On the Volume of Macromolecules

Abstract: The surface accessible to the solvent and the volume contained within this surface are key notions both for the general description of proteins and other macromolecules, and for studying their interactions with the solvent and protein stability under a change of various parameters. There are many geometrical recipes for assigning a surface area and a volume to a protein. These recipes are not equivalent, and involve parameters whose definition is not always unique. We discuss the relative merits of the molecular and Voronoi definitions of protein volume, and their relation to experimentally measurable quantities like “intrinsic compressibility.” The molecular volume of a protein has the contradictory feature of not reflecting the packing of solvent molecules around it, being at the same time a function of the probe radius. The Voronoi volume does take into account modifications occurring in the solvent due to the presence of the protein and to possible external perturbations. Using the Voronoi volume results in a much better approximation of the experimental intrinsic compressibility. We suggest an explanation for the incorrect behavior of the molecular volume as a function of pressure, and propose using a pressure-dependent probe radius as a remedy in such cases when the molecular volume computation might be preferable. © 1997 John Wiley & Sons, Inc. Biopoly 41: 785–797, 1997

Keywords: Protein volume; protein surface; protein–solvent interface; molecular simulation

INTRODUCTION

In describing proteins, one often has to speak about volumes and surfaces. First, proteins are very dense systems, which means that it makes sense to speak of packing and of the fraction of volume occupied by a protein in a solution. Moreover, arguing in terms of packing may help one understand protein folding: the packing density requirement, combined with the constraints imposed by the presence of a linear covalent structure (a chain with branches), must pose severe limitations on the number of ac-
ceptable three-dimensional structures. The choice of secondary and tertiary structures can often be controlled by simple space-filling requirements.\cite{1} Second, the protein–solvent interface, which is the arena of protein interactions with the solvent, is well known to play a no less important role in the protein folding mechanism, both directly (surface effects) and via volume (packing density) changes. In this context it is natural that one often feels the need to speak about a protein surface. Surface and volume changes occurring when a protein is exposed to denaturing agents could partially explain the loss of stability of secondary and tertiary structures. Recent results\cite{2} confirm that packing forces play an important role in protein stability and specificity.

However, while the notions of volume and surface area are perfectly defined for macroscopic objects, it is not trivial to define them on the scale of a molecule, even for a macromolecule.

If we want to define the volume of a macroscopic object, two procedures are available: the occupied volume can be measured by observing the change of the volume of a fluid in which the object is immersed, or else the object’s volume can be measured as the geometric volume contained inside the object’s surface. For macroscopic solids, this surface is usually easy to define. If fluid–solid interactions do not change the geometry of the solid, and if the surface of the solid is connected (there are no closed holes inside the solid object), these two quantities have the same value.

For a molecule, even for a big molecule such as a protein, the second procedure—finding the volume contained inside a surface—is not well defined: there is no such thing as a clear-cut surface of a macromolecule. And in whichever way we define the molecule’s surface, the geometric volume contained in it may still be different from the volume obtained by immersion (occupied volume).

This is because chemical interactions with the solvent, and at higher concentrations, protein–protein interactions, may change the geometry of the interface, which, in turn, entails not only a change in the geometric volume, but also a possible change in the local density of the solvent near the interface.

The rearrangement a protein–solvent system undergoes under pressure, or under some other change of parameters (temperature, pH, etc.), can be described exactly by giving all the old and new atomic positions (easy to do in a simulation, very difficult in experiment). However, this is not usually what one wants. One would rather prefer to drastically reduce the number of variables, and describe the change in terms of volumes, surfaces, and compressions. In doing so, one also feels that it would be nice to separate the quantities (volumes) and the effects of pressure (compressions) into two parts: a solvent-dependent (hydrational) and solvent-independent (intrinsic).

However, one can easily see that such a splitting may be in part a matter of terminology. When a pressure is applied, several things happen:

(a) There are changes in the protein’s interior. Clearly, this effect unambiguously deserves to be called an intrinsic change.

(b) The protein’s atoms in contact with the solvent change their positions, which results in a change of the shape and the size of the surface, and, therefore, of the degree of its roughness. There are two sources of ambiguity in describing this effect. First, the change of the volume depends on how we define the surface (i.e., how we construct the envelope of those atoms that are in direct contact with the solvent). Second, this effect is due to hydration, but concerns the protein itself: should we call it hydrational or intrinsic?

(c) The solvent atoms come, on the average, closer to the macromolecule. This results in an overall change of the volume. Should it be attributed to the macromolecule or to the solvent?

(d) The solvent gets reorganized. There is no ambiguity here: if this results in a density change, it should be attributed entirely to the solvent.

At the same time, experimentalists have already developed ways of extracting what they call “intrinsic” compressibility from experimental data.\cite{3,4}

Thus one faces the following problems:

(1) to find out which definition of volume based on atomic position best agrees with the experimental results on intrinsic compressibility;

(2) to see if using this volume does, indeed, mean that a separation of intrinsic and hydrational contributions has been achieved;

(3) to see if other definitions of volume are to be rejected, or if they can be used at the price of introducing some modifications.

The answer to (1) was already given in Ref. 5: the volume that best describes the experimental results is that of Voronoi. Here we shall reiterate and expand the arguments in favor of that statement.
As we shall see, the answer to (2) is that no reasonably defined volume is entirely intrinsic in nature, and that, therefore, experimental intrinsic compressibility is not either. In order to answer to (3), we shall be forced to dwell on the subject of probe radius for molecular surface, and we shall see that the concept of molecular surface and volume can be saved if one considers the probe radius as a pressure-dependent fitting parameter.

DEFINITIONS OF PROTEIN VOLUME

Apparent Volume

The easiest to measure is occupied volume, also called the apparent volume. This volume cannot be exactly related to the geometry of the protein. Rather, it is defined as the volume of the solution minus the volume of the solvent in the absence of solute. It is easy to establish that apparent volume is equal to partial volume in the limit of infinite dilution. The values of apparent volume reported in literature are usually extrapolated to infinite dilution, so that in fact they correspond to partial volume.

Geometric Definitions of Volume

Various geometric definitions of volume of a protein have been proposed. These geometrical volumes are microscopic observables, in the sense that their value is a function of the positions of all atoms of the protein. The most widely used definitions of the protein’s volume are the van der Waals (vdW), the molecular, the excluded, and the Voronoi volumes. They can be determined either from experimental data, i.e., from the average position of each atom of the protein as determined by x-ray and nmr spectroscopy, or from a simulation, using instantaneous configurations (frames) of a molecular dynamics trajectory.

Electron Density Volumes. One can define the volume of a molecule as the geometric volume of those regions of space where the electron density is greater than some preassigned value. While this definition looks very physical, it is extremely difficult to implement, and the volume thus defined depends on the cutoff we choose. Moreover, the only effect of a pressure change on this volume will be due to a small change of distances between atoms in contact with each other. A rearrangement leading to a better packing (decreasing the total volume of cavities and voids) will not be noticed.

A different, cutoff-independent definition of an electron-density volume follows from Bader’s separation of a molecule into atoms defined as basins of attraction of the ∇ρ field (where ρ is the charge density; see Ref. 7 for details). This volume shares many features of the Voronoi volume defined below, it is parameter independent and it reacts to packing rearrangement. There are essentially two drawbacks: first, one would need detailed information on the electron charge density for all atoms of a macromolecule, and second, even then, a numeric implementation of the method would be very computationally expensive due to the necessity to compute exactly the boundaries of the ∇ρ basins of attraction for each atom.

The van der Waals Volume. This is a reasonable approximation to the above definition. Any macromolecule can be seen as a collection of intersecting spheres of arbitrary size. If each atom is represented by a sphere of the atom’s van der Waals radius, the vDW volume (surface) of a molecule is defined as the volume (surface) of the body consisting of the intersecting set of such spheres. The main inconvenience of this volume is the same as that of the electron-density one: it does not take into account cavities and voids, and, therefore, the quality of packing of the protein’s atoms.

A word of caution should be given concerning the very concept of a vDW radius. Quoting from the work by Richards,1 “The bases on which the radii of the individual atoms are derived and the uses to which they are put are different. There may be no simple set of correct values.” One often relates the atomic vDW radii to the nonbonded potential energy function, whose parameters, for a given system, are not uniquely defined. Even within the same parameterization, at least two definitions of the vDW radius can be encountered in literature: the sum of the radii of a pair of atoms is equated either to the value of the interatomic separation at the minimum or to the smaller value where the potential is zero. This ambiguity results in a difference of about 12%. In general, vDW radii of same atoms accepted by different authors vary by about 10%. We shall show (Results and Discussion) that while such differences affect the volumes, they have a much smaller effect on the volumes’ behavior under a pressure change.

The values of a vDW radius obtained from the potential energy function should be practically pres-
sure independent. This stems from the fact that the electron densities and covalent bond lengths are not affected by pressure up to several GPa, and that the potential energy function is a phenomenological approximation to the quantum-mechanical energy obtained from electron density distributions. However, one may envisage a purely parametric definition of vdW radii, in which these are fit in such a way as to yield a “correct” value for a vdW radius-dependent volume (e.g., molecular volume). Evidently, one must first decide what one considers as the “correct” value. If one decides that the “correct” volume is that of Voronoi, one may want to reparameterize the vdW radii in such a way as to make the molecular and Voronoi volumes equal to each other. The vdW radii thus defined will obviously be pressure dependent. We do not think that one could gain much in adopting that policy. The topic will be discussed in some more detail in Results and Discussion.

Molecular Volume. The term molecular volume was introduced by Richards. It is the volume enclosed by the surface (molecular surface) defined by rolling a spherical probe on the vdW surface of the molecule. (Evidently, this surface and volume depend on the probe’s radius.) The molecular surface consists of two parts: the contact surface, which is the part of the vdW surface of each atom accessible to the probe, and the reentrant surface, consisting of the inward-facing parts of the probe sphere when it is simultaneously in contact with more than one atom. The probe radius commonly used ranges between 1.4 and 1.6 Å (see discussion below). It is clear that starting from some (small enough) radius of the probe, the protein surface thus defined becomes simply connected, so that there are no wormholes (narrow tunnels between the atoms going through the body of the protein). From this point on, one can suggest two different versions of the definition of the molecular volume: in one, one would simply take the volume contained inside this surface, while in the other, one will search for internal cavities and take their volume off the total. The latter definition is the one usually implemented in standard simulation and graphics software. In fact, the first definition, in which closed cavities are a part of the volume, allows one to see a stronger reaction to pressure changes. We used the second definition, but since there were no cavities in our molecules, this made no difference. Anyway, all interatomic voids and smaller cavities (unable to accommodate a probe sphere) are now included into the volume. Therefore, one can expect a much greater reaction to pressure-induced internal rearrangements than in the case of vdW volume. Some surface effects are also taken into account, but in a rather complicated way. The molecular surface reflects the instantaneous configuration of the protein, but not at all the instantaneous configuration of the solvent. The surface consists of accessible, rather than accessed points, so that it does not reflect the instantaneous or average number of solvent molecules in contact with the protein. This property would not be a defect per se if we tried to interpret molecular surface and volume as “intrinsic” quantities. What prevents them, however, from being seen as intrinsic, is their dependence on the radius of the probe.

Physically, it is not reasonable to think that water molecules hit a wall whenever they come to a distance of 1.4 Å + r_{vdW} from a protein atom whose vdW radius is r_{vdW}. In a simulation one can see a distribution of distances between water oxygens and nonhydrogen protein atoms with a lower cutoff of about 2.4 Å, which is significantly lower than the sum of vdW distances. Notice that this lower cutoff has very little pressure dependence, while the distribution of distances has a stronger dependence on pressure. We shall discuss this issue in more detail in Results and Discussion.

Excluded Volume. The solvent-excluded volume is defined as the volume that is inaccessible to the centers of solvent particles. Thus it differs from the molecular volume by the volume of a r_{probe} thick layer. The added term is grossly proportional to the molecular surface. However, this term is usually much smaller than the molecular volume, so that there using molecular or excluded volume does not result in a significant change in behavior.

Voronoi Volume. The Voronoi procedure consists of assigning a certain well-defined polyhedron to each point of a discrete set of points. The Voronoi polyhedron can be seen as a generalization of the Wigner-Seitz cell for a nonperiodic set of points. Voronoi polyhedra, like Wigner-Seitz cells, are primitive cells in the sense that they fill all space without overlapping or leaving voids. (In this sense the Voronoi volume can be seen as a very rough implementation of Bader’s electron-density construction.) In application to a protein in solution, the points in question are atoms, and the set can be thought as consisting of subsets—residues, domains, or the whole protein. The Voronoi volumes are additive, so...
the volume of a subset is the sum of volumes associated with each site of the subset, and the volume of the set is the sum of subset volumes.

There seems to be a problem: for some points of the set, the corresponding polyhedron can have an infinite volume and surface. This is a common problem for all methods that do not use artificial cutoffs, i.e., in which an atom’s boundary is the boundary between that atom and another one. One could intuitively think that the problem concerns only those atoms that are close to the surface of the molecule. However, in the case of the superoxide-dismutase dimer, only $\sim 650$ out of 2198 atoms do not give rise to infinite polyhedra.

In fact, this difficulty can be very easily overcome. For a protein in a solvent, even one layer of solvent molecules suffices to make finite all polyhedra corresponding to atoms of the protein.

In the Voronoi procedure, the surface accessible to the solvent (the protein Voronoi surface) is defined as the sum of the surfaces of those faces of the protein’s polyhedra that are shared with polyhedra corresponding to the solvent’s atoms. One can see that contrary to the molecular surface case, “accessible” means here “instantaneously accessed”: the Voronoi surface reflects instantaneous configuration of the solvent.

We would expect that the Voronoi volume reacts to pressure-induced internal (packing) rearrangements in the same way as the molecular volume with a reasonable probe radius. As for the surface effects, they are taken into consideration in a slightly different fashion: a part of the changes occurring in solvent packing get translated into a change in the protein Voronoi volume. Namely, the further away the nearest water neighbor of a protein atom, the larger this atom’s Voronoi volume. The question whether this effect should be considered as a part of “intrinsic” or “hydration” volume change is a matter of pure convenience: our task is to compare definitions, not to discuss their linguistic propriety. (See also discussion in Ref. 14.)

One other point of caution in what concerns the Voronoi volume is that the original Voronoi procedure (which we adopt) does not at all take into consideration individual properties of atoms. The volume of an atom’s Voronoi polyhedron is assigned in a purely geometrical way, whatever its chemical properties and the size of its electron cloud may be. To remedy this situation, a number of modifications have been suggested so as to position the dividing plane between two atoms according to preassigned atom radii (see, e.g., Refs. 1 and 15). This procedure, one of the most frequently used versions of which is called Richard’s method B, is an excellent algorithm for studying volumes (and their fluctuations) of individual atoms, because the original Voronoi procedure leads to an overestimation of the variance in the volume for any particular atom type. However, it was found in Ref. 14 that a hybrid method, in which method B is used for the protein’s interior and the original Voronoi procedure (bisection) to position the dividing plane involving water molecules, gives very good results. Since we are interested only in compression, and therefore in the volume of the whole molecule, rather than its individual atoms, and since the total volume of the protein’s interior does not change whether one uses method B or bisection, we have no reason to use a relatively more complicated method B.

We want to make it perfectly clear that the Voronoi volume thus computed need not have much to do with the electron-density, or any other “physical” volume of an atom. On the contrary, all cavities and gaps—i.e., regions of space with low electron density—are translated into Voronoi volume. This is why the Voronoi volume is an excellent tool for describing the quality of packing.

For the surface atoms we have the usual choice of treating hydrogens separately or including them into an all-atom construction. As we shall see, this choice affects the total volume, but has practically no effect on compressibility.

**COMPRESSIBILITY AND COMPRESSION**

To each volume, one can associate a compression and a compressibility. The compression $k_{P_2}^{P_1}$ is defined as the fractional change in volume caused by the pressure increasing from $P_1$ to $P_2$:

$$k_{P_2}^{P_1} = \frac{\langle V_{P_2} \rangle - \langle V_{P_1} \rangle}{\langle V_{P_1} \rangle}$$  \hspace{1cm} (1)

where $\langle V_P \rangle$ is the average volume at pressure $P$.

Compressibility $\beta(P)$ is defined as $-(\partial \ln V / \partial P)$ and is related to compression by

$$\beta(P) = -\lim_{\Delta P \to 0} \frac{\ln(\langle V_{P+\Delta P} \rangle / \langle V_P \rangle)}{\Delta P} = -\lim_{\Delta P \to 0} \ln(1 - k_{P_2}^{P_1+\Delta P}).$$ \hspace{1cm} (2)

**Intrinsic Compressibility**

*Experimental Intrinsic Compressibility.* Strangely enough, there seem to be more compressibilities...
than volumes. One of the objects most frequently encountered in literature is intrinsic compressibility.\textsuperscript{3,4,16} What volume does it correspond to? It is not at all clear that in a different context the words “intrinsic compressibility” mean the same thing. What is measured in, e.g., sound velocity experiments,\textsuperscript{3,16,17} is the adiabatic compressibility of the solution $\beta_s = (\rho u^2)^{-1}$ ($\rho$: density, $u$: velocity of sound). This gives an access to the adiabatic partial compressibility of the protein $\beta_P$, which one would then like to separate the intrinsic and hydration contributions: $\beta_P = \beta_i + \beta_h$. Skipping the details, two main lines of arguing have been advanced so far: the “regression” and the “additive” methods.\textsuperscript{4} In the first, one assumes a linear dependence of the hydration term on the specific surface (surface per unit volume) of the protein molecule, with a negative coefficient, and then obtains $\beta_i$ from a graph of $\beta_P$ as a function of the specific surface (or rather of $M^{-1/3}$, where $M$ is the protein’s mass) for various proteins: $\beta_i$ is the value of intercept of the straight line drawn through the points on the graph, and the y axis. The obtained “pseudoadiabatic” intrinsic compressibility, which in this case is defined averaged over the proteins figuring in the graph, can then be used to obtain the isothermal intrinsic compressibility.

The second, “additive” method uses the assumption that hydration contributions of protein’s atomic groups are the same for any substance including these groups, and that these contributions are additive. This method does give access to individual compressibilities of different proteins.

We are not going to analyze here the appropriateness of the hypotheses lying in the foundation of these two methods. For us, the main justification of the results lies in the good agreement between the two. In fact, the regression method yields the same average isothermal intrinsic compressibility, the actual configuration of the solvent, it is solvent of the results lies in the good agreement between As we have already mentioned, even though molecular (or excluded) volume does not depend on surface effects. The electron-density and vdW volumes are

Geometric Definitions. It is clear that a proper definition of any compressibility should be phrased in terms of volumes. Even though it seems reasonable to demand that the volume whose compression we call intrinsic does not implicate the solvent, it is not evident that one can produce such a definition for the entire protein. It is relatively easy to construct volumes so as to describe intrinsic compressibility of the core of a protein. One can propose several such definitions, all equivalent as long as we speak only about a protein’s interior. For example, one can think of a scale-dependent quantity $V(r)$ obtained by averaging volumes of all polyhedra whose vertices are atoms situated approximately at the distance $r$ from the center of mass of the molecule. This is a simpler construction than the Voronoi method. If, for a given protein, the compressibility $\beta(r)$ obtained from $V(r)$ does not depend on $r$, or stops depending on $r$ as $r$ grows, we can speak about a genuine intrinsic core compressibility of a protein. The problem with this definition is that it does not work for the entire protein, with its surface atoms: the surface does not lie approximately at the same distance from any point inside the protein, so that we cannot construct the same kind of definition for polyhedra whose vertices are surface atoms: different choices of surface atoms would result in drastically different volumes.

A uniform and isotropic compression could also be obtained from contraction of interatomic distances $d$: $\Delta V/V = 3\Delta d/d$. However, it was found that $\Delta d/d$ is scale dependent,\textsuperscript{5} which means that protein compression is anisotropic and/or nonuniform, and, therefore, cannot be obtained in this simple fashion [even if we introduce a scale-dependent quantity of $3\Delta d(r)/d(r)$, it would not directly correspond to contraction of a well-defined volume].

As we have already mentioned, even though molecular (or excluded) volume does not depend on the actual configuration of the solvent, it is solvent dependent via the dependence on the probe radius.

Finally, one can think of two volumes that share the merit of being completely solvent independent, but are very far from what one would reasonably call a protein’s volume: (a) the spherical volume with the radius equal to the gyration radius of the molecule, and (b) the volume of the inertia ellipsoid of the molecule. It is clear that we cannot use those two volumes when we pose questions concerning a particular protein.

We have seen that molecular, excluded, and Voronoi volumes depend to a various degree on surface effects. The electron-density and vdW volumes are
not suitable definitions of intrinsic volume because they do not convey the idea of a molecule’s volume as the volume inaccessible to solvent. The ‘‘truly intrinsic’’ definitions added in the previous paragraphs suffer from other drawbacks. Namely, $V(r)$ cannot be extended to the entire molecule, and the spherical and elliptical volumes we have just mentioned have very little to do with the volume of a concrete protein.

So rather than defining ‘‘truly intrinsic’’ quantities, we should better ask again the question as to which of the definable volumes gives rise to a compressibility whose behavior is closest to that experimentally measurable.

It was found that the Voronoi compressibility comes closest to the experimental intrinsic compressibility as defined above. Since the Voronoi construction is not entirely solvent independent, this may mean that the experimental intrinsic compressibility is not either. This should not be considered as a drawback per se, so long as our theoretical construction reproduces experimental results. Indeed, hydration-induced rearrangements of surface protein atoms are inseparable of the solvent rearrangement. We shall see, however, that this solvent dependence of the Voronoi volume is weak.

RESULTS AND DISCUSSION

A natural question that comes to mind is why the pressures used were so high, and why we do not observe protein denaturation at pressures that must normally lead to it. If the second question can be answered, the first one becomes trivial: the bigger the pressure jump, the lesser the error on compression. So should we observe denaturation at 1000 and 2000 MPa?

When high pressure is suddenly applied to a protein, the latter undergoes first elastic relaxation, and then conformational relaxation. In the first, elastic phase, all changes in the protein are of reversible nature. This can be explained by the existence of multiple energy minima (substates), the evidence for which has been amply discussed in the literature (e.g., Refs. 24 and 25). (In fact, the situation is slightly more complicated: the energy minima are organized hierarchically, so that the notions of state and substate become relative, but for our purposes it suffices to speak of only two hierarchical levels.) Upon a pressure jump, the protein is driven into a neighbor substate, belonging to the same general conformational state (in our case, the native state). This change involves a change in interatomic distances, but all secondary and tertiary structures remain intact. However, at sufficiently high pressure this substate, along with all other substates of the native state, becomes metastable, so that with time the protein falls into a different energy minimum, no longer corresponding to the native state. This second phase is called conformational relaxation, and is irreversible. It is very important for our discussion that the second phase takes many orders of magnitude longer time than the first one. The relaxation functions observed in complex systems are usually not exponential in time and cannot be characterized by a single rate coefficient. The typical times depend on the protein and temperature, but are generally on the order of $10^{-3}$–$10^{-4}$ s.

The simulations were 120 ps long. It was found that this time is sufficient for good convergence of all quantities measured along the dynamics trajectories. However, it is most certainly insufficient for observing a conformational relaxation onset. This means that the compressions and compressibilities we obtain from
for any meaningful definition of the latter, and that the intrinsic and hydrational contributions into apparent compressibility must be comparable in size. This means that using the Voronoi volume as intrinsic results in a good agreement with the expected hydrational effect, while using the molecular volume (like in Refs. 27 and 28) would result in the compressibility of hydration water lower than that of bulk water, in contradiction with experimental results on low weight molecular compounds. 4

Even though we have stressed that the Voronoi compressibility is not quite solvent independent, and that the experimental intrinsic compressibility it approximates may not be either, it would be interesting to compare their values with the already mentioned solvent-independent quantities resulting from the compression of (a) a spherical volume whose radius equals the gyration radius of the molecule, (b) the inertia ellipsoid of the molecule, (c) the C₆₁ distances. Rather than comparing compressibilities, we shall compare compressions. Computing these between 0.1 and 1000 MPa, we obtain, respectively, 12.5, 7.6, and 7.9% (see Table I), to be compared with the Voronoi result of 9.76% and the molecular result of 2.4%. One should not be looking here for more than an order-of-magnitude agreement or disagreement, since these three quantities cannot serve as more than very rough estimates of the intrinsic compression. However, we see that they agree much better with the Voronoi compression than with the molecular one. (Notice that all are larger than the apparent compression, as should be). Thus, the experimental intrinsic compressibility and the Voronoi volume compressibility describe a property whose nature is mostly, even though not entirely, intrinsic. This statement can be corroborated by the fact that the average Voronoi volume of a residue is practically independent of its position with respect to the solvent (buried, exposed), which is remarkable given that solvent plays a crucial role in computing the Voronoi volume.

All these results have been obtained for Voronoi volumes computed without hydrogens. What would be the effect of taking hydrogens into account? We performed a test computation of volumes, adding exactly the same hydrogens (i.e., polar) as in the molecular volume case. At all the three values of pressure, the effect is a net increase of the Voronoi volume by ≈ 6%, with no effect on compressibility within the uncertainty limits.

Volumes: Discussion

We see that while the Voronoi volume has a proper pressure behavior to describe experimental intrinsic

---

8The apparent volume is computed as the average volume of the solution minus the average per molecule volume of pure SPC water multiplied by the number of water molecules in the solution. Therefore, it contains a contribution due to counterions. The volume of pure SPC water was computed from a series of simulations of 686 water molecules in the same thermodynamic conditions as the solution.
compressibility, the molecular volume is a rather poor candidate for the same purpose. This means that attempts at comparing results concerning experimental intrinsic compressibility with calculations of compressibility of molecular volume \(^{27,28}\) are not quite appropriate. Contrary to the opinion of the authors of these papers, molecular volume has little to do with experimental intrinsic compressibility. And as we have seen, this means that it has little to do with any kind of intrinsic compressibility.

Why is that so? Let us take a more detailed look at the volumes we have defined.

We have already said that we would not like to study the volume of the space occupied by the protein matter: this is the electron-density volume, and it is not a very interesting quantity as regards its reaction to pressure. The same applies to the vdW volume. The apparent volume, though easiest to measure, sweeps under the carpet the most interesting hydration-related effects. The only alternative left is to use the volume that the solvent fails to occupy. Both the molecular and the Voronoi volumes can be considered as candidates.

The Voronoi volume is a very simple construction. It attributes a space to each atom, and the closer its neighbors, the smaller is this space. This applies to surface atoms, which have solvent molecules among its neighbors as well. So the closer water molecules come to the protein, the smaller its volume. This property conveys the idea that the boundary between the macromolecule and the solvent depends on the geometry of both. It also conveys the intuitive feeling that when a water molecule is “pressed into” the protein, the space occupied by the latter gets smaller. (In other words, smaller volume reflects tighter packing \(^{14}\).

Not so with the molecular volume. It does not depend on the actual configuration of the solvent. It assumes that the solvent occupies all space available to hard spheres with a given radius—to be more precise, swept by all imaginable configurations of such hard spheres. But in fact water does not behave as a collection of hard spheres. At any pressure, there is a distribution of distances between water molecules and their nearest nonhydrogen protein neighbors. It was found by simulation \(^{29}\) that this distribution is pressure dependent, with an almost pressure-independent lower cutoff of about 2.4 Å. As pressure increases, the number of water–protein hydrogen bonds goes up. This means that not only does the hydration layer’s structure undergo significant changes at high pressure, but also that water comes closer to the protein. This should be reflected in a decrease of the protein’s volume, but the very construction of the molecular volume prevents it from directly reacting to this water–protein boundary shift. Why then, nevertheless, does the molecular volume go down as pressure goes up? Because whether or not we take into account surface changes, the interior gets always compressed. However, molecular volume is inappropriate to describe even that effect, because it does take into account changes in the surface geometry of the protein; but in a smoothed-out, probe-radius-dependent, not quite coherent way.

### Table 1 Cell and SOD Dimer Compressions, According to Different Definitions of Protein Volume, Computed as Averages Over Molecular Dynamics Trajectories

<table>
<thead>
<tr>
<th>Volume Type</th>
<th>(k_1^{1000}) (%)</th>
<th>(k_2^{2000}) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>vdW volume</td>
<td>0.40 ± 0.01</td>
<td>0.39 ± 0.01</td>
</tr>
<tr>
<td>Molecular volume ((r = 1.4 \text{ Å}))</td>
<td>2.4 ± 0.2</td>
<td>2.1 ± 0.2</td>
</tr>
<tr>
<td>Voronoi volume</td>
<td>9.76 ± 0.07</td>
<td>4.77 ± 0.06</td>
</tr>
<tr>
<td>Apparent volume</td>
<td>5.6 ± 0.3</td>
<td>3.7 ± 0.2</td>
</tr>
<tr>
<td>Gyration radius sphere</td>
<td>12.5</td>
<td>6.0</td>
</tr>
<tr>
<td>Inertial ellipsoid</td>
<td>7.6</td>
<td>5.0</td>
</tr>
<tr>
<td>(C_a-C_a) distances result</td>
<td>7.9</td>
<td>5.1</td>
</tr>
</tbody>
</table>

### Redefining Molecular Volume

Do we mean to say that the molecular volume should not be used at all? This would be very inconvenient, given that most existing software compute exactly that volume. New simulation techniques based on continuous electrostatic models for the solvent (see, e.g., Ref. 30) also require volumes that could be computed without considering the solvent microscopically. This calls for a redefinition of the molecular volume so as to correct its pressure dependence.
Using the fact that the molecular volume depends on the probe radius, we can find at which probe radii, at various pressures, it coincides with the Voronoi volume.

The most commonly used probe radius is 1.4 Å. The idea of representing water as a sphere of 1.4 Å radius comes from the observation that the oxygen–oxygen nearest neighbor distance in ice (at atmospheric pressure) is 2.8 Å. Even though ice and water have very little in common with densely packed hard spheres, this figure also fits in the general concept of almost spherical water molecules with a radius slightly larger than the vdW radius of oxygen (1.35 Å). This would be an excellent argument if we took seriously the concept of the vdW surface of a macromolecule. What does not allow us to do so is the simple fact we have already mentioned that water oxygens come as close as 2.4 Å to nonhydrogen protein atoms, and that the typical distance between hydrogen-bonded water oxygens and peptide nitrogens is 2.8 Å. The commonly accepted vdW radii of peptide atoms are: 1.65 Å for nitrogen, 1.6–1.7 Å for oxygen, and 1.8–1.9 Å for carbon. This means that water is not distributed around the protein as if the latter were surrounded by a vdW wall. In fact, the whole concept of molecular surface does not seem to correspond to any physical reality.

One possible way to “save” the molecular volume concept is to consider the probe radius as a parameter that can be adjusted so as to obtain the same value for the molecular volume as for the Voronoi volume, well knowing that these two volumes do not have the same shape.

A simple examination of Figure 2 tells us that in order to satisfy $V_{mol} = V_{Voronoi}$ at atmospheric pressure, the probe radius must be about 1.5–1.6 Å. At 1000 MPa, it should be 1.0–1.1 Å.

Evidently, the last result does not mean that the probe sphere shrinks so as to become smaller than a water molecule. A smaller radius simply expresses the fact that at higher pressure, more water molecules come closer to protein atoms, and that neither behave as hard spheres. (In this regard, see also the discussion of accessibility in Ref. 31; a rather complete discussion of accessibility-related and some other aspects of the vdW and molecular volumes can be found in Ref. 32. A very good discussion of water radius can be found in Ref. 14.)

In this way one can still use $V_{mol}$ adjusting the probe radius according to pressure.

Another set of parameters that could in principle be used for the same purpose of making molecular volume artificially equal to Voronoi volume at all pressures is the set of vdW radii of the protein’s atoms. We have already examined the concept of a vdW radius in the section on the vdW volume. As we have already mentioned in the section dedicated to the vdW volume, physically vdW radii should not depend on pressure. As long as we use pressure-independent vdW radii, the ambiguity in assigning a vdW radius to an atom does not result in a noticeable change of the dependence of the molecular volume on pressure. In Figure 3 we show how the
Table II  Roughness Computed as a Ratio of the Radius of a Sphere with the Same Surface Area as that of the Protein, to the Radius of a Sphere with the Same Volume as that of the Protein, for the vdW, Molecular, and Voronoi Surfaces and Volumes*

<table>
<thead>
<tr>
<th>Pressure</th>
<th>0.1</th>
<th>1000</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPa</td>
<td>MPa</td>
<td>MPa</td>
</tr>
<tr>
<td>vdW</td>
<td>2.477</td>
<td>2.437</td>
<td></td>
</tr>
<tr>
<td>Molecular (r = 0.7)</td>
<td>1.955</td>
<td>1.847</td>
<td></td>
</tr>
<tr>
<td>Molecular (r = 1.4)</td>
<td>1.580</td>
<td>1.543</td>
<td>1.526</td>
</tr>
<tr>
<td>Molecular (r = 2.1)</td>
<td>1.411</td>
<td>1.395</td>
<td></td>
</tr>
<tr>
<td>Voronoi</td>
<td>1.646</td>
<td>1.684</td>
<td>1.691</td>
</tr>
</tbody>
</table>

*Roughness of an ideal sphere is 1.

protein’s molecular surface and volume change if one uniformly modifies all vdW radii of protein atoms. One can see that as the vdW radius grows, the molecular surface goes down (it becomes smoother) and the volume goes up; however, looking at curbs corresponding to two different pressures, one can see that the molecular compressibility depends very little on the vdW radius change. Certainly, one could introduce an artificial pressure dependence of the vdW radii, rather than the probe radius, so as to achieve the same result as above \( V_{\text{Voronoi}}(p) = V_{\text{mol}}(p) \). We do not see much sense in doing so, because it is usually much simpler to change the probe radius (an adjustable parameter in most existing computational programs). Also, while adjusting the probe radius does reflect some physics (water molecules coming closer to protein), adjusting the vdW radii does not (the minimum distance between neighbor atoms does not depend on pressure).

**Roughness**

In this context, it is interesting to look at roughness. Roughness can be defined as \( \rho = R_s/R_e \), where \( R_e \) (\( R_s \)) is the radius of a sphere whose surface area (respectively, volume) equals that of the protein.

Evidently, one would like roughness to be related to the effective exposure of the protein to the solvent.\(^1^0\) However, it clearly depends on the definition of the volume and surface. For the molecular volume and surface, roughness goes down as the probe radius goes up, so as to reach a value \( \approx 1 \) (which depends on the global shape of the protein) in the limit of an infinite probe radius.

Table II gives the Voronoi and molecular roughness at different pressures. We can see that as pressure goes up, so does the Voronoi roughness, while the molecular roughness goes down. We suggest that the Voronoi roughness behavior is more physically appropriate. Roughness is defined through surface area, which is not a measurable quantity, so one could think that the choice of the type of roughness is a matter of taste. However, even though surface area of a protein cannot be measured, the concept of surface has been the subject of active discussion in the literature (e.g., Refs. \( 10 \) and \( 33 \)). Among more sophisticated definitions, we can cite Ref. \( 33 \), where it was proposed that a natural boundary of a protein in a solvent is the second shell of water molecules around it. All geometrically defined volumes give rise to corresponding surfaces. The very principle of the Voronoi construction (the fact that the volume envelope reflects the accessed, rather than accessible, surface) makes the Voronoi surface a reasonable a priori candidate. The correct pressure dependence of the Voronoi volume makes this hypothesis even more plausible.

Physically, as pressure increases, so does the energy of protein–solvent interaction (which is partly expressed in the increase of the number of hydrogen bonds). One would expect that this effect corresponds to an increase of the interface area. So, without being able to really prove it, we are tempted to say that the Voronoi roughness is a correct construction.

One should be warned against trying to interpret the increase of the Voronoi roughness as a sign of denaturation onset. Even though it seems plausible that the passage to denaturation phase at higher pressure is due to the increase of the relative importance of the protein–solvent interaction, there is no denaturation during simulations on the 100 ps scale. Denaturation onset should be marked by cracks and fissures capable of accommodating a water molecule. Such surface defects would result in an increase of both the Voronoi and molecular roughness. In the light of our discussion of the probe radius, it is possible that molecular roughness reacts to the denaturation onset with a certain retardation; however, this happens most certainly far beyond the scale of the simulations we use.

**CONCLUSIONS**

We have shown that much care is necessary when dealing with various definitions of the volume of proteins (and macromolecules in general). The of-
ten used ‘‘molecular volume’’ of a protein is a construction that is instantaneous for the protein, but statistical for the solvent. The ‘‘molecular surface’’ consists of points that could be points of contact with hard-sphere like water molecules. It does not consist of points that are in contact with those spheres at any particular moment. Therefore, the molecular construction does not directly reflect the packing of water around the protein. At the same time, it depends on a would-be solvent-related parameter: the probe radius. Moreover, the very idea that lies behind the ‘‘molecular’’ construction, according to which the space available to solvent molecules is bounded by the vdW surface of the macro-molecule, does not quite correspond to physical reality, as can be seen from the distribution of distances between the water and protein atoms: atoms do come closer to each other than the sum of their empirical vdW radii. It is not strange, therefore, that the pressure dependence of the molecular volume fails to correctly reflect the experimental data, i.e., the intrinsic compressibility as measured in the sound velocity experiments. Moreover, the compressibility obtained from the molecular volume is lower than the apparent compressibility, which is known (at least at the room temperature) to include a negative contribution due to hydronal effects.1 All this means that the pressure dependence of the molecular volume does not properly reflect actual changes in density and intramolecular distances.

In this regard, the Voronoi volume works much better. Rather than being a rough polyhedron approximation of the molecular surface as stated in Ref. 34, the Voronoi volume, by definition, reflects the actual positions of both the protein’s and the solvent’s atoms, which results in a much more reasonable definition of the protein–solvent interface. Moreover, the Voronoi volume is best for studying the relative packing efficiency of protein substructures, whether they are or are not in contact with the solvent, and, therefore, pressure-induced local packing modifications. This idea is corroborated by studying the pressure dependence of the Voronoi and molecular volumes of a protein. The correct behavior of the Voronoi volume under a pressure change is demonstrated by a good agreement between the Voronoi compressibility and the experimental intrinsic compressibility. The increase of the Voronoi roughness with pressure agrees with the change of the water–protein distance distribution and with the increase of the number of water–protein hydrogen bonds found in simulation. The roughness increase suggests that as pressure goes up, small crevices and packing defects appear on the protein surface. These crevices, ignored by a 1.4 Å probe commonly used in the ‘‘molecular’’ construction, facilitate nevertheless water–protein interactions at higher pressure. Apparently, despite its very complicated and somewhat hypothesis-dependent definition, the experimental intrinsic compressibility actually mostly reflects the compressibility of the protein’s interior (even though some dependence on surface-related effects, present in the Voronoi compressibility, cannot be excluded). This can be seen from the general agreement between this quantity, the Voronoi compressibility, and various estimates of intrinsic compressibilities defined in different ways.

REFERENCES