The “rules” governing protein structure and stability are still poorly understood. Important clues have come from proteins that operate under extreme conditions, because these clarify the physical constraints on proteins. One obvious extreme is pressure, but so far little is known of the behavior of proteins under pressure, largely for technical reasons. We have therefore developed new methodology for calculating structure change in solution with pressure, using NMR chemical shift changes, and we report the change in structure of lysozyme on going from 30 bar to 2000 bar, this being the first solution structure of a globular protein under pressure. The α-helical domain is compressed by approximately 1%, due to tighter packing between helices. The interdomain region is also compressed. By contrast, the β-sheet domain displays very little overall compression, but undergoes more structural distortion than the α-domain. The largest volume changes tend to occur close to hydrated cavities. Because isothermal compressibility is related to volume fluctuation, this suggests that buried water molecules play an important role in conformational fluctuation at normal pressures, and are implicated as the nucleation sites for structural changes leading to pressure denaturation or channel opening.

© 2003 Elsevier Science Ltd. All rights reserved

Keywords: pressure; chemical shift; lysozyme; buried water; compression

---

**Introduction**

Recent protein structures from extremophiles have provided clues as to how proteins can adapt to extreme temperatures or high salt, and thereby helped unravel details of the physical constraints governing the structure and evolution of proteins. Temperature causes simultaneous changes in volume and thermal energy, which are difficult to disentangle. By contrast, pressure causes changes only in volume, and therefore the change in energy is better defined thermodynamically, which makes it easier to study theoretically. However, it remains difficult experimentally: there are only two crystal structures at high pressure and other spectroscopic techniques have to date produced only low-resolution or very site-specific data. Pressure effects are of interest because they help us understand how macromolecules behave not only at high pressure but also under normal conditions, since protein compressibility is directly related to the structural and conformational fluctuations of proteins at normal atmospheric pressure. Therefore, changes in structure in response to pressure can provide detailed insight into local volume fluctuations, such as those that lead to channel opening or ligand exchange. Pressure is also of practical interest because it is an important environmental variable: some deep-sea organisms live at pressures of over 1 kbar (100 MPa), and space exploration will require a detailed understanding of the effects of pressure on organisms.

We therefore report here a novel method for calculating the change in structure of a protein based on the changes in NMR chemical shifts measured on applying pressure, and its application to lysozyme. This report provides the first information at an atomic level on pressure effects on proteins in solution.
Figure 1. (bottom) Difference distance plot for Cα atoms. Orange and red contours denote parts of the structure that have moved closer together, while cyan and blue contours denote parts that have moved apart. Regions of secondary structure are indicated as α, α-helix; β, β-sheet; L, loop. (top) Structure of lysozyme, color coded as in the difference distance plot to show regions of the structure that have moved most relative to the rest of the structure. Data are averaged over all regions of the structure that have moved most relative to the rest of the structure. Parts of L1, β3/L4 and L7 have moved closer to the body of the protein, while most of β3 and L6 have moved away, in agreement with crystallographic results. It is interesting to note that some of the largest changes are close to the ligand-binding site of lysozyme. Thus, residues 59, 62 and 63 interact with the sugar in binding site C, while the active site nucleophile, Asp52, is located in the L4/β4 region, which undergoes the largest deformation in the structure.

Results

Structural changes

The key to the method is that chemical shift changes are much more powerful and reliable than the shifts themselves. This is because of the nature of chemical shifts, which are very short-range, and therefore act well as “refinement” tools for modifying an existing good structure, but poorly as a method for ab initio structure calculation. Therefore, even though current methods for the calculation of chemical shift in proteins have limited accuracy, a structural change from a high quality starting structure can be calculated reliably and accurately, even when the changes involved are (as here) no more than 0.2 Å. This implies that the best way to calculate a structural change is to start from a good structure for which the experimental and calculated shifts agree exactly. The only way in which this can be realized in practice is to generate a “good” structure (i.e. one that has good covalent geometry and is at an energy minimum in any force field to be used subsequently), and to calculate its chemical shifts. These calculated shifts can then be adopted as the “experimental starting shifts” (in our case, the low-pressure shifts), and the changed shifts (here, the high-pressure “experimental” shifts) are generated by adding the experimental changes in shift on going from low pressure to high pressure to the experimental starting shifts. This procedure achieves the aim of using only the change in shift as a real experimental restraint. In order to ensure that the small changes in structure are meaningful, we carried out two sets of calculations in parallel: from the starting structure, we applied the high-pressure experimental shifts as restraints, but we also applied the experimental starting shifts as restraints in a parallel calculation, so that all conditions were as similar as possible for the two structures. A sufficient number of structures was calculated for each set that random effects were very small.

A crystal structure of hen egg-white lysozyme was energy refined and minimized using X-PLOR, to produce a starting structure that has no systematic structural drift in the X-PLOR forcefield. The chemical shifts for this structure were calculated, using X-PLOR. The structure was then subjected to two calculations, in one of which the chemical shifts were restrained to stay constant, and in the other of which they were restrained to change to the same extent as caused experimentally by a change in pressure from 30 bar to 2000 bar, to produce “low-pressure” and “high-pressure” structures, respectively. The resultant sets of 50 structures had chemical shift errors of 0.000(±0.004) ppm for the unchanged shifts, and 0.002(±0.007) ppm for the changed shifts. The structures had moved 0.118(±0.015) Å and 0.212(±0.026) Å (backbone RMSD), respectively, from the starting structure. The root-mean-square difference in backbone coordinates between the two resultant structures was 0.21 Å. The changes arising from the application of pressure are thus small, but they are significant, and moreover they are not evenly distributed around the molecule. Analysis of structural change by best-fit superimposition and subsequent measurement of distance between corresponding atoms is sensitive to the atoms used for superimposition. We therefore used a distance matrix approach, which is not biased in the same way. Distances were calculated between all pairs of Cα atoms in the high-pressure structure, and subtracted from the corresponding distances in the low-pressure structure. The resultant difference distance matrix (Figure 1) gives a good indication of regions that have moved relative to the rest of the protein. Parts of L3, β3/L4 and L7 have moved closer to the body of the protein, while most of β3 and L6 have moved away, in agreement with crystallographic results. It is interesting to note that some of the largest changes are close to the ligand-binding site of lysozyme. Thus, residues 59, 62 and 63 interact with the sugar in binding site C, while the active site nucleophile, Asp52, is located in the L4/β4 region, which undergoes the largest deformation in the structure.
Interdomain movement

The structure of lysozyme (Figure 1, top) can be divided into two domains: an α-helical domain, comprising residues 1–39 and 89–129, and a β-sheet domain, comprising residues 40–88. The α-domain has altered rather little in its position and orientation, while the β-domain has been more extensively deformed, particularly in the β-hairpin comprising strands β3 and β4, and in the spatially adjacent long loop L6. The deformation is essentially a twist of the β-hairpin, bringing opposing corners of the sheet closer together. The large β-domain deformation was not observed in the crystallographic study, possibly as a consequence of crystal contacts, which are known to be important for the β-hairpin. This loop is also one of the most variable parts of the structure.

Because the interdomain region, which forms the hinge between the two domains, is also the region that undergoes the greatest compression, we analyzed the structure for hinge bending around an axis along the interdomain region. Any hinge bending should be revealed as approximately equal, but opposite in sign, rotations around the hinge. However, analysis (Figure 2) shows that a unimodal distribution was observed for the intact protein and for both domains, indicating absence of hinge bending. The peak is not exactly at 0°, due to a small movement of the axis between structures.

We also analyzed the structures to see whether on compression, atoms moved towards the center of each domain independently, or whether they moved towards the center of mass of the protein as a whole. The results (Figure 3) show that there is little difference between the behavior of each domain separately or the protein as a whole: in each case, there is a general tendency to move towards the center of mass, but with a considerable spread of directions.

Compression

Despite the largest deformation occurring in the β-domain, the main compression is in the interdomain region and in the α-domain (Table 1). The results indicate that the overall volume of the protein is reduced by about 1.5%, with proportionally smaller changes in the surface area and in the radius of gyration. The volume of the α-domain is reduced by approximately 1%, while the β-domain volume is almost unchanged. This corresponds to a compressibility coefficient of approximately 5/Mbar for the α-domain, 0 for the β-domain, and 7.5 for the whole protein. In comparison, the crystallographic study obtained values of 5.7, 0, and 4.7, respectively. The results are therefore in good agreement. We note that in itself this is a good justification of the calculation, which used only chemical shifts as restraints and therefore contained no explicit compression forces.
We have undertaken an analysis of the location of the compression. First, we note that the compression in the \( \alpha \)-domain arises mainly from a closer packing between adjacent helices, whose mean interhelical C\( ^\alpha \)–C\( ^\alpha \) distance is reduced by 0.033(\( ^{\pm} 0.058 \)) \( \text{Å} \). This suggests that the compression involves a reduction in the volume of packing defects. The volumes of internal cavities were therefore measured using the program VOIDOO.\(^{16} \) In order to obtain consistent results, we used a small grid size, and repeated the calculations using at least ten random rotations of the

![Figure 3](image)

**Figure 3.** Shifts towards center of mass. Continuous line, all atoms, protein center of mass. Broken line, \( \alpha \)-domain atoms and center of mass. Dotted line, \( \beta \)-domain. A random distribution of directions would produce a horizontal line at \( f = 1 \), while movement towards the center of mass produces angles biased towards 0°.

<table>
<thead>
<tr>
<th>Measure</th>
<th>30 bar</th>
<th>2 kbar</th>
<th>Percentage change (30 → 2000 bar)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>Whole protein</td>
<td>( \alpha )-Domain</td>
<td>( \beta )-Domain</td>
</tr>
<tr>
<td>( R_g ) (( \text{Å} ))(^{a} )</td>
<td>13.82</td>
<td>11.75</td>
<td>9.98</td>
</tr>
<tr>
<td>( I_{xx} )(^{b} )</td>
<td>7.473</td>
<td>3.962</td>
<td>1.190</td>
</tr>
<tr>
<td>( I_{yy} )(^{b} )</td>
<td>9.038</td>
<td>6.026</td>
<td>2.187</td>
</tr>
<tr>
<td>( I_{zz} )(^{b} )</td>
<td>21.52</td>
<td>7.450</td>
<td>4.129</td>
</tr>
<tr>
<td>Vol. (( \text{Å}^3 ))(^{c} )</td>
<td>15,325</td>
<td>9283</td>
<td>5378</td>
</tr>
<tr>
<td>Vol. (( \text{Å}^3 ))(^{d} )</td>
<td>14,957</td>
<td>9130</td>
<td>5161</td>
</tr>
<tr>
<td>Surface area (( \text{Å}^2 ))(^{e} )</td>
<td>6382</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cavity vol. ((1.2 \text{ Å probe})^{f} )</td>
<td>1187 ± 98</td>
<td>1129 ± 93</td>
<td>–4.8</td>
</tr>
<tr>
<td>Cavity vol. ((0 \text{ Å probe})^{f} )</td>
<td>1990 ± 523</td>
<td>1705 ± 330</td>
<td>–14.3</td>
</tr>
</tbody>
</table>

\( a \) Radius of gyration, calculated as mass-weighted root mean square distance from center of mass.

\( b \) Principal moments of inertia, in \( \text{Da} \text{Å}^2 \times 10^{-4} \) (calculated using X-PLOR).

\( c \) Calculated using GRASP, with a probe radius of 1.1 Å.

\( d \) Calculated from Voronoi volumes.\(^{38} \)

\( e \) Calculated using the CCP4 routine AREAIMOL. Areas of domains are not meaningful.

\( f \) Calculated using VOIDOO.\(^{16} \)
coordinates. Calculations were carried out using probe radii both of 1.2 Å (representing cavities large enough potentially to contain water molecules) and of 0 Å (representing not only these larger cavities but also “true” packing defects). Results for the whole protein are presented in Table 1, from which it can be seen that “large” cavities (i.e. those measured using a 1.2 Å probe) decrease in volume by approximately 5%, whereas the total cavity volume decreases by 14%. Indeed, subtraction of the 1.2 Å-probe volume from the 0Å-volume provides an estimate of the packing defect volume, which decreases from 803 Å³ to 576 Å³, a reduction of 28%. It is therefore concluded not only that there is a large reduction in cavity volume, but also that the reduction is much more significant for the smaller packing defects.

As a further measure of compression, we calculated the radial distribution function of the protein. The results are shown in Figure 4. There are large peaks at 1.1 Å (bonds to hydrogen), 1.6 Å (heavy atom bonds) and 2.2 Å (non-bonded contacts). The distributions for the low and high-pressure structures are similar (Figure 4(a)). However, the difference between the two distributions (Figure 4(b)) shows that at every peak in the radial distribution function, there is an increase in the number of shorter contacts and a decrease in the number of longer contacts. Thus, the compression appears to be very broadly distributed, appearing as reductions in directly bonded distances as well as in non-bonded contacts and cavity contractions.

Hydrogen bond lengths within the structure are on average slightly shorter at higher pressure, but with considerable variation: backbone hydrogen bond lengths are reduced by 0.011 Å but with a standard deviation of 0.107 Å. This change is of the same order of magnitude as seen in our previous studies, which saw on average a change of 0.01 Å/kbar, and is also consistent with spectroscopic studies and demonstrates that compression due to shortening of hydrogen bonds is of similar magnitude to the overall compression. A detailed study of the distribution of changes in hydrogen bond length showed no obvious correlation with structure. The change in HN chemical shift receives approximately equal contributions from a shortening of the hydrogen bond and from the hydrogen bond becoming more linear (i.e. from an improvement in the geometry of the hydrogen bond). Backbone dihedral angles show changes of the same order of magnitude as those seen in the crystallographic study: ϕ changes by 4.0(±3.9)°, while ψ changes by 4.1(±3.5)°, compared to 3.2° and 3.1°, respectively, in the crystal study. Interestingly, there is a good correlation (R = -0.80) between the change in ϕ of one residue and the change in ψ of the preceding residue, implying that many of the angle changes are a result of peptide plane reorientation.

Discussion

Compression

Compression is not uniform. The α-domain is compressed to a much larger extent than the β-domain, in agreement both with the previous crystallographic study and with measurements of compressibility of a wide range of proteins. We note, however, that hydrogen bond lengths in the β-domain are compressed to a similar extent as those in the α-domain. Despite this, there is no overall compression of the β-domain. This clearly implies that there must be regions of local expansion as well as compression. This is indeed true, as shown in Figure 1: again, this matches the conclusions seen from the crystallographic study. This result is of importance for understanding pressure denaturation. In the most common model, cavities in the protein become filled with water molecules, which leads to hydration of the interior of the protein and subsequent denaturation. Our results confirm that cavities can be enlarged under pressure (since this is the only way in which the β-domain can expand locally), hence presenting a means by which hydration of the interior can be accomplished.

The structural changes presented here as a consequence of increase of pressure are broadly comparable to those seen in previous crystallographic studies. The major difference is the
The role of buried water molecules

The most interesting result of this study arises from examining the locations of residues that have large changes in volume. Figure 5 shows that the largest changes occur close to the buried water molecules. Interestingly, the volume changes are both positive and negative, again demonstrating that the application of pressure can cause local expansion. The association between volume change and buried water molecules can also be seen from Figure 1, where most of the regions of large structural change are next to water molecules: L6 close to the single isolated water, and L3 and L7 close to the cluster of water molecules in the region between the two domains. This is also the region where the greatest overall compression occurs. It is noteworthy that we previously observed a small number of large and non-linear pressure-dependent changes in chemical shift in lysozyme for residues close to cavities large enough to contain buried water molecules. We have also observed in a range of proteins that non-linear pressure-dependent changes in chemical shift are correlated with the density of cavities large enough to contain buried water molecules, and suggested that the non-linear changes might derive from more compact low-lying excited states centered on these cavities. Here, we have shown that the same regions of the protein undergo large structural changes, thus directly implicating the regions close to buried water molecules in structural plasticity, and strengthening the argument for the low-lying excited states. It is interesting to note that large cavities (≥ 1.2 Å radius) are less compressible than the smaller packing defects (Table 1), but yet are the nuclei for pressure-dependent structural change. One can speculate that this is a result of the conformational freedom of water molecules, which is expected to be greater than that of protein.

It is striking that the buried water molecules are mainly located in the “hinge” region between the two domains. However, there is no significant hinge bending under pressure. We conclude that the buried water molecules do not contribute to any concerted hinge bending motion: rather, they provide an increase in local compressibility.
The compressibility of a protein is related to volume fluctuations by the equation:

\[
\langle (\delta V)^2 \rangle = kT \beta_V
\]

where \(\langle (\delta V)^2 \rangle\) is the average squared volume fluctuation, \(k\) is the Boltzmann constant, \(T\) the absolute temperature, \(V\) the volume, and \(\beta_V\) the isothermal compressibility coefficient. Although this equation is strictly valid only for macroscopic properties, it indicates a likely relationship between local compressibility (i.e., atomic volume changes) and local structural fluctuation. The local volume fluctuation can be quite large; the equation translates to a RMS fluctuation for the whole protein of approximately 70 Å² at 20 °C, which is likely to be concentrated in the region around buried water molecules because these are the regions of highest local compressibility. Thus, the results presented here suggest that volume fluctuation is very significant in the vicinity of hydrated cavities, which are thereby implicated as nucleation sites for structural change, including pressure-dependent shift differences and as “mobile defects” permitting solvent exchange and the transient opening of channels.

Materials and Methods

Pressure was applied to hen egg-white lysozyme using an on-line pressure cell, and 2D NMR spectra were recorded, and used to measure 1H chemical shifts as a function of pressure. For 240 protons with a linear pressure dependence, structural restraints were generated as the change in chemical shift between 30 bar and 2 kbar, and are listed in the Supplementary Material.

Crystal structure 2lym was selected as the initial structure, being the low-pressure structure used for the crystallographic study. Protons were added by X-PLOR, except for the protons on the five buried water molecules (numbers 130, 131, 132, 133, 134), which were added manually using standard geometries. The structure was refined in a full forcefield, being the low-pressure structure used for the crystallographic study. First, a hinge axis was defined, between the Cα atoms of residues 38 and 97, and the Cartesian coordinates of the low and high-pressure structures were separately converted to cylindrical coordinates, with the hinge axis forming the z axis, and atoms were analyzed for the distribution of their change in rotation angle \(\theta\) around the axis. If there is motion around the hinge, a bimodal distribution of angles is expected, because each domain will require a separate value of \(\theta\) to align it with the corresponding domain in the other coordinate system. By contrast, if there is no hinge motion, both domains will have identical values of \(\theta\).

The direction of movement of all backbone atoms was analyzed by superimposing the low and high-pressure structures on their centers of mass (using all heavy atoms). For each atom, the direction of movement was calculated with respect to the vector to the center of mass of the whole protein or of each domain. Directions of motion \(\langle \theta \rangle\) were divided into \(9\) bins, and the frequencies were smoothed and normalized by multiplying by \(180/2\pi\sin(\theta)\), where \(n\) is the number of atoms.

Atomic coordinates

Low and high-pressure coordinates have been deposited with the Research Collaboratory for Structural Bioinformatics Protein Data Bank (accession codes 1gxv and 1gxx, respectively).

Acknowledgements

We thank Professors T. Asakura and K. Gekko for helpful discussions, and Drs T. J. Palmer, C. J. Craven and L. Hosszu for help with the computing. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan and a project grant from the Biotechnology and Biological Sciences Research Council of the UK.

References


Edited by M. F. Summers

(Received 23 October 2002; received in revised form 27 January 2003; accepted 29 January 2003)

Supplementary Material comprising one Table is available on Science Direct