information about the solvent structure, as seen in Color Plate 1d. Water molecules remain in the first hydration shell much longer than the second; lifetimes in the second shell are only slightly longer than those in bulk water, indicating that these second shell waters are moving and exchanging rapidly with bulk solvent.

N-Methyl acetamide

N-Methyl acetamide (NMA) is an example of a relatively simple molecule that is both a hydrogen bond donor and acceptor. The mean dipolar vector images in Color Plate 2a and b clearly show the hydrogen-bonding patterns associated with both the hydrogen bond donor (N–H) and acceptor (C=O) groups of NMA (Color Plate 2a). The geometry of hydrogen bond donation by the N–H group is, as expected, more restricted than the geometry of hydrogen bonds accepted by the carbonyl (Color Plate 2b).

18-Crown-6

The crown ether 18-crown-6 is an example of a compound for which simple pairwise measurements fail to describe hydration adequately. The closely associated waters that donate hydrogen bonds to ether oxygens on either face of the ring typically interact with more than one ether group and often interfere or interact with one another, creating a complex, threefold-symmetric hydration shell. The network of water–solute and water–water hydrogen bonds in this shell was first observed in a Monte Carlo study,[9] and our data replicate their observation of "bridging" water conformations. In these configurations, one water molecule sits just above the plane of the ring and donates two hydrogen bonds to ether oxygens. It accepts a hydrogen bond from a higher, bridging water that also hydrogen bonds to the third ether oxygen on the same face of the ring. The overall solvent structure around 18-crown-6 is shown in Color Plate 2c, and the high-probability regions, which correspond to waters making at least one hydrogen bond to the ring oxygens, are shown in more detail in Color Plate 2d.

(Hydroxymethyl)phenols

The free energy of transfer from water to toluene is much more favorable (about 3 kcal/mol) for 1,3-(hydroxymethyl)phenol than either the 1,2- or 1,4-isomer. Ben-Naim[10] has attributed this to the presence of a bridging water that forms strong hydrogen bonds to both hydroxyl groups of the 1,3-isomer (Color Plate 3a). The probability density from our MD simulations, however, shows no region of significant density that would correspond to this bridging water for the 1,3-(hydroxymethyl)phenol (Color Plate 3b and c), even when compared with the 1,4-isomer (Color Plate 3d). This confirms the free energy calculations of Sun and Kollman,[11] who established that the free energies of transfer from the gas phase to water are similar for all three phenols (within 1.0 kcal/mol), discounting any anomalous hydration of the 1,3-form.

DISCUSSION

Our water probability density data clearly describe the average structure of the hydration shell of a solute. In addition to these data, we have managed to collect information describing both the average orientation of solvent waters and approximate lifetime information, yielding a detailed picture of solvent structure and dynamics. The utility of this picture is shown by its application to a question of physical chemistry—the anomalous transfer free energies of (hydroxymethyl)phenols.

These tools for graphical visualization of the mean solvent structure calculated from molecular dynamics simulation were developed for two reasons. First, it is difficult to extract information about the structure of the hydration shell from instantaneous coordinates of a molecular dynamics trajectory (for example, see Color Plate 1a). In addition, traditional radial distribution functions are inadequate for describing the hydration of complex solutes, especially large, moderately polar solutes such as proteins and nucleic acids. Consequently, we have developed grid-based methods that capture the average structure of the hydration shell but are also extensible to permit the study of large, biomolecular solutes. The software described above, with which data are collected at every MD time step, is too memory- and compute-intensive to use with large solutes. In collaboration with T. Cheatham (NIH), we have developed a postprocessing utility that performs similar analysis and data collection on a previously calculated molecular dynamics trajectory. Application of these tools and methods to a 1-ns simulation of a DNA decamer clearly show the minor groove "spine of hydration" observed in high-resolution X-ray crystallographic studies of DNA.[12] Similar agreement is seen between molecular dynamics simulations and high-resolution structures of RNA.[13]

The programs developed in this article are available from the authors.

ACKNOWLEDGMENTS

We are grateful for research support from the NSF (Grant CHE-94-17458) and for the use of the resources of the UCSF Computer Graphics Laboratory; Tom Ferrin, PI, is supported by NIH RR-1081.

REFERENCES

1 Lounnas, V., Pettitt, B., Findsen, L., and Subramanian, S. A microscopic view of protein solvation. *J. Phys. Chem.* 1992, **96,** 7157–7159; Rudnicki, W., and Pettitt, B. *Biopolymers* 1997, **41,** 107–119

2 Subramanian, P.S., and Beveridge, D.L. A theoretical study of the aqueous hydration of canonical B d(CGCGAATTCGCG): Monte Carlo simulation and comparison with crystallographic ordered water sites. *J. Biomol. Struct. Dyn.* 1989, **6,** 1093–1122

3 Meng, E., and Kollman, P. Molecular dynamics studies of the properties of water around simple organic solutes. *J. Phys. Chem.* 1996 **100,** 11460–11470

4 Pearlman, D., Case, D., Caldwell, J., Ross, W., Cheatham, T., Ferguson, D., Seibel, G., Singh, U., Weiner, P., and Kollman, P. *AMBER*, version 4.1. University of California, San Francisco, San Francisco, California, 1995
5 Cornell, W., Cieplak, P., Bayly, C.I., Gould, I.R., et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 1995, **117,** 5179–5197
6 Cornell, W., Cieplak, P., Bayly, C.I., and Kollman, P. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* 1993, **115,** 9620–9631
7 *MidasPlus*. Computer Graphics Laboratory, University of California, San Francisco, 1994. Supported by NIH RR-01081
8 Huang, C.C., Pettersen, E.F., Klein, T.E., Ferrin, T.E., and Langridge, R. Conic: A fast renderer for space-filling molecules with shadows. *J. Mol. Graphics* 1991, **9,** 230–236
9 Ranghino, G., Romano, S., Lehn, J.M., and Wipff, G. Monte Carlo study of the conformation-dependent hydration of the 18-crown-6 macrocycle. *J. Am. Chem. Soc.* 1985, **107,** 7873–7877
10 Ben-Naim, A. Solvent-induced interactions: Hydrophobic and hydrophilic phenomena. *J. Chem. Phys.* 1989, **90,** 7412–7425; Ben-Naim, A. Solvent effects on protein association and protein folding. *Biopolymers* 1990, **29,** 567–596
11 Sun, Y., and Kollman, P. Are there water-bridge-induced hydrophilic interactions. *J. Phys. Chem.* 1996, **100,** 6760–6763
12 Cheatham, T., Crowley, M., Fox, T., and Kollman, P. A molecular level picture of the stabilization of A-DNA in mixed ethanol–water solutions. *Proc. Natl. Acad. Sci. U.S.A.* 1997, **94,** 9626–9630
13 Miller, J. Unpublished results