Insights into Protein Compressibility from Molecular Dynamics Simulations

Voichita M. Dadarlat and Carol Beth Post*

Department of Medicinal Chemistry and Molecular Pharmacology, 1333 RHPH Bldg., Purdue University, West Lafayette, Indiana 47907-1333

Received: July 6, 2000; In Final Form: October 9, 2000

Isothermal compressibility based on molecular dynamics simulations in a normal temperature and pressure (NTP)−Gibbs ensemble is estimated for five solvated globular proteins (bovine pancreatic trypsin inhibitor, trypsin, ribonuclease A, HEW lysozyme, and α-lactalbumin), as well as bulk water, using the TIP3P model. Protein intrinsic isothermal compressibilities were calculated from molecular total volume fluctuations and averages using the statistical definition of compressibility. A new and efficient method was developed for calculating protein total molecular volume based on an atomic van der Waals radius extension algorithm. The calculated isothermal compressibilities are in good agreement with experimental data (the correlation coefficient is 0.94). The main source of volume fluctuation is the free volume inside the protein, whereas variations in overlap of atomic van der Waals volume are less of a factor. Proteins with low packing density tend to have high compressibility, but packing density alone cannot explain the differences in the compressibility among globular proteins. A simple approach to assess the contribution to solution compressibility from hydration waters suggests small differences between hydration and bulk water compressibility. Estimated bulk water compressibility is in excellent agreement with experimental data. Two criteria for overcoming finite-size effects in bulk water molecular dynamics simulation are a simulation time longer than 300 ps and a system size larger than 260 water molecules.

1. Introduction

Protein molecular volume in solution is a fluctuating quantity due to the internal, thermal motion of secondary structure elements, loops, and side chain orientations, and the interaction of the surface side chains with the solvent. The magnitude of the volume fluctuations as well as the equilibrium value of the molecular volume determine the isothermal compressibility and the effect of pressure on protein structural stability.[1,2] Such dynamical properties of a protein system contain direct information[3] on the nature of forces that govern the structure and dynamics of the protein molecule, whereas knowledge of the mean properties of the system—Gibbs free energy, enthalpy, and entropy—is useful for comparing two states of the molecule (e.g., native and denatured).

Isothermal compressibility, β_T, is defined as the relative change in volume with respect to a change in pressure, when the system is kept at constant temperature. Experimental data show an empirical relationship between β_T in the native state of globular proteins and the change in protein entropy on unfolding, ΔS_{unf}: the higher the compressibility, the smaller the entropy change between native and unfolded states.[4] Furthermore, as a reflection of atomic fluctuations, experimental compressibilities are of interest in regard to understanding hydrogen exchange,[5] protein denaturation,[6] and changes in enzyme activity and modified protein stability by single amino acid mutations.[7] An increase in the thermal stability of a virus on binding antiviral compounds has also been associated with a calculated change in the compressibility of the complexed virus and entropic stabilization.[4,8] Recent experiments confirmed the entropic basis for viral stabilization.[9]

Thermodynamic quantities such as free energy, enthalpy, and entropy change, as well as compressibility, can be evaluated experimentally, but specific contributions to these quantities are difficult to separate. The experimental determination of the protein compressibility accounts for protein intrinsic compressibility, as well as a hydration contribution due to protein−water interaction. Both the hydrational contribution to the compressibility and the intrinsic compressibility are thought to be less than the compressibility of bulk water.[10] The main difficulty in determining the intrinsic isothermal compressibility from the experimentally determined apparent molar adiabatic compressibility is the need to evaluate specific contributions to this quantity, such as the hydrational compressibility due to the interaction of surface atomic groups with the solvent. The possibility that the hydrational contribution to the compressibility is proportional to the specific solvent accessible surface area (SAS) has been explored by Kharakoz and Sarvazyan.[11] They concluded that, for globular proteins, the larger the specific SAS, the larger the hydrational contribution and the smaller the intrinsic compressibility.

Crystallography, fluorescence spectroscopy, nuclear magnetic resonance spectroscopy (NMR), hole burning, and ultrasound velocity measurement are the experimental techniques used to determine protein compressibility. Isothermal compressibilities calculated from the measured partial adiabatic compressibilities of a number of globular proteins[6,10,11] range from 5×10^{-6} to 15×10^{-6} atm^{-1}. By comparison, bulk water has a compressibility of 45×10^{-6} atm^{-1}, organic liquids have compressibilities between 45 and 174 $\times 10^{-6}$ atm^{-1}, and organic crystal compressibilities range from 14 to 35 $\times 10^{-6}$ atm^{-1}. Globular proteins have higher apparent compressibilities than nonglobular biological molecules perhaps because of the larger intrinsic compressibility and smaller specific solvent accessible surface.

* Corresponding author (telephone: 765-494-5980; Fax: 765-496-1189; e-mail: voichi@cc.purdue.edu, cbp@cc.purdue.edu).

10.1021/jp0024118 CCC: $20.00
Published on Web 12/30/2000

The protein intrinsic compressibility, β_T^{int}, can be calculated from molecular dynamics (MD) simulation by using the statistical definition of compressibility and the evolution in time of the protein total volume. Earlier evaluations of the compressibility from MD simulations took a finite difference approach that involved estimating the protein volume change on compression at high pressure.[12,13] In this report, we describe a numerical approach to estimate the isothermal compressibility for globular proteins from protein molecular volume fluctuations and averages obtained from a single trajectory. We introduce an efficient method to calculate the protein total volume so that evaluation of each coordinate snapshot in the simulation trajectory is practical.

The empirical correlation between compressibility and the change in entropy on unfolding suggests that the variation in β_T^{int} is useful in understanding relative stability of globular proteins. To this end, calculation of compressibility using MD simulations is valuable by allowing simultaneous evaluation of factors such as packing density. In addition, simulations provide insight into the nature of volume fluctuations arising from alterations in van der Waals volume or free volume within the protein molecular volume.[14]

2. Computational and Simulation Methods

Isothermal compressibility is a measure of the relative volume change with the change in the pressure of the system kept at constant temperature:

$$\beta_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \tag{1}$$

Statistically, the isothermal compressibility of a system of N particles in equilibrium at constant temperature and pressure, in a normal temperature and pressure (NTP)−Gibbs ensemble is directly related to the volume fluctuation around its average value:[1]

$$\beta_T = \frac{1}{k_B T}\frac{\langle \Delta V^2 \rangle_{NTP}}{\langle V \rangle_{NTP}} \tag{2}$$

where k_B is Boltzmann's constant, T is the temperature of the system, $\langle \Delta V^2 \rangle_{NTP}$ is the average volume fluctuation and $\langle V \rangle_{NTP}$ is the average volume. In a system at constant V and T, isothermal compressibility is calculated from particle number averages and fluctuations:

$$\beta_T = \frac{V}{k_B T}\frac{\langle \Delta N^2 \rangle_{VT}}{\langle N \rangle^2_{VT}} \tag{3}$$

An equivalent formula is obtained when the fluctuation and average of the particle number density, $\rho = N/V$, are used.

Molecular Dynamics Simulations in an NTP Ensemble. MD simulations of proteins in aqueous solution were calculated at constant P and T. The crystallographic protein structures were subjected to in vacuo energy minimization and solvated by repeated overlays of an equilibrated cubic volume of water molecules to fill a truncated octahedron box (Figure 1). The total system volume provided a minimum of four water molecules between the protein surface and the octahedron edge. A 10 ps trajectory allowed for the rearrangement of the water molecules around the fixed protein atoms. The systems were then equilibrated with no constraints for another 100 ps before starting the production run. The number of TIP3P[15] water molecules (4000 to 6000) and the size of the simulation boxes (box axes between 55 and 65 Å) varied with the protein size (58−223 residues) and shape. Periodic boundary conditions were imposed using the CRYSTAL facility in CHARMM.[16] Constant temperature ($T = 300$ K) and pressure ($P = 1$ atm) conditions were applied using the Nose−Hoover method of coupling to a heat bath[17,18] and extended system algorithms for controlling the pressure of the system[19,20] implemented in CHARMM.[21,22] In the constant pressure method, the volume of the system is a dynamic variable whose magnitude is controlled by a generalized force that is proportional to the difference between the internal pressure of the system and the external, fixed pressure of 1 atm. The effect of this constant pressure control is a dynamic change in the volume of the system and a spacial scaling of the position of each atom in the system. Covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm[23] to allow for a time step of 2 fs. A nonbonded cutoff of 12 Å and shifted forces[16] were used in the calculation of Lennard−Jones potentials. The nonbond pair lists were updated every 10 steps. The electrostatic forces and energies were computed using the particle mesh Ewald (PME) method[24,25] with a charge grid spacing of 0.7 Å and direct sum tolerance of 4×10^{-6} for interpolation. Structures for analysis were saved every 0.1 ps. The simulations were carried out on a 8-node parallel machine IBM/SP2 and required 4.5−6 h of CPU time for each 10 picosecond of MD run.

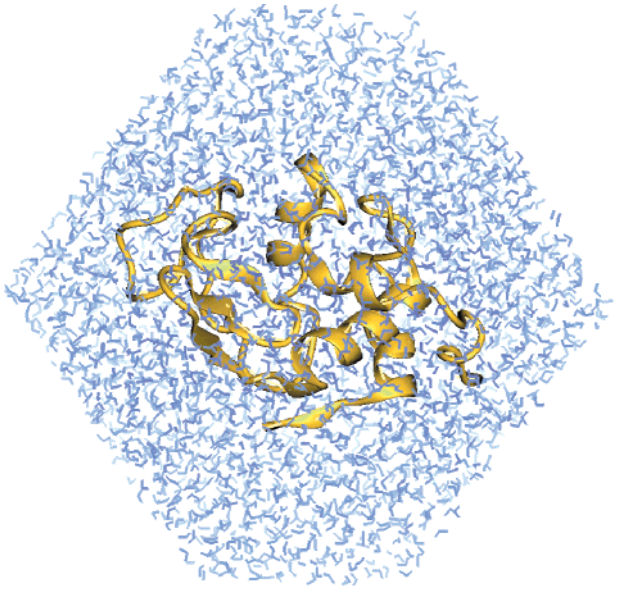

Figure 1. Solvated α-lactalbumin in a truncated octahedron simulation box filled with TIP3P model water molecules.

Protein Molecular Volume Calculation. We describe a new and efficient method for protein total volume calculation based on a grid point analysis. A regular grid large enough to include the full protein is generated in a cubic lattice with a spacing of 0.2 Å. The total molecular volume of the protein comprises the sum of volumes within van der Waals radii of the atomic centers, V_{vdW}, plus the interstitial volume, V_{free}. Thus, V_{vdW} is a sum of all the grid points within the atomic van der Waals radii multiplied by the volume of a cubic pixel. The atomic van der Waals radii used in this study are those in the CHARMM22 parameter set. Figure 2A is a schematic representation of a slab through a protein, where the gray disks represent the van der Waals volume. The total protein molecular volume, V_{prot}, is a sum of the van der Waals and unoccupied interstitial volume within the molecular boundary, $V_{prot} = V_{vdW} + V_{free}$. V_{free} includes internal cavities and packing defects. Figure 2C shows

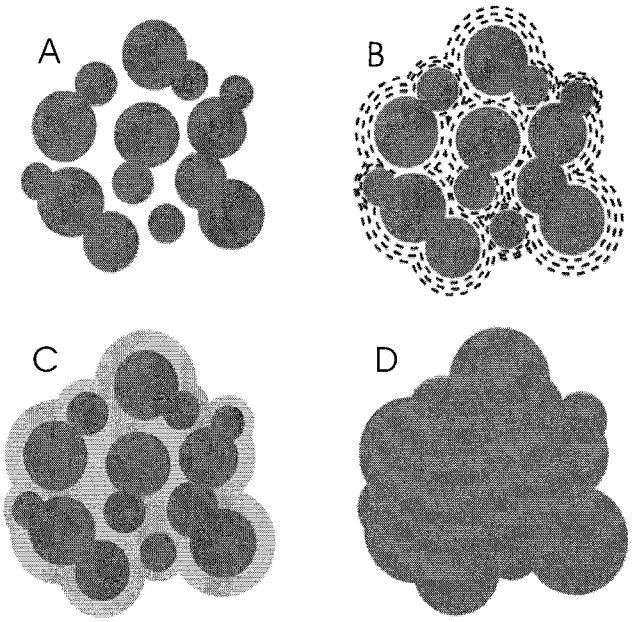

Figure 2. A schematic of a slice through a space-filling representation of a protein to illustrate the molecular volume calculation: (A) van der Waals volume shown in dark gray is the space within atomic van der Waals radius; (B) to obtain the protein total volume, the atomic radii are expanded until the free volume inside the protein is filled; (C) van der Waals volume (light gray) plus free volume (light gray); (D) molecular protein volume as defined in this paper.

the free (light gray) and van der Waals (dark gray) volumes of the protein, the sum of which is the total, extended volume (Figure 2D). The protein molecular packing density is the ratio of the van der Waals and total volumes, $P_d = V_{vdW}/V_{prot}$.

To obtain the value of V_{prot} from a grid-based calculation, the interstitial volume must be recognized as part of the space occupied by the protein. This recognition is accomplished by artificially extending the atomic van der Waals radii (Figure 2B). We introduce an "extension parameter", γ, which is defined as the ratio of the expanded atomic radius, r^i, and the van der Waals radius, r^i_{vdW}, of each atom i in the protein molecule. Protein volumes, calculated as a function of γ, are shown in Figure 3. The volume increases rapidly with γ up to a value of 1.2 to 1.22 and then more slowly for larger values of γ. We interpret the behavior in the first part of the graph, for γ ranging from 1 to 1.2−1.22, to be "filling" of cavities and packing defects inside the protein. In the second part of the graph, $\gamma >$ 1.22, the volume increases more slowly with γ because all the interstitial spaces between protein atoms have been filled and the increase in volume is due only to external layers added at the protein molecular surface. The value of γ at the intersection of the upper and lower regression lines is the "filling" parameter characteristic to the protein and related to protein packing density. Thus, to calculate the total or extended volume of the protein, the radius of each atom is multiplied by a value γ near this intersection. A constant value of γ, based on the behavior of the volume calculated for the crystallographic average structures in Figure 3, was used for analyzing the MD trajectories. To allow for volume fluctuations during the simulation at 300 K, a value of γ equal to 1.3, slightly larger than 1.22, was used. The average width of the external layer of volume introduced by atomic extension at the protein molecular surface is 0.54 Å (based on an average atomic radius of 1.8 Å). Therefore, according to the algorithm used in this study for the protein total volume calculation, the protein surface lies between

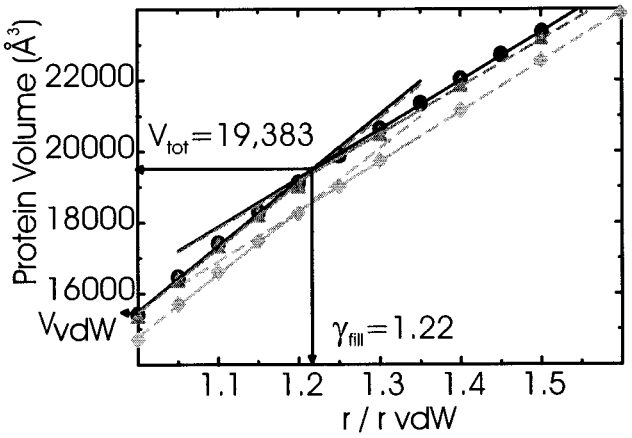

Figure 3. Protein volume as a function of $\gamma = r^i/r^i_{vdW}$. Key: (● and solid lines) α-lactalbumin; (△ and gray dashed lines) HEW lysozyme; (◇ and gray dot−dashed lines) RNase A. The total volume of each protein, V_{tot}, is the value at the intersection of the regression lines for the upper and lower regimes of volume behavior. At $\gamma = 1.0$, $r^i = r^i_{vdW}$ and the calculated volume is the van der Waals volume, V_{vdW}, of each protein (the y-intercept of the graph). The value of $\gamma = \gamma_{fill}$ at the regression lines intercept is characteristic to each protein and is related to the protein packing density. The characteristic V_{tot}, V_{vdW}, and γ_{fill} for α-lactalbumin are shown in the graph.

its molecular surface and the solvent accessible surface obtained using a 1.4 Å radius probe. The total protein volume calculated here is smaller than the volume enclosed within the solvent-accessible surface. This technique for protein total volume calculation has some similarities to that used for estimating partial molar volume of small, approximately spherical solutes in infinitely dilute hard-sphere binary mixture models.[26] For these molecules, fitting of the model with experimental data yields a width of the "extra layer" at the molecular surface ranging from 0.3 to 0.6 Å. With the atomic expansion of $\gamma =$ 1.3, the protein interior is essentially filled, so that the packing defects and small cavity volume are shared by the adjacent atoms; in a 10 Å cubic box centered at the protein center of mass, there is <0.1% free volume. Thus, no large cavities in the protein interior remain in the protein volume calculation.

Compressibility Calculation. *Method 1: β_T Calculated from Protein Molecular Volume Fluctuations and Averages.* Globular protein compressibilities were calculated from the protein molecular volume average and fluctuations from MD simulations using the NTP statistical relationship of eq 2. Protein molecular volumes were calculated using the method described in the previous section with $\gamma = 1.3$ from coordinate sets at 0.4 ps intervals, yielding a total number of 2000 values for the length of the simulation run. Figure 4 shows the protein volume histogram for α-lactalbumin, the time average of the total volume, $\langle V_{prot} \rangle$, and the standard deviation in the volume distribution, σ_V. The parameter σ_V^2, is the total volume fluctuation: $\sigma_V = \langle \Delta V_{prot}{}^2 \rangle^{1/2}$. The volume histograms are well defined, and the compressibilities converged in ~600 ps.

Method 2: *β_T Calculated from Particle Number Fluctuations and Averages.* Isothermal compressibility can be calculated from particle density fluctuations in a box of fixed volume (eq 3). This method[4] involves immersing the region of interest in a fixed-volume grid of specified size and is applicable to a homogeneous system. It is used here to estimate β_T for pure water. The time average of the number and fluctuations of the occupied grid points are used to determine the compressibility.

Method 3: β_T Calculated from Whole Simulation System Volume Fluctuations and Averages. The thermodynamic system

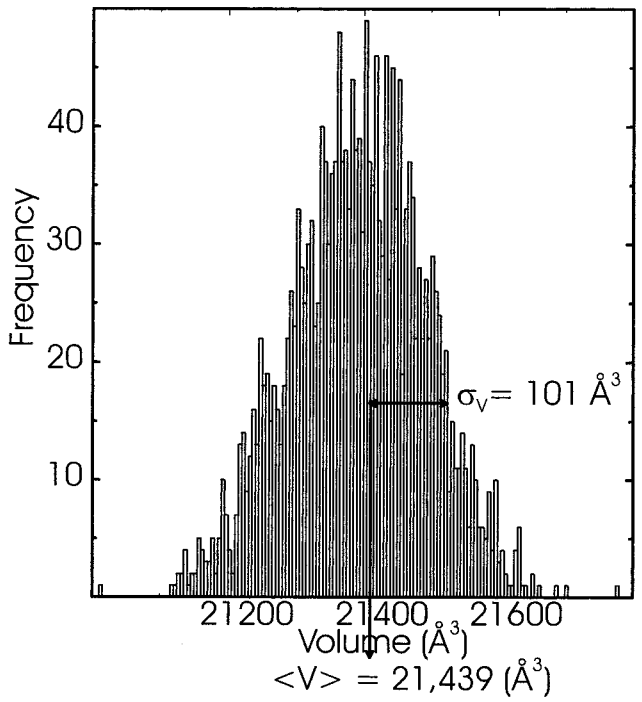

Figure 4. Volume histogram for α-lactalbumin. The compressibility, β_T^{calc}, is calculated using $\langle V\rangle = \langle V_{prot}\rangle$ and $\sigma_V^2 = \langle \Delta V^2\rangle = \langle \Delta V_{prot}^2\rangle$ in eq 2.

that is represented by the solvated proteins in bulk water will have a solution compressibility, β_T^{sol}, that can be estimated directly from MD simulations at constant P and T. The volume of the full truncated octahedron systems, including water and protein, is calculated to maintain constant pressure by coupling the size of the simulation box to the trace of the pressure tensor. The solution compressibility is estimated from these values for the octahedron volume by calculating the whole system volume average and fluctuations and using eq 2. This method is also used to estimate bulk water compressibility.

Error Estimates. *Errors in Volume Calculation.* The primary source of error in the protein total volume calculation is the external layer of volume introduced by atomic expansion, $\gamma = 1.3$, at the protein molecular surface. The average width of this layer is 0.54 Å, based on an average atomic radius of 1.8 Å. The proteins are assumed to be spherical, with radius equal to the time-average radius of gyration from the MD simulation, $\langle R_{gyr}\rangle$, listed in Table 1. The relative error in the volume calculation, $\delta V/V$, is estimated as:

$$\frac{\delta V}{V} = \frac{(\langle R_{gyr}\rangle + 0.54)^3 - \langle R_{gyr}\rangle^3}{\langle R_{gyr}\rangle^3} \tag{4}$$

The estimated error in the total volume calculation is 10% for trypsin and 12% for α-lactalbumin, HEW lysozyme, and RNase A. For trypsin inhibitor, the estimated error in the volume calculation is 15%. However, not all the volume of the external layer will constitute an error. As Lee pointed out in 1983,[4] a "border" of empty space exists between a solute molecule and a solvent. This border will have a variable width, depending on the type of protein−water interaction.

Errors in Volume Fluctuation Calculation. Errors in the volume fluctuations are calculated assuming that the protein volume fluctuates between two R_{gyr} values: $\langle R_{gyr}\rangle$ and $(\langle R_{gyr}\rangle + \Delta R_{gyr})$. The term ΔR_{gyr} is the fluctuation or standard deviation in the protein radius of gyration from MD simulations. The real fluctuation in volume is

$$\Delta V_{real} = (\langle R_{gyr}\rangle + \Delta R_{gyr})^3 - \langle R_{gyr}\rangle^3 \tag{5}$$

and the calculated fluctuation that includes the external 0.54 Å wide layer is

$$\Delta V_{calc} = (\langle R_{gyr}\rangle + \Delta R_{gyr} + 0.54)^3 - (\langle R_{gyr}\rangle + 0.54)^3 \tag{6}$$

The error in the fluctuation due the volume calculation technique used is:

$$\frac{\delta(\Delta V)}{\Delta V} = \frac{\Delta V_{calc} - \Delta V_{real}}{\Delta V_{real}} \tag{7}$$

Therefore:

$$\frac{\delta(\Delta V)}{\Delta V} = \frac{(\langle R_{gyr}\rangle + \Delta R_{gyr} + 0.54)^3 - (\langle R_{gyr}\rangle + 0.54)^3}{(\langle R_{gyr}\rangle + \Delta R_{gyr})^3 - \langle R_{gyr}\rangle^3} \tag{8}$$

The error (overestimate) introduced in total volume fluctuation values is 7% for trypsin, 8% for α-lactalbumin, HEW lysozyme, and RNase A, and 9% for trypsin inhibitor. Therefore, our method for protein volume calculation is more accurate for larger proteins than for the smaller ones.

Errors in Compressibility Calculation. The error in the compressibility calculation is estimated using the error propagation formalism with eq 2 and the already estimated errors in the volume and volume fluctuation calculation:

$$\frac{\delta\beta_T}{\beta_T} = 2\frac{\delta(\Delta V)}{\Delta V} - \frac{\delta V}{V} \tag{9}$$

According to the relationship in eq 9 and the estimated errors in the volume and volume fluctuations for each protein in particular, β_T is underestimated by 3% for α-lactalbumin, HEW lysozyme, RNase A, and trypsin, and by 6% for trypsin inhibitor. The very reasonable approximation of the protein compressibility (3−6% error) indicates that even if the total volume distribution (histogram) is shifted toward larger than real values, the compressibility calculated from total volume fluctuations

TABLE 1: Total Molecular, V_{Prot}, van der Waals, V_{VdW}, and Partial Specific, $\bar{v}^0$, Volumes and Radii of Gyration, R_{Gyr}, for the Protein X-ray Crystallographic Structures[a]

parameter	BPTI[b]	trypsin	RNase A	HEW lyso	α-lacta
N_{res}[b]	58	223	124	129	123[c]
R_{gyr}	11.33	16.14	14.39	14.01	14.14
$\bar{v}^0$ (mL/g)[d]	0.718	0.719	0.704	0.712	0.736
M_w (amu)[e]	6518	23 200	13 700	14 320	13 635
SAS/res	70.69	42.75	57.55	53.34	58.35
V_{vdW}	6974	25 704	14 716	15 379	15 420
V_{exp}[f]	7815	27 617	16 100	17 003	17 576
V_{prot}^{calc}[g]	9515	32 094	18 298	19 296	19 383
V_{prot}^{corr}[h]	7992	28 884	16 138	17 018	17 095
P_d[i]	0.89	0.93	0.91	0.90	0.88

[a] R_{gyr} is reported in Å, SAS/res in Å^2, and the volumes in Å^3. [b] Number of residues in each protein structure. [c] The last six residues in the α-lactalbumin X-ray structure are not defined. [d] From Gekko and Hasegawa.[10] [e] Protein molecular weight. [f] Calculated from experimental $\bar{v}^0$ and M_w. [g] Protein molecular volume as defined in the text. [h] Obtained from the V_{prot}^{calc} by correcting for the error introduced by the external layer volume (see text). [i] $P_d = V_{vdW}/V^{exp}$.

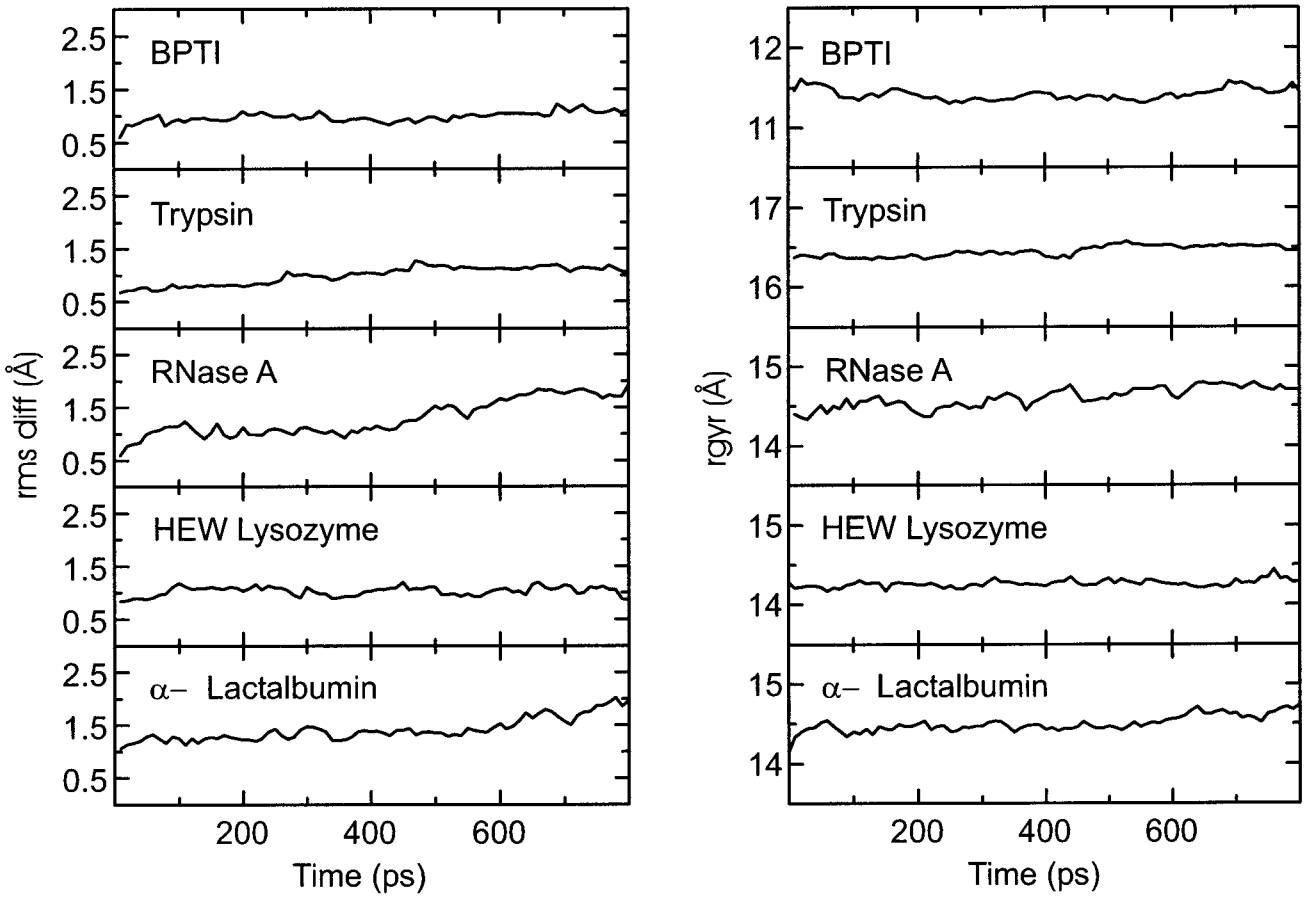

Figure 5. Stability of the five simulations. Time evolution of root mean square differences, rms diff, between main chain atoms of 10 ps average structures in the MD simulation and the crystallographic coordinates is in the left column. Time evolution of the radius of gyration, R_{gyr}, is in the right column.

and protein average volume using this technique is close to its correct value.

3. Results and Discussion

This study reports results for five globular proteins: bovine pancreatic trypsin inhibitor (BPTI; 1bpi),[27] trypsin (2ptn),[28] ribonuclease A (5rsa),[29] HEW lysozyme (1lzt),[30] and α-lactalbumin (1hfz).[31] The coordinates for bovine α-lactalbumin, 1hfz at 2.3 Å resolution, were kindly provided for us by K. R. Acharya prior to their publication. Lysozyme (129 residues), α-lactalbumin (123 residues), and RNase A (124 residues) are globular proteins close in molecular weight and have approximately equal SAS area (Table 1). Thus, these three proteins should have approximately the same contribution from hydration to their partial compressibility. Trypsin (223 residues) is a larger globular protein and BPTI is smaller one with only 58 residues. For comparison, the isothermal compressibility of bulk water (TIP3P model) is also calculated from particle density fluctuations in a box of fixed volume (eq 3) and whole simulation system volume fluctuations (eq 2) from MD simulation of 5944 water molecules, using the same simulation conditions as the five solvated proteins.

Simulation Characteristics. Average root-mean-square differences (rms diff) in atomic coordinates for the main chain atoms, N, C_α, and C, from the starting, energy minimized crystallographic structures range from ~1 Å for BPTI, trypsin, and HEW lysozyme to 1.4 Å for α-lactalbumin and RNase A (Figure 5). The radii of gyration of the proteins are almost constant during the simulation, but an increase of 0.83 to 1.5% with respect to the corresponding radius of the starting crystallographic structure is observed (Figure 5 and Tables 1 and 2).

TABLE 2: Time-Average Volumes, Radius of Gyration, Volume Fluctuations, and Compressibilities Calculated from MD Trajectories[a]

parameter	BPTI[b]	trypsin	RNase A	HEW lyso	α-lacta
$\langle R_{gyr}\rangle$	11.45	16.47	14.62	14.24	14.45
$\langle V_{vdW}\rangle$	7200	25 757	14 928	15 681	15 666
$\langle \Delta V^2_{vdW}\rangle^{1/2}$	11.9	25.7	20.9	21.2	21.2
$\langle V_{prot}\rangle$	9960	32 620	20 400	21 191	21 439
$\langle \Delta V^2_{prot}\rangle^{1/2}$	45.0	91.8	76.3	93.3	101.0
$\langle V_{prot}\rangle^{corr}$	8366	29 358	17 952	18 648	18 867
P_d[c]	0.86	0.88	0.83	0.84	0.83
β_T^{calc}[d]	6.2	6.4	7.5	9.9	11.5
$\Delta\beta_T/\beta_T$ (%)[e]	−6	−3	−3	−3	−3
β_T^{corr}[f]	6.6	6.5	7.8	10.2	11.8
β_T^{exp}[g]	5.152	5.16	5.48	7.73	12.4

[a] The radius of gyration is in Å, the volume and volume fluctuations are in Å^3, and the compressibility unit is 10^{-6} atm^{-1}; $\langle\ \rangle$ designates a time average. [b] BPTI compressibility estimated from independent simulation data.[13,33] [c] Ratio between the average van der Waals, $\langle V_{vdW}\rangle$, and the corrected average total protein volumes, $\langle V_{prot}\rangle^{corr}$. [d] Isothermal compressibility calculated using eq 2 and Method 1. [e] Estimated relative error in compressibility calculation, eq 9. [f] Corrected compressibility calculated from β_T^{calc} and $\Delta\beta_T/\beta_T$ (%). [g] Experimental compressibility estimated from ultrasound velocity measurements.[10]

Molecular Volume Calculation. Table 1 summarizes the main characteristics of the protein crystallographic structures considered in this study: the number of residues, N_{res}, radius of gyration, R_{gyr}, molecular weight, M_w, and the experimental

partial specific volumes,[10] v^0. The SAS of the proteins was calculated using Richards' method[32] with a 1.4 Å probe radius. The experimental total molecular volume, V^{exp}, was calculated from v^0 and M_w. The protein total molecular volume, V_{prot}^{calc}, and van der Waals volume, V_{vdW}, were estimated using the technique just described. The corrected total molecular volume, V_{prot}^{corr}, is obtained from the V_{prot}^{calc} by correcting for the error introduced by the external layer volume (see *Error Estimates* in *Methods*). The V_{prot}^{corr} values for the five proteins are very close to the V^{exp} values (Table 1). The term V^{exp} includes any hydration contribution to the partial specific protein volume, yet these contributions appear to be negligible because the volume change on protein denaturation is small.[6,26] Thus, the similarity between V^{exp} and V_{prot}^{corr} values is reasonable. That V_{prot}^{corr} for α-lactalbumin is smaller than its experimental molecular volume reflects the fact that the crystal structure has six undefined residues. Adding an average volume per residue of 139 Å^3 calculated from the α-lactalbumin total volume for 123 residues, the corrected total volume becomes 17 929 Å^3, which is in good agreement with the experimental total molecular volume of 17 576 Å^3. The calculated average volume per residue of 139 Å^3 is also close to literature values for average residue volumes.[33]

Compressibility Calculation for Bulk Water. The bulk water simulation (5944 water molecules, TIP3P model) corresponds to average values for energy and density that are in good agreement with the reported values[15] from Monte Carlo simulations. This section gives results for bulk water compressibility calculated from its MD trajectory following methods 2 and 3 already outlined. An alternative approach based on excess volumetric properties of solvated molecules was developed by Lockwood and Rossky[34] and gives a value of 28×10^{-6} atm^{-1}.

Isothermal compressibility was calculated for bulk water using both the localized density (eq 3, Method 2) and the whole system volume (eq 2, Method 3). For Method 2, the grid was varied in size from 1000 to 64 000 Å^3, positioned at the center (0,0,0), or off the center (0,10,10) of the simulation octahedron. Finite-size effects are observed in the convergence of the compressibility value for short simulation times (Figure 6A, all but top curve) and small box sizes (Figure 6B). Two criteria for overcoming finite size effects in bulk water MD simulations are a simulation time longer than 300 ps (or 600 ps for complete convergence) and a number of water molecules in the simulation system larger than 267, which corresponds to box sizes of 8000 Å^3. The compressibility calculation converged in ~600 ps for box volumes <8000 Å^3 to 46×10^{-6} atm^{-1} (Figure 6 A). No effect was observed on the position of the box where the density fluctuation is monitored. For Method 3, the compressibility converged in 600 ps to 61×10^{-6} atm^{-1} (Figure 6A, top curve). The whole simulation system (the truncated octahedron shape; Figure 1) volume average for water is 179 027 Å^3 (Table 3).

Isothermal compressibility for bulk water at 300 K has been determined experimentally through ultrasound velocity measurements and is 45×10^{-6} atm^{-1}.[35] There is an excellent agreement between the experimental and calculated compressibility of 46×10^{-6} atm^{-1} using Method 2. On the other hand, water compressibility estimated from whole system volume fluctuations (61×10^{-6} atm^{-1}) is 35% higher than the experimental compressibility. The average volume of the simulation box divided by the number of water molecules is 30.1 Å^3, similar to standard average volume per water molecule reported elsewhere.[33] This agreement suggets that the error in the compressibility calculation using Method 3 is due to errors in the estimated total volume fluctuations. A possible explanation is that fluctuation value is sensitive to the numerical errors in estimating an exact geometrical volume for a specified system of water molecules. Errors in estimating whole system volume fluctuations are recognized for constant pressure and temperature MD simulations; constant pressure and temperature MD simula-

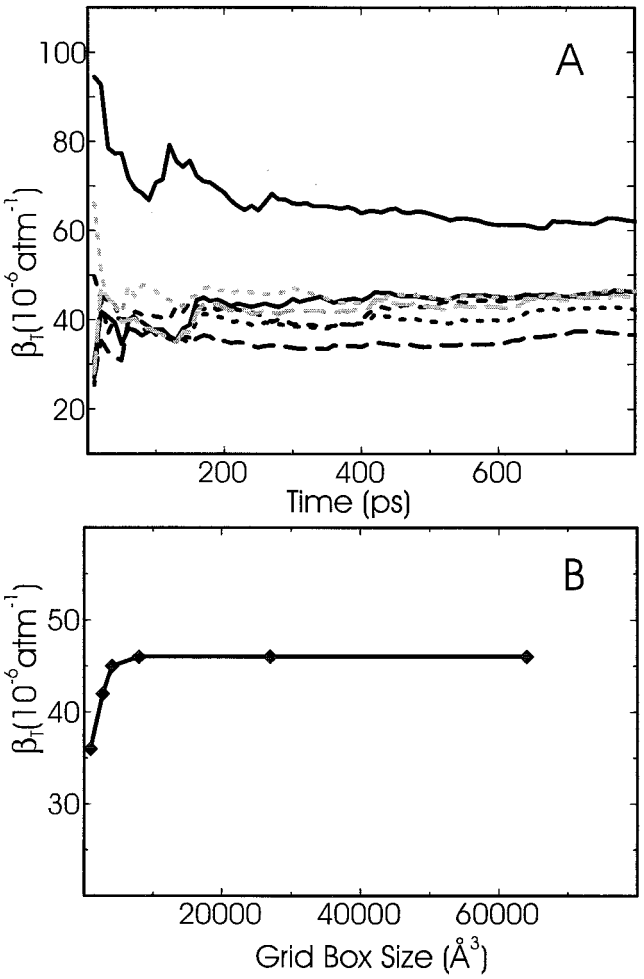

Figure 6. (A) Time evolution of the calculated compressibilities for TIP3P water model from particle density fluctuations in a cubic, fixed box volume (lower curves). Key: (dotted gray line) 64 000 Å^3; (solid black line) 27 000 Å^3; (short dashed black line) 8000 Å^3; (dashed gray line) 5832 Å^3; (dotted black line) 4096 Å^3; (dashed black line) 1000 Å^3; (upper curve, black, solid line) whole simulation octahedron volume fluctuations and average. (B) Bulk water (TIP3P model) isothermal compressibility from the limiting value in (A) as a function of the grid box size.

TABLE 3: Solution Compressibilities, β_T^{sol}, Calculated from Simulation Box Volume Fluctuations, $(\Delta V^2)_{sol}^{1/2}$, and Average $\langle V \rangle_{sol}$[a]

parameter	BPTI	trypsin	RNase A	HEW lyso	α-lacta	water
$\langle V \rangle_{sol}$	123 643	179 659	121 679	120 124	121 947	179 027
$(\Delta V^2)_{sol}^{1/2}$	560	632	531	511	526	672
β_T^{sol}	63(61)[b]	55	54	52	53	61
N_{water}	3830	4901	3487	3170	3465	5944
Φ^p (%)[c]	7	16	15	16	15	0
$\beta_{T_{pred}}^{sol}$[d]	57.2	52.3	53.0	52.8	53.6	NA
$\Delta\beta_T^{sol}$[e]	3.8	2.7	1.0	−0.8	−0.6	NA

[a] Volumes are reported as Å^3 and compressibilities as 10^{-6} atm^{-1}. [b] After 1800 ps MD simulation for BPTI. [c] $\Phi^p = \langle V_{prot} \rangle^{corr} / \langle V \rangle_{sol}$. [d] Solution compressibilities as in eq 10, using the calculated protein compressibility, $\beta_T^p = \beta_T^{corr}$ (Table 2), bulk water compressibility, $\beta_T^w = 61 \times 10^{-6}$ atm^{-1}, and volume fractions, Φ^p and $\Phi^w = 1 - \Phi^p$. [e] $\Delta\beta_T^{sol} = \beta_T^{sol} - \beta_{T_{pred}}^{sol}$.

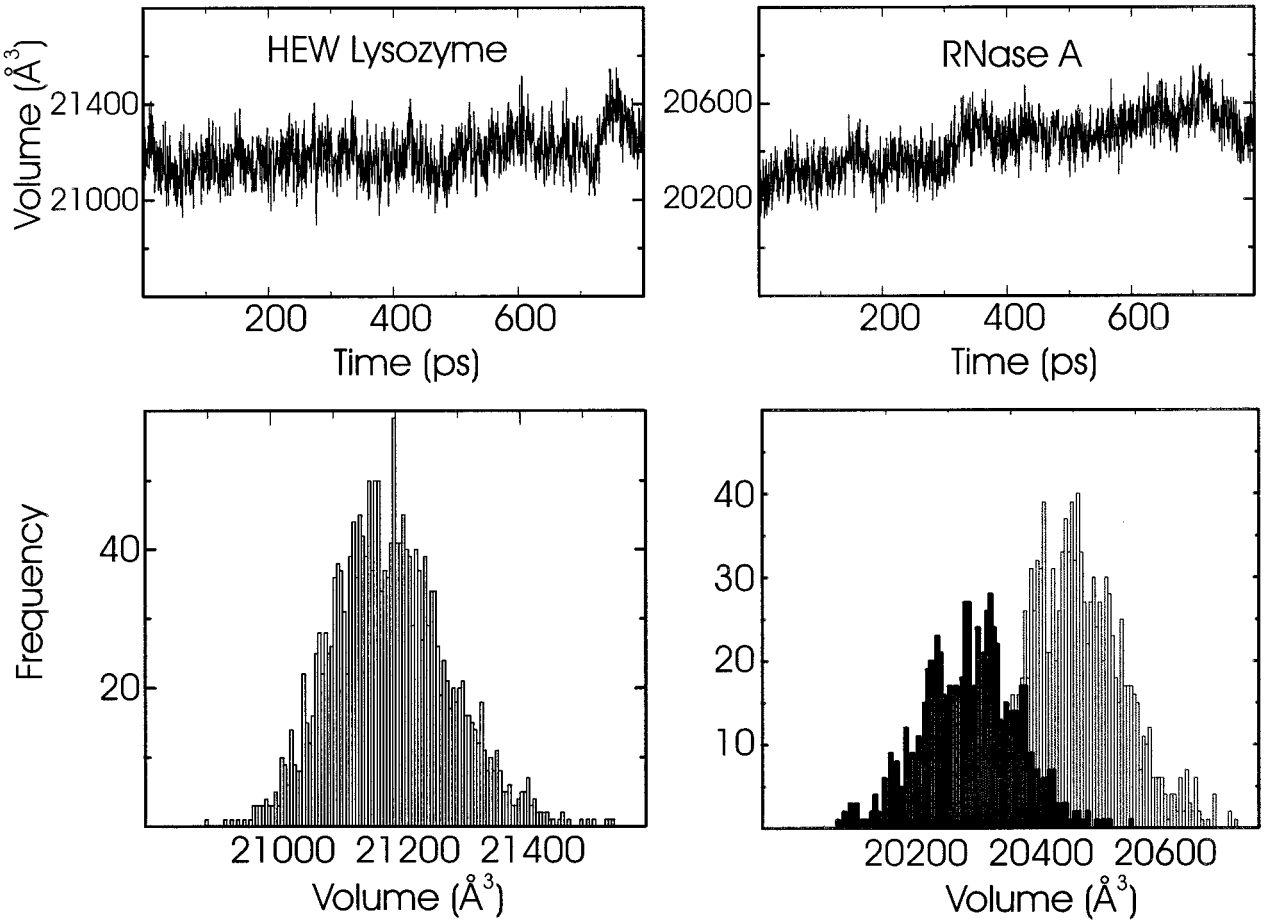

Figure 7. Total volume in time and volume histograms for HEW lysozyme and RNase A. The volume distribution is unimodal for HEW lysozyme (bottom, left column) and bimodal for RNase A (bottom, right column).

tions of water in a cubic shaped unit cell[12] reported statistical errors in the volume fluctuations of 30%.

The results reported here for TIP3P water compressibility differ from those reported by Jorgensen and co-workers.[15] These authors calculated isothermal compressibility from 1 500 000 steps (configurations) generated through Monte Carlo simulations in an NTP ensemble at 25 °C and 1 atm for systems of 125 monomers, but the volume fluctuations were reported not to have converged. The reported estimate for isothermal compressibility for the TIP3P model was 18×10^{-6} atm^{-1}. The low estimate in the compressibility calculation for TIP3P from Monte Carlo simulations is likely due to number of configurations and the system size, based on the finite size effects demonstrated here in Figure 6.

Protein Compressibility Calculation and Correlation with Experimental Data. The β_T of the five proteins was calculated according to Method 1 (eq 2) from protein molecular volume averages and fluctuations. The evolution of the molecular volumes in time for BPTI, trypsin, HEW lysozyme, and α-lactalbumin displayed approximately unimodal volume distributions (Figure 7). For ribonuclease A, the evolution of the molecular volume in time reveals a bimodal distribution with corresponding compressibilities of 7.75×10^{-6} atm^{-1} (from 1 to 300 ps) and 7.41×10^{-6} atm^{-1} for the second part of the trajectory (300−800 ps). The calculated compressibility for RNase A is a weighted average of these two values. Figure 7 shows a comparison between the molecular volume in time and volume histogram for HEW lysozyme (a unimodal distribution) and RNase A (bimodal distribution).

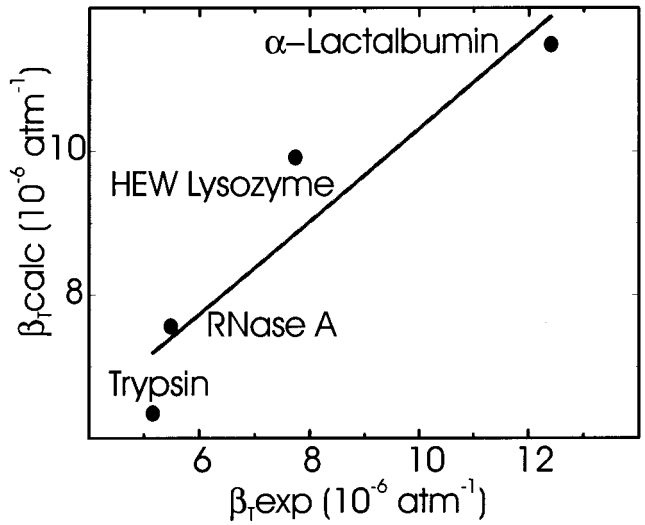

Figure 8. Calculated versus experimental compressibility for globular proteins. Experimental compressibilities are estimated from sound velocity measurements.[10] Calculated compressibilities are from Method 1, molecular volume fluctuation and average.

Table 2 shows the calculated, β_T^{calc}, and experimental, β_T^{exp}, compressibilities of the five proteins. There is a good correlation between the calculated and experimental compressibilities, as shown in Figure 8 (the correlation coefficient is 0.94 and the regression coefficient is 0.697) for the four globular proteins for which experimental values of compressibility have been

determined from ultrasound velocity measurements.[10] According to eq 6, the value of β_T^{calc} is underestimated by 6% for BPTI and 3% for the other four proteins. The corrected compressibilities, β_T^{corr}, do not improve the correlation with the experimental data. The correlation between β_T^{corr} and other experimental values for compressibility[14] is less satisfactory. The experimentalists have no explanation for the discrepancies between measured values. Comparison for BPTI is made with earlier simulation studies that reported volume fluctuations[33] and a finite volume difference from two MD simulations at different pressures.[13] Compressibility values obtained with their results are 5.86 and 4.43 $\times$ 10^{-6} atm^{-1}, respectively. The calculated β_T value reported here is 6.22 $\times$ 10^{-6} atm^{-1}. Paci and Marchi[12] estimated the compressibility of HEW lysozyme from finite difference calculations to be 9.1 $\times$ 10^{-6} atm^{-1}, which is close to value calculated here from protein molecular volume fluctuations of 9.92 $\times$ 10^{-6} atm^{-1}.

As a general trend, the calculated compressibilities, β_T^{calc}, are larger than the experimental values, β_T^{exp}. The experimental estimate of compressibility measures the intrinsic protein component (contributed by the volume occupied by the protein molecule) and the component from the hydration layer of water molecules. The hydration contribution is generally considered to depend on the protein SAS.[11,26] Nonetheless, the variation in experimental values of β_T cannot be accounted for by solvent accessible surface alone, as was pointed out early on.[5] The specific SAS, SAS/res, are seen from Table 1 to differ by almost 30 $Å^2$ between BPTI (70.69 $Å^2$) and trypsin (42.75 $Å^2$), whereas their compressibilities are estimated to be almost the same. Ribonuclease A, lysozyme and α-lactalbumin have very close SAS/res values (57.55, 53.34, and 58.35 $Å^2$, respectively), yet their compressibility varies from 5 to 12 $\times$ 10^{-6} atm^{-1}. Because of their similar SAS as well as similar distribution of polar and nonpolar surface area,[36] differences in the compressibility among the three proteins are likely due to intrinsic properties of the protein rather than hydration. Together, these results suggest that the intrinsic protein compressibility is a significant factor in the observed variation in β_T among proteins.

The protein total volume fluctuations for the five proteins studied range from 45 to 112 $Å^3$, representing between 0.3 and 0.47% of the total protein volume. The fluctuations are due to local, small, and rapid (picosecond time scale) changes in bond lengths and angles and torsion angles of the protein. If these volume fluctuations were concentrated in one region of the protein structure, the "cavity" created could accommodate 1.5 to 4 solvent (water) molecules. The volume change would also be large enough to allow channel formation for the exchange between buried water molecules and solvent molecules.[5,34] These volume fluctuations of globular proteins are in line with other estimates from experimental data.[3,5,10]

Compressibilities calculated from density fluctuations in a fixed-volume grid, Method 2, are consistent with the macroscopic compressibility of homogeneous systems (e.g., bulk water). This is not the case for inhomogeneous systems (e.g., proteins) where compressibilities calculated from fixed-volume density fluctuations are related to the specific, local structure of the part of the protein included in the grid. For the five proteins studied here, compressibilities calculated from fixed-volume density fluctuations varied with the grid size and position. We calculated compressibilities in cubic boxes with sides of 10, 14, and 20 Å centered at the protein mass center. The calculated values of the compressibility increased with the box size, ranging from 4.5 to 9 for trypsin, from 7 to 11 for RNase A, from 9 to 13 for HEW lysozyme, and from 10 to 13($\times$$10^{-6}$ atm^{-1}) for α-lactalbumin. Thus, Method 2 is not reliable for estimating absolute protein molecular compressibility as the values are not uniquely determined.

Compressibility calculation in regions of the protein molecule containing particular types of secondary structure elements (α-helix, β sheath, or loops) show that they have different contributions to the total compressibility of the protein. Compressibility in mostly α regions of the proteins is lower than compressibility in mostly β regions, and this is lower than compressibility in the protein loops. The calculation has been performed for the HEW lysozyme MD-generated configurations and included any solvent molecules that fall in the grid region. The calculated compressibilities are 6.6 $\times$ 10^{-6} atm^{-1} for the mostly α region, 7.7 $\times$ 10^{-6} atm^{-1} for the mostly β region, and 12.3 $\times$ 10^{-6} atm^{-1} for the loop-containing region. This result is in qualitative agreement with the experimental results on pressure-induced amide ^{15}N chemical shifts in BPTI, which show the same trend in the magnitude of the local compressibilitie[38] and changes in the local configuration of HEW lysozyme protein crystals when high pressure is applied.[39]

Source of Protein Compressibility. The source of total molecular volume change can be assessed from the detailed information contained in an MD trajectory. van der Waals volume fluctuations, $\langle \Delta V^2_{vdW} \rangle^{1/2}$, account for ~25% of the total volume fluctuations, $\langle \Delta V^2_{prot} \rangle^{1/2}$; thus, changes in the free volume inside the protein are the major source of volume variation (Table 2). A compressibility value calculated from van der Waals volume fluctuations is ~1 order of magnitude smaller than that resulting from the total molecular volume fluctuation. For example, the HEW lysozyme average van der Waals volume is 15 681 $Å^3$ and the van der Waals volume fluctuation is 21 $Å^3$ (Table 2). The corresponding isothermal compressibility (eq 2 and Method 1) is 0.69 $\times$ 10^{-6} atm^{-1}. For the same protein, the total volume average is 21 191 $Å^3$ and the volume fluctuation is 93.3 $Å^3$, leading to a calculated isothermal compressibility of 9.92 $\times$ 10^{-6} atm^{-1}. Thus, the compressibility calculated from total volume average and fluctuations is 14 times higher than the compressibility calculated from corresponding van der Waals values. The protein compressibility is due to fluctuations in the unoccupied free volume of the interstitial space in the protein interior defined by the protein tertiary structure. The large decrease in compressibility on protein unfolding[36] can be thus related to the loss of protein secondary and tertiary structure.

Compressibility of the Protein−Water Solution and Hydration Effects. The compressibilities of the protein−water solutions and bulk water systems were estimated from the whole truncated octahedron average volume, $\langle V \rangle_{sol}$, and fluctuations, $(\Delta V^2)^{1/2}_{sol}$, (Table 3), as described in Method 3. This approach allows a direct assessment of the hydration water contribution to solution compressibility because fluctuations in the total simulation box volume are due to fluctuations in the volume occupied by water as well as fluctuations in the protein molecular volume. The protein−water solution compressibilities, β_T^{sol} (Table 3), are between 53 and 63 $\times$ 10^{-6} atm^{-1}, compared with the compressibility of bulk water calculated from whole system volume fluctuations of 61 $\times$ 10^{-6} atm^{-1}. Figure 9 shows the evolution in time of the calculated solution compressibilities for the five proteins (lower curves) and bulk water (top curve).

The detailed information provided by MD simulations is exploited to investigate the extent to which the compressibility of hydration water differs from that of bulk water. The hydration water compressibility has been estimated in the literature to be between 18 and 39 $\times$ 10^{-6} atm^{-1}.[5,11,40] These estimates are based on a hydration shell model for proteins and experimental

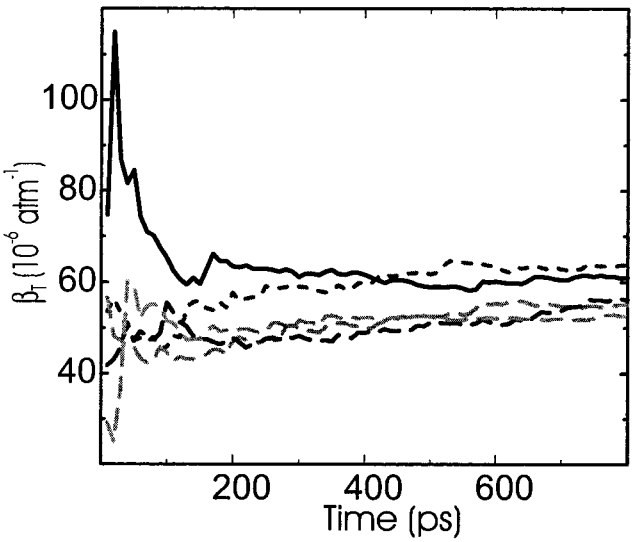

Figure 9. Time evolution of the calculated compressibilities for TIP3P water model and globular protein solutions from Method 3, using the whole simulation system to obtain volume fluctuations and averages. Top curve is β_T^{sol} for bulk water, and the lower curves are for BPTI (....), trypsin (gray; - — - -), RNase A (black; - — - -), α-lactalbumin (gray; ●—●—).

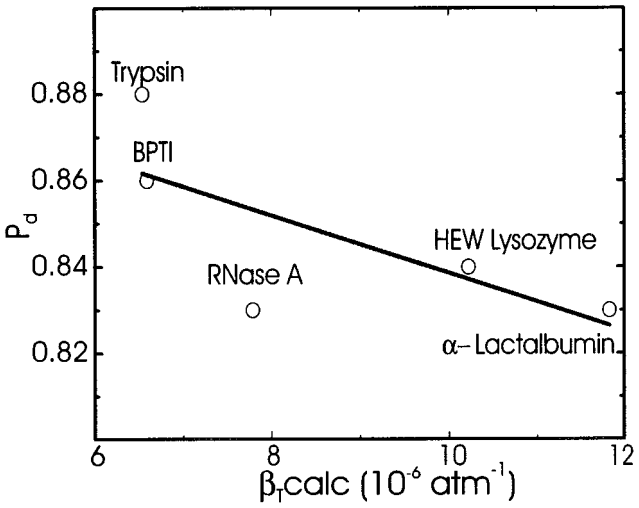

Figure 10. Isothermal compressibility versus packing density.

compressibilities from small solutes. Recent evaluation of solvation effects from different solute functional groups finds that the solvent perturbation is highly localized,[41] a result which lends support to the use of the small solute compressibilities to estimate protein hydration effects. A simple approach is proposed here in which water and protein are assumed to be independent components of the protein−water system. We ask how well pure water compressibility and protein compressibility account for the solution compressibility. In a first approximation, the compressibility of a system of two independent and noninteracting components can be expressed as a function of the individual compressibilities and their respective volume fractions. Applying this principle to the protein−water solution, the compressibility of the system can be predicted from the protein compressibility, β_T^p, the compressibility of the bulk water, β_T^w, and the volume fractions, Φ^p and Φ^w, of the protein and bulk water, respectively:

$$\beta_{T_{pred}}^{sol} = \Phi^p \beta_T^p + \Phi_w \beta_T^w \qquad (10)$$

A detailed description of the assumptions involved and the limitations of this expression for protein solutions is given by Lee.[3] The number of water molecules in the simulation boxes, N_{water}, and the average volume of the whole simulation boxes, $\langle V \rangle_{sol}$, are different in the six simulations (Table 3). The protein volume fraction, Φ^p, defined as the average total protein volume divided by the average simulation box volume, $\Phi^p = \langle V_{prot} \rangle^{corr} / \langle V \rangle_{sol}$, is only 7% for trypsin inhibitor, whereas for trypsin and lysozyme it is 16% and for RNase A and α-lactalbumin it is 15%. The predicted solution compressibilities, $\beta_{T_{pred}}^{sol}$, determined with $\beta_T^p = \beta_T^{corr}$ and $\beta_T^w = 61 \times 10^{-6}$ atm^{-1} according to eq 10, are shown in Table 3 along with solution compressibilities calculated from the whole simulation box volume fluctuations, β_T^{sol}. The deviation between predicted (eq 10) and calculated (from MD) solution compressibilities, $\Delta\beta_T^{sol}$, is largely due to the difference between the bulk and hydration water compressibility with the analytical expression for this difference being $(\beta_T^w - \beta_T^h)\Phi^h$ (Φ^h is the hydration water volume fraction). The inherent limitation in estimating compressibility from whole system fluctuations (Method 3) is recognized from the error in bulk water compressibility. Nevertheless, $\Delta\beta_T^{sol}$ should be a reasonable indicator of the hydration water contribution to solution compressibility given that it is an internally consistent determination. $\Delta\beta_T^{sol}$ range from 3.8×10^{-6} atm^{-1} for BPTI to -0.83×10^{-6} atm^{-1} for HEW lysozyme (Table 3). These values indicate that hydration water compressibility does not differ greatly from bulk water compressibility, in agreement with the estimates from simple model systems[3,11,40] that suggest the compressibility of the hydration shell is 78% of the bulk water compressibility. To provide a definitive value for the magnitude of the hydration water compressibility from $\Delta\beta_T^{sol}$ would require a model of protein hydration to estimate Φ^h, an exercise left to the reader. The variation in $\Delta\beta_T^{sol}$, including the difference in sign, is noteworthy. The variation exists even after taking into account the differences in total volume of hydration for the different size proteins. Such a range in $\Delta\beta_T^{sol}$ implies a nonuniform contribution from the protein surface that could arise from differences in the polarity of surface groups or even in more subtle chemical differences. For example, methyl and methylene groups have been reported to contribute to compressibility with opposite sign.[41] A better understanding of this apparent variation in compressibility of hydration layers around proteins is worth pursuing in future work.

Protein Compressibility Related to Structural Factors. Packing density, P_d, reflects the strength of the atomic interactions in the interior of the protein and would appear intuitively to be linked with compressibility. P_d, defined as the ratio between the average van der Waals, $\langle V_{vdW} \rangle$, and the corrected average total volumes, $\langle V_{prot} \rangle^{corr}$, is plotted against β_T in Figure 10. The packing densities obtained in this study, 0.83 to 0.88, are higher than the average value of 0.75 obtained by other authors[2,32] because the van der Waals radii in the CHARMM22 parameter set are larger than the radii used in these studies. The scatter shown in Figure 10, giving a low correlation coefficient of 0.72, suggests that packing density alone cannot explain the differences in the compressibility among globular proteins. Another factor that likely contributes to the amplitude of the volume fluctuations is the character of the forces inside the protein. The balance between hydrophobic and polar interactions determines the strength of the net forces inside the protein and thus its conformational mobility and molecular volume fluctuations. Residues must not only have room to move

(i.e., low packing density), but also the freedom to do so (i.e., not engaged in strong, directional interactions). Other factors that could influence the magnitude of the compressibility are the hydration effect and the number of disulfide bonds. Values for the SAS/res are given in Table 1. The number of disulfide bonds per total number of amino acids for BPTI, α-lactalbumin, lysozyme, ribonuclease A, and trypsin are 3/58, 4/123, 4/129, 4/124, and 6/223, respectively.

The isothermal compressibility of globular proteins does not appear to be explained by the correlation with a single variable. Rather, the compressibility is a complex function that may result from a combination of factors, such as protein packing density, hydrophobicity, polarity, specific solvent accessible area, number of disulfide bonds per amino acid, and the type of secondary structure elements.

4. Conclusions

In this paper, we describe an algorithm to estimate the intrinsic compressibility for globular proteins from MD simulations in an NTP ensemble using the statistical definition of compressibility and an efficient method to calculate protein molecular volumes.

A high correlation between experimental compressibilities and calculated isothermal compressibilities for globular proteins is found. Compressibility of globular proteins results largely from imperfect packing, which allows for large fluctuations in the interstitial space in the protein interior. van der Waals volume fluctuations account for only 10% of the protein compressibility. Variations in protein compressibility appear to be related to differences in packing densities, although other factors, such as the specific solvent accessible areas and the strength and character of the average interatomic forces, are likely to contribute. A simple approach to assess the contribution to solution compressibility from hydration waters suggests a small difference between hydration and bulk water compressibility.

Calculated isothermal compressibility for bulk water, TIP3P model, is in excellent agreement with the experimental data. Two criteria to overcome finite-size effects in bulk water simulation emerged: convergence required a simulation time longer than 300 ps (600 ps for complete convergence of compressibility calculation) and a simulation size of ∼260 or more water molecules.

Acknowledgment. This work was supported in part by a DOE/Sloan Fundation Postdoctoral Fellowship to V. M. D., NIH grant AI39639 to C. B. P., and a Research Career Development Award GM 00661 to C. B. P. The Structural Biology group at Purdue University received support from the Lucille P. Markey Foundation and the Purdue Academic Reinvestment Program. The authors thank Drs. B. K. Lee, T. Ichiye, and B. Lucier for helpful suggestions and discussions.

References and Notes

(1) Hill T. L. *An Introduction to Statistical Thermodynamics*; Addison-Wesley: Reading, MA, 1960; pp 37−38.
(2) Richards, F. M.; Lim, W. A. *Quart. Rev. Biophys.* **1994**, *26*, 423.
(3) Lee, B. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 622.
(4) Phelps, D.; Post, C. B. *J. Mol. Biol.* **1995**, *254*, 544.
(5) Eden, D.; Matthew, J. B.; Rosa, J. J.; Richards, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 815.
(6) Chalikian, T. V.; Gindikin, V. S.; Breslauer, K. J. *J. Mol. Biol.* **1995**, *250*, 291.
(7) Gekko, K.; Tamura, Y.; Ohmae, E.; Hayashi, H.; Kagamiyama, H.; Ueno, H. *Protein Sci.* **1996**, *5(3)*, 542.
(8) Phelps, D.; Rossky, P.; Post, C. B. *J. Mol. Biol.* **1998**, *276*, 331.
(9) Tsang, S. K.; Danthi, P.; Chow, M.; Hogle, J. M. *J. Mol. Biol.* **2000**, *296*, 335.
(10) Gekko, K.; Hasegawa, Y. *Biochemistry* **1986**, *25*, 6563.
(11) Kharakoz, D. P.; Sarvazyan, A. P. *Biopolymers* **1993**, *33*, 11.
(12) Paci, E.; Marchi, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11609.
(13) Kitchen, D. B.; Reed, L. H.; Levy, R. M. *Biochemistry* **1992**, *31*, 10083.
(14) Chalikian, T. V.; Totrov, M.; Abagyan, R.; Breslauer, K. J. *J. Mol. Biol.* **1996**, *260*, 588.
(15) Jorgensen, W. L.; Chandarsekhar, J.; Medura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(16) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminatha, S.; Karplus, M. it *J. Comput. Chem.* **1983**, *4*, 187.
(17) Nosé, S. *J. Chem. Phys.* **1984**, *81(1)*, 511.
(18) Hoover, W. G. *Phys. Rev. A* **1985**, *31(3)*, 1695.
(19) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384.
(20) Nosé, S.; Klein, M. L. *Mol. Phys.* **1983**, *50*, 1055.
(21) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103(11)*, 4613.
(22) Zhang, Y. H.; Feller, S. E.; Brooks, B. R.; Pastor, R. W. *J. Chem. Phys.* **1995**, *103(23)*, 10252.
(23) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(24) Essmann, E.; Perrera, L.; Berkovitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(25) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(26) Lee, B. *J. Phys. Chem.* **1983**, *87*, 112.
(27) Parkin, S.; Rupp, B.; Hope, H., unpublished results.
(28) Walter, J.; Steigemann, W.; Singh, T. P.; Bartunik, H.; Bode, W.; Huber, R. *Acta Crystallogr. B* **1982**, *38*, 1982.
(29) Wlodawer, A.; Borkakoti, N.; Moss, D. S.; Howlin, B. *Acta Crystallogr. B* **1986**, *42*, 379.
(30) Hodsgon, J. M.; Brown, G. M.; Sieker, L. C.; Jensen, L. H. *Acta Crystallogr. B* **1990**, *46(5)*, 54.
(31) Pike, A. C.; Brew, K.; Acharya, K. R. *Structure* **1996**, *4*, 691.
(32) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151.
(33) Gerstein, M.; Tsai, J.; Levitt, M. *J. Mol. Biol.* **1995**, *249*, 955.
(34) Lockwood, D. M.; Rossky P. J. *J. Phys. Chem. B* **1999**, *103*, 1982.
(35) Dorsey, N. E. *Properties of Ordinary Water Substances*; Rheinhold: New York, 1940.
(36) Chalikian, T. V.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 1012.
(37) Cooper, A. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *92*, 2740.
(38) Akasaka, K.; Li, H.; Yamada, H.; Li, R.; Thorensen, T.; Woodward, C. K. *Protein Sci.* **1999**, *8*, 1946.
(39) Kundrot, C. E.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 157.
(40) Chalikian, T. V.; Sarvazyan, A. P.; Breslauer, K. J. *Biophys. Chem.* **1994**, *51*, 89.
(41) Lockwood, D. M.; Rossky P. J.; Levy, R. M. *J. Phys. Chem. B* **2000**, *104*, 4210.