J. Mol. Biol. (1995) **252**, 672–708

JMB

Acid and Thermal Denaturation of Barnase Investigated by Molecular Dynamics Simulations

Amedeo Caflisch and Martin Karplus*

Department of Chemistry
Harvard University
Cambridge, MA 02138, USA

**Corresponding author*

The transition in barnase from the native state to a partially unfolded conformation has been studied by molecular dynamics simulations with explicit water molecules at 360 K and low pH(450 ps), and at 600 K and neutral pH (three simulations of 120, 250 and 200 ps each). The use of several simulations provides evidence that the results are not sensitive to initial conditions. To mimic low pH conditions, the acidic sidechains in barnase were neutralized and the two histidine residues were doubly protonated. Runs at 300 K showed that the solvated structures at low pH (300 ps) and neutral pH (310 ps) are very similar. The main structural differences involved the acidic residues, histidine residues, and the β-turn connecting strands 4 and 5. When the temperature is raised to 360 K at low pH and to 600 K at neutral pH the barnase molecule begins to unfold. The molecule rapidly expands (R_g changes from 13.9 Å to 15.3 Å in 450 ps at 360 K and from 13.7 Å to between 15.1 and 15.5 Å in 120 ps at 600 K). However, the expansion is not uniform. In all the simulations, the chain termini, loops and the N-terminal parts of the main α-helix (helix 1) show a continuous and progressive unfolding. An essential step in the denaturation process is that the major α-helix (helix 1) separates from the β-sheet; this is coupled to the exposure of the principal hydrophobic core, many of whose non-polar sidechains become solvated by hydrogen-bonded water molecules. The barnase-water interaction energy improves during unfolding at the expense of the barnase self-energy. The deterioration of the intramolecular van der Waals energy suggests that the rupture of the tight packing during the initial unfolding phase contributes to the energy barrier of the denaturation process. The mutationally well-analyzed Asp8-Arg110-Asp12 double salt-bridge on the barnase surface is found to be marginally stable in the folded form in the simulations. A Poisson-Boltzmann calculation indicates that the salt-bridge is unstable; this is probably due to an overestimate of the solvation energy.

A detailed analysis of the main hydrophobic core reveals that increase in solvent-accessible surface area and penetration of water molecules are simultaneous in the high-temperature simulation; at lower temperatures there is significant cavity formation and the entrance of the water molecules is somewhat delayed. The cavities occur in the neighborhood of the hydrophobic sidechains; the region formed by the sidechains of Val10, Leu14, Leu20, Tyr24, Ala74, Ile76 and Tyr90 is involved. The loosening of the core packing is coupled to an increase in the number of dihedral transitions.

Water molecules connected in chains or clusters penetrate the cores by participating in hydrogen bonds with polar groups; e.g. on tyrosine and tryptophan sidechains. The hydrogen-bonding propensity of the water molecules tends to be satisfied throughout the denaturation process. This is evident also in the denaturation of the secondary structural elements, where water molecules compete with the interstrand and intrahelical hydrogen bonds. In all simulations, the β-sheet disruption starts at the edges and is coupled to an increase in the twist. In agreement with experimental results, α-helix 1 starts to denature at the N-terminal end. Helices 2 and 3 undergo early unfolding at 600 K and neutral pH, while they are stable at 360 K and low pH. These differences arise from the greater entropic contribution to the unfolded state at high temperature.

0022–2836/95/400672–37 $12.00/0

The overall similarity between the low-pH and high-temperature simulations indicates that the present results are representative of the barnase unfolding process. Experimental tests of the role of the solvent in barnase unfolding are proposed.

Keywords: barnase; acid denaturation; thermal denaturation; molecular dynamics; finite difference Poisson-Boltzmann equation

Introduction

The mechanism of protein folding and unfolding is among the major unsolved problems of biology. There has been experimental work in this area for many years, but detailed structural information has begun to appear only rather recently (for reviews, see Creighton, 1990; Jaenicke, 1991). Of particular importance are NMR measurements and mutation experiments. Because of the complexity of the problem, theoretical studies and computer simulations are necessary to supplement the experimental data. Reviews of recent work have been published by Karplus & Shakhnovich (1992) and by Caflisch & Karplus (1994b). Although simplified models can play an important role in obtaining general ideas concerning possible folding mechanisms (Levitt & Warshel, 1975; Šali *et al*., 1994), all-atom simulations are important for a more detailed understanding. This is true particularly for an elucidation of the role of the solvent in protein folding and unfolding. A number of high-temperature all-atom molecular dynamics simulations of protein denaturation exist. Fan *et al*. (1991) did a high-temperature vacuum simulation to study the molten globule state of α-lactalbumin. They found that constraining the α-helices does not stabilize the hydrophobic core, whereas constraining the hydrophobic core does stabilize the α-helices. Daggett & Levitt (1992, 1993) investigated the unfolding of the reduced bovine pancreatic trypsin inhibitor (BPTI) by high-temperature molecular dynamics in the presence of explicit water molecules. The authors used a low water density in their constant-volume, high-temperature simulations, and concluded that the observed BPTI expansion was not related to solvent penetration. A simulation of the thermal denaturation of lysozyme with explicit solvent has been reported by Mark & van Gunsteren (1992). No difference between the unfolding of the two lysozyme domains was observed in the simulation; this disagrees with experiment. From the number of helical mainchain hydrogen bonds as a function of time, the authors concluded that during unfolding the disruption of secondary structure takes place while the protein molecule is still highly compact. No discussion of the role of water was given. Molecular dynamics simulations of apomyoglobin in aqueous solution were carried out by Brooks (1992) at 312 K, and by Tirado-Rives and Jorgensen (1993) at 298 K and at 358 K (pH values of 6 and 4 were used at the higher temperature). In both simulations helices A, E, G and H were stable, while the other helices showed different behavior. Brooks (1992) found that helices B, C, D and F also existed during the entire simulation, with B as stable as A, E, G and H, while C, D and F were more mobile and moved into the heme cavity. By contrast, Tirado-Rives & Jorgensen (1993) found that helix B unfolded in part, and helices C, D and F unfolded completely in a relatively short time. This difference is not surprising, since more destabilizing conditions were used in the latter simulations. The differential stability of β-sheets and α-helices upon thermal denaturation of β-lactamase has been investigated by Vijayakumar *et al*. (1993) in a molecular dynamics simulation at 600 K.

High-temperature molecular dynamics simulations (600 K) were performed recently (Caflisch & Karplus, 1994a) to study the unfolding of barnase, a 110 amino acid residue ribonuclease excreted from *Bacillus amyloliquefaciens*, for which a large body of experimental work on folding and unfolding is available from Fersht and co-workers (Fersht, 1993). Here, we report more detailed results on several high-temperature simulations and complement them with simulations of the denaturation at 360 K under acid conditions.

Barnase was chosen for study because it is a monomeric enzyme that is of particular interest as a folding model. Both its crystal structure (Mauguen *et al*., 1982; Baudet & Janin, 1991) and its structure in solution are known (Bycroft *et al*., 1991). Barnase consists of three α-helices and a five-stranded β-sheet that are stabilized by three hydrophobic cores (Figure 1). There are no disulfide bonds to constrain the unfolded state and the three proline residues are all *trans* in the native conformation. The transition state and pathway of barnase folding and unfolding have been investigated in detail by protein engineering (Fersht *et al*., 1992a; Serrano *et al*., 1992a,b; Matouschek *et al*., 1992a; Meiering *et al*., 1992) and NMR hydrogen-exchange trapping experiments (Bycroft *et al*., 1990; Matouschek *et al*., 1992b). The two techniques, which yield complementary results (Serrano *et al*., 1992c; Fersht *et al*.,

---

Present address: A. Caflisch, Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland.

Abbreviations used: BPTI, bovine pancreatic trypsin inhibitor; LPB, linearized Poisson-Boltzmann; RMSD, root-mean-square deviation.

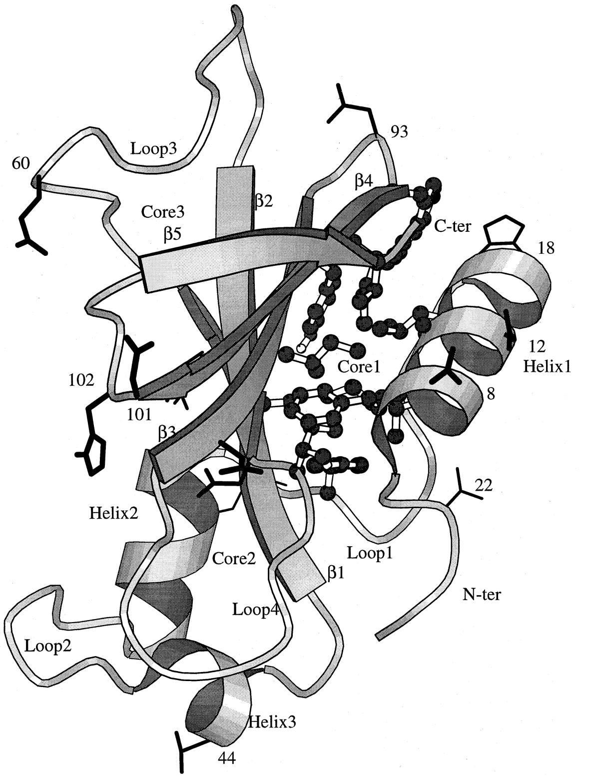


Figure 1. A representation of the backbone atoms of barnase, emphasizing the secondary structural elements. Sidechains of hydrophobic core 1 are plotted in a ball and stick representation with thin lines. The sidechains that are affected by lowering the pH are shown in a stick representation with depth-dependent thickness; they are: Asp8, Asp12, His18, Asp22, Glu29, Asp44, Asp54, Glu60, Glu73, Asp75, Asp86, Asp93, Asp101 and His102. The secondary structural elements include the following residues: N terminus (1 to 5), helix 1 (6 to 18), loop 1 (19 to 25), helix 2 (26 to 34), loop 2 (35 to 40), helix 3 (41 to 45), type II β-turn (46 to 49), strand 1 (50 to 55), loop 3 (56 to 69), strand 2 (70 to 75), loop 4 (76 to 84), strand 3 (85 to 90), type I β-turn (91 to 94), strand 4 (95 to 100), type III' β-turn (101 to 104), strand 5 (105 to 108), C terminus (109 to 110). The figure was made with the MOLSCRIPT program (Kraulis, 1991).

1992b), have shown that the rate-determining step for both folding and unfolding involves the crossing of a free energy barrier near the native state. Recently, a detailed analysis of the NMR spectrum of acid-denatured barnase has been published (Arcus *et al*., 1994).

Based on the high-temperature simulation, a mechanism was suggested for the solvation of the main hydrophobic core and for the dissolution of secondary structural elements (Caflisch & Karplus, 1994a). The main α-helix moved away from the β-sheet and exposed the principal hydrophobic core, many of whose non-polar sidechains, beginning with those near the surface, became solvated by hydrogen-bonded water molecules. Simultaneously, partial destruction of the α-helices and the outer strands of the β-sheet was observed with water molecules replacing the hydrogen bonds of the secondary structural elements.

In the previous study, barnase unfolding at neutral pH was accelerated by using a 600 K temperature; the activation free energy at 25°C and pH 6.3 is equal to about 20 kcal / mol (Matouschek *et al*., 1990). The free energy of unfolding of barnase at pH 3 and 298 K is about 4 kcal / mol (Pace *et al*., 1992), relative to the value of about 9 kcal / mol at pH 6. This suggests that the activation energy for barnase unfolding is significantly smaller at low pH than at neutral pH. Hence, a low-pH environment would be expected to promote rapid barnase unfolding at a lower simulation temperature than that used at neutral pH. To mimic low-pH conditions, the nine Asp and three Glu were protonated and the two His were doubly protonated; standard CHARMM parameters (Brooks *et al*., 1983) were used for these forms. With 14 Lys and Arg, and the positively charged N terminus and neutral C terminus, this leads to a net charge of +17, in contrast to the charge of +2 at neutral pH. In fact, the initial stages of the acid-mediated denaturation process were observed during a 450 ps simulation time at 360 K after 600 ps at 300 K and 327 K; the structural changes were small during the first 600 ps at low temperature.

The low-temperature, acid-mediated simulation makes it possible to compare the denaturation behavior under very different conditions (pH 7 and 600 K *versus* pH 3 and 360 K). Overall, the results concerning the protein and solvent behavior are very similar in the two sets of simulations, although certain specific differences are found. Moreover, the slower time-scale of the acid-mediated denaturation at 360 K led to a larger separation of the times of certain events, which allowed a more detailed investigation of the initial stages of the cooperative unfolding mechanism. Several high-temperature simulations were also performed to test the sensitivity of the results to the initial conditions.

The results for all of the simulations and their analysis are presented below. First, the overall behavior of the protein and its energetics during the unfolding process are presented. This is followed by an analysis of the behavior of the hydrophobic cores and the secondary structural elements during denaturation. The role of the solvent is examined in some detail. Similarities and differences of the low and high-temperature denaturation are analyzed. The results provide information that can be compared with experimental data.

Results and Analysis

We first examine the low-pH and the neutral-pH simulations at 300 K. We then describe the global behavior of barnase during the initial stages of the unfolding process. The energetic changes that occur on denaturation are also presented. Finally, details are given concerning the unfolding of the hydrophobic cores and the secondary structural elements with emphasis on the role of water. Table 1 lists the simulations that were done.

Table 1. Simulations performed

	Started from	Temperature (K)	pH	Length (ps)	Name
Control run	Crystal	300	Neutral	310	NpH300
Acid denaturation	Minimized crystal	300–360[a]	Low	1050	LpH300–LpH360
Thermal denat 1	4 ps NpH300	600	Neutral	120	A600
Thermal denat 2	100 ps NpH300	600	Neutral	250	R600
Thermal denat 3	258 ps NpH300	600	Neutral	200	S600
Cooled A600	90 ps A600	300	Neutral	160	B300

[a] The acid denaturation was run at 300 K from 1 to 300 ps (LpH300), at 327 K from 301 to 600 ps, and at 360 K from 601 to 1050 ps (LpH360).

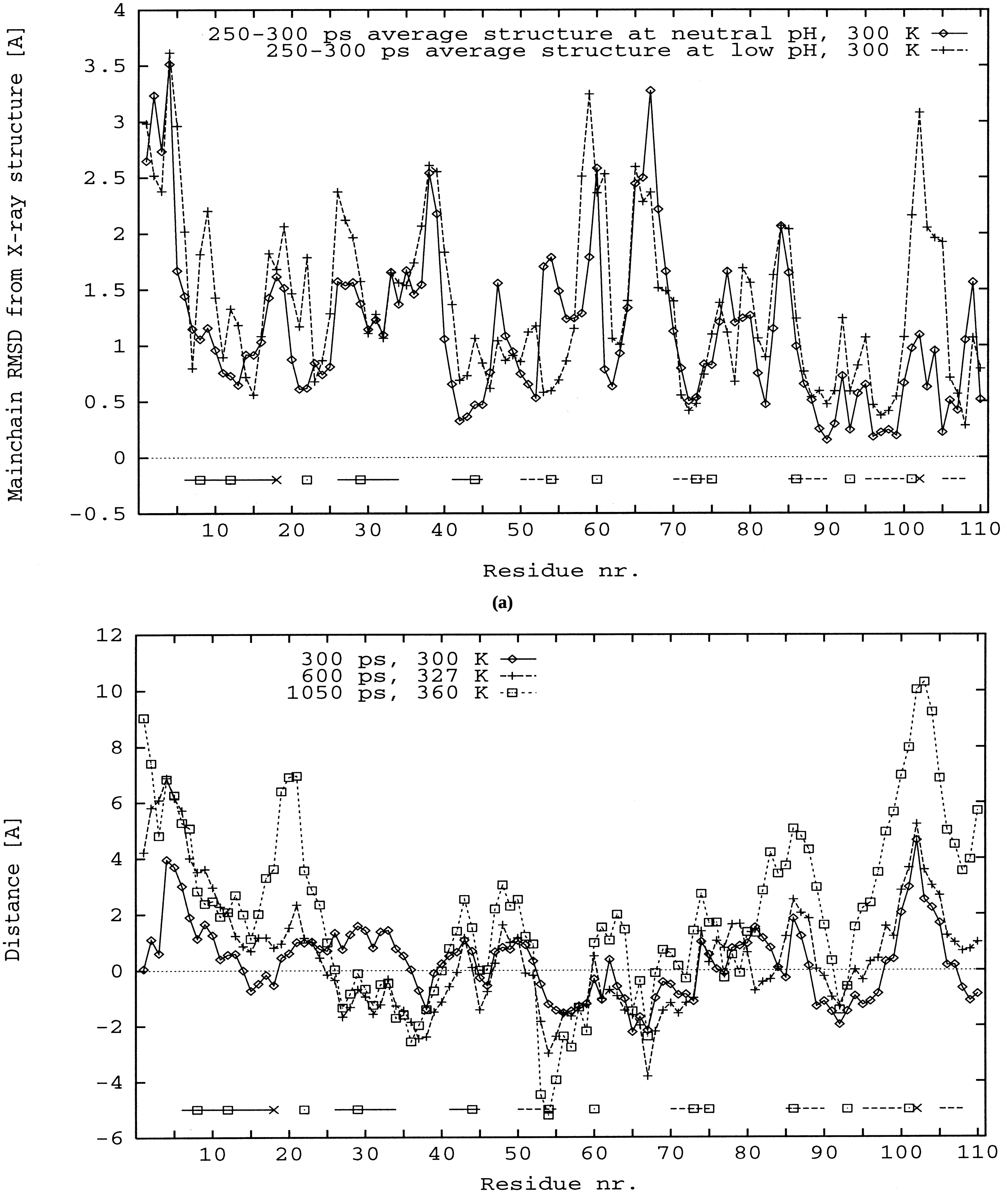
Room temperature simulations

To provide controls and to analyze the effect of pH on the barnase structure and dynamics, the 250 to 300 ps average structure at low pH and 300 K (LpH 300), and the 250 to 300 ps average structure of the control run at neutral pH and 300 K (NpH 300) were calculated and minimized (Figure 2). Both simulations appear to be stable over the time-range of the simulations. This can be seen from the behavior of the radius of gyration (Figure 3(a) and (c)) and the RMSD (Figure 3(b) and (d)) relative to the X-ray structure. The LpH 300 simulation is stable over the 300 ps, even though experimentally it is unstable. This indicates that the barrier to unfolding is such that the unfolding rate at 300 K is slower than the simulation time. For NpH 300, corresponding behavior is observed over the simulation time of 310 ps. Here the system is stable experimentally.

It is of interest to note that core 2 contains four water molecules within 7 Å of the core center after 200 ps of the (NpH 300) simulation. These core water molecules have positions that correspond to those observed in the X-ray structure. This is an important result, since the simulation was started without any crystal water by immersing the X-ray structure of barnase (after deletion of the crystal water molecules) in a sphere of equilibrated solvent. No other water molecule enters the core, again in agreement with the X-ray results. This indicates that the balance of protein-protein and protein-solvent interactions is correct. The N terminus, loop 2 and loop 3 show the largest mainchain deviation from the X-ray structure in both simulations (Figure 2(a)). Overall, these averaged structures show similar mainchain deviation from the crystal structure, apart from the β-turn connecting strands 4 and 5, which deviates much more in LpH300. This turn involves the doubly protonated His102 and the neutralized Asp101 at low pH. The same β-turn shows the largest displacement from the center in the LpH 360 simulation (see Figure 2(b)). This result concerning the effect of a charged His on the loop structure is similar to that observed in an NMR study of interleukin 4 at low pH (C. M. Dobson, personal communication).

Most of the sidechains of residues 1 to 15 and 30 to 80 show similar deviations from the X-ray structure in the two 300 K simulations. The LpH 300 structure is characterized by larger deviations at residues 16 to 29 and at the C-terminal part of the chain (residues 80 to 110). For every sidechain, the differential RMSD from the X-ray structure (i.e. the sidechain RMSD of the LpH 300 structure substracted from the corresponding value for the NpH 300 average structure) was calculated. As expected, among the six sidechains whose differential sidechain RMSD is larger than 1.9 Å there are two Glu (29 and 60), one Asp (54) and the two His sidechains (18 and 102). The largest differential RMSD is found for Glu29 (2.5 Å). The Arg59 sidechain has the second largest differential RMSD (2.4 Å), which originates from its vicinity to Glu60. Somewhat surprisingly, Asp54 (see also below) and Glu60 sidechains in LpH 300 show a smaller RMSD from the X-ray structure than in NpH 300.

The difference in solvent-accessible surface area between corresponding sidechains in the LpH 300 and NpH 300 structures was also calculated. The largest difference is found for the His102 sidechain (on the β-turn connecting strands 4 and 5, Figure 1), which has 83 Å^2 additional exposed surface in LpH 300. The remaining ten sidechains with a differential exposed surface larger than 20 Å^2 are: Gln2, −38 Å^2; Asp54, −35 Å^2; Lys108, −29 Å^2; His18, 29 Å^2; Ser85, −28 Å^2; Trp94, −24 Å^2; Ile109, −24 Å^2; Thr6, 23 Å^2; Lys27, −22 Å^2; Arg59, 20 Å^2 (a negative value means a larger exposed surface in NpH 300). The values for the His18, His102 and Asp54 sidechains are consistent with the fact that adding a net unit charge on a sidechain (His at low pH) yields a preference for the exterior of a protein, whereas removing a net unit charge (Asp and Glu at low pH) is likely to result in increased burial of the sidechain. Visual analysis of the NpH 300 and LpH 300 structures by QUANTA (Molecular Simulations, Inc.) confirms that the His102 side-chain is pointing towards the solvent at low pH, whereas at neutral pH its imidazole ring occupies a pocket formed by the Tyr103 aromatic ring, the C^{α} and C^{β} atoms of Ser85 and the guanidinium group of Arg87 (data not shown). At low pH, the His18 sidechain is also almost completely solvated and a water molecule bridges its N^{δ} hydrogen atom to the NH of the Trp94 sidechain through hydrogen bonds. The His18 sidechain acts as C-cap for helix 1 at neutral pH, so that its exposed surface is slightly reduced. The stabilizing effect of the histidine-aromatic interactions observed by Fersht and co-workers (Loewenthal *et al*., 1992) are not present in the simulation because of the lack of



(a)

(b)

Figure 2. Simulation properties as a function of residue number. (a) Mainchain heavy-atom RMSD of the 250 to 300 ps average structure from the X-ray structure. Control run at 300 K (NpH 300; diamonds and continuous line); low-pH simulation at 300 K (LpH 300; crosses and broken line). The X-ray structure does not determine the positions of the N-terminal residues 1 to 3; they were added by modeling. At the bottom of the plot, the continuous segments represent α-helices, the broken segments β-sheet strands, the boxes Asp and Glu positions, and the crosses His positions. (b) Distance of C^{α} atoms from the origin in the low-pH simulations relative to the distance in the X-ray structure: at 300 ps and 300 K (diamonds and continuous line), at 600 ps and 327 K (crosses and broken line), and at 1050 ps and 360 K (boxes and broken line). The origin corresponds to the center of the X-ray structure, defined as the average value of all atomic coordinates of the polar hydrogen structure.

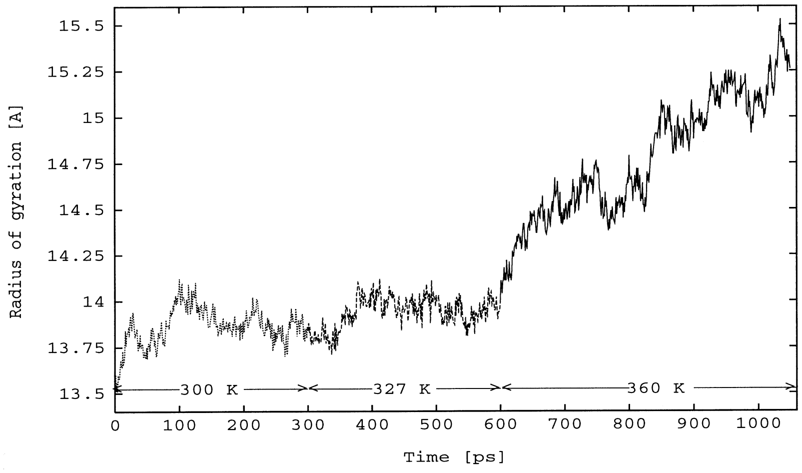
the electrostatic quadrupole moment on the aromatic rings in the polar hydrogen model. At low pH, the Asp54 sidechain CO acts as hydrogen bond acceptor for the Glu73 sidechain OH, whereas it protrudes into the solvent at neutral pH, because of the repulsion between these two negatively charged sidechains.

Summarizing the above, we note that the comparison of the LpH 300 and NpH 300 simulation at room temperature shows small but significant structural differences that can be related to the change in the charge states of certain residues. The most important alterations are near the C-terminal end, involving His102. It would be interesting to have a low-pH X-ray or NMR structure to confirm these results. That the NpH 300 structure is stable and remains close to the X-ray structure over 300 ps provides evidence that the potential function used is adequate. Further, the lack of drastic disruptions in the structure of LpH 300 suggests that there is still an activation barrier to unfolding, even though the native structure may be thermodynamically unstable; i.e. longer simulations would be required to observe the denaturation. An NMR analysis of the denatured low-pH (pH2) structure has been given recently (Arcus *et al*., 1994). However, as yet, the structural changes in the native state as a function of pH have not been determined.

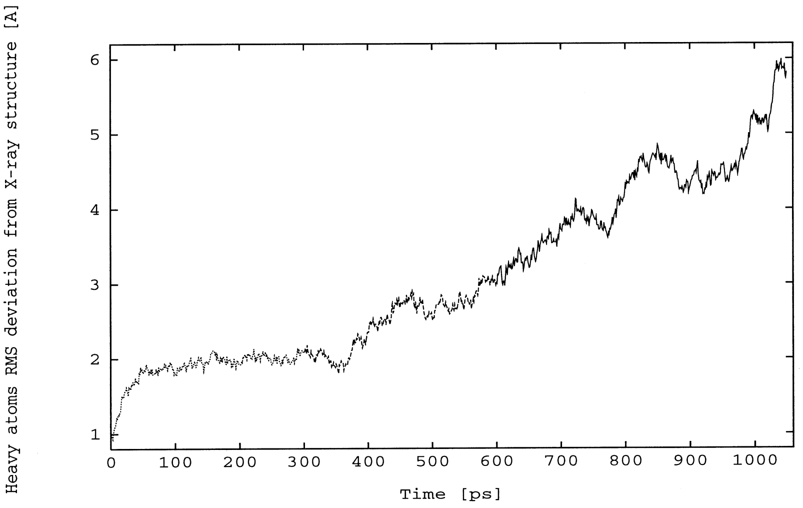
Global unfolding behavior

Since some structural results for the high-temperature simulation have been presented (Caflisch & Karplus, 1994a,b), we here focus on the low-pH, moderate-temperature unfolding simulation (LpH 360). The radius of gyration (R_g) and heavy-atom RMSD from the X-ray structure as a function of time are shown in Figure 3 for all of the simulations (see Table 1).

Only the early stages of the unfolding process were simulated in LpH 360 because the denaturation process is slower at low temperature. In R600, which was run for 250 ps at 600 K, the barnase molecule reached a higher degree of unfolding, as is evident from the R_g and RMSD values (Figure 3(c) and (d)).

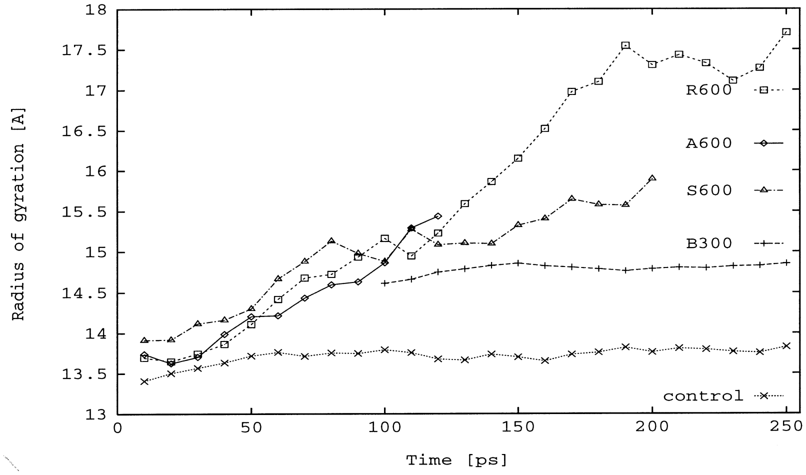


(a)

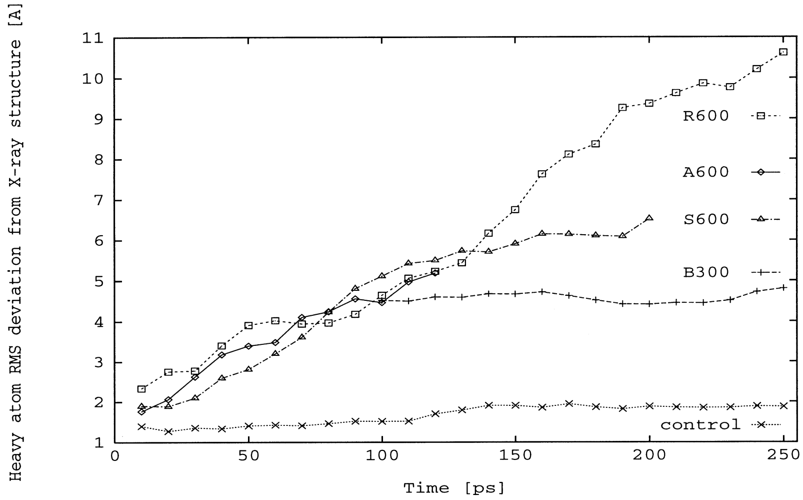


(b)

Figure 3(a–b) (*legend opposite*)



(c)



(d)

Figure 3. Global properties from simulations. (a) Radius of gyration (R_g) plotted as a function of time in the low-pH simulation (LpH 360). Dotted line, 300 K (1 to 300 ps); broken line, 327 K (301 to 600 ps); continuous line, 360 K (601 to 1050 ps). (b) Heavy-atom RMSD from the X-ray structure as a function of time for LpH 360. Line styles are the same as in (a). (c) R_g as a function of simulation time averaged over 10 ps intervals: (+) B300; (◇) A600; (□) R600; (△) S600; (×) control run at 300 K. (d) RMSD from the X-ray structure as a function of simulation time averaged over 10 ps intervals. Symbol are the same as in (c).

At 1050 ps, which corresponds to 450 ps at 360 K, the R_g value of LpH 360 is about equal to that found after 110 ps at 600 K. A600 and S600 ended with an intermediate degree of unfolding.

In the low-pH simulation at 300 K, after an initial small swelling during the first 100 ps, R_g decreases and has an average value of 13.9 Å from 200 to 300 ps (the R_g of the X-ray structure is 13.6 Å). At 327 K, there is a small increase from 350 to 400 ps and then R_g stabilizes at about 14.0 Å from 400 to 600 ps (Figure 3(a)). When the temperature is raised to 360 K at 600 ps (LpH 360), the R_g value increases over most of the trajectory, although the increase seems to proceed in stages, with R_g relatively constant for certain periods. The heavy-atom RMSD from the X-ray structure is about 2.0 Å from 100 to 300 ps (Figure 3(b)). After the temperature is raised to 327 K, it increases over most of the simulation to a final value of about 6 Å after 1050 ps. There are two transient periods (730 to 780 ps and 850 to 890 ps) where there is a small decrease in R_g and RMSD. During this time the hydrophobic core 2 (730 to 780 ps) and core 3 (850 to 890 ps) underwent a partial contraction and water was expelled (see below). The RMSD has an almost constant value of 4.4 Å during the 870 to 970 ps period when R_g remains essentially constant.

The present results confirm the observation made in the 600 K simulations that the RMSD begins to increase earlier than R_g. Here, the former starts at 327 K when there is no increase in R_g, and there also is a lag of R_g *versus* RMSD later in the simulation. One possible explanation would be that the barnase polypeptide chain is searching for a conformation in the neighborhood of the native structure that will permit R_g to increase and denaturation to proceed.

This corresponds to the expected behavior for finding the transition region in a high-dimensional configuration space (Elber & Karplus, 1990; Becker & Karplus, unpublished results).

To further explore the nature of the structural changes, we examine the behavior of LpH 360 on a per residue basis. It is clear that the expansion is not uniform (Figure 2(b)). The N and C-terminal segments and loop 1 undergo the largest expansion, whereas strand 1 and part of loop 3 actually move towards the barnase center of geometry. The expansion behavior is sustained during the entire simulation for most of the chain, apart from helix 2 and the C terminus. Helix 1 moves away from the rest of the protein during most of the simulation, though it partially repacks its N-terminal turn in the last 200 ps. Helix 2 undergoes a separating motion at 300 K and contracts at 327 K and 360 K, whereas the C terminus experiences a slight contraction at 300 K and expands at 327 K and 360 K (Figure 2(b)). A similar non-uniform expansion is seen in the 600 K simulations (Caflisch & Karplus, 1994a,b).

The C^{α} trace of four representative snapshots from the LpH 360 dynamics is shown in Figure 4(a). The structure at 600 ps and 327 K is similar to the X-ray structure, the main differences being in the N terminus, loop 2, the first half of loop 3 (residues 71 to 75) and the β-turn connecting strands 4 and 5 (see above). When the temperature is increased to 360 K, the termini, loops and the N-terminal part of helix 1 show a continuous and progressive unfolding. Strands 4 and 5, and the β-turn connecting them, separate from the rest of the β-sheet. The sliding motion of loop 3, relative to the β-turn connecting strands 4 and 5, is coupled to the movement of strands 1 and 2, and strands 4 and 5, which results in a slight increase in the β-sheet twist. The main helix (helix 1) has partially separated from the β-sheet at 800 ps. The unfolded N-terminal part then repacks before the end of the simulation. The accessibility of core 1 to water (see below) is coupled with the relative movement of helix 1 and the β-sheet. The separating motion of loop 1 (which contributes the Leu20 and Tyr24 sidechains to core 1), the denaturation of the N-terminal part of helix 1 (Phe7) and the denaturation of the protein C terminus (Ile109) contribute significantly to the unfolding of the main hydrophobic core. Helix 2 starts to partially unfold at the N terminus after the first 620 ps, while helix 3 is stable.

In the 600 K simulations the N terminus, loop 1 and loop 2 began to unfold during the first 30 ps. This was followed by partial denaturation of the hydrophobic cores; core 2 unfolded relatively rapidly, followed by core 1 and core 3 in A600; in R600 and S600 the expansion of the three cores was almost simultaneous. The solvation of hydrophobic core 1 was coupled with a significant distortion of helix 1 and of the edge strands of the β-sheet. Both helix 1 and helix 2 lost about half of the native α-helical hydrogen bonds; helix 3 unfolded after about 20 ps in A600 and R600 and after about 150 ps in S600. The behavior of helices 2 and 3 represents the main difference between the low-pH simulation and the 600 K simulations (Caflisch & Karplus, 1994a); i.e. at low pH, as mentioned above, they are almost stable, whereas in the 600 K runs they unfold almost completely in the first 120 ps. A representation of the denaturation behavior for the S600 simulation is shown in Figure 4(b).

As already noted in the high-temperature simulations (Caflisch & Karplus, 1994a), the edge strands of the β-sheet are less stable than the central strands during the low-pH simulation. This is in agreement with experiments on barnase (Matouschek *et al*., 1992a) and with NMR studies of β-sheets in variable antibody fragments (Constantine *et al*., 1992; Freund *et al*., 1994).

Energetics of denaturation

To obtain a better understanding of the denaturation process, we have examined the behavior of the energy terms as a function of time. We consider only

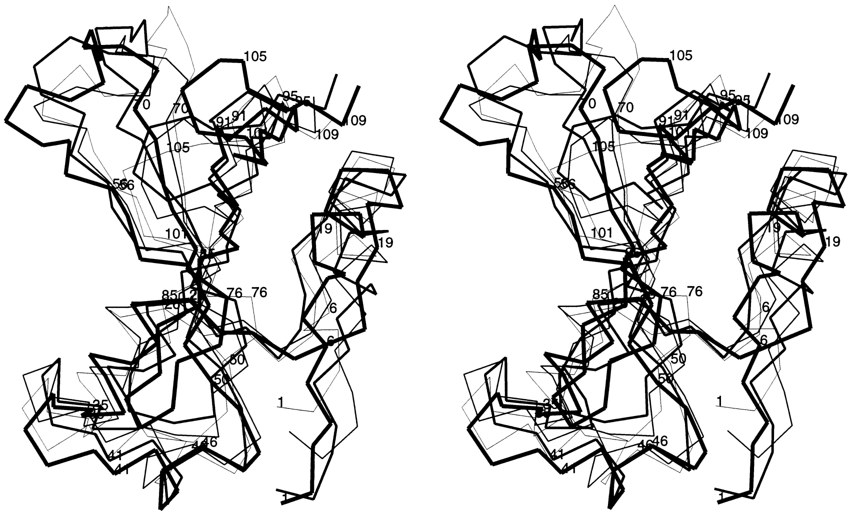
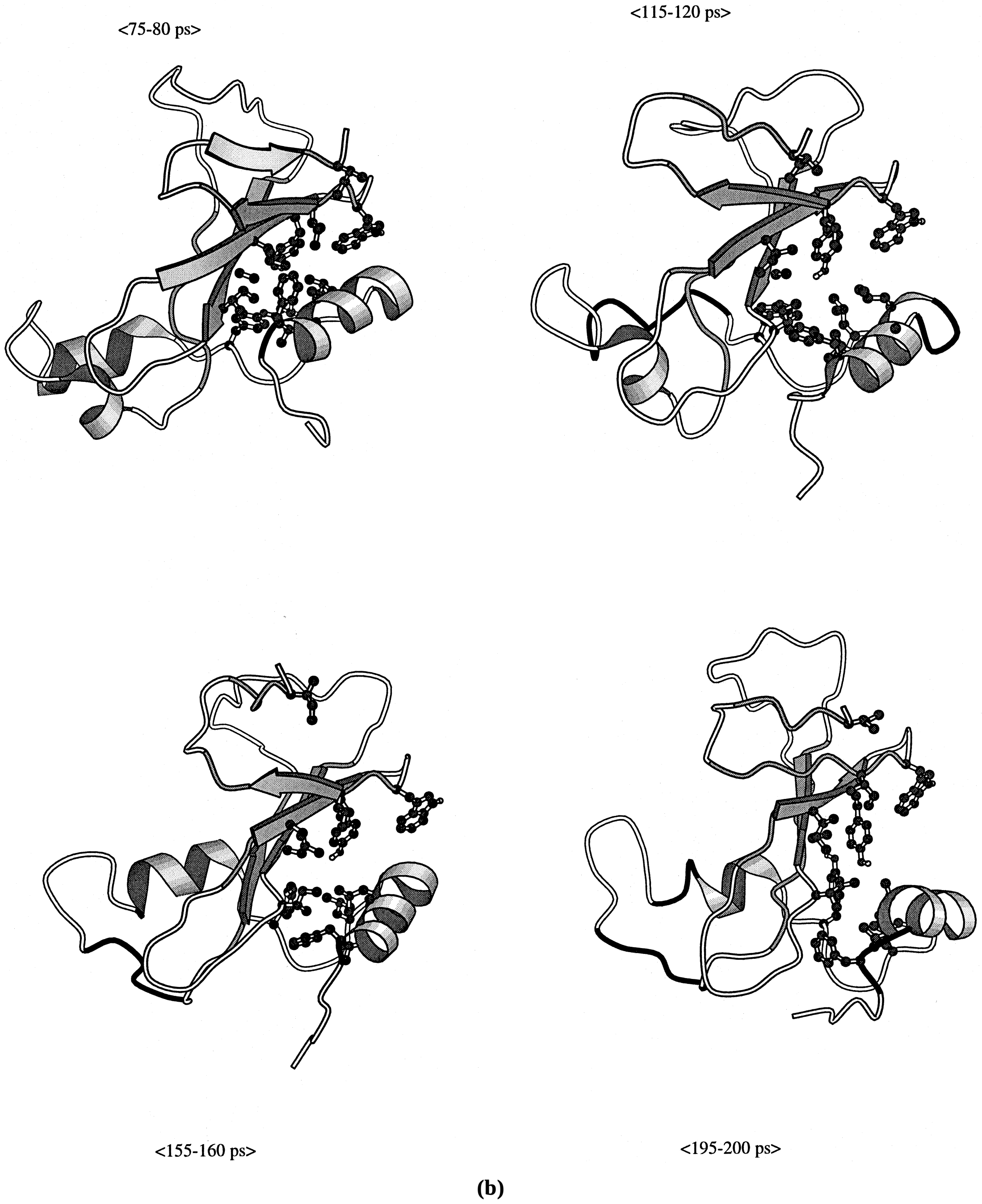


Figure 4(a) (*legend opposite*)


<75-80 ps>

<115-120 ps>

<155-160 ps>

<195-200 ps>

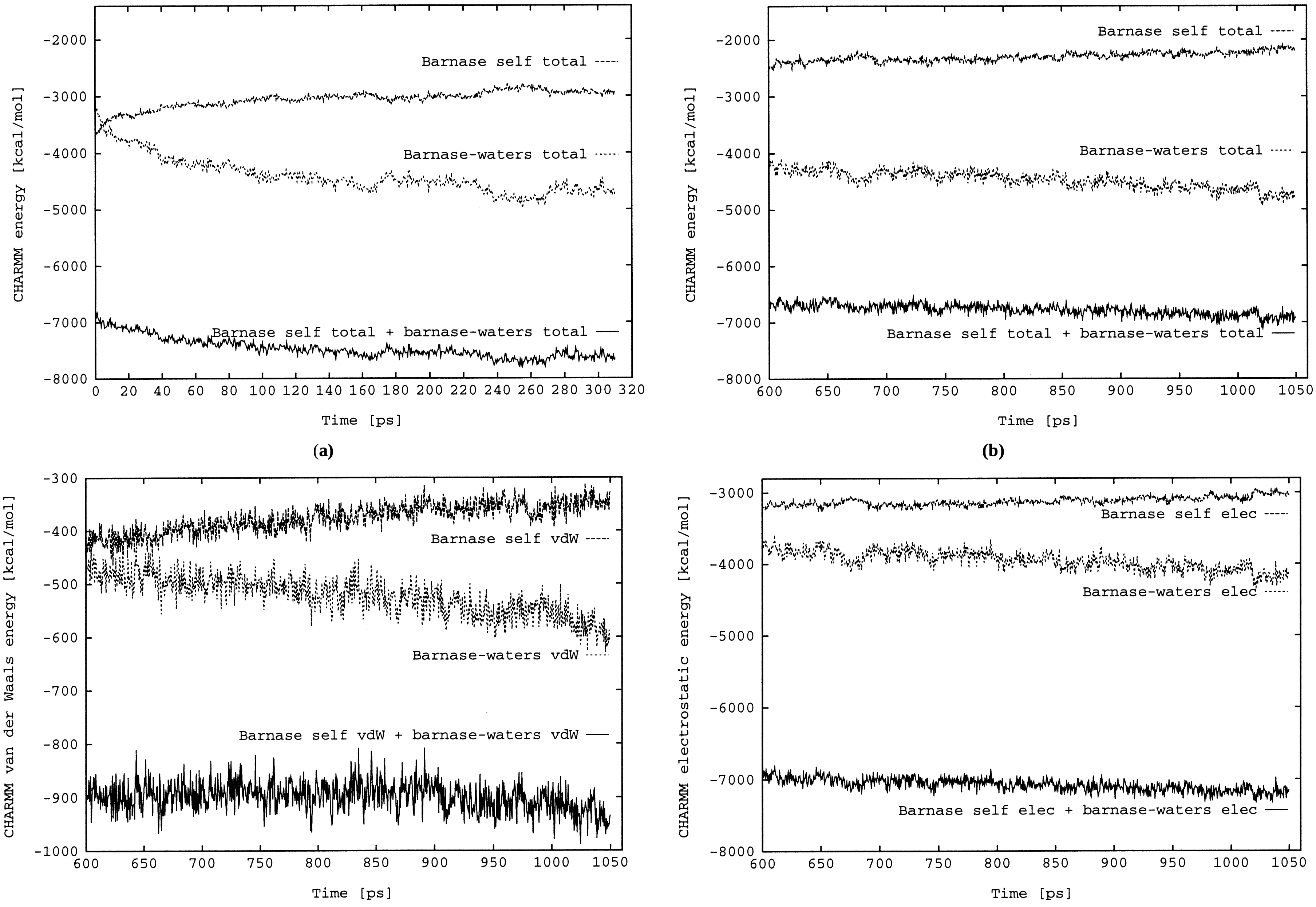
(b)

Figure 4. (a) Denaturation of barnase. Stereo view of the barnase C^{α} atoms. Thin line and labels, 1 ps; medium line, 600 ps; thicker line, 800 ps; thick line and labels, 1050 ps. (b) A representation of the backbone atoms of barnase at different stages of the S600 simulation. Unfolded segments that correspond to native α-helical (β-strand) structure are in black (gray). Sidechains of hydrophobic core 1 are plotted in a ball and stick representation. Figures made from 5 ps averages with the MOLSCRIPT program (Kraulis, 1991).

the barnase intramolecular energy and the barnase-water interaction energy. There is an additional term corresponding to water-water interaction energy, which is not easily obtained from a simulation because of the slow convergence of the mean water energy and its sensitivity to the details of the water structure. As has been shown formally from thermodynamic arguments, the water-water energy is important for the solvation enthalpy and entropy but cancels in the solvation free energy (Ben-Naim, 1975; Yu & Karplus, 1988).

Barnase internal energy and barnase-solvent interaction energy

In Figure 5, we show results for the potential energy as a function of time calculated from the simulations. For each simulation, the total barnase



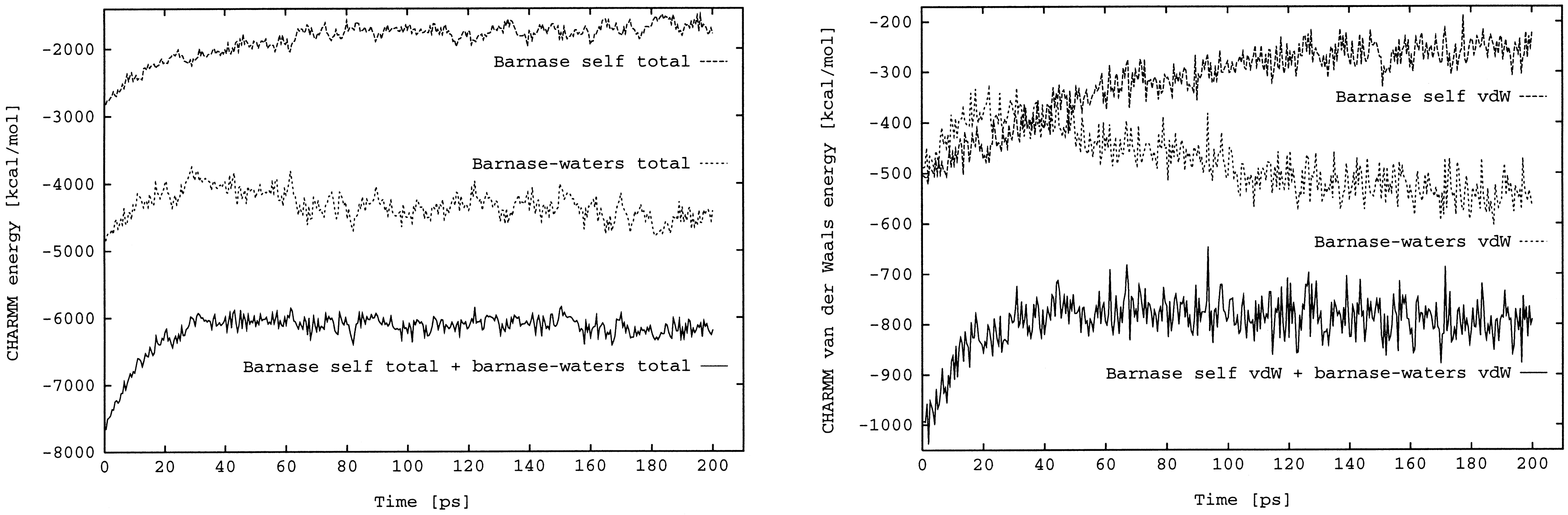
(a)

(b)

(c)

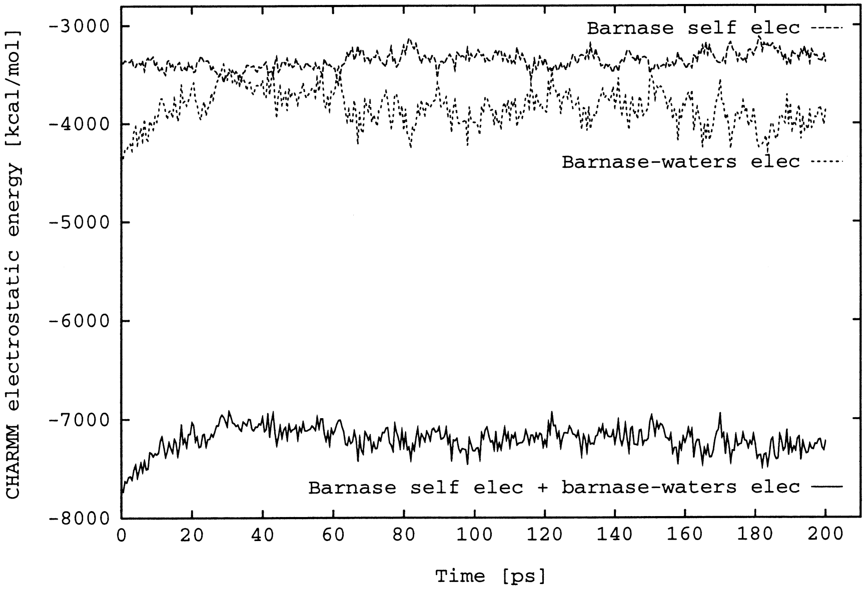
(d)

Figure 5(a–d) (*legend opposite*)



(e)

(f)



(g)

Figure 5. CHARMM energies plotted as a function of simulation time. (a) Total energy during the control run at 300 K (NpH 300: barnase self energy, broken line; barnase-water interaction energy, dotted line; sum, continuous line. (b) Same as (a) for LpH 360. (c) van der Waals energy during the LpH 360 simulation: barnase self energy, broken line; barnase-water interaction energy, dotted line; sum, continuous line. (d) Electrostatic energy during the LpH 360 simulation: barnase, broken line; barnase-water, dotted line; sum, continuous line. (e) to (g) Same as (b) to (d) for thermal denaturation run S600.

self energy (i.e. the internal bonding and non-bond energy), the barnase-water interaction energy and the sum of the two are presented. Figure 5(a), which is the control run (NpH 300), shows the usual behavior in a stable simulation; i.e. the total potential energy initially decreases due to the improved barnase-water interaction energy, in part at the expense of the self energy. This is expected, since the initial structure, which corresponds to the crystal where the barnase has a protein and solvent environment, is adapting to a pure solvent environment during the simulation. The self terms reach equilibrium rather rapidly (in about 160 ps), but the barnase-water contribution (particularly that arising from the electrostatic term) requires about 260 ps for equilibration.

The NpH300 results can be compared with those obtained for barnase under acid denaturation conditions (LpH 360). Figure 5(b) shows the results starting at 600 ps when the temperature was raised to 360 K. For earlier times, the system energy behaved in a manner similar to the NpH 300 simulation results; i.e. it reached equilibrium in terms of the barnase self energy and the barnase-water interaction energy by about 200 ps at 300 K. When the temperature was raised to 360 K, unfolding began almost immediately, as indicated by the reduction in favorable self energy and the more favorable barnase-water energy (see Figure 5(b)). The van der Waals (Figure 5(c)) and the electrostatic (Figure 5(d)) terms show corresponding behavior with the latter contribution nearly ten times the former.

Corresponding results for the high-temperature run (S600) are shown in Figure 5(e), (f) and (g). The bonding term (not shown) is essentially constant once the system reaches the equilibrium high-temperature value, which is somewhat less than the factor of 2 expected from a harmonic model; i.e. the bonding terms have a value of 1880($\pm$50) kcal/mol at 600°C, relative to 1028($\pm$20) kcal at 300°K. The S600 run, which was started from the 258 ps structure of the control run, shows a large deterioration of the solvation energy during the first 30 ps of the trajectory. This behavior suggests that the adaption to the water environment in the 258 ps of simulation (NpH 300) is partly destroyed when the temperature is raised. This effect is not observed in the initial phase of LpH 360 (dotted line in Figure 5(b)) when the temperature is raised only to 360 K, which corresponds to much milder unfolding condition. Also, the other high-temperature simulations (A600 and R600) showed a smaller effect because they were started at earlier times and so were less well adapted to the solvent. After this initial period, the behavior of S600 is similar to the LpH 360 simulation in that the barnase-water interaction energy decreases and the barnase self term increases on denaturation. The van der Waals term shows larger changes than for LpH360 due to the greater structural disruption of S600. The magnitude of the electrostatic energy in the low-pH and neutral-pH system are very different (see Figure 5(d) and (g)). Due to the total charge of +17, the barnase electrostatic self energy is less favorable in the low-pH simulation than in the neutral-pH runs (total charge of +2). It increases from −3880($\pm$24) kcal/mol in the first 10 ps to −3023($\pm$20) kcal/mol in the last 10 ps at low pH. This can be compared with the S600 simulation increase from −3375($\pm$36) to −3308 ($\pm$46) kcal/mol.

An analysis of the enthalpy of protein denaturation (Lazaridis *et al*., 1995) has been made based on the measurements made by Makhatadze and Privalov (1993). It was shown that $\Delta H_{\mathrm{N}}^{\mathrm{U}}(\mathrm{sol})$, the enthalpy of unfolding in solution, could be written as:

$$\Delta H_{\mathrm{N}}^{\mathrm{U}}(\mathrm{sol}) = \Delta H_{\mathrm{N}}^{\mathrm{U}}(\mathrm{vac}) + \Delta H_{\mathrm{vac}}^{\mathrm{sol}}(\mathrm{U}) - \Delta H_{\mathrm{vac}}^{\mathrm{sol}}(\mathrm{N})$$

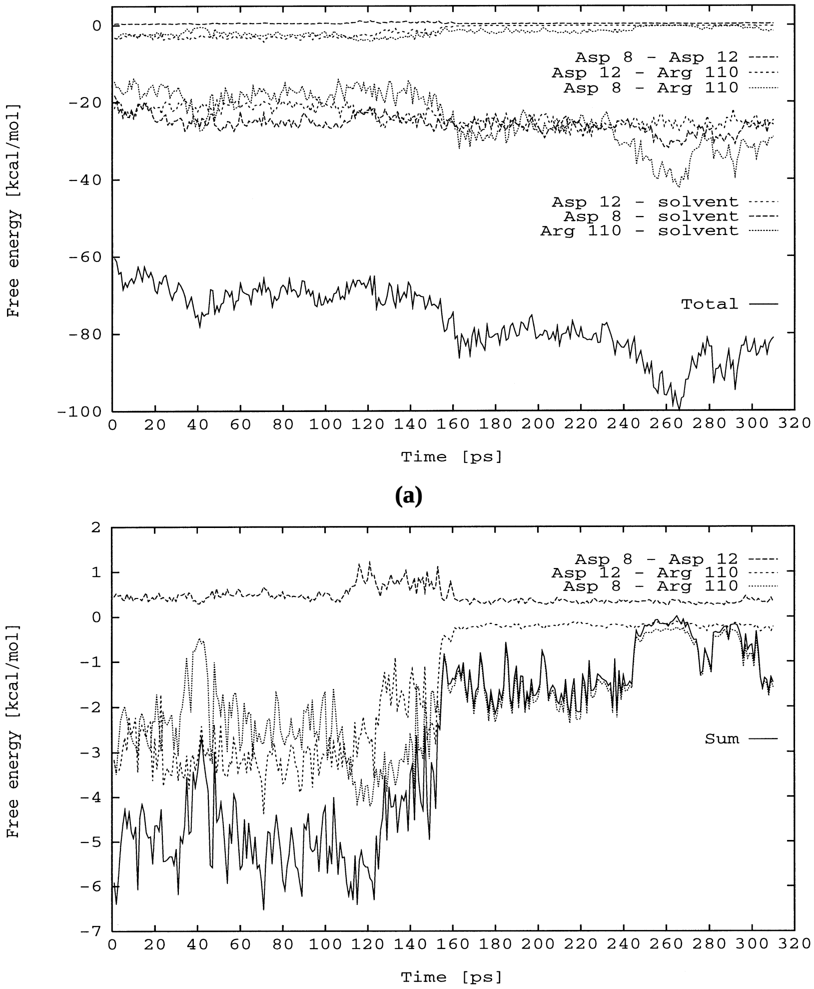
where $\Delta H_{\mathrm{N}}^{\mathrm{U}}(\mathrm{vac})$ is the unfolding enthalpy for the protein in vacuum and $\Delta H_{\mathrm{vac}}^{\mathrm{sol}}(\mathrm{N})$ and $\Delta H_{\mathrm{vac}}^{\mathrm{sol}}(\mathrm{U})$ are the enthalpy of solvation of the protein in the native (N) and unfolded (U) states, respectively. The quantity $\Delta H_{\mathrm{N}}^{\mathrm{U}}(\mathrm{vac})$ is equal to the barnase self energy shown in Figure 5(b) and (e); i.e. it is the enthalpy of the protein alone with structures corresponding to that native and unfolded state in solution. The solvation terms $\Delta H_{\mathrm{vac}}^{\mathrm{sol}}(\mathrm{N})$ and $\Delta H_{\mathrm{vac}}^{\mathrm{sol}}(\mathrm{U})$ correspond to the barnase-water interaction energy shown in Figure 5(b) and (e) plus a cavity term arising from the change in the water-water enthalpy on solvating the native (N) or denatured (U) protein. From the analysis presented by Lazaridis *et al*. (1995), as well as earlier work (Ben-Naim, 1975; Yu & Karplus, 1988), it is clear that the cavity term plays a significant role in the energy and the enthalpy of denaturation. For a variety of proteins, the enthalpy contribution of non-polar groups stabilizes the native state while that of polar groups destabilizes the native state when the cavity term is included in both cases (Lazaridis *et al*., 1995). Comparing the total energy (barnase self plus barnase-water) at 600 ps (5 ps average) with that at 1050 ps (5 ps average) as shown in Figure 5(b), we obtain a stabilization of the (partly) unfolded form that is equal to 227 kcal. For barnase, the measurements made by Griko *et al*. (1994) yield a value of $\Delta H_{\mathrm{N}}^{\mathrm{U}}(\mathrm{sol}) = 73.4$ kcal/mol at pH 2. If the simulation results are assumed to correspond to the transition from the folded to the unfolded state, this means that the cavity enthalpy contribution is 300 kcal/mol stabilizing the native state. This value is of the appropriate magnitude, since theory indicates that the cavity term is approximately equal to the negative of one-half of the average of the solute-solvent interaction energy in polar systems (Yu *et al*., 1991; Archontis & Karplus, unpublished results). The calculated change in the barnase-water interaction energy in the (partial) unfolding process is −538 kcal/mol, from which the expected cavity term is about 270 kcal/mol.

Double salt-bridge: Poisson-Boltzmann and molecular dynamics results

The stability of the double salt-bridge Asp8-Arg110-Asp12 on the surface of barnase was analyzed at neutral pH by solving the finite-difference LPB equation for the set of structures at 1 ps intervals from the NpH 300 simulation (see Methods); in presenting the results below, we give average values and RMS fluctuations. The LPB results are compared with the simulation results, including the CHARMM energies, where appropriate.

Figure 6(a) and (b) show the results of the Poisson-Boltzmann calculations for NpH 300. During the first 150 ps of NpH 300 the double salt-bridge was present (Figure 6(a)), apart from the transient insertion of a water molecule between the Arg110-Asp8 salt-bridge in the 38 to 46 ps period. The value of the Arg110-Asp8 and Arg110-Asp12 electrostatic free energy of interaction was −2.55(±0.78) kcal/mol (mean ± standard deviation in the first 150 ps) and −2.86(±0.66) kcal/mol, respectively; the repulsion between Asp8 and Asp12 corresponds to a free energy of 0.50(±0.18) kcal/mol. These values cannot be compared directly with the salt-bridge contribution to the stabilization of the protein measured by site-directed mutagenesis (Sali *et al*., 1991), since it would be necessary to compare the folded and unfolded protein and include the solvation self energy term (see below); for an example of such calculations for a variety of proteins, see Hendsch & Tidor (1994). The present results are concerned directly with the stability of the salt-bridge in the folded state.

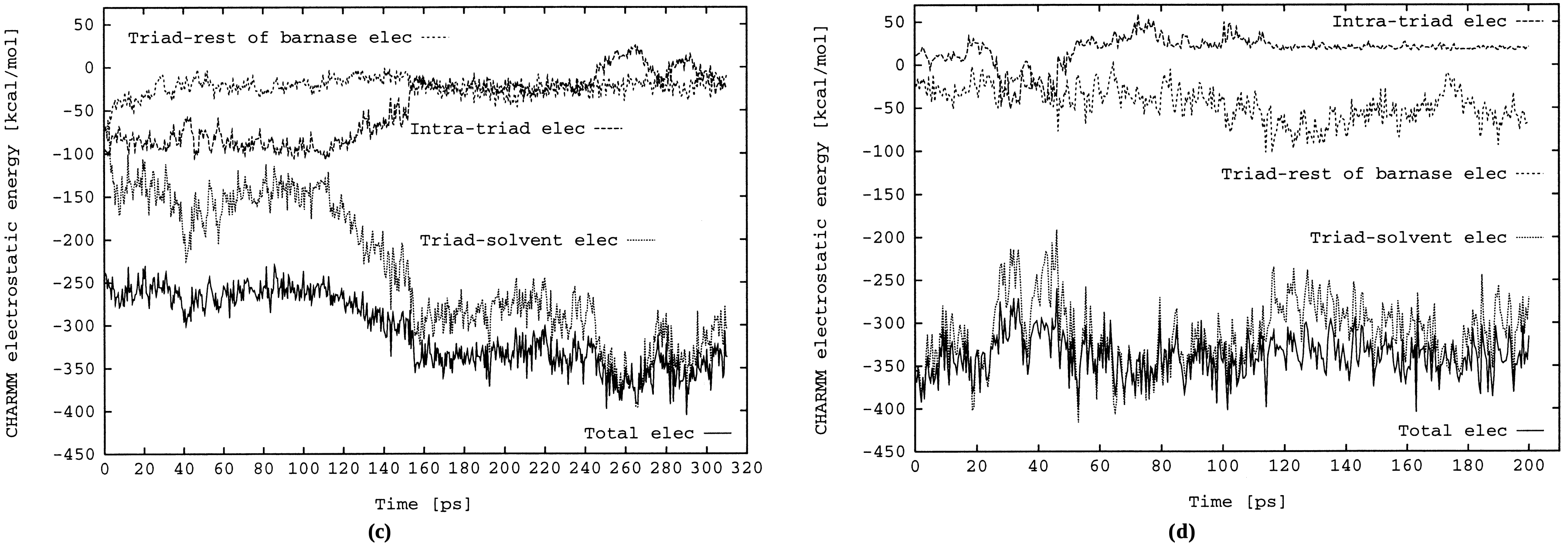
It is of interest to examine the magnitudes of the terms that contribue the free energy of the salt-bridge. During the first 150 ps when the salt-bridge existed, the self energy of Asp8, Asp12 and Arg110 that arises from their solvation in the partly aqueous environment were −24.51(±1.71) kcal/mol, −21.71(±1.58) kcal/mol, and −18.23(±2.60) kcal/mol, respectively. These values are much smaller in magnitude than the solvation free energy of isolated acetate or guanidinum groups, indicating that the side-chains contributing to the salt-bridge are shielded by the protein environment. After about 155 ps the Arg110-Asp12 bridge is broken (Figure 6(b) and 7). This resulted in a slightly greater solvation of both the Asp12 sidechain and the Arg110 sidechain; the self energies during the 156 to 245 ps period are −27.09(±2.36) kcal/mol (from −18.23(±2.60) kcal/mol); the self energy of Asp12 improved dur-



(a)

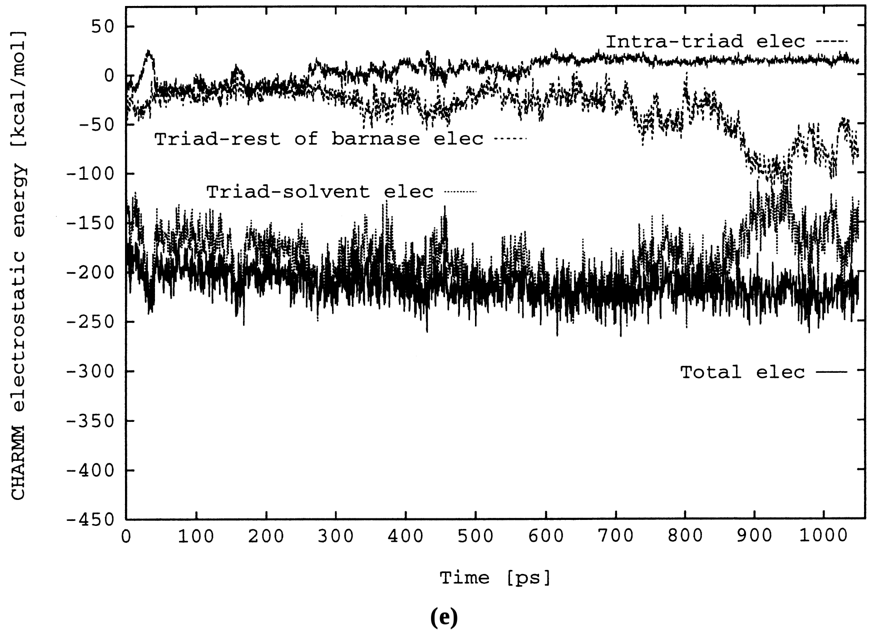
(b)

Figure 6(a–b) (*legend overleaf*)



(c)

(d)



(e)

Figure 6. (a) and (b) Poisson-Boltzmann, and (c) to (e) CHARMM electrostatic energies of Asp8-Arg110-Asp12 salt-bridge. (a) Control run at 300 K (NpH 300). Top lines: electrostatic free energies of interaction; Asp8-Asp12 (broken line, long dashes); Arg110-Asp8 (dotted line); Arg110-Asp12 (broken line, short dashes). Lines in the middle part of plot: self-energies; Asp8-solvent (broken line, long dashes); Asp12-solvent (broken line, short dashes); Arg110-solvent (dotted line). Sum of the six contributions (continuous line). (b) Expanded version of (a) showing the electrostatic free energies of interaction. (c) to (e) CHARMM electrostatic interaction energy. Intra-triad (broken line, long dashes), triad-rest of barnase (broken line, short dashes), triad-solvent (dotted line), sum of these three contributions (continuous line). (c) NpH 300; (d) S600; (e) LpH 360.

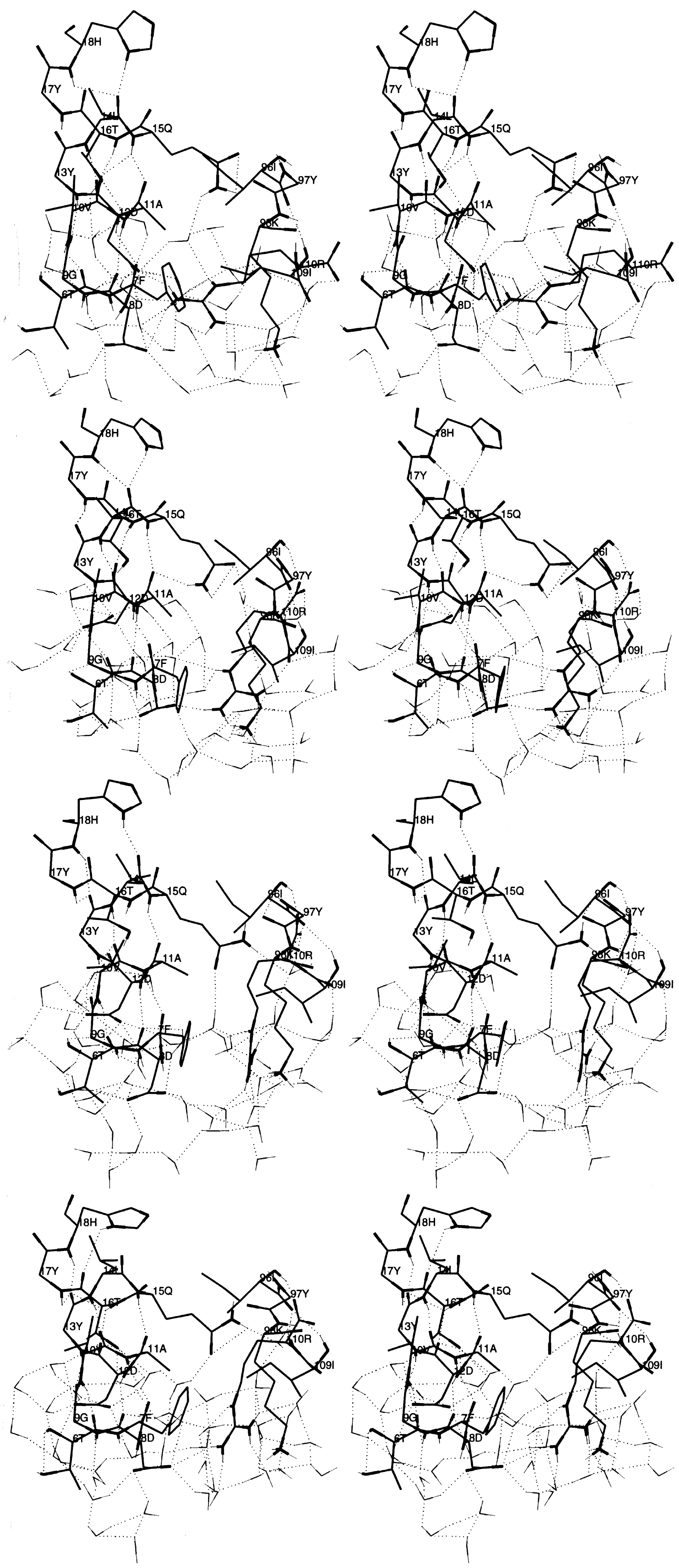


Figure 7. Stereo drawing of the double salt-bridge during the 300 K control run (NpH 300). Helix 1 and part of strand 4 are also shown. Nitrogen and oxygen atoms are thick, hydrogens are thin and hydrogen bonds are dotted (acceptor to hydrogen distance smaller than 2.5 Å; no angular criterion); water molecules within 4 Å of any atom of the Asp8, Asp12 and Arg110 sidechains are shown. The Tyr13 and Tyr17 sidechains are omitted for clarity. From top to bottom: 150 ps, 200 ps, 250 ps and 300 ps.

ing the 156 to 310 ps period to a value of –25.58(±1.25) kcal/mol. A minor improvement, from –24.51(±1.71) to –26.18(±1.17) is found for the Asp8 side-chain during the 156 to 240 ps period when it is still part of the salt-bridge to Arg110. Due to the absence of the Arg110-Asp12 salt-bridge, the strength of the Arg110-Asp8 salt-bridge was reduced to a value of –1.70(±0.39) kcal/mol during this period (Figure 6(b)). That there is coupling between the strengths of the two salt-bridges is in agreement with mutagenesis results (Horovitz *et al*., 1990). After about 245 ps, a water molecule inserted between the Arg110 and Asp8 sidechains for about 50 ps. The third picture from the top of Figure 7 shows the instantaneous structure at 250 ps; the inserted water molecule donates a hydrogen bond to an oxygen atom of the Asp8 sidechain and accepts from the Arg110 sidechain and another water molecule. Each donor group on the Arg110 guanidinium moiety donates one or two hydrogen bonds to water oxygen atoms. The bottom picture of Figure 7 shows that the Arg110-Asp8 salt-bridge has partially reformed at 300 ps, although most of the Arg110 and Asp8 sidechains are surrounded by water molecules. During the 250 to 270 ps period, in which the double salt-bridge is broken, the sum of the sidechain self energies (Asp8, –29.72(±1.51) kcal/mol; Asp12, –25.55(±1.30) kcal/mol; Arg110, –38.33 (±2.40) kcal/mol), and the electrostatic free energies of interaction (Arg110-Asp8, –0.31(±0.04) kcal/mol; Arg110-Asp12, –0.15(±0.03) kcal/mol; and Asp8-Asp12, 0.32(±0.04) kcal/mol) is more negative (continuous line in Figure 6(a)). At the end of NpH 300 (i.e. from 305 to 310 ps) the Arg110-Asp8 salt-bridge reforms (–1.55(±0.13) kcal/mol) and the solvation energy of both Asp8 and Arg110 is reduced to –26.14(±1.00) kcal/mol and –30.07(±0.89) kcal/mol, respectively. This results in a less favorable electrostatic free energy of the triad (continuous line in Figure 6(a)).

If the solvation free energies calculated from the Poisson-Boltzmann treatment are accurate, the Arg110-Asp12 bridge is not stable, even in the presence of Asp8, and neither is the Asp8-Arg110-Asp12 double salt-bridge. From the results of the first 150 ps, the total free energy of the salt-bridge is –68.8 kcal/mol. This is to be compared with the result of –93.7 kcal/mol for the 250 to 270 ps period when the salt-bridge is broken. From the simulation it is not clear whether the salt-bridge is unstable and takes a long time to come apart or whether the Poisson-Boltzmann calculations, which do not take account of water structure, are inaccurate. The fact that the Arg110-Asp8 salt-bridge reforms suggests that the equilibrium between the existing and broken salt-bridge is in the folded state much closer to unity than would be expected from the Poisson-Boltzmann calculations.

At 600 K, water molecules often insert transiently between the charged sidechains of the double salt-bridge. The Arg110 sidechain separates definitively from Asp8 and Asp12 at about 40 ps, 110 ps and 50 ps in A600, R600 and S600, respectively. During the first 110 ps of R600 the Arg110-solvent, Arg110-Asp8 and Arg110-Asp12 electrostatic energies show significant fluctuations. Despite several transient ruptures, the Arg110-Asp8 salt-bridge lasts during the first 100 ps, while the Arg110-Asp12 bridge comes apart after about 80 ps. At 110 ps, all three charged sidechains are fully solvated. There is transient formation of a strong polar interaction between the guanidinium group of Arg110 and the hydroxyl group of Tyr97 at about 240 ps.

To complement the Poisson-Boltzmann analysis of the double salt-bridge and to compare with the low-pH simulation, in which there are no charge-charge interactions between Asp8, Arg110 and Asp12, plots of the CHARMM electrostatic energy are shown in Figure 6(c) to (e). There is qualitative agreement in the time-dependence of the free energy of the Poisson-Boltzmann results (Figure 6(a)) and the energy from the CHARMM results (Figure 6(c)). However, as expected, the range of values is much larger in the latter. It is found that the CHARMM Coulombic interaction energy between the triad and the rest of barnase is almost constant. The behavior of the triad-solvent Coulombic energy is similar to the Arg110 free energy of solvation (compare dotted lines in Figure 6(a) and (c)). Moreover, the sum of the Coulombic triad-solvent, intra-triad and triad-rest of barnase energies (continuous line in Figure 6(c)) shows the same decreasing and increasing intervals as the Poisson-Boltzmann free energy value (continuous line in Figure 6(a)). In the S600 run (Figure 6(d)) the CHARMM intra-triad electrostatic energy is slightly unfavorable during the first 25 ps and from about 50 to 200 ps, the double salt-bridge being broken after 258 ps of the control run, i.e. at the beginning of S600. From 25 to 50 ps of S600 there is a transient reformation of the double salt-bridge. Interestingly, the electrostatic interaction energy between the triad and the rest of barnase improves during the first 120 ps. It then fluctuates around an almost constant value until the end of the simulation.

At low pH during the first 260 ps, the hydrogen bonds and polar interactions between the Asp8 and Arg110 sidechains (and to a smaller extent between Asp12 and Arg110) are stabilizing, apart from some transient periods when some of the hydrogen bonds are broken (Figure 6(e)). From 260 to 560 ps the three sidechains come apart, though there is transient formation of hydrogen bonds mainly between the Asp12 and Arg110 sidechains. When these residues separate, the intra-triad electrostatic energy reaches a slightly unfavorable plateau value of about 15 kcal/mol (Figure 6(e)). The electrostatic interactions between the triad sidechains and the rest of barnase show a behavior opposite to the triad-solvent electrostatic energy. The latter has a value of about –200 kcal/mol (from 500 to 700 ps at low pH, when these sidechains are solvated, Figure 6(e)), which has to be compared with a value of about –350 kcal/mol at neutral pH (Figure (6c) at ~260 ps and (d)). At about 900 ps, a strong hydrogen bond

Table 2. Solvent-accessible surface area calculated for the different sidechains of core 1 using the algorithm of Lee & Richards (1971) within CHARMM

		Control run at 300 K (NpH300)			
Residue	X-ray	Average[b]	SD	Min	Max
Phe7	26	43	8	22	63
Val10	0	0	1	0	5
Ala11	0	0	1	0	9
Leu14	0	0	1	0	6
Leu20	10	29	7	15	54
Tyr24	2 (2)[a]	0 (0)	1 (1)	0 (0)	7 (7)
Ala74	0	0	0	0	0
Ile76	0	2	2	0	9
Ile88	0	0	0	0	0
Tyr90	10 (6)	6 (6)	3 (3)	0 (0)	17 (15)
Trp94	68 (59)[a]	51 (47)	10 (9)	25 (25)	102 (93)
Ile96	0	0	0	0	5
Ile109	20	21	10	2	48

[a] Values in parentheses exclude the OH group of Tyr and the indole NH of Trp.

[b] In the dynamics, surface area calculations were done each picosecond. Average refers to the average over the dynamics and SD to the standard deviation; Min and Max refer to the minimum and maximum values observed during the simulation, respectively.

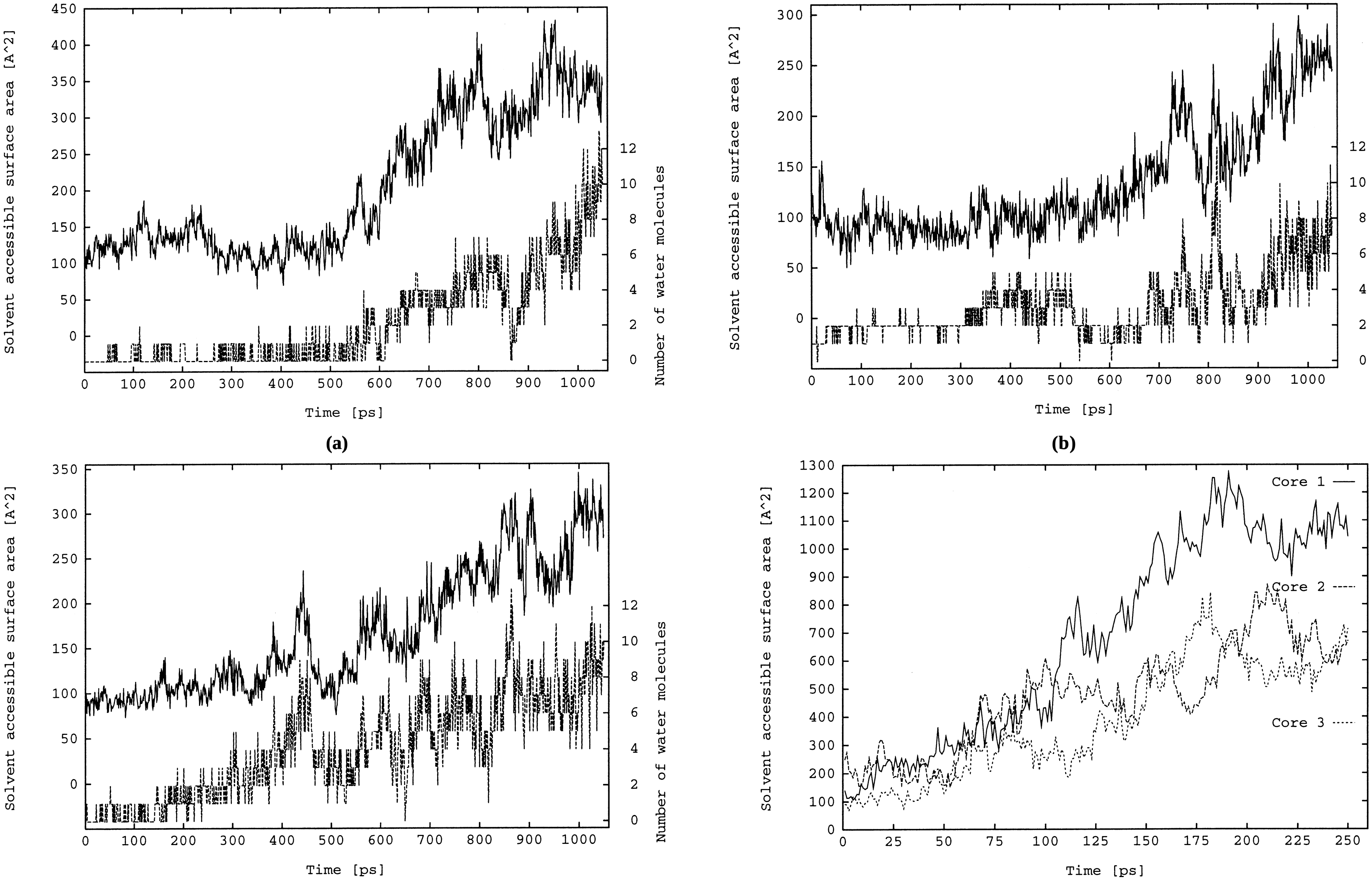
is formed between the Asp8 and Lys98 sidechains; it lasts until the end of the run (see Figures 6(e) and 10(f)).

Denaturation of hydrophobic cores

The main hydrophobic core is formed by the packing of helix 1 against the β-sheet (Figure 1) and it is thought to be the major stabilizing element of barnase (Mauguen *et al*., 1982). The core consists of 13 hydrophobic sidechains; its center is at the Ile88 sidechain (see the legend to Figure 8 for a list of core residues). Analysis of the exposed surface of individual core 1 sidechains in the X-ray structure and during the 310 ps control trajectory indicates that residues Val10, Ala11, Leu14, Tyr24, Ala74, Ile76, Ile88 and Ile96 are completely buried. Residues Phe7, Leu20, Tyr90, Trp94 and Ile109 are partially exposed to the solvent; the exposed surface areas and their fluctuations during the simulation are given in Table 2, which includes the X-ray structure values for comparison. Fersht and co-workers considered Val10 and Ile76 as belonging to the edges of core 1, though these residues are fully buried; we prefer to divide core 1 residues into ‘‘completely buried’’ and ‘‘partially exposed’’. There are also two smaller hydrophobic cores consisting of nine and seven residues. Core 2 belongs to an almost separated subdomain in barnase formed by α-helices 2 and 3, loop 2 (which connects them) and the β-turn consisting of residues 46 to 49 (Figures 1 and 12). It contains several water molecules (Serrano *et al*., 1992c). The third core (core 3) is formed by packing of the Phe56 and Leu63 sidechains in loop 3 against the β-sheet and the sidechain of Tyr103 (β-turn, Figure 1). Cores 1 and 3 do not contain any water molecules in the crystal structure. The solvent-accessible surface of the barnase hydrophobic cores and the individual sidechain contributions in the X-ray structure are shown in Tables 1 and 2 of Serrano *et al*. (1992a). In what follows, we discuss the general denaturation behavior of the different cores and describe in some detail the role of the water molecules in the denaturation process. We realize that what is found in the present simulations represents only an example of what can happen. Other simulations would be expected to show a different sequence of events. Nevertheless, the results provide insights into the dynamics of water solvating a hydrophobic region. Of particular interest is the general observation that the hydrogen-bonding capacities of the water molecules tend to be satisfied throughout the denaturation process. Also, the similarity between the low and high-temperature results is important for validating the analysis.

Core 1

Figure 8(a) shows the time-dependence of the solvent-accessible surface area of the main hydrophobic core (core 1) and the number of water molecules that had penetrated in LpH 360. The corresponding plot for the A600 and B300 simulations is shown in Figure 3 of Caflisch & Karplus (1994a). At 327 K, core 1 undergoes a transient partial opening at about 560 ps, due to the transient increase in exposure of the Phe7 sidechain, which is already partly exposed in the control simulations and the X-ray structure. A similar transitory phenomenon is seen at 440 ps in core 3, due to the transient increase in exposed surface of the Tyr103 sidechain (see below). Increase in accessible surface area and partial penetration of the water molecules are nearly simultaneous for the three cores and begin almost immediately at 600 ps when the temperature is raised to 360 K. The essentially simultaneous opening of the cores also took place in S600 and R600 at about 50 ps; the results for R600 are shown in Figure 8(d). In A600, core 2 started to unfold at about 15 ps followed by core 1 (at about 50 ps) and core 3 (at about 70 ps). In LpH 360 the core 1 exposed surface area increases during the 600 to 800 ps period. This was due mainly to the motion of the Phe7, Leu20 and Ile109 sidechains, all of which are partly exposed in the control. After 800 ps, core 1 undergoes a slight contraction. This is followed by a steady period (from 830 to 900 ps) characterized by about 290 Å^2 of accessible surface area. There are three to four water molecules inserted between the edge sidechains and the center of the core, which is still compact. In LpH 360 during the 830 to 870 ps period, an unoccupied solvent-accessible volume of up to about 60 Å^3 is present; the value is derived by considering a spherical cavity and adding 1.4 A to the radius of the pseudosphere representing the volume accessible to the center of the probe sphere. This volume could hold about two water molecules; at normal density in liquid water the volume per water molecule is 30 Å^3. The behavior of the empty volume (data not shown) is similar to that of the



(a) (b)

(c) (d)

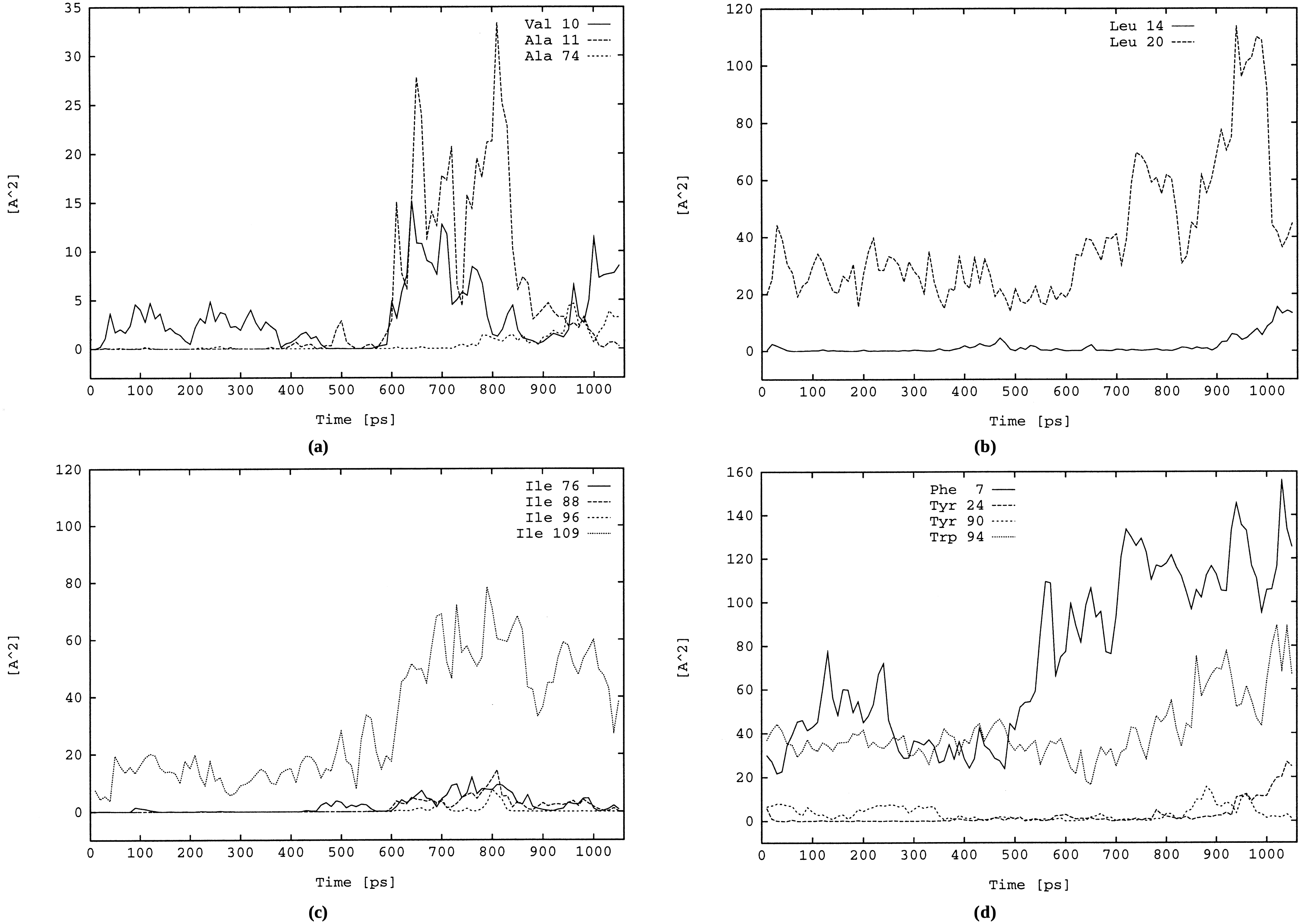
Figure 8. (a) to (c) Solvent-accessible surface area and number of water molecules as a function of time. (a) Core 1 at low pH; (b) core 2 at low pH; (c) core 3 at low pH. For the exposed surface area, shown in the upper curve with the scale on the left, the Lee & Richards (1971) algorithm (CHARMM implementation) and a probe sphere of 1.4 Å radius were utilized. Only the carbon atoms of the hydrophobic sidechains are taken into account. For the number of water molecules, shown in the lower curve with the scale on the right, those within 7 Å (6 Å) of the center of geometry of core 1 (cores 2 and 3) were included. The following residues were taken into account: Phe7, Val10, Ala11, Leu14, Leu20, Tyr24, Ala74, Ile76, Ile88, Tyr90, Trp94, Ile96 and Ile109 for core 1; Ile25, Ala30, Leu33, Trp35, Leu42, Val45, Ala46, Ile51 and Tyr78 for core 2; Phe56, Leu63, Trp71, Leu89, Tyr97, Tyr103 and Phe106 for core 3. (d) Solvent-accessible surface area of the barnase cores plotted as a function of time during the R600 simulation.

exposed surface; for definitions of solvent-accessible volume and empty volume, see Methods. Core 1 expands again from 900 to 950 ps, reaching a peak of about 400 Å^2 in accessible surface with six to seven water molecules in the core. In the last 100 ps of the simulation the accessible surface slightly diminishes to a value of about 340 Å^2, whereas the number of water molecules increases to about 8 to 11 and no cavity remains.

The solvent-accessible surface area of individual core 1 sidechains was determined as a function of time for the LpH 360 simulation. The partial separation of the Phe7 sidechain is evident in the first 600 ps (Figure 9(d)), though the unfolding process has not yet begun. The Leu20, Trp94 and Ile109 residues at the edges of the core are partially exposed in the first 600 ps (and in the control run at 300 K and neutral pH, Table 2), while the Val10, Ala11 and Trp94 sidechains begin to unfold after 600 ps; Val10 and Ala11 do so only transiently. Significant oscillations, originating from the slightly higher temperatures and the loosening of the core packing, are found after 600 ps for the exposed surface of the Ala11 sidechain and to a smaller extent for the Phe7, Leu20, Trp94 and Ile109 sidechains. The separation of the Leu14 and Tyr24 sidechains begins at about 900 ps. The sidechains of Ile88 and Ile96 are completely buried during most of the trajectory (Figure 9(c)). These results are consistent with the protein engineering analysis of core 1 (mutations at Val10, Leu14, Ile76, Ile88, Ile96 and Ile109; Matouschek *et al*., 1992a), which indicates that interactions among residues in the core center (Leu14, Ile88 and Ile96) are weakened by about 40% in the refolding intermediate, while those made by residues at the edge of core 1 (Val10, Ile76 and Ile109) are weakened by 70% to 100%. Thus, there is a correspondence between the simulation and mutation results for these residues. Ile76, however, is almost completely buried during the entire low-pH simulations; in the analysis of Matouschek *et al*. (1992a), Ile76 is as unfolded in the transition state and intermediate state as in the fully unfolded state.

To illustrate the role of solvent in the denaturation of the main hydrophobic core in LpH 360, selected snapshots are shown in Figure 10. At 600 ps, core 1 is compact and is surrounded by water molecules (Figure 10(a)). The sidechains of Asp8 and Lys98 participate in hydrogen bonds with a cluster of water molecules near Phe7 and Ile109. This results in the almost complete solvation of the Phe7 sidechain and the partial solvation of the Ala11 and Ile109 sidechains at 700 ps; the latter interacts also with the alkane chain of Lys98. Water molecules are attacking the CO groups of helix 1 residues 7 to 9, and 16 to 18. This is coupled to the separating motion of helix 1 from the β-sheet. At 800 ps, the Phe7, Ala11 and Ile109 sidechains are almost completely solvated by water molecules (Figure 10(b)). The beginning of the denaturation of the N-terminal part of helix 1 (α to 3_{10} transition, see also Table 3A) and the unfolding of the protein C terminus (residues 109 to 110) occurs at the same time. A chain of water molecules separates the Phe7 sidechains from the rest of the core; the end water molecules are hydrogen-bonded to the OH of Tyr24 on one side and the NH_3 of Lys98 on the other. Moreover, the polar groups of the protonated Asp8 sidechain participate in hydrogen bonds with water molecules, which surround the methyl group of Ala11. The central part of the core is still compact. The Leu20 sidechain is now encircled by a chain of water molecules, whose ends participate in hydrogen bonds with the polar groups of the Tyr90 and Trp94 sidechains. The partial distortion in the upper left part of the core is coupled with the opening of a cavity of about 32 Å^3 (value derived by considering a spherical cavity as mentioned above; the volume of the cavity is larger than the volume accessible to the center of the probe sphere) in the neighborhood of the sidechains of residues Val10, Leu14, Leu20, Tyr24, Ala74, Ile76 and Tyr90; a water molecule donates a hydrogen bond to the hydroxyl group of Tyr90 (Figure 10(b)). This cavity is larger (43 Å^3) at 836 ps (Figure 10(c)) and reaches its maximal size (59 Å^3) at 860 ps (Figure 10(d)). In addition, there is a transient empty space on the surface of the core near the Ile109 idechain (Figure 10(c)). Partial repacking of the Phe7, Ala11, Leu20, Tyr90 and Ile109 sidechains against the rest of the core is visible at 865 ps (Figure 10(e)); this is consistent with the decrease in the solvent-accessible surface of core 1 and the number of core water molecules from 800 to about 870 ps (Figure 8(a)). Shrinking of the cavity (20 Å^3 at 865 ps) is caused mainly by the rotation of the Tyr24 aromatic ring. The Tyr24 OH is pulling a chain of water molecules towards the core edge; it also promotes the partial solvation of the Ile76 sidechain. At 940 ps, the Phe7 sidechain and most of the Ile109 sidechain are again solvated; moreover, the Leu20 and Tyr24 sidechains are separated from the rest of the core by a layer of hydrogen-bonded water molecules, which have filled the aforementioned cavity. A water molecule is inserted in the first turn of helix 1; it donates to the Co of Phe7 and accepts from the NH of Ala11. At the end of the simulation, water molecules solvate the external residues of the core; but the central part (Val10, Ala11, Leu14, Ala74, Ile76, Ile88 and Ile96) is still compact (Figure 10(f)), with less than 20 Å^2 exposed surface per sidechain (Figure 9(a) to (d)). The N terminus of helix 1 is unfolded and the CO group of the Asp8 sidechain acts as acceptor for the Lys98 sidechain from 870 to 1050 ps (Figure 10(f)).

The early phase of solvation of core 1 in the acid-mediated denaturation resembles the behavior seen in the high-temperature simulations. The edge residues, particularly those at or near the N-terminal part of helix 1 (Phe7 and Ile109), are less stable and unfold first. The polar groups of the tyrosine and tryptophan sidechains and the Lys98 NH_3^+ group are important in gathering solvent molecules on the surface of the core and bringing them into the core. Water molecules tend to penetrate the core in chains or clusters.

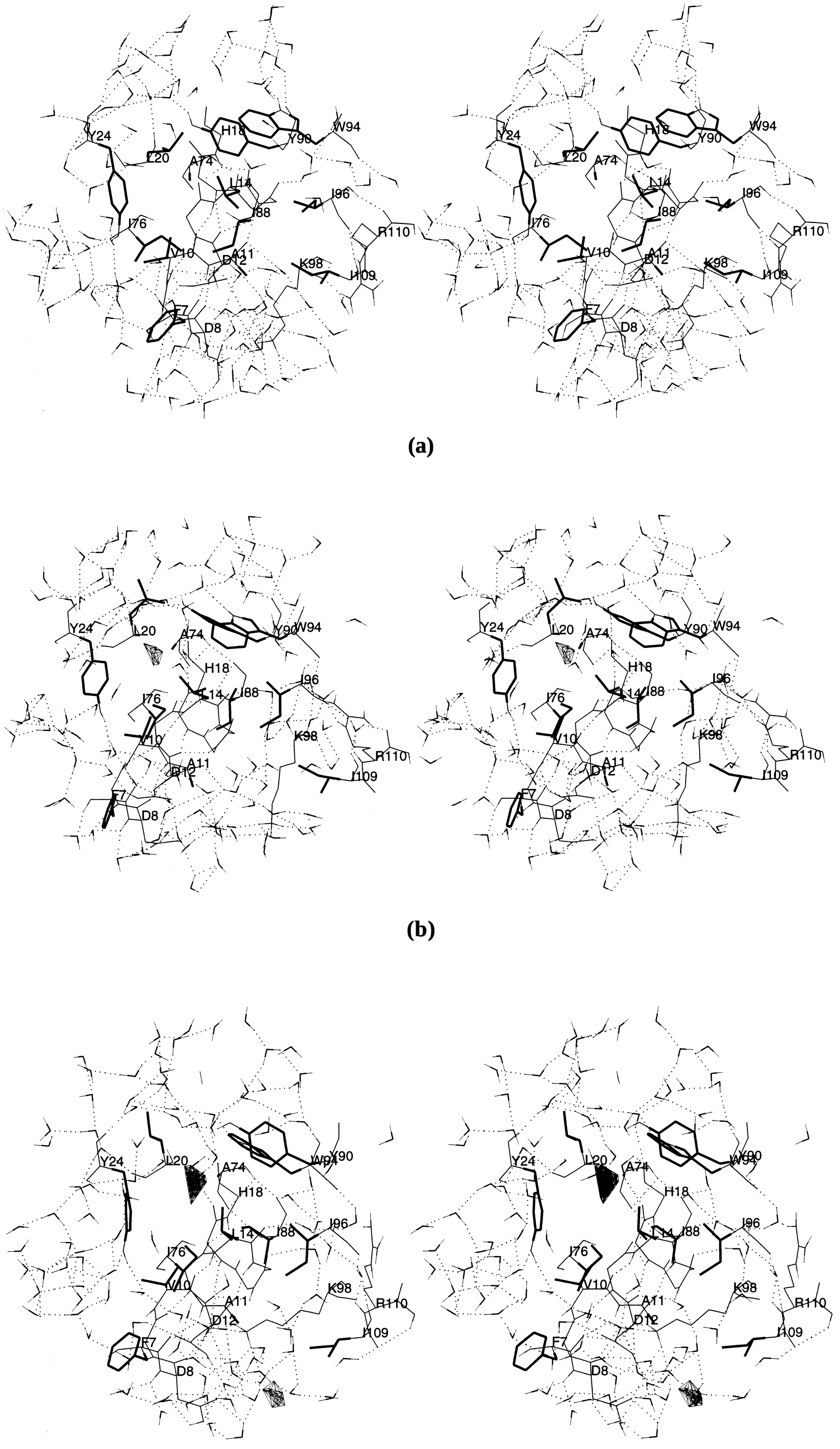


(a) (b)

(c) (d)

Figure 9. Solvent-accessible surface area of individual core 1 sidechains as a function of simulation time averaged over 10 ps intervals (low pH). Only the carbon atoms are taken into account. Residues are grouped by size and characteristics; (a) small (Ala and Val), (b) medium (Leu), (c) large (Ile), and (d) aromatic (Phe, Tyr and Trp).

In the acid-mediated denaturation (LpH 360) the initial opening of hydrophobic core 1 is associated with the penetration of a smaller number of water molecules than in the high-temperature simulations. For example, there are about three to four water molecules in the core during LpH 360 simulation at 700 ps when the exposed surface area has increased to 290 Å^2 (Figure 8(a)), while there are seven water molecules in A600 at 60 ps when there is approximately the same (250 Å^2) exposed surface

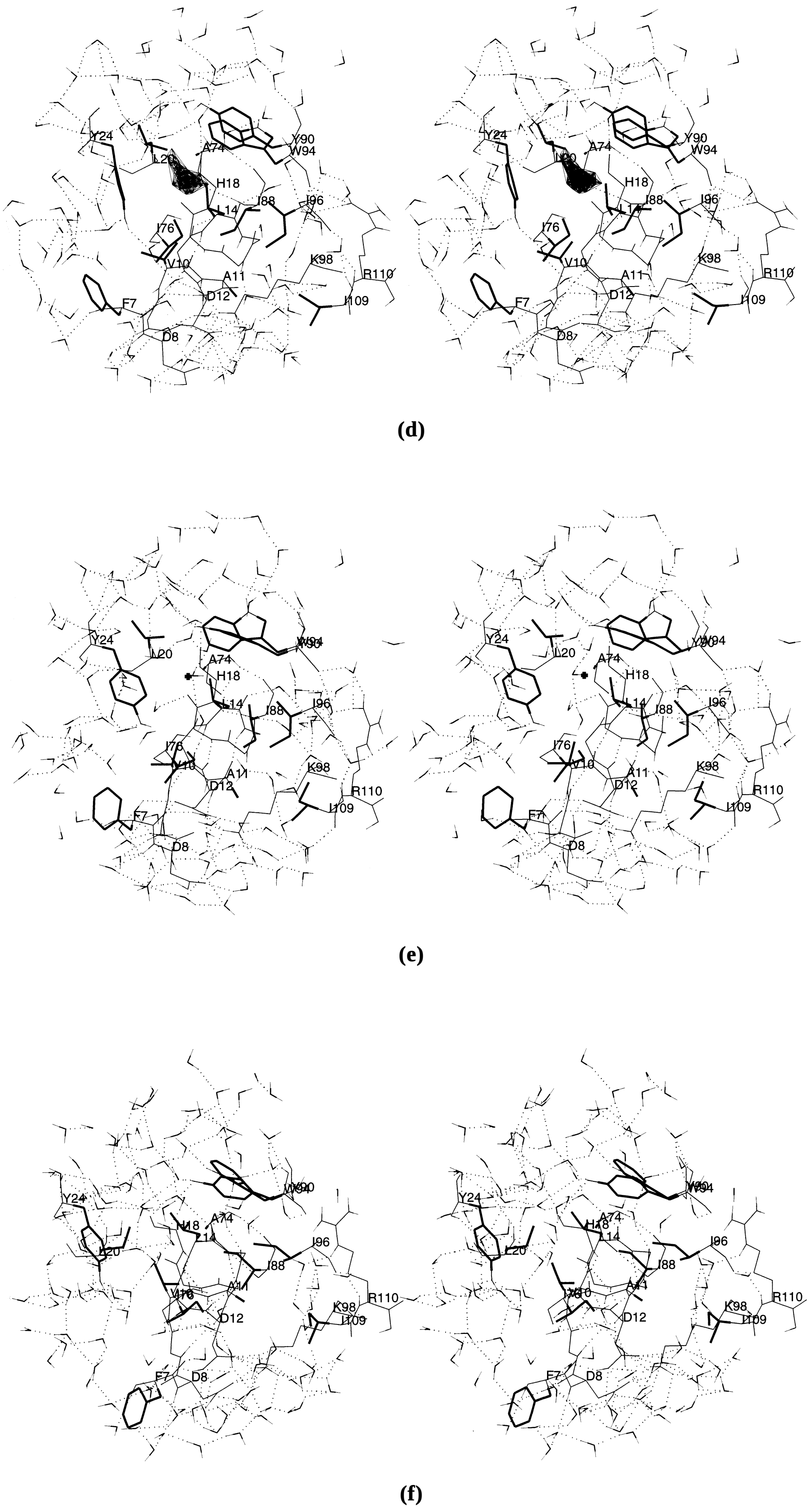


(a)

(b)

(c)

Figure 10(a–c) (*Legend overleaf*)



(d)

(e)

(f)

Figure 10. Stereo pictures of hydrophobic core 1 and its water molecules during the LpH 360 trajectory: (a) 600 ps; (b) 800 ps; (c) 836 ps; (d) 860 ps; (e) 865 ps; (f) 1050 ps. Hydrophobic sidechains of core 1 are shown as thick lines; α-helix backbone (residues 7 to 18), Asp8, Asp12, Lys98 and Arg110 sidechains are shown as thin lines. Water molecules within 13 Å of the center of geometry of hydrophobic core 1 are included. Hydrogen bonds are dotted (acceptor-hydrogen distance smaller than 2.5 Å; no angular criterion). Solvent-accessible volume is shown as a polygon with thin lines; the cavities first appear at 836 ps in LpH 360. The cross in (e) represents a cavity consisting of a single grid point.

Table 3. Helical types for the mainchain hydrogen bonds

A. During the 450 to 1050 ps period of the low-pH simulation

Donor	450	500	550	600	650	700	750	800	850	900	950	1000	1050
10	α	α	α	α	α	α	α	α	3	α	w1914	w35	w35
11	α	α	α	α	α	α	α	α	3	3	3	3	w753
12	α	α	α	α	α	α	α	α	3	W795	w795	w354	w354
13	α	α	α	α	α	α	α	α	α	α	α	α	α
14	α	α	α	α	α	α	α	α	α	α	α	α	α
15	α	α	α	α	α	α	α	α	α	α	α	α	α
16	α	α	α	α	α	α	α	α	α	α	α	α	α
17	α	π	π	α	π	π	π	π	π	π	π	π	π
18	α	α	π	π	π	π	π	π	π	π	π	π	π

B. During the A600 simulation

Donor	10	20	30	40	50	60	70	80	90	100	110	120
10	α	α	α	α	α	w2639		3	α	3	α	α
11	α	α	α	α	α		α	α	α	α	α	α
12	α	α	α	α	α			w1331	α	w1857	α	α
13	α	α	α	α	α	α			α	3	α	
14	α	α	α	α	α	α	α	α	3	α	α	α
15	α	α	α	3	3	3	α		w838		3	3
16	α	3	α	α	α	α	α		W2144	W1224	3	3
17	α	α	α	α	α	α	α	α	α	α	α	α
18	α	α	α	w672								

C. During the B300 simulation

Donor	100	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250
10	w1461	w1461	w1461	w1461	w1461	w1461	w1461	w1461	w1015	w1015	w1015	w1015	w1015	w1015	w1015	w1015
11	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
12	3	3	3		3	3	α	α	α	α	α	α	α	α	α	α
13	3	3		α	α	α	α	α	α	α	α	α	α	α	α	α
14	α	α	α	α	3	α	α	3	3	3			3	3	3	3
15	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838	w838
16	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
17	α	α	α	α	α	w1530	α	α	α	α	α	α	α	α	α	α
18	w1975	w1975	w1975	w1975	w1975	w1975	w2806	w2806	w2523	w2806	w2806	w2806	w2794	w2794	w551	w2626

D. The R600 simulation (1 to 150 ps)

Donor	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
10	α	α	α	α	α	α	α	3	3	α	α				
11	α	α	α	α	α	α	α	α	3	α	3	3	w965	w880	w880
12	α	α	α	α	α	α	α	α	α	α	α	3			
13	α	α	α	α	α	α	α	α	α	α	α	α	3	α	α
14	α	α	α	α	α	α	α	α	α	α	α	α	3	α	α
15	α	α	α	α	α	α	α	α	3	3	α	α	α	α	α
16	α	α	α	α	α	α	α	α	α	α	α	α	w209	3	α
17	α	α	π	α	α	π	π	α	α	α	π			3	α
18	α	π	α	α	α	3				w616	w2539		W1910		w2845

continued overleaf

Table 3—*continued*

The simulation R600 (151 to 250 ps)

Donor	160	170	180	190	200	210	220	230	240	250
10										
11	w880									
12				w2882		w363			w2926	
13	α	w783			3	3			w1486	
14	α	w2818	3	α	α	3	3	3	3	3
15	α	α	α	3	α	3	α	α	α	α
16	3	α	α	3	3	α	α	α	3	α
17	3	α	α	α	α	α	π			
18	w991	w16	w16							

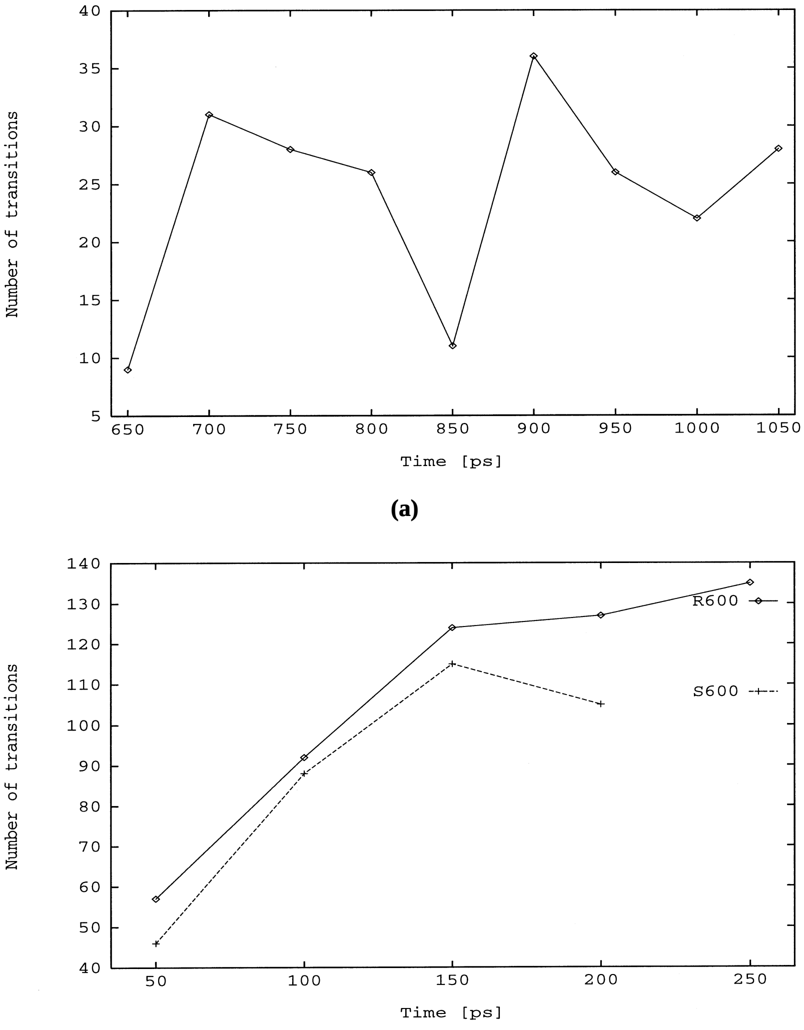
E. The S600 simulation (1 to 150 ps)

Donor	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
10	α	α	α	α	α	α	3	3	3	3	w901	α	α	α	α
11	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
12	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
13	α	α	α	α	α	α	α	3	α	α	α	α	α	α	α
14	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
15	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
16	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
17	α	α	α	α	α	π	α	α	α	π	π	π	π	π	π
18	π	π	π	α	α	π	π	α	α	π	π	π	π	π	π

The simulation S600 (151 to 200 ps)

Donor	160	170	180	190	200
10	3	3	α	α	
11	α	3	3	3	w218
12	α		3	3	
13	3	3	α	α	α
14	α	α	α	α	α
15	α	α	3	3	α
16	α	α	3	3	α
17	α	α	α	3	α
18	π	α	α	3	α

The hydrogen bond from each donor NH is represented in columns corresponding to consecutive 10 ps intervals of the simulation (50 ps intervals for the low-pH simulation) by α, 3 or π depending on whether the α-helical, the 3_{10} -helical type or the π-helical type was predominant. Hydrogen bonds to water molecules occurring during 50% or more in the interval are specified by a w followed by the number of the water molecule. A hydrogen bond was considered formed if the donor to acceptor distance was less than 3.5 Å and the donor-hydrogen-acceptor angle was larger than 120°.



(a)

(b)

Figure 11. Number of dihedral transitions of core 1 sidechains in 50 ps intervals: (a) LpH 360; (b) R600 (continuous line and diamonds), S600 (broken line and crosses).

area. This difference may be due to the higher pressure associated with the boundary potential simulation at 600 K. In the last 100 ps of the LpH 360 denaturation, water penetration occurs without any further expansion of the core (Figure 8(a)). This is possible because of the presence of cavities surrounded by hydrophobic side-chains in the early phase of core 1 expansion. In the LpH 360 simulation the cavities are found to exist for up to about 50 ps.

In the 600 K simulations there is cavity formation in core 1 in the same sites as in the acid-mediated denaturation, i.e. in the upper left part of the core and on the surface, near the Ile109 sidechain. The main difference is that at 600 K the opening of a cavity is a very brief phenomenon; it tends to last less than a picosecond. During the early phase of unfolding (first 50 ps at 600 K) there is only minor distortion in the core and a cavity can shrink and disappear as a consequence of sidechain repacking (the core is still dry), while later on, because of the larger distortion, any cavities are immediately filled by water molecules. It is interesting to note that there also is cavity formation in the partially exposed core during the B300 run, though it is mainly a transient phenomenon that lasts less than 5 ps; there are about 12 to 15 water molecules in core 1 during most of B300.

Dihedral transitions of the core 1 sidechains

The number of sidechain transitions is a measure of the packing of the core. Figure 11(a) shows the number of dihedral angle transitions for the low-pH simulation as a function of time; each symbol represents the number of transitions (see Methods) during the given 50 ps interval. There are only eight sidechain transitions from 600 to 650 ps when the core is still almost completely compact. From 650 to 1050 ps of LpH 360 there are about 28 transitions every 50 ps intervals apart from the 800 to 850 ps period (only 11 transitions); this decrease is concomitant with the loss of core 1 exposed surface area (Figure 8(a)). As a basis of comparison, Loncharich & Brooks (1990) counted 181 sidechain transitions during the last 50 ps of myoglobin dynamics at 340 K. Since myoglobin has about 11

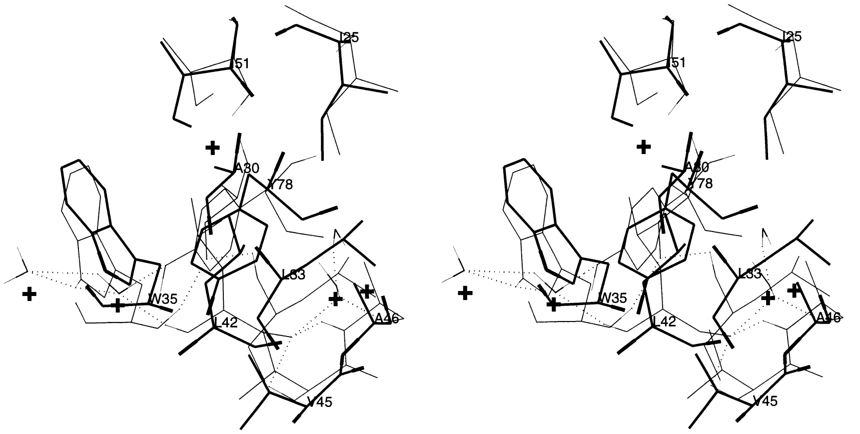


Figure 12. Stereo view of the core 2 residues in the X-ray structure (thick lines) and in the conformation at 200 ps of the low-pH simulation (thin lines). Water molecules of the X-ray structure (bold crosses) and of the low-pH 200 ps conformation (thin lines) are shown, along with the hydrogen bonds of the 200 ps structure (dotted lines).

times more residues than the core 1 of barnase, there are nearly twice as many transitions per residue in LpH 360. This originates from the loosening of the core packing. Specific sidechains that undergo many transitions are Phe7 (32 χ_2 flips from 601 to 1050 ps), Tyr24 (16 χ_2 flips), and the χ_2 of Leu20 (20 transitions), Ile88 (27 transitions) and Ile109 (33 transitions). Ile88 is the central residue of the core, Tyr24 is almost completely buried, while Phe7, Leu20 and Ile109 are on the surface of the core.

Similar, though not identical, behavior is seen at 600 K (Figure 11(b)). The number of transitions increases during the first 150 ps and then reaches a plateau value, indicating that the core is almost completely exposed and rotation of its sidechains is not hindered by the packing. In S600, the decrease of the total transitions in going from the 100 to 150 ps period to the 150 to 200 ps interval is due mainly to the Ile109 sidechain, which undergoes 22 and 11 transitions in the 100 to 150 ps period and the 150 to 200 ps period, respectively. Another example of the loosening of the core packing is provided by χ_2 of Ile88. It experiences 11 transitions during the last 100 ps of B300 and only one during the first 100 ps of the control run at 300 K. All of this is in accord with the conclusion that the conformation of core 1 reached after about 90 ps dynamics at 600 K (A600) is characterized by a looser packing.

One aspect of the denaturation process that has been discussed by Shakhnovich & Finkelstein (1989) is the formation of a relatively compact, partly ordered globule (Ptitsyn, 1992), in which sidechains have considerably more rotational freedom than in the native state. The results of our simulations are consistent with the theory that the looser packing of the globule is an important aspect of the thermodynamics of denaturation. From the simulation of the kinetics of unfolding, the looser packing of the core and its solvation are essentially simultaneous. The theory of Shakhnovich & Finkelstein (1989) is concerned with the thermodynamics of denaturation and, therefore, does not provide information on the kinetics.

Core 2

During the first 300 ps at 300 K and low pH, core 2 has an average accessible surface area of 91 Å^2. The major contribution comes from the exposed surface of the Leu33 sidechain; in the X-ray structure, the Leu33 sidechain has 81 Å^2 of exposed surface. Core 2 contains two water molecules within 6 Å of its center of geometry (Figure 8(b)). One of these donates a bifurcated hydrogen bond to the CO groups of Leu42 and Ala46, and accepts from another water molecule within 7 Å of the core center (Figure 12). The other donates to the OH group of Tyr78 and to another water within 7 Å of the core center, and accepts from the mainchain NH of Leu42 and the sidechain NH of Trp35.

Core 2 expands from 600 to 730 ps due to the separation of the Tyr78 sidechain (loop 4) and the partial solvation of the six-membered ring of the Trp35 sidechain (loop 2); see Figure 1. It then undergoes a significant contraction from 750 to 790 ps (the Leu33, Tyr78 and Trp35 sidechains repack against the rest of the core) followed by a transient solvation of the Tyr78 sidechain at about 810 ps and a quasi-steady period from 820 to 900 ps. The 750 to 790 ps contraction of core 2 and the core 1 contraction from 800 to 850 ps contribute to the stability of the R_g value from 750 to 830 ps (Figure 3(a)). Core 2 then expands with a concomitant influx of water molecules from 850 to 1050 ps. A water molecule in the core accepts from the NH of Ile51 and donates to one or more other water molecules from about 800 ps until the end of the simulation. This corresponds to another core water molecule in the X-ray structure (the one at the top of Figure 12). During the last 50 ps of the trajectory, core 2 contains about seven water molecules within 6 Å of its center of geometry and an average solvent-accessible surface of 256 Å^2. At the end of the simulation, the Leu33, Trp35, Leu42, Val45 and Tyr78 sidechains are partially solvated and the rest of the core is significantly distorted.

At 600 K, core 2 starts to denaturate at the onset of the simulations (Figure 8(d)), with a concomitant

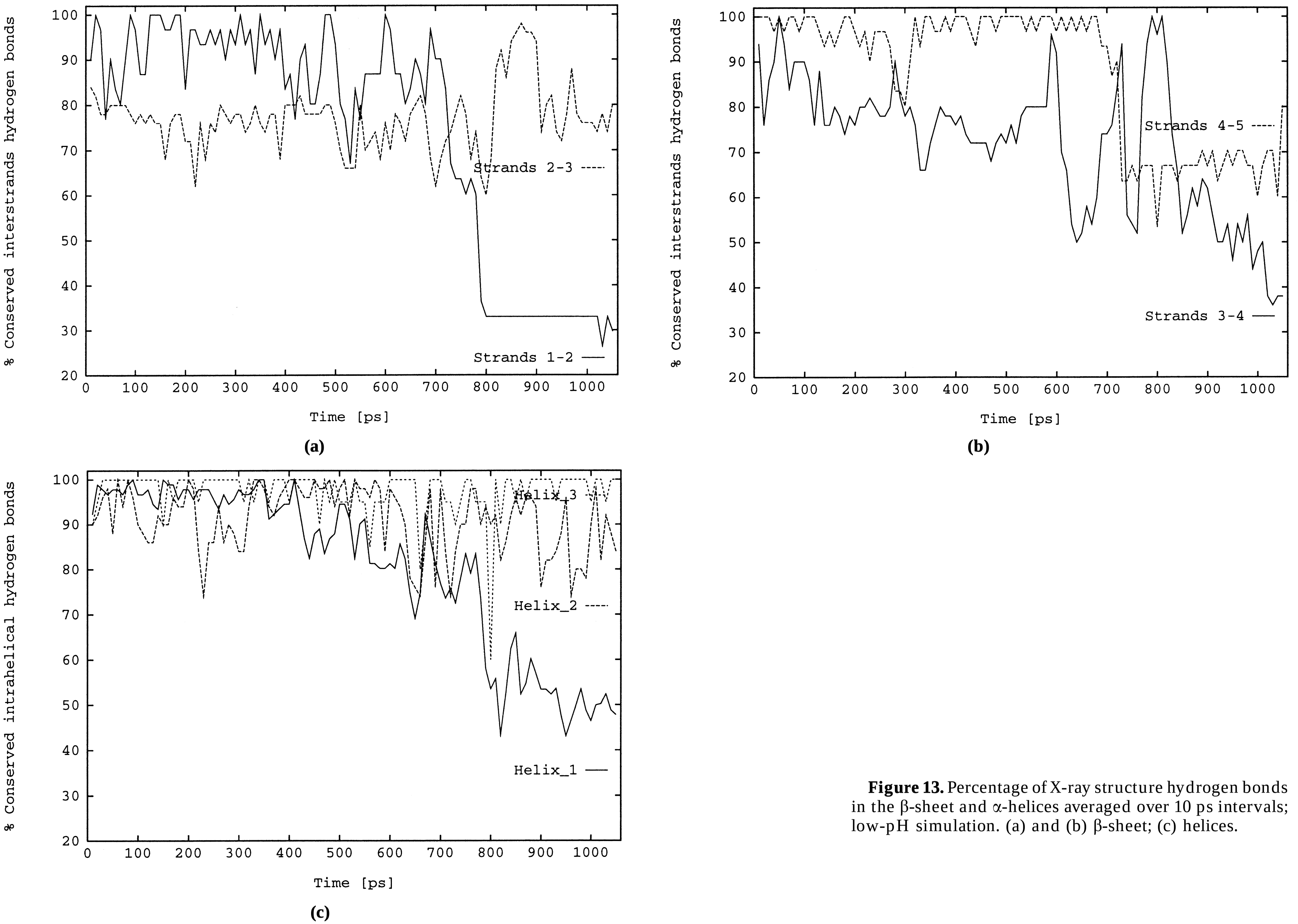


Figure 13. Percentage of X-ray structure hydrogen bonds in the β-sheet and α-helices averaged over 10 ps intervals; low-pH simulation. (a) and (b) β-sheet; (c) helices.

influx of up to 12 water molecules in the first 100 ps. Protein engineering analysis of mutations at Ile25, Val45, Ile51 and Tyr78 (Matouschek *et al*., 1992a) indicates that the interactions made by these residues are broken early during unfolding, in agreement with all the simulations.

Core 3

The Tyr103 sidechain, which is in the β-turn connecting strands 4 and 5, transiently separates from the other core 3 sidechains at about 440 ps and 600 ps, resulting in transient increments of both accessible surface and core water molecules (Figure 8(c)). This is likely to originate from the motion of the β-turn (see above), which contains the Asp101 and His102 sidechains. Core 3 is almost completely compact at about 650 ps; it then undergoes a partial unfolding. The sidechain of Tyr103 is completely solvated at about 750 ps. This is followed by transient solvation of the Phe56 sidechain at about 850 ps and 900 ps, and the Leu63 sidechain at about 900 ps. From 900 to about 950 ps there is a slight contraction, and at 950 ps core 3 is compact, apart from the Tyr103 sidechain, which is surrounded by water molecules. In the last 50 ps, core 3 contains about eight water molecules within 6 Å of its geometric center and has an average exposed surface of 299 Å^2. It is distorted with solvation of the Tyr103 sidechain and part of the Leu63 and Phe106 sidechains. The Leu89 sidechain and the C^{δ} atoms of Leu63 pack against a compact nucleus formed by the aromatic sidechains of Phe56, Trp71, Tyr97 and Phe106. The tight packing of the Leu89 sidechain during the initial phase of unfolding is in agreement with the protein engineering analysis; for core 3, only Leu89 mutants were studied (Matouschek *et al*., 1992a).

In the 600 K simulations core 3 is stable for 50 ps or more (Figure 8(d)). In A600, core 3 is the last core to unfold (after about 75 ps), while in R600 and S600 the three cores begin to denaturate almost simultaneously after 50 ps. Among the core 3 sidechains, Tyr103 is solvated first in agreement with the LpH 360 results. In S600, the exposed surface of the apolar part of Tyr103 increases to 124 Å^2 at 70 ps. This is followed by solvation of the Phe106 sidechain (76 Å^2 at 130 ps) and the Leu63 sidechain (77 Å^2 at 140 ps). In agreement with the mutant data, the Leu89 sidechain is fully buried at the end of S600 and has a very small exposed surface (14 Å^2) at the end of A600. Core 3 is fully solvated after 230 ps of R600, while in A600 and S600, the Trp71, Leu89 and Tyr97 sidechains still form a compact hydrophobic nucleus at the end of the simulation.

Behavior of the β-sheet

Barnase is an α + β protein with three α-helices in the first 45 residues and a five-stranded antiparallel β-sheet in the last 65 residues (Figure 1). In the native structure, the β-sheet has a regular twisted shape, apart from a β-bulge in strand 1 between residues 53 and 54. The twist (defined as the angle formed by the N terminus to C terminus vector of strands 1 and 5) is about 90°. Figure 13(a) and (b) show the percentage of conserved interstrand hydrogen bonds in the low-pH simulation calculated at the end of each picosecond and averaged over 10 ps intervals; there are a total of three, five, five and three hydrogen bonds between strands 1 and 2, 2 and 3, 3 and 4, and 4 and 5 in the X-ray structure. Transient rupture of some 1 to 2 and 4 to 5 interstrand hydrogen bonds take place during the initial 600 ps. In the central part of the β-sheet, apart from the 76 NH . . . CO 86, four of the five hydrogen bonds between strands 2 and 3 are conserved during most of the simulation. Only the hydrogen bond 87 NH . . . CO 99 (strands 3 to 4) is broken during most of the first 600 ps. This broken hydrogen bond is likely to originate from the structural changes associated with the neutralization of the Asp86 charge at low pH.

When the temperature is raised to 360 K (at 600 ps), there is a progressive destabilization of the hydrogen bonds between strands 1 and 2 (Figure 13(a)), and a significant oscillating behavior in those between strands 3 and 4 (Figure 13(b)). The hydrogen bonds between strands 4 and 5 begin to come apart after about 700 ps. In the last 50 ps of the simulation, from 1000 to 1050 ps, only the 54 CO . . . NH 73 hydrogen bond is present between strands 1 and 2, while the percentage of conserved strands 3 to 4 and 4 to 5 hydrogen bonds oscillates around 45% and 65%, respectively. The rupture of several hydrogen bonds between strands 3 and 4 is coupled to the separating motion of strands 4 and 5 from the rest of the β-sheet (Figure 4).

At 600 K the hydrogen bonds between strands 1 and 2 come apart early in the unfolding process. The most stable interactions appear between strands 2 and 3, and strands 3 and 4. In the R600 simulation strands 2 and 3 separate transiently from 140 to 180 ps, and strands 3 and 4 come apart at about 130 ps.

To illustrate the role of the solvent in the denaturation of the β-sheet, some snapshots from the molecular dynamics simulation of LpH 360 are shown in Figure 14. Water molecules that spent more than 200 ps within 3 Å of any mainchain atom of the β-sheet are labeled (from A to M). As in the high-temperature simulations, solvent insertion begins at the edges. At 620 ps, water C donates to the CO groups of residues 51 (strand 1) and 73 (strand 2), and accepts from the NH of residue 54 in the β-bulge (Figure 14(a)). Water molecule L has inserted between strands 4 and 5; it acts as acceptor for the NH of residue 100 (strand 4) and donates to the CO of residue 105 (strand 5). At the N-terminal part of strand 2, water molecule A donates to the CO of residue 71, while a hydrogen atom of water molecule B makes a bifurcated bond to the CO groups of residues 70 and 90. Water molecule D donates to the CO of residue 76 at the C-terminal part. There are also several unlabeled water molecules (i.e. waters that spent from 1 to 200 ps within 3 Å of any mainchain atom of the β-sheet) that are hydrogen-bonded to the terminal polar groups of all five

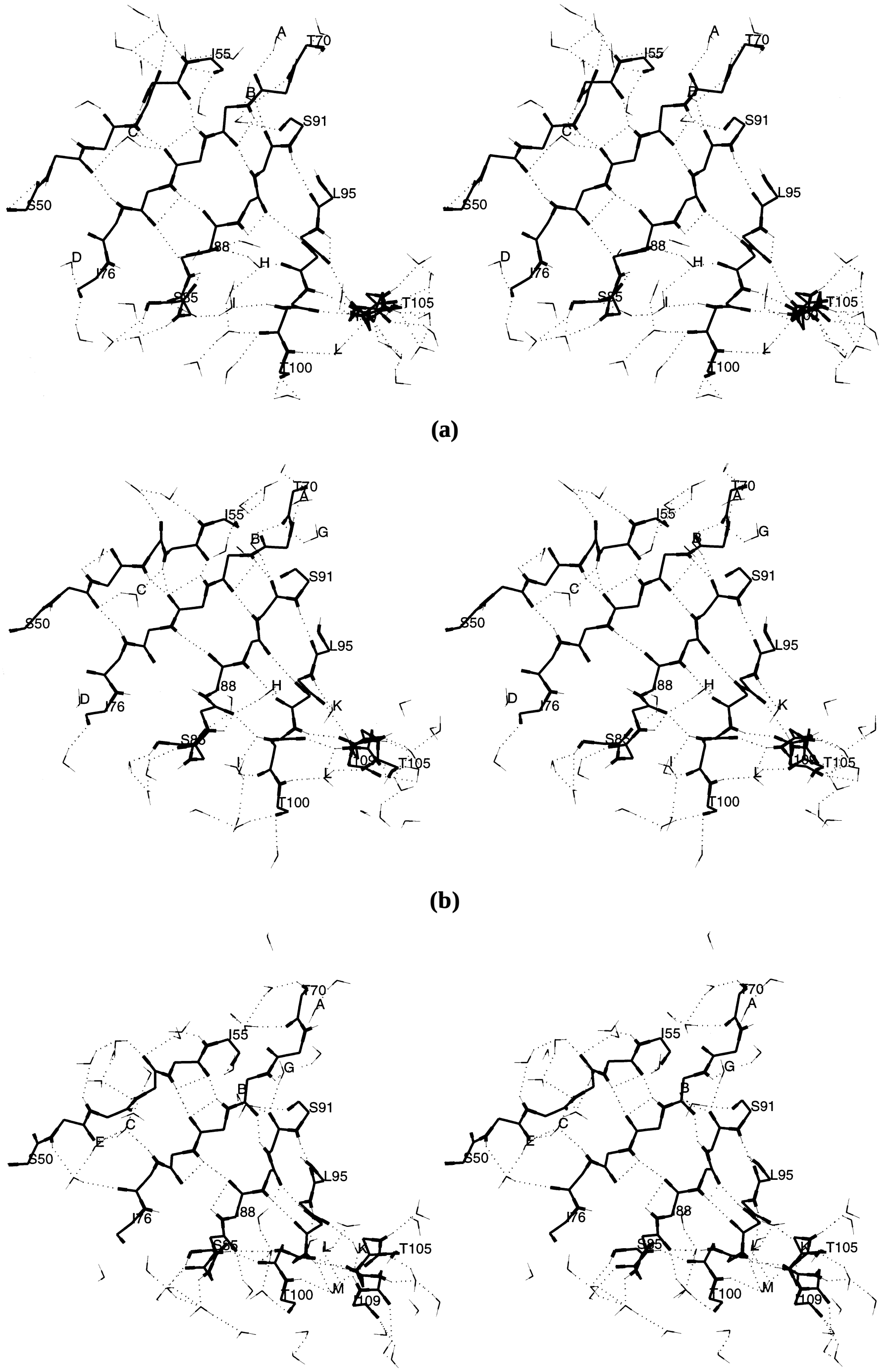


Figure 14. Stereo view of water penetration into the β-sheet during low-pH dynamics: (a) 620 ps; (b) 700 ps; (c) 800 ps. Mainchain N and O atoms are thick lines, hydrogen atoms are thin lines, and hydrogen bonds are dotted lines; water molecules within 3 Å of any mainchain atom of the β-sheet are shown. Water molecules that spent more than 200 ps within 3 Å of any mainchain atom of the β-sheet are labeled from A to M and discussed in the text. The orientation is chosen to clearly show the first four strands so that strand 5 is nearly perpendicular to the page.

strands. A cluster of water molecules penetrates between strands 3 and 4; among these, water molecule H donates to the CO groups of residues 87 (strand 3) and 97 (strand 4), and accepts from water molecule I, which acts as acceptor for the NH of residue 99 (strand 4).

At 700 ps, certain of the water molecules between strands 3 and 4 are no longer present, although water molecule H is still donating hydrogen bonds to the CO groups of residues 87 and 97 (Figure 14(b)). Water molecule B, which previously donated a bifurcated hydrogen bond to the CO groups of residues 70 (strand 2) and 90 (strand 3), now bridges these CO groups.

Two water molecules have inserted between strands 1 and 2 at 800 ps (Figure 14(c)); water molecule C accepts now from the NH of residue 75 and donates to the CO of residues 51 and 52; an unlabeled water molecule accepts from the NH of residue 51 and donates to the CO of residue 75. Water molecule E bridges these two solvent molecules. The five native hydrogen bonds between strands 3 and

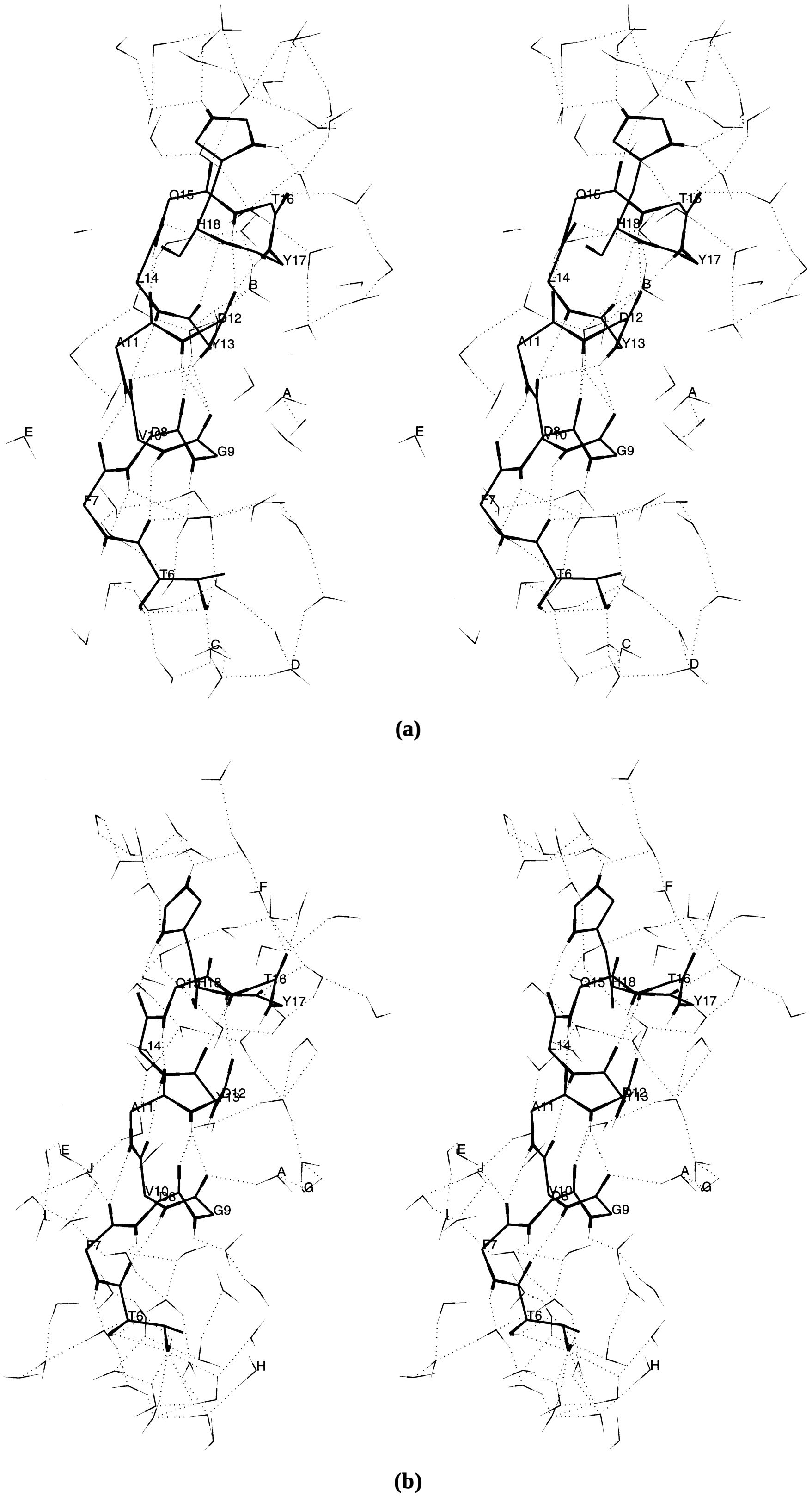
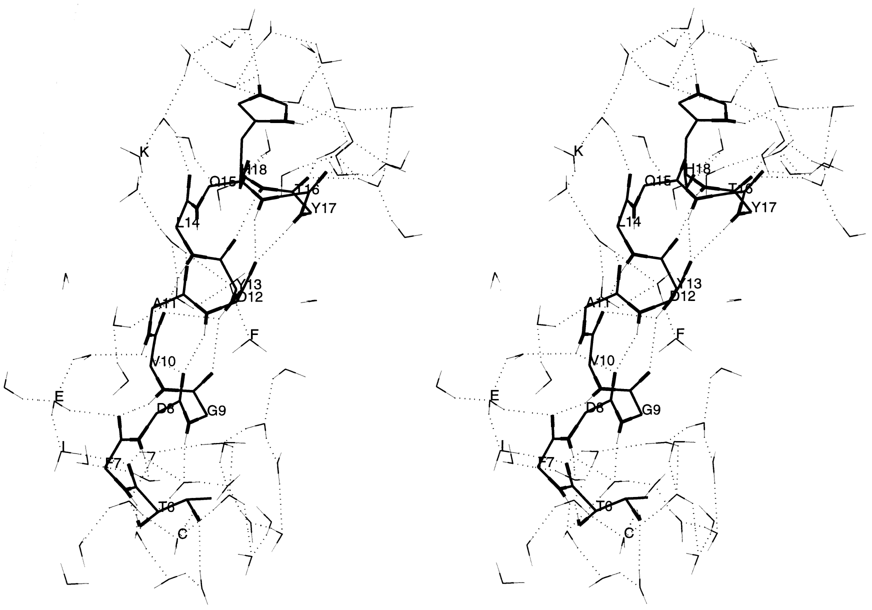


Figure 15(a–b) (*legend opposite*)



(c)

Figure 15. Stereo view of water around helix 1 during dynamics: (a) 600 ps; (b) 700 ps; (c) 800 ps. Mainchain N and O atoms are thick lines, hydrogen atoms are thin lines, and hydrogen bonds are dotted lines. The mainchain of the helix is shown along with the sidechains of the N and C-cap residues (Thr6 and His18). Water molecules within 4.5 Å of any atom of helix 1 (mainchain and sidechain of terminal residues) are shown. Water molecules that spent more than 150 ps within 4.5 Å of any mainchain atom of the α-helix, are labeled from A to M and discussed in the text.

4 are all reformed, although transiently (see Figures 14(c) and 13(b)).

Figure 14(a) to 14(c) illustrate how water molecule C inserts between strands 1 and 2 by utilizing the free NH of residue 54 in the β-bulge; they show also how water molecule L is displaced from strands 4 and 5 by water molecule M. Another insertion pattern is seen with water molecule G, which approaches the β-sheet by donating to the CO of residue 70 (Figure 14(b)); it then inserts between strands 2 and 3 by donating to the CO of residue 90 and accepting from the NH of residue 72 (Figure 14(c)).

The water insertion between edge strands is sustained; at 900 ps clusters of three to five water molecules are visible between strands 1 and 2, 3 and 4, and 4 and 5. The central strands are not significantly distorted apart from the insertion of water molecule J, which accepts from the NH of residue 87 (strand 3) and donates to the CO of residue 99 (strand 4). The 76 NH . . . CO 86 hydrogen bond at the edge of strands 2 and 3 has transiently reformed so that the five native hydrogen bonds between strands 2 and 3 are present.

Towards the end of the simulation, the β-sheet twist increases by about 30° and strand 1 has almost completely separated from strand 2 (apart from the 54 CO . . . NH 73 hydrogen bond). Furthermore, most of the 3 to 4 and 4 to 5 interstrand hydrogen bonds are replaced by water molecules. The interactions between strands 2 and 3 are preserved.

The above results indicate several possible pathways of water insertion, most of which involve mainchain polar groups at the edges of the β-sheet. Water molecules that participate in two or more hydrogen bonds with polar groups on adjacent strands (e.g. C, E, H, L and M), are particularly stable and can interact with mainchain polar groups on the β-sheet for more than 200 ps at 360 K. This is not always true; e.g. the unlabeled water molecule between strands 1 and 2 in Figure 14(c) moves away from the β-sheet in less than 40 ps.

The early stage of the β-sheet solvation is similar in the LpH 360 simulation and the previous high-temperature simulations (Caflisch & Karplus, 1994a,b). At 600 K the β-sheet disruption started near the β-bulge (residues 53 and 54 on strand 1) and at the edges (strands 1 and 5). It was associated with an increase in the twist and an influx of water molecules, some of which inserted between adjacent strands and participated in hydrogen bonds as both donors and acceptors with the mainchain polar groups.

Behavior of the α-helices

The first α-helix (residues 6 to 18) has three turns and a regular pattern of *i* to *i* + 4 hydrogen bonds in the crystal structure. The sidechain of Thr6 acts as the N-cap residue; its sidechain OH group accepts from the mainchain NH of residue 9. The C-cap residue is His18, whose sidechain NH group acts as donor for the mainchain CO group of residue 15 (Šali *et al.*, 1988). α-Helix 2 (residues 26 to 34) has two turns and packs against the edge of the antiparallel β-sheet, loop 2 and the small one-turn α-helix 3 (residues 41 to 45). Figure 13(c) shows the percentage of native *i* to *i* + 4 intrahelical hydrogen bonds during the low-pH dynamics run. α-Helix 1 starts to fray after about 400 ps; it loses five out of its nine hydrogen bonds by the end of the simulation. Helices 2 and 3 show some fluctuations, but between 80 and 100% of the native H bonds remain intact.

The hydrogen bonds involving the cap residues of α-helix 1 are marginally stable in the simulation. The C-cap hydrogen bond, which consists of His18 sidechain $N^{\epsilon}H$ . . . CO 15, transiently breaks and reforms several times during the first 600 ps. It is then replaced by the His18$N^{\delta}H$ . . . CO 15 hydrogen bond of the doubly protonated histidine from 710 to 760 ps and from 810 to 830 ps, and the His18 $N^{\delta}H$ . . . CO 14 from 530 to 550 ps and from 690 to 760 ps. The N-cap hydrogen bond breaks in the first 10 ps and does not reform. Instability of the N and C-cap hydrogen bonds is seen also in the high-temperature simulations.

To focus on the nature of motions of water molecules interacting with helix 1 during the early stage of the barnase unfolding process in LpH 360, some snapshots from the molecular dynamics simulation are presented in Figure 15. Most of the water molecules that spent more than 150 ps within 4.5 Å of any mainchain atom of helix 1 or sidechain atom of its caps, are labeled (from A to M). At 600 ps, there is little distortion in the intrahelical hydrogen-bonding pattern, whereas the N and C-cap sidechains are completely solvated; their polar groups participate in hydrogen bonds with water molecules (Figure 15(a)). At the C terminus, the NH of residue 18 donates a bifurcated hydrogen bond to the CO of residues 13 and 14. The hydrophilic side of the helix is in contact with water molecules (some of which donate to the CO groups of residues 6, 8, 15 and 16), whereas the hydrophobic side (Phe7, Val10, Ala11 and Leu14) does not interact with the solvent (the orientation in Figure 15 is the same as in Figure 1; hence, the main hydrophobic core, which is omitted for clarity, is located on the left of the helix).

At 700 ps, water molecules E, A and F act as donors for the CO groups of residues 7, 9 and 16, respectively (Figure 15(b)). Some unlabeled water molecules accept from the NH of residues 6 to 9. The 9 CO . . . NH 13 hydrogen bond has transiently broken. The C-terminal hydrogen bond (14 CO . . . NH 18) is lost and does not reform; the NH of residue 18 donates to the CO of residue 13. This i–$i+5$ hydrogen bond at the C terminus is present from 500 ps to the end of the simulation (Table 3A). Along with the 12 CO . . . NH 17 hydrogen bond, it forms one half turn of a π-helix, which is stable. The C-cap has transiently reformed.

Partial unfolding of the N-terminal turn is visible at 800 ps (Figure 15(c)), due to the solvation of the CO group of residues 6 and 7 by water molecules I and E, which are connected by a hydrogen bond. Further, the 7 CO . . . NH 10 3_{10} hydrogen bond has formed. Complete unfolding of the N-terminal part is evident at 900 ps, except for the 3_{10} hydrogen bond 6 CO . . . NH 9 (Table 3A). Near the main hydrophobic core, water molecule J accepts from the NH of Ala11 and donates to an unlabeled water molecule, which in turn donates to the CO of Phe7. At the C terminus, water molecule K donates two hydrogen bonds, characterized by a poor geometry, to the CO groups of residues 14 and 15; it also donates a good hydrogen bond to another water molecule, which is part of a chain of water molecules surrounding the imidazole ring of His18.

Toward the end of the simulation the N-terminal part of helix 1 is completely surrounded by water molecules, whereas the i–$i+4$ hydrogen bonds 9–13, 10–14, 11–15, 12–16 and 13–17 are stable (CO 13 accepts also from the NH of residue 18). Water molecule D donates to the CO group of Thr6 at 1050 ps. An unlabeled water molecule has inserted into the first turn; it donates to the CO group of residue 6 and accepts from the NH of residue 10. The helix axis has rotated clockwise by about 30°, bringing its unfolded N-terminal part closer to the rest of the protein.

The behavior of helix 1 in LpH 360 is similar to that observed in the 600 K simulations; however, in the latter it began to unfold at both termini (Table 3B to D). Further, several α to 3_{10} and 3_{10} to α conversions took place in both the acid-mediated and the high-temperature denaturations. Often, but not always, this transition precedes the insertion of a water molecule. In addition, the NH groups of residues 17 and 18 form a short piece of π-helix in all simulations except A600. This element of secondary structure is stable at low pH and 360 K (Table 3A); it is rapidly solvated or reconverts to α-helical geometry at 600 K and neutral pH (Table 3D and E).

The α-helices 2 and 3 behave differently in the LpH 360 and high-temperature simulations. At low pH, the N-terminal part of α-helix 2 shows some instability after the first 620 ps (Figure 13(c)); the 26 CO . . . NH 30 hydrogen bond comes apart and reforms several times, while the rest of it is stable. α-Helix 2 experiences a substantial degree of unfolding in the first 150 ps at 600 K; it denatures completely towards the end of R600, the longest high-temperature simulation. The one-turn α-helix 3 is conserved throughout the low-pH simulation apart for a short transient rupture of the 41 CO . . . NH 45 hydrogen bond from 793 to 799 ps duje to the insertion of a water molecule. This is different from the 600 K simulations, where helix 3 unfolds in the first 30 ps of A600 and R600, and after 150 ps of S600.

The early unfolding of the N terminus of α-helices 1 and 2, and the stability of α-helix 3 in the low-pH simulation is in agreement with the experimental data reported by Fersht and co-workers (Serrano *et al*., 1992c). In the high-temperature simulations, entropic effects may have played a role in the early denaturation of helices 2 and 3, which are on the edges of the barnase subdomain formed by residues 25 to 49. At such high temperatures a flexible polypeptide chain, with its polar groups participating in hydrogen bonds to water, has a lower free energy than a more rigid helix with enthalpically stronger intra-helical hydrogen bonds.

Conclusions

A molecular dynamics study of the initial stages of the low-temperature (360 K), acid-mediated unfolding of barnase has been made to compare with high-temperature simulations (600 K); some of the results of the high-temperature simulations have been reported previously (Caflisch & Karplus, 1994a). The low-temperature results are expected to be more realistic and to increase the time-scale of the transition. It is found that the unfolding behavior is very similar in the two sets of simulations. The cooperative nature of the initial stages of the denaturation process and the important role of the water molecules for the unfolding of both secondary structural elements and hydrophobic cores found at high temperature are confirmed in the low-temperature denaturation. The cooperative behavior in the low-pH simulation is manifested in the solvation of the edge sidechains of core 1 (Phe7, Leu20, Tyr90, Trp94 and Ile109), which is coupled to the separating motion of helix 1 from the β-sheet and the essentially simultaneous denaturation of its N-terminal part (residues 6 to 10), the movement of loop 1 (residues 19 to 25), and the unfolding of the protein C terminus (residues 109–110). The disruption of the three hydrophobic cores when the temperature is raised to 360 K is nearly simultaneous.

The early unfolding of the N terminus, loops and the external strands of the β-sheet, along with the early penetration of water molecules into the cores are similar to the high-temperature simulations, though a smaller amount of water penetrates the main hydrophobic core during the initial phase of the expansion. The stability of helices 2 and 3 constitutes the most significant difference between the low-pH, low-temperature simulation and the high-temperature simulations. Entropic effects may have played a role in the early unfolding of helices 2 and 3 at high temperature. In addition, since the denaturation process is faster at high temperatures, the barnase molecule reached a higher degree of unfolding in the 600 K simulations than at 360 K.

Several interesting aspects are revealed by an analysis of the barnase self energy and barnase-solvent energetics during unfolding. First, the barnase bonding energy reflects quasi-harmonic behavior; i.e. the energy in the bonding terms increases proportionally to the increase in temperature. Second, in all simulations the total barnase self energy is characterized by an initial sharp increase in less than 50 ps, with a slower increase during the remainder of the simulation. The latter originates from the intra-barnase non-bonding interactions. Third, the barnase-water interactions improve during unfolding at the expense of the intra-barnase non-bonding energy. Finally, the behavior of the van der Waals energy suggests that much of the energy increase during the early phase of unfolding originates from the rupture of the tight packing of the protein molecule. This is in accord with the suggestions made by Shakhnovich & Finkelstein (1989).

Finite-difference Poisson-Boltzmann calculations indicate that the Asp8-Arg110-Asp12 double salt-bridge is unstable in the native structure of barnase. From the marginal stability of the salt-bridge in the room-temperature simulation, it appears that the solvation free energy estimated from the Poisson-Boltzmann equation is too large.

The detailed analysis of the acid-mediated unfolding simulation reveals a number of characteristics that may play a more general role in protein denaturation. Some external sidechains of the cores undergo a ‘‘breathing’’ motion during the initial phase of unfolding. In particular, the aromatic rings of Phe7 in core 1, Tyr78 in core 2 and Tyr103 in core 3 separate and then partially repack one or more times. In addition, the partial distortion of the main hydrophobic core induces the formation of cavities among apolar sidechains. These are large enough to accommodate at least one water molecule and can last up to about 50 ps at 360 K. There are two sites in core 1 that are affected by cavity formation. The first is the region formed by the sidechains of Val10, Leu14, Leu20, Tyr24, Ala74, Ile76 and Tyr90, and the second is on the core surface, between the Phe7 and Ile109 sidechains. Water molecules tend to penetrate the cores in chains or clusters. The polar groups of Tyr and Trp sidechains play an important role in the penetration of water, as was found previously in the high-temperature simulations; i.e. the penetrating water molecules make hydrogen bonds to these groups. This suggests that mutants with a core Tyr substituted by Phe, for example, could have a higher activation free energy for unfolding, if the polar group sidechain is not involved in essential hydrogen bonds with the rest of barnase in its native structure. The only available experimental result (the Tyr78 → Phe mutation) leads to a decrease of 1.4 kcal/mol in the activation free energy (Matouschek *et al*., 1989), since the Tyr78 OH is hydrogen-bonded to the mainchain NH and CO of Gly81 in the native structure.

The β-sheet disruption starts at the edges and is coupled to an increase in the twist. The N-terminal part of α-helix 1 unfolds early and the helix caps are unstable, even in the room-temperature simulation. The essential factor in the denaturation of the β-sheet and α-helix is that water molecules compete with the existing hydrogen bonds of the secondary structure. The water molecules attack primarily the carbonyl groups. Some of the water molecules insert and replace interstrand hydrogen bonds in the β-sheet and intrahelical hydrogen bonds in the helix. The inserted water molecules act as both donors and acceptors. In the β-sheet, they can insert by donating hydrogen bonds to CO groups on adjacent strands.

Tests of some of the results reported here would be of great interest. Photochemically induced nuclear polarization (Kaptein *et al*., 1978), nuclear Overhauser effects of water interactions with specific residues (Otting & Wüthrich, 1989), chemical markers (Ghelis, 1980), and NMR of ^{19}F substituted amino acids (Frieden *et al*., 1993) are possible approaches. In addition, the early unfolding behavior could be investigated by nanosecond laser photolysis (Jones *et al*., 1993) and nanosecond temperature jump experiments (M. Grübele, personal communication).

Methods

The method used is essentially the same as that employed in the previous high-temperature simulations (Caflisch & Karplus, 1994a). The CHARMM (param19, Brooks *et al*., 1983) polar hydrogen approximation (aliphatic and aromatic hydrogen atoms are considered as part of the carbon atoms to which they are covalently bound) was used and the modified TIP3P model was employed for the water molecules (Jorgensen *et al*., 1983; Reiher, 1985). Polar hydrogen atoms were added to the coordinates of the X-ray crystal structure of barnase (Dr A. Cameron and Professor G. Dodson kindly provided a structure originating from data collected on crystals at pH 9) with the HBUILD option (Brünger & Karplus, 1988) of the CHARMM program. To reproduce low pH conditions, the acidic residues (Asp and Glu) and the C terminus were neutralized by adding a hydrogen atom to one of the two sidechain oxygen atoms; for the position of the ionizable residues in the barnase chain, see Figure 1. HBUILD was run twice for each acidic residue to separately protonate each sidechain oxygen atom. This made it possible to select the oxygen atom for protonation on the basis of the more favorable interaction energy with the rest of the protein. The two His sidechains were double-protonated. This yields a barnase total charge of +17 electronic charges; at neutral pH, the net charge is +2. The resulting structure was minimized by 200 steps of the steepest descent algorithm; its RMSD from the pH 9 X-ray structure was 0.16 Å for the backbone heavy-atoms and 0.19 Å for all heavy-atoms. This is similar to the neutral pH minimized structure, which deviated by 0.16 Å (backbone) and 0.20 Å (all) from the pH 9 X-ray structure.

Standard molecular dynamics (Brooks *et al*., 1988) was used to simulate barnase (1106 atoms) centered in a sphere of 30 Å radius consisting of 3006 water molecules. A deformable boundary potential was used (Brooks & Karplus, 1983, 1989). The temperature of the system was controlled by weak coupling to an external bath (Berendsen *et al*., 1984); the coupling parameter was 5 ps. After 300 ps of equilibration at 300 K, the temperature was increased to 327 K (the experimental denaturation temperature at neutral pH; Kellis *et al*., 1989) for 300 ps. Since denaturation did not take place, the temperature was raised to 360 K and the trajectory was run for another 450 ps (the 600 to 1050 ps trajectory at low pH and 360 K is called LpH 360). Thus, the low-pH trajectory was run for a total of 1050 ps; a time-step of 2 fs was employed with the leapfrog integration algorithm and structures were saved every 0.5 ps.

The non-bonding cutoff distance was increased from 8 Å to 11 Å at 300 ps, with the idea that it might accelerate unfolding because of the large net positive charge of the molecule. This was not the case; actually the larger cutoff did not affect the barnase electrostatic self energy, but it did increase the barnase-water electrostatic energy. The barnase molecule was almost completely compact until 600 ps. At that time the temperature was increased to 360 K and the non-bonding cutoff distance was returned to 8 Å; nearly twice as much CPU time was required with an 11 Å cutoff than with an 8 Å cutoff.

Three denaturation simulations were performed at 600 K and neutral pH (A600, 120 ps; R600, 250 ps; S600, 200 ps; A600 and R600 were described briefly by Caflisch & Karplus, 1994a) and a 300 K control trajectory at neutral pH was run for 310 ps; the latter is important for demonstrating that the protein is stable under the appropriate conditions. In these simulations, both His18 and His102 were singly protonated at N^{δ}. The acidic residues and C terminus were deprotonated, and the Lys, Arg and N terminus were protonated. The A600, R600 and S600 simulations were started after 4 ps, 100 ps and 258 ps of the control run. A comparison of these simulations makes it possible to determine whether the results are sensitive to the initial conditions. The A600 simulation was branched at 90 ps; the system was cooled to 300 K and the simulation was continued for 160 ps at 300 K (B300). This simulation was performed to investigate whether a stable intermediate structure could be trapped.

To calculate the solvent-accessible surface area, the CHARMM implementation of the algorithm of Lee & Richards (1971) and a probe sphere of 1.4 Å radius were utilized. The formation of cavities in the main hydrophobic core (core 1) was investigated by determining the solvent-accessible volume with a finite grid (COOR SEARCH command within CHARMM). The gridsize was 0.2 Å and the grid dimensions were selected to encompass all core 1 sidechain atoms. The van der Waals radii of all solute and solvent atoms were augmented by 1.4 Å for the grid search and all empty grid points were determined. In analogy with the definition of solvent-accessible surface area, the solvent-accessible volume is the space in which the center of a sphere of radius 1.4 Å and volume of 11.5 $Å^3$ corresponding to a water molecule can move without contacting any occupied grid point. In addition, the total empty volume was calculated to investigate the presence of microcavities; for this purpose, the original (i.e. non-augmented) van der Waals radii were used.

Transitions of the torsion angles of the core 1 sidechains were enumerated by using CHARMM (MONI DIHE command). A transition is defined as a change in the dihedral angle that results in the passage from one well of the torsion potential to another well, and that involves a displacement of at least 30° beyond the position of the potential maximum. As an example, for rotation about a bond between isolated tetrahedral carbon atoms, the minima are at 60°, 180° and −60°, while the maxima are at 0°, 120° and −120°. For an initial angle of 45° there is a transition if the angle becomes >150° or < −30°. The original minimum was at 60°, and the new minimum would be 180° or −60°, respectively. The angle can change by as much as 120° or as little as 60° in going from one well to the next using this counting method. It gives a conservative measure of the barrier-crossing frequency. The criteria used are those of Loncharich and Brooks (1990).

In addition to the simulation analysis of the Asp8-Arg110-Asp12 salt-bridge, the electrostatic contribution to the free energy was calculated by numerical solution of the linearized Poisson-Boltzmann (LPB) equation. Studies have shown that the finite difference method (Press *et al*., 1986) applied to solve the LPB equation represents a useful procedure for estimating electrostatic interactions in proteins (Warwicker & Watson, 1982; Gilson & Honig, 1988; Bashford & Karplus, 1990; Davis *et al*., 1991). It approximates the differential equation by a set of finite difference equations on a grid. The latter are solved on a computer by iterative adjustments of the value of the potential at each grid point. The UHBD software package (Davis & McCammon, 1989, 1990; Davis *et al*., 1991) was utilized for solving the finite difference LPB equation in this study. The CHARMM param19 partial charges (on heavy-atoms and polar hydrogen atoms) and van der Waals atomic radii were used. Lim *et al*. (1991) have shown that the solvation free energy of models of polar and ionizable sidechains calculated with the LPB method and the CHARMM param19 set agree reasonably well with experiments. UHBD places the charges on a grid according to the trilinear weighting method developed by Edmonds *et al*. (1984). Coulombic potentials for each point charge were used to set the boundary potential. First, a cubic grid of $50 \times 50 \times 50$ points was used with a grid spacing of 2.0 Å; this yields a layer of solvent (high dielectric, $\epsilon = 78$) of at least 20 Å around the barnase structure (low dielectric, $\epsilon = 4$). A second calculation, focused on the individual sidechains involved in the double salt-bridge, was performed with a cubic grid of 60 points per side and a grid spacing of 0.5 Å. Focusing has been shown to dramatically reduce the error originating from the distribution of the charges on the grid points (Gilson *et al*., 1988; Mohan *et al*., 1992). The protein value of 4 for the dielectric constant accounts, in an average way, for the effects of electronic polarizability and orientational polarizability of the permanent dipoles (Gilson & Honig, 1986; Harvey, 1989). The value of 78 for the solvent represents the dielectric constant of bulk water at 298 K. To define the surface of the low dielectric region, a radius of 1.4 Å was chosen for the solvent probe. Furthermore, dielectric boundary smoothing was used (linear interpolation of the permittivities at the midpoints between grid points intersecting the dielectric boundary), since it provides a more accurate description of the potential near the discontinuity region and has been shown to improve convergence (Davis & McCammon, 1991; Mohan *et al*., 1992). Values of 300 K for the temperature, 100 mM for the ionic strength, and 2.0 Å for the Stern layer were selected. LPB calculations were done for the coordinate sets obtained from the neutral pH control run (NpH 300) at 300 K; barnase coordinates at 1 ps intervals were used.

The electrostatic energy of a molecular assembly is influenced in two ways by the surrounding solvent. First the interactions of the solute charges are screened by the solvent, and second the solvent ineracts directly with each solute charge (self-energy). To calculate the electrostatic contribution to the free energy, the sidechain of Arg110 was charged (CHARMM param19 partial charges) and the protein was considered as a neutral region of low dielectric, which displaces the solvent; i.e. no other charge was included. The calculation was performed first with an exterior dielectric of 78 and then with an exterior dielectric constant of 1, and the energies were subtracted to obtain the self-energy of the Arg110 sidechain. The same was done for the Asp8 and Asp12 side-chains. There the field generated by the positively charged Arg110 sidechain at the positions of the Asp8 and Asp12 sidechains was multiplied by the partial charges on the sidechains of Asp8 and Asp12 (both negatively charged). This yields the electrostatic interaction energy between Arg110 and Asp8, and between Arg110 and Asp12 as mediated (screened) by the solvent; i.e. the electrostatic free energy of interaction. The same was done for the field of Asp8 at the position of the Asp12 partial charges. Indirect contributions due to interactions with other charges of the protein are not included. Test calculations performed every 10 ps yielded negligible values for these contributions, in agreement with the experimental and theoretical results presented by Loewenthal *et al*. (1993), which showed that the effect of long-range charge-charge interactions on the Asp8-Arg110-Asp10 salt-bridge is small. In a recent analysis of salt-bridge strengths by Hendsch & Tidor (1994), the salt-bridge Asp12 to Arg110 in barnase is considered (see their Table 1). They find a large protein contribution to the stabilization; from our calculations this is almost entirely due to Asp8.

Acknowledgements

The work was supported in part by the National Science Foundation and a gift from Molecular Simulations, Inc. A.C. was supported by the Ciba-Geigy Jubilaeums Stiftung and the Roche Research Foundation.

We thank Dr A. Blondel, Dr L. Caves and Dr M. Schaefer for helpful discussions. The atomic coordinates of barnase were kindly provided by Dr A. Cameron and Professor G. Dodson of the University of York. The calculations were performed on a DEC Alpha 3000, an IBM RISC 6000/550, a SGI 340 GTX and on the CRAY C90 at the Pittsburgh Supercomputing Center.

References

Arcus, V. L., Vuilleumier, S., Freund, S. M. W., Bycroft, M. & Fersht, A. R. (1997). Toward solving the folding pathway of barnase: the complete backbone ^{13}C, ^{15}H and ^{1}H assignments of its pH-denatured state. *Proc. Natl Acad. Sci. USA*, **91**, 9412–9416.

Baudet, S. & Janin, J. (1991). Crystal structure of a barnase-d(GpC) complex at 1.9 Å resolution. *J. Mol. Biol.* **219**, 123–132.

Bashford, D. & Karplus, M. (1990). The pK_a's of ionizable groups in proteins: atomic details from a continuum electrostatic model. *Biochemistry*, **29**, 10219–10225.

Ben-Naim, A. (1975). Hydrophobic interaction and structural changes in the solvent. *Biopolymers*, **14**, 1337–1355.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.

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

Brooks, C. L., III & Karplus, M. (1983). Deformable stochastic boundaries in molecular dynamics. *J. Chem. Phys.* **79**, 6312–6325.

Brooks, C. L., III, Karplus, M. & Pettitt, B. M. (1988). Proteins: a theoretical perspective of dynamics, structure and thermodynamics. In *Advances in Chemical Physics*, vol. 71, pp. 1–259, John Wiley & Sons, New York.

Brooks, C. L., III & Karplus, M. (1989). Solvent effects on protein motion and protein effects on solvent motion. *J. Mol. Biol.* **208**, 159–181.

Brünger, A. T. & Karplus, M. (1988). Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins: Struct. Funct. Genet.* **4**, 148–156.

Bycroft, M., Matouschek, A., Kellis, J. T., Jr, Serrano, L. & Fersht, A. R. (1990). Detection and characterization of a folding intermediate in barnase by NMR. *Nature*, **346**, 488–490.

Bycroft, M., Ludvigsen, S., Fersht, A. R. & Poulsen, F. M. (1991). Determination of the three-dimensional solution structure of barnase using nuclear magnetic resonance spectroscopy. *Biochemistry*, **30**, 8697–8701.

Caflisch, A. & Karplus, M. (1994a). Molecular dynamics simulation of protein denaturation: solvation of the hydrophobic cores and secondary structure of barnase. *Proc. Natl Acad. Sci. USA*, **91**, 1746–1750.

Caflisch, A. & Karplus, M. (1994b). Molecular dynamics studies of protein and peptide folding and unfolding. In *The Protein Folding Problem and Tertiary Structure Prediction* (Merz, K. M., Jr & LeGrand, S. M., eds), pp. 193–230, Birkhauser, Boston.

Constantine, K. L., Goldfarb, V., Jeffrey, P. D., Sheriff, S. & Mueller, L. (1992). Recombinant anti-digoxin antibody V_L domain. *Biochemistry*, **31**, 5033–5043.

Creighton, T. E. (1990). Protein folding. *Biochem. J.* **270**, 1–16.

Daggett, V. & Levitt, M. (1992). A model of the molten globule state from molecular dynamics simulations. *Proc. Natl Acad. Sci. USA*, **89**, 5142–5146.

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

Davis, M. E. & McCammon, J. A. (1989). Solving the finite difference linearized Poisson-Boltzmann equation: a comparison of relaxation and conjugate gradient methods. *J. Comput. Chem.* **10**, 386–391.

Davis, M. E. & McCammon, J. A. (1990). Calculating electrostatic forces from grid-calculated potentials. *J. Comput. Chem.* **11**, 401–409.

Davis, M. E. & McCammon, J. A. (1991). Dielectric boundary smoothing in finite difference solutions of the Poisson equation: an approach to improve accuracy and convergence. *J. Comput. Chem.* **12**, 909–912.

Davis, M. E., Madura, J. D., Luty, B. A. & McCammon, J. A. (1991). Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Comp. Phys. Commun.* **62**, 187–197.

Edmonds, D. T., Rogers, N. K. & Sternberg, M. J. E. (1984). Regular representation of irregular charge distributions: application to the electrostatic potentials of globular proteins. *Mol. Phys.* **52**, 1487–1494.

Elber, R. & Karplus, M. (1990). Enhanced sampling in molecular dynamics: use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. *J. Am. Chem. Soc.* **112**, 9161–9175.

Fan, P., Kominos, D., Kitchen, D. B., Levy, R. M. & Baum, J. (1991). Stabilization of α-helical secondary structure during high-temperature molecular-dynamics simulations of α-lactalbumin. *Chem. Phys.* **158**, 295–301.

Fersht, A. R. (1993). Protein folding and stability: the pathway of folding of barnase. *FEBS Letters*, **325**, 5–16.

Fersht, A. R., Matouschek, A. & Serrano, L. (1992a). The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* **224**, 771–782.

Fersht, A. R., Matouschek, A., Sancho, J., Serrano, L. & Vuilleumier, S. (1992b). Pathway of protein folding. *Faraday Discuss.* **93**, 183–193.

Freund, C., Ross, A., Plückthun, A. & Holak, T. A. (1994). Structural and dynamic properties of the F_v fragment and the single-chain F_v fragment of an antibody in solution investigated by heteronuclear three-dimensional NMR spectroscopy. *Biochemistry*, **33**, 3296–3303.

Frieden, C., Hoeltrli, S. D. & Ropson, I. J. (1993). NMR and protein folding: equilibrium and stopped-flow studies. *Protein Sci.* **2**, 2007–2014.

Ghelis, C. (1980). Transient conformational states in proteins followed by differential labeling. *Biophys. J.* **32**, 503–511.

Gilson, M. K. & Honig, B. H. (1986). The dielectric constant of a folded protein. *Biopolymers*, **25**, 2097–2119.

Gilson, M. K. & Honig, B. H. (1988). Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies and conformational analysis. *Proteins: Struct. Funct. Genet.* **4**, 7–18.

Gilson, M. K., Sharp, K. A. & Honig, B. H. (1988). Calculating the electrostatic potential of molecules in solution: method and error assessment. *J. Comput. Chem.* **9**, 327–335.

Griko, Y. U., Makhatadze, G. I., Privalov, P. L. & Hartley, R. W. (1994). Thermodynamics of barnase unfolding. *Protein Sci.* **3**, 669–681.

Harvey, S. C. (1989). Treatment of electrostatic effects in macromolecular modeling. *Proteins: Struct. Funct. Genet.* **5**, 78–92.

Hendsch, Z. S. & Tidor, B. (1994). Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* **3**, 211–226.

Horovitz, A., Serrano, L., Avron, B., Bycroft, M. & Fersht, A. R. (1990). Strength and cooperativity of contributions of surface salt bridges to protein stability. *J. Mol. Biol.* **216**, 1031–1044.

Jaenicke, R. (1991). Protein folding: local structures, domains, subunits, and assemblies. *Biochemistry*, **30**, 3147–3161.

Jones, C. M., Henry, E. R., Hu, Y., Chan, C., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J. & Eaton, W. A. (1993). Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl Acad. Sci. USA*, **90**, 11860–11864.

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

Kaptein, R., Dijkstra, K. & Nicolay, K. (1978). Laser photo-CIDNP as a surface probe for proteins in solution. *Nature*, **274**, 293–294.

Karplus, M. & Shakhnovich, E. (1992). Protein folding: theoretical studies of thermodynamics and dynamics. In *Protein Folding* (Creighton, T. E., ed.), pp. 127–195, W. H. Freeman, San Francisco.

Kellis, J. T., Jr, Nyberg, K. & Fersht, A. R. (1989). Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry*, **28**, 4914–4922.

Kraulis, P. (1991). MOLSCRIPT, a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.

Lazaridis, T., Archontis, G. & Karplus, M. (1995). The enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics. *Advan. Protein Chem.* In the press.

Lee, B. & Richards, F. M. (1971). The interpretation of protein structures; estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.

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

Lim, C., Bashford, D. & Karplus, M. (1991). Absolute pK_a calculations with continuum dielectric methods. *J. Phys. Chem.* **95**, 5610–5620.

Loewenthal, R., Sancho, J. & Fersht, A. R. (1992). Histidine-aromatic interactions in barnase. Elevation of histidine pK_a and contribution to protein stability. *J. Mol. Biol.* **224**, 759–770.

Loewenthal, R., Sancho, J., Reinikainen, T. & Fersht, A. R. (1993). Long-range surface charge-charge interactions in proteins: comparison of experimental results with calculation from a theoretical method. *J. Mol. Biol.* **232**, 574–583.

Loncharich, R. J. & Brooks, B. R. (1990). Temperature dependence of dynamics of hydrated myoglobin. *J. Mol. Biol.* **215**, 439–455.

Makhatadze, G. I. & Privalov, P. L. (1993). Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration. *J. Mol. Biol.* **232**, 343–350.

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.

Matouschek, A., Kellis, J. T., Jr, Serrano, L. & Fersht, A. R. (1989). Mapping the transition state and pathway of protein folding by protein engineering. *Nature*, **340**, 122–126.

Matouschek, A., Kellis, J. T., Jr, Serrano, L., Bycroft, M. & Fersht, A. R. (1990). Transient folding intermediates characterized by protein engineering. *Nature*, **346**, 400–445.

Matouschek, A., Serrano, L. & Fersht, A. R. (1992a). The folding of an enzyme. IV. Structure of an intermediate in the refolding of barnase analysed by a protein engineering procedure. *J. Mol. Biol.* **224**, 819–835.

Matouschek, A., Serrano, L., Meiering, E. M., Bycroft, M. & Fersht, A. R. (1992b). The folding of an enzyme. V. H^2H exchange-nuclear magnetic resonance studies on the folding pathway of barnase: complementarity to and agreement with protein engineering studies. *J. Mol. Biol.* **224**, 837–845.

Mauguen, Y., Hartley, R. W., Dodson, E. J., Dodson, G. G., Bricogne, G., Chothia, C. & Jack, A. (1982). Molecular structure of a new family of ribonuclease. *Nature*, **297**, 162–164.

Meiering, E. M., Serrano, L. & Fersht, A. R. (1992). Effect of active site residues in barnase on activity and stability. *J. Mol. Biol.* **225**, 585–589.

Mohan, V., Davis, M. E., McCammon, J. A. & Pettit, B. M. (1992). Continuum model calculations of solvation free energies: accurate evaluation of electrostatic contributions. *J. Phys. Chem.* **96**, 6428–6431.

Otting, G. & Wüthrich, K. (1989). Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules. *J. Am. Chem. Soc.* **111**, 1871–1875.

Pace, C. N., Laurents, D. V. & Erickson, R. E. (1992). Urea denaturation of barnase: pH dependence and characterization of the unfolded state. *Biochemistry*, **31**, 2728–2734.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986). *Numerical Recipes*, Cambridge University Press, Cambridge.

Ptitsyn, O. B. (1992). The molten globule state. In *Protein Folding* (Creighton, T. E., ed.), pp. 243–300, W. H. Freeman, San Francisco.

Reiher, W. E., III (1985). Theoretical studies of hydrogen bonding. PhD thesis, Harvard University.

Šali, D., Bycroft, M. & Fersht, A. R. (1988). Stabilization of protein structure by interaction of α-helix dipole with a charged side chain. *Nature*, **335**, 496–500.

Šali, D., Bycroft, M. & Fersht, A. R. (1991). Surface electrostatic interactions contribute little to stability of barnase. *J. Mol. Biol.* **220**, 779–788.

Šali, A., Shakhnovich, E. & Karplus, M. (1994). Kinetics of protein folding. *J. Mol. Biol.* **235**, 1614–1636.

Serrano, L., Kellis, J. T., Jr, Cann, P., Matouschek, A. & Fersht, A. R. (1992a). The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* **224**, 783–804.

Serrano, L., Matouschek, A. & Fersht, A. R. (1992b). The folding of an enzyme. III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure. *J. Mol. Biol.* **224**, 805–818.

Serrano, L., Matouschek, A. & Fersht, A. R. (1992c). The folding of an enzyme. VI. The folding pathway of barnase: comparison with theoretical models. *J. Mol. Biol.* **224**, 847–859.

Shakhnovich, E. I. & Finkelstein, A. V. (1989). Theory of cooperative transitions in protein molecules. I. Why denaturation of globular proteins is a first order phase transition. *Biopolymers*, **28**, 1667–1680.

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

Vijayakumar, S., Vishveshwara, S., Ravishanker, G. & Beveridge, D. L. (1993). Differential stability of β-sheets and α-helices in β-lactamase: a high temperature molecular dynamics study of unfolding intermediates. *Biophys. J.* **65**, 2304–2312.

Yu, H. & Karplus, M. (1988). A thermodynamic analysis of solvation. *J. Chem. Phys.* **89**, 2366–2379.

Yu, H.-A., Karplus, M. & Pettitt, B. M. (1991). Aqueous solvation of *N*-methyl acetamide conformers: comparison of simulations and integral equation theories. *J. Am. Chem. Soc.* **113**, 2425–2434.

Warwicker, J. & Watson, H. C. (1982). Calculation of the electric potential in the active site cleft due to α-helix dipoles. *J. Mol. Biol.* **157**, 671–679.

Edited by A. R. Fersht

(Received 30 June 1994; accepted in revised form 27 June 1995)