PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Computational Investigations of Protein Denaturation: Apomyoglobin and Chaotrope-Arene Interactions

William L. Jorgensen, Erin M. Duffy and Julian Tirado-Rives

Phil. Trans. R. Soc. Lond. A 1993 **345**, 87-96
doi: 10.1098/rsta.1993.0120

Computational investigations of protein denaturation: apomyoglobin and chaotrope-arene interactions

By William L. Jorgensen, Erin M. Duffy and Julian Tirado-Rives

Department of Chemistry, Yale University, New Haven, Connecticut 06511, U.S.A.

Molecular dynamics (MD) and Monte Carlo (MC) simulations are being used to investigate protein denaturation. The calculations use the AMBER/OPLS force field with explicit representation of the solvent via the TIP3P and TIP4P models of water. The thermal denaturation of apomyoglobin has been followed in two 500 ps MD simulations at 85 °C. The resultant structures provide a detailed model of a molten globule, and close agreement is obtained with experimental data on the helical content of both native apomyoglobin and the low pH folding intermediate. The mechanism of protein denaturation by chaotropic agents is also being pursued. The possibility of direct contact between the chaotropes and aromatic sidechains is supported by MC computations of free energy profiles for the approach of urea and guanidinium ion to aromatic hydrocarbons in water.

1. Introduction

Pathways for protein folding and denaturation are only beginning to be understood. Folding intermediates have now been detected for several proteins, and characterization of the physical properties and structures of these 'molten globules' is advancing. Hydrogen isotopic exchange experiments in conjunction with two-dimensional (2D) NMR have proven particularly valuable in characterizing the protection of backbone NH protons during refolding (Kim & Baldwin 1990). Strong correspondences have been found between the sites protected in the early stages of refolding and the locations of α-helices and β-sheets in native proteins including apomyoglobin (Hughson *et al.* 1990), cytochrome C (Roder *et al.* 1988), and barnase (Matouschek *et al.* 1992). For apomyoglobin, the helicity has also been determined by circular dichroism (CD) to be reduced from *ca.* 55% for the native protein to *ca.* 35% for an unfolding intermediate (I state) observed at pH 4.2. The I state has been further characterized by the isotopic exchange technique to retain moderately stable helices at the positions of the A, G, and H helices in native myoglobin, illustrated in figure 1 (Hughson *et al.* 1990).

Besides the structures of folding intermediates, another fundamental, but poorly understood, issue is the mechanism of protein denaturation by common chaotropic agents such as urea and guanidinium chloride. Direct probes are hard to formulate. Model studies on the solubility of hydrocarbons and peptides support a bulk solvent effect by which the chaotropes make water a less hydrophobic medium or direct interactions between the denaturants and polypeptide, which favour the unfolded state (Robinson & Jencks 1965; Roseman & Jencks 1975). Support for direct

Phil. Trans. R. Soc. Lond. A (1993) **345**, 87–96
Printed in Great Britain

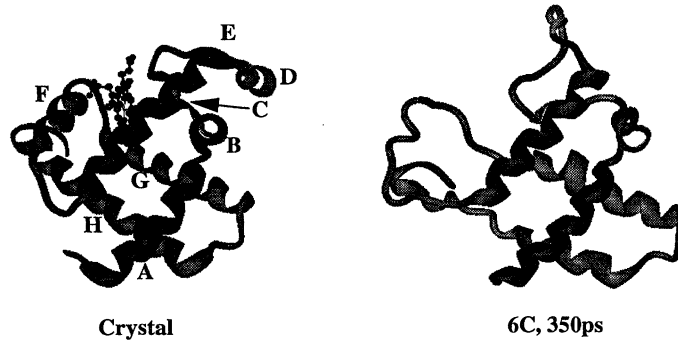

Figure 1. The crystal structure of sperm whale myoglobin (left) and the final structure from the molecular dynamics simulation of apomyoglobin at 25 °C (right).

interactions has been mounting slowly through CD (Tiffany & Krimm 1973), surface tension (Breslow & Guo 1990), and calorimetric measurements (Makhatadze & Privalov 1992).

These problems can also be pursued computationally via current simulation methodology, molecular dynamics (MD) and Monte Carlo statistical mechanics (MC), with atomic detail for both the solutes and water solvent. A summary of some of our initial efforts aimed at better understanding protein denaturation is presented here. In particular, pathways for thermal denaturation of apomyoglobin have been studied by MD, and possible complex formation for urea, guanidinium ion, and tetramethylammonium ion with aromatic hydrocarbons has been examined with MC free energy calculations.

2. Unfolding apomyoglobin to a molten globule

(a) *Computational procedure*

The MD simulations were carried out with the AMBER 3.0a program (Singh *et al.* 1986) using the AMBER/OPLS force field (Jorgensen & Tirado-Rives 1988) and the TIP3P model of water (Jorgensen *et al.* 1983). Hydrogens are explicitly included on heteroatoms and aromatic carbons (Jorgensen & Severance 1990). Full technical details have been described elsewhere (Tirado-Rives & Jorgensen 1993), though it is noted that bond lengths and H–H distances in water were kept fixed, the time step was 2 fs, and interactions were truncated at 9 Å using a residue-based cutoff.

Since an X-ray structure is not available for apomyoglobin, the initial coordinates were derived from the crystal structure of sperm whale myoglobin (Kuriyan *et al.* 1986). Three simulations were run after extensive preparation of the system through energy minimizations, addition of *ca.* 5350 TIP3P water molecules in a periodic cell (62 Å × 58 Å × 53 Å), addition of neutralizing chloride counterions, and a sequence of short MD runs that equilibrated the system and brought it to 25 °C and 1 atm. Two simulations were designed to correspond to pH 6 and were run at 25 °C (simulation 6C) and 85 °C (6H); in these cases, His-24 and His-119 were unprotonated (Cocco *et al.* 1992). The third simulation (4H), also run at 85 °C, was meant to mimic the pH 4 conditions, so all histidines were protonated and two extra chloride ions were equilibrated near the protein surface. The 6C simulation was then run for 350 ps to compute a structure for native apomyoglobin, while the simulations at 85 °C were extended to 500 ps to allow substantial unfolding.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

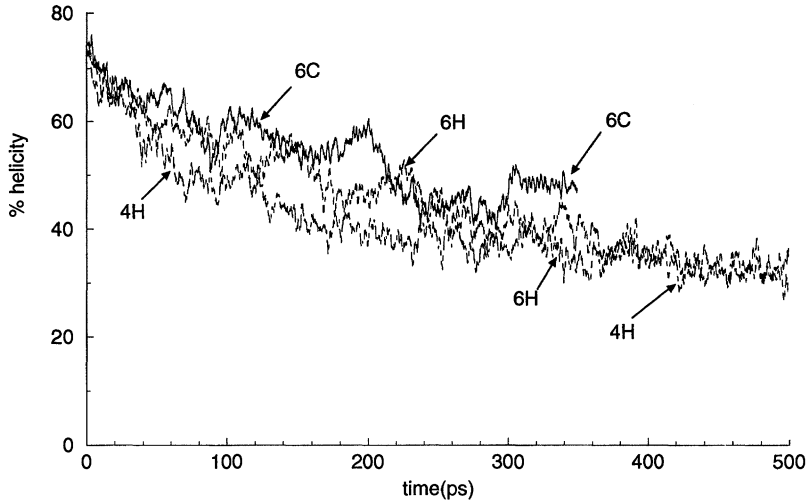

Figure 2. Helical content of apomyoglobin from the three molecular dynamics simulations as a function of time.

(*b*) *Structural results*

Numerous thermodynamic and structural indices were monitored during the simulations including the root mean square (RMS) deviation of the instantaneous structure for the backbone atoms from the initial myoglobin crystal coordinates and the percent helicity. The RMS deviations at the ends of the 6C, 6H, and 4H runs were 5.4, 6.2, and 9.1 Å respectively; the deviations are 4–5 Å after 200 ps in each case owing to the initial motions that are responding to the loss of the heme. The criterion of Daggett & Levitt (1992) defining helicity was adopted and depends on the backbone dihedral angles for three consecutive residues. The variations in helicity for the three runs are shown in figure 2.

Notably, the 6C run starts with the 75% helicity of myoglobin and levels off for the last 100 ps at 50% helicity, which agrees well with the CD estimate of 55% for native apomyoglobin (Hughson *et al.* 1990). The final structure from this simulation is illustrated in figure 1 in ribbon form. The apoprotein retains much of the helical character of the holoprotein including most of the A, E, G, and H helices and part of the B helix. One turn of the C helix remains sporadically, a gap occurs near the C-terminus of the H helix, and the short D and F helices completely decay. Experimentally, nearly all of the slowly exchangeable protons for the holoprotein were also found to have protection factors of 100 to over 100000 in apomyoglobin at pH 6 (Hughson *et al.* 1990). However, probes were only available at sites in the A, B, E, G and H helices at one site in the C helix. The computed structure now explains the observed, significant reduction in overall helicity primarily through loss of the unprobed D and F helices along with fraying at other helical ends, which is also apparent in diminished protection factors. Comparison of contact maps for the final structure of simulation 6C and the myoglobin crystal structure provides some information on the changes in tertiary structure. Naturally, the unfolded D and F helices lose many contacts with the rest of the protein. Also, the CD loop closes with helix G and especially the FG loop to fill the cavity left by the heme.

Turning to the simulations at elevated temperature, the percent helicity levels off over the last 100 ps in both cases to about 33%. This is close to the figure of 35%

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

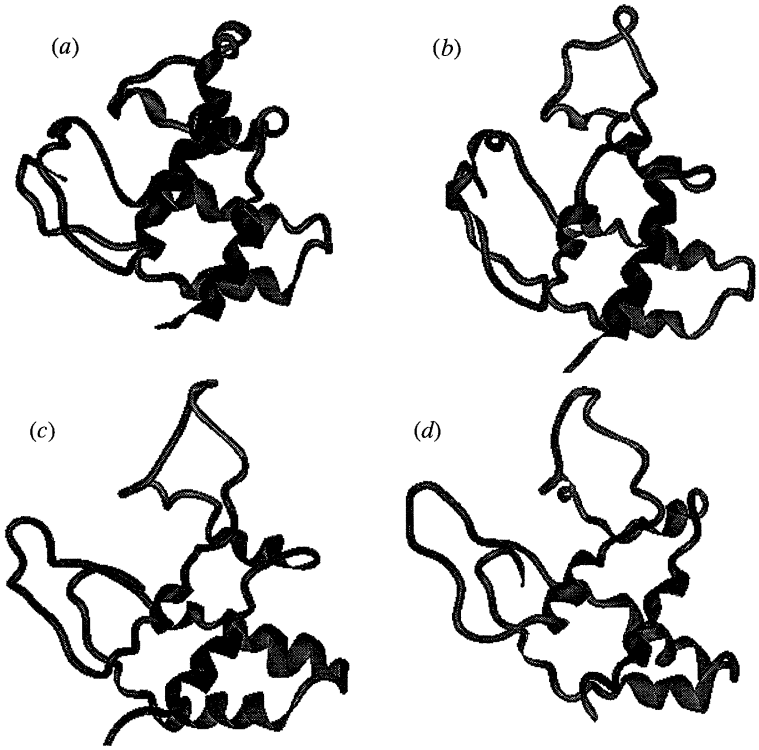

Figure 3. Four instantaneous structures from simulation 4H for apomyoglobin at 85 °C. $t = 100$ ps (*a*), 250 ps (*b*), 400 ps (*c*), 500 ps (*d*).

from the CD measurements of Hughson *et al.* (1990) for the I state. Furthermore, the properties of the average computed structures over the last 50 ps are notably similar to those attributed experimentally to molten globules (Christensen & Pain 1991). They are less compact than the folded proteins with radii of gyration *ca.* 10% larger than for the crystal structure. There is, of course, reduced secondary structure and the loss in backbone–backbone hydrogen bonding is largely offset by increased hydrogen bonding with sidechain atoms and water.

The structures from the two MD runs are qualitatively similar; four snapshots from the 4H simulation are shown in figure 3. The final structure only has significant fragments of the A, B, E, G and H helices intact, while the C, D and F helices are completely unfolded. Notably, the region near the junction of the A, G and H helices is the best preserved. These results largely parallel the observations from the hydrogen exchange experiments on the I state, in particular, only portions of the A, G and H helices and a single residue of helix B show protection factors of 10 or more (Hughson *et al.* 1990). It was proposed that the I state primarily retains helicity in the vicinity of the junction of the A, G and H helices with B–E unfolded. This is the case for the final structure in figure 3 with the additions that the unmonitored helix F is also unfolded and that 1–2 turns of helix B are preserved along with the C-terminal region of helix E which contacts G and H. The same basic picture emerges from simulation 6H, though helix F now retains 1–2 turns and helix B has decayed more.

The MD simulations appear to have led to detailed views of structures for native apomyoglobin and the low-pH folding intermediate, which are consistent with

almost all of the available experimental data on helix content and sites of protection. This opens an avenue for further interplay between computation and experiment on characterizing folding intermediates. A potential problem is that the observed molten globules appear to occur early in the folding process (Kuwajima 1989), while the simulations are providing structures that occur early in unfolding. The consistency between the experimental and computational views of the secondary structure for apomyoglobin is then, perhaps, surprising. Further work is needed to establish whether apomyoglobin is a special case and/or there are, in fact, substantial differences between reality and the computed model when comparisons go beyond the gross features of secondary structure. It should be noted that a 500 ps simulation of apomyoglobin has also been carried out by Brooks (1992) at 39 °C with CHARMM potential functions in TIP3P water; the helix structure is better preserved compared to myoglobin than from the present low-temperature simulation.

3. Interactions of chaotropes with benzene and naphthalene

(*a*) *Computational procedure*

As mentioned in the introduction, a key question is whether or not direct interactions occur between denaturing agents and protein constituents. Thermodynamically, this translates into determining if there are minima in free energy profiles or 'potentials of mean force' near contact for the approach of a chaotrope to another solute (Jorgensen 1991). This has been pursued initially for the interactions of urea, guanidinium ion, and tetramethylammonium ion (TMA) with benzene and for urea with napthalene. The potentials of mean force (PMFS) were computed using Monte Carlo simulations with statistical perturbation theory in a manner analogous to that previously utilized for benzene dimerization in water (Jorgensen & Severance 1990). The reaction coordinate is defined as the distance between molecular centres, which are taken to be the carbon of urea and guanidinium ion, the nitrogen in TMA, the ring centre of benzene, and the middle of the fusion bond (C9–C10) in naphthalene.

Some key details of the calculations should be noted. All-atom OPLS potential functions have been used for urea (Duffy *et al.* 1993*a*), guanidinium ion (Jorgensen & Tirado-Rives 1988), and benzene and naphthalene (Jorgensen & Severance 1990), while united-atom methyl groups were used in a five-site model for TMA (Jorgensen & Gao 1986; Buckner & Jorgensen 1989), and the TIP4P model was taken for water (Jorgensen *et al.* 1983). The individual molecules are rigid except the two-fold torsions for the amino groups in urea and guanidinium ion were included with the potential energy described by

$$V(\phi) = \tfrac{1}{2}V_2(1-\cos 2\phi), \tag{1}$$

where the values for V_2 were determined from ab initio 6-31G*//6-31G* calculations as 17.7 and 15.7 kcal mol^{-1} respectively (Duffy *et al.* 1993*b*). The two solutes were surrounded by 740 TIP4P water molecules in periodic cells with dimensions *ca.* 25 Å × 25 Å × 37.5 Å. Cutoffs of 9.0 Å and 12 Å were invoked for the water–water and water–solute interactions, based on the separation of the molecular centres, and the interactions were quadratically feathered to zero over the last 0.5 Å. The simulations were run in the isothermal isobaric (NPT) ensemble at 25 °C and 1 atm. Solute moves were attempted every 75 configurations and consisted of translating the two

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

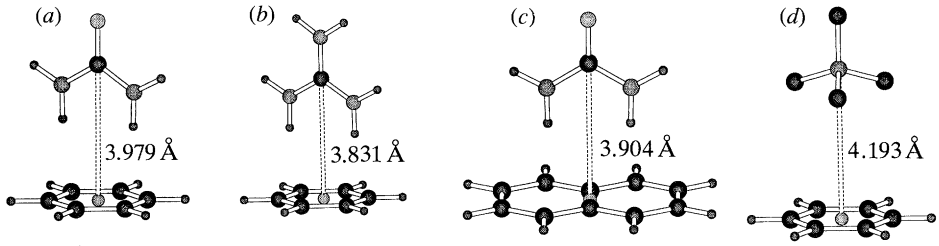

Figure 4. Optimized structures in the gas phase for the complexes of (*a*) urea with benzene ($\Delta E = -5.59$ kcal mol^{-1}), (*b*) guanidinium ion with benzene ($\Delta E = -9.42$ kcal mol^{-1}), (*c*) urea with naphthalene ($\Delta E = -7.74$ kcal mol^{-1}), and (*d*) tetramethylammonium ion with benzene ($\Delta E = -6.61$ kcal mol^{-1}).

molecules in tandem, rotating one of them at random about its molecular centre, and attempting any internal torsions of the amino groups. This procedure maintains the reaction coordinate value. The ranges for the solute, solvent, and volume moves were chosen to yield roughly 40% acceptance rates for new configurations. Reaction coordinate values from 3.0 to 8.0 Å were covered in steps of 0.125 Å with double-wide sampling; e.g. a simulation was performed at 3.125 Å with perturbations to 3.000 and 3.250 Å, then the next simulation was run at 3.375 Å with perturbations to 3.250 and 3.500 Å, etc. Thus, twenty full simulations were required to cover the 5 Å for each PMF. Each simulation consisted of 2.0×10^6 configurations of equilibration followed by 4.0×10^6 configurations of averaging. Many free energy increments were checked by running forwards and backwards simulations, extending the simulations, and by insisting that the standard deviations for the increments stay below 0.1 kcal mol^{-1}, as computed from separate averages over blocks of 2×10^5 configurations. All calculations were performed with the BOSS program (Jorgensen 1993).

MC simulations have advantages over MD for these calculations mostly stemming from the use of internal rather than cartesian coordinates. In our hands, the MC calculations are also easier to set-up and lead to less noise for computed free energy changes than we have obtained with MD for comparable amounts of computer time. Furthermore, the reason for the switch from the TIP3P to the somewhat preferable TIP4P model of water arises from difficulties in MD simulations of dealing with the negatively charged site in the TIP4P model, which is at a massless point near the oxygen.

(*b*) *Energetic and structural results*

For reference, the optimal structures for the complexes in the gas phase were obtained from the potential functions via gradient optimizations with BOSS (figure 4). The three complexes with urea and guanidinium ion have C_{2v} symmetry with amino hydrogens perpendicular to the ring planes. This is electrostatically reasonable in view of the partial negative charges on the carbons of aromatic CH units; analogous structures are found for benzene with a water molecule (Jorgensen & Severance 1990; Suzuki *et al.* 1992). The attraction for urea increases with increasing size of the arene from -5.6 kcal mol^{-1} for benzene to -7.8 kcal mol^{-1} for naphthalene. Not surprisingly, the optimal interaction with guanidinium ion is still stronger at -9.6 kcal mol^{-1} and the ring-centre – carbon separation is about 0.1 Å shorter at 3.83 Å. In view of the successful testing of the potential functions for many condensed phase properties of, particularly, urea and benzene (Jorgensen & Severance 1990; Duffy *et al.* 1993), the present results are expected to be reliable. A

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

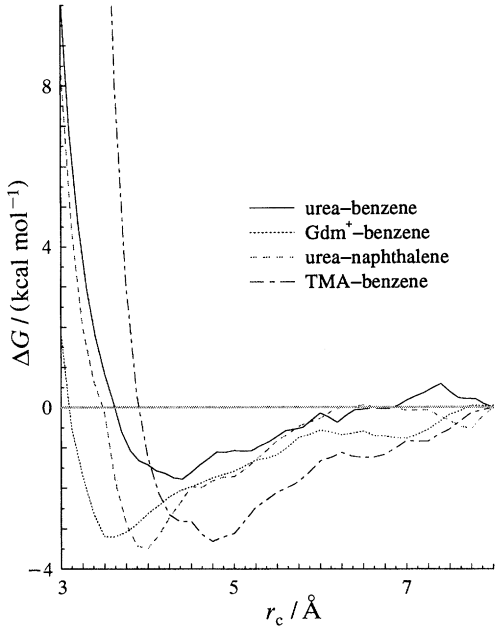

Figure 5. Computed free energy profiles as a function of intermolecular separation for the four complexes in water at 25 °C.

similar structure is obtained for the TMA-benzene complex with an interaction energy of −6.6 kcal mol^{-1} for this prototypical cation–π interaction. The potential functions for TMA (Jorgensen & Gao 1986) have not been as well tested as for urea. However, the optimal energy for the TMA-water complex from the potential functions, −9.5 kcal mol^{-1}, agrees well with an experimental interaction enthalpy of −9.0 kcal mol^{-1} from high-pressure mass spectroscopy (Meot-Ner & Deakyne 1985). On the other hand, an interaction enthalpy of −9.4 kcal mol^{-1} is reported in that same study for the TMA-benzene complex, significantly more attractive than the present value of −6.6 kcal mol^{-1}. Since the computed interaction energies for TMA with water and benzene with water, −2.3 kcal mol^{-1} (Jorgensen & Severance 1990), are essentially correct, it seems odd that the optimal interaction for benzene-TMA could be off by 3 kcal mol^{-1}.

The computed potentials of mean force for the four systems in water are shown in figure 5. The curves have been zeroed at the largest separation studied, 8.0 Å, since only relative free energies are obtained from the perturbation calculations. The PMFs except possibly for TMA-benzene appear to be flattening out beyond 6 Å, keeping in mind the intrinsic noise of *ca.* 0.05 kcal mol^{-1} in each increment. The overall uncertainty in the PMFs is estimated to be about 0.5 kcal mol^{-1}. The striking result is that in each case a single attractive well is obtained at separations corresponding to direct contact between the arene and additive.

For urea with benzene, the well-depth is −1.8 kcal mol^{-1} at 4.4 Å separation. This is a larger distance than the optimal one of 4.0 Å in figure 4. In fact, display of configurations from the simulation, as in figure 6, shows that the urea is typically oriented at the shorter separations with one syn NH hydrogen towards the benzene ring. Optimizations reveal that this orientation is also an energy minimum in the gas phase with an optimal separation of 4.29 Å and energy of −4.50 kcal mol^{-1}. The

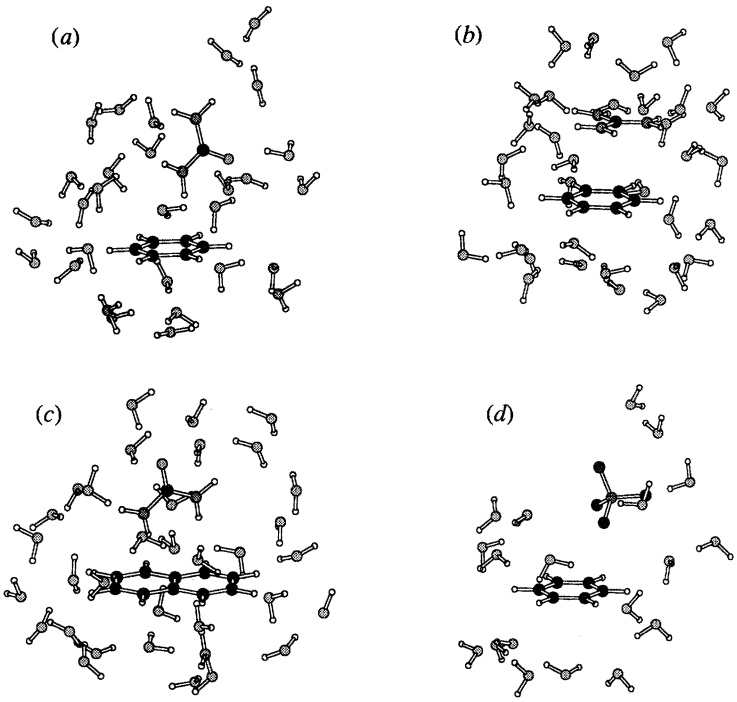

Figure 6. Snapshots from the Monte Carlo simulations for the four complexes in water at 25 °C. Only water molecules with an atom within 3.0 Å of a solute atom are shown. The complexes are as in figure 4*a–d* and the intermolecular separations correspond to the positions of the minima in figure 5. r_c = 4.40 Å (*a*), 3.50 Å (*b*), 4.00 Å (*c*), and 4.75 Å (*d*).

1 kcal mol^{-1} loss in the urea–benzene interaction is presumably compensated by better hydrogen bonding of the urea with water through the two anti NH groups (figure 6), which are no longer shielded by the benzene ring. An estimate of the association constant, K_a, for the complex can be obtained by integrating the PMF, $w(r)$, to a geometric limit for association, c, according to (2) (Jorgensen 1991). With c = 6, 7 and 8 Å, the computed K_a is 0.11, 0.38 and 0.65 M^{-1}

$$K_a = 4\pi \int_0^c r^2 \exp\left(-w(r)/k_B T\right) \mathrm{d}r. \qquad (2)$$

An experimental value of 0.16 M^{-1} can be inferred for urea with toluene from the data of Robinson & Jencks (1965), assuming that the increased solubility of toluene results entirely from formation of 1:1 complexes with urea. In any event, complexation of benzene and urea is clearly computed to occur, but to be weak.

The complexation of urea and napthalene is computed to be stronger in figure 5; the well-depth is now −3.5 kcal mol^{-1} at 4.0 Å separation. The position of the minimum is similar to the optimal distance for the complex in figure 4. Display of configurations from the simulations in the vicinity of the minimum do show frequent population of such geometries (figure 6). The stronger intrinsic napthalene–urea interaction is more competitive with the urea–water interactions than in the case of benzene. Experimentally, it is also known that urea has a greater solubilizing effect on naphthalene than toluene (Robinson & Jencks 1965; Roseman & Jencks 1975).

Guanidinium ion is also computed to form a complex with benzene; the well-depth in the free energy profile is −3.2 kcal mol^{-1} at 3.5 Å separation (figure 5). At such

distances, the simulations reveal a tendency for the guanidinium ion to lie parallel to the plane of the benzene ring (figure 6). In the gas phase, constrained optimizations show the parallel arrangement to have an optimal interaction energy of -4.72 kcal mol^{-1} at 3.59 Å separation. This is 5 kcal mol^{-1} weaker than for the global minimum in figure 4; however, the parallel geometry allows the guanidinium hydrogens to be unfettered for hydrogen bonding with water. The present results firmly support the proposal of Breslow & Guo (1990) that there is a direct interaction between guanidinium ion and benzene in water. The observed increased surface tension in aqueous solutions of guanidinium chloride impedes cavity formation and hinders solubilization of hydrocarbons. This effect is overcome by the salting-in contribution from the guanidinium–benzene complexation.

Finally, the TMA–benzene PMF reveals complex formation with a well-depth of -3.3 kcal mol^{-1} at 4.75 Å in figure 5. Cation–π interactions involving ammonium and guanidinium fragments with aromatic rings in amino acid sidechains and ligands are receiving increased attention as important contributors to binding (Burley & Petsko 1986; Dougherty & Stauffer 1990; Sussman *et al.* 1991; Waksman *et al.* 1992). The present calculations support this notion even in the absence of the lower dielectric constant environment of an enzyme's interior.

4. Conclusion

The present examples illustrate the ability of modern simulation methods to provide detailed insights into fundamental structural and energetic issues regarding protein denaturation. Great opportunities exist for related studies characterizing unfolding pathways and the structures of folding intermediates for different classes of proteins. Model studies of the interactions of chaotropes with polar functionality as in peptide backbones should also be considered, and eventually the unfolding of proteins in the presence of denaturants can be modelled. The complementarity of computations and experiment needs to be stressed. On the theoretical side, concern for the validity of the potential functions and for adequate configurational sampling requires vigilance and active pursuit of opportunities for quantitative comparisons with experimental data.

Gratitude is expressed to the National Science Foundation for support of this work and to Mr Dongchul Lim and Dr Paul J. Kowalczyk for computational assistance.

References

Breslow, R. & Guo, T. 1990 Surface tension measurements show that chaotropic salting-in denaturants are not just water-structure breakers. *Proc. natn. Acad. Sci. U.S.A.* **87**, 167–169.

Brooks, C. L. III 1992 Characterization of native apomyoglobin by molecular dynamics simulation. *J. mol. Biol.* **227**, 375–380.

Buckner, J. K. & Jorgensen, W. L. 1989 Energetics and hydration of the constituent ion pairs of tetramethylammonium chloride. *J. Am. chem. Soc.* **111**, 2507–2516.

Burley, S. K. & Petsko, G. A. 1986 Amino-aromatic interactions in proteins. *FEBS Lett.* **203**, 129–143.

Christensen, H. & Pain, R. H. 1991 Molten globule intermediates and protein folding. *Eur. Biophys. J.* **19**, 221–229.

Cocco, M. J., Kao, Y,-H., Phillips, A. T. & Lecomte, J. T. J. 1992 Structural comparison of apomyoglobin and metaquomyoglobin: pH titration of histidines by NMR spectroscopy. *Biochemistry* **31**, 6481–6491.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Daggett, V. & Levitt, M. 1992 Molecular dynamics simulations of helix denaturation. *J. mol. Biol.* **223**, 1121–1138.

Dougherty, D. A. & Stauffer, D. A. 1990 Acetylcholine binding by a synthetic receptor: implications for biological recognition. *Science, Wash.* **250**, 1558–1560.

Duffy, E. M., Severance, D. L. & Jorgensen, W. L. 1993*a* Urea: potential functions, log P, and free energy of hydration. *Isr. J. Chem.* **33**. (In the press.)

Duffy, E. M., Kowalczyk, P. J. & Jorgensen, W. L. 1993*b* Interaction of denaturants with aromatic hydrocarbons in water. *J. Am. chem. Soc.* **115**. (In the press.)

Hughson, F. M., Wright, P. E. & Baldwin, R. L. 1990 Structural characterization of a partly folded apomyoglobin intermediate. *Science, Wash.* **249**, 1544–1548.

Jorgensen, W. L. 1991 Computational insights on intermolecular interactions and binding in solution. *Chemtracts Org. Chem.* **4**, 91–119.

Jorgensen, W. L. 1993 BOSS version 3.4. Yale University, New Haven, Connecticut, U.S.A.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. 1983 Comparison of simple potential functions for simulating liquid water. *J. chem. Phys.* **79**, 926–935.

Jorgensen, W. L. & Gao, J. 1986 Monte Carlo simulations of the hydration of ammonium and carboxylate ions. *J. phys. Chem.* **90**, 2174–2182.

Jorgensen, W. L. & Severance, D. L. 1990 Aromatic-aromatic interactions: free energy profiles for the benzene dimer in water, chloroform, and liquid benzene. *J. Am. chem. Soc.* **112**, 4768–4774.

Jorgensen, W. L. & Tirado-Rives, J. 1988 The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. chem. Soc.* **110**, 1657–1666.

Kim, P. S. & Baldwin, R. L. 1990 Intermediates in the folding reactions of small proteins. *A. Rev. Biochem.* **59**, 631–660.

Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. A. 1986 X-ray structure and refinement of carbon-monoxy (Fe II)-myoglobin at 1.5 Å resolution. *J. mol. Biol.* **192**, 133–154.

Kuwajima, K. 1989 The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure. *Proteins Struct. Funct. Genet.* **6**, 87–103.

Makhatadze, G. I. & Privalov, P. L. 1992 Protein interactions with urea and guanidinium chloride, a calorimetry study. *J. mol. Biol.* **226**, 491–505.

Matouschek, A., Serrano, L., Meiering, E. M., Bycroft, M. & Fersht, A. R. 1992 The folding of an enzyme. V. H/^{2}H exchange – nuclear magnetic resonance studies on the folding pathway of barnase. *J. mol. Biol.* **224**, 837–845.

Meot-Ner, M. & Deakyne, C. A. 1985 Unconventional ionic hydrogen bonds. 1. $CH^{\delta+}\cdots X$. Complexes of quaternary ions with n- and π-donors. *J. Am. chem. Soc.* **107**, 469–474.

Roder, H., Elove, G. A. & Englander, S. W. 1988 Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature, Lond.* **335**, 700–704.

Roseman, M. & Jencks, W. P. 1975 Interactions of urea and other polar compounds in water. *J. Am. chem. Soc.* **97**, 631–640.

Singh, U. C., Weiner, P. K., Caldwell, J. & Kollman, P. A. 1986 AMBER 3.0a. University of California, San Francisco, U.S.A.

Sussman, J. L., Harel, M., Frolow, F. *et al.* 1991 Atomic structure of acetylcholinesterase from torpedo californica: prototypic acetylcholine-binding protein. *Science, Wash.* **253**, 872–879.

Suzuki, S., Green, P. G., Bumgarner, R. E., Dasgupta, S., Goddard, W. A. III & Blake, G. A. 1992 Benzene forms hydrogen bonds with water. *Science, Wash.* **257**, 942–945.

Tiffany, M. L. & Krimm, S. 1973 Extended conformations of polypeptides and proteins in urea and guanidine hydrochloride. *Biopolymers* **12**, 575–587.

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

Waksman, G., Kominos, D., Robertson, S. C. *et al.* 1992 Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature, Lond.* **358**, 646–653.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES