Review
High pressure simulations of biomolecules
Emanuele Paci *

Laboratoire de Chimie Biophysique, ISIS, Université Louis Pasteur, 4 rue Blaise Pascal, 67000 Strasbourg, France
Received 12 September 2001; accepted 18 October 2001

Abstract

Pressure is a thermodynamic variable which is particularly suitable for exploration of the properties of biological macromolecules. For proteins, in particular, denaturation induced by pressure is different from that induced by temperature or denaturants. The response of proteins to pressure changes can provide information on properties of their native and non-native states. This review focuses on molecular dynamics studies of the effect of pressure on detailed atomic models of proteins. It also reports on other theoretical approaches, such as Monte Carlo simulations, which have been used to study simplified models. Another purpose of this review is to try to point out potential future studies that may be both interesting and feasible, with constantly increasing computing power. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Protein; Volume; High pressure; Compressibility; Denaturation; Non-native state

1. Introduction

Exploring phase diagrams of biological systems, even outside biologically relevant regions, can provide useful information. For proteins, it is particularly interesting to explore those properties which make these molecules very peculiar heteropolymers; the most interesting property is that of possessing a unique stable and functional conformation in a specific range of temperatures and pressures and to be disordered and biologically inactive outside this range.

The present review deals primarily with theoretical approaches and molecular simulations as tools to understand how proteins react to a change in external pressure. A molecular simulation is a ‘numerical experiment’ in which the system being modeled is represented as a set of classical particles interacting through an empirical energy function, and the thermodynamic and kinetic properties are computed from the simulated time evolution of that system. The principal interest of molecular simulations is a very high resolution in time (1 fs) and space (0.01 nm).

Proteins are complex molecules. At or close to physiological conditions, transitions from native to non-native states and vice versa take place on the time scale of microseconds to seconds. Considering the number of degrees of freedom involved and the astronomical number of possible conformations, this time scale is quite short. Yet, it is far too large to be simulated, at least for realistic, detailed atomic models of proteins. Spontaneous folding and unfolding that take place in a test tube upon a change in the denaturant concentration, temperature, or pressure cannot be studied using molecular dynamics. Therefore, only limited regions of the conformational

* Present address: Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.
Fax: +41-1-635-5712.
E-mail address: paci@bioc.unizh.ch (E. Paci).
space close to the initial conformation (e.g. the native state) can be explored in a statistically significant way, while the whole conformational space remains too big to be explored during the simulation time. Such regions are confined by kinetic barriers which prevent the exploration of other conformations (e.g. various denatured states). Within such a limited region the system can be assumed to be approximately ergodic.

For the purpose of understanding the effect of pressure on proteins, the folding landscape of a protein can be schematically represented as in Fig. 1. The ordinate represents the free energy and the abscissa some arbitrary reaction coordinate. At physiological temperature and pressure the free energy has an absolute minimum which corresponds to the native state (N in Fig. 1).

The main effect of a moderate pressure [1] is to shift the position of the absolute minimum of the free energy (N′ in Fig. 1). The stable state of the system is mildly affected by the external pressure in such a way that the protein and the surrounding solvent occupy a smaller volume. The structural response of the protein–solvent system to moderate pressure can be called ‘elastic’ since it does not involve a transition from native to non-native state (there is no denaturation of the protein). From a structural point of view, elastic relaxation is characterized by a moderate change in interatomic distances, reversible when the pressure is released, but not necessarily homogeneous throughout the protein structure.

At pressure exceeding 200–500 MPa1 [2,3], the native minimum shifts further from its original position (N″ in Fig. 1), but a more interesting effect is that N″ is no longer the absolute minimum. A different absolute minimum appears (D in Fig. 1), corresponding to a set of completely different conformations (the high pressure-denatured state). The denaturing effect of pressure is quite different from that of temperature. In some cases, pressure plays a role opposite to that of temperature: it has been observed in several proteins that a moderate pressure stabilizes again heat denaturation [3]. Also, experimental studies [4,5] have shown that pressure-induced unfolding at room temperature is considerably slower than temperature-induced unfolding.

Simulations, even when performed at pressures large enough to destabilize the native state, usually show only elastic relaxation to a relative minimum (N″ in Fig. 1). This state (or set of states) is usually metastable. The time scales that can be currently explored by molecular dynamics simulations are too short to allow for a full relaxation from a metastable native-like minimum to the denatured state (D in Fig. 1).

To simulate transitions between states and to explore non-native states, various physical or non-physical perturbations or very high temperature were used, which either drive the system across barriers, or flatten those barriers, or both. For example, thermal unfolding in water of hen egg white lysozyme was simulated at 500 K [6], that of bovine pancreatic trypsin inhibitor (BPTI) at 423 K and

---

1 Atmospheric pressure corresponds to about 0.1 MPa = 1 bar. The Pascal (Pa) is the SI unit for pressure, 1 Pa = 1 N/m². The pressure below 10 000 m of water (as in the deepest part of the oceans) is about 100 MPa.
498 K [7], and that of the protein barnase at 498 K and 600 K [8,9]. Other perturbations, mostly involving application of non-physical forces, have been used to unfold proteins or to study conformational preferences at non-native conditions (see e.g. [10–12]).

Since high pressure does not result in a sufficient decrease of energy barriers and, moreover, slows down the kinetics, it is clearly not a suitable perturbation to unfold proteins on the simulation time scale. What high pressure simulations allow is the study of the elastic relaxation after a pressure jump, i.e. elastic properties of both native and non-native states. Naturally, the best studied cases are native states, for which X-ray and NMR experimental techniques provide structures which are used as initial configurations in molecular dynamics simulations.

Non-native states are less well characterized, both because of their structural heterogeneity and because they have been considered of less interest from a biological perspective. However, recent experimental results have underlined the interest of some non-native states like molten globules and other folding intermediates: their elastic response to a pressure perturbation has been shown to be a good way to detect them [13] and also to understand how they differ from the native state.

2. Denaturation of proteins under pressure

According to the Le Châtelier principle [14], the effect of a large, denaturing pressure on a protein is determined by the change in solution volume before and after the unfolding; i.e. at high pressure and normal temperature a protein denatures because the denatured state occupies a smaller volume than the native one. This behavior can be contrasted with what happens at normal pressure (below 100–200 MPa), at which denatured proteins occupy a larger volume than native states.

The ‘hydrophobic model’ [15], i.e. the model of a folded protein stabilized by hydrophobic interactions, succeeded in explaining a large amount of experimental calorimetric data on protein denaturation. However, as Kauzmann [16] pointed out, the hydrophobic model apparently contradicts the experimentally observed high pressure protein denaturation. Indeed, liquid hydrocarbons have an opposite behavior, i.e. a transfer from non-polar to aqueous environment is accompanied by a volume decrease at atmospheric pressure and by a volume increase at higher pressures.

Theory and simulation have been used to attempt to explain the apparent failure of the hydrophobic model to explain the pressure-induced unfolding of proteins, by focusing on the pressure dependence of hydrophobic interactions [17,18]. Monte Carlo simulations [17] have been used to study the effect of high pressure on the potential of mean force between two simple hydrophobic solutes. The expression of the free energy of association in terms of partial volume and isothermal compressibility shows that the change in volume favors solute association, but the change in compressibility favors the dissociation. This means that the application of pressure can disperse the hydrophobic core of a protein in water, not because the solution containing the unfolded protein occupies a smaller volume than that containing the native state, but because it is more compressible.

Hummer and coworkers [18] suggested a solution of the pressure denaturation puzzle by considering the transfer of water into the hydrophobic core as opposed to the transfer of non-polar residues into water. Using a wide range of theoretical and simulation approaches, they computed the potential of mean force between two methane-like particles in water. Their results led to the conclusion that pressure destabilizes the configuration where two non-polar groups are in contact relative to that where a water molecule is in between the two non-polar groups (i.e. at higher pressure it is less profitable for non-polar groups to be in direct contact). This suggests that pressure denaturation is accompanied by penetration of water molecules inside the protein, rather than by a transfer of the hydrophobic interior into water, as happens in high temperature denaturation.

Using Monte Carlo simulations, Hillson et al. [19] investigated the free energy profiles and the folding kinetics of an off-lattice minimalist model of a protein. Their model included the pressure dependence of the potential of mean force between two hydrophobic solutes found by Hummer et al. [18]. It was shown that the presence of a desolvation barrier in-
creasing with pressure results in an increase of the roughness of the energy landscape; as a consequence, folding time increases even if the intrinsic thermodynamic free energy barriers get lower. These results agree well with the experimental findings \cite{4,5,20} that high pressure denatures proteins by slowing down the rate of folding more than that of unfolding.

Despite the extreme interest of theories and models attempting to explain the hydrophobic effect, it must be stressed that these are based on using simple non-polar spherical objects. Applying conclusions from these studies to real systems such as proteins would ignore the fact that connectivity, polar interactions, hydrogen bonds and specific packing all play an essential role in determining the stability of the native state.

3. Molecular dynamics at high pressure

3.1. Methods for high pressure simulations

Rigorous statistical mechanical studies of properties of complex molecular systems, at given pressure and temperature, have been made possible by a series of important theoretical and methodological advances that followed the seminal work of Andersen in 1980 \cite{21}. In molecular dynamics, the equations of motion of a collection of particles in a fixed volume are integrated numerically. Energy, volume, and number of particles are conserved, and time averages equal microcanonical ensemble averages (if the system is ergodic, which can be assumed to be approximately the case at least in a limited region of the phase space). Andersen \cite{21} introduced the so-called extended Lagrangian method where a fictitious Lagrangian with one extra degree of freedom is introduced: in the specific case addressed by Andersen, the volume in which the particles are confined becomes a dynamical degree of freedom coupled to the particles' positions. The volume varies with time so that the internal pressure equals, on the average, the external fixed pressure. The trajectory deriving from the extended Lagrangian samples an ensemble different from microcanonical: in the original case studied by Andersen, it is the iso-enthalpic–isobaric ensemble (constant enthalpy and pressure). This formalism was extended to obtain the canonical (constant temperature, or NVT) and isobaric–isothermal (constant temperature and pressure, or NPT) ensembles by Nose \cite{22} and Hoover \cite{23}.

Other methods have also been devised to control the temperature and the pressure of the system. The most popular, and still widely used, is the weak coupling method of Berendsen \cite{24}, which, however, does not sample any standard statistical mechanical ensemble. The most serious consequence is that statistical mechanical relationships between fluctuations (of the energy, of the volume, etc.) and second derivatives of the thermodynamic potential (specific heat, compressibility, etc.) do not hold. The extended Lagrangian and weak coupling methods have been thoroughly compared for proteins \cite{25}; the convergence of the total volume and of its fluctuations (related to the system compressibility) and the statistical efficiency of the two methods were studied. The most interesting finding was that observables such as volume and enthalpy obtained by extended Lagrangian and weak coupling simulations are within statistical error of each other. The same agreement was found for the compressibility and specific heat computed with finite differences by performing simulations at different pressures. Thermodynamic derivatives from fluctuation formulas in the NPT ensemble, which can only be evaluated for an extended Lagrangian simulation, agree within statistical error with results obtained from finite differences.

An improved extended Lagrangian method for performing constant pressure molecular dynamics simulations has been proposed \cite{26}. The original equation of motion for the volume variable is replaced by a Langevin equation, thus eliminating the observed non-physical 'ringing' of the volume variable \cite{27} which might induce problems of convergence.

A change in volume implies a change in the atomic positions; when dealing with molecules, one has the choice of coupling the volume change to each atomic degree of freedom or to each molecular center of mass \cite{28}. In the former case, the dynamical adjustment of the volume corresponds to a space scaling that affects the position of each atom. In the latter case, the scaling occurs only for the positions of the molecular centers of mass, thus affecting only inter-molecular distances. An intermediate scaling scheme
has also been proposed [29] in which the centers of mass of arbitrary subsets of atoms are rescaled. This scheme has the molecular and atomic scaling as limiting cases. If the system is ergodic and the molecules do not dissociate, all the above scaling procedures have been demonstrated to be equivalent for flexible molecules [30]. In principle, different ways of coupling volume and coordinates should not affect the results of a constant pressure simulation, but rather differ in how efficiently the system relaxes and equilibrates at a certain pressure. In practice, since simulations are not exactly ergodic, the way in which the pressure is applied may affect the results. Zhu and Averbak [31] presented a modification of the Andersen method which is particularly suitable for non-homogeneous systems such as solutions containing biological macromolecules.

Molecular dynamics simulations of a large protein in explicit water with periodic boundary conditions remain extremely demanding in terms of computation time. Other approaches have been proposed where the protein is immersed in a droplet whose boundaries adjust to keep pressure constant. One such approach is the solvent boundary potential developed by Beglov and Roux [32] which accounts for the van der Waals and electrostatic interactions of an infinite bulk system surrounding simulated solute and water molecules in a sphere of variable radius; the method was improved [33] to simulate non-spherical systems by using a half-harmonic restraining potential to prevent water molecules from escaping. More recently, Louamas et al. [34] proposed an alternative method which accounts for the dielectric properties of the bulk solvent by subjecting water molecules to forces approximating van der Waals and dipole-dipole interactions with the implicit surrounding bulk solvent. However, it has not been proved that these methods sample the NPT ensemble.

3.1.1. Are current force fields appropriate for high pressure simulations?

Current force fields used to simulate proteins are designed to reproduce specific properties at atmospheric pressure and room temperature. Marchi and Akasaka [35] compared two different all-atom force fields, AMBER [36] and CHARMM [37]. Structural changes occurring in hydrated BPTI at high pressure (500 MPa) were studied by performing 1 ns long constant pressure simulations. Comparing the simulation results with chemical shift measurements performed up to 200 MPa pressure [38,39] they concluded that the accuracy of available force fields is not sufficient to quantitatively study structural changes occurring in biomolecules. However, another possible reason for this disagreement is that the experimental result might contain large error bars due to a very small change in chemical shifts observed between the atmospheric and 200 MPa pressure.

Similarly, common water models were elaborated to reproduce bulk water properties at normal pressure and temperature, not at high ones. For one particular model (SPC/E [40]), a detailed study of the pressure-dependent properties [41] suggests that some experimental properties of 'real' water, like the diffusion maximum at 150 MPa (1.5 kbar), are not reproduced. For another water model (TIP3P [42]) it has been observed (Paci, unpublished results) that density and compressibility strongly depend on how interactions are truncated. Simulations of systems solvated in water would gain in reliability if the potential energy functions for water were transferable to a broader pressure and temperature range.

3.2. Sub-nanosecond simulations at high pressure

The first molecular dynamics simulation of a solvated protein was presented less than 10 years ago: Kitchen et al. [43] performed a simulation of BPTI in solution at a pressure of 1000 MPa, using the AMBER energy function [44] for the protein and the SPC [45] for water and the weak coupling constant pressure technique. The structural and energetic properties of the solvated protein at atmospheric and high pressure where compared along two relatively short (100 ps) simulations. No evidence of pressure-induced unfolding was observed during the high pressure simulation.

A comparison of various attempts to unfold proteins during molecular dynamics simulations using a variety of perturbations, including high pressure, has been presented by van Gunsteren and collaborators [10,46] (using the GROMOS force field [47] for the protein and the SPC model for the water, and the weak coupling method to keep pressure constant [24]). A pressure of 1000 MPa applied for a few
hundred picoseconds affects the protein only marginally. Not even the onset of unfolding is detected on such a short time scale. On the other hand, the authors also observe that alternative methods, which lead to unfolding on a sub-nanosecond time scale, usually strongly bias the results; i.e. different sequences of unfolding events are observed when different perturbations are used.

Using the same force field and the same method to keep pressure constant, Wroblowski et al. [48] performed a molecular dynamics simulation of BPTI in water. Starting from atmospheric pressure, they applied pressure jumps, at 1000, 1500 and 2000 MPa, every 200 ps. The protein remained globular, but the authors noticed a conformational change that was larger than that observed in previous simulations [10,43,49]. The authors claim that this change, which mainly involves disruption of secondary structure and exposure of the hydrophobic core to water, is in good agreement with the experimental results of Goossens et al. [50] (that show denaturation of BPTI between 800 and 1400 MPa). However, other results [4] clearly show that such high pressure unfolding cannot occur on the time scale used. Another piece of evidence that simulation was too short to effectively show high pressure unfolding is that the root mean square deviation (RMSD) was never larger than 2.5 Å (i.e. the protein is very native-like even at the end of the 2000 MPa simulation); also, the RMSD was steadily increasing with time, suggesting that the equilibrium was not reached.

Paci and Marchi [51] reported a simulation of a unit cell of lysozyme containing eight proteins [52] and of superoxide dismutase (SOD) in solution at atmospheric, 1000, and 2000 MPa pressure. They used the CHARMM force field [53] for the protein and TIP3P for water [42] and a rigorous statistical mechanical method for sampling the isothermal-isobaric ensemble (see Section 3.1); this paper mostly focused on the change in the apparent and intrinsic protein volume following elastic relaxation discussed in Section 4.

Floriano et al. [54] performed a simulation of metmyoglobin up to 1200 MPa, using the AMBER force field and NPT ensemble. Instead of considering water explicitly, as in all simulations listed above in this section, they treated it as a continuum characterized by a pressure-dependent dielectric constant estimated using the Bradley and Pitzer equation [55]. They observe full unfolding passing through a molten globule-like state at about 700 MPa. It is not clear whether this result is due to the strong approximations in the solvent energy function or else to the absence of a viscous solvent. The fact that the latter increases with pressure is one of the causes of the slow kinetics of the pressure-induced unfolding both in experiments and in simulations of explicitly solvated proteins.

In the following of this section the most interesting properties observed in the high pressure MD simulations performed so far will be discussed.

3.2.1. Energetics

Simulations of solvated proteins generally show that at high pressure protein-solvent interaction is energetically more favorable than at atmospheric pressure, while the protein-protein bonded interaction energy increases (due to the moderate deformation of the protein’s bond distances and angles). The more favorable interaction between protein and solvent at high pressure is accompanied by an increase in the number of protein-solvent hydrogen bonds.

3.2.2. Hydrogen bonds

At high pressure (1000 MPa), the number of protein-to-protein hydrogen bonds was found [56] to increase by 2% for SOD and by 6% for lysozyme, while those between water and protein increase more (19% and 14%, respectively). In a simulation [43] of BPTI in solution a similar trend for both protein–protein and protein–solvent hydrogen bonds after a similar pressure jump was observed. Another simulation of the same protein (using a different force field) gave a similar increase in protein–solvent hydrogen bonds but a decrease in the protein–protein hydrogen bonds. On the other hand, high pressure crystallography [52] on lysozyme shows a decrease of 7% in the number of water-to-protein hydrogen bonds. This result is, however, not directly comparable with the simulation since only crystallographic water molecules were included in the estimate. The number of hydrogen bonds between solvent water molecules increases by 14 and 11% in the presence of SOD and lysozyme, respectively, while in pure water the increase of hydrogen bonds is larger (19%) [56]. The considerable increase in the number
of protein-to-solvent hydrogen bonds at high pressure is related not only to a higher density of donors and acceptors due to the larger compressibility of the water relative to the protein, but rather to a clear preference of water to make hydrogen bonds with the protein rather than water itself.

Despite the changes in the number of hydrogen bonds due to the increase in pressure, the average distance between acceptor and donor remains very close to 2.8–2.9 Å at all the pressures studied [35,43,51,56]. This is apparently in contradiction with the results of Akasaka and collaborators [57] which show (in BPTI) that hydrogen bonds formed by amide groups either with carbonyls or with water are shortened by pressure.

3.2.3. Atomic displacement and radius of gyration

In all the high pressure simulations reported so far the displacement of atomic positions with pressure, as measured by the RMSD from the experimental native structure, is moderate, confirming that the high pressure is not denaturing the protein over this time scale. As expected, the radius of gyration always decreases at high pressure, reflecting the increased compactness of the high pressure state. However, the variations of the radius of gyration do not provide a reliable estimate of the compressibility of the protein interior [58].

3.2.4. Atomic mobility

Most simulations of proteins at high pressure show that positional fluctuations decrease with increasing pressure, which is consistent with the intuitive argument that increasing pressure would have a damping effect on atomic fluctuations, analogous to that produced by a decreasing temperature. Brunne and van Gunsteren [49] observed the decrease of the positional fluctuations of atoms with increasing pressure; this effect was more important for side chains than for the main chain atoms, while the mobility of the backbone $\phi$- and $\psi$-angles was not affected. The amplitude of oscillation of side chain dihedral angles remains nearly the same for both pressures, but the rate of torsional angle transitions decreases on the average by 30% when the hydrostatic pressure increases to 5000 atm (about 500 MPa).

3.2.5. Structure modifications

Simulation allows the precise evaluation of the structural modifications occurring upon elastic relaxation after an increase of pressure. Experimentally, only the structure of lysozyme [52] has been determined at moderately high (100 MPa) pressure. When structures are available, the behavior of distinct regions within the protein under pressure changes can be singled out by computing the change in the interatomic distances $d_{ij}$. Simulation of the lysozyme crystal at 1000 MPa [56] led to an average estimation of $d_{ij} = -0.30 \pm 0.06$ Å as compared to $-0.068$ Å for the X-ray structures [52], suggesting that the deformation is very small and almost linear in pressure up to 1000 MPa. The computed and experimental $d_{ij}$ show some common features. Lysozyme is made of two domains, one containing mostly $\alpha$-helices and the other, smaller, containing a three stranded $\beta$-sheet; the latter is essentially rigid, while helices, $\alpha_1$, in particular, approach the $\beta$-domain. In general, inter-domain distances decrease much more than intra-domain ones.

In one of the high pressure simulations of BPTI [48] a structural change occurred at 1000 MPa involving the central part of the $\beta$-sheet and letting few hydrophobic residues exposed to the solvent. As for the helical structure, it reacts to a change in pressure less dramatically. Another simulation of BPTI at 1000 MPa [43] also showed a very small deformation of the helices, while there was a more dramatic change in the loops.

For SOD (Paci, unpublished results), structural changes at 1000 MPa were moderate but it is interesting to note that the contact surface between the two monomers decreased by about 16%, while the total solvent-accessible surface diminished by only 2.3%. This effect seems to indicate a reduction of the dimer stability, which might prelude to a solvent-mediated dissociation prior to unfolding of the monomers.

The major structural changes observed in metmyoglobin [54] at or above 900 MPa mainly consist in the disruption of about 70% of the native secondary structure.
4. Compressibility

The apparent volume of a solute is the only directly measurable volume of a protein. It is defined as the volume change observed when a solute is added to the pure solvent:

$$V_p^a = (V - V_s^0)$$

The partial volume is a well defined thermodynamic quantity which is equal to the partial derivative, at constant particle number and temperature, of the chemical potential with respect to the pressure:

$$\varphi_p = \left( \frac{\partial \mu_p}{\partial P} \right)_{NT}$$

In the limit of infinite dilution, partial and apparent volumes are equal.

The apparent volume of proteins, as any apparent property of a molecule in solution, depends on properties of the molecule only (which are thus called intrinsic properties), and on the non-ideal behavior of the solvent hydrating the molecule (which can be considered a surface effect). Apparent specific volume of native globular proteins falls into a narrow range around 0.72 cm$^3$/g revealing a marginal dependence on the extent of the solvent–protein interface.

Compressibility is defined as

$$\beta = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)$$

where $V$ is the volume of a microscopic sample and $P$ is the pressure. If the derivative is measured at constant temperature or entropy (and, in all cases, constant particle number), $\beta$ is isothermal ($\beta_I$) or adiabatic ($\beta_S$) compressibility, respectively.

Adiabatic compressibility of a solution, at a given concentration, can be obtained in density and sound velocity measurements [59,60]. The apparent or partial compressibility can be extracted [61,62] from the compressibilities of the solution at different concentrations which must include zero concentration (pure solvent).

The apparent and partial volumes and, consequently, apparent and partial compressibilities, can be written as the sum of an intrinsic protein term plus a solvation-related correction due to non-ideal behavior of the solvent. The intrinsic volume can, in turn, be written as the sum of the van der Waals volume plus the volume of voids due to an imperfect packing of a set of covalently connected spheres. Both decompositions can be reasonable and useful for specific purposes, but are essentially arbitrary. Several different and incompatible evaluations of the intrinsic compressibility can be found in the literature. Their difference is related to different definitions of what protein intrinsic volume is. In an important paper, Kharakoz and Sarvazyan [63] reviewed the partial compressibilities of globular proteins in water. They evaluated the contribution of hydration and intrinsic compressibilities to experimental partial compressibility using two independent methods. The first method assumes that the non-ideal behavior of the solvent in the presence of a protein can be evaluated by assuming that each group on the surface of the protein modifies the compressibility of the solvent as it does in small molecules and that this effect is additive. The second is based on the analysis of correlation between partial compressibility and the solvent-accessible surface (which is assumed to be proportional to $M^{2/3}$, where $M$ is the mass of the protein). They found that the isothermal intrinsic compressibility is about $14 \times 10^{-5}$ MPa$^{-1}$. (The volume of an object characterized by a constant compressibility of $14 \times 10^{-5}$ MPa$^{-1}$ decreases by 1.4% when the pressure increases by 100 MPa. In general, compressibility is not constant and decreases with increasing pressure.) This value is similar to that of solid organic polymers; it is about one third the isothermal compressibility of water at room temperature (45.2$ \times 10^{-5}$ MPa$^{-1}$) and varies little for different proteins. The average compressibility of water in the hydration shell of proteins was found to be 20% less than that of pure water. If one takes into account the fact that proteins are usually not spherical and uses $M^{0.73}$ as a better approximation for the protein surface, the protein intrinsic compressibility becomes about one half of the compressibility of the water [64].

Using the Voronoi volume [65–67] as a definition of the intrinsic volume, Paci and Marchi [51] found very similar compressibilities for two globular proteins possessing very different secondary structures, SOD and lysozyme. On a larger set of proteins (Paci, unpublished), including barnase, lactalbumin and potato carboxypeptidase inhibitor (PCI), using pressure
differences ranging from 500 to 2000 MPa, the Voronoi to pure water compressibility ratio was found to range from 0.46 to 0.54.

There are many geometrical recipes for assigning a volume to a protein. These recipes are not equivalent, and involve parameters whose definition is not always unique. The advantages of Voronoi volume is that the volume of each atom of the protein and the solvent is defined and does not depend on any parameter. It provides a plausible definition of the intrinsic volume of a protein because the compressibility estimated through it is close to the most reliable experimental estimates [61,63]; also, the compressibility of different proteins falls into a narrow range, which is expected since globular native proteins, from the point of view of condensed matter physics, are very similar objects [68].

Among other commonly used definitions of protein volume are the van der Waals volume (the volume of a collection of intersecting spheres of the atoms’ van der Waals radii, therefore not taking into account cavities and voids), the molecular volume [69], which is the volume enclosed by the contact surface of a probe sphere rolled on the van der Waals envelope of the molecule, or the excluded volume [70], which is the volume excluded to the center of the probe sphere. As discussed by Paci and Velikson [58], the above definitions of protein’s intrinsic volume have the contradictory feature of not reflecting the packing of solvent molecules surrounding the protein, being at the same time a function of the van der Waals radii of the protein’s atoms and of the probe radius. The molecular volume of a protein, which has been used to evaluate the intrinsic protein compressibility on some occasions [43,52,58], turned out to provide an estimation of the intrinsic compressibility (about 1/10 of the water compressibility) which seems to be strongly underestimated (and comparable to the compressibility of a metal). Unlike the molecular volume, the Voronoi volume does take into account modifications occurring in the solvent due to the presence of the protein and to possible external perturbations. This is possibly a reason why it provides estimates close to the experimental ones.

Another experimental approach to estimate the compressibility of proteins consists in monitoring changes in specific protein interatomic distances under pressure and assuming that the volume behaves as the distance to the third power (the assumption would be exact in the case of homogeneous systems). This compressibility would, in principle, be an intrinsic quantity in the sense that it does not include the effect of hydration. On the other hand, simulations have revealed that, on average, short distances compress less than longer ones [51]. This non-uniform contraction induced by high pressure contrasts with the spatially uniform expansion observed at high temperature [71]. Simulation [51] has shown how, by randomly picking a pair of atoms, compressibilities ranging from very large positive to negative values can be obtained. This explained why fluorescence spectroscopy measurements [72], which probe donor-acceptor distances in heme proteins as a function of pressure, found a protein intrinsic compressibility comparable to that of bulk water. It also provides an explanation to the low compressibility (10^-4 MPa^-1) obtained by the hole burning [73] technique; since average near-neighbor distances are probed, this technique is likely to lead to an underestimation of the intrinsic compressibility of proteins.

4.1. Volume fluctuations

Volume fluctuations are sometimes estimated by relating them to protein isothermal compressibility, \( \beta_T \), with the equation [74]:

\[
\delta V^2 = \frac{V k_B T \beta_T}{N^2 v^2}
\]

where \( k_B \) is the Boltzmann constant, \( T \) the temperature, \( V \) and \( \delta V^2 \) are the protein’s volume and the volume fluctuations squared. This relationship assumes either that the protein alone samples the NPT ensemble distribution, which is not true in general, or that the coupling between the protein and solvent volumes is small. It was shown by Paci and Marchi [51] that for Voronoi volumes the latter assumption is not even approximately true. The protein and solvent volume cross-fluctuations are of a magnitude comparable to the protein and solvent volume fluctuations alone, due to the small size of proteins and the resulting large surface-to-volume ratio. As a consequence, using Eq. 4 to derive \( \beta_T \) generally leads to an underestimation. Nevertheless, volume fluctuations and volume compressions of proteins and subunits of proteins are correlated. For example, the
cross-correlations between the relative volume fluctuations of residue $i$, $\delta V_i/V_i$, and its compression range from +0.65 to +0.80 for a set of five proteins (Paci, unpublished data). Thus, although there is no exact relation between volume fluctuations and compressibility of residues, these quantities are strongly correlated: highly compressible residues will show larger volume fluctuations and vice versa.

The volume fluctuations of a series of proteins have been recently used by Dadarlat and Post [75] to estimate the intrinsiccompressibility of a series of proteins. As protein intrinsic volume they consider the van der Waals volume where the radius of each atom is extended so that cavities and packing defects are filled (similar to the definition of excluded volume). The compressibility they report is about one fifth of the compressibility of water, thus below that obtained using Voronoi volumes and finite differences, but larger than the value obtained, always by finite differences, using the molecular volume [43,52]. The reason for the low compressibility obtained is probably attributable to the use of a relation (Eq. 4) between volume fluctuation and compressibility which is only approximately true for small systems. Also, the excluded volume as definition of intrinsic volume, similar to van der Waals and molecular volumes, tends to underestimate the compressibility as noted above.

4.2. Local compressibility

Another advantage of Voronoi volumes is that they are additive and well defined for every atom of every component of the solution. It is thus possible to attribute exact additive volumes to atoms, amino acids and substructures of a protein.

4.2.1. Substructures

For various proteins, the following results have been found ([56]; Paci, unpublished results). In lysozyme, the two domains have very similar compressibility. The less compressible structure is the $\beta$-sheet, while the three helices are more compressible than the whole protein. For SOD, the compressions of the regular and bent $\beta$-sheets do not deviate from the averaged compression. The most and least compressible structures are the two loops. The least compressible loop contains important residues for the regulation of the protein catalytic activity; its low compressibility suggests that the active site in SOD is ‘screened’ [1] from the effect of pressure by the deformation of its surrounding regions. In barnase, $\alpha$-helices, loops, $\beta$-strands and turns show almost the same compressibility while the most and least compressible structures are the two loops. In lactalbumin, $\beta$-strands are markedly less compressible than all other structural elements.

The analysis of the Voronoi compression, on the limited set of proteins studied, does not reveal a simple relation between compressibility and structure. Different compressibility of various structural elements of a protein seems to be related to the overall protein’s architecture and function rather than to the type of secondary structure elements present.

4.2.2. Residues

The compressibility of all amino acids is the same, within 20%. For a set of four proteins analyzed ([56]; Paci, unpublished results), a clear trend was found: most compressible amino acids are the hydrophobic Val, Leu and Ile, while least compressible are the charged amino acids Glu and Asp. For the four proteins, the calculated compressibilities are very similar and increase from charged to polar and to non-polar residues. These findings are in accord with the results, obtained by molecular modeling and simulation, that the free energy cost of forming atomic size cavities is significantly lower in non-polar regions of proteins than on the surface [76]. As non-polar regions contain larger portions of empty space, it is not surprising that they are more compressible. This agreement also supports the hypothesis that the free energy of cavity formation serves as a fine probe of the compressibility of the protein medium as a whole, and in particular of selected regions within the protein. Quite in disagreement with this is a result from normal mode analysis [77] (see Section 4) that the intrahelix compressibility in sperm whale deoxymyoglobin is markedly more compressible than hydrophobic clusters, even though there are large cavities in these clusters.

4.3. Normal mode calculations

Gö and collaborators studied the effects of pressure on proteins using normal modes and other tech-
niques than molecular dynamics. The common result of the various studies is that the intrinsic compressibility of proteins (using the excluded volume definition as intrinsic protein volume) may be larger than was thought due to low frequency modes which are overlooked in short molecular dynamics trajectories. Studying volume fluctuation by normal mode and strain tensor analysis [77] in sperm whale deoxymyoglobin showed that interhelical regions are more compressible than the intrahelical ones. The pressure-induced deformation was shown to be heterogeneous, and, most interestingly, the contribution from low frequency normal modes was more important than that from the high frequency ones. The same group also studied the mechanical response of a TIM-barrel protein to pressure [78] by assuming the volume change to be a linear function of normal mode variables. By Delaunay tessellation [79], a procedure analogous to the Voronoi construction, the space occupied by protein atoms was partitioned and attributed to various secondary structures in the protein. The compressibility of various secondary structural elements in this protein was obtained from volume changes of respective regions in two structures with and without applied pressure; regions inside the β-barrel and between secondary structural elements appeared to be compressible, while the helix, strand, and loop segments themselves were found to be quite incompressible.

Tama et al. [80] presented an alternative approach to study volume fluctuations of proteins (they also assume that the intrinsic protein volume is the excluded volume). They estimated volume fluctuation of human lysozyme from a previous molecular dynamics simulation [81] using a formalism based on principal component analysis. The isothermal intrinsic compressibility was found to be larger than that reported by most experiments (from their data one can deduce the value $10^9 \text{ MPa}^{-1}$ while the reported value, $10^3 \text{ MPa}^{-1}$, is probably a type). The difference between this result and others obtained by molecular dynamics (for example that of [51]) probably is due to different definitions of intrinsic protein volume, which, as we have stressed already, is a delicate issue in itself [58]. It is nevertheless interesting that the spectral analysis shows that the main contribution to fluctuation comes from low frequency modes related to large and slow motions of proteins. The authors conclude that long-time dynamics simulations are necessary to describe volume fluctuations of proteins, and the relatively short trajectories reported so far (including the 100 ps trajectory of [51]) might not be sufficient.

### 4.4. Compressibility of non-native states

The compact fold of native globular proteins is accompanied by a large number of non-covalent interactions which are responsible for both its structure and its physical properties. It is thus interesting to understand what discriminates, in terms of physical properties, native states from non-native ones.

Recent experimental results have shown that certain stable intermediates of various proteins (like the low pH molten globules) are characterized by an increase in the apparent compressibility as compared to the native state. This contrasts with the decrease of the apparent compressibility that occurs on protein denaturation.

Nölting and Sligar [82] first reported the measurement of the apparent adiabatic compressibility for ‘molten globule’ states of cytochromes $c$ and $b_{562}$. They showed that the intrinsic compressibility in the molten globule state differs only slightly from that in the native protein.

Breslauer and coworkers [83] measured apparent volume and compressibility of native, molten globule, and unfolded states of cytochrome $c$ continuously modifying the pH of the solution from 6 to 2. They found that, while the apparent volume shows only one broad transition, the apparent compressibility shows three distinguishable transitions. Based on their measurements and several assumptions, they conclude that the solvent-inaccessible core, which is 40% of the intrinsic volume of cytochrome $c$, is preserved in the molten globule and that the compressibility of this core is about 4 times larger than the intrinsic compressibility of the whole native protein. Further studies belonging to the same group of a set of globular proteins [13] led to the conclusion that changes in the apparent compressibility for single domain globular proteins correlate with the type of transition being monitored, independent of the protein. This seems to be a universal thermodynamic feature, just like the increase in the heat capacity upon unfolding [84].
Tamura and Gekko [85] also studied volume and compressibility changes upon thermal and guanidine hydrochloride-induced denaturation of ribonuclease A, concluding that apparent volume and compressibility sensitively reflect the different conformational changes induced by temperature or GuHCl. Denatured states have clearly different properties but the difficulty of splitting cavity and hydration contributions limits the possibility for a quantitative discussion of the results.

Kharakoz and Bychkova [86] measured partial specific volume and compressibility of human $\alpha$-lactalbumin using standard densitometric and ultrasonic techniques. Measured variations of apparent volume and compressibility led the authors to conclude that the molten globule is liquid-like (the intrinsic compressibility is twice larger than that of the native molecule) and that it contains a large amount of water (like in the wet molten globule model of Finkelstein and Shakhnovich [87]).

The experimental results summarized above, although they do not always provide a clear and coherent picture, always show that compact non-native states have a larger apparent compressibility relative to the native state; this in spite of the decrease expected from the increased hydration. This increase is not negligible: it ranges from 20 to 50% of the native value. This means that the molten globule (or the ensemble of states which fall under this name) has certain physical properties which clearly distinguish it from a native or an unfolded protein. However, the transition from native to molten globule cannot be compared to melting of organic crystals, where a 3-fold increase of the compressibility is observed. Simulations might play an important role in interpreting experimental results and challenging their interpretation (often based on assumptions which are difficult to check).

Molecular dynamics simulations (Paci, manuscript in preparation) at atmospheric and high pressure have been performed on native and non-native conformations of several proteins. Non-native conformations have been generated by high temperature molecular dynamics for barnase [8] and $\alpha$-lactalbumin (M. Buck, personal communication), and by reducing disulfide bonds for PCI (as described by Martí et al. [88]). In addition, polyalanines in $\alpha$-helical and extended conformations have been studied. For barnase and $\alpha$-lactalbumin, the intermediates considered have been selected as relatively stable conformations which, for $\alpha$-lactalbumin, have several of the experimentally observed properties of the molten globular state [89].

Molecular dynamic simulation shows that both apparent and intrinsic volumes of proteins are only marginally affected by partial denaturation. In contrast with experimental results mentioned above, the apparent compressibility always decreases upon partial unfolding. The estimated increase of the intrinsic compressibility (estimated from Voronoi volume changes) is moderate (13% at most, for the barnase compact intermediate) and very far from that (3-fold) expected from a solid to liquid phase transition.

The completely denatured form of PCI, lacking any secondary structure, shows the same moderate increase in the intrinsic compressibility as the compact intermediates of barnase and lactalbumin. Also, the $\alpha$-helical and extended polyalanines have identical intrinsic compressibility. Preliminary simulations do not show a strong dependence of the intrinsic compressibility on the conformation, when the Voronoi volume is used as intrinsic protein volume. A possible reason for the disagreement of the preliminary simulation results discussed here with the experiment is that the relatively stable intermediates selected at high temperature (or high temperature and low pH) molecular dynamic trajectories do not have the exact features of the molten globules studied experimentally.

5. Concluding discussion

Molecular dynamics simulations of solvated proteins show that even at pressure as high as 2000 MPa, the time scales that are computationally possible are too short to observe pressure-induced denaturation of proteins. Conformational states sampled by molecular dynamic simulations are not fully relaxed in the thermodynamic sense, but have undergone only elastic relaxation. Nonetheless, studying these pre-denatured states has provided useful information about the response of native proteins to external pressure and, as a consequence, some clues about pressure-induced denaturation.

On the other hand, studies based on simplified
models [17–19] have been shown to be very powerful tools to provide answers to questions which are important in understanding hydrophobic interactions and the role that they might play in pressure-induced unfolding of proteins. These studies show that the hydrophobic model for the protein interior explains both protein stability and the experimental evidence of protein denaturation at high pressure.

While a simple picture of a protein core as an oil drop in water can explain both stability at atmospheric pressure and denaturation at high pressure, the interior of a globular protein in the native state is characterized by a density [90,91] and a compressibility typical of a crystal [61,63]. This is so in spite of the constraints imposed by the covalent connectivity, heterogeneity and lack of long range order. Simulations, while confirming crystal-like packing and compressibility of proteins, also show that it would be simplistic to picture native globular proteins as crystal-like objects: proteins do not respond to a change in pressure in a uniform and homogeneous way. In particular, amino acids of different chemical character compress differently; polar and charged amino acids compress less than non-polar ones. The larger compressibility of hydrophobic residues is related to the much larger probability of cavity formation in hydrophobic cores than in other regions of the protein. Also, different regions or substructures within the protein compress differently; this does not seem to be correlated in a clear way to the structural patterns but depends more on the protein’s specificity.

As the protein volume decreases, no systematic relative increase in the solvent-accessible surface of the non-polar residues is found during elastic relaxation, nor is there a noticeable penetration of water molecules into the hydrophobic core. However, the situation may be quite different for the high pressure-denatured state, not observable in simulation, for which it is believed that non-polar residues are transferred to the solvent from the apolar core of the protein or that water is conveyed into the protein interior.

Proteins respond to changes in pressure by increasing their affinity to form hydrogen bonds with water. The increase in the number of protein–solvent hydrogen bonds is larger than that of solvent–solvent and solvent–protein hydrogen bonds. At the same time, the protein surface accessible to the solvent decreases while protein–solvent interaction energy decreases and the solvent–solvent energy increases.

A certain amount of experimental data on high pressure properties [92] of membranes await simulations to complement the experimental picture. Some studies on volumetric properties of nucleic acids were reviewed by Chalikian and Breslauer [93]. Among the results cited, some show how compressibility data can characterize hydration properties of DNA and RNA. From such results, information on conformational and binding properties can be extracted. Simulations at various pressures of these systems would be useful to clarify the relation between measured properties, which include the effect of hydration, and intrinsic ones, which are in principle related to purely structural properties.

Much remains to be done in understanding the properties of non-native states. The increasing amount of experimental data on such states, together with new theoretical approaches to determine their structures [94], call for further simulation work aimed at understanding the physical properties of those states, and in particular their packing and compressibility.

Longer high pressure simulations (10–100 ns time scale) of solvated proteins have become possible now; performed with a correct treatment of long range electrostatics and with a rigorous statistical–mechanical implementation of constant pressure, these simulations would also constitute a reference to test the results obtained from simplified approaches. Promising techniques include solvent boundary potentials [95] and implicit solvents [96], which may allow access to time scales one or two orders of magnitude longer. These time scales may not be sufficient to directly simulate the high pressure unfolding. However, longer simulations will allow the investigation of important issues related to the convergence and to the role of transitions between different conformational substates in the estimation of compressibility and volume fluctuations.

**Acknowledgements**

The author thanks Annette Salmeen, Michael Gross and Boris Velikson for critically reading the
manuscript, Michele Vendruscolo and Martin Karplus for useful discussions, Matthias Buck and Mare Martí for providing initial non-native conformations and Guy de Riencourt for hospitality while most of this review was written.

References


