Molecular dynamics simulation of protein denaturation: Solvation of the hydrophobic cores and secondary structure of barnase

AMEDEO CAFLIŠCH AND MARTIN KARPLUS

Department of Chemistry, 12 Oxford Street, Harvard University, Cambridge, MA 02138

Contributed by Martin Karplus, September 27, 1993

ABSTRACT The transition in barnase from the native state to a compact globule has been studied with high-temperature molecular dynamics simulations. A partial destruction of the α-helices and the outer strands of the β-sheet is observed with water molecules replacing the hydrogen bonds of the secondary structural elements. Simultaneously, the main α-helix moves away from the β-sheet and exposes the principal hydrophobic core, many of whose nonpolar side chains, beginning with the ones near the surface, become solvated by hydrogen-bonded water molecules. This step involves a significant increase in the solvent-exposed surface area; the resulting loss of stability due to the hydrophobic effect may be the major source of the activation barrier in the unfolding reaction. The detailed mechanism described here for the first stage of the denaturation of barnase, including the essential role of water molecules, is likely to be representative of protein denaturation, in general.

The mechanism of protein folding is one of the major unsolved problems of biology. Although considerable progress has been made in experimental studies of the folding and unfolding transitions (1), our knowledge is limited by the difficulty of obtaining structural data. Simplified models (e.g., the use of reduced representations for the amino acids, effective interaction potentials, and discretized conformational space) are yielding useful insights (2). However, more detailed studies with atomic potentials are needed, particularly for determining the role of the solvent. In this paper, we describe the results of molecular dynamics simulations of the initial stages of unfolding of barnase (a 110-amino acid RNase from Bacillus amyloliquefaciens) at high temperature in the presence of water. This protein is a particularly good system for folding studies (3). The crystal (4) and solution (5) structures are known. Barnase consists of three α-helices and a five-stranded β-sheet that are stabilized by three hydrophobic cores in the native structure (Fig. 1), there are no disulfide linkages to constrain the unfolded state, and the three prolines are in a trans configuration. Transition states and pathways of barnase folding and unfolding have been investigated by protein engineering (6–8) and NMR hydrogen-exchange experiments (9). The rate-determining step for both folding and unfolding involves the crossing of a free-energy barrier near the native state (6, 8). The simulations yield information that can be compared with the experiment and provide a mechanism for solvent denaturation of the secondary structural elements and the hydrophobic cores.

METHODS

The barnase denaturation simulations used a deformable boundary potential (10, 11) and standard molecular dynamics (12) methodology. The system consisted of 1091 protein atoms and 3003 water molecules in a sphere of 30-Å radius.

Two denaturation simulations were performed at 600 K (A600, 120 ps and R600, 230 ps) and a 300 K control trajectory was run for 250 ps. The 600 K temperature was used to speed up the unfolding transition. An increase of ∼100 relative to the experimental denaturation temperature of 327 K (13) is expected since the activation energy for unfolding is 20 kcal/mol (1 cal = 4.184 J) (14). The deformable boundary leads to an increase of the pressure as the temperature is raised. This may speed up unfolding and increase water penetration. Two initial structures with different random velocities were studied to evaluate the effect of the initial conditions on the unfolding behavior. The A600 simulation was started after 4 ps of solvent equilibration from the minimized x-ray structure at 300 K, while the R600 run was initiated after 100 ps of simulation at 300 K. 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). After 150 ps of simulation of R600 some side-chain atoms of Lys-39, Lys-62, and Leu-63 reached the edge of the spherical shell. The simulation was stopped and the entire system (barnase and water) was centered in a larger sphere of water molecules (36-Å radius). Water molecules of the larger sphere overlapping any atom of the original system were removed; this yielded a total of 5647 water molecules. The atoms of barnase were then fixed during 200 steps of steepest descent minimization followed by 4 ps of Langevin dynamics of the water molecules at 300 K. The constraints were then removed for a 200-step minimization and a 1-ps heating to 600 K by conventional molecular dynamics; no discontinuity in the barnase behavior was observed. In all simulations, the temperature of the system was controlled by weak coupling to an external bath (15); a time step of 2 fs was employed.

RESULTS

The radius of gyration ($R_g$) and heavy-atom rms deviation (RMSD) from the x-ray structure as a function of time are given in Fig. 2. In A600 and R600, $R_g$ starts to increase after 30 ps, whereas the RMSD increases immediately. Both $R_g$ and RMSD then increase over most of the simulation. However, the increase is not uniform; e.g., in A600, $R_g$ is nearly constant for 20 ps between 45 and 65 ps; this may be indicative of an intermediate (16).

In the control simulation at 300 K, $R_g$ increases to 13.7 Å, relative to the x-ray value of 13.6 Å; the RMSD from the x-ray structure is 1.9 Å (the main-chain atom RMSD is 1.5 Å) during the last 50 ps. The overall conformation, hydrophobic core compactness, and secondary structural elements are stable. There is no water penetration into the protein; there are zero, two, and one molecules in hydrophobic cores 1, 2, and 3, respectively; experiments suggest that core 2 contains three water molecules (8).

In the A600 simulation (120 ps) and the first half (115 ps) of the R600 simulation, there are similar structural changes. The

Abbreviation: $R_g$, radius of gyration; RMSD, rms deviation.
The B300 simulation has a nearly constant $R_g$ equal to an average value of 14.8 Å; the volume increase from the minimized x-ray structure is 29%. The percentages of native interstrand hydrogen bonds at the edges of the sheet (strands 1–2 and 4–5) and at the center of the sheet (strands 2–3 and 3–4) stabilize at ≈50% and 70%, respectively. Interactions with water molecules replace the helical and β-sheet hydrogen bonds. Fersht and coworkers (7) have shown that there exists an intermediate on both the folding and unfolding pathway. It has some of the properties of B300 (e.g., unfolded N terminus, loops 1, 2, and 4; distorted secondary structure elements; solvated core 2; and weakened hydrophobic interactions in cores 1 and 3), but no $R_g$ measurement is available.

Solvation of the Main Hydrophobic Core. Core 1, which is an important stabilizing element of barnase (4, 8), is formed by the packing of helix 1 against the β-sheet and is centered around the side chain of Ile-88 (Fig. 1). Fig. 3 shows the time dependence in A600 of the solvent-accessible surface area of the side chains of core 1 and the number of water molecules in the core; similar behavior is seen during the first half of the R600 trajectory. Increase in accessible surface area and water penetration are nearly simultaneous and begin at ≈35 ps. Sixteen water molecules have penetrated at 82 ps; this falls to 10 between 89 and 98 ps and increases to 17 in the period from 111 to 120 ps. Many of the solvating waters make hydrogen bonds to waters outside the core. The accessibility of core 1 to water is coupled with the relative motion of helix 1 and the β-sheet (Fig. 4); i.e., they begin to move apart at ≈30 ps and their separation is continuous during the 30- to 80-ps period; between 80 and 100 ps, there is a small closing movement in accord with the decrease of water in the core, followed by expansion for the remainder of the simulation. During the B300 simulation, the number of water molecules in core 1 and the solvent-accessible surface area of its side chains are nearly constant (see Fig. 3); the average number of water molecules is 14. This steady state of solvation of the core is correlated with the nearly constant number of hydrogen bonds in helix 1 and the central part of the β-sheet (strands 2–3 and 3–4).

To illustrate the solvent role in the denaturation of the principal hydrophobic core in A600, some snapshots are shown in Fig. 5, although the specific details may be unique to this simulation. At 1 ps, helix 1 has a regular shape, the double salt link (Asp-8—Arg-110—Asp-12) is present, and there are no water molecules within the core (Fig. 3). A few water molecules are able to penetrate prior to major solvation of the core; e.g., at 9 ps (Fig. 5A), a solvent molecule is

![Figure 1](image1.png)

**Fig. 1.** Schematic diagram of the backbone of barnase, emphasizing the secondary structural elements; side chains of hydrophobic core 1 are plotted in a ball and stick representation. The structural elements include the following residues: N terminus, aa 1–5; helix 1, aa 6–18; loop 1, aa 19–25; helix 2, aa 26–34; loop 2, aa 35–40; helix 3, aa 41–45; type II β-turn, aa 46–49; strand 1, aa 50–55; loop 3, aa 56–69; strand 2, aa 70–75; loop 4, aa 76–84; strand 3, aa 85–90; type I β-turn, aa 91–94; strand 4, aa 95–100; type III β-turn, aa 101–104; strand 5, aa 105–108; C terminus, aa 109 and 110.

![Figure 2](image2.png)

**Fig. 2.** (A) $R_g$ as a function of simulation time averaged over 10-ps intervals. (B) RMSD from the x-ray structure as a function of simulation time averaged over 10-ps intervals. ◦, A600; +, B300; □, R600; ×, control run at 300 K.
FIG. 3. Solvent-accessible surface area and number of water molecules for core 1 as a function of time. Solid line, A600; dashed line, B300. For the exposed surface area, shown in the upper curve with the scale on the left, the Lee and Richards algorithm (CHARMM implementation) and a probe sphere of 1.4-Å radius were utilized. For the number of water molecules, shown in the lower curve with the scale on the right, those within 7 Å of the center of the core (the instantaneous center of geometry of the carbon atoms of the side chains of Phe-7, Val-10, Ala-11, Leu-14, Leu-20, Tyr-24, Ala-74, Ile-76, Ile-88, Tyr-90, Trp-94, Ile-96, and Ile-109) were included.

located between the hydrophobic side chains of the top part of core 1 and donates a hydrogen bond to the CO of residue His-18 in helix 1 and to another water molecule in a cluster. Also, a water molecule has penetrated from the bottom part of the core and is located between the Ile-109 and Arg-110 side chains. By 23 ps, a water molecule has inserted between the Asp-8 and Arg-110 side chains and a second water molecule has inserted between the Asp-12 and Arg-110 side chains initiating the rupture of the double salt link. At 39 ps, active penetration of the core has begun (Fig. 5B) and helix 1 and the β-sheet have started to move apart (see Fig. 4). The polar groups of the Tyr-24, Tyr-90, and Trp-94 side chains become engaged in hydrogen bonds with water molecules. Tyr-90 and Trp-94 side chains move toward the center of core 1 and a cluster of waters penetrates from the top. However, the stacking interaction of the Tyr-90 and Trp-94 is preserved and there are no water molecules between the two aromatic rings. Two water molecules in the same cluster donate hydrogen bonds to the CO groups of residues 14, 15, and 17 of helix 1. The side chain of Ile-109 has moved away from the aromatic ring of Phe-7 and the Asp-8, Arg-110, and Asp-12 side chains are solvated. (In R600 the double salt bridge

FIG. 4. Stereoview of the barnase Cα atoms to illustrate the relative helix/β-sheet motion during A600. Thin line and labels, 1 ps; medium line, 70 ps; thicker line, 90 ps; thick lines, 120 ps.

FIG. 5. Stereoviews of A600 dynamics. (A) 9 ps. (B) 39 ps. (C) 69 ps. (D) 115 ps. Hydrophobic side chains of core 1 are shown as thick lines; α-helix backbone (residues 7–18), Asp-8, Asp-12, Lys-98, and Arg-110 side chains are shown as thin lines. Water molecules within 12 Å of the center of geometry of core 1 are included. Hydrogen bonds are dotted (acceptor–hydrogen distance smaller than 2.5 Å; no angular criterion). Waters discussed in text are shown with thick lines.

Asp-8—Arg-110—Asp-12 is broken after ≈100 ps.) Also, a cluster of 8–10 hydrogen-bonded water molecules is located at Phe-7, Ala-11, Ile-88, Ile-96, and Ile-109; some of these waters accept hydrogen bonds from the charged groups of the Lys-98 and Arg-110 side chains. At 46 and 51 ps, some water molecules linked to Lys-98 move into the center of the core; the inward motion of Lys-98 is coupled to the outward motion
of the Arg-110 side chain. The exterior strands of the β-sheet are being solvated by water molecules that replace some of the hydrogen bonds between strands 1 and 2 and between strands 4 and 5 (Fig. 6). At 69 ps (Fig. 5C), pentagonal arrays (17) of water molecules cover the Leu-20 side chain and separate it from the remainder of the core. The OH groups of Tyr-24 and Tyr-90 participate in the pentagonal structures. At 88 ps, some waters escape from the bottom part of core in conjunction with the outward motion of Lys-98 and the simultaneous inward motion of Phe-7 and Ile-109. Also, helix 1 moves closer to the β-sheet and the exposed surface area decreases (Figs. 3 and 4). Analysis of the time series of the solvent-accessible surface area of each hydrophobic side chain in the core reveals that during the 85- to 105-ps period of A600, the central part of core 1 (Leu-14, Ala-74, Ile-88, and Ile-96) is compact and does not contain any water molecules, whereas the Leu-20, Tyr-24, Tyr-90, Trp-94, and Ile-109 side chains are exposed to solvent. In the 85- to 105-ps structures a layer of water molecules separates the hydrophobic side chains on the core surface from the central part of the core. Similar behavior is seen during the 70- to 105-ps period of R600. The larger stability of the residues in the core center is in agreement with protein engineering results (7). The system expands again over the subsequent 10 ps and at 115 ps the core is solvated (Fig. 5D) with 21 water molecules within 7 Å of the center (Fig. 3).

Denaturation of Secondary Structural Elements. The β-sheet in barnase is composed of five antiparallel β-strands (Fig. 1). In the native structure, it has a regular shape, apart from a β-bulge in strand 1 between residues 53 and 54 (Fig. 6). There is a twist (defined as the angle formed by the N terminus to C terminus vector of strands 1 and 5) of ~90° from strand 1 to strand 5. The main α-helix of barnase (helix 1) has three turns, helix 2 has two turns, and helix 3 has one turn.

The initial phase of solvation of core 1 is coupled with the partial distortion of helix 1 and the edge strands of the β-sheet during both the A600 and R600 runs. With the steep increase in solvent-accessible surface area of core 1 (40- to 70-ps period in both A600 and R600; see Fig. 3), there occurs an abrupt decrease in the number of native 1-4 hydrogen bonds in helix 1 and a pronounced distortion of the native interstrand hydrogen bonds between strands 1 and 2 and between strands 4 and 5 of the β-sheet (60-70 ps in A600 and 40-70 ps in R600; Fig. 6). Because a number of studies have been made of isolated helix denaturation (18-20) and partial unfolding in helical proteins (21, 22), we concentrate here on the β-sheet results. Solvent insertion begins at the edges of the β-sheet (Fig. 6). The C-terminal part of strand 1 (near the β-bulge) is rapidly solvated in both A600 and R600; this is likely to originate from the instability of this irregular part of secondary structure. At 60 ps, several solvent molecules are located between strands and participate in hydrogen bonds with the main-chain NH and CO groups as donors, acceptors, or both. In R600, the hydrogen bonds between strands 1 and 2 are broken at 60-70 ps (in A600 the strands separate after ~100 ps). This is followed by separation (with water insertion) of strands 3 and 4 at 120 ps and strands 2 and 3 at 130 ps; strands 4 and 5, although distorted, maintain between one and two interstrand hydrogen bonds. The twist in the β-sheet increases during denaturation; this is coupled to the distortion of the strands and the insertion of water molecules interacting with the polar groups of the main chain. In R600, the twist increases to ~120° after 70 ps and reaches a value of almost 180° at 150 ps. At this point, several waters have inserted between strands 1 and 2, strands 2 and 3, and strands 4 and 5, while strands 3 and 4 interact through weak polar interactions between nearby NH and CO groups (Fig. 6 Bottom). A similar behavior is seen in A600 and in an in vacuo simulation. This suggests that β-sheet denaturation at high temperatures is initiated by the twisting motion that promotes the loss of interstrand hydrogen bonds.

Helix 2 and helix 3, which are small and accessible to solvent, unfold during the first 70 ps in both A600 and R600. The main helix (helix 1) starts to fray at the termini; the C-terminal hydrogen bond (18-NH ••• CO-14) is lost after ~30 ps and does not reform. The three N-terminal hydrogen bonds are broken after ~50 ps; they reform in the later stages of the simulation. Reformation of the N-terminal hydrogen bond in helix 1 (at 90 ps and 110 ps) occurs through a 3p hydrogen bond; this has been observed in simulations of oligopeptides (18-20) and is in accord with x-ray studies (23). The helical hydrogen bonds between 15-NH ••• CO-11 and 16-NH ••• CO-12 are broken after ~70 ps; they convert into a 3p form in the last 20 ps of A600. During the last 120 ps of R600, helix 1 becomes fully solvated by water molecules and undergoes a rigid body rotation of ~120° about its axis so that its hydrophobic side chains are exposed to solvent. At 220 ps, only the 10-CO ••• NH-14, 11-CO ••• NH-15, and 12-CO ••• NH-16 helical hydrogen bonds are still present.

CONCLUDING DISCUSSION

Simulations of the initial stages of barnase unfolding by high-temperature molecular dynamics have demonstrated the cooperative nature of the process and the essential role of water molecules. Two simulations at 600 K started from
different initial conditions gave similar results. The N terminus and the loops unfolded first. This was coupled to a small distortion of regular secondary structural elements and partial unstacking of the hydrophobic cores. Water molecules penetrated core 2 and the edge residues of core 1. Core 2 solvation promoted the destabilization of helix 2 and helix 3, while partial solvation of core 1 initiated the disruption of helix 1 and the detachment of the edge strands from the $\beta$-sheet. The cooperative nature of the denaturation process was evident in the solvation and rupture of the principal hydrophobic core (core 1), which is coupled to the relative motion of helix 1 and the $\beta$-sheet. The resulting increase in solvent-exposed surface area of hydrophobic groups may give rise to the activation barrier for the unfolding reaction (3), though it is possible that it occurs earlier when $R_2^*$ begins to increase. Use of a relation between the hydrophobic free energy and the solvent-exposed surface area (25 kcal/mol$^2$), in spite of questions about the validity of such a simple formula (24), yields a destabilization of 33 kcal/mol between the native state and that reached after 70 ps in A600 (see Fig. 3). This value can be compared with the measured activation energy of 20 kcal/mol (14), though corrections for electrostatic contributions and an increase in conformational entropy need to be considered.

A number of detailed results of the present analysis are of interest because they may play a role in protein denaturation, in general. The polar OH and NH groups of tyrosines and tryptophan side chains, respectively, are important in the penetration of water (from the top part of core 1), while the motion of a lysine side chain helps the water molecules reaching the center of core 1 from the bottom part. Clusters of hydrogen-bonded water molecules surrounding hydrophobic side chains participate in hydrogen bonds with polar groups of the backbone and/or of the side chains. The $\beta$-sheet disruption starts near the irregular element ($\beta$-bulge at residues 53 and 54) and at the edges (strands 1 and 5); it is promoted by an increase in the twist and an influx of water molecules, some of which insert between adjacent strands and participate in hydrogen bonds as both donors and acceptors with the main-chain polar groups. Water molecules act mainly as hydrogen-bonding donors in the initial phase of solvation of the main helix; sometimes they insert and replace the helical hydrogen bond, as previously found in the fragment studies (20).

Only qualitative comparisons can be made with the experiments of Fersht and coworkers (6–9), because they used mutation studies to deduce the properties of certain states (native, transition, intermediate, and unfolded states) while the present simulations follow the time course of the unfolding. Since it is not clear how the time dependence and the position along the reaction coordinate are related, we restrict the comparison to what is observed to happen early or late in the simulations and/or experiments. The overall behavior during the simulations (early unwinding of the N terminus, loop 1, loop 2, and loop 4; partial conservation of the secondary structure; solvation of core 2; and weakening of the hydrophobic interactions of the core 1 edge residues) is consistent with the experimental results (6–9). Moreover, solvation of the $\beta$-sheet began rather than actually at the N terminus as found experimentally (7), but the C-terminal unfolding was a transient phenomenon in R600. Residual helicity at residues 11–15 is present at the end of the R600 simulation, although the protein is almost fully denatured. This is consistent with mutagenesis results for barnase (7) and NMR measurements of its N-terminal fragment of residues 1–36 (25). An intermediate found in the simulations (and stabilized at 300 K) has characteristics corresponding to those of the refolding intermediate studied by protein engineering and NMR (7, 8).

The present results suggest a possible mechanism for the solvation of hydrophobic cores and for the dissolution of secondary structural elements in protein denaturation. Testing of this mechanism is a challenge for experimentalists. Photochemically induced nuclear polarization (26), nuclear Overhauser effects of water interactions with specific residues (27), chemical markers (28), and NMR of 19F substituted amino acids (29) are possible approaches.

We thank L. Caves, A. Fersht, E. Shakhnovich, and S. Wodak for helpful discussions and P. Kraulis for use of the program MOLSCRIPT. The work was supported in part by the National Institutes of Health and a gift from Molecular Simulations, Inc. A.C. was supported by a grant from the Schweizerischer Nationalfonds. The coordinates of the protein were provided by Dr. A. Cameron and Prof. G. Dodson. The calculations were performed on an IBM RISC/6000 550, a SGI 340 GTX, a DEC Alpha 3000, and on the CRAY C90 at the Pittsburgh Supercomputing Center.