Computation of the Mean Residence Time of Water in the Hydration Shells of Biomolecules

Angel E. García[1]* and Lewis Stiller[2]

[1]*Theoretical Biology and Biophysics Group, T-10, MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, and* [2]*Department of Computer Sciences, Johns Hopkins University, Baltimore, Maryland 21218*

Received 16 November 1992; accepted 21 June 1993

An algorithm is presented within the context of the calculation of the time-relaxation behavior of the hydration shells around atomic sites in biomolecules. We report a calculation of the time-relaxation behavior of the first and second hydration shells of polar, hydrophobic, and charged groups in a protein, crambin. The water mean residence times around protein groups are obtained from averages over configurations sampled during a 325-ps molecular dynamics simulation of crambin in solution. A convolution arising in the calculation of the mean relaxation time is implemented using a parallel prefix operator. A new characterization is given of the parallel prefix operator as a linear transformation, and this formulation enables us to derive efficient factorization of the convolution as a product of two parallel prefix operations. The parallel prefix operations are implemented in logarithmic time.

INTRODUCTION

There has been a continuing interest in studying the structure[1–5] and dynamics[6] of water molecules around biomolecules. Raman,[7] Brilloiuin,[8,9] nuclear magnetic resonance (NMR),[6,10–11] infrared (IR),[12,13] and inelastic and quasielastic neutron scattering[14,15] techniques give information about the time relaxation and energetics of water molecules in the first and second hydration shells around biomolecules. High-resolution neutron diffraction[2,16,17] and X-ray crystallography[3–5,18] give detailed information about the structure of water surrounding the protein surface. Recently, multidimensional NMR techniques have been used to study the structure and time relaxation of bound water molecules on the surface and interior of proteins.[10,11] The structural details of buried water molecules of BPTI in solution are in complete agreement with X-ray diffraction data of crystals. However, surface water molecules observed in crystals were not detected in solution. This emphasizes the importance of considering the different time resolution of various techniques. Brillouin, Raman, IR, and quasielastic neutron scattering techniques can detect coupling between water molecules and linear excitations of the system (phonons) ranging from a tenth of a ps to over a few hundred ps.[7–9] Multidimensional NMR techniques in solution will detect couplings with relaxation times longer than a few hundred picoseconds,[10,19] while diffraction techniques will detect time average phenomena over long (days) periods of time.[2,4,16,17]

Time-dependent information about structural water can be obtained from measurements that depend on relaxation (i.e., time dependent) mechanisms. We define the relaxation time, $\tau_{R_{s1}}$, of coordinated water molecules as the average time of residence of a water molecule in a hydration shell around a protein site before it is replaced by other water molecules coming from the bulk or other hydration shells. This time constant is related to the water diffusion time and, in some instances, is of the same order of magnitude as the Debye relaxation time for water.[19,20]

The water relaxation time can be measured by multidimensional NMR techniques.[10,11] In 2-D NMR experiments, a pair of spins (e.g., a protein proton and a hydrogen atom from a coordinated water molecule) can interact by means of dipolar coupling if the distance between spins is less than 5 Å.[21] The decay of the excitation will depend, among other things, on the exchange of the coordinated water molecule with solvent water molecules. The ability to measure this coupling will depend on the characteristic tumbling time of the protein in solution, that is, the relaxation phenomena to be studied should relax in a time slower than the characteristic relaxation time of the technique. 2-D NMR experiments on BPTI[10,11] detect resonance between protein hydrogen atoms and *bound internal* water molecules, not with *bound* (i.e., coordinated) surface waters. This suggests that all water molecules on the protein surface have exchanged with bulk water molecules in a time scale faster than the characteristic tumbling time of a protein in solution, approximately 300 ps for BPTI.[10]

*Author to whom all correspondence should be addressed.

CCC 0192-8651/93/111396-11

Protein structures studied by diffraction methods show ordered water molecules around some charged,[2,16,17] polar,[2,16,17] and nonpolar[3,4] (hydrophobic) amino acids, although the latter are only observed when they are at the interface between symmetry-related molecules in the crystal. Water networks surrounding the protein have been observed for small proteins.[3] Crystallographically observed water molecules are referred to as *bound* structural water molecules. During crystallographic data analysis, *bound* water molecules are often treated as though they belong to the protein. Recent developments in the treatment of the *bulk* solvent contribution to the low-order diffraction data allow a better evaluation of the surface structure of the protein and a better localization of *bound* waters.[2,16] The mobility of *bound* waters is studied by means of Debye–Waller factors and occupancy.[2,14,16] The bulk solvent has relatively high disorder (liquid), which is represented by liquidity factors.[16] Water layers surrounding the protein have little mobility. Diffraction methods do not provide any information about the time behavior of structural water molecules. At first sight, diffraction and NMR results seem to be contradictory. We should keep in mind that diffraction experiments measure a long-time-averaged occupancy of surface sites by water molecules at different unit cells but that 2-D NMR experiments can detect only those water molecules in the vicinity of a protein site that reside for a period of time longer than the tumbling time of the molecule in solution. Considering these facts, it is possible that both results are consistent with the structure and dynamics of the solvation shells around proteins.

The structure and relaxation (in the picoseconds time scale) of the solvation shells may be properly described by molecular dynamics (MD) simulations.[19,20] The MD simulation results can be qualitatively compared to diffraction and NMR experimental data.[22–26]

In this work, we study the relaxation of water molecules in first and second hydration layers around a protein atom by means of a survival probability function, $p_{\alpha,j}(t, t + t')$. $p_{\alpha,j}$ is a binary function that adopts a value of *one* if the water molecule labeled j has been in the referred hydration shell around site α from time t' to time $t + t'$, without getting out in between this time interval, and zero otherwise. The survival function for a hydration shell of a protein site consists of the average over time and over all water molecules of $p_{\alpha,j}$:

$$P_\alpha(t) = \sum_{j=1}^{N_{\text{water}}} \sum_{t'} p_{\alpha,j}(t', t' + t) \qquad (1)$$

This function is a time correlation function. $P_\alpha(t = 0)$ equals the average number of water molecules belonging to the hydration shell of site α (i.e., the coordination number), and $P_\alpha(t)$ gives the average number of water molecules that still remain in the hydration shell after a time t from when they first entered the shell. The averages are calculated over configurations sampled during the MD simulation.

Similar functions have been defined by Sciortino et al.[27,28] to study the distribution of bond lifetimes in simulations of liquid water. In particular, Sciortino's definition of $P_\infty(\tau)$ corresponds to our definition of $P_\alpha(t)$ but with a different *yes* or *no* question asked. There, Sciortino and Forlini[27] use a definition for a bond between water molecules, while here we will use the residence of a water molecule in the hydration shell of a protein site to study the lifetime of water molecules in the hydration shells. We have adopted a definition used previously by Impey, Madden, and McDonald[29] to study the residence time of water molecules in the first coordination shell of ions. Impey et al.[29] define a junction $P_j(t, t_n, t^*)$ that takes a value of *one* if molecule j lies within the first coordination shell of an ion at both time steps t_n and $t + t_n$ but that in the iterim does not leave the coordination shell for any continuous period longer than t^*; otherwise, it takes a value *zero*. Our definition is the equivalent of taking $t^* = 0$ (actually, $t^* = \Delta t$, where Δt is the time interval between configurations). The code discussed here could be easily modified to use any value of t^*. The residence time of water molecules in the hydration shells of ions has also been reported by Bopp.[30] In general, a function like $P(t)$ may be used to study the statistics of any two-state system, like the formation of hydrogen bonds[31] in a biomolecule or the transitions of dihedral angles.[32] A relatively long-time MD calculation is necessary to obtain statistically significant time correlation functions to use in describing any hydration shell's relaxation rates that may extend to over 100 ps.

The calculation of the survival probability function described above should not be confused with the calculation of the diffusion coefficient previously reported.[22,24] Previous calculations of the diffusion coefficient of water molecules around hydrophobic, polar, and charged atoms in a peptide depend on the assumption that the mean diffusion length, $l^2 \sim Dt_{\text{run}}$, during the simulation time or the averaging time is less than a water molecular diameter. Therefore, the definition of average positions of a water molecule and the subsequent identification of this water molecule with a site in the peptide is justified. The interpretation of fast relaxation measurements of the diffusion coefficient, determined by the broadening of the NMR resonance signal, requires the assumption of small diffusion lengths during the measurement. To compare molecular dynamics simulation results with the NOESY and ROESY NMR data in solution,[10,11] we must take into account that the resolution time of the signal is determined by the mean tumbling time of the solute, which ranges in the hundreds of picoseconds. For this time scale, the

mean diffusion length will be, on average, much larger than the mean water diameter of the order of the protein length. As a consequence, a water molecule that samples a large volume will occupy positions close to the domains around various atoms, regardless of their electrostatic properties.

In this article, we will show that the survival probability function used to describe the time-relaxation behavior of the hydration shells around biomolecules can be efficiently calculated by using an algorithm that makes use of a parallel prefix operation.[33–35] We give a new characterization of the parallel prefix operator as a linear transformation. This formulation enables us to derive an efficient factorization of the convolution as a product of two parallel prefix operators.

FORMULATION OF THE COMPUTATIONAL PROBLEM

Given a vector of distances between a water molecule, j, and a protein site, α, over the number of configurations, N, obtained from a molecular dynamics simulation (with a configuration saved every $\Delta t = s \times dt$, where dt is the time integration step and s is the number of time integration steps between saved configurations), find a statistical function measuring the time a water molecule remains within a hydration shell around the site, that is, get the number of time intervals of length t such that the water is *continuously* inside the hydration radius for that time interval. First, we must determine the time (binary) sequence, $A_{\alpha,j}$ of length N that answers *yes or no* (i.e., *one* or *zero*) if a water molecule indexed j resides within the specified hydration shell around site α for each interval corresponding to a configuration during the simulation. A simple definition of $P_\alpha(t)$ is given by

$$P_\alpha(t) = \sum_{j=1}^{N_{\text{waters}}} \frac{1}{N-t} \sum_{t_o=1}^{N-t} \prod_{k=t_o}^{t_o+t} A_{\alpha,j}(k) \qquad (2)$$

Notice that this operation is similar to a convolution where the products of a function, $A(k) \times A(t-k)$, are replaced by the product of the function at all intervals encompassing the two points, $A(t-k) \times A(t-k-1) \times \ldots \times A(k)$. This is not the actual algorithm used in the serial computer. We will follow another approach that consists in performing an operation on $A_{\alpha,j}$ to produce a sequence $B_{\alpha,j}$, where the ith element of $B_{\alpha,j}$ is the number of sequences of i contiguous 1s contained in $A_{\alpha,j}$. For example, given $A = (1,1,0,0,1,1,1)$, then $B = (5,3,1,0,0,0,0)$. From the set of $B_{\alpha,j}$, we can calculate $P_\alpha(t)$ in eq. (1) for all water molecules j as

$$P_\alpha(t) = \sum_j \frac{B_{\alpha,j}(m)}{N-m} \qquad (3)$$

where $t = m\delta t$ and $B_{\alpha,j}(m)$ is the mth component of $B_{\alpha,j}$. Next, we describe the parallelization of the implementation of this approach by using a parallel prefix operation.

A parallel prefix operation is the *scan* of a vector, V, of partial sums of V. The ith element of the *scan* is the sum of all elements of V with index greater than or equal to i, that is,

$$\text{scan}(V)_i = \sum_{j=N-i}^{N} V_j \qquad (4)$$

For example, given $V = (1,3,5,7,9)$, then $\text{scan}(V) = (25,24,21,16,9)$. This parallel prefix has been proposed as a primitive for a wide range of parallel computations. In particular, it is supported by the Connection Machine (CM-200 and CM-5) hardware and microcode. To simplify the presentation of our results, the parallelization of the algorithm to calculate $P_\alpha(t)$ using eq. (3) is described in Appendix A. There, we show that the parallel prefix operation can be represented as a linear transformation. The square of this transformation relates $B_{\alpha,j}$ to $A_{\alpha,j}$. Therefore, two consecutive applications of the parallel prefix operation are necessary to obtain $P_\alpha(t)$. The parallel prefix operations can be implemented in logarithmic time (see Appendix B). The implementation of this algorithm for the analysis of the relaxation times of the hydration shells around a biomolecule was done in slicewise CM FORTRAN 1.1. All parallel computations were done on a CM-200 computer at the Advanced Computing Laboratory (ACL) at Los Alamos National Laboratory. The source codes for the parallel and serial codes are available upon request. The functions $P(t)$ were computed for all atomic sites in a protein, crambin.

DESCRIPTION OF THE BIOMOLECULAR SYSTEMS

To study the structure and relaxation of hydration shells around a biomolecule, we have simulated the dynamics and hydration of a small hydrophobic protein, crambin, in solution, by means of a constant-temperature molecular dynamics simulation using AMBER[36] (v. 3.1). Crambin is a 46-amino-acid amphipathic protein for which high-resolution X-ray[18] and neutron diffraction[3] data are available. The water structure of crambin in the crystal environment has been reported by Teeter.[3,4] In addition, 1-D/2-D NMR studies of crambin in both a crystal environment[37,38] and water/ethanol solution[39,40] have also been reported. Crambin has also been the subject of many theoretical calculations. These calculations include normal mode analysis,[41] *in vacuo* MD simulations and quasiharmonic analysis,[42] energy minimization in a crystal–water environment,[43] energy minimization *in vacuo*,[44] and molecular dynam-

ics simulations in solution.[45] The amphipatic character of the protein and the fact that the protein retains its secondary and tertiary structures in a water/ethanol solution[39,40] provide a unique opportunity to study the hydration of polar, charged, and hydrophobic groups in the same simulation.

During the simulation, crambin was immersed in a box of dimensions 42.11 × 36.85 × 29.34 Å containing 1315 TIPS3[46] water molecules. The protein was oriented along the principal moments of inertia axes (after the rotation, the x-axis indicates the direction of the principal moment of inertia). The initial conformation of the protein was obtained from the crystallographic coordinates reported by Hendrickson and Teeter.[18] Hydrogen atoms of charged and polar groups were added to the X-ray heavy-atom structure. The *united atoms* force field of Weiner et al.[47] was used. The resulting system contains 327 heavy atoms, 69 hydrogens, and 12 lone pairs attached to the S_γ atoms of cysteine residues. All bond length distances involving hydrogen atoms were constrained using the SHAKE algorithm.[48] Periodic boundary conditions were used for Coulombic and van der Waals interactions between atoms. A residue-based spherical cutoff distance of 12.5 Å was used to truncate all nonbonding interactions. The large cut-off distance of 12.5 Å (relative to the hydration shells radius of 3.5–5.5 Å) used to truncate all nonbonding interactions in the simulations should yield a better description of the hydration around protein sites up to three hydration shells, that is, the first and second hydration shells will be similar to those obtained in a system in which the long-range interactions are not truncated. A third hydration shell may be affected by indirect correlations with other water molecules not interacting with the parent site. Further details about these calculations will be published elsewhere.

The system temperature was constrained to 300 K using a weak coupling of the system to a heat bath.[49] The system was heated to 300 K at intervals (1, 10, 50, 100, 150, 250, and 300 K) during the first 7 ps of simulation. For the next 17 ps, the velocities of the protein and solvent were rescaled to balance the instantaneous temperature of each of the two components to 300 K. During the 15- to 24-ps time period, the average temperature for the protein, solvent (water), and the total system (solute and solvent) were 277 ± 10, 304 ± 3, and 300 ± 3 K, respectively. The fact that the protein exhibits larger fluctuations (±10 K) than the solvent (±3 K) should not be related to differences in specific heat but rather to a size effect. For example, a subset (408 water molecules) of solvent molecules during the same period showed an average temperature of 303 ± 7 K.

The simulation has been extended for 325 ps and configurations of the system collected at a rate of 20/ps. The integration step was 2 fs. Configurations from the first 24 ps are not included in any averaging.

RESULTS

Description of the Protein Dynamics

The overall secondary and tertiary structures of the protein are preserved during the simulation. The average main-chain (Cα, C, O, N) atoms' rms distance[50] between the initial structure and all structures during 216 ps is 1.18 ± 0.08 Å. The largest contributors to these fluctuations are residues 1 and 33–39. These residues belong to beta sheet regions (residues 1–4 and 32–35) and to coil regions (residues 36–40). The secondary and tertiary structures are preserved during the simulation, although the beta sheets regions are more flexible in the alpha helical regions. The mean square displacements for the C_α atoms are shown in Table I. The mean square displacements of all atoms around their average position during the first 216 ps of averaged configurations is 0.946 Å^2. The motions around the average structure are bounded (i.e., exhibit oscillations, although nonharmonic). We have reported[45] a description of the dynamics of the protein in terms of multibasin nonlinear dynamics.

All side-chain to main-chain and side-chain to side-chain hydrogen bonds present in the crystal structure are also present during the simulation. In addition, a strong hydrogen bond between the charged side chains arginine, R_{17}, and glutamic acid, E_{23}, is formed. This *ion pair* is not formed in the crystal structure. We speculate that the absence of excess ions in our simulation may have enhanced the probability of forming this ion pair. In addition, hydrogen bonds are formed between $N_{\delta 2}$ of N_{46} and O of C_4 (77% of the time) and $N_{\delta 2}$ of N_{46} and O of P_5 (23% of the time). Table II lists some of the hydrogen bonds between side-chain to main-chain atoms and side-chain to side-chain atoms. We indicate the presence of a hydrogen bond as a percentage of the time that a bond is formed (with donor-to-acceptor distance of less than 3.2 Å, and angle formed by the donor-hydrogen-acceptor greater or equal than 120°) to emphasize the dynamic nature of the hydrogen bonds. For example, the donor-to-acceptor distance in the hydrogen bond formed between $N_{\delta 2}$ of N_{46} and O of C_4 varies from 2.8 to 4.8 Å, exhibiting distances near 3.0 Å in most configurations and a time average distance near 3.0 Å. The breaking of this bond (23% of the time) is mostly due to variations in the donor-hydrogen-acceptor angles from 80 to 180°, with frequent jumps out of the 120–180° range. The average lifetime of this hydrogen bond is around 2.0 ps.

Description of Water Mean Residence Time around Crambin Side Chains

We computed the survival time statistics described before for various radii for all 408 protein sites in crambin and for Na^+ and Cl^- ions in water solutions. The mean residence time of water molecules around

Table I. rms fluctuations of Cα atoms of crambin during 216 ps of molecular dynamics simulation.

Residue	rms distance (Å)
T_1	1.79
C_3	0.77
P_5	0.68
I_7	0.56
A_9	0.45
S_{11}	0.45
F_{13}	0.37
V_{15}	0.42
R_{17}	0.41
P_{19}	0.69
T_{21}	0.55
E_{23}	0.55
I_{25}	0.61
A_{27}	0.52
Y_{29}	0.57
G_{31}	0.62
I_{33}	1.07
I_{35}	1.14
G_{37}	1.47
T_{39}	1.17
P_{41}	0.80
D_{43}	1.34
A_{45}	0.66
T_2	1.02
C_4	0.65
S_6	0.58
V_8	0.57
R_{10}	0.40
N_{12}	0.46
N_{14}	0.40
C_{16}	0.39
L_{18}	0.59
G_{20}	0.79
P_{22}	0.58
A_{24}	0.70
C_{26}	0.41
T_{28}	0.61
T_{30}	0.57
C_{32}	0.64
I_{34}	1.14
P_{36}	1.58
A_{38}	1.46
C_{40}	0.75
G_{42}	1.10
Y_{44}	0.99
N_{46}	0.71

Notice that most of the fluctuations occur for residues 1 and 33–39. These residues belong to beta sheet regions (residues 1–4 and 32–35) and coil regions (residues 36–40).

the first hydration shell of atoms (or groups) belonging to the sidechains of residues in crambin range from a few picoseconds for polar and nonpolar groups to over 50 ps for charged side chains. To simplify the interpretation of our results and compare them to similar calculations by others, Table III lists the mean residence time of water and some monovalent ions. $M(R1)$ is the number of coordinated water molecules around the first hydration shell of each ion or molecule. The first hydration shell radii, $R1$, is defined as the separation between the reference atom and a water oxygen at which the radial distribution function, $g_{\alpha,O}(r)$, has its first

Table II. Hydrogen bonds between side- and main-chain atoms.

Donor	Acceptor	X-ray distance (Å)	Simulation % occurrence
Nε (R_{10})	OXT (N_{46})	2.83	59.0
NH1 (R_{10})	O (N_{46})	2.85	97.0
NH2 (R_{10})	O (T_2)	2.99	53.0
Oγ (S_{11})	O (I_{17})	3.28	83.0
Oγ (T_{30})	O (C_{26})	2.74	93.0
Nδ2 (N_{46})	O (C_4)	5.51	77.0
Nδ2 (N_{46})	O (P_5)	6.00	23.0
Nε (R_{17})	Oε2 (E_{23})	3.97	84.0
NH1 (R_{17})	O (T_{21})	4.32	92.0
NH2 (R_{17})	Oε2 (E_{23})	5.16	93.0

The observed hydrogen bonding distance (when present) in the X-ray structure[18] is indicated under the column labeled X-ray distance. The presence of a hydrogen bond in the simulation is indicated by the percentage of the time that the bond is present in the sampled configurations. The criteria for hydrogen bonding used here is that the donor-to-acceptor distance is less than 3.2 Å and the angle is greater than 120°. We indicated the presence of a hydrogen bond as a percentage to emphasize the dynamic nature of the bonds.

minimum. The integration over space, $M(R1) = 4\pi \int_0^{R1} r^2 g_{\alpha,O}(r)dr$, gives the number of coordinated water molecules on the shell. τ_R describes their mean residence time. The value of τ_R is obtained from a single exponential fitting of $P_\alpha(t)$, $M(R1) = P_\alpha(0)$. The columns under $M(R1)$ and τ_R show the results obtained in our calculations using a TIPS3[42] model for water and the ion energy parameters of Jorgensen et al.[51] The columns under τ(Bopp) and τ(Impey) give the mean residence time obtained by Bobb[30] and Impey et al.[29] Bopp used a Tip4p model[46] for water, while Impey et al. used the MCY model[52] for water. The values in the column under τ(exp) are the experimental values for the mean residence time reported by Bopp. Our results for τ_R are smaller than those reported by others. As mentioned before, our

Table III. Mean residence time in the hydration shells of water and simple ions.

Ion	M(R1)	τ_R	τ(Bopp)	τ(Impey)	τ(exp.)
H_2O	5.0	1.4	2.5	1.8	2.5–10.0
Cl^-	8.0	4.0	5.0	4.5	3.0–10.0
Na^+	7.0	20.0	15.0	9.9	5.0–30.0
F^-	—	—	20.0	12.0	6.0–60.0
Li^+	—	—	33.0	85	8.0–40.0

The mean residence times are measured for the first hydration shell of each molecule or ion. M(R1) is the number of water molecules coordinated in the first hydration shell of the groups. The values τ(Bopp) refer to the values reported by Bopp[27]; τ(Impey) refer to values reported by Impey, Madden, and McDonald.[26] The experimental values are taken from Bopp.[27] Our calculations for Cl^- were done at infinite dilution (216 TIPS3 waters and 1 ion). The calculations for Na^+ were done for Na^+ counterions in a 9-bases DNA hairpin + 8 Na^+ + 1678 TIPS3 water molecules (A.E. García, unpublished work). τ_R is the mean life of water molecules in the first hydration shells of the referenced ions. All times are given in ps.

Table IV. Mean residence time in the hydration shells of positively charged residues in crambin.

Residue	3.5 Å		4.5 Å		5.5 Å	
	τ_R	τ_1	τ_R	τ_1	τ_R	τ_1
N H_3–end	2.7	3.4	7.6	22.0	4.2	41.0
R_{10}	0.3	0.0	2.0	4.5	5.7	27.0
R_{17}	0.3	0.1	1.5	3.5	4.0	13.0

The mean residence times are measured for shells of radii 3.5, 4.5, and 5.5 Å. The first hydration shell radii are 4.5 Å for the N H_3–end and 5.5 Å for the arginine residues. These distances are measured from the N and Cζ atoms, respectively. τ_R is the mean relaxation time of water molecules in the first hydration shells of the reference atoms in the charged groups. τ_1 refers to the mean time in which only one water molecule still remains in the hydration shells of the group. All times are given in ps.

definition is: $P(t) = P_{\text{Impey}}(t, t = 0)$. Impey et al.[29] used $t^* = 2.0$ ps. A calculation with $t^* = 0$ will yield smaller values of τ_R except when $\tau_{\text{ion}} \gg \tau_{\text{solvent}}$, where $\tau_{\text{solvent}} = \tau_{\text{water}}$. Such is the case for Na^+. Notice that all calculated values are within experimental estimates. Table III also shows τ_R for Li^+ and F^-, which were not calculated by us.

Tables IV–VIII show the calculated values of τ_R for different layers of water around positive charges (Table IV), negative charges (Table V), polar (Table VI), and nonpolar (Tables VII and VIII) residue side chains in crambin. These values should be compared with the results listed in Table III. Tables III–VIII also show a mean time, τ_1, defined as the mean time after which only one water molecule remains in the hydration layer [i.e., $\tau_1 = \tau_R \log_e P(0)$].

Table IV shows the mean residence time (MRT) for water molecules around positive-charge groups in crambin.[53] The hydration layers are centered around the N atom in the NH_3^+-terminal group, and around Cζ atoms in arginine. The first and second hydration shells of the N-terminal group have approximate radii of 4.5 and 6.0 Å, respectively. The corresponding radii for Cζ in arginine are 5.5 and 7.5 Å. Notice that τ_{R1} for these ions ($\tau_{R1} \sim 4$–5.7 ps) is shorter than the values corresponding to Na^+ ions ($\tau_{R1} \sim 20$ ps). This may be due to the fact that these are molecular ions with a larger distance of close approach for water than simple ions. From Table III, we notice that Li^+ ions, which have a closer approach to water molecules, have a larger MRT than Na^+.

τ_{R1} for R_{17} is smaller than for R_{10}. This is an effect of the local environment on the behavior of the hydration shells of various groups. For example, R_{17} forms an ion pair with E_{23} while R_{10} forms three

Table V. Mean residence time in the hydration shells of negatively charged residues in crambin.

Residue	3.5 Å		4.5 Å		5.5 Å	
	τ_R	τ_1	τ_R	τ_1	τ_R	τ_1
E_{23}	0.4	0.3	6.8	26.0	4.5	29.0
D_{43}	0.4	0.5	8.6	18.0	4.6	27.0
CO_2^--terminal	0.5	0.3	12.0	21.0	5.2	35.0

The mean residence times are measured for shells of radii 3.5, 4.5, and 5.5 Å. The radii of the first hydration shells are ~ 4.5 Å when measured relative to the Cδ of E_{23}, Cγ of D_{43}, and C of the charged C-terminal group. τ_R is the mean relaxation time of water molecules in the first hydration shells of the reference atoms in the charged groups. τ_1 refers to the mean time in which only one water molecule still remains in the hydration shells of the group. All times are given in ps.

Table VI. Mean residence time in the hydration shells of hydroxyl oxygens in threonine, serine, and tyrosine residues in crambin.

Residue	4.0 Å			5.0 Å		
	$M(R1)$	τ_R	τ_1	$M(r_2)$	τ_R	τ_1
T_1	5.1	1.0	1.9	11.0	2.5	8.4
T_2	3.3	1.6	2.3	7.0	4.4	12.0
T_{21}	3.3	1.9	2.4	6.5	3.5	9.9
T_{28}	3.6	1.7	2.7	8.5	2.2	8.5
T_{30}	0.0	0.0	0.0	0.3	0.4	0.0
T_{39}	5.0	1.2	2.3	11	2.1	7.8
S_6	4.3	1.3	2.6	8.8	3.4	10.6
S_{11}	4.2	1.2	2.5	9.6	2.4	8.4
Y_{29}	6.0	0.9	1.8	12.7	1.9	8.4
Y_{44}	5.3	1.1	2.0	10.7	2.1	6.1

The mean residence times are measured for layers of radii 4.0 and 5.0 Å. The hydration shells around these groups are calculated relative to Oγ of serine, Oγ1 of threonine, and Oη of tyrosine. The first hydration shell radii of Oγ and Oγ1 are 4.0 Å for S_6, S_{11}, T_2, and T_{28}, 5.0 Å for T_1, T_{21}, and T_{39}; nonexistent for T_{30}; and 5.0 Å for Oη of Y_{29} and Y_{44}. τ_R is the mean relaxation time of water molecules in the first hydration shells of the reference atom in the groups. τ_1 refers to the mean time in which only one water molecule still remains in the hydration shell of the group. All times are given in ps. $M(R1)$ and $M(r_2)$ are the number of water molecules coordinated in a sphere, with the center in the group's oxygen atom, with radii 4.0 and 5.0 Å, respectively.

Table VII. Mean residence time in the hydration shells of methyl groups in threonine residues in crambin.

Residue	$M(R1)$	τ_R	τ_1
T_1	9.4	2.0	5.8
T_2	3.6	1.7	2.5
T_{21}	5.1	2.5	5.1
T_{28}	12.0	2.2	6.5
T_{30}	1.6	1.4	0.5
T_{39}	12.6	2.0	6.11

The mean residence times are measured for shells of radius 5.0 Å, which corresponds to the radii of the first hydration shell. τ_R is the mean relaxation time of water molecules in the first hydration shells of the reference atom in the groups. τ_1 refers to the mean time in which only one water molecule still remains in the hydration shell of the group. All times are given in ps. $M(R1)$ is the number of water molecules coordinated in the first hydration shell of the groups.

Table VIII. Mean residence time in the hydration shells of methyl groups in isoleusine residues in crambin.

Residue	$M(R1)$	τ_R	τ_1
I_7 Cγ_2	9.1	1.9	5.1
I_7 Cγ_1	12.2	2.0	6.1
I_{25} Cγ_2	7.7	2.1	5.5
I_{25} Cγ_1	10.4	1.8	5.0
I_{33} Cγ_2	10.4	1.9	5.8
I_{33} Cγ_1	6.9	1.4	3.25
I_{34} Cγ_2	3.9	2.4	3.8
I_{34} Cγ_1	1.9	1.0	0.6
I_{35} Cγ_2	5.8	1.8	3.8
I_{35} Cγ_1	3.1	1.2	1.4

The mean residence times are measured for shells of radius 5.0 Å. $R1$ corresponds to the radius of the first hydration shell. τ_R is the mean relaxation time of water molecules in the first hydration shells of the reference atoms in the groups. τ_1 refers to the mean time in which only one water molecule still remains in the hydration shell of the group. All times are given in ps. $M(R1)$ is the number of water molecules coordinated in the first hydration shell of the groups.

hydrogen bonds to the two O atoms of the charged C-terminal and the carbonil oxygen of T_2.

The corresponding MRT for water molecules around negative-charge groups are listed in Table V. The hydration shells for the groups are taken relative to the Cδ atoms of E_{23}, Cγ of D_{43}, and the C atom of the charged C-terminal. The radii of the first hydration shells are 4.0 Å for Cδ and Cγ atoms and 4.5 Å for C at the C-terminal. Notice that τ_{R1} ($\tau_{R1} \sim 6.8$–12.0 ps) are larger than the corresponding values for Cl^- ($\tau_{R1} = 4.0$ ps), especially for the C-terminal ion. This is an effect of the conformation of CO_2^-, which forms hydrogen bonds with C_4 and R_{10}. As mentioned before, E_{23} forms an ionic pair with R_{17}.

Table VI shows the MRT of hydroxyl groups in threonine, serine, and tyrosine residues of crambin. The hydration layers around these groups are calculated relative to Oγ of serine, Oγ1 of threonine, and Oη of tyrosine. The first hydration shell radii of Oγ and Oγ1 are 4.0 Å for S_6, S_{11}, T_2, and T_{28}; 5.0 Å for T_1, T_{21}, and T_{39}; nonexistent for T_{30}; and 5.0 Å for Oη of Y_{29} and Y_{44}. All the τ_R for the hydroxyl groups are similar to the corresponding values for bulk water.

Tables VII and VIII show τ_R around Cγ methyl groups in threonine and isoleucine residues. The first hydration shells of these groups have a radius of 5.0 Å. Notice that the coordination numbers for these groups vary drastically from 1.6 water molecules to 12.6. However, the values for τ_R are similar to the values for bulk water and for the hydroxyl groups (i.e., $\tau_{R1} \sim 1$–2.4 ps for Cγ atoms and 1.4 ps for water). This suggests that the time-dependent behavior and the stability of the hydration shells of nonpolar groups are determined by water–water interactions.

The results for the MRT presented above have been obtained by means of a single exponential fitting of the survival functions, $P_\alpha(t)$. Figure 1 shows a log–log plot of $P(t)$ for the first ($R1 = 5.5$ Å), and second ($R_2 = 7.5$ Å) hydration shells around the Cζ of R_{10}. First, notice that these curves are not monotonic. Second, notice that the long-time behavior decays slower than an exponential. This is indicative that these curves do not follow a single exponential decay. These curves can be fitted, among other things, to a stretch exponential function, $P(t) = A \exp(-(t/\tau)^\beta)$, where A, β, and τ are the fitting parameters; or to a multiple exponential function, $P(t) = \Sigma_i A_i \exp(-t/\tau_i)$, where two parameters, A_i and τ_i, are introduced with every new exponential.

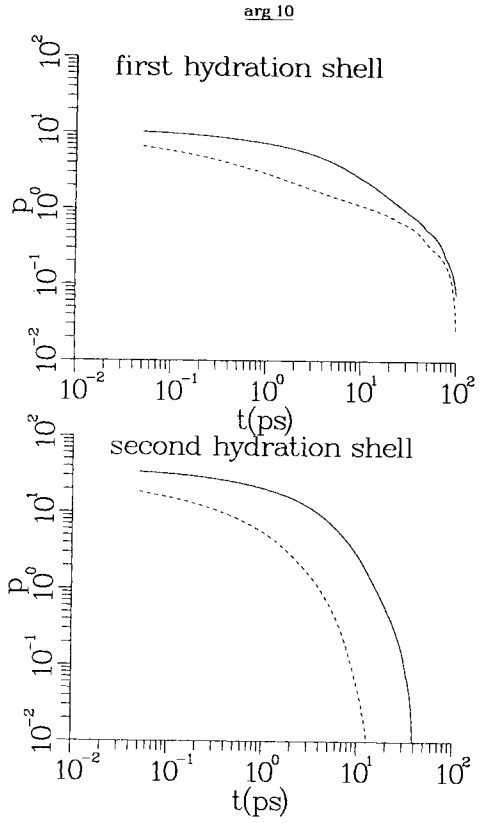

Figure 1. Log–log plot of the water residence time (or survival function), $P(t)$, for the first and second hydration shells around the Cζ atom of the arginine 10 residue ($R10$) in crambin, as a function of time. Time is given in ps. The radii of the *first* and *second* hydration shells are 5.5 and 7.5 Å, respectively (A. García, unpublished results). The solid lines show the function $P(t)$ for the occupancy of water molecules in the first and second hydration shells. The broken lines show a similar function, but with two simultaneous conditions: (1) The water molecules occupy the volume defining the hydration shell; and (2) the energy of interaction of such a water molecule with the charged $R10$ side chain is negative (i.e., attractive). Notice that these curves are nonmonotonic (i.e., there are changes in the sign of the second derivative of the curves). In addition, these curves show a long time decay, slower than an exponential curve. Table IX shows the coefficients obtained by fitting the water residence time curves to multiple exponential curves.

The stretch exponential will account for the long-time behavior of $P(t)$, but it will not account for the nonmonotonic behavior. A multiple exponential fitting will account for both properties at the expense of a larger number of parameters. The interpretation of a multiple exponential fitting is easier, although it may lead to overinterpretation. Table IX shows the coefficients obtained from a multiple exponential fitting of the curves in Figure 1. The number of exponential terms in the fitting was sequentially increased, one at a time, until the average residual of the fitting was less than 1×10^{-3} parts per point. If the reader accepts that this fitting has a physical interpretation in terms of groups of water molecules exhibiting different time-relaxation behavior, we offer the following description. First, notice that of the 10.5 ($\Sigma_i A_i = 10.5$) water molecules around the first hydration shell of Cζ 4.67 show mean residence times ($\tau_{R1} \sim 0.11$–2.7 ps) below the single exponential time ($\tau_{R1} \sim 4.0$–5.7 ps) shown in Table IV. Four water molecules show a similar mean residence time ($\tau_{R1} = 7.3$ ps) as the single exponential fitting. Approximately two water molecules show a mean residence time larger than 40.0 ps.

Of interest is the residence time of water molecules in layers of different widths around the whole biomolecule's surface,[54] as shown in Figure 2. The hydration layers' time of residence curves show that a layer of water thicker than 3.5 Å, as measured from the protein surface, contains at least 10 water molecules that do not exchange with bulk water during a period of time as long as 100 ps. Some of these water molecules reside next to the R_{10} residue, which shows two *bound* water molecules within a 100-ps time scale. Table X shows a multiple exponential fitting to the curves for layers of water within 3.25 and 4.0 Å. The curve for the 4.00-Å layer shows that at least 11 water molecules do not exchange with waters in the remainder of the system for the duration of the simulation. Given that the residence times of water around each individual site for $R1 \leq 5.5$ Å is not any longer than ~100 ps, we can conclude that some water molecules diffuse on the surface of the protein.

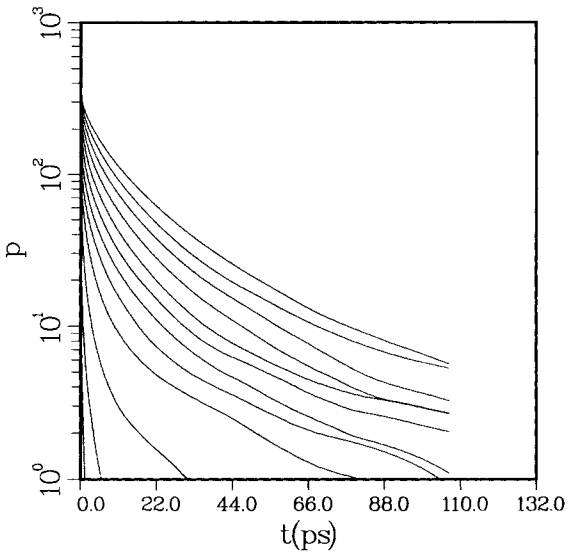

Figure 2. Semilog plot of the residence time of water molecules in layers of different width around the whole surface of crambin, as a function of time. A layer of radius R is defined as the volume including all points whose minimum distance to *any* protein atom is less than or equal to R. Curves for layers of radii $R = 2.25$ Å to $R = 5.25$ Å (at $\Delta R = 0.25$ Å intervals) are shown. The curve with the lowest value at the origin, $P(t = 0)$, corresponds to the layer of smaller radius, $r = 2.25$ Å, and each curve with the next largest value of $P(0)$ corresponds to the layer with the next larger radius. Similar to the curves in Figure 1, these curves are nonmonotonic and cannot be accurately described by a single exponential decay curve. (A single exponential decay curve is described as a straight line in a semilog plot shown here.) Notice that all water molecules in layers with $R \geq 3.0$ Å are exchanged by other water molecules in times less than 50.0 ps. Water molecules in layers with $R \geq 3.25$ Å are exchanged slower, and at least 10 water molecules do not exchange with other molecules in times shorter than half the simulation runtime. Multiple exponential fits to these curves are shown in Table X.

Table IX. Multiple-exponents fitting of $P(t)$ for arginine 10 in crambin.

	First hydration shell		Second hydration shell	
i	A_i	$\tau_i^{(1)}$	B_i	$\tau_i^{(2)}$
1	0.67	0.11	4.9	0.19
2	1.6	0.43	6.1	8.71
3	2.4	2.7	19.0	3.6
4	4.1	7.3	5.4	8.3
5	1.8	40.5	—	—

The curves for $P(t)$ are best fitted by a sum of exponentials than by a single exponential. The multiple-exponential fitting is $P_{R10}^{(1)} = \Sigma_{i=1}^{5} A_i \exp(-t/\tau_i^{(1)})$ and $P_{R10}^{(2)} = \Sigma_{i=1}^{4} B_i \exp(-t/\tau_i^{(2)})$. The residual of the fitting is less than 10^{-3} units per point.

CONCLUSIONS

In general, the residence times of water molecules around side chains range from a few picoseconds for polar groups in threonine, serine, and asparagine to over 50 ps for charged groups. The mean residence time of water around polar groups is longer than the mean residence time around water in bulk water. On the other hand, the mean residence time of water around positively charged groups are shorter (by a factor of four) than for isolated simple ions like Na^+ in water. The mean residence time of water around negatively charged groups is larger (by a factor of two) but of the same order of magnitude as those

Table X. Multiple-exponential fitting of the mean residence time of water molecules in hydration layers around crambin.

	3.25 Å		3.5 Å		3.75 Å		4.00 Å	
Exponent	A_i	τ_i	A_i	τ_i	A_i	τ_i	A_i	τ_i
1	43	0.1	54	0.17	50	0.24	34	0.23
2	20	0.45	29	0.55	47	0.74	63	0.84
3	38	1.5	39	2.1	49	2.8	55	10.9
4	25	5.1	36	7.1	40	9.1	53	31.4
5	8	35.9	8	33.3	13	35.0	11	$>t_{run}$
M_{layer} (R)	134		166		199		216	

The curves of $P(t)$ are best fitted by a sum of exponentials than by a single exponential. The fitted curves correspond to layers around crambin of various widths. The layers are defined as the volume contained within the surface of the protein and a point in space in which all water molecules are closer to a protein atom than the layer width. Notice that for a layer of 4.0 Å some molecules do not leave during the averaging portion of the simulation. M_{layer} (R) is the average number of water molecules inside a layer thickness R, measured from the protein surface.

of isolated chlorine ions in water. We can generalize our results by ranking the MRT in the order $\tau_{charged} > \tau_{polar} > \tau_{nonpolar} \sim \tau_{bulk\ water}$.

The calculation of the residence times around all sites of a protein has been computed thanks to an efficient algorithm that makes use of parallel prefix operators. A new characterization of the parallel prefix operator as a linear transformation has enabled us to derive efficient factorization of the time convolution used to calculate the survival function, $P(t)$, as a product of two parallel prefix operations. The parallel prefix operations are implemented in the Connection Machine, CM-200, in logarithmic time. Without this algorithm and a model of a connection machine, it would have been ineffective to calculate the water survival functions around each protein site or for a multitude of layers of different radii. The computation of the sequences $B_{\alpha,j}$ for a single protein site, α, and all waters, required about 7 h (CPU) on a serial computer (Convex C220). Two hours were required for reading the configurational data and generating the $A_{\alpha,j}$ arrays. The CM-200 compute time, using one quarter of the machine processors (16-K processors), is around 35 s plus about 20 s to output the results. This is not an accurate picture of the bandwidth differential between the two machines because the handling of 400 Mb of configuration data sets forced the serial machine to be I/O bound, while the CM-200 can handle up to 8 Gb of core memory.

The parallel algorithm presented here opens a new window of opportunity to analyze large simulation data sets in highly parallel computers. The evaluation of time correlation functions from configurational data involves a highly repetitive sequence of computations. These kinds of computations are ideal for single-instruction-multiple-processor machines, like the CM-200.

APPENDIX A: CHARACTERIZATION OF THE PARALLEL PREFIX AS A LINEAR TRANSFORMATION

The parallel prefix operation described above can be represented as a matrix multiplication operation. Let I be the identity matrix and U be an upper triangular matrix

$$U_{ij} = \begin{cases} 1 & \text{if } i \le j \\ 0 & \text{otherwise} \end{cases}$$

View V as a column vector. Then, scan$(V) = UV$. In the example given in the text, eq. (4),

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 25 \\ 24 \\ 21 \\ 16 \\ 9 \end{pmatrix}$$

By using the matrix representation of the parallel prefix operation, we will show that that operation in eq. (2), which gives $B_{\alpha,j}$ from $A_{\alpha,j}$, can be represented as two parallel prefix operations. Given the binary sequence $A_{\alpha,j}$, previously defined, we form the auxiliary sequence, E, such that $E(i)$ is the number of contiguous substrings of 1s in A that are bounded by 0s or the ends of the string. For example, if A = (1,1,0,1,1,0,1,1,1,0,0,1,1,1,1) then E = (0,2,1,1,0,0,0,0,0,0,0,0,0,0,0). Notice that a substring of contiguous 1s of length i and bounded by 0s or by the end of the string contains i substrings (not necessarily bounded by 0s) of length 1, $i - 1$ substrings of length 2, etc. Then, the following relation, pointed out by Burton Wendroff, holds that

$$B(k) = \sum_{i=k}^{N} (i - k + 1)E(i)$$

Thus, B is the matrix–vector product ME where

$$M_{ji} = \begin{cases} i - j + 1 & \text{if } j \le i \\ 0 & \text{otherwise} \end{cases}$$

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & . & . \\ 0 & 1 & 2 & 3 & 4 & 5 & . & . \\ 0 & 0 & 1 & 2 & 3 & 4 & . & . \\ 0 & 0 & 0 & 1 & 2 & 3 & . & . \\ 0 & 0 & 0 & 0 & 1 & 2 & . & . \\ 0 & 0 & 0 & 0 & 0 & 1 & . & . \end{pmatrix}$$

But, $M = U^2$ and can therefore be computed with two logarithmic time *scans*. In a serial machine, this

evaluation should take $O(N^2)$ operations. The relationship between $P_a(t)$ and the *scan* function is the basis for the parallelization of the calculation of $p_a(t)$.

APPENDIX B: COMPLEXITY OF MULTIPLICATION BY A FIXED MATRIX

Let M be an $N \times N$ matrix. We would like a measure of the complexity of matrix–vector multiplication, MV, for arbitrary V. Call this complexity $\rho(M)$. Here, we assume an N processor machine and elements of V distributed one per processor.

If M is the identity matrix I, then $MV = V$ and can be computed in constant time. Thus, $\rho(I) = 1$.

If a is a scalar, then we have $\rho(aM) \leq \rho(M) + 1$. Similarly, $\rho(M + M') \leq \rho(M) + \rho(M') + 1$.

Interestingly, $\rho(MM') \leq \rho(M) + \rho(M')$.

Each basis element is viewed as directly communicating with some set of other basis elements. This gives a set of permutation matrices with unit ρ.

Let S be the shift matrix with no wraparound, $\delta_{i,j-1}$.

For example, for $N = 4$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 0 \end{pmatrix}$$

Because of the hypercube connections, $\rho(S^{2^i}) = 1$. (Actually, due to certain overheads on the CM, it would be fair to call this value some small constant $k \geq 1$.)

Let $W_i = I + S^{2^i}$; then, $\rho(W_i) = 3$. Note that $U = I + S + S^2 + \cdots + S^N$.

Suppose $N = 2^m$. Then, by simple algebra, because

$$1 + x^2 + x^3 + \ldots + x^{(n-1)} =$$

$$(1 + x)(1 + x^2)(1 + x^4) \ldots (1 + x^{(n/2)})$$

$$U = W_1 W_2 \ldots W_m$$

Hence, $\rho(U)$ is $O(\log N)$. This places the demonstration that scans are logarithmic in the matrix algebra framework.

Although we did not use this fact in our code, all the standard logarithmic operators on the CM can be written as a matrix–vector product. These include reductions (CM) FORTRAN instincts (sum), spreading, and discrete Fourier transform. It is well known that the FFT may be derived by similar matrix factorization techniques as above (L. Stiller, unpublished results).

This work was supported by the U.S. Department of Energy. The authors thank the Advanced Computer Laboratory at Los Alamos National Laboratory for allowing the use of their facilities. They also thank Burton Wendroff for suggestions regarding the algorithms presented here and C.A. Angell, J. Berendzen, and B. Goldstein for helpful discussions.

References

1. G. Nemethy and H.A. Scheraga, *J. Phys. Chem.*, **66,** 1773 (1962).
2. B.P. Schoenborn, *J. Mol. Biol.*, **201,** 741 (1988).
3. M.M. Teeter, *Proc. Natl. Acad. Sci. USA*, **81,** 6014 (1984).
4. M.M. Teeter, *Annu. Rev. Biophys. Biophys. Chem.*, **20,** 577 (1991).
5. H. Berman, *Curr. Opin. Struct. Biol.*, **1,** 423 (1991).
6. S.D. Kennedy and R.G. Bryant, *Biopolymers*, **29,** 1801 (1990).
7. Y. Tomonaga, M. Shida, K. Kubota, H. Urabe, Y. Nishimura, and M. Tsuboi, *J. Chem. Phys.*, **83,** 5972 (1985).
8. N.J. Tao, PhD thesis, Arizona State University, Tempe, AZ, 1988.
9. N.J. Tao and S.M. Lindsay, *Biopolymers*, **26,** 171 (1987).
10. G. Otting, E. Liepinsh, and K. Wüthrich, *Science*, **254,** 974 (1991).
11. G. Otting and K. Wüthrich, *J. Am. Chem. Soc.*, **111,** 1871 (1989).
12. R.H. Austin, M.W. Robertson, and P. Mansky, *Phys. Rev. Lett.*, **62,** 1912 (1989).
13. S.M. Lindsay, J.W. Powell, and A. Rupprecht, *Phys. Rev. Lett.*, **53,** 1853 (1984); for a review, see S.M. Lindsay, In *Structure and Dynamics of Nucleic Acids, Proteins, and Membranes*, E. Clementi and S. Chin, Eds., Plenum, New York, 1987, p. 239.
14. H. Grimm, H. Stiller, C.F. Majkrzak, A. Rupprecht, and U. Dahlborg, *Phys. Rev. Lett.*, **59,** 1780 (1987).
15. W. Dorster, S. Cusack, and W. Petry, *Nature*, **337,** 754 (1989).
16. X. Cheng and B.P. Schoenborn, *Acta Cryst.* **B46,** 195 (1990).
17. X. Cheng and B.P. Schoenborn, *J. Mol. Biol.*, **220,** 381 (1991).
18. W.A. Hendrickson and M.M. Teeter, *Nature*, **290,** 107 (1981).
19. J.D. Simon, *Acc. Chem. Res.*, **21,** 128 (1988).
20. M. Maroncelli, J. MacInnis, and G.R. Fleming, *Science*, **243,** 1674 (1989).
21. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John Wiley, New York, 1986.
22. P.J. Rossky and M. Karplus, *J. Am. Chem. Soc.*, **101,** 1913 (1979).
23. C.L. Brooks III, M. Karplus, and B. Montgomery-Pettit, *Adv. Chem. Phys.*, **71,** 137 (1988).
24. C.L. Brooks III and M. Karplus, *J. Mol. Biol.*, **208,** 159 (1989).
25. M. Levitt and R. Sharon, *Proc. Natl. Acad. Sci USA*, **85,** 7557 (1988).
26. M. Levitt, *Chem. Scripta*, **29A,** 197 (1989).
27. F. Sciortino and S. Fornili, *J. Chem. Phys.*, **90,** 2786 (1989).
28. F. Sciortino, P.H. Poole, H.E. Stanley, and S. Havlin, *Phys. Rev. Lett.*, **64,** 1686 (1990).
29. R.W. Impey, P.A. Madden, and I.R. McDonald, *J. Chem. Phys.*, **87,** 5071 (1983).
30. P. Bopp, In *The Physics and Chemistry of Ionic Solutions*, M.C. Bellissent-Funnel and G.W. Neilson, Eds., D. Reidel Publishing, Amsterdam, 1987, p. 217.
31. A.E. García, *J. Biomol. Struct. Dyn.*, **8,** A59 (1991).
32. E. Helfand, *Physica*, **118A,** 123 (1983).
33. R.E. Ladner and M.J. Fischer, *J. ACM*, **27,** 831 (1990).
34. G. Blelloch, *Vector Models for Data-Parallel Computing*, MIT Press, Cambridge, MA, 1990.

35. G.S. Almasi and A. Gottlieb, *Highly Parallel Computing*, Benjamin/Cummings, New York, 1989.
36. P.K. Weiner and P.A. Kollman, *J. Comp. Chem.*, **2,** 287 (1981).
37. M. Llinás, A. de Marco, and J.T.J. Lecomte, *Biochemistry*, **19,** 1140 (1980).
38. M.G. Usha and R.J. Wittebort, *J. Mol. Biol.*, **208,** 669 (1989).
39. J.A.W.H. Vermeulen, R.M.J.N. Lamerichs, L.J. Berliner, A. de Marco, M. Llinás, R. Boelens, J. Alleman, and R. Kaptein, *FEBS Lett.*, **219,** 426 (1987).
40. R.M.J.N. Lamericks, L.J. Berliner, R. Boelens, A. de Marco, M. Llinás, and R. Kaptein, *Eur. J. Biochem.*, **171,** 307 (1988).
41. M. Levitt, C. Sander, and P.S. Stern, *J. Mol. Biol.*, **181,** 423 (1985).
42. M.M. Teeter and D.A. Case, *J. Phys. Chem.*, **94,** 8091 (1990).
43. W.L. Jorgensen and J. Tirado-Riva, *J. Am. Chem. Soc.*, **110,** 1657 (1988).
44. M. Whitlow and M.M. Teeter, *J. Am. Chem. Soc.*, **108,** 7163 (1986).
45. A.E. García, *Phys. Rev. Lett.*, **68,** 2696 (1992).
46. W.L. Jorgensen, *J. Am. Chem. Soc.*, **103,** 339 (1981); see also W.L. Jorgensen, J. Chandrasekher, and J.D. Madura, *J. Chem. Phys.*, **79,** 926 (1983).
47. S.J. Weiner, P.A. Kollman, D.T. Nguyen, and D.A. Case, *J. Comp. Chem.*, **7,** 230 (1986).
48. W.F. van Gunsteren and H.J.C. Berendsen, *Mol. Phys.*, **34,** 1311 (1977).
49. H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, and J.R. Kaak, *J. Chem. Phys.*, **81,** 3654 (1984).
50. J. McLachalan, *J. Mol. Biol.*, **128,** 49 (1979).
51. J. Chandrasekhar, D.C. Spellmeyer, and W.L. Jorgensen, *J. Am. Chem. Soc.*, **106,** 903 (1984).
52. O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.*, **64,** 1351 (1976).
53. For a definition of single-letter amino-acids codes and atoms' names, see G.E. Schulz and R.H. Schirmer, *Principles of Protein Structure*, Springer-Verlag, New York, 1978.
54. W. van Gunsteren, H.J.C. Berendsen, R.G. Geurtsen, and H.R.J. Zwinderman, *Ann. NY Acad. Sci.*, **482,** 297 (1986).