PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Investigation of Protein Unfolding and Stability by Computer Simulation

W. F. Van Gunsteren, P. H. Hunenberger, H. Kovacs, A. E. Mark and C. A. Schiffer

Phil. Trans. R. Soc. Lond. B 1995 **348**, 49-59
doi: 10.1098/rstb.1995.0045

Investigation of protein unfolding and stability by computer simulation

W. F. VAN GUNSTEREN, P. H. HÜNENBERGER, H. KOVACS, A. E. MARK AND C. A. SCHIFFER

Laboratory of Physical Chemistry, ETH Zürich, 8092 Zürich, Switzerland

SUMMARY

Structural, dynamic and energetic properties of proteins in solution can be studied in atomic detail by molecular dynamics computer simulation. Protein unfolding can be caused by a variety of driving forces induced in different ways: increased temperature or pressure, change of solvent composition, or protein amino acid mutation. The stability and unfolding of four different proteins (bovine pancreatic trypsin inhibitor, hen egg white lysozyme, the surfactant protein C and the DNA-binding domain of the 434 repressor) have been studied by applying the afore-mentioned driving forces and also to some artificial forces. The results give a picture of protein (in)stability and possible unfolding pathways, and are compared to experimental data where possible.

1. INTRODUCTION

In the biologically active form, proteins generally assume a specific, folded conformation. Knowledge of native protein conformations has been derived from a variety of experiments. Three-dimensional structural models have been derived from X-ray diffraction of crystals and from multi-dimensional nuclear magnetic resonance (NMR) experiments. By comparison, very little is known about the denatured state of a protein or about the processes of denaturation and folding. Experimentally, it is very difficult to obtain structural information about these processes because of the atomic length scale and the short timescale of these processes. Partial structural information on intermediate states of the folding or unfolding pathways of proteins may be obtained by monitoring different spectroscopic properties of amino acid side chains or particular protein atoms or labelling groups as a function of time after induction of folding or denaturation (Dobson *et al.* 1994). Partly denatured states may be trapped and studied in atomic detail (Neri *et al.* 1992*b*).

In view of the difficulty of the study of protein folding and denaturation by experimental means, one may turn to the method of computer simulation of these processes as a means of elucidating their nature in atomic detail. Theoretical methods to predict protein properties can be classified as follows.

1. Methods to assess the correctness of a given protein structure (see, for examples, Novotny *et al.* 1984; Jones *et al.* 1992; Lüthy *et al.* 1992; Maiorov & Crippen 1992; Ouzounis *et al.* 1993; Kocher *et al.* 1994), and for the prediction of protein structure (see, for examples, Dill 1985, 1990; Covell & Jernigan 1990; Unger & Moult 1993; Kolinski & Skolnick 1994).

2. Methods to determine the relative stability of protein mutants or different protein conformations in terms of relative free energies. It has been shown (Shi Yun-yu *et al.* 1993) that free-energy simulation techniques cannot currently be used to reliably predict protein stability.

3. Methods to simulate the process of protein folding or denaturation. Here, one thinks primarily of molecular dynamics (MD) simulation or of Monte Carlo (MC) simulation in cases where the MC step and acceptance criterion are of a physical nature (see, for examples, Levitt & Warshel 1975; Sali *et al.* 1994).

The simulation of protein folding is severely limited by two factors.

1. The size of the conformational space to be searched for the native state of low free energy.

2. The lack of structural information with respect to the denatured (possibly unfolded) state of a protein, which may be characterized by manifold conformations.

To reduce the size of the conformational space, very simple protein models (e.g. models using one interaction site per amino acid residue) are used (Seetharamulu & Crippen 1991), or the accessible protein conformations can be restricted to those fitting on a regular spatial lattice (Covell & Jernigan 1990). The lack of knowledge of the structures that constitute the denatured state can be compensated by repeating the folding process starting from different random conformations. However, owing to the limited accuracy of very simple protein models and the heuristic nature of most search methods used, reliable simulation of protein folding is still out of reach.

The process of protein destabilization or denaturation can be simulated much more easily on a computer, because atomic detail of the initial conformation is known for many proteins; but even these simulations are limited.

1. Because the process of protein unfolding (and folding) is driven by the interplay between protein–

Phil. Trans. R. Soc. Lond. B (1995) **348**, 49–59
Printed in Great Britain

protein, protein–solvent and solvent–solvent interactions at a given temperature, pressure, pH, etc., the inclusion of explicit solvent molecules in an unfolding simulation is mandatory to obtain useful results.

2. Simulation of a protein in explicit solvent requires substantial computing power, which currently limits the timescale of such simulations to nanoseconds.

3. The interatomic interaction function used should give the correct relative strength of protein–protein, protein–solvent and solvent–solvent interactions, a feature which is not necessarily present when, for example, protein and solvent force fields of different type or parameterization are combined.

In this paper we study protein stability and the onset of denaturation by MD simulation under denaturing conditions in the presence of explicit solvent molecules.

The unfolding of a protein can be induced by different driving forces. These include: amino acid mutation (e.g. the change of cysteine to alanine residues); increase in temperature; increase in pressure; change in solvent type (e.g. from a polar one, such as water, to a less polar one, such as chloroform); and change in the pH or ionic strength of the solution containing the protein.

Artificial driving forces, such as the application of a constant radial (from the protein centre of mass) force or a constant radial temperature gradient, have also been used (Hao *et al.* 1993; Hünenberger *et al.* 1995), but have been shown to introduce considerable bias to the unfolding process (Hünenberger *et al.* 1995).

Technically, the unfolding pathway of a protein can be studied by two different types of simulation.

1. By performing equilibrium simulations of the protein in solution under different conditions of temperature, pressure, pH, etc., or for different mutants and solvent composition. In this case the equilibrium properties averaged over the different equilibrium simulations are compared, and it is assumed that the state points at which the protein is simulated are representative for points along its denaturation pathway.

2. By performing, in parallel, equilibrium and non-equilibrium simulations starting from a single (equilibrated) structure. Equilibrium of the system is perturbed by switching on one of the mentioned driving forces and its relaxation towards the (new) equilibrium, which may be a denatured state, is monitored. In this case the change of properties over time, and not the averaged properties, is analysed.

Both types of simulation give information about protein stability. A protein would be considered more stable if it displays structural integrity alongside a small degree of fluctuation at equilibrium under various (extreme) external conditions (temperature, pressure, etc.). When performing non-equilibrium simulations the protein would be considered more stable the slower it reacts or gives in to the perturbing forces.

A variety of structural properties and some energetic properties can be monitored (out of equilibrium) or averaged (in equilibrium) and subsequently compared between simulations under different conditions or with experimental data. The radius of gyration of the protein can be compared to light scattering data and changes in the polar and non-polar solvent accessible area of the protein can be used to estimate the change in heat capacity upon unfolding. A changing amount of helicity may be compared with circular dichroism data. Atom–atom distances may be compared to nuclear overhauser enhancement (NOE) intensities obtained by NMR spectroscopy. The deviation of the protein structure from its crystalline X-ray structure may be analysed. Presence or absence of hydrogen bonds may be correlated with data from hydrogen–deuterium exchange experiments. Protein energies may be compared with experimental melting temperatures. Yet, because the experimental data on protein properties at different points along the pathways of denaturation are scarce, the reliability of the denaturation simulations cannot be established by a comparison of simulated properties with these experimental data. The confidence in the denaturation simulations will largely stem from the correct reproduction of native protein properties under native conditions using the same interaction function and similar simulation protocols.

Simulation studies of protein stability and denaturation in explicitly simulated solvents have only recently become possible with the increasing power of computers. Mark & van Gunsteren (1992) studied thermal unfolding of hen egg white (HEW) lysozyme using non-equilibrium MD at high temperature (500 K). Daggett & Levitt (1992, 1993) studied thermal unfolding of bovine pancreatic trypsin inhibitor (BPTI) and its reduced form in equilibrium at different temperatures (423 K, 498 K). A similar type of study was performed by Daggett (1993) for the C-terminal fragment (CTF) of the L7/L12 ribosomal protein, and by Caflisch & Karplus (1994) for the protein barnase. Tirado-Rives & Jorgensen (1993) studied the effect of a change in pH by comparing simulations of myoglobin in which histidine residues were differently protonated. Simulations at high pressure have been reported by Kitchen *et al.* (1992) and Brunne & van Gunsteren (1993) for BPTI, and by Hünenberger *et al.* (1995) for HEW lysozyme. The effect of different cysteine to alanine mutations on the stability of BPTI at room temperature has been studied by Schiffer & van Gunsteren (1995). The effect of a change of solvent has been studied for the surfactant lipoprotein C by Kovacs *et al.* (1995) and for the DNA binding domain of the 434 repressor by Schiffer *et al.* (1995).

Here we briefly review the four last-mentioned studies to obtain a general picture of protein stability and the onset of denaturation under the influence of different driving forces. Because the computational set-up and details of analysis have been reported elsewhere, we only describe the simulations and results in a global manner.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

2. STRUCTURAL STABILITY OF DISULPHIDE MUTANTS OF BPTI

Experimentally, the folding pathway of BPTI has been studied extensively (Weissman & Kim 1991; Darby *et al.* 1992; Goldenberg 1992). In its native form (see figure 1) BPTI has three disulphide bonds (5–55, 14–38, 30–51). The three forms of BPTI having two native disulphide bridges maintain the structure and function of the native three disulphide species. However, the three forms of BPTI having only one native disulphide bridge show different structural properties. Using NMR it has been found that the (5–55) species is folded in the native conformation. It can inhibit trypsin. The (14–38) species is not likely to adopt a well defined structure because this disulphide bond lies at the surface of the protein (see figure 1), and is the first to break and the last to form. The (30–51) species maintains native-like structure for residues 19–36 (β-sheet) and 42–56 (α-helix), the N-terminal part and the loop between residues 37 and 41 being disordered (van Mierlo *et al.* 1992). The fully reduced BPTI shows no stable secondary or tertiary structure.

In the light of this experimental information we decided to study the relative structural stability of the native form, the (5–55) species, the (30–51) species and the no-disulphide (all Cys to Ala) mutant of BPTI using non-equilibrium MD simulations. The four species were each simulated in aqueous solution at room temperature (at 300 K for 150 ps, then at 320 K for 100 ps), starting from the same equilibrated solution structure of native BPTI. Each system consisted of one BPTI molecule in a periodic truncated octahedron containing 2371 water molecules. The Gronigen molecular simulation (GROMOS) protein force field (van Gunsteren & Berendsen 1987) and the SPC/E water model (Berendsen *et al.* 1987) were used. Computational details are given in Schiffer & van Gunsteren (1995).

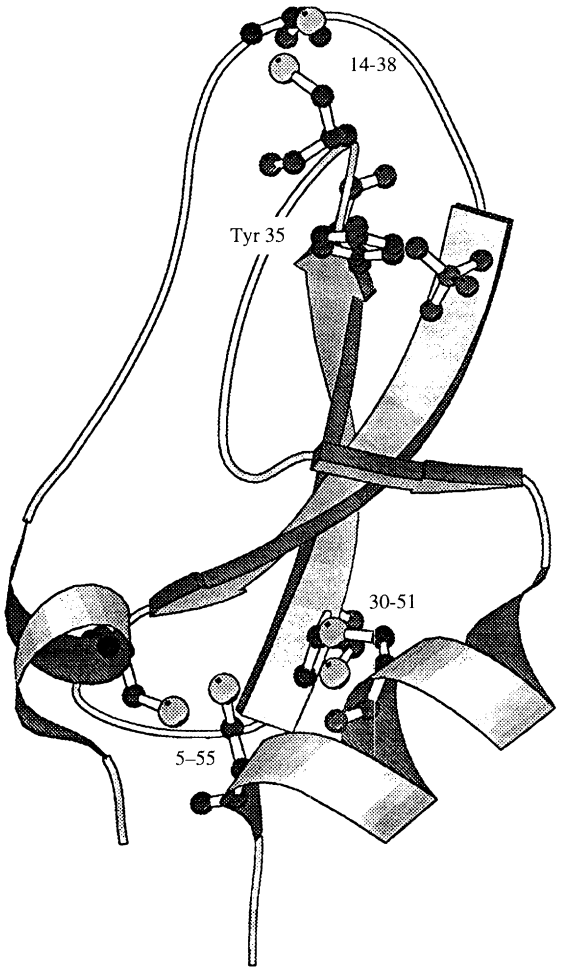

Figure 1. Schematic drawing of the crystal structure of BPTI showing the location of the three native disulphide bridges, the helices and the β-sheet.

Figure 2 shows a comparison of the common starting structure with the structure after 250 ps of MD simulation of the four different species. The native BPTI (figure 2*a*) maintains its native structure, apart from a few N- and C-terminal residues which are also found to be disordered experimentally. The (5–55) species (figure 2*b*) essentially maintains the native structure, although slight deformations in the loop regions are observed. The (30–51) species (figure 2*c*) shows structural changes for residues 1–17 and residues 38–43 with respect to the rest of the molecule, which corresponds directly to the part of this mutant observed to be disordered experimentally. The non-disulphide species (figure 2*d*) shows much more deformation over the 250 ps simulation period.

These results show that parallel MD simulations of different protein mutants starting from the same initial structure may be used to obtain indications of the relative structural stability of the protein mutants.

3. STRUCTURAL STABILITY OF HEW LYSOZYME UNDER CONDITIONS OF HIGH TEMPERATURE OR PRESSURE

The folding and unfolding of HEW lysozyme has been studied extensively by a great variety of experimental methods (Dobson *et al.* 1994). Most studies concern temperature or denaturant-induced unfolding, but pressure-induced unfolding data have also been reported (Samarasinghe *et al.* 1992). In its native state (see figure 3*a*) HEW lysozyme is characterized by two structural domains: one consisting primarily of α-helix (α-domain) and the other dominated by a section of triple stranded β-sheet (β-domain). It has been shown that the two domains behave differently under different denaturing conditions, and are largely independent of each other during folding and unfolding processes. These factors make lysozyme a particularly interesting system in which to study such processes. At pressures up to 5 kbar, a low pH of 3.9 and a temperature of 342 K (which is close to the thermal denaturation temperature of about 350 K at 1 bar and pH of 7.0) the maximum degree of denaturation observed by NMR was about 53%. In X-ray diffraction studies of lysozyme crystals at high pressure (Kundrot & Richards 1987) deformation was observed primarily in the α-domain. In contrast, hydrogen exchange studies during refolding from a denaturant-induced unfolded state indicate

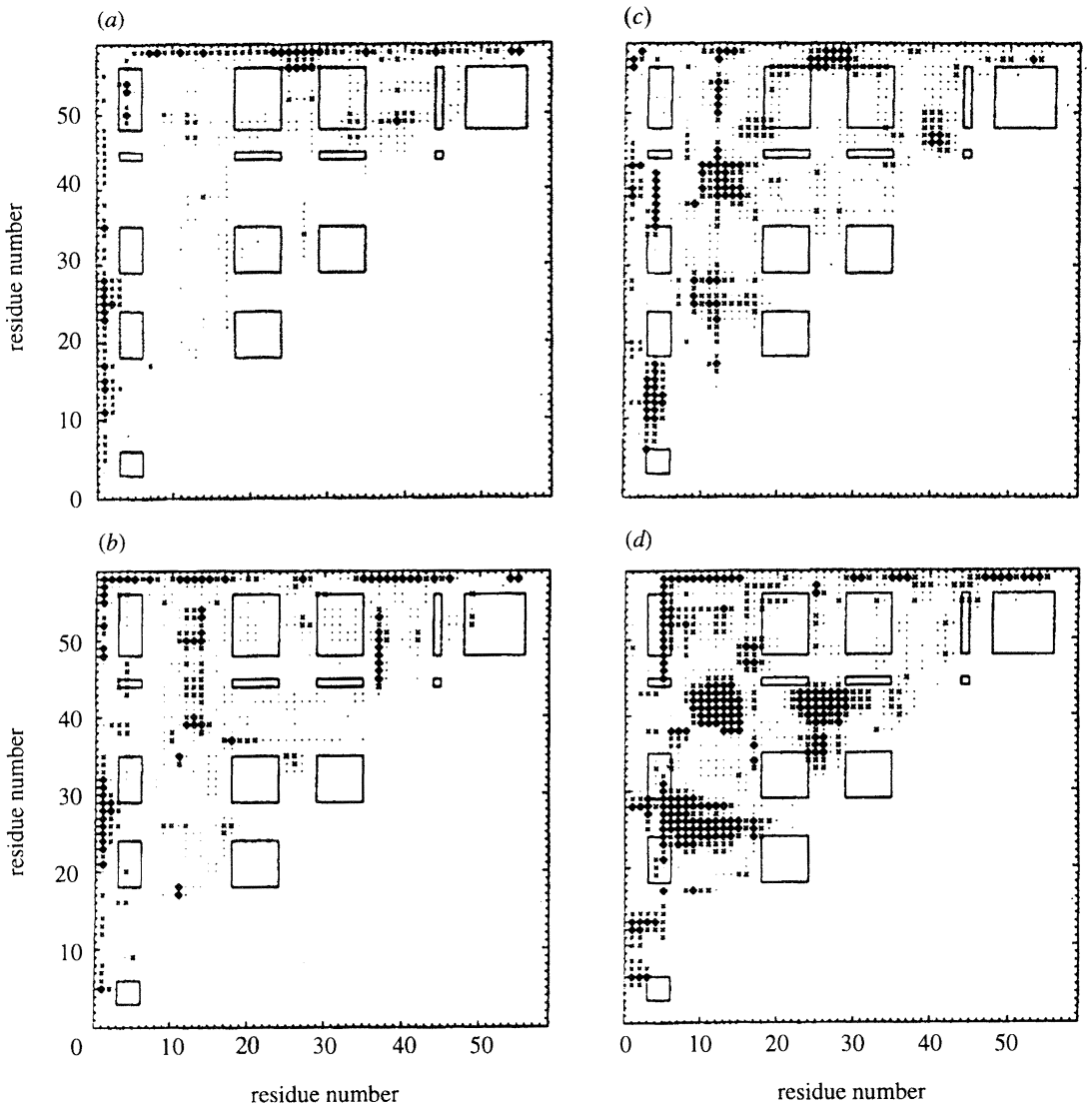

Figure 2. Residue–residue matrices showing the change in C_α–C_α atom distances for four BPTI species occurring in four MD simulations of 250 ps length at room temperature starting from an equilibrated structure of native BPTI in aqueous solution. If a C_α–C_α distance has changed less than 0.1 nm over the 250 ps, the residue–residue matrix displays a blank; if the distance has changed by 0.1–0.2 nm, it shows a dot; a change of 0.2–0.3 nm is indicated by a cross; changes of more than 0.3 nm are indicated by a rhombus. (*a*) native BPTI; (*b*) (5–55) species; (*c*) (30–51) species; (*d*) no disulphide species. Secondary structure elements and their relative position in the amino acid sequence are indicated by boxes.

preferential stabilization of the α-domain. The helices in the α-domain (see figure 3*a*) are protected against hydrogen exchange from the onset of folding, whereas the β-sheet amide hydrogens are only protected in the final stages of denaturant-induced refolding (Radford & Dobson, this volume).

In view of this wealth of experimental information, we decided to study the denaturation of HEW lysozyme in aqueous solution by non-equilibrium MD simulation using different driving forces. The system consisted of one HEW lysozyme molecule in a rectangular periodic box of $4.9 \times 5.3 \times 6.8$ nm^3 containing 5345 water molecules. The GROMOS protein force field (van Gunsteren & Berendsen 1987) and the SPC water model (Berendsen *et al.* 1981) were used: computational details are given in Mark & van Gunsteren (1992) and Hünenberger *et al.* (1995).

Figure 3*a* shows the solution structure of HEW lysozyme equilibrated for 50 ps at 300 K and 1 bar pressure; this should be compared with the structure obtained after an additional 210 ps of MD simulation at 342 K and a pressure of 10 kbar (see figure 3*c*). Within 210 ps, only slight denaturation is observed: the D helix unfolds. A comparison of the compressibility of the two domains of HEW lysozyme shows that the β-domain is essentially incompressible whereas the α-domain and the interdomain hinge region contract under pressure, a result which agrees with an X-ray diffraction study of HEW lysozyme crystals under 1 kbar pressure (Kundrot & Richards 1987).

Increasing the temperature to 500 K after 50 ps MD simulation at 300 K and 1 bar pressure, results in rapid unfolding of lysozyme. Figure 3*b* shows the structure after 120 ps of high temperature motion and figure 3*d* that which is obtained after 180 ps. Figure 4 shows the presence of different types of (secondary) structure elements as a function of time. With an increase in temperature from 300 to 500 K at 50 ps, a differential

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

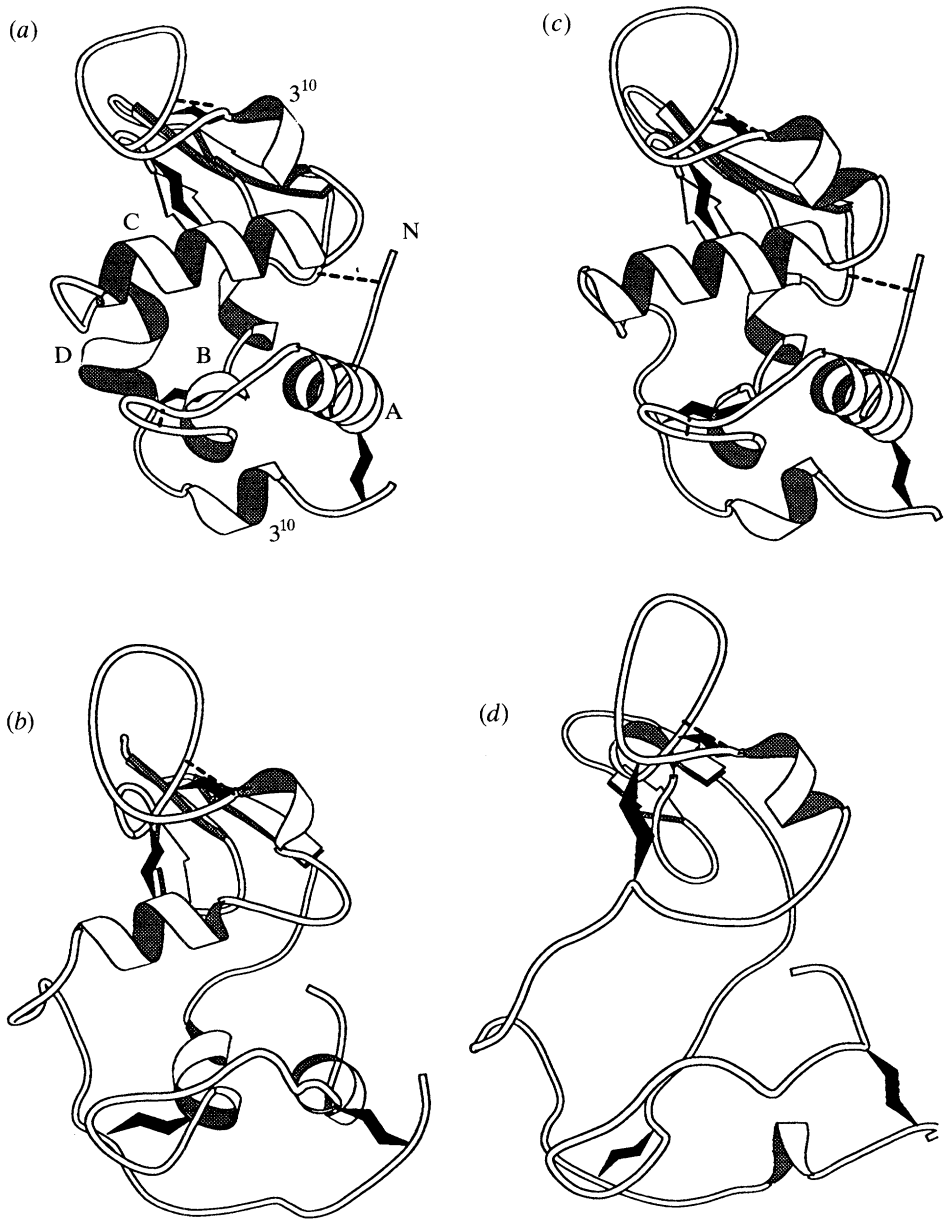

Figure 3. Ribbon models for the different structures of HEW lysozyme after: (*a*) 50 ps equilibration at 300 K and 1 bar (*b*)120 ps at 500 K (*c*)210 ps at 342 K and 10 kbar (*d*)180 ps at 500 K.

denaturation of the different parts of the molecule is observed. The B helix breaks down, but refolds shortly after 110 ps. The D helix and terminal 3^{10} helix do not seem to be very stable. The C helix and nearby 3^{10} helix resist denaturation for the longest time, followed by a part of the B helix and then the A helix. The β-sheet is very resistant to denaturation. The fluctuating nature of the intra-protein hydrogen bonds displayed in figure 4 shows that different secondary structure elements may be formed along the (un-)folding pathway. The occurrence of stable hydrogen bonds outside secondary structure elements shown in figure 4 suggests that protection of a backbone amide hydrogen may result from a transient (un-)folding conformation and need not necessarily result from a direct involvement in secondary structure.

These results of pressure- and temperature-induced denaturation of HEW lysozyme show that computer simulation of these processes can be used to obtain insight, at the atomic level, into the differential structural stability of parts of a protein and with respect to the initial stages of the denaturation process.

4. STRUCTURAL STABILITY OF THE SURFACTANT PROTEIN C IN CHLOROFORM, METHANOL AND WATER

The stability of a folded protein will depend not only on its amino acid sequence, the temperature and pressure, but also on the nature of its solvent or membrane environment. The stability of an α-helix is especially sensitive to its environment. For example, trifluoroethanol is known to induce helicity of peptides in solution, and peptides with stretches of hydrophobic residues are assumed to adopt a helical conformation in a lipid environment. In these cases the stability of a given helix will depend on an energetic balance between interactions between spatially neighbouring amino acids in the polypeptide chain and interactions

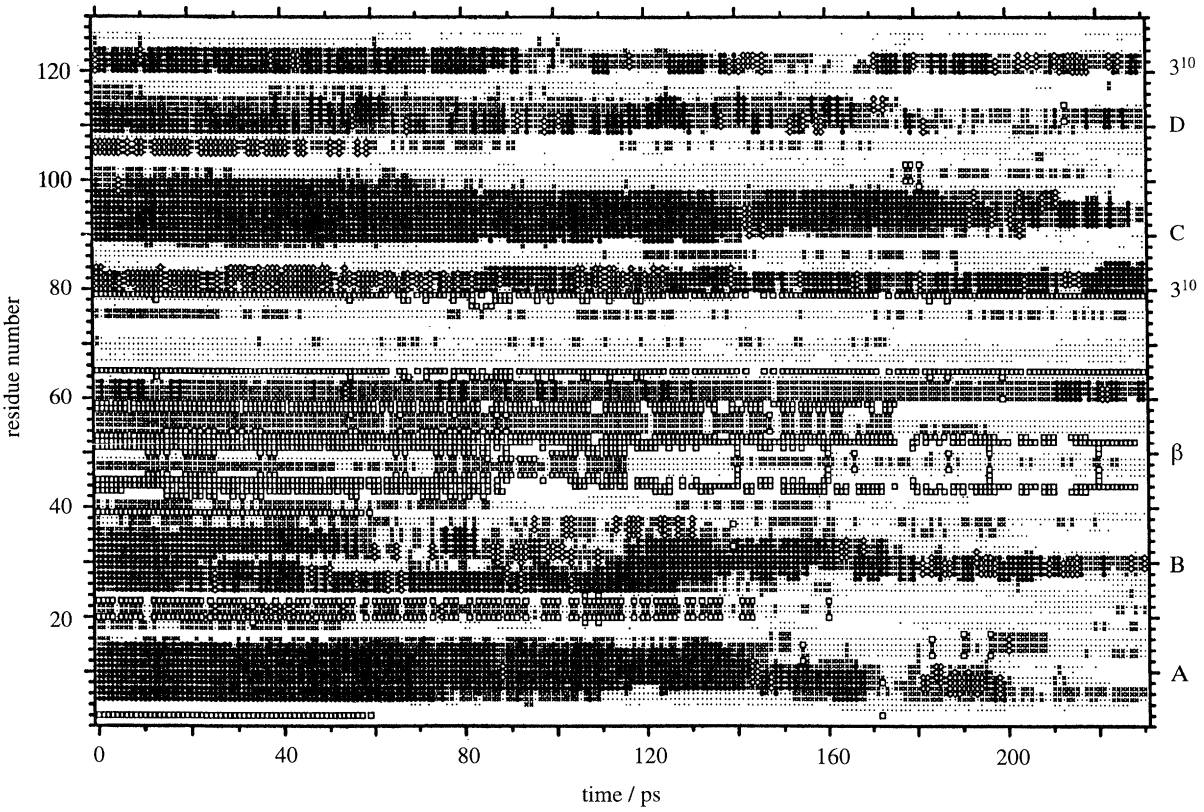

Figure 4. Secondary structure of HEW lysozyme as a function of time. Until 50 ps the temperature is 300 K, afterwards it is 500 K. α-helix: ●; β-bridge or sheet: □; 3^{10}-helix: ◇; π-helix: ◆; hydrogen bonded turn: ×; bend: ·; all according to the DSSP program (Kabsch & Sander 1983).

Table 1. *Physical properties of the solvents at 298 K (Lide 1993)*

	dipole moment	density	molar volume	dielectric constant	viscosity	polarizability $(\alpha_0/4\pi\epsilon_0)$
	Debye	(g cm^{-3})	(cm^3 mol^{-1})		cP	10^{-24} cm^3
chloroform	1.04 (gas)	1.480 (exp)	80.7	4.7	0.54	9.5
(Dietz *et al.* 1984, 1985)	1.10 (MD)	1.484 (MD)				
methanol	1.70 (gas)	0.787 (exp)	40.7	32.7	0.54	3.3
(Stouten 1989)	2.32 (MD)	0.791 (MD)				
water	1.85 (gas)	0.997 (exp)	18.1	78.4	0.89	1.5
(Berendsen *et al.* 1981)	2.27 (MD)	1.000 (MD)				

to the solvent environment. Stability also depends on the rearrangement of the solvent medium that is induced by the presence of the protein. An interesting case is presented by the lung surfactant protein C (SP-C), which is an essential component of pulmonary surfactant, a mixture composed mainly of phospholipids and a few specific proteins which reduces the surface tension at the alveolar air–liquid interface in the lung thereby preventing alveolar collapse at end expiration (Johansson *et al.* 1994*a*). The 35 residue SP-C contains stretches of seven (15–21) and four (23–26) consecutive valine residues. It maintains an α-helical form between residues 9 and 34 in a 1:2 mixture chloroform:methanol containing 5% 0.1 M HCl. The eight N-terminal residues lack a stable regular secondary structure (Johansson *et al.* 1994*b*).

To investigate the effect of environment suggested by this experimental information we studied the relative stability of SP-C in three different pure solvents – chloroform, methanol and water – by equilibrium MD simulations. The use of pure solvents avoids the difficulty of obtaining a proper equilibration of a mixed solvent within the simulation timespan and enables a more straightforward comparison of protein–solvent interactions. The physical properties of the three solvents are quite different (see table 1), and it was ensured that the solvent models used in the simulations correctly reproduced the pure liquid properties of the solvents. The GROMOS protein force field (van Gunsteren & Berendsen 1987) with the correction indicated by Mark *et al.* (1994) was used. A characterization of the simulated systems is given in

Table 2. *Characterization of the simulated systems: the reported energies are an average over the last 60 ps of simulation*

	number of atoms	temperature	initial box	length of simulation	protein–solvent energy	
					van der Waals	electrostatic
		K	nm^3	ns	(10^3 kJ mol^{-1})	(10^3 kJ mol^{-1})
chloroform	9935	300	261	1.02	−1.73	−0.57
methanol	11362	300	261	1.02	−0.91	−3.47
water	19966	300	200	0.30	−0.64	−3.89
truncated SP-C in water at:						
300 K	5624	300	60[a]	1.03	−0.53	−2.97
500 K	11462	500	116	0.90	−0.45	−2.46

[a] A rectangular box was used instead of the truncated octahedron used in the other simulations.

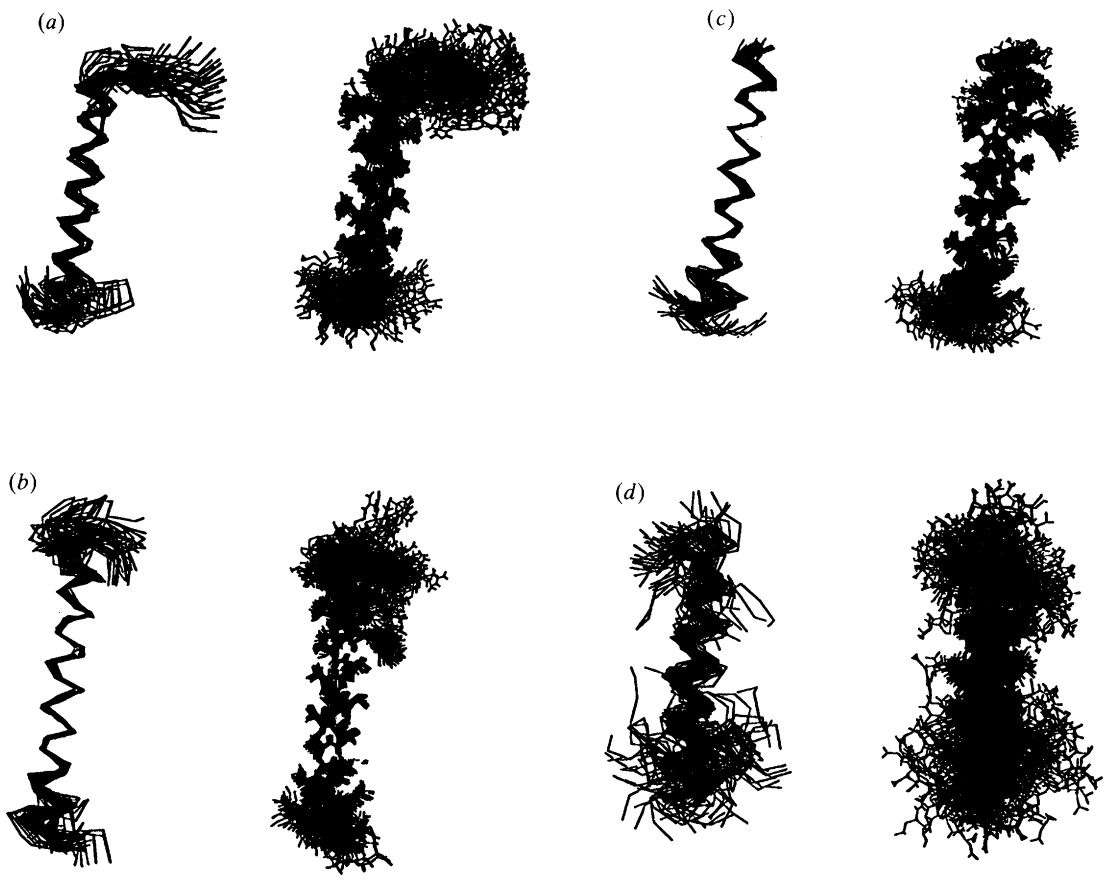

Figure 5. Superposition of averaged coordinates from 15 ps consecutive intervals from the 1 ns trajectories of the surfactant protein SP–C. Backbone C_α atoms (left) and all atoms (right) are displayed. (*a*) SP-C in chloroform, 300 K; (*b*) SP-C in methanol, 300 K; (*c*) truncated SP-C in water, 300 K; (*d*) truncated SP-C in water, 500 K. Only the C_α atoms of residues 9 to 28 of the complete peptide, LRIPCCPVNLKRLLVVVVVVVLVVVVIVGALLMGL, were included in the superposition. The molecules are depicted with the N-terminus up and C-terminus down. In truncated SP-C the first five residues are absent.

table 2. Owing to the small size of a water molecule, the simulation of SP-C in water is very expensive. Therefore, we decided to omit the first five N-terminal residues of SP-C, which were experimentally shown to be disordered, and simulated this truncated SP-C in water for a full nanosecond. A simulation of truncated SP-C in water at 500 K served as a control experiment and allowed a comparison with earlier high temperature unfolding studies of other proteins. The starting structure contained an extended chain (residues 1–8 and 35) and a helical part (residues 9–34). Further computational details are given by Kovacs *et al.* (1995).

Figure 5 shows a superposition of SP-C structures obtained from the four 1 ns MD simulations. At 300 K the α-helical conformation of SP-C is stable in all three solvents. It is most stable in water with a root mean square (r.m.s.) positional fluctuation of the C_α atoms of 0.069 nm, and of all atoms of 0.141 nm (residues 9–28). SP-C was least stable in chloroform with r.m.s. positional fluctuations of 0.126 nm for the C_α atoms and 0.290 nm for all atoms.

The simulations show that contrary to what might have been expected based on previously published values for the helix propensity of valines, the α-helical fold of the valyl-rich, predominantly hydrophobic

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Figure 6. Residue–residue matrices showing the change in C_α–C_α atom distances for the DNA binding domain of the 434 repressor occurring in six MD simulations of 250 ps length starting from an equilibrated structure of the protein in aqueous solution. The symbols indicating the size of the changes are defined in the legend of figure 2. (*a*, *b*) Standard solvent interaction at (*a*) 300 K and (*b*) 350 K. (*c*, *d*) Modified Coulomb protein–solvent interaction: (*c*) 10% increased (SPM) and (*d*) 10% decreased (SPL). (*e*, *f*) Modified Coulomb solvent–solvent interaction: (*e*) 10% increased (SSM) and (*f*) 10% decreased (SSL).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

peptide SP-C is remarkably stable in water, methanol and to a lesser extent in chloroform. This stability is primarily due to the additive effect of van der Waals interactions caused by the close packing of the branched aliphatic side-chains. This also prevents the solvent molecules from interacting with the protein backbone, especially in the poly-valyl part of Val-15–Val21. In the more polar solvents the interaction between hydrophobic side-chains is enhanced, producing increased helix stability. Even at elevated temperatures the poly-valyl stretch in the middle of the helical part of SP-C does not unfold. This does not mean that the α-helical fold of SP-C corresponds to the free-energy minimum in water. This cannot be determined from the current simulation. The results do demonstrate that the helix propensity of a given amino acid, in particular that of valine, can be highly sequence and environment specific.

5. STRUCTURAL STABILITY OF THE DNA BINDING DOMAIN OF THE 434 REPRESSOR IN DENATURING SOLVENTS

The three-dimensional structure of the 63-residue DNA binding domain of the 434 repressor in aqueous solution has been determined by NMR spectroscopy (Neri *et al.* 1992*a*). It consists of five α-helices and is not stabilized by disulphide bonds or metal ions. Yet, the folded conformation of this molecule is very stable, and even in 7 M urea a residual hydrophobic cluster is observed for residues 53–60 by NMR (Neri *et al.* 1992*b*).

In view of this experimental information, we decided to study the denaturation of the DNA-binding domain of the 434 repressor, under the influence of a denaturing solvent using non-equilibrium MD simulations. A simulation of a urea–water mixture would take a long time to equilibrate and sample the spatial distribution of the molecules properly. In short simulations, salt molecules would not have sufficient time to diffuse and equilibrate around the protein. A way around this limitation is to simulate a homogeneous aqueous environment but mimic the effect that the denaturant has on the protein conformation by changing the water–water or the water–protein interaction. First, the protein, immersed in a periodic truncated octahedron containing about 2000 water molecules, was equilibrated for 380 ps at 300 K and 1 bar. The corrected GROMOS force field and the SPC/E water model were used, and the equilibrated system was used to branch off five different simulations. In four of these the Coulomb interactions were modified as follows: the solvent–protein interaction was increased (SPM) or decreased (SPL) by 10%, or the solvent–solvent interaction was increased (SSM) or decreased (SSL) by 10%. For comparison, the equilibrium simulation was continued at 300 K, and a control simulation at 350 K was also performed: computational details are given in (Schiffer *et al.* 1995).

Figure 6 shows a comparison of the common starting structure with the structure after 250 ps of MD simulation of the six different systems. In the natural solvent (figure 6*a*) the protein maintains its native structure, apart from the C-terminal residue and a slight shift of the C-terminal ends of helices I and III with respect to each other. Of the four simulations with modified solvent interactions (*c*–*f*), the one in which the protein–solvent interaction was increased by 10% (SPM) shows by far the largest extent of protein denaturation. The denaturation is much larger than that obtained by raising the temperature to 350 K (figure 6*b*). This picture is confirmed if the simulations are extended by another 250 ps. The SPM denaturation process distorts the whole molecule, except for the five helices and the relative position of helix I with respect to helix V (figure 6*c*). Experimentally, the latter part of the molecule is seen to maintain residual structure in a high salt solution.

The simulations show that the protein is most easily denatured by increasing the protein–solvent interaction.

6. CONCLUSIONS

Using the proteins BPTI, HEW lysozyme, surfactant protein C and the DNA-binding domain of the 434 repressor as examples, it has been illustrated how an atomic picture of the onset of protein denaturation can be obtained by MD computer simulation. Denaturation can be induced by different driving forces, such as an increase of the temperature or pressure, a change of protein or solvent composition, or a modification of particular interactions in the molecular system. In all cases the process of protein denaturation is governed by the interplay of protein–protein, protein–solvent and solvent–solvent interactions. By monitoring the unfolding of a protein, the relative stability of the different parts of the protein can be determined and an indication of the responsible atomic interactions can be obtained.

When simulating non-equilibrium processes, there is generally little or no experimental information available with which the simulated results can be compared to establish the reliability of the latter. In addition, a direct comparison with results of hydrogen exchange labelling and other related experiments is made difficult by the large difference between simulated and experimental timescales, and by the possibility that unfolding and folding follow different pathways. This makes an unambiguous assessment of unfolding simulations difficult. Conversely, such simulations are useful because they give access to atomic details which are inaccessible by experiment. In cases where several models for a process have been proposed, the compatibility with the simulated results can be taken as supportive evidence in favour of a particular model.

REFERENCES

Berendsen, H.J.C., Grigera, J.R. & Straatsma, T.P. 1987 The missing term in effective pair potentials. *J. phys. Chem.* **91**, 6269–6271.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. & Hermans, J. 1981 Interaction models for water in relation to protein hydration. In *Intermolecular forces* (ed. B. Pullman), pp. 331–342. Dordrecht: Reidel.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Brunne, R.M. & van Gunsteren, W.F. 1993 Dynamical properties of bovine pancreatic trypsin inhibitor from a molecular dynamics simulation at 5000 atm. *FEBS Lett.* **323**, 215–217.

Caflisch, A. & Karplus, M. 1994 Molecular dynamics simulation of protein denaturation: Solvation of the hydrophobic cores and secondary structure of barnase. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1746–1750.

Covell, D.G. & Jernigan, R.L. 1990 Conformations of folded proteins in restricted spaces. *Biochemistry* **29**, 3287–3294.

Daggett, V. 1993 A model for the molten globule state of CTF generated using molecular dynamics. In *Techniques in protein chemistry IV* (ed. R.H. Angeletti), pp. 525–532.

Daggett, V. & Levitt, M. 1992 A model of the molten globule state from molecular dynamics simulations. *Proc. natn. Acac. Sci. U.S.A.* **89**, 5142–5146.

Daggett, V. & Levitt, M. 1993 Protein unfolding pathways explored through molecular dynamics simulations. *J. molec. Biol.* **232**, 600–619.

Darby, N.J., van Mierlo, C.P.M., Scott, G.H.E., Neuhaus, D. & Creighton, T.E. 1992 Kinetic roles and conformational properties of the non–native two-disulphide intermediates in the refolding of bovine pancreatic trypsin inhibitor. *J. molec. Biol.* **22**, 905–911.

Dietz, W. & Heinzinger, K. 1984 Structure of liquid chloroform. A comparison between computer simulation and neutron scattering results. *Ber. BunsenGes. phys. Chem.* **88**, 543–546.

Dietz, W. & Heinzinger, K. 1985 A molecular dynamics study of liquid chloroform. *Ber. BunsenGes. phys. Chem.* **89**, 968–977.

Dill, K.A. 1985 Theory for the folding and stability of globular proteins. *Biochemistry* **24**, 1501–1509.

Dill, K.A. 1990 Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155.

Dobson, C.M., Evans, P.A. & Radford, S.E. 1994 Understanding how proteins fold: the lysozyme story so far. *TIBS* **19**, 31–37.

Goldenberg, D.P. 1992 Native and non-native intermediates in the BPTI folding pathway. *TIBS* **17**, 257–261.

Hao, M.-H., Pincus, M.R., Rackovsky, S. & Scheraga, H.A. 1993 Unfolding and refolding of the native structure of bovine pancreatic trypsin inhibitor studied by computer simulation. *Biochemistry* **32**, 9614–9631.

Hünenberger, P.H., Mark, A.E. & van Gunsteren, W.F. 1995 Computational approaches to study protein unfolding: Hen egg white lysozyme as a case study. *Proteins*. (In the press.)

Johansson, J., Curstedt, T. & Robertson, B. 1994*a* The proteins of the surfactant system. *Eur. Resp. J.* **7**, 372–391.

Johansson, J., Szyperski, T., Curstedt, T. & Wüthrich, K. 1994*b* The NMR structure of the pulmonary surfactant-associated polypeptide sp–c in an apolar solvent contains a valyl-rich α-helix. *Biochemistry* **33**, 6015–6023.

Jones, D.T., Taylor, W.R. & Thornton, J.M. 1992 A new approach to protein fold recognition. *Nature, Lond.* **358**, 86–89.

Kabsch, W. & Sander, C. 1983 Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.

Kitchen, D.B., Reed, L.H. & Levy, R.M. 1992 Molecular dynamics simulation of solvated protein at high pressure. *Biochemistry* **31**, 10083–10093.

Kocher, J.-P., Rooman, M.J. & Wodak, S.J. 1994 Factors influencing the ability of knowledge-based potentials to identify native sequence-structure matches. *J. molec. Biol.* **235**, 1598–1613.

Kolinski, A. & Skolnick, J. 1994 Monte Carlo simulations of protein folding. I. Lattice model and interaction scheme. *Proteins* **18**, 338–352.

Kovacs, H., Mark, A.E., Johansson, J. & van Gunsteren, W.F. 1995 The effect of environment on the stability of an integral membrane helix: Molecular dynamics simulations of surfactant protein C in chloroform, methanol and water. *J. molec. Biol.* (In the press.)

Kundrot, C.E. & Richards, F.M. 1987 Crystal structure of hen egg-white lysozyme at a hydrostatic pressure of 1000 atmospheres. *J. molec. Biol.* **193**, 157–170.

Levitt, M. & Warshel, A. 1975 Computer simulation of protein folding. *Nature, Lond.* **253**, 694–698.

Lide, D.R. 1993 *Handbook of chemistry and physics*. Boca Raton, Florida: CRC Press.

Lüthy, R., Bowie, J.U. & Eisenberg, D. 1992 Assessment of protein models with three-dimensional profiles. *Nature, Lond.* **356**, 83–85.

Maiorov, V.N. & Crippen, G.M. 1992 Contact potential that recognizes the correct folding of globular proteins. *J. molec. Biol.* **227**, 876–888.

Mark, A.E. & van Gunsteren, W.F. 1992 Simulation of the thermal denaturation of hen egg white lysozyme: Trapping the molten globule state. *Biochemistry* **31**, 7745–7748.

Mark, A.E., van Helden, S.P., Smith, P.E., Janssen, L.H.M. & van Gunsteren, W.F. 1994 Convergence properties of free energy calculations: α-cyclodextrin complexes as a case study. *J. Am. chem. Soc.* **116**, 6293–6302.

Neri, D., Billeter, M. & Wüthrich, K. 1992*a* Determination of the nuclear magnetic resonance solution structure of the DNA-binding domain (residues 1 to 69) of the 434 repressor and comparison with the X-ray crystal structure. *J. molec. Biol.* **223**, 743–767.

Neri, D., Billeter, M., Wider, G. & Wüthrich, K. 1992*b* NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science, Wash.* **257**, 1559–1563.

Novotny, J., Bruccoleri, R. & Karplus, M. 1984 An analysis of incorrectly folded protein models. Implication for structure prediction. *J. molec. Biol.* **177**, 787–818.

Ouzounis, C., Sander, C., Scharf, M. & Schneider, R. 1993 Prediction of protein structure by evaluation of sequence–structure fitness. Aligning sequences to contact profiles derived from three-dimensional structures. *J. molec. Biol.* **232**, 805–825.

Sali, A., Shakhnovich, E. & Karplus, M. 1994 Kinetics of protein folding. A lattice model study of the requirements for folding to the native state. *J. molec. Biol.* **235**, 1614–1636.

Samarasinghe, S.D., Campbell, D.M., Jonas, A. & Jonas, J. 1992 High-resolution NMR study of the pressure-induced unfolding of lysozyme. *Biochemistry* **31**, 7773–7778.

Schiffer, C.A. & van Gunsteren, W.F. 1995 Structural stability of disulfide mutants of BPTI: A molecular dynamics study. (Submitted.)

Schiffer, C.A., Dötsch, V. & van Gunsteren, W.F. 1995 Exploring the role of solvent in the denaturation of a protein: A molecular dynamics study on the DNA binding domain of the 434 repressor. *Biochemistry*. (Submitted.)

Seetharamulu, P. & Crippen, G.M. 1991 A potential function for protein folding. *J. math. Chem.* **6**, 91–110.

Shi Yun-yu, Mark, A.E., Wang Cun-xin, Huang Fuhua, Berendsen, H.J.C. & van Gunsteren, W.F. 1993 Can the stability of protein mutants be predicted by free energy calculations? *Protein Engng.* **6**, 289–295.

Stouten, P.F.W. 1989 Small molecules in the liquid, solid and computational phases. Dissertation, University of Utrecht, The Netherlands.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Tirado-Rives, J. & Jorgensen, W.L. 1993 Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry* **32**, 4175–4184.

Unger, R. & Moult, J. 1993 Genetic algorithms for protein folding simulations. *J. molec. Biol.* **231**, 75–81.

van Gunsteren, W.F. & Berendsen, H.J.C. 1987 *Groningen molecular simulation* (GROMOS) library manual. Groningen, The Netherlands: Biomos.

van Mierlo, C.P.M., Darby, N.J. & Creighton, T.E. 1992 The partially folded conformation of the Cys-30 Cys-51 intermediate in the disulfide folding pathway of bovine pancreatic trypsin inhibitor. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6775–6779.

Weissman, J.S. & Kim, P.S. 1991 Reexamination of the folding of BPTI: Predominance of native intermediates. *Science, Wash.* **253**, 1386–1393.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES