Properties of Water Molecules in the Active Site Gorge of Acetylcholinesterase from Computer Simulation

Richard H. Henchman,* Kaihsu Tai,* Tongye Shen,† and J. Andrew McCammon*
*Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, and Department of Pharmacology; and †Department of Physics, University of California, San Diego, La Jolla, California 92093-0365 USA

ABSTRACT A 10-ns trajectory from a molecular dynamics simulation is used to examine the structure and dynamics of water in the active site gorge of acetylcholinesterase to determine what influence water may have on its function. While the confining nature of the deep active site gorge slows down and structures water significantly compared to bulk water, water in the gorge is found to display a number of properties that may aid ligand entry and binding. These properties include fluctuations in the population of gorge waters, moderate disorder and mobility of water in the middle and entrance to the gorge, reduced water hydrogen-bonding ability, and transient cavities in the gorge.

INTRODUCTION

Acetylcholinesterase (AChE) catalyzes the hydrolysis of the neurotransmitter acetylcholine (ACh), terminating signaling across cholinergic synapses (Soreq and Seidman, 2001). It is one of the fastest known enzymes, operating close to the diffusion controlled rate in water (Rosenberry, 1975; Hasinoff, 1982). Therefore, considerable work has gone into understanding how the enzyme is able to achieve this. The catalytic machinery of AChE, a serine hydrolase, is estimated to increase the rate by 10^7–10^{11}-fold (Fuxreiter and Warshel, 1998; Vagedes et al., 2000). In addition, electrostatic steering of the ligand into the active site is believed to be responsible for a further 10^2–10^3-fold increase (Zhou et al., 1996). However, one of the peculiarities of the enzyme revealed from the x-ray structure (Sussman et al., 1991; Bourne et al., 1995) is that the active site resides in a deeply buried gorge whose entrance is not even wide enough for the entrance of the substrate ACh. Such an arrangement does not seem conducive to high turnover since the substrate has to not only pass through the narrow passage, but also displace and move through water molecules that are themselves slowed down in the gorge. To address the problem of the narrow passage, molecular dynamics simulations (Wlodek et al., 1997; Tai et al., 2001) have shown that the width of the entrance is not fixed, but fluctuate. The aim of this work is to determine the properties of the water in the gorge from a recent 10-ns computer simulation of mouse AChE (Tai et al., 2001) to understand how water might influence substrate diffusion. A second reason to study the water properties in the active site is related to ligand design. The moderate inhibition of AChE is clinically important to prolong the action of ACh in diseases such as Alzheimer's, in which there is a deficiency of ACh (Grutzendler and Morris, 2001) or myasthenia gravis, in which there is a deficiency of ACh receptors on the postsynaptic membrane (Wittbrodt, 1997). There is increasing evidence that including waters in the modeling of enzyme ligand complexes is necessary to obtain more accurate binding predictions (e.g., Lam et al., 1994; Ni et al., 2001).

Properties of water around proteins may be determined by a number of methods. High-resolution x-ray and neutron diffraction methods are able to reveal water positions, their extent of disorder, orientation in the case of neutron diffraction, and occupancy (Savage and Wlodawer, 1986; Schoenborn et al., 1995; Carugo and Bordo, 1999). An example of relevance to this study is that a recent study of a number of AChE x-ray structures has shown variations in the packing arrangement of water molecules with different ligands (Koellner et al., 2000). NMR methods provide not only structural information on hydration sites, but also the residence times of waters in these sites (Otting and Liepinsh, 1995; Denisov and Halle, 1996; Wiesner et al., 1999). Confined waters are found to have residence times in the range 10^{-9}–10^{-3} s, while for the more mobile waters on the protein surface residence times are only 10–50 ps. For water in large cavities in proteins, the 10–20 waters in the binding cavity of intestinal fatty acid binding protein are estimated to have a 1-ns residence time (Wiesner et al., 1999). Molecular dynamics simulations, subject to their limitations of approximate modeling and timescale of 10 ns–1 μs, have the advantage of being able to probe many water properties more directly, including diffusion coefficients, water paths, radial distribution functions, time correlation functions, residence times, hydrogen bonds, water density and orientation, and free energy (Wong and McCammon, 1986; Brooks and Karplus, 1989; Brunne et al., 1993; Roux et al., 1996; Zhang and Hermans, 1996; Helms and Wade, 1998; Garcia and Hummer, 2000; Makarov et al., 2000; Likic and Prendergast, 2001).

No simulations have reported the properties of water in the active site gorge of AChE as far as we are aware. However, previous simulations have provided information about how waters and ligands may enter and exit the gorge.

Submitted September 6, 2001, and accepted for publication January 2, 2002.

Address reprint requests to Dr. Richard H. Henchman, 9500 Gilman Dr., Mail Code 0365, La Jolla, CA 92093. Tel.: 858-822-1469; Fax: 858-534-7042; E-mail: rhenchma@mccammon.ucsd.edu.

0006-3495/02/05/2671/12 $2.00

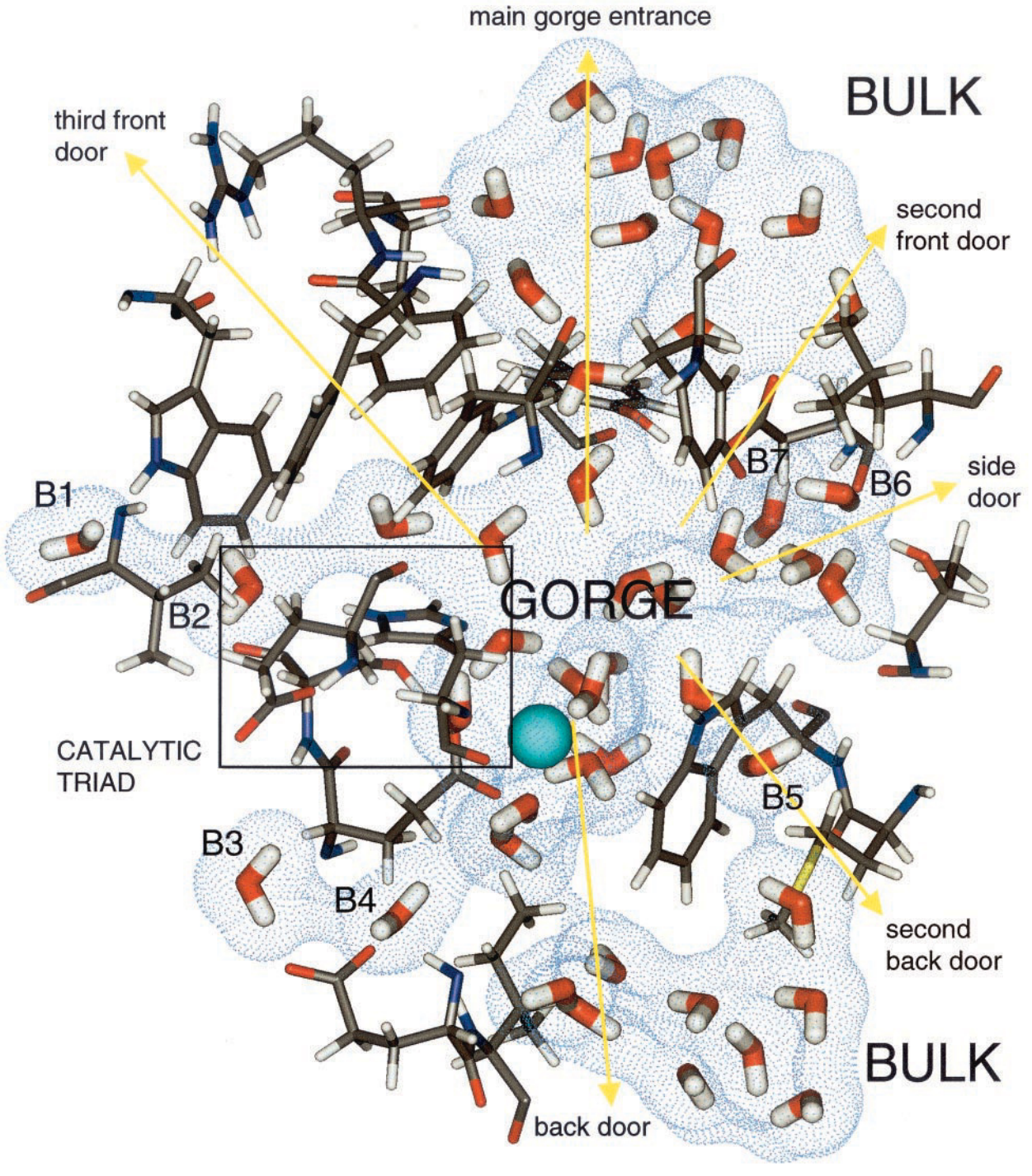


FIGURE 1 Location of the entrances into the gorge. Clockwise from noon, they are the main gorge entrance, second front door, side door, second back door, back door, and third front door. Water molecules are V-shapes and the blue sphere is the sodium ion. The seven waters marked B1–B7 are sometimes or always buried in the protein.

In addition to the main gorge entrance seen in the x-ray structure, a number of possible entrances to the gorge that open wide enough for water passage have been observed, but without any water transits into the gorge (Gilson et al., 1994; Wlodek et al., 1997; Tara et al., 1999a,b). An illustration of the gorge and its possible entrances is shown in Fig. 1. The most frequently observed alternative path in simulations has been the back door at the bottom of the gorge. A second neighboring back door was also observed here and in a simulation of *Torpedo californica* AChE

(TcAChE) dimer complexed with tacrine (Wlodek et al., 1997). The side door has only been seen when AChE was complexed with ligands such as huperzine A in mouse AChE (Tara et al., 1999a) (mAChE) and tacrine in TcAChE (Wlodek et al., 1997). Free energy calculations (Enyedy et al., 1998) have also suggested that a third front door is favorable for passage of negatively charged acetate.

When analyzing water around proteins, it is not the average properties of an individual water molecule that are desired but the average properties of water in a given location. Therefore, the water's properties have to be assigned not to the water molecule, but to some coordinate frame such as a grid. However, there is a problem trying to map information from thousands of simulation frames onto a single reference frame for dynamic systems. Since a protein is dynamic, a protein reference frame will move around in time, making a water molecule buried in the protein appear to move even when it is possibly fixed with respect to its binding contacts in the protein. This not only smears out the water density, but may also lead to misleadingly short residence times in a given position. Therefore, we adopted a coordinate frame local to nearby residues for each water using a method termed averaged residue coordinates (ARC) fully described elsewhere (R. H. Henchman and J. A. McCammon, in preparation).

In this work, we first examine how water enters and exits the gorge. To examine the gorge water structure more closely, we extract a water density averaged over a 10-ns simulation of mouse AChE using ARC. This density is mapped to a single protein coordinate frame for analysis. Hydration sites are taken from the maxima in the density. The properties of the hydration sites examined are occupancy, residence time, hydrogen bonds, and dipole moment. The complementary properties of gorge volume and empty space are also studied. Conclusions are then drawn about how these properties may influence the function of AChE.

MATERIALS AND METHODS

Simulation protocol

The details of how the system is set up are given in a previously reported 1-ns simulation (Tara et al., 1999b). In summary, the system consists of mouse AChE (Bourne et al., 1995) (Protein Data Bank identification code 1MAH) in a box of 25205 SPC/E waters and 10 sodium ions to maintain system neutrality. For the water setup in particular, there were no crystal water molecules in the 1MAH structure. Therefore, waters were placed inside the protein using the GRID program (Goodford, 1985). The criteria for water placement in a cavity was that the water's energy be <-46 kJ mol^{-1} and that it make at least two hydrogen bonds. The protein was then solvated in a cubic box (edge 96 Å) of preequilibrated waters. Waters closer than 2.6 Å to any protein heavy atom were removed. Nine sodium ions were placed in the solvent at ~5 Å from the protein surface and one sodium ion was placed in the choline binding pocket of the active site gorge. The full 10-ns equilibrated MD trajectory used in this analysis has been reported previously by Tai et al. (2001). Their simulation was run on a Cray T3E using the NWChem program (Straatsma et al., 2000), the AMBER94 force field (Cornell et al., 1995), particle mesh Ewald summation (Darden and York, 1993), SHAKE (Ryckaert et al., 1977) for bonds involving hydrogens, and constant temperature and pressure reservoirs (Berendsen et al., 1984), set to 298.15 K and 1 atmosphere, respectively. Snapshots for analysis were saved every 1 ps.

Water analysis

For most subsequent analyses, water molecules are treated as spheres centered at the oxygen atom. Gorge waters are defined as those waters in the active site chamber separated from the bulk by the narrow bottlenecks. Two definitions are used to identify such waters: 1) all waters within 4 Å of any other gorge water and having the fewest number of 4 Å contacts with bulk water. This places the boundary at the narrowest part of any gorge entrance. To reduce noise arising from the fuzzy boundary at the bottleneck, a water only leaves or enters the gorge if the change is maintained consecutively for 10 ps. All other waters are either bulk or buried in the protein. An open passageway connecting the gorge and bulk exists when a bulk and gorge water are within 4 Å of each other. An alternative means of defining these passageways used in previous work (Tai et al., 2001) is when the two regions are connected by a solvent-accessible surface for a water-sized 1.4 Å probe. 2) The second definition of gorge waters is more precise and defines the gorge waters as all waters in a set of predetermined sites that lie inside the main chamber within the bottlenecks. These sites are determined using a recently developed method that places sites at peaks in the water density calculated in frames of reference that are averaged over all surrounding residues (R. H. Henchman and J. A. McCammon, in preparation).

In brief, the construction of the water density is a two-stage process. First, the time-averaged position (TAP) of each water is calculated in the reference frame of each neighboring residue. At any time, the TAP of the water may be calculated in the protein frame as the average position with respect to each neighboring residue weighted by the number of times that that residue was a neighbor of the water. This coordinate system is termed averaged residue coordinates (ARC). The TAP is continually updated as long as the water remains within 2.8 Å of the TAP averaged over the three previous frames. This averaging provides a more robust check of whether the water has left the TAP. When a water leaves a TAP, it forms a new TAP. Second, to enable clustering of the TAPs, the water density is constructed from all the TAPs on a 0.5 Å grid covering the gorge region by projecting the TAPs in ARC onto a single protein frame. The first frame of the 10 ns is chosen for this construction. Each TAP is placed on the grid with height given by the lifetime of the TAP. All maxima above 25 times the average density become sites, and any of these sites closer than 2 Å are merged to give the final site definitions. All TAPs are then placed in the closest site within 2.8 Å, a slightly larger distance than the TAP spacing to ensure that all nearby TAPs are placed in one site. The properties of each site are derived from the properties of each TAP, weighted by the lifetime of that TAP. Residence times are calculated from the survival function, $S(t)$ (Impey et al., 1983) given by

$$S(t) = 1/N_{\text{water}} \sum_{i=1}^{N_{\text{water}}} P_i(t)$$

where $S(t)$ gives the fraction of total waters, N_{water}, that remain in a site after a given time, t. $P_i(t)$ is a binary function that equals 1 if water i is still in the site after time t, and zero otherwise. A single expontial fit to $S(t) \propto \exp(-t/\tau)$ yields a residence time, τ, for that site. One other property of the gorge that is examined is unoccupied volume. Volume is measured over a 0.5 Å grid laid over all desired gorge sites. With the first frame as a reference, all subsequent frames are aligned by minimizing the root mean square deviation (RMSD) of the C_α values of 43 residues lining the gorge. For a given probe size, any grid point outside the sum of the van der Waals radius and the probe size for all atoms creates accessible volume for that probe. Each such grid point is assumed to occupy one grid box, and volume is estimated as the sum of these. This definition of volume is only

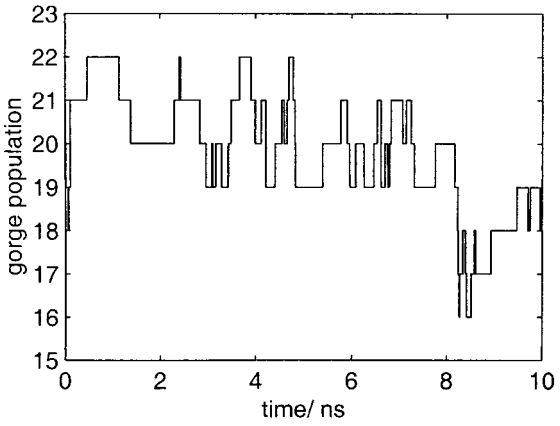


FIGURE 2 Population of gorge waters over 10 ns (according to definition 1 described in the text).

approximate, but is sufficient given that volume is not a precisely determined quantity due to boundary effects, etc. Two different types of volume are examined: 1) the gorge volume as defined by the protein ignoring water, and 2) empty space. In each case the volume is measured over a fixed region in space that encompasses the gorge. For the gorge volume, this region is the total volume enclosed by spheres of radius 2 Å placed at gorge water sites according to the definition two of gorge sites described above. For empty space a larger region is examined, which is the total volume enclosed by spheres of radius 2 Å placed at all the hydration sites considered, including those outside the gorge. Fig. 1 was made predominantly with Insight II 2000 and Figs. 4 and 5 were made with OpenDX, using the chemistry modules (Gillilan and Wood, 1995).

RESULTS

Fluctuations in gorge water population

One property related to ligand entry is how the population of the gorge varies over time. The variation in gorge water population according to definition 1 over the full 10 ns is shown in Fig. 2. It can be seen that the population is not static, but ranges quite substantially from 16 to 22 waters with an average of 19.8. The GRID program, by comparison, placed 19 waters in the gorge at the start of the run, which suggests the GRID does a reasonable job placing waters. A total of 36 unique waters at one time or another reside in the gorge. The sodium ion appears stable and always remains present in the gorge; 69 single changes in gorge water population occur during the entire 10 ns. This variation indicates that water stability in the gorge is marginal when the population is around the average, and that the free energy change for exchanging a water with the bulk is small. The relative free energies, ΔG, between population i and population j may be estimated using the equation

$$\Delta G = -RT \ln(n_i/n_j)$$

where n_i and n_j are the numbers of times the two respective populations occur in the simulation. There is only a 3 kJ mol^{-1} variation in free energy in the range 18–22 waters.

TABLE 1 Percentage of time each passage is open

Passage	Connecting Waters	Connecting Surface*
Main Gorge Entrance	66.0	71.8
Front Door 2	0.4	—
Back Door	0.3	0.8
Back Door 2	0.1	—

*Determined with a probe of radius 1.4 Å.

The populations 16 and 17 are higher in free energy, but occur too infrequently for a reliable estimate. This difference of six water molecules in the gorge population is significant as it is comparable to the volume of a typical ligand. Thus, when the population is low enough, a ligand should be able to enter the gorge without waters having to simultaneously exit through an alternative exit. Interestingly, in a comparable study of water in the active site of fatty acid binding protein (Likic and Prendergast, 2001) the fluctuation in water population was much larger, ranging from 14 to 33. Given the larger size of typical fatty acids versus ACh, these fluctuations in water population may contribute to the size-selectivity of ligand binding proteins.

The changes in gorge population arise either from waters exchanging with bulk water or with the interior of the protein. There are a number of passages leading from the gorge to the bulk. In addition to the main gorge entrance observed in the original crystal structure (Sussman et al., 1991), five other passages have been observed in simulations of AChE, as discussed earlier. Three of these alternative passages were observed in this simulation: the back door, the second back door, and the second front door (Fig. 1). In this simulation it is found that all 53 transits observed between the gorge and bulk go through the main gorge entrance, and none through the other passages. Where the transit occurs correlates strongly with the amount of time that these passages are open wide enough for water molecules.

Two definitions for an entrance being "open" are used here: 1) a continuous chain (oxygen-oxygen distance <4 Å) of waters leading from the gorge to the bulk via that passage, and 2) the solvent-accessible surface (probe size 1.4 Å) directly connecting the gorge and bulk. The calculation of the solvent-accessible surface is described in an earlier work (Tai et al., 2001). Table 1 shows the percentage of time each entrance is open using the two definitions. On the timescale of the simulation, only the main gorge appears open long enough to allow waters to pass through, a process that occurs once every ~200 ps. All other passages are open for only a small fraction of the simulation time. Nevertheless, this does not rule out their role in water or product release on a longer timescale that is comparable to the turnover time of ~0.5 ms for mouse AChE (Radić et al., 1997). Furthermore, the presence of ligands in the gorge may have a perturbing influence on these passages, such as the appearance of the side door with huperzine A or tacrine

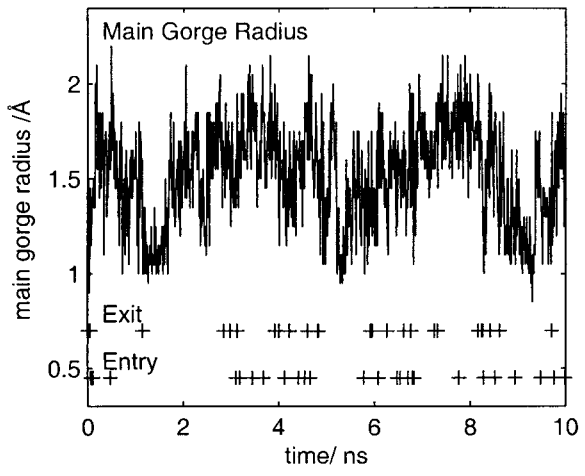


FIGURE 3 Correlation of entry and exit water transits (*crosses*) between the gorge and bulk with the radius of the main gorge radius (*solid line*).

(Wlodek et al., 1997; Tara et al., 1999a). Indeed, the sodium ion present in this simulation, by complexing with neighboring waters and impeding their mobility, may be acting as an effective plug to the back door.

The probability of a transit is likely to depend both on the width of the main gorge entrance and the gorge water structure itself. The width of the main gorge may be quantified by the main gorge radius (Tai et al., 2001). This is the maximum probe radius at which the solvent-accessible surface connects the gorge to the bulk. The main gorge radius may be approximately visualized for the snapshot in Fig. 1 as the narrowest part of the dotted blue surface through the main gorge entrance. The importance of the main gorge width may be seen in Fig. 3, which shows the value of the main gorge radius when an entry or exit event occurs. Transitions are evidently more likely when the gorge entrance is wider than the average radius, 1.5 Å. The average radius when waters enter or exit is 1.6 Å, slightly larger than the average, and larger than 1.4 Å, the radius of a water molecule. Only 10 of the 53 transits occur when the radius is nominally $<$1.4 Å. Such transits are still possible because transiently larger openings are masked by the 1-ps resolution. However, a wide enough gorge entrance alone is not the only factor governing water transits. The water structure and dynamics, discussed below, must also contribute to the frequency of transits. The existence of a relationship between water population and main gorge radius is also examined to determine whether a larger radius is due to more water molecules in the gorge. However, a correlation coefficient of only 0.16 between these two variables indicates that any relationship is weak.

The alternative means of water entering the gorge is transits between the gorge and isolated pockets in the protein. Twelve of these transits are somewhat artifactual and arise when the movement of a neighboring gorge water strands neighboring water molecules in the protein. Thus they reflect subtle changes in gorge water packing rather than a change in gorge population. Fig. 1 shows the location of these sites. Site B1 is next to Val-407, B6 lies between Leu-76 and Thr-83, and B7 is between Asp-74 and Leu-76. Four transits involve the displacement of a water molecule between the gorge and isolated cavity in the protein. The location of the sites participating in these transits is shown in Fig. 1. B2 lies between the catalytic Ser-203 and Val-407, and B5 between Trp-86 and Met-85. Each site fills and empties once during the simulation, indicating that the timescale for waters exchanging one way with these sites is ~5 ns. The waters in two other sites near the gorge, B3 and B4, remain buried in the protein for the duration of the simulation.

Gorge hydration sites

A more detailed analysis of gorge water structure and dynamics is necessary to understand the population fluctuations and get insight into how a ligand may behave in the gorge. This is achieved by identifying hydration sites in the gorge from maxima in the water density. The water density in the gorge is shown in Fig. 4 *a*. The density in the figure has been constructed from slightly broadened TAPs using Gaussians of width 1 Å to aid visualization. The larger and redder the contoured volume, the more localized the site. The density plot clearly shows distinct regions where waters predominate. The most ordered regions are those close to the protein, especially the buried regions. The density in the middle and the entrances to the gorge is more diffuse.

By extracting sites from the maxima in water density, the gorge is found to contain 19 hydration sites and one sodium site, as the sodium remains in the same location throughout the simulation. This is consistent with an average water population of 19.8 reported above using the first gorge water definition. A plot of gorge population (not shown) also yields a very similar result to that in Fig. 2. Note that the sites obtained from this water density are characteristic of the protein frame used to construct them. If a different frame were to be used, a slightly different arrangement of sites is obtained, molding to the altered protein shape. These 20 gorge sites are illustrated in Fig. 4 *b* by the orange points, together with selected residues also in orange. To verify the presence of the sites, a comparison may be made with x-ray crystal waters. However, this comparison was not possible using the mAChE x-ray structure because in the highest-resolution mAChE structure available to date only six waters were resolved inside the gorge (Bourne and Marchot, personal communication). A recent 1.8 Å structure of TcAChE (Protein Data Bank identification code 1EA5) contains 18 gorge waters (M. Harel et al., M. Weik, I. Silman, and J. L. Sussman, in preparation), almost as many as the number of sites found in the simulation. The two structures were superimposed, minimizing the RMSD of the C_α values of 43 residues lining the gorge. The agreement

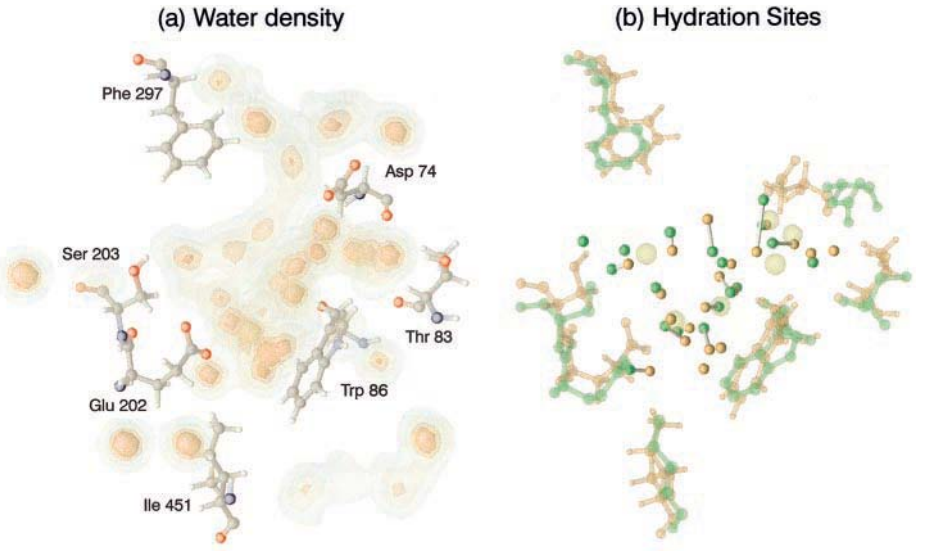


FIGURE 4 (*a*) Contour plot of the water density in and around the gorge. The reddest regions show where water density is most highly localized. (*b*) Comparison of hydration sites from simulation and x-ray structures. The 20 orange points are simulation sites with matching green protein residues; the 18 green points are x-ray waters sites from TcAChE with matching orange protein residues; the gray bars connect corresponding simulation and TcAChE waters; the faint yellow circles are the six x-ray waters from mAChE.

between the sites is reasonable, with 16 waters matching each other with an RMSD of 1.4 Å (an alignment using C_α values of all protein residues gave slightly worse agreement of 1.6 Å). The simulation (*orange*) and x-ray sites (*green*) overlaid are shown in Fig. 4 *b*, together with the corresponding positions of selected residues. The 16 matching waters are each connected by a gray bar. The unmatched sites lie in three regions. In the first area 2 x-ray sites lie in the region between Ser-203 and Phe-297. Waters in the simulation do sometimes appear in these positions, but not often enough to form a distinct site. Three simulation sites are unmatched in the second area in the choline binding site, and one simulation site is unmatched in the third area next to Thr-83. All these discrepancies may be due to the different species of enzyme, particularly disparate in structure around Asp-74 (72 in TcAChE numbering), or to different conformational states due to crystal packing effects for x-ray sites and force field and inadequate sampling effects for simulation sites.

Site occupancy

Entry of a ligand into the gorge requires empty space and/or moving waters. The hydration site occupancies, residence times, and traffic between sites offer a convenient way to measure this. The first property discussed is occupancy. The average occupancy of the 20 sites in the gorge is 1.01, as would be hoped for a meaningful definition of a site. However, the occupancies of certain sites are found to fluctuate from single occupancy. The numbered sites and their occupancy are shown in Fig. 5 *a*. Sites 1–20 are in the main gorge chamber, with sodium in site 12. Also shown are some selected sites outside the gorge: 21 and 22 are the two buried sites, B2 and B5, mentioned earlier, that exchange with the gorge; 23–27 are outside the main gorge entrance; 28–30 are beside the back door. The coloring shows the site occupancies. Blue sites have occupancy that is always close to 1, red sites are doubly occupied >15% of the time, green sites are empty >20% of the time, while yellow sites may be either doubly occupied or empty using the same criteria. Triple occupancy of sites is negligible.

Four of the sites in the gorge may be unoccupied. Of these, sites 1 and 2 lie at the gorge entrance. The fact that sites 1 and 2 may be empty may be helpful to the entrance of a ligand by occasionally making space for it to enter. The other two unoccupied sites, 4 and 5, sit adjacent to the carboxylate oxygens of Glu-202. The vacancy of these sites may indicate water instability here. Displacement of these unstable waters may be favorable to ligand binding; these sites do make up the binding site of where ACh and other ligands bind. Their vacancy may also reflect a dynamic piston-like motion of Glu-202 moving in and out of the gorge. The other partially unoccupied sites are the buried sites 21 and 22, as noted earlier, and sites 23, 28, and 29 outside the gorge. Deviations from single occupancy outside the gorge are much more likely given side chain mobility and less resolved sites from the smoother water density. This is particularly the case for site 28 near the back door, which may be unoccupied or doubly occupied. Six gorge sites are doubly occupied 15% of the time. Three of these sites, 13–15, are in the middle of the gorge, while sites 2 and 3 are near Ser-203. This too is important, as it suggests that there is significant empty space in these regions that may be occasionally occupied by waters. Such space might also aid ligand mobility diffusing into the gorge. The sodium ion in site 12 also appears as doubly occupied. This is more artifactual and arises due to overlapping sites because the neighboring waters sites are very close. This effect can be reduced by using a shorter cutoff when placing waters in the sodium site, but this complicates the protocol and makes sites dependent on whether or not

they contain sodium, and is therefore not used. The remaining sites in the gorge predominantly display single occupancy. Knowledge of site occupancies now makes possible a more detailed interpretation of the gorge population in Fig. 2. A comparison of the time series of site occupation with total gorge population indicates where in the gorge the water changes are occurring. For example, the maximum at 1 ns in Fig. 2 corresponds to double occupancy of sites 2, 3, and 14, while the minima at 8.5 ns corresponds to the vacancy of sites 1, 2, 4, and 5.

Care must be taken in interpreting the occupancies. The occupancy may be some measure of the stability of a water in a site. However, empty or double occupancy does not necessarily imply an empty cavity or two tightly packed waters in the gorge, as will be seen later in the accessible volume analysis. Rather, it may be evidence of less common alternative water packing, a slightly expanded or contracted total gorge volume due to protein motion, or a mobile side chain. Hence, occupancy should perhaps be interpreted as the frequency of existence of a site. Another caveat concerning occupancy is that the crossing back and forth of some waters across site boundaries may create some artificial site occupancies. For example, this may be causing the deviation in occupancy for sites 1, 2, and 15. To check that this is not an artifact of waters crossing site boundaries, there is a 24% chance that the cluster of sites 1, 2, and 15 is only doubly occupied, supporting the observation that at least one of them is empty at this time. Such a problem would get worse as the water density becomes more smoothed moving away from the protein surface, but at least in the gorge, this effect is expected to be small given the fairly sharp density profile seen in Fig. 4 *a*. These features illustrate the usefulness of ARC coordinates. In the protein frame, these differences in water structure would be much more averaged over and consequently harder to extract than in ARC coordinates (R. H. Henchman and J. A. McCammon, in preparation).

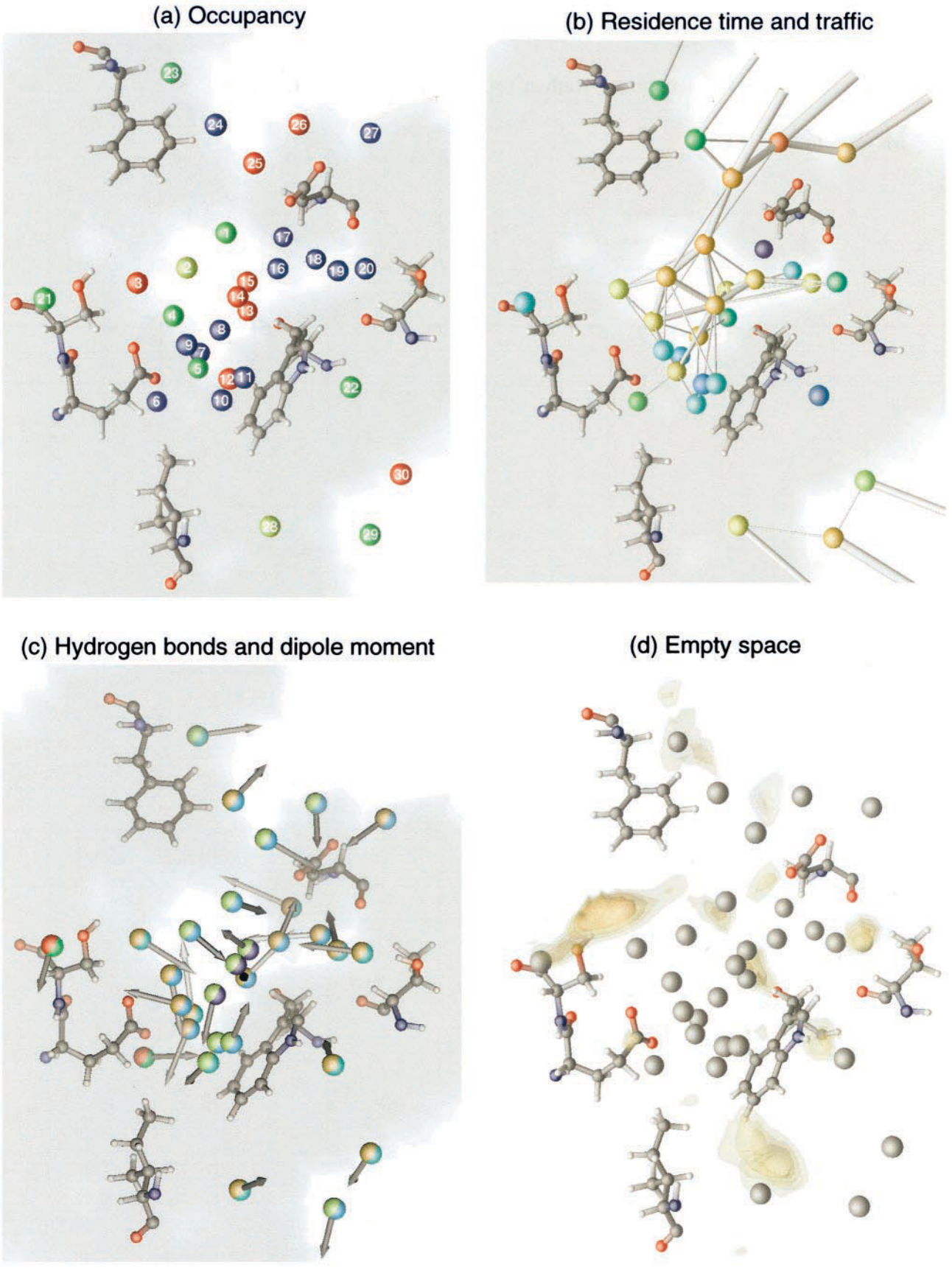
Site mobility

Site mobility may be inferred from two properties. One is the residence time of sites. The other is the number of transits between adjacent sites. The residence time, τ, is the decay time for the survival function, $S(t)$, of each site. A transit takes place when a water moves into a new TAP that is in a different site. Both of these properties are displayed in Fig. 5 *b*. Residence times in the gorge vary over a wide range from ~100 ps (*red sites*), ~200 ps (*orange*), ~500 ps (*yellow*), ~1 ns (*green*), ~2 ns (*cyan*) to the full simulation time of 10 ns (*blue*). These times are much longer than residence times on the outer surface of the protein, which are commonly 10–50 ps and demonstrate the confining effect the protein has on water properties. Other notable statistics for gorge waters are that the average TAP time for all gorge waters combined is 2.0 ns, on 35 occasions a water remains in the same position for >1 ns, and that the water in site 17 and the sodium ion in site 12 gorge each remain in the same site throughout the simulation. The distribution of gorge TAP times, given in Fig. 6, shows that the most common TAP times are in the range 100–300 ps. Fig. 5 *b* indicates that the most mobile areas are the gorge center and entrance, while the least mobile are the sites 17–20 in the pocket between Trp-86 and Asp-74, sites 9–12 in the choline binding region, and sites 6 and 7 adjacent to Tyr-133 (not shown) near the back of the gorge. This reflects the general principle that the fewer neighboring waters and more surrounding residues a water has, the lower its mobility. The clear exception to this is the main gorge entrance lined with many aromatic residues. Presumably these residues with their moderately mobile and featureless side chains provide little opportunity either sterically or electrostatically to lock in waters, and they therefore facilitate water diffusion despite the narrowness of the passage. It should also be pointed out that the reduced mobility of waters in the choline binding region may be due to the presence of the sodium ion, which coordinates with and locks in these waters.

As well as knowing how long a water stays in a particular place, it is useful to know to where the waters go. The sites between which waters move are also shown in Fig. 5 *b* by the bars connecting the sites. Waters that leave the gorge area for the bulk are shown by bars pointing nonspecifically into the bulk. The thickness of the bar scales with the number of water transits. The thickest bar from site 26 to the bulk has 78 transits, while the thinnest bars have as few as 3. There are numerous bars representing less than three transits and these are omitted for clarity. This figure displays a number of interesting features. The first is that transits are much more common outside the gorge than inside. This reflects the greater mobility of water outside the gorge than inside. In particular, the sites immediately outside the entrance to the gorge exchange frequently with more bulk-like regions. Hence the water here would not be expected to slow ligands down significantly. The second feature is that for transits between the gorge and bulk, there appears to be a dominant route through the main gorge entrance; 27 of the 53 gorge/bulk transits follow a single route between sites 1 and 25. Finally, of those transits inside the gorge, most take place between sites in the middle. Few waters diffuse directly between sites along the gorge surface.

Hydrogen bonds

Another property of interest is the number of hydrogen bonds the waters make in each site. A hydrogen bond is defined as existing when a polar hydrogen is within 2.5 Å of any oxygen or nitrogen. Fig. 5 *c* shows the number of water-residue and water-water hydrogen bonds. The coloring of the left hemisphere of the site indicates the number of


(a) Occupancy
(b) Residence time and traffic
(c) Hydrogen bonds and dipole moment
(d) Empty space

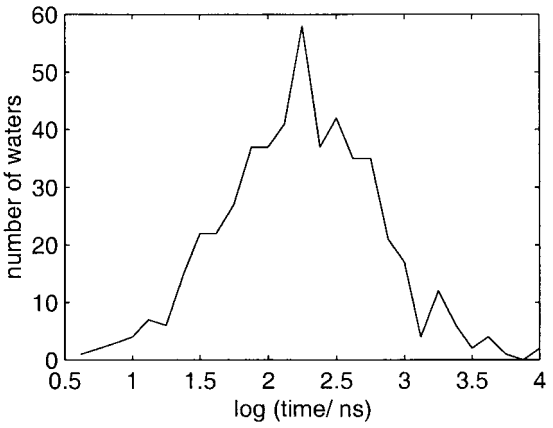


FIGURE 6 Logarithmic distribution of TAP times in the gorge.

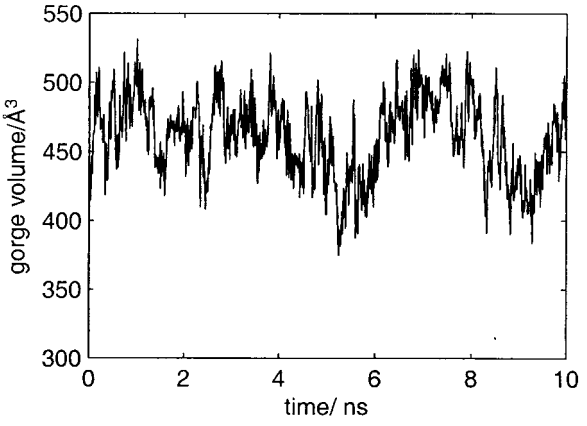


FIGURE 7 Gorge accessible volume as a function of time. A 1-Å probe is used for the accessible volume. Data are smoothed to a 10-ps resolution to remove noise.

water-residue hydrogen bonds, from yellow being zero to red being two. The right hemisphere shows water-water hydrogen bonds, from green being zero to dark blue being three. The results are fairly intuitive. Most waters lining the gorge have one to two hydrogen bonds with residues, while those in the middle and main gorge entrance have few, if any. The average for all gorge waters is 1.2. The reverse trend is true for water-water hydrogen bonds. Most waters in middle have two to three hydrogen bonds, while those lining the protein have one to two. The average for all gorge waters this time is 2.2. The exception to this trend is in the sites lining the main gorge entrance, which are largely unable to form hydrogen bonds with the protein. This possibly contributes to the higher mobility of waters in this region. Overall, each water is fairly similar and has three to four hydrogen bonds with an average of 3.4, less than the number in bulk water, which is around four. This number of hydrogen bonds is typical of water in internal cavities of proteins (Williams et al., 1994).

Dipole moment

The final property of sites to be examined is their average dipole moment. This is shown in Fig. 5 *c* by the arrows pointing out from each site. The direction of the dipole moment indicates the average direction, while the extent of grayness shows the average magnitude. The whitest arrows have dipole moment of 2.4 D, namely that of SPC/E water. The darkest arrows are 1.3 D and below. The lower the size of the average dipole moment, the more disordered the water in the site. The dipole moment of each site except the sodium site has its own preferred direction and is not completely averaged out, as would be the case in bulk water. The protein evidently structures waters not only in their position, but also in their orientation. Some local correlation does seem present such as in the clusters near Ser-203 or Thr-83, and the dipoles around sodium clearly point away from the sodium site. However, there is no single preferred direction of dipole moment for all sites, and the ability of the protein to orient the water molecules is reduced in the center of the gorge whose sites have more smaller and more disordered average dipoles.

Accessible volume

The accessible volume also reveals interesting features of the water structure in and around the gorge. The first type of volume considered is the volume of the gorge cavity formed by the protein, ignoring water molecules. The time series of accessible gorge volume for a 1 Å probe over the 10 ns is shown in Fig. 7. These data were smoothed to 10 ps reso-

FIGURE 5 (*a*) The occupancy of hydration sites in and near the gorge. Blue indicates sites that are usually singly occupied; red indicates sites that are doubly occupied >15% of the time; green indicates sites that are empty >20% of the time; yellow indicates both red and green (none and double occupancy). The gray shading is a protein cross-section. (*b*) Site residence time (*colored spheres*) and the traffic between them (*bars*) in and around the gorge. The color of the site indicates the average residence time of the site (red ~50 ps to blue ~10 ns). The thickness of the bar indicates the number of waters that pass along the bar (thinnest <10 to thickest ~50; bars with less than three transits are excluded for clarity). The gray shading is a protein cross-section. (*c*) Number of hydrogen bonds (*colored hemispheres*) and average dipole moment (*arrows*) of each site. The coloring of the left hemisphere of the site indicates the number of water-residue hydrogen bonds, from yellow being zero to red being two. The right hemisphere shows water-water hydrogen bonds, from green being zero to dark blue being three. The shading of the arrow indicates the magnitude of the dipole moment with the lightest being 2.4 D, while black is 1.3 D and below. The gray shading is a protein cross-section. (*d*) Empty space in the gorge. The contours indicate the probability of finding a 1-Å spherical cavity at that point and are spaced at 10, 15, 20, and 30% intervals. The black spheres are the hydration sites.

lution to remove noise. The information of most interest in this figure is how the volume changes with time and how this correlates with the gorge water population in Fig. 2. While the correlation is not perfect, there are a number of features that the two figures share. Both graphs have local minima around 0, 2, 5, and 8.5 ns and local maxima around 1, 4, 7, and 10 ns. Thus the gorge volume appears to be a contributing factor toward the total gorge population, although the correlation coefficient between gorge volume and population is still small at 0.28. However, a striking resemblance was noticed between the time series of the gorge volume and the main gorge radius in Fig. 3. This relationship has a correlation coefficient of 0.47. This confirms the fairly intuitive thought that as the main gorge gets wider, the gorge increases in size.

The second volume property considered is empty space. Fig. 5 *d* illustrates the predominant locations of empty space for a 1-Å probe. Two large cavities are revealed. The largest is in the pocket next to Ser-203 and is open to the probe 50% of the time. This cavity is also visible in the gray protein cross-sections in Fig. 5, *a–c*. Being so close to the catalytic serine indicates that this cavity may have a functional significance. By keeping out waters, it may be conducive to the binding of some other functional group such as a methyl group, for example, on the acetate part of ACh. Waters are seen to transiently exist here, but they evidently appear to be less stable with few polar residues nearby with which to form hydrogen bonds. Also of note is that it is in this region where there is disagreement with the x-ray waters seen in the TcAChE structure (M. Harel et al., M. Weik, I. Silman, and J. L. Sussman, in preparation) (Fig. 4 *a*). In that case, there was a water in this place. Such a difference may be species-dependent, but this may only be confirmed by a high-resolution structure of mAChE capable of seeing all gorge waters. A second large cavity is observed immediately outside the gorge by the back door. The role of the back door in ACh hydrolysis is still unclear. This cavity may be functionally significant, either by making waters if not ligands more mobile here, or may even favor ligand binding outside the back door. There is also a series of small cavities running from inside the main gorge entrance to the outside that may assist in maintaining a high level of water mobility in the confined main gorge entrance. Other smaller cavities are present whose significance, if any, is less obvious. One is above Trp-86 between Gly-82 and Trp-439; another lies at the buried site 22, which is only partially occupied; one lies lower down by the Trp-86 carbonyl oxygen between Gly-126 and Leu-130. For a second contour plot made with a 1.4-Å probe (not shown), only the first cavity mentioned by Ser-203 is still significant, showing that it is indeed large enough for a water molecule.

DISCUSSION

A detailed study of water properties in the active site gorge has revealed a number of interesting properties that may influence how a ligand behaves in the gorge. Water molecules move in and out of the gorge through the main gorge entrance once every 200 ps, on average. Waters are also found to exchange with two small cavities in the protein on a 5-ns timescale. Fluctuations in gorge volume also influence the number of waters in the gorge. By these means, the gorge water population fluctuates from 16 to 22 molecules. The size and frequency of these fluctuations may influence how quickly and what size ligand may enter the gorge. The roles of the second front door and two back doors remain unclear, but the fact that they do occasionally open wide enough suggests that waters are likely to be able to pass through.

Twenty hydration sites were found to exist in the gorge, and these compare closely with the 18 hydration sites observed from the x-ray structure of TcAChE. Of these 20 sites, there appear to be two main types. The waters in the middle, the main gorge entrance, and near Ser-203 and Glu-202 appear to be more disordered in density and dipole moment orientation, display more variable occupancy and cavities, have lower residence times with more intersite traffic, and in the case of the main gorge entrance, fewer hydrogen bonds. However, the waters in the choline binding site and pocket between Asp-74 and Trp-86 have the opposite properties, namely they are more ordered, have single occupancy, longer residence times, and little intersite traffic. These differences may be attributed to the more confining shape of the Asp-74/Trp-86 pocket and to the presence of the sodium in the choline binding pocket. However, this difference is significant because the more mobile class of water makes up most of the binding site of the natural substrate ACh and provides a path for it to enter. The more mobile and less structured waters would permit faster ligand diffusion. The tendency of certain water sites to possess occupancies below 1 and the presence of cavities suggest a reduced stability of waters here. Displacement of these waters or cavities by a ligand may lead to stronger binding. This may be particularly so for the large cavity observed adjacent to Ser-203. Such an effect, though, can only be confirmed by free energy calculations. The one part of the ACh binding region that is relatively immobile is the choline binding site, in which sodium appears quite stable. Water appears unable to displace sodium from this region, but choline, being also positive, may have a greater ability.

Ultimately, water properties are most useful when observing their direct effect on the binding of ligands, and so the properties derived here at this stage may only be used as a guide, because they only provide half the story. The complete effect of water on ligands may only be observed by comparing water behavior with and without ligands present. Nevertheless, the 10-ns molecular dynamics simu-

lation has provided valuable insight into how water behaves in the active site of an enzyme, information that is currently beyond the scope of experimental techniques.

We are grateful to Dr. Tjerk Straatsma for his continuing help with the molecular dynamics components of the NWChem software, Accelrys for providing the Insight II 2000 software, Drs. Nathan Baker, Stephen Bond, and Palmer Taylor for helpful discussions, and Drs. Yves Bourne and Pascale Marchot for providing the latest x-ray structure of mAChE.

K.T. holds a fellowship from the La Jolla Interfaces in Science program, which is supported in part by the Burroughs Wellcome Fund. This work has also been supported in part by grants from the NSF, NIH, and the San Diego Supercomputer Center. Additional support has been provided by NBCR and the W. M. Keck Foundation.

REFERENCES

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Bourne, Y., P. Taylor, and P. Marchot. 1995. Acetylcholinesterase inhibition by fasciculin: crystal structure of the complex. *Cell.* 83:503–512.

Brooks III, C. L., and M. Karplus. 1989. Solvent effects on protein motion and protein effects on solvent motion: dynamics of the active site region of lysozyme. *J. Mol. Biol.* 208:159–181.

Brunne, R., E. Liepinsh, G. Otting, K. Wuthrich, and W. F. van Gunsteren. 1993. Hydration of proteins: a comparison of experimental residence times of water molecules solvating the bovine pancreatic trypsin inhibitor with theoretical model calculations. *J. Mol. Biol.* 231:1040–1048.

Carugo, O., and D. Bordo. 1999. How many water molecules can be detected by protein crystallography? *Acta Crystallogr.* D55:479–483.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A 2nd generation force-field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Darden, T. D., and P. L. York. 1993. Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98: 10089–10092.

Denisov, V. P., and B. Halle. 1996. Protein hydration dynamics in aqueous solution. *Faraday Disc.* 103:227–244.

Enyedy, I. J., I. M. Kovach, and B. R. Brooks. 1998. Alternate pathways for acetic acid and acetate ion release from acetylcholinesterase: a molecular dynamics study. *J. Am. Chem. Soc.* 120:8043–8050.

Fuxreiter, M., and A. Warshel. 1998. Origin of the catalytic power of acetylcholinesterase: computer simulation studies. *J. Am. Chem. Soc.* 120:183–194.

Garcia, A. E., and G. Hummer. 2000. Water penetration and escape in proteins. *Proteins.* 38:261–272.

Gillilan, R. E., and F. Wood. 1995. Visualization, virtual reality, and animation within the data flow model of computing. *Comput. Graphics.* 29:55–58.

Gilson, M. K., T. P. Straatsma, J. A. McCammon, D. R. Ripoll, C. H. Faerman, P. H. Axelsen, I. Silman, and J. L. Sussman. 1994. Open "back door" in a molecular dynamics simulation of acetylcholinesterase. *Science.* 263:1276–1278.

Goodford, P. J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 28:849–857.

Grutzendler, J., and J. C. Morris. 2001. Cholinesterase inhibitors for Alzheimer's disease. *Drugs.* 61:41–52.

Hasinoff, B. B. 1982. Kinetics of acetylthiocholine binding to electric eel acetylcholinesterase in glycerol/water solvents of increased viscosity: evidence for a diffusion-controlled reaction. *Biochim. Biophys. Acta.* 704:52–58.

Helms, V., and R. C. Wade. 1998. Hydration energy landscape of the active site cavity in cytochrome P450cam. *Proteins.* 32:381–396.

Impey, R. W., P. A. Madden, and I. R. McDonald. 1983. Hydration and mobility of ions in solution. *J. Phys. Chem.* 87:5071–5083.

Koellner, G., G. Kryger, C. B. Millard, I. Silman, J. L. Sussman, and T. Steiner. 2000. Active-site gorge and buried water molecules in crystal structures of acetylcholinesterase from *Torpedo californica*. *J. Mol. Biol.* 296:713–735.

Lam, P. Y. S., P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C. H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, and S. Ericksonviitanen. 1994. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science.* 263:380–384.

Likic, V. A., and F. G. Prendergast. 2001. Dynamics of internal water in fatty acid binding protein: computer simulations and comparison with experiments. *Proteins.* 43:65–72.

Makarov, V. A., B. K. Andrews, P. E. Smith, and B. Pettitt. 2000. Residence times of water molecules in the hydration sites of myoglobin. *Biophys. J.* 79:2966–2974.

Ni, H., C. Sotriffer, and J. A. McCammon. 2001. Ordered water and ligand mobility in the HIV-1 integrase-5CITEP complex: a molecular dynamics study. *J. Med. Chem.* 44:3043–3047.

Otting, G., and E. Liepinsh. 1995. Protein hydration viewed by high-resolution NMR spectroscopy: implications for magnetic resonance image contrast. *Accounts Chem. Res.* 28:171–177.

Radić, Z., P. D. Kirchhoff, D. M. Quinn, J. A. McCammon, and P. Taylor. 1997. Electrostatic influence on the kinetics of ligand binding to acetylcholinesterase: distinctions between active center ligands and fasciculin. *J. Biol. Chem.* 272:23265–23277.

Rosenberry, T. L. 1975. Catalysis by acetylcholinesterase: evidence that the rate-limiting step for acylation with certain substrates precedes general acid-base catalysis. *Proc. Natl. Acad. Sci. U.S.A.* 72:3834–3838.

Roux, B., M. Nina, R. Pomes, and J. C. Smith. 1996. Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys. J.* 71:670–681.

Ryckaert, J.-P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341.

Savage, H., and A. Wlodawer. 1986. Determination of water-structure around biomolecules using x-ray and neutron-diffraction methods. *Methods Enzymol.* 127:162–183.

Schoenborn, B. P., A. Garcia, and R. Knott. 1995. Hydration in protein crystallography. *Prog. Biophys. Mol. Biol.* 64:105–119.

Soreq, H., and S. Seidman. 2001. Acetylcholinesterase: new roles for an old actor. *Nat. Rev. Neurosci.* 2:294–302.

Straatsma, T. P., M. Philippopoulos, and J. A. McCammon. 2000. NWChem: exploiting parallelism in molecular simulation. *Comp. Phys. Commun.* 128:377–385.

Sussman, J. L., M. Harel, F. Frolow, C. Oefner, A. Goldman, L. Toker, and I. Silman. 1991. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. *Science.* 253: 872–879.

Tai, K., T. Shen, U. Börjesson, M. Philippopoulos, and J. A. McCammon. 2001. Analysis of a 10-nanosecond molecular dynamics simulation of mouse acetylcholinesterase. *Biophys. J.* 81:715–724.

Tara, S., V. Helms, T. P. Straatsma, and J. A. McCammon. 1999a. Molecular dynamics of mouse acetylcholinesterase complexed with huperzine A. *Biopolymers.* 50:347–359.

Tara, S., T. P. Straatsma, and J. A. McCammon. 1999b. Mouse acetylcholinesterase unliganded and in complex with huperzine A: a comparison of molecular dynamics simulations. *Biopolymers.* 50:35–43.

Vagedes, P., B. Rabenstein, J. Aqvist, J. Marelius, and E. W. Knapp. 2000. The deacylation step of acetylcholinesterase: computer simulation studies. *J. Am. Chem. Soc.* 122:12254–12262.

Wiesner, S., E. Kurian, F. G. Prendergast, and B. Halle. 1999. Water molecules in the binding cavity of intestinal fatty acid binding protein: dynamic characterization by water ^{17}O and ^{2}H magnetic relaxation dispersion. *J. Mol. Biol.* 286:233–246.

Williams, M. A., J. M. Goodfellow, and J. M. Thornton. 1994. Buried waters and internal cavities in monomeric proteins. *Protein Sci.* 3:1224–1235.

Wittbrodt, E. T. 1997. Drugs and myasthenia gravis: an update. *Arch. Intern. Med.* 157:399–408.

Wlodek, S. T., T. W. Clark, L. R. Scott, and J. A. McCammon. 1997. Molecular dynamics of acetylcholinesterase dimer complexed with tacrine. *J. Am. Chem. Soc.* 119:9513–9522.

Wong, C. F., and J. A. McCammon. 1986. Computer simulation and the design of new biological molecules. *Israel J. Chem.* 27:211–215.

Zhang, L., and J. Hermans. 1996. Hydrophilicity of cavities in proteins. *Proteins.* 24:433–438.

Zhou, H. X., J. M. Briggs, and J. A. McCammon. 1996. A 240-fold electrostatic rate-enhancement for acetylcholinesterase-substrate binding can be predicted by the potential within the active site. *J. Am. Chem. Soc.* 118:13069–13070.