Lifetime of a Hydrogen Bond in Aqueous Solutions of Carbohydrates

Tim Astley,*,†,§ Gordon G. Birch,‡ Michael G. B. Drew,† and P. Mark Rodger†

Chemistry Department, University of Reading, Whiteknights, Reading, RG6 6AD, U.K., and Department of Food Science and Technology, University of Reading, Whiteknights, Reading, RG6 6AP, U.K.

Received: January 4, 1999

In this paper we present a novel method of calculating hydrogen bond lifetimes that circumvents many of the problems encountered with existing methods. The method is based on a three-level model in which the pair of molecules are hydrogen bonded, not hydrogen bonded, or in an intermediate state; hydrogen bonds are identified within the intermediate state according to how the dimer enters and leaves the intermediate state. An effective dimer interaction energy is used to define these three states. The hysteresis introduced by this definition makes it possible to distinguish between large-amplitude vibrations in the hydrogen-bonding coordinates and real bond-breaking events. The method is applied to a dilute aqueous glucose solution and shown to generate hydrogen bond lifetimes that are in accord with both chemical intuition and experiment.

Introduction

Most discussions of the structure of liquid water emphasize the dominant influence of hydrogen bonds and many have suggested, following Bernal and Fowler's seminal work,[1] that the instantaneous structure of the liquid can be modeled by a random hydrogen-bonded network. Thermal excitation of the network may be represented[2] either by broken bonds, as in the several "discrete" models, or in terms of intact but variously distorted bonds as in the "continuum" models. Thus, when a solute which can potentially form hydrogen bonds is introduced, it is desirable to ascertain the strength and durability of any hydrogen bonds formed between it and water. In our present research into the mechanism of the sweet taste response,[3] we have studied the interaction between monosaccharides and water, and in this paper we address especially the nature of the hydrogen bonds formed.

The most commonly encountered H-bond network is that of liquid water;[4,5] this is roughly characterized by a tetrahedral arrangements of bonds, where on average about 50% of the molecules engage in four H-bonds, 30% engage in three bonds, and smaller percentages participate in one, two, or five bonds. Any dynamic properties are, to a large extent, manifestations of the underlying dynamics of the H-bond network. Direct studies of the H-bond dynamics are thus of considerable interest. In the past, such studies focused primarily on the mean H-bond lifetime, t_{HB}. As yet, no rigorous way of deducing t_{HB} from experimental measurements has been developed. For liquid water, t_{HB} has been estimated from the line width of the broader Lorentzian in the depolarized Rayleigh light-scattering spectrum.[6] At room temperature a value of t_{HB} of $\sim$0.8 ps has been obtained, in agreement with the asymptotic exponential decay rate obtained from molecular dynamics simulation.[7] But it must be stressed that H-bonding is only one of many interactions that affect Rayleigh scattering line widths, and in general there is no unambiguous or rigorous definition of the mean H-bond lifetime. At a molecular level one can understand this as arising from the difficulty in deciding when fast librational motion breaks an H-bond, and when it merely distorts it. Since the H-bond is a weak bonding interaction, such ambiguous distortions occur frequently, so that the precise definition of when the bond breaks can markedly alter the resulting lifetime. Several definitions of H-bond time correlations have been developed that take into account this breaking and re-forming of the bonds in different ways,[8−10] but none has resolved the problem satisfactorily.

In furthering the understanding of the dynamic behavior of H-bonds, simulations of many systems have been investigated. There is the potential for very accurate calculations of the lifetime of H-bonds using techniques such as molecular dynamics, but to succeed they require a workable definition for an H-bond. The two most common definitions for a hydrogen bond are based on either energetic or geometric criteria. In the energetic case, two molecules are considered to be H-bonded if their interaction energy is lower than a given value, E_{HB}. However, the distribution of interaction energies is not always bimodal[11] and so the choice of E_{HB} remains somewhat subjective. (The distribution obtained in this present work, is given in Figure 2, and is discussed later). In the geometric case, two molecules are considered bonded if the values of pertinent internal dimensions of the dimer are within appropriate ranges. For a water dimer, for example, such dimensions are usually chosen as the oxygen−oxygen distance, R_{OO}, the donor oxygen−hydrogen bonding hydrogen-acceptor oxygen angle θ_H, and/or the lone pair(LP) oxygen−oxygen angle θ_{LP}[12] (see Figure 1). Some earlier work also considered the dihedral angle, between the H−O−O and LP−O−O planes, but this has now been shown to be of little importance.[14] To define these dimensions, the atom (or lone pair) participating in the bond is identified as the hydrogen atom (or lone pair) on the donor (or acceptor) water closest to the oxygen atom of the acceptor (or donor) water. As with the energy definition, the choice of limits for the range of acceptable values is rather arbitrary, although in the geometric case results do not show much sensitivity to the precise value of these limits when they are close to the commonly accepted values, viz., $R_{OO} < 3.5$ Å and $\theta_H < 120°$.

Although these definitions have been useful for analysis of static systems, both cause considerable problems for dynamic

† Chemistry Department.
‡ Department of Food Science and Technology.
§ Present address: School of Chemistry, University of Tasmania, GPO Box 252-75, Hobart, Tasmania 7001, Australia.

10.1021/jp9900279 CCC: $18.00
Published on Web 06/11/1999

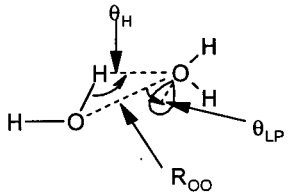

Figure 1. Water dimer interaction showing geometry definitions used to define a hydrogen bond.

properties. Water molecules show a librational motion on a time scale of less than 10^{-12} s, superimposed on slower diffusion and rotational motions, which causes complex time variations in dimer interaction parameters. This results in very large and rapid fluctuations in the geometry and energy of the H-bonded dimers, such that H-bonds defined using either the energetic or geometric definitions appear to break and re-form on an unrealistically short time scale.[7,14,15] The problem here is the time resolution implicit in any definition of a bond: can a bond be said to have broken if it re-forms, say, only 1 fs later, especially if there have been no intervening measurements? This question is important in its own right, but pragmatically one must ignore bond-breaking events that are too short to resolve and that lead to no observable difference in the system. This does still leave a problem in that simulations can resolve events on a femtosecond time scale whereas experiments are limited to ~nanosecond time scales for H-bond dynamics in solution. Thus, one must still consider carefully how to treat events that occur orders of magnitude faster than they can be measured and how this should affect a working definition of the H-bond for use with simulations.

There have been various attempts at tackling this problem. One method[14,16−18] is to measure the length of time, t_C, that a bond remains in "continuous" existence, i.e., how long until the first breakage was observed. This tends to give rise to very short lifetimes and can be affected by the time interval between calculating the state variables. Values of t_C in the range 0.05−0.3 ps have been reported.[15] Another method is to calculate the length of time, t_I, before long-term rearrangements occurred, i.e., elapsed time until the final breakage of the bond occurred, ignoring all prior breakages and re-formations regardless of duration. This definition of a lifetime as an "intermittent" bond is essentially independent of the time interval between measurements, but the values will depend on the total time of observation since an interaction that breaks and re-forms very much later will be recorded as just one very long interaction. Intermediate positions have also been adopted, in which "trivial" H-bond breakages are ignored,[19] without an objective rationale for defining "trivial". Values of t_I between 1 and 10 ps have been reported.[15,20−23] The overall problem with all these methods has been well summarized by Mountain when he wrote: "The Mezei-Beveridge criterion (using a geometry based definition) for hydrogen bonding is a geometrical condition that does not take the lifetime of a configuration into account. It, therefore, counts short duration configurations that are intuitively not consistent with the idea of a bond."[17]

Thus, it is clear a new definition is required that allows for transient breaks in an interaction without assigning them as bond breaks, while allowing for longer-lived distortions to be recognized as legitimate bond breakages. In this work we propose a new definition for the hydrogen bond which successfully accomplishes the above requirements and therefore provides reliable estimates for hydrogen bond lifetimes in aqueous carbohydrate solutions.

Methodology

Molecular dynamics simulations were performed using DL_POLY[24] on SG Indy PCs. In all cases the system consisted of a single α-D-glucose molecule surrounded by 248 water molecules in a cubic box of length 20 Å^3, and using periodic boundary conditions to eliminate edge effects. The potential used for the solute was one specifically developed for carbohydrates[25] and the SPC[26] force field was used to model water. The SPC force field is like most simple models of water in that the water molecule is treated as a rigid body. For the entire system, the parameters in the pairwise additive intermolecular potential incorporate some of the nonpairwise additive effects of the surrounding medium in an average sense and so should depend to some extent on the thermodynamic state of the system. For a pair of sites, α and β, belonging to different molecules, the potential is of the form

$$V_{\alpha\beta}(r) = V^{sr}_{\alpha\beta}(r) + \frac{q_\alpha q_\beta}{4\pi\epsilon_0 r}$$

where r is the intersite separation, q_α is the partial charge on the αth site, and $V^{sr}_{\alpha\beta}$ is the short-range interaction term, in our case a 12-6 Lennard-Jones potential, describing the atomic core repulsion and attractive dispersion interaction. A common feature of models of H-bonding and other polar liquids is a rescaling of partial charges in order to obtain an enhancement of the electric dipole moment over its gas-phase value. This accounts for some of the effects resulting from the molecular polarizability.

All hydrogen atoms were included explicitly in the simulations; that is, "united" or "extended" atom representations were not used. V^{sr} was modulated by a switching function so that long-range interactions vanished smoothly for separations in the range 7.0−8.5 Å.[27] Electrostatic interactions were calculated using an EWALD summation.[28]

Equations of motion were integrated using the Verlet algorithm[29] with a step size of 0.5 fs and saving coordinates every 50 steps. Larger time steps were found to give an inaccurate integration of the hydrogen motion on the carbohydrates and hence were inappropriate for a study of H-bonding. The SHAKE algorithm[30] was used to ensure that all chemical bond lengths involving a hydrogen atom and all solvent bond angles remained constant throughout each simulation.

To generate the simulation box a solute molecule was placed in the center of a box of water molecules whose positions were obtained by taking a configuration from a simulation of pure liquid water and then minimizing the energy with respect to the position of all the water molecules. Any water molecule in which the oxygen atom overlapped any atom in the solute was removed. An overlap was deemed to occur when, to within a given tolerance, two atoms were closer than the sum of their van der Waals radii; the tolerance was then defined to ensure the system gave a density close to that previously reported for SPC systems of 0.971 g cm^{-3}.[31] The resulting system was again minimized to remove any "hot spots" caused by close contacts. Equilibration was then ensured by alternately freezing the solvent then solute while allowing the rest of the system to relax over about 5 ps, this being followed by a period where the entire system was allowed to reach equilibrium (defined by constant total energy and temperature) before the production run was undertaken. The final simulation was 100 ps in length and took one week to generate on a Silicon Graphics R4600 IndyPC.

Calculation of H-Bond Energies. We have adopted an energetic definition of the H-bond as the initial basis of this

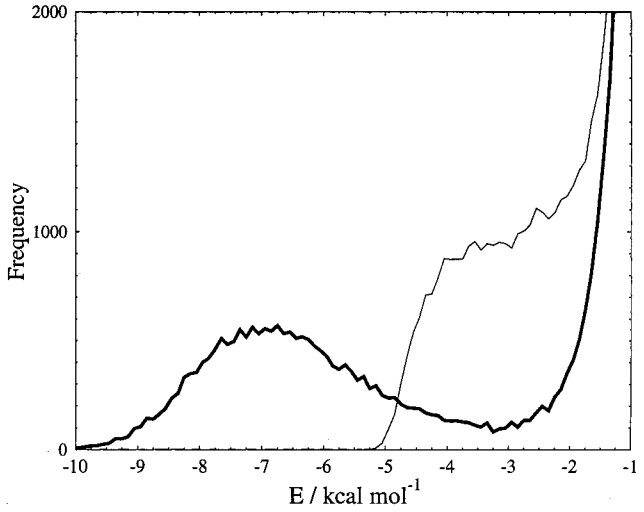

Figure 2. Frequency histogram for interaction energy between each carbohydrate C−OH moiety and each water molecule. Donor (dark) and acceptor (light) hydrogen bond interactions are defined by geometry.

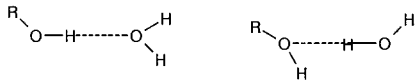

Figure 3. Geometry of a donor and acceptor hydrogen bond as defined for this study. According to the Geometry definition, a H-bond exists when $R_{OO} < 3.5$ Å, and one of the angles θ_D or θ_A is in the range $120 < \theta < 180°$.

study. Further, since the solute contains many potential H-bonding sites, it is ultimately desirable to analyze each such site separately. We have therefore chosen to monitor H-bonds through a pseudo dimer potential energy. The pseudo dimer was defined as a water molecule and a COH fragment from the solute. For the purposes of calculating an energy, the charge associated with the C was then adjusted to give a neutral COH group. The energies were then calculated for every COH−H_2O dimer geometry found in the simulations. All interactions were classified as either donor or acceptor, based on whether the distance between the H on the carbohydrate and the water oxygen was less than both of the water hydrogen to carbohydrate oxygen distances: if so, the interaction was classed as a donor, otherwise it was deemed an acceptor. This definition will be adhered to for the remainder of this work.

The frequency distribution of donor and acceptor pseudo dimer energies was calculated from our simulations and is depicted in Figure 2. It is clear that the donor interaction is more stable than the acceptor interaction. This arises as a consequence of the charges on the atoms. The charge on the hydrogen atoms in the SPC water (+0.41 *e*) and the COH groups of the carbohydrate (+0.40 *e*) are essentially equal, but there is a substantial difference between the charges on the oxygens: −0.82 *e* for SPC water and −0.55 *e* for COH. The more negative the charge on the H-bonding oxygen, the lower the interaction energy and thus it is the H_2O···HOC "bond" that is stronger, i.e., when the carbohydrate is the donor (see Figure 3 for diagram of acceptor and donor H-bonds). An extensive search of the Cambridge Crystallographic Database revealed that the donor interactions are on average 0.1 Å shorter than acceptor interactions (Table 1) consistent with the donor bond being a stronger interaction. Also, the minimum acceptor distance is only 2.689 Å (out of 298 examples) whereas 13% of the donor hydrogen bonds were shorter than 2.689 Å, with the shortest being 2.615 Å.

TABLE 1: H-Bond Lengths between Carbohydrates and Water from the Cambridge Crystallographic Database[a]

	donor					acceptor				
atom	mean	SD	min	max	no.	mean	SD	min	max	no.
O1	2.809	0.078	2.623	2.977	43	2.836	0.104	2.703	3.078	39
O2	2.768	0.090	2.620	3.192	46	2.878	0.137	2.701	3.192	82
O3	2.769	0.082	2.615	3.145	59	2.895	0.135	2.689	3.192	110
O4	2.779	0.010	2.762	2.790	8	3.067	0.000	3.067	3.067	1
O6	2.788	0.097	2.597	2.985	47	2.842	0.123	2.709	3.175	64

[a] All measurements in Å; atom numbering as in Figure 4.

Ab initio calculations, using GAUSSIAN94,[32] were also used to study these effects. Two models of glucose were constructed, each with a water molecule hydrogen bonded to O(4). In the first model, the carbohydrate acted as an acceptor and in the second as a donor. HF calculations were performed with the 6-31G* basis set and the structures were geometry-optimized. The water molecules were then removed and the carbohydrates were geometry-optimized. The difference in energy between the hydrogen-bonded structure and the individual carbohydrate and water molecules were then calculated. The energy difference E(carbohydrate+H_2O) − E(carbohydrate) − E(H_2O) was −0.0144 E_h (−9.04 kcal mol^{-1}) for the carbohydrate acceptor and −0.0090 E_h (−5.65 kcal mol^{-1}) for the donor, consistent with our modeling results.[33] It was also noted that the O···O distance was significantly shorter for the acceptor: 2.900 Å compared with 2.952 Å for the donor. Charges on the atoms in the hydrogen-bonded complex were calculated via the electrostatic potential method and values were consistently greater for the water oxygen (−0.90 *e*) compared with the carbohydrate oxygen atoms (−0.80 *e*). While these values are consistent with our methodology we did not feel it appropriate to impose these charges on the molecules in our simulation because the values would not be consistent with the other parameters in the force field, which had been optimized with the original charges.

Results and Discussion

The problem of finding a workable definition of a H-bond for dynamic situations has evaded solution for 20 years. Our approach has been to examine the time evolution of specific dimers and to identify regions that, to any reasonable person, must be H-bonds. From these intuitively clear examples we have developed a set of rules for hydrogen bond formation and breakage; the new rules have then been tested against specific dimer interactions for which the classification is more problematic.

To monitor each specific dimer we used the five conventional parameters used in either the geometric or energetic definitions: E_D the dimer interaction energy, R_{OO}, the oxygen−oxygen distance, and the three O−H···O angles θ_D, θ_{A1}, and θ_{A2}, defined in Figure 3. In the following figures (Figures 5−10) each of these parameters is plotted throughout the entire duration of the simulation for selected C−OH···H_2O dimers. Of the 1240 separate C−OH···water interactions in the simulation under investigation we have chosen six to illustrate some of the different behavior observed. These six have been selected as instructive examples against which to prove our hydrogen bond definition. The atom numbering for the sugar is taken from Figure 4, while the different water molecules involved are differentiated by the use of a subscript letter.

Clear-cut Examples. *Example 1.* Figure 5 shows the interaction between O4 and water molecule W_a. From the energy trace this trajectory can be divided into three main sections: 2−55 ps and 75−100 ps in which E_D is close to zero, and 55−75 ps

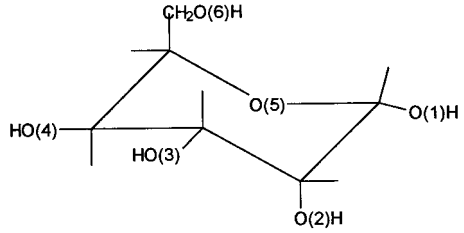

Figure 4. General structure of monosaccharides with oxygen numbering system highlighted. No attempt has been made to represent the stereochemistry.

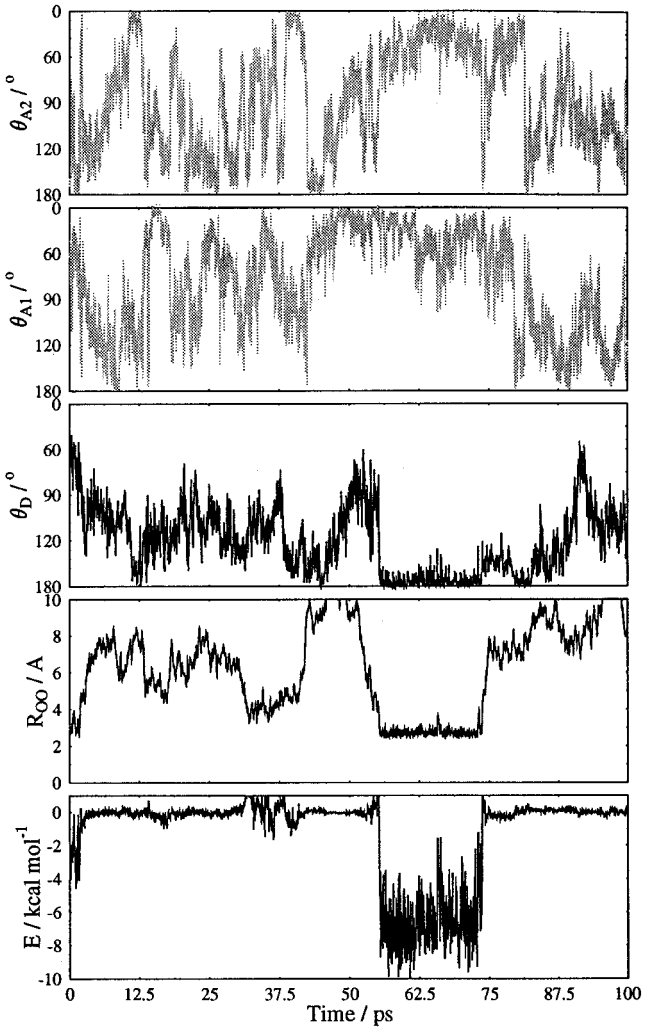

Figure 5. Descriptors for the interaction between O4 and water molecule W_a. Angles and distances as defined in Figure 3. Dark traces refer to parameters that indicate H-bonds in this example.

in which E_D fluctuates in the range -2 to -10 kcal mol^{-1}. From Figure 2 it has already been noted that H-bonding should be associated with dimer energies below ca. -1 kcal mol^{-1}, and so we might expect H-bonding to be associated with the second region only. In general terms this is supported by the geometric data which confirm a donor hydrogen bond during the period 55−75 ps.

In the context of defining the dynamics of H-bonding it is useful to look more closely at Figure 5. Intuitively, it is clear that there should be a single H-bond lasting 20 ps in the 55−75 ps region. Unfortunately, none of the current geometric or energetic definitions will give this. Instead, depending on the choice of threshold values, either they give several short-lived bonds in the 55−75 ps region that are separated by nonbonded

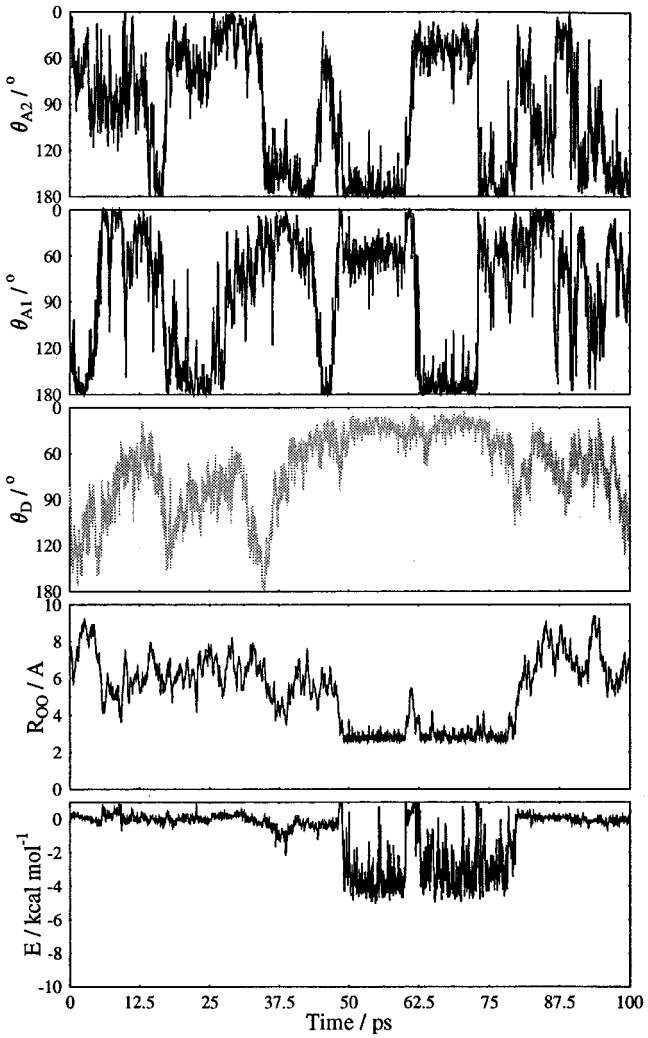

Figure 6. Descriptors for the interaction between O6 and water molecule W_b. Angles and distances as defined in Figure 3. Dark traces refer to parameters that indicate H-bonds in this example.

periods that last for less than 10^{-13} s, or they identify other very short-lived hydrogen bonds associated with the fluctuations around 30−40 ps of the trajectory. For example, with the energy definition, E_{HB} would need to be < -4 kcal mol^{-1} to avoid identifying the early fluctuations as hydrogen bonds, but > -1.75 kcal mol^{-1} to retain the 55−75 ps region as a single H-bond. Similarly, an unrealistically large R_{OO} threshold (3.9 Å) would be needed to make the interpretation of the 55−75 ps period in line with chemical intuition.

Example 2. The next dimer chosen is that between O6 and water molecule W_b; shown in Figure 6. Again, there are two long periods, 0−48 and 80−100 ps, during which $E_D \sim 0$. Between 48 and 80 ps, E_D oscillates between -5 and 1 kcal mol^{-1}. If an H-bond occurs only when $E_D < -1.0$ kcal mol^{-1}, as indicated in Figure 2, then broadly speaking, two H-bonds would occur during this period. This is supported by the geometric traces, which indicate two acceptor H-bonds during this period. A break occurs around 60 ps and this can be associated with the water molecule moving away from the solute, rotating, and moving back toward the solute to re-form an H-bond, this time, however, donating the other H atom. This rearrangement of the H-bond is evident in the traces of θ_{A1} and θ_{A2}. The question of whether this rearrangement should be considered as one long H-bond interaction or two distinct H-bonds will be discussed later. The size of the fluctuations in

R_{OO} and both θ_As is such that any strict geometric definition would lead to this period being broken up into a number of short-lived H-bond interactions. For example, during the period 48−80 ps but ignoring the obvious peak associated with the rearrangement of the H-bond geometry, R_{OO} exceeds 4.0 Å on five occasions. Outside this time period R_{OO} drops briefly below 4.0 Å three times, further highlighting the unsuitability of focusing on R_{OO} to define H-bond lifetimes. The same problems occur with any strict choice for the angles involved. The fluctuations in E_D also prevent an ideal selection of E_{HB}: during the 48−80 ps period (we reiterate that this period should reasonably be analyzed to give at most two separate hydrogen bonds), E_D exceeds −1 kcal mol^{-1} 15 times.

Some Consequences for a Dynamic Hydrogen Bond Definition. It is clear from the rapid and large fluctuations in all the properties normally used to define H-bonds that some sort of data smoothing is required to produce a sensible definition of the H-bond lifetime. The most obvious method of achieving this is to introduce a rolling average of the properties over some time interval to be identified. This method has been used in the past, but as we show in this paper, does not on its own provide an adequate solution to the fluctuation problem: when the time interval is large enough to smooth the fluctuations, it tends to merge H-bonds that should be distinct events; when the interval is chosen to be small enough to resolve separate H-bonds adequately, it still leaves unacceptably large fluctuations in energy, distance, and angles.

In order to circumvent this problem, we have adapted a procedure that has been found to be successful in the theory of solvent effect on reaction kinetics.[34] To explain this model we may consider the formation of a H-bond as analogous to a chemical reaction

$$\text{sugar} + \text{water} \rightarrow \text{sugar}\cdots\text{water}$$

In such a scheme it is useful to identify an intermediate in addition to the product and reactants; some type of reaction coordinate must then be identified to define the reactant, intermediate and product zones. The progress of a reaction must involve the reactants entering the intermediate zone. A reaction event may then be either successful, in which case it continues into the product zone, or unsuccessful, in which case it returns to the reactant zone. The intermediate state may therefore be identified with either the reactant or product according to how it both enters and leaves this intermediate zone.

From a detailed analysis of the individual dimer interactions, we have identified the dimer energy as a suitable candidate for the "reaction coordinate". Non-H-bonding interactions are clearly associated with dimer energies close to zero, whereas H-bonds are associated with low (i.e., large negative) dimer energies. To implement this three-level model we must identify an intermediate range of energies, $E_{on} < E_D < E_{off}$, in which the dimer may or may not be H-bonded. An H-bond is then defined to form when $E_D < E_{on}$, and terminate when $E_D > E_{off}$. Such a definition introduces a hysteresis into the formation of an H-bond that will prove to be very useful in resolving the ambiguities in H-bond lifetimes observed in previous simulations.

Problematic Examples. *Example 3.* Figure 7 shows the traces associated with the interaction between O4 and W_c. The E_D trace again shows two distinct types of behavior. For the first and last 30 ps of the trajectory the energy oscillates about 0, indicative of no hydrogen bonding. In the middle 40 ps the energy oscillates much more widely, between +1 and −4 kcal mol^{-1}, suggestive of a weak acceptor H-bond. The geometric

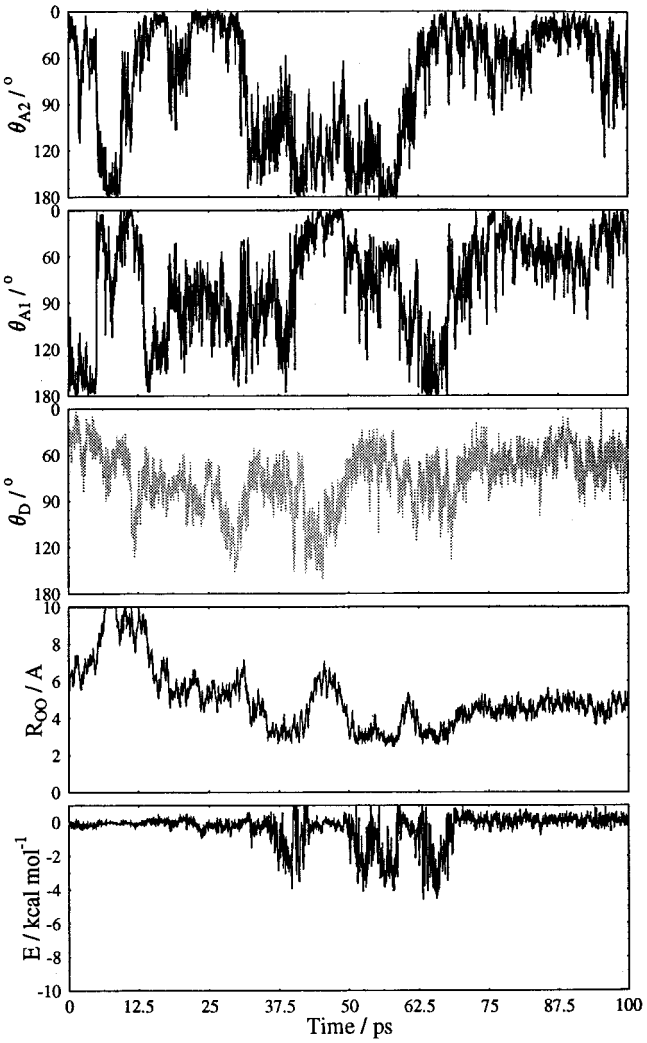

Figure 7. Descriptors for the interaction between O4 and water molecule W_c. Angles and distances as defined in Figure 3. Dark traces refer to parameters that indicate H-bonds in this example.

traces support this assignment and also highlight the weak nature of the interaction, this being especially evident from the lack of a persistent period where either of the θ_A is close to 180°. Even so, there is strong evidence that there are three periods of acceptor H-bonds: 35−40, 50−57, and 60−65 ps. The range of choices of E_{on} is constrained by this example to be between −1.8 and −3.8 kcal mol^{-1}. The upper limit prevents any of the energy spikes prior to 35 ps registering as an H-bond while the lower limit is the minimum E_{on} could be, since that is the minimum energy measured in the first "bonding" period. No values of $E_{off} < 0$ would prevent all three "H-bonds" being terminated prematurely. In each instance E_D exceeds zero at least once, but only for a very short duration and while still retaining O−O distances that are small enough to indicate an H-bond.

Example 4. Figure 8 depicts the energy and geometric data for the interaction between O3 and the water molecule W_d. From both the energy and distance traces it is clear that the significant interactions occur only during the interval 35−70 ps, with no evidence of H-bonding outside this period. All parameters are consistent with a well-defined acceptor H-bond during 35−50 ps and again from 60 to 65 ps. Between these periods the descriptors indicate that the H-bond breaks and re-forms through the alternate H a short time later. This disruption of the H-bond is seen first in the E_D, which drifts up from −4 kcal mol^{-1}

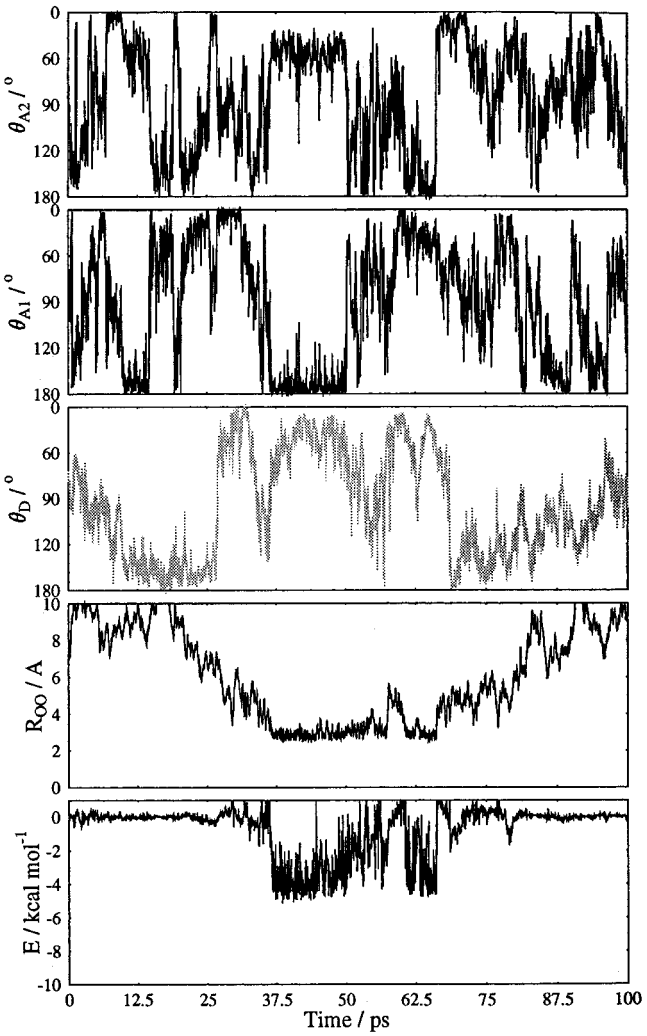

Figure 8. Descriptors for the interaction between O3 and water molecule W_d. Angles and distances as defined in Figure 3. Dark traces refer to parameters that indicate H-bonds in this example.

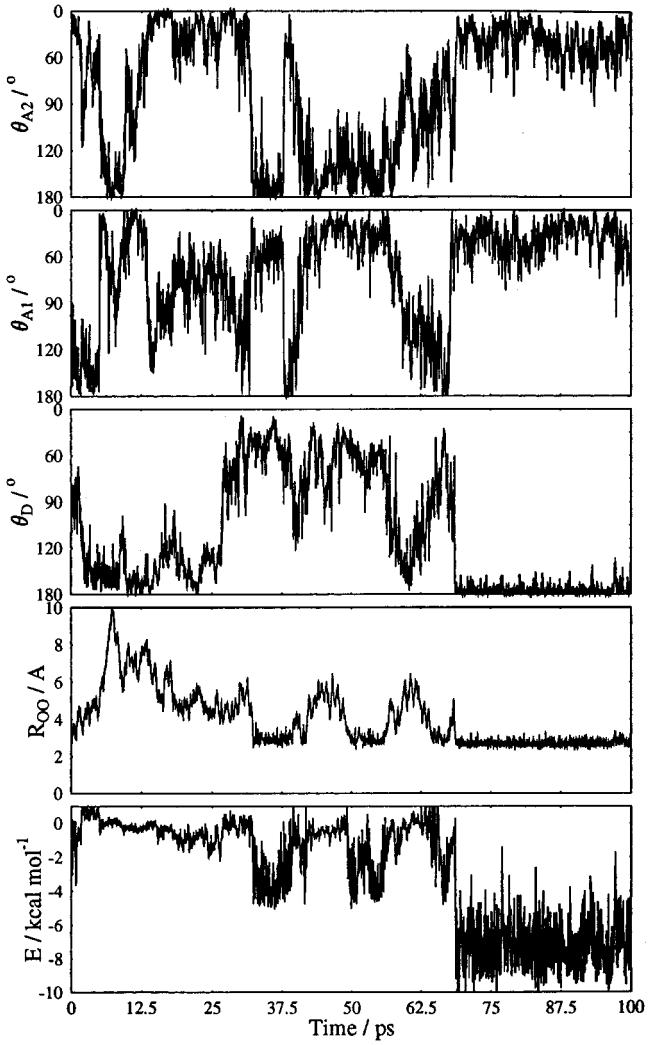

Figure 9. Descriptors for the interaction between O3 and water molecule W_c. Angles and distances as defined in Figure 3. Dark traces refer to parameters that indicate H-bonds in this example.

toward 0, and in θ_{A1} which decreases below 120° simulataneously. After a short delay (~5 ps) there is a corresponding increase in R_{OO} to about 5.5 Å before it returns to ca. 3 Å at about 60 ps. The duration and scale of these deviations from the conventional H-bonding limits is sufficiently large to force this trajectory to be interpreted as generating two separate H-bonding events.

This example further constrains the choice of E_{on}. Beyond the 70 ps mark there are two occasions on which E_D approaches -2 kcal mol^{-1} but with R_{OO} parameters that remain in excess of 4 Å. It is interesting to note that these two low-energy regions persist for periods in excess of 1 ps, and so would survive as low-energy interactions in any analysis based purely on the rolling time average of the H-bond energy; they do not, however, cause any difficulties for the three-level H-bond model advocated here.

Example 5. Figure 9 shows the interaction between O3 and water molecule W_c, the same water molecule as that in the interaction used for the third example (Figure 7). The E_D trace can be broken into three periods. For the first 30 ps the energy remains above -2 kcal mol^{-1}; while this indicates a prolonged favorable interaction between the water and glucose, none of the geometrical parameters are consistent with hydrogen bonding during this time interval. From 30 to 68 ps the energy oscillates between -5 and $+1$ kcal mol^{-1}, with a few periods where it is predominantly at one extreme or the other. In the final section, 68−100 ps, the energy drops to average about -7 kcal mol^{-1} with oscillations between -2 and -10 kcal mol^{-1}. The geometry trace can be divided into three similar periods. These periods can be associated with no interaction, a series of acceptor interactions, and then a donor interaction. As was shown in the discussion of the first clear-cut example (Figure 5), it should be relatively easy to find a suitable choice of geometric or energetic threshold to define the donor H-bond, i.e., $E_{off} < -1.2$ kcal mol^{-1}. However, the acceptor interaction(s) is far less well-defined, with the oscillations in the midst of a bond being substantial. Coupled with the fact that E_D drops below -2 kcal mol^{-1} during the no-interaction period, this will help constraint both E_{on} and E_{off}.

By comparing Figures 7 and 9 it is evident that the water molecule W_c is interacting more strongly with O3 than O4 (E_D has a lower energy minima, smaller oscillations in R_{OO} and θ_A, and the H-bond interactions occur slightly earlier with O3 than O4). It is also interesting to note that when the water molecule bonds in a donor fashion with O3 (Figure 9) this is also evident in the R_{OO} trace for O4 (Figure 7) by a relatively stable, but long, distance of about 4.5 Å; a stable R_{OO} distance of this magnitude was not observed in any other distance traces. Thus, while the water molecule does not seem to be forming a strong acceptor bond with O3, it is interacting with both O3

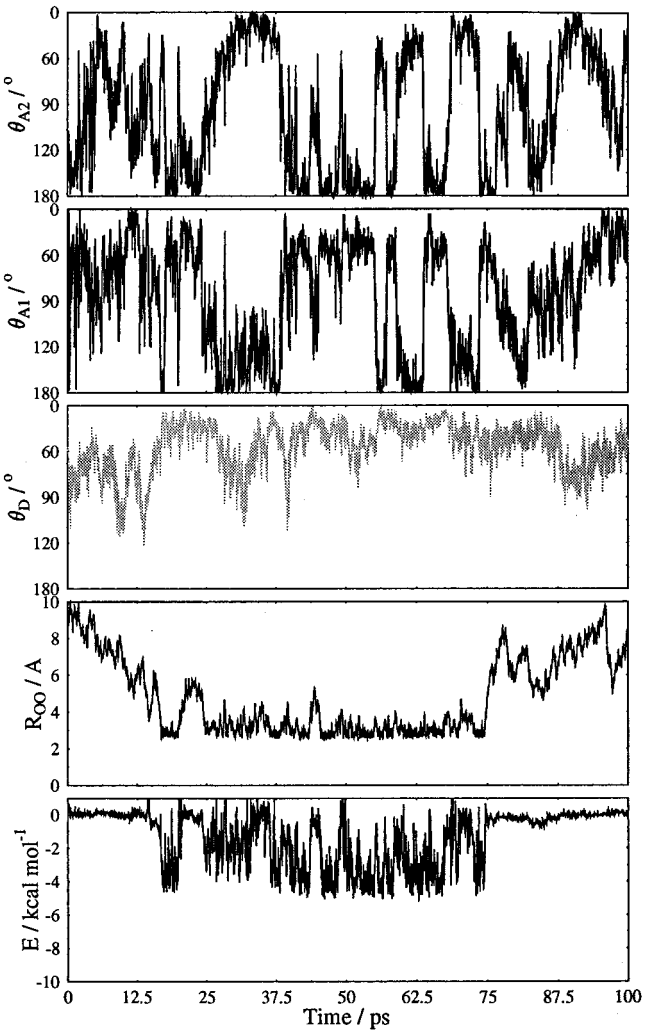

Figure 10. Descriptors for the interaction between O2 and water molecule W_e. Angles and distances as defined in Figure 3. Dark traces refer to parameters that indicate H-bonds in this example.

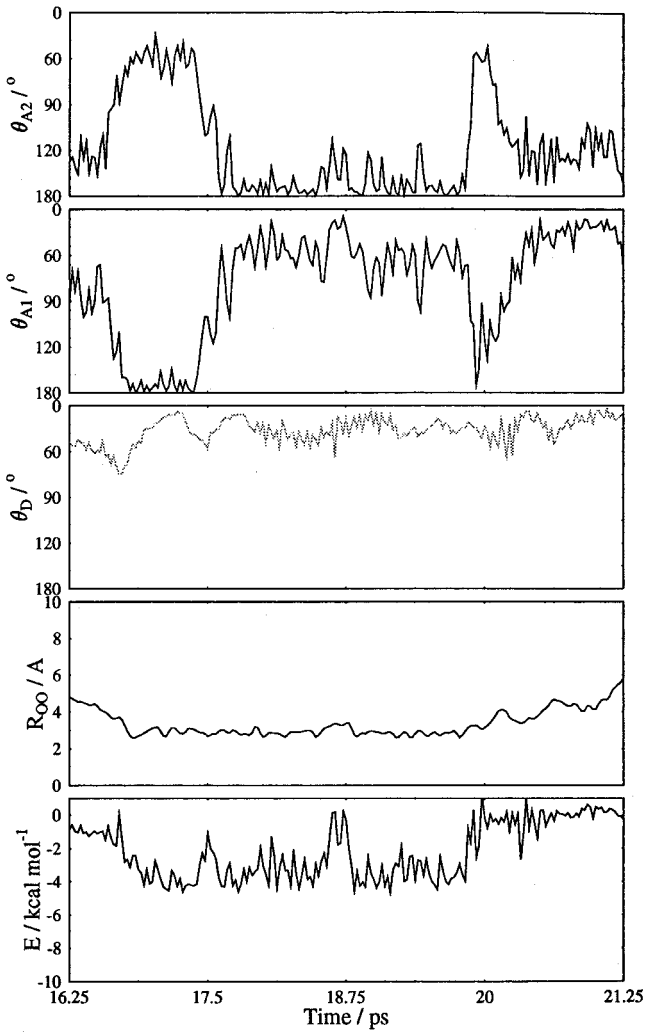

Figure 11. Descriptors for the interaction between O2 and water molecule W_e for a segment of the run. Angles and distances as defined in Figure 3.

and O4 and is favoring O3 during the majority of this simulation. This explains the relative weakness of the interactions in example 3.

Example 6. The next example is shown in Figure 10. This is an interaction between O2 and water molecule W_e. In this example the E_D is ~0 for both the first 15 ps and the final 25 ps of the trajectory. Between times it oscillates in the range −5 to +1 kcal mol^{-1}, suggesting an acceptor hydrogen bond interaction. The geometry traces confirm this, with R_{OO} varying between 2.5 and 6 Å, while θ_{A1} and θ_{A2} appear to be anticorrelated, with one or the other being in the range $\theta >$ 120° most of the time. There seems to be a distinct break in the interaction, from both an energetic (a period where $E_D \sim 0$) and geometric (R_{OO} in excess of 5 Å) definition, during 19−25 ps and again at 43 ps. Both the energetic and the geometric definitions for a hydrogen bond, if applied rigidly, would associate a large number of extremely short-lived interactions with this period, when it is clear from Figure 10 that a more realistic definition should find at most only a couple of interactions.

To highlight one of the problems with defining a hydrogen bond, a short period from 16.25 to 21.25 is shown in greater detail in Figure 11. The main points of interest in this figure are the changes in both θ values around 17.5 ps. It is clear from the traces that the water molecule in question is rotating about an axis perpendicular to the O−O axis such that the hydrogen atom involved in the H-bond is changed without any distinct increase in R_{OO}. There is a short-lived rise in the E_D to ca. −1 kcal mol^{-1}, but this is very short and weak when compared to other energy "spikes" later in the trajectory, e.g. the spikes about 18.75 ps. Both strict geometric and energetic definitions for a H-bond would interpret this as a breakage and subsequent reformation of a H-bond. However, the same definitions would cause the remainder of the interaction (to $t = 20$ ps) to be broken up into a number of very short-lived bonds. On the other hand, less restrictive energetic or geometric definitions would cause too many fortuitous long-range interactions to be counted as H-bonds.

A Three-Level Energetic Definition. To obtain a workable definition for E_{on} and E_{off}, it is still necessary to take a rolling average of E_D. This removes the very short-lived fluctuations that complicate the choice for E_{off}. Figure 12 shows the effect of averaging over 5, 10, 25, and 50 fs the dimer interaction shown in Figure 11. Obviously the larger the time interval incorporated into a rolling average the fewer the spikes, but at the same time there is concomitant loss of fidelity in the signal. Rolling averages generated in excess of 25 fs must be rejected, since they merge the two regions that were identified above as separate H-bonds. Similarly, rolling averages over 10 fs or less do not remove the high-energy spikes that lead to excessively

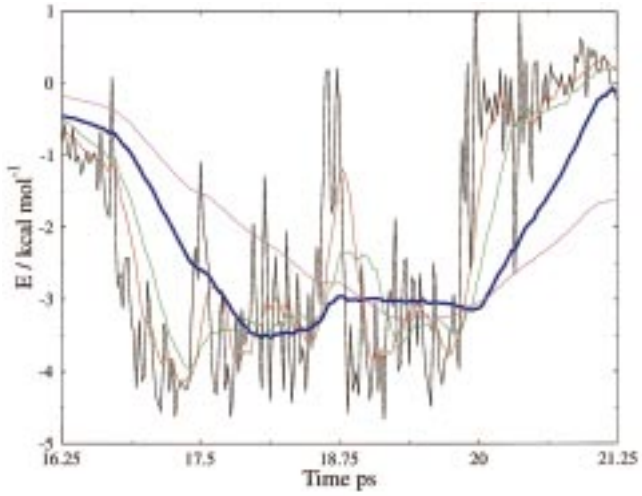

Figure 12. Plot of the interaction between O2 and water molecule W_e. Original energy (black) and that averaged over 5 (red), 10 (green), 25 (thick blue), and 50 (magenta) fs.

TABLE 2: Average Lifetimes of Acceptor (a) and Donor (d) Hydrogen Bonds around Each Oxygen of Glucose in a 100 ps Simulation, Using Different Length Rolling Averages[a]

	length of rolling average (fs)			
	5	12.5	25	50
O1(a)	2.26	3.05	5.13	6.96
O1(d)	17.74	17.76	20.89	21.89
O2(a)	2.43	3.87	5.62	9.26
O3(a)	2.47	4.57	6.03	7.34
O3(d)	11.58	13.62	14.05	16.38
O4(a)	2.89	4.76	5.44	7.09
O4(d)	12.58	14.55	20.17	20.47
O6(a)	2.92	4.72	7.43	9.23
O6(d)	8.35	9.99	16.24	16.59

[a] Atom numbering as per Figure 4. All times in ps. Note: For O2 only 2 donor H-bonds were detected and thus it is omitted since the average is not statistically valid.

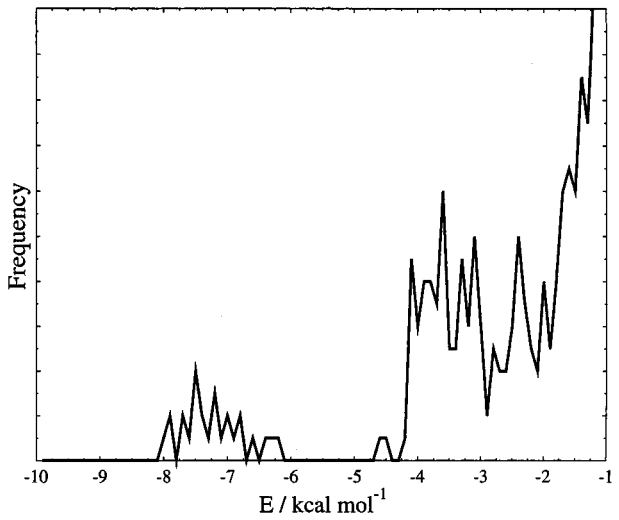

Figure 13. Frequency histogram for the lowest interaction energy observed between the carbohydrate hydroxyl moieties and water molecules.

short H-bond lifetimes. We have therefore adopted a rolling average over 25 fs as an appropriate compromise between these two extremes. Table 2 shows the effect of the time interval chosen on the average H-bond lifetimes eventually calculated. While the lifetime of the acceptor interactions continued to rise

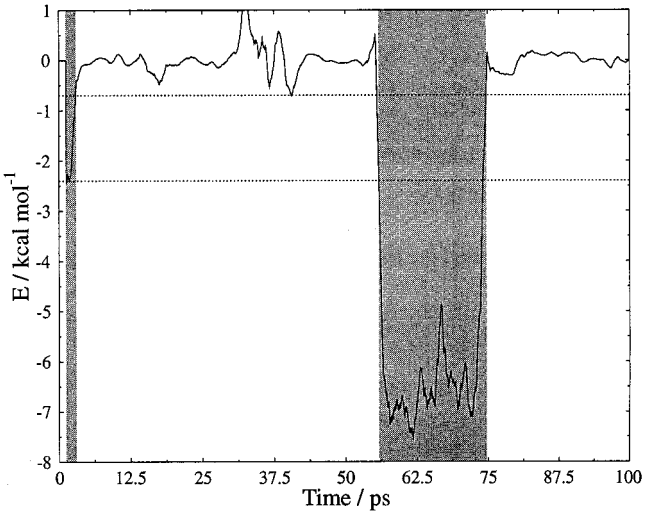

Figure 14. Interaction energy between O4 and water molecule W_a. Energy has been averaged over 25 fs. Shaded areas indicate H-bond interactions.

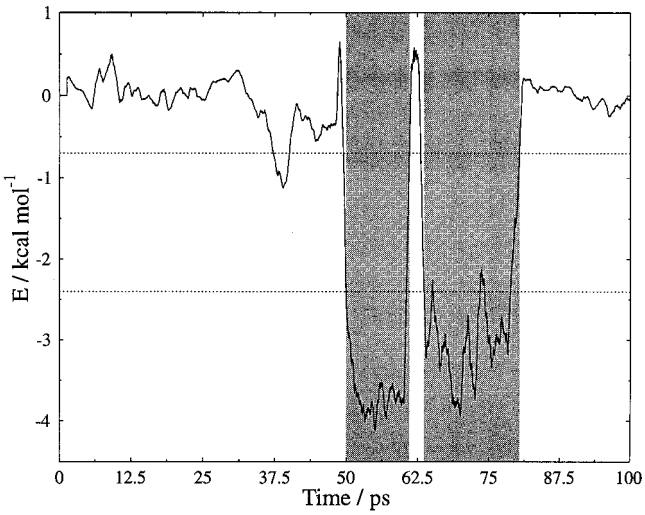

Figure 15. Interaction energy between O6 and water molecule W_b. Energy has been averaged over 25 fs. Shaded areas indicate H-bond interactions.

with larger time intervals for averaging, the donor interactions reach a plateau at about 25 fs, supporting the choice of 25 fs made above.

Having adopted a standard for the rolling average, it is still necessary to define the bounds E_{on} and E_{off}, for the intermediate region in our three-level model. To assist in their assignment, the minimum interaction energy observed for each carbohydrate OH with water during the simulation was calculated, and the resulting frequency distribution is presented in Figure 13. It is clear that there are three clusters of minimum interaction energies. One cluster occurs between -8 and -6 kcal mol^{-1} and can be associated with donor interactions. The next cluster is between -4.75 and -2.5 kcal mol^{-1}, although the upper bound is not well-defined; this can be associated with acceptor interactions. The final cluster is by far the most common and is associated with the normal long-range non-H-bonded interactions. The clear distinction between donor and acceptor interactions allows an unequivocal assignment of any interaction that drops below -4.75 kcal mol^{-1} as belonging to a donor interaction. Combining the limits for the first and second cluster and following the points identified in the discussion of problematic examples, we have defined an H-bond as starting

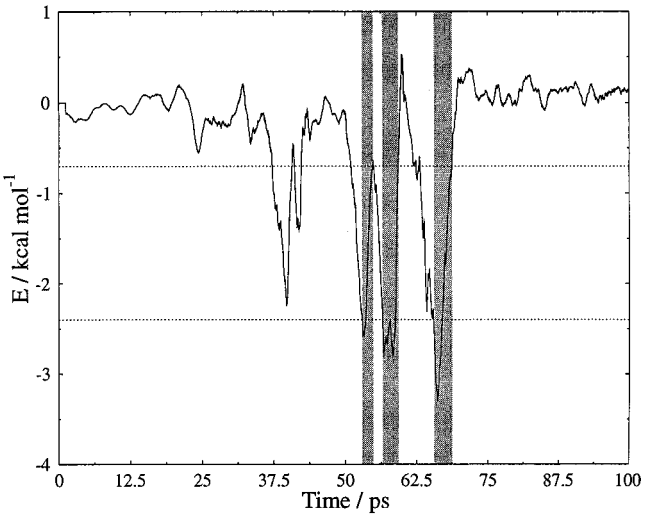

Figure 16. Interaction energy between O4 and water molecule W_c. Energy has been averaged over 25 fs. Shaded areas indicate H-bond interactions.

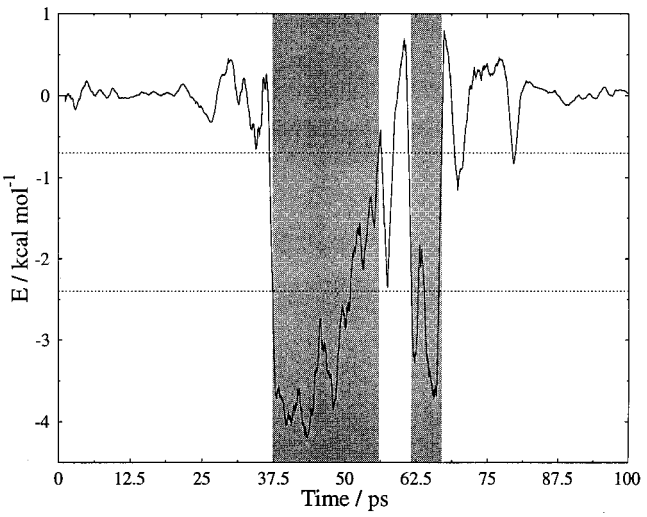

Figure 17. Interaction energy between O3 and water molecule W_d. Energy has been averaged over 25 fs. Shaded areas indicate H-bond interactions.

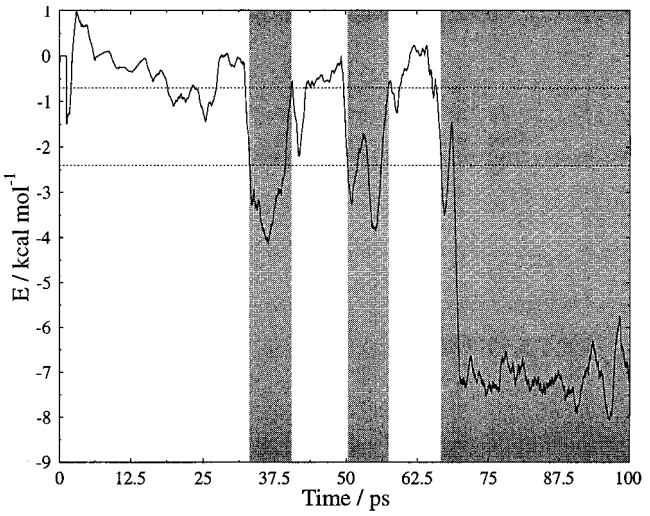

Figure 18. Interaction energy between O3 and water molecule W_c. Energy has been averaged over 25 fs. Shaded areas indicate H-bond interactions.

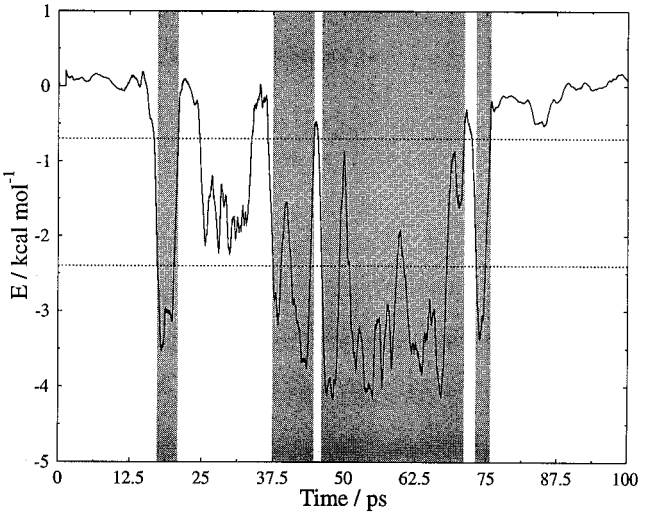

Figure 19. Interaction energy between O2 and water molecule W_e. Energy has been averaged over 25 fs. Shaded areas indicate H-bond interactions.

when the energy dropped below -2.4 kcal mol^{-1} (E_{on}) and stopping when the energy rose above -0.7 kcal mol^{-1} (E_{off}). If at any time in this period the energy dropped below -4.75 kcal mol^{-1} the bond was assigned as a donor, otherwise it was defined to be an acceptor.

To validate this choice of energy bounds, the energy traces shown in figures 5 to 11 were individually analyzed to assess whether the results obtained were consistent with the detailed discussion given above, and in particular, whether the number and duration of lifetimes were reasonable. The results of these calculations are shown in Figures 14−19; an H-bond interaction is shown by the gray shaded portion of the plot and the two energy thresholds are indicated by horizontal dotted lines. In each case the new definition succeeds in generating lifetimes

TABLE 3: Lifetimes of Acceptor (a) and Donor (d) Hydrogen Bonds around Each Oxygen of Glucose in a 100 ps Simulation[a]

	geometry definition				three-level model			
	number	average	minimum[b]	maximum	number	average	minimum	maximum
O1(a)	583	0.18	0.025	4.800	18	5.13	1.325	11.075
O1(d)	217	0.48	0.025	16.300	5	20.89	9.600	27.575
O2(a)	906	0.16	0.025	3.780	27	5.62	1.180	24.925
O2(d)	102	0.80	0.025	35.425	2	50.01	18.66	81.55
O3(a)	1030	0.16	0.025	7.650	26	6.03	1.025	18.580
O3(d)	122	0.51	0.025	10.550	8	14.05	3.300	33.280
O4(a)	608	0.24	0.025	6.600	26	5.44	0.925	13.900
O4(d)	109	0.88	0.025	34.525	5	20.17	3.550	51.350
O6(a)	880	0.20	0.025	17.050	20	7.43	1.080	22.225
O6(d)	200	0.41	0.025	10.950	6	16.24	8.700	45.225

[a] Atom numbering as per Figure 4. All times in ps. [b] All oxygens had numerous interactions that lasted just one step of the trajectory so the minimum was limited to the size of each step of the trajectory.

that are completely consistent with the above discussion, and so we conclude that our new definition succeeds in giving sensible H-bonds with intuitively reasonable lifetimes.

Two of the more difficult aspects identified above are worthy of further discussion. In example 5 we identified what appeared to be a "series" of acceptor H-bonds in the period 35−60 ps; from our new analysis we conclude that there are in fact just two acceptor H-bonds, lasting from 35 to 50 ps and 50−60 ps, respectively. A third possible acceptor bond at about 65 ps is actually part of the ensuing strong donor H-bond. Again, in example 6 we noted that all the primary H-bond descriptors were very noisy in the period 35−80 ps and that this made it very difficult to identify the duration of H-bonds in this trajectory using conventional definitions. From our analysis it becomes clear that this is dominated by one long acceptor H-bond that is separated from two shorter duration bonds by breaks of about 2 ps.

As a third example we return to the question of whether a hydrogen bond should be deemed to have broken when the dynamics indicate a change in which hydrogen is involved in hydrogen bond. In this case our results indicate that the answer is sometimes yes, and sometimes no. We first raised the question in connection with example 2, and our subsequent analysis indicates that in this case the perturbations associated with the rearrangement of H atoms are sufficiently large to ensure that the hydrogen bond must first break before the rearrangement can occur. However, the other examples provide at least two instances in which such rearrangements occur without breaking the H-bond. The first of these is in example 5 (at ca. 65 ps) when an apparently short-lived acceptor bond is subsumed within a strong donor bond without breaking. The second occurs in example 6 (16−21 ps) where the water hydrogens interchange without breaking the acceptor H-bond: the rearrangement involves no lengthening of R_{OO} and no significant change in E_D. Such events emphasize the advantage of choosing an H-bond descriptor such as energy, which is not affected by such permutations of the atoms involved in the hydrogen bond.

Numerical values for the H-bond lifetimes calculated using both the standard geometry definition ($R_{OO} = 3.5$ Å and $\theta_H = 120°$) and our definition are given in Table 3. It is clear that a substantial increase in the perceived lifetime of the hydrogen-bonding interactions has been realized. The new definition gives an average of 5.93 ps for an acceptor interaction and 19.81 ps for a donor interaction, compared with 0.19 and 0.62 ps when using the conventional geometry definition. As outlined in the Introduction, lifetimes in the subpicosecond range are unrealistically small.

Conclusion

The definition and calculation of H-bond lifetimes have always posed a problem for computational chemists. By looking individually at each of the parameters commonly used to define a H-bond it has proved possible to understand more fully the problems encountered using previous definitions. Using the novel idea of a three-level energy definition, we have avoided the problem of large-amplitude vibrations in the hydrogen bond coordinates being identified as a break in the bond. At the same time, this new definition does retain sufficient resolution to identify real short-lived H-bonds. The lifetimes gained using our new definition are commensurate with the experimental lifetimes, which are in the order of picoseconds. The new definition presented here is sufficiently robust and precise to resolve differences in behavior between different solute oxygens and thus should prove an invaluable tool in better understanding the interaction between hydrogen-bonding solutes and water.

One possible inconvenience with our new definition is that the choice of parameters (E_{on} and E_{off}) will be system dependent. Where the typical strength of the hydrogen bond interactions differ from those between glucose and water, different values for these parameters will need to be derived. However, the derivation of such parameters is unlikely to be a major problem. We have already noted that the lower energy bound, E_{on}, is usually identifiable from the probability distribution of dimer interaction energies. We note further that since the uncertainty in H-bond definition arises from thermal fluctuations in the dimer geometry, then one must expect the difference between E_{on} and E_{off} to be about 2−3 RT (i.e., to encompass 95−99% of thermal fluctuations). Thus, good initial estimates of the defining parameters are readily available. Thus, we conclude that the three-level model for H-bond lifetimes is viable to implement in other systems, and reiterate that it gives a conceptually simple resolution to the problem of short-lived fluctuations in calculating H-bond lifetimes.

Acknowledgment. This work was supported by the E.C (AIR program CT94-2107) and the U.K. B.B.S.R.C. grant 45/F02510.

References and Notes

(1) Bernal, J. D.; Fowler, R. H. *J. Chem. Phys*. **1933**, *1*, 515.
(2) Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Oxford University Press: London, 1969.
(3) Astley, T.; Birch, G. G.; Drew, M. G. B.; Rodger, P. M.; Wilden, G. R. H. *Food Chem.* **1996**, *56*, 231.
(4) Beveridge, D. L.; Mezei, M.; Mehrotra, P. K.;, Marchese, F. T.; Ravi-Shanker, G., et al. *Adv. Chem. Ser.* **1983**, *204*, 297−351.
(5) Geiger, A.; Stillinger, F. H.; Rahman, A. *J. Chem. Phys.* **1979**, *70*, 4185−93.
(6) Chen, S.-H.; Teixeira, J. *Adv. Chem. Phys.* **1986**, *64*, 1−45, and references therein.
(7) Geiger, A.; Mausbach, P.; Schnitker, J.; Blumberger, R. L.; Stanley, H. E. *J. Phys. (Paris)* **1984**, *45*, C7−13.
(8) Stanley, H. E.; Blumberg, R. L.; Geiger, A.; Mausbach, P.; Teixeira, J. *J. Phys.* **1984**, *45*, C7−3.
(9) Speedy, R. J.; Madura, J. D.; Jorgensen, W. L. *J. Phys. Chem*. **1987**, *91*, 909−913.
(10) Sciortino, F.; Fornili, S. L. *J. Chem. Phys*. **1989**, *90*, 2786.
(11) Stillinger, H. *Science* **1980**, *209*, 451.
(12) Sceat, M. G.; Rice, S. A. *J. Chem. Phys*. **1980**, *72*, 3237.
(13) Mezei, M.; Beveridge, D. L. *J. Chem. Phys*. **1981**, *74*, 622.
(14) Rapaport, D. C. *Mol. Phys*. **1983**, *50*, 1151.
(15) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys*. **1986**, *84*, 2184.
(16) Rapaport, D. C.; Scheraga, H. A. *Chem. Phys. Lett.* **1981**, *78*, 491.
(17) Mountain, R. D. *J. Chem. Phys.* **1995**, *103*, 3084.
(18) Haughney, M.; Ferrario, M.; McDonald, I. R. *J. Phys. Chem*. **1987**, *91*, 4934.
(19) Rerrario, M.; Haughney, M.; McDonald, I. R.; Klein, M. L. *J. Chem. Phys.* **1990**, *93*, 5156.
(20) Luzar, A.; Chandler, D. *J. Chem. Phys*., **1993**, *98*, 8160. Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 6560.
(21) Tanaka, H.; Ohmine, I. *J. Chem. Phys*. **1987**, *87*, 6128.
(22) Zichi, D. A.; Rossky, P. J. *J. Chem. Phys.* **1986**, *84*, 2814.
(23) Sciortino, F.; Fornili, S. L. *J. Chem. Phys.* **1989**, *90*, 2786.
(24) DLPOLY, Smith, W.; Forester, T. R.; DRAL Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK.
(25) Kouwijzer, M. L. C. E.; van Eijk, B. P.; Kooijman, H.; Kroon, J. *Acta Crystallography* **1995**, *B51*, 209.
(26) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dodrecht, The Netherlands, 1981; p 331.
(27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(28) Ewald, P. *Ann. Phys.* **1921**, *64*, 263.

(29) Verlet, L. *Phys. Res.* **1967**, *159*, 98.
(30) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.
(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926.
(32) *Gaussian94*, Revision B.3; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head.-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 1995.
(33) Basis set superposition errors should be similar in these two difference calculations.
(34) Northrup, S. C.; Hynes, J. T. *J. Chem. Phys.* **1980**, *73*, 2700.