Dynamics and unfolding pathways of a hyperthermophilic and a mesophilic rubredoxin

THEMIS LAZARIDIS,¹ IRWIN LEE,¹ AND MARTIN KARPLUS^{1,2}

¹Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

²Laboratoire de Chimie Biophysique, Institut le Bel, Université Louis Pasteur, 67000 Strasbourg, France

(RECEIVED May 7, 1997; ACCEPTED September 18, 1997)

Abstract

Molecular dynamics simulations in solution are performed for a rubredoxin from the hyperthermophilic archaeon *Pyrococcus furiosus* (RdPf) and one from the mesophilic organism *Desulfovibrio vulgaris* (RdDv). The two proteins are simulated at four temperatures: 300 K, 373 K, 473 K (two sets), and 500 K; the various simulations extended from 200 ps to 1,020 ps. At room temperature, the two proteins are stable, remain close to the crystal structure, and exhibit similar dynamic behavior; the RMS residue fluctuations are slightly smaller in the hyperthermophilic protein. An analysis of the average energy contributions in the two proteins is made; the results suggest that the intraprotein energy stabilizes RdPf relative to RdDv. At 373 K, the mesophilic protein unfolds rapidly (it begins to unfold at 300 ps), whereas the hyperthermophilic does not unfold over the simulation of 600 ps. This is in accord with the expected stability of the two proteins. At 473 K, where both proteins are expected to be unstable, unfolding behavior is observed within 200 ps and the mesophilic protein unfolds faster than the hyperthermophilic one. At 500 K, both proteins unfold; the hyperthermophilic protein does so faster than the mesophilic protein. The unfolding behavior for the two proteins is found to be very similar. Although the exact order of events differs from one trajectory to another, both proteins unfold first by opening of the loop region to expose the hydrophobic core. This is followed by unzipping of the β -sheet. The results obtained in the simulation are discussed in terms of the factors involved in flexibility and thermostability.

Keywords: flexibility; hyperthermostability; molecular dynamics; rubredoxin; unfolding pathways

Most living organisms have optimal temperatures for growth near room temperature and are referred to as “mesophilic.” Proteins from such organisms tend to have optimal function near 40 °C and denature at somewhat higher temperatures. Of course, there are proteins in mesophilic organisms that have higher thermal stability, such as the bovine pancreatic trypsin inhibitor, that denatures near 100 °C (Makhatadze et al., 1993). Other microorganisms (“thermophilic”) grow optimally at higher temperatures (50–60 °C). Since the early 80s (Stetter, 1982), the temperature range over which life is sustainable has been extended by the discovery of “hyperthermophilic” microorganisms that can thrive at temperatures as high as 100 °C (Adams & Kelly, 1992, 1995; Rees & Adams, 1995). Sometimes the term “extreme thermophile” is used for those microorganisms (like the *Thermus* species) that have optimal growth temperatures between 70 and 80 °C. Many species have adapted to very low temperatures (<20 °C); these are referred to as “psychrophilic.” In addition to temperature adaptation, there are microorganisms that thrive under extreme conditions of pH or ionic strength (acidophilic, alkaliphilic, halophilic) (Jaenicke, 1981; Jaenicke & Zavodszky, 1990).

Research on proteins derived from such “extremophilic” organisms has intensified in recent years due to both their practical and fundamental significance. On the practical side, certain enzymes with high thermostability, either natural or engineered, have important technological applications, including detergent manufacturing, production of high-fructose corn syrup, and PCR (Adams & Kelly, 1995). On the fundamental side, comparison of extremophilic proteins with their normal counterparts can help to increase our understanding of the physical basis of protein stability. The activity of homologous enzymes from mesophiles and thermophiles appears to be similar at the optimal conditions for growth of each organism, whereas the thermophilic enzyme may have little activity under mesophilic conditions (Wrba et al., 1990; Volkl et al., 1994). This has been ascribed to the greater rigidity of the thermophilic enzyme (Zuber, 1988; Wrba et al., 1990), although the evidence is not compelling. Comparisons of homologous thermophilic and mesophilic enzymes can aid in isolating the factors contributing to the efficiency of enzyme catalysis.

The three-dimensional structures of proteins from moderately thermophilic bacteria, such as *Bacillus stearothermophilus*, have been known for a long time [for example, thermolysin (Matthews et al., 1972) or glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Biesecker et al., 1977)]. More recently, the structures of proteins from hyperthermophiles have become available, e.g., rubredoxin (Day

Reprint requests to: Martin Karplus, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge Massachusetts 02138; e-mail: marci@tammy.harvard.edu.

et al., 1992), GAPDH (Korndorfer et al., 1995), aldehyde ferredoxin oxidoreductase (Chan et al., 1995), histone HMFb (Starich et al., 1996), glutamate dehydrogenase (Yip et al., 1995), and DNA binding proteins Sso7d (Baumann et al., 1994) and Sac7d (Edmondson et al., 1995). The structures of homologous proteins from thermophilic and hyperthermophilic sources are generally very similar to their mesophilic and psychrophilic counterparts. For example, the RMS difference between rubredoxin from the hyperthermophile *Pyrococcus furiosus* and the mesophile *Clostridium pasteurianum* is 0.47 Å for the main chain (Day et al., 1992) and that between GAPDH from the hyperthermophile *Thermotoga maritima* and the thermophile *B. stearothermophilus* is 0.57 and 0.83 Å for each of the two domains (Korndorfer et al., 1995). This is not surprising because the sequence identity is usually in the range 45–65% and there are few significant additions or deletions. From the work of Chothia and Lesk (1986), a sequence identity in the neighborhood of 50% is expected to yield an RMS difference of less than 1 Å in the core region of a pair of homologous proteins.

The availability of homologues with high sequence identity and structural similarity but greatly different thermal stability, as well as results from the effect of site-directed mutagenesis on stability (Alber, 1989; Serrano et al., 1992; Matthews, 1993), have led to various suggestions concerning the interactions contributing to protein stability [for a recent analysis, see Vogt and Argos (1997)]. These include the overall amino acid composition (Argos et al., 1979; Menendez-Arias & Argos, 1989), site-specific amino acid replacement, such as proline in turns (Matthews et al., 1987; Watanabe et al., 1991), salt bridges (Perutz & Raidt, 1975), hydrogen bonding (Day et al., 1992; Tanner et al., 1996), hydrophobic interactions and packing (Yamada et al., 1990; Chan et al., 1995; Spassov et al., 1995), electrostatic interactions (Spassov et al., 1995), increased helix stability (Davies et al., 1993; Russell et al., 1994; Hennig et al., 1995), or disulfide bonds and metals (Frommel & Sander, 1989). The range of factors that have been cited suggests that incremental contributions that vary from one protein to another are involved. This would not be surprising because the individual interactions that contribute to protein stability are weak (Lazaridis et al., 1995) and the overall free energy change between the denatured and native protein is only on the order of 0.1 kcal/mol per residue. Further, an increase in the denaturation temperature from 50 to 100°C involves a change in the free energy of unfolding of only a few kcal/mol (Rees & Adams, 1995).

Molecular dynamics simulations have provided detailed information on the dynamics of native proteins (McCammon & Harvey, 1987; Brooks et al., 1988). Recently, high-temperature simulations in solution have been used to study protein unfolding (Mark & van Gunsteren, 1992; Daggett & Levitt, 1993; Caflisch & Karplus, 1994a, 1994b, 1995; Li & Daggett, 1994). Although such simulations at different temperatures do not address thermodynamic stability directly, unless some free energy evaluation technique is employed (Kollman, 1996), they can provide information on native state fluctuations and the unfolding behavior.

It is our purpose in this paper to compare simulation results for a mesophilic protein with those for a homologous hyperthermophilic protein. The systems chosen for study are two rubredoxins, small proteins consisting of 45–54 residues with a redox center, but unknown function. The crystal and solution structures of rubredoxin from the hyperthermophilic archaeon *P. furiosus* (Pf) has been determined recently (Blake et al., 1992a; Day et al., 1992) and several crystal structures from mesophilic sources are known to high resolution (Sieker et al., 1994): *C. pasteurianum* (PDB code

5RXN) (Watenpaugh et al., 1979, 1980), *Desulfovibrio desulfuricans* (PDB code 6RXN) (Stenkamp et al., 1990), *Desulfovibrio vulgaris* (Dv) (PDB codes 7RXN and 8RXN) (Adman et al., 1991; Dauter et al., 1992), and *Desulfovibrio gigas* (PDB code 1RDG) (Frey et al., 1987). Here we examine the dynamic behavior of the hyperthermophilic rubredoxin from Pf and one of its mesophilic homologues, the rubredoxin from Dv, by molecular dynamics simulations in solution. We refer to the former as RdPf and to the latter as RdDv. Although the stability of RdDv has not been measured, it is expected to be significantly lower than that of RdPf. Relative to the mesophilic proteins, the rubredoxin from Pf has more surface salt bridges and a more extensive β -sheet due to a substitution of a Pro by Glu (Blake et al., 1992a; Day et al., 1992). This substitution also gives rise to electrostatic interactions involving the Glu side chain. Calorimetry showed that the irreversible denaturation temperature for this protein is greater than 100°C (Klump et al., 1994). It was also found that it retains its spectroscopic properties (visible absorption) after incubation at 95°C for 24 h (Blake et al., 1991). Stability studies of the other rubredoxins are limited. 5RXN denatures rapidly at 80°C (Lovenberg & Sobel, 1965) and 1RDG loses 25% of its visible absorption after 1 h at 80°C (Papavassiliou & Hatchikian, 1985). A study of pH and ionic strength effects (Cavagnero et al., 1995) showed that the stability of the hyperthermophilic rubredoxin decreases as the pH decreases, but, even at pH 2, the protein is quite stable (it denatures at 70°C at pH 2). This means that the additional salt bridges could make some but not the dominant contribution to the stability of this protein. The rubredoxin from *C. pasteurianum* has been cloned and some site-directed mutagenesis results have been reported (Richie et al., 1996; Eidsness et al., 1997). The results indicated that the extension of the β -sheet does not contribute significantly to the stability of the hyperthermophilic protein. This is in agreement with earlier results where extension of an α -helix by substitution of Pro by other amino acids did not alter the stability of T4 lysozyme (Alber et al., 1988). The study of chimeras of Pf and *C. pasteurianum* rubredoxins suggested that the interactions between the β -sheet and the loop residues are more important (Eidsness et al., 1997).

The hyperthermophilic rubredoxin from Pf has been a subject of some theoretical studies. Wampler et al. (1993) (see also Bradley et al., 1993) built homology models for it based on the structures of three mesophilic rubredoxins and performed molecular dynamics simulations in vacuum with the AMBER program and a distance-dependent dielectric. They noticed that the molecular mechanics energy was always lower than that of the mesophilic rubredoxins, but this is probably due to the lack of solvent in the simulations, i.e., the hyperthermophilic rubredoxin has a larger number of salt bridges that have large interaction energies in vacuum. From the simulations, it was concluded that the mesophilic proteins sample a more complex conformational space with two or more minima, but it is not clear how this is related to thermal stability. Ichiye and coworkers have studied the redox properties of a mesophilic rubredoxin (5RXN) and of RdPf by energy minimization and molecular dynamics (Shenoy & Ichiye, 1993; Yelle et al., 1995; Swartz & Ichiye, 1996), but did not concern themselves with stability. Jung et al. (1997) performed 30-ps molecular dynamics simulations in solution for Pf rubredoxin, five of its mutants, and three mesophilic rubredoxins at 25 and 100°C. At room temperature, the calculated atomic fluctuations over a period of 10 ps were similar for all proteins, whereas, at 100°C, they were smaller for Pf rubredoxin. The length of these simulations is too short for definitive conclusions.

Simulations for the two proteins, RdPf and RdDv, in explicit solvent at a number of temperatures are reported in this paper: one at 300 K, one at 373 K, two at 473 K, and one at 500 K (see Table 1). The first objective of the simulations is to evaluate the flexibility of the two proteins at room temperature by comparing the atomic fluctuations obtained in the simulations. In addition, we compare the unfolding behavior at a series of temperatures; the lowest (373 K) is such that the mesophilic protein should be unstable, whereas the hyperthermophilic protein should be stable. To speed up unfolding, we do additional simulations at 473 K and 500 K, temperatures in the neighborhood of those used in other unfolding simulations (e.g., Li & Daggett, 1994). Such unfolding studies are of particular interest because it has been suggested recently that RdPf is "kinetically trapped," i.e., that it unfolds more slowly than its mesophilic analogues (Cavagnero et al., 1997).

The aligned amino acid sequences of RdPf and RdDv are shown in Figure 1 and a MOLSCRIPT representation of their structure is shown in Figure 2A. The numbering of residues is based on RdDv because RdPf is the only known rubredoxin that lacks the N-terminal methionine. Residues that differ in the two proteins are shown in bold in Figure 1. Among the mesophilic rubredoxins of known structure, RdDv has the largest number of residues (34) that are identical with RdPf; there is 65% identity and there are no insertions or deletions. The backbone RMS deviation (RMSD) of residues 2–52 between the structures for the two proteins is 0.58 Å. Rubredoxin has relatively little regular secondary structure. There is an antiparallel β -sheet comprising three strands: β 1 at the N terminus (residues 2–7), β 2 immediately following it (residues 10–15), and β 3 at the C terminus (residues 48–52). The β -sheet is interrupted by a β -bulge between residues 10 and 11. The iron-sulfur cluster is located at one end of the β -sheet; Cys residues 6, 9, 39, and 42 are coordinated with the iron. The rest of the protein is comprised of two loops; they are L1 (residues 16–29) and L2 (residues 30–47), and they contain a number of turns and short 3_{10} -helical segments. The hydrophobic core is comprised of residues from the β -sheet (residues 4, 11, 13, 49), loop 1 (residue 24), and loop 2 (residues 30, 33, 37, 46) (see Fig. 3). Residue 4 (Tyr in RdDv and Trp in RdPf) interacts with Phe 30. The hydroxyl groups of Tyr 4 and Tyr 11 are surface exposed, whereas that of Tyr 13 hydrogen bonds to the backbone at residue 28. Lys 46 lies flat on the surface of the protein, and its nonpolar part interacts with

residues 37, 30, and 49. Leu 33 interacts with residues 24, 13, and 37, and Phe 30 with residues 13, 37, 49, and 4.

A significant difference between RdPf and its mesophilic analogues is close to the N terminus, where Pro 15 is replaced by Glu. This allows extension of the β -sheet and introduction of a number of electrostatic interactions involving the Glu 15 side chain (Blake et al., 1992a; Day et al., 1992). The major change in the hydrophobic core is the replacement of Tyr 4 in RdDv by Trp in RdPf. The substitution of Val 24 by Ile in RdPf versus RdDv is also seen in some mesophilic rubredoxins. An additional salt bridge in RdPf is observed between Lys 7 and Glu 50. All of these differences could confer higher thermal stability to RdPf versus RdDv, but there is no evidence concerning the contributions of any one of them. Clearly, site-directed mutagenesis would be of interest to determine the effects of individual substitutions.

Results

300 K Simulations

Structural aspects

The RMSD from the crystal structure for RdPf and RdDv during the 400-ps simulations at 300 K is shown in Figure 3A. It appears to have stabilized by the end of the simulations, where the RMSD is about 1.5 Å (backbone) and 2 Å (all atoms) for both proteins. These deviations are typical of those found in solution simulations of this and other proteins. For example, a 60-ps simulation of rubredoxin from *C. pasteurianum* gave an RMSD of 1.06 Å or 1.42 Å for the backbone, depending on whether the reduced or oxidized protein was simulated (Yelle et al., 1995). For barnase, the backbone RMSD was 1.5 Å after 250 ps (Caflisch & Karplus, 1994a). The deviations are smaller for the β -strands and larger in the loop region. At the end of the simulations, the backbone RMSD of the β -sheet (residues 1–7, 10–15, 48–51) is 1.06 Å for RdDv and 0.83 Å for RdPf, whereas that of the loop region is 1.54 Å and 1.75 Å, respectively. The radii of gyration (Rg, Fig. 3B) remain close to those in the crystal structures (9.67 Å for RdDv and 9.86 Å for RdPf). There is a small increase in the period 200–300 ps for RdPf and in the period 150–300 ps for RdDv, but then Rg decreases again. The backbone RMSD of residues 2–52 between the

Table 1. Summary of simulations

	Starting structure	Heating	Equilibration	Production
RdPf 300 K	Crystal	50 → 300 K in 2 ps	8 ps	400 ps
RdDv 300 K	Crystal	50 → 300 K in 2 ps	8 ps	400 ps
RdPf 373 K	15 ps of 300 K sim	300 → 373 in 2 ps	8 ps	600 ps
RdDv 373 K	15 ps of 300 K sim	300 → 373 in 2 ps	8 ps	1,020 ps
RdPf 473 K-set 1	100 ps of 373 K	373 → 473 in 3 ps	12 ps	200 ps
RdDv 473 K-set 1	120 ps of 373 K	373 → 473 in 3 ps	12 ps	200 ps
RdPf 473 K-set 2	40 ps of 373 K	373 → 473 in 3 ps	12 ps	260 ps
RdDv 473 K-set 2	40 ps of 373 K	373 → 473 in 3 ps	12 ps	340 ps ^a
RdPf 500 K	Crystal	50 → 500 K in 10 ps	20 ps	260 ps
RdDv 500 K	Crystal	50 → 500 K in 10 ps	20 ps	260 ps

^a At about 180 ps, RdDv collided with the boundary. At that point, the protein was resolvated in a larger sphere (25 Å radius). The sphere had density 0.0322 Å⁻³ and was added five times, at different orientations, with overlapping waters deleted. Final system size was 6,025 atoms (1,855 water molecules). Buffer region for the stochastic boundary simulation was set to 23–25 Å. Water was thermalized for 2 ps with the protein constrained and the production run was continued from that point.

	1		11		21		31		41		51
RdDv	M KKYVCTVCG	YEYD P AEGDP	DNGV K PGTSF	D DLPADWVCP	VCGAPKSEFE	A A					
RdPf	A KWVCKICG	YIYDE D AGDP	DNGI S PGTKF	E ELPDDWVCP	ICGAPKSEFE	K LED					
	*	*	*	*	*	*	*	*	*	*	*

Fig. 1. Aligned sequences of RdPf and RdDv (single-letter code). Bold letters denote the residues that differ in the two proteins. Star denotes the residues that are part of the hydrophobic core.

two proteins at 400 ps is 0.6 Å, similar to the difference between the crystal structures (0.58 Å).

Comparing the structures at the end of the 400-ps simulations with the crystal structures, we observe a small decrease in the

number of hydrogen bonds (see the Materials and methods for definitions). The number of hydrogen bonds in the crystal structures of RdPf and RdDv is 35 and 31, respectively, whereas this number at 400 ps is 26 and 27, respectively. A single coordinate set was used for comparing the hydrogen bonds because nearly all of the hydrogen bonds are either present or absent during the last 10 ps in these structures (see Fig. 4). Table 2 lists the hydrogen bonds in the crystal structure in the two proteins, showing which ones have disappeared by 400 ps, and the new hydrogen bonds that are formed in the molecular dynamics simulation. Many of the hydrogen bonds are included in the list of native contacts monitored throughout the simulation and are discussed further below. Because the criteria for the presence of a hydrogen bond are more strict than the criterion for the presence of a native contact (see the Materials and methods), some hydrogen bonds may be listed as broken, whereas the corresponding contact is listed as present. One example is the Glu 15–Ala 2 backbone hydrogen bond, which is

Fig. 2. Structure of rubredoxin. Coordinates of RdDv are used; RdPf has an identical fold. **A:** Overall structure. Positions of the iron atom (Fe) and of a few residue numbers are indicated. **B:** Hydrophobic core side chains. This view is rotated by 90° around the vertical axis from that in Figure 2A. Figures were drawn with the program MOLSCRIPT (Kraulis, 1991).

Fig. 3. RMSD from the crystal structure (upper panel) and radius of gyration (lower panel) during the control simulation at 300 K. Upper pair of curves is the all-atom RMSD and lower pair of curves is the backbone RMSD. Dark lines correspond to RdPf and light lines to RdDv.

Fig. 4. Contact plots for the control simulation at 300 K. A dot means that a particular contact is present at that time. For these simulations, a contact is assumed to be present if the interatomic distance is less than 1.3 times that in the crystal structure. Contacts are defined in Table 4. **A:** RdDv. **B:** RdPf.

unique to RdPf. Many of the broken hydrogen bonds in Table 2A and B involve ionic side chains. In some cases, the broken hydrogen bonds are replaced by others. For example, the Lys 7–Glu 50 salt bridge in RdPf is replaced by Lys 7–Glu 48. It is interesting that a few 3_{10} -helical hydrogen bonds are replaced by α -helical hydrogen bonds. For example, the Leu 33–Phe 30 hydrogen bond in both proteins is replaced by a hydrogen bond between residues 33 and 29. The Tyr 13–Thr 28 hydrogen bond was not present in the crystal structure of RdPf because the Tyr OH hydrogen was not built in the direction of residue 28. This hydrogen bond forms during the simulation. A number of the protein–protein hydrogen bonds are replaced by protein–water hydrogen bonds in the simulation, e.g., for both proteins, the NH group of Gly 10 and the CO group of Cys 6, which hydrogen bond to each other in the crystal structures, turn toward the solvent and hydrogen bond with water molecules. There is no significant difference between RdPf and RdDv with regard to hydrogen bonding behavior.

Figure 4A and B show the evolution of the selected native contacts during the 300 K simulations. In these plots, a dot is

printed when a contact exists at a particular time. As explained in Materials and methods, the cutoff in these two figures is set to 1.3 (rather than the usual 1.5) to reveal small fluctuations in the distances. The behavior of RdPf and RdDv is similar in that most contacts exist throughout the simulations and those that fluctuate are for the most part the same for both structures. Contact #5 (a β 1– β 2 hydrogen bond near the β bulge) disappears in RdPf and flickers in RdDv; it is absent during the last 10 ps in both simulations (see discussion of Table 2 above). Contact #9, a β 1– β 3 hydrogen bond, flickers in both proteins. Contact #14 is a salt bridge in RdPf and a Thr–Glu contact in RdDv. This contact flickers in both proteins. The fact that the salt bridge is not stable (no more stable than the Thr–Glu contact) suggests that it does not contribute significantly to the stability of the molecule. This is in accord with the general analysis of salt bridges by Hendsch and Tidor (1994). Contact #25, an L1–L2 hydrophobic contact, flickers more in RdPf than in RdDv. Contact #26 breaks in both proteins; it is the Asp 19–Trp 37 indole hydrogen bond near the rim of the hydrophobic core. Contact #41 also involves Asp 19: it is a hydrogen bond with Asn 22, at the edge of the hydrophobic core (Fig. 3), and it breaks in both proteins, although it reappears in the RdPf simulation. The instability of contacts #25, 26, and 41 indicates a slight opening of the hydrophobic core at its rim, where the two loops meet. In RdPf, a crystal water molecule inserts itself and interacts simultaneously with Trp 37, Asp 19, and Asn 22. In RdDv, Asp 19 moves too far away for a water-mediated interaction with Asn 22 and Trp 37. Instead, a hydrogen bond is formed between Asn 22 and Trp 37 in the latter part of the trajectory. In RdDv contact #2, a β 1– β 2 hydrogen bond in RdPf and a Pro–Lys N tertiary contact in RdDv, flickers substantially. Contacts #48 and 49, which are hydrogen bonds of the Lys 46 side chain to the backbone, disappear in RdDv. This Lys side chain has a different conformation in the crystal structures of the two proteins and this is mirrored in the simulations. In RdPf, as well as in the other three crystal structures of mesophilic rubredoxins, the Lys side chain is extended and hydrogen bonds to the carbonyl of residues 30 and 33. In RdDv, it is bent and hydrogen bonds to the carbonyl of residues 35 and 37 and a sulfate ion, which has been deleted in this study. During the RdDv simulation, the Lys 46 side chain moves away from 35 and 37 and toward residues 30 and 33. This brings it closer to the conformation it has in RdPf and the other crystal structures.

Dynamic aspects

It is frequently stated that thermophilic proteins at room temperature are more rigid than their mesophilic counterparts (Zuber, 1988; Jaenicke & Zavodszky, 1990). Evidence for this greater rigidity is obtained from hydrogen exchange (Wrba et al., 1990), resistance to proteolysis (Daniel et al., 1982; Fontana, 1988), and fluorescence quenching (Varley & Pain, 1991). A measure of the flexibility of a structure is provided by the RMS fluctuations of each atom relative to its average position during the dynamics trajectory. Such fluctuations have been calculated for the two proteins in the 300 K simulations. For each residue, the backbone C, N, and CA average RMS fluctuation is given. Figure 5 shows the RMS fluctuations per residue over the production run (400 ps). As expected, the N-terminal Met of RdDv and the two C-terminal residues of RdPf are most mobile because they do not interact with other residues. Residues 2–25 show similar fluctuations in the two proteins, but residues 25–50 exhibit higher fluctuations in RdDv. The

Table 2. Protein-protein hydrogen bonds in crystal and after molecular dynamics at 300 K for RdDv and RdPf^a

A. Hydrogen bonds in the crystal structure for RdDv						B. Hydrogen bonds in the crystal structure for RdPf					
Tyr	4	N-H	13	O-C	Tyr	Ala	2	N-HT1	15	OE1-CD	Glu
Val	5	N-H	50	O-C	Glu	Trp	4	N-H	13	O-C	Tyr
Cys	6	N-H	11	O-C	Tyr	Val	5	N-H	50	O-C	Glu
Thr	7	N-H	48	O-C	Glu	Cys	6	N-H	11	O-C	Tyr
Gly	10	N-H	6	O-C	Cys	Lys	7	N-H	48	O-C	Glu
Tyr	13	N-H	4	O-C	Tyr	Lys	7	NZ-HZ3	50	OE1-CD	Glu
Tyr	13	OH-HH	28	O-C	Thr	Gly	10	N-H	6	O-C	Cys
Ala	16	N-H	14	OD1-CG	Asp	Tyr	13	N-H	4	O-C	Trp
Glu	17	N-H	14	O-C	Asp	Glu	15	N-H	2	O-C	Ala
Gly	18	N-H	14	O-C	Asp	Asp	16	N-H	14	OD1-CG	Asp
Asp	19	N-H	24	O-C	Val	Ala	17	N-H	14	O-C	Asp
Asn	22	N-H	19	O-C	Asp	Gly	18	N-H	14	O-C	Asp
Asn	22	ND2-HD22	19	OD2-CG	Asp	Asp	19	N-H	24	O-C	Ile
Gly	23	N-H	20	O-C	Pro	Asn	22	N-H	19	O-C	Asp
Val	24	N-H	19	O-C	Asp	Asn	22	ND2-HD22	19	OD1-CG	Asp
Gly	27	N-H	15	O-C	Pro	Gly	23	N-H	20	O-C	Pro
Thr	28	N-H	25	O-C	Lys	Ile	24	N-H	19	O-C	Asp
Thr	28	OG1-HG1	25	O-C	Lys	Gly	27	N-H	15	O-C	Glu
Asp	32	N-H	29	OG-CB	Ser	Phe	30	N-H	15	OE2-CD	Glu
Asp	32	N-H	29	O-C	Ser	Glu	31	N-H	31	OE1-CD	Glu
Leu	33	N-H	30	O-C	Phe	Glu	32	N-H	29	O-C	Lys
Trp	37	N-H	34	O-C	Pro	Leu	33	N-H	30	O-C	Phe
Trp	37	NE1-HE1	19	OD2-CG	Asp	Asp	36	N-H	36	OD1-CG	Asp
Cys	39	N-H	44	O-C	Ala	Trp	37	N-H	34	O-C	Pro
Gly	43	N-H	39	O-C	Cys	Trp	37	NE1-HE1	19	OD2-CG	Asp
Lys	46	NZ-HZ2	37	O-C	Trp	Cys	39	N-H	44	O-C	Ala
Lys	46	NZ-HZ3	35	O-C	Ala	Gly	43	N-H	39	O-C	Cys
Glu	48	N-H	45	O-C	Pro	Lys	46	NZ-HZ2	33	O-C	Leu
Phe	49	N-H	46	O-C	Lys	Lys	46	NZ-HZ3	30	O-C	Phe
Glu	50	N-H	5	O-C	Val	Ser	47	OG-HG	48	OE2-CD	Glu
Ala	52	N-H	3	O-C	Lys	Glu	48	N-H	45	O-C	Pro
New hydrogen bonds formed after molecular dynamics at 300 K						Phe	49	N-H	46	O-C	Lys
Thr	7	OG1-HG1	50	OE1-CD	Glu	Glu	50	N-H	5	O-C	Val
Ala	16	N-H	14	OD2-CG	Asp	Leu	52	N-H	3	O-C	Lys
Asn	22	N-H	19	OD2-CG	Asp	Glu	53	N-H	51	O-C	Lys
Asn	22	ND2-HD22	11	OH-CZ	Tyr	New hydrogen bonds formed after molecular dynamics at 300 K					
Gly	23	N-H	19	O-C	Asp	Lys	3	NZ-HZ2	14	OD2-CG	Asp
Thr	28	N-H	15	O-C	Pro	Lys	7	NZ-HZ1	48	OE1-CD	Glu
Leu	33	N-H	29	O-C	Ser	Tyr	13	OH-HH	28	O-C	Thr
Lys	46	NZ-HZ1	30	O-C	Phe	Ala	17	N-H	14	OD1-CG	Asp
Lys	46	NZ-HZ1	33	O-C	Leu	Asn	22	N-H	19	OD1-CG	Asp
Lys	46	NZ-HZ2	30	O-C	Phe	Gly	23	N-H	19	OD1-CG	Asp
						Thr	28	OG1-HG1	25	O-C	Ser
						Lys	29	N-H	32	OE2-CD	Glu
						Leu	33	N-H	29	O-C	Lys

^a— indicates those that disappear after molecular dynamics at 300 K.

results show that the backbone RMS fluctuations of RdPf are somewhat smaller than those of RdDv; the average value for all residues is 0.8 Å for RdPf and 0.9 Å for RdDv; the more significant values, leaving out the three terminal residues, are 0.76 Å for RdPf and 0.88 Å for RdDv. The experimental values of the RMS fluctuations estimated from the crystallographic temperature factors for the two proteins (see the Materials and methods) do not show a significant difference. Moreover, they are much more uniform than the calculated values. In part, this is likely to be due to crystal contacts (Phillips, 1990), which are expected to be particularly important for a small protein, and, in part, to errors in the analysis of the X-ray data (Kuriyan et al., 1986).

Principal components analysis of the trajectories was performed (Caves et al., 1998) and the projections of the trajectories onto the first two principal components (PC1, PC2) is shown in Figure 6. The first two principal components (lowest frequencies) account for 50–55% of the total RMS displacement. PC1 is a low-frequency motion where loops 1 and 2 slide relative to each other, and PC2 a motion where the loops move away from each other. On these projections, we see three main clusters of conformations corresponding to three different values of the PC1 coordinate. The cluster on the left corresponds to the first part of the simulation. The clusters are better defined in RdPf than in RdDv. For RdPf, the transition to the second cluster occurs at about 90 ps and the

Fig. 5. Residue RMS fluctuations for RdPf and RdDv at 300 K averaged over the 400 ps simulations. Solid line corresponds to RdPf; dashed line corresponds to RdDv.

transition to the third cluster at about 280 ps. In RdDv, the clusters in the PC1–PC2 plane are more diffuse, which is consistent with the idea of higher flexibility for this protein. The transition to the second cluster occurs at about 80 ps and the transition to the third cluster at about 210 ps.

The 300 K simulations show that both proteins are stable at this temperature and do not drift significantly from the X-ray structure. There is no striking difference in behavior, but the smaller fluctuations of RdPf relative to RdDv appear to be significant and are consistent with the principal components analysis.

Energetic aspects

The last 100 ps of the 300 K simulations were used to perform a comparison of the energetics of the native state of the two systems. Table 3 lists the average protein–protein, protein–solvent, and solvent–solvent energies. Because the two systems contain different numbers of protein atoms and water molecules, the values on a per atom or per water molecule basis are reported in addition to the total energies. The protein–solvent energy is essentially identical in the two systems on a per atom basis, whereas the protein–protein energy is more negative for RdPf, both on a total and per atom basis. Both the van der Waals and electrostatic components contribute to this difference, although the former is larger than the latter. One source of the electrostatic energy difference is that RdPf has 18 charged residues compared with 14 for RdDv. This is probably also the reason why the water–water energy is slightly higher in the RdPf system, because the ionic groups are expected to disrupt some water–water interactions. The more favorable interactions of ionic groups in the native state of RdPf could be compensated by ionic group–water interactions in the unfolded state. The primary origin of the lower van der Waals component of the protein–protein energy in RdPf relative to RdDv is the larger number of atoms in the former. However, the van der Waals energy per atom is also lower in RdPf. This suggests that there may be better packing in the hydrophobic core of RdPf, which is likely to contribute to its stability.

Fig. 6. Projection of the motions from the room-temperature trajectories onto the first two principal components (lowest frequency quasiharmonic modes). PC1 and PC2 are the values of these principal components in arbitrary units. **A:** RdDv. **B:** RdPf.

373 K Simulations

These simulations are done at a temperature where experimental data indicate that RdPf is stable, whereas RdDv is expected to be unstable. RdPf did not show any appreciable increase in RMSD from the crystal structure (Fig. 7A). RdDv, on the other hand, starts unfolding at 300 ps and exhibits a rather steady increase in RMSD over the 1,020-ps simulation period. The RMSD decreases somewhat at about 650 ps, but rises back to about 3 Å (backbone) and 4 Å (all atoms) at the end of the simulation. The Rg plots show similar behavior (Fig. 7B). The Rg of RdPf fluctuates around the crystal value, whereas that of RdDv exhibits a slow but steady increase. As will be expounded in the Discussion, the results do not show that RdPf is stable at this temperature because kinetics as well as thermodynamics may be involved, i.e., it could be that in this simulation RdPf unfolds more slowly than RdDv and that a longer simulation would show unfolding. However, the results are suggestive of the expected difference in behavior of the two proteins.

Figure 8 shows the evolution of native contacts for RdDv and RdPf as a function of time. Many contacts flicker on and off. As expected from the RMSDs, the disappearance of contacts is more

Table 3. Energetics from the 300 K simulations^a

	RdPf	RdDv
Natoms	488	460
Nwaters	1,390	1,418
E^{PP} , Total ^b	$-1,459 \pm 6$ (-2.99)	$-1,195 \pm 8$ (-2.60)
Elec	$-1,651 \pm 6$ (-3.38)	$-1,390 \pm 10$ (-3.02)
vdW	-299 ± 0.6 (-0.613)	-266 ± 2 (-0.579)
E^{PS} , Total	$-3,049 \pm 15$ ($-449 \cdot 10^{-5}$)	$-2,930 \pm 20$ ($-449 \cdot 10^{-5}$)
Elec	$-2,716 \pm 16$ ($-400 \cdot 10^{-5}$)	$-2,613 \pm 18$ ($-401 \cdot 10^{-5}$)
vdW	-333 ± 0.7 ($-49.1 \cdot 10^{-5}$)	-317 ± 2 ($-48.6 \cdot 10^{-5}$)
E^{SS}	$-13,668 \pm 10$ (-9.833)	$-14,134 \pm 10$ (-9.967)

^aAverages over the last 100 ps of the simulations. All values in kcal/mol. Errors were estimated by averaging over five 20-ps subintervals. E^{PP} is the protein-protein energy, E^{PS} the protein-solvent energy, and E^{SS} the solvent-solvent energy. In parentheses are the protein-protein energy per atom, the protein-solvent energy per atom and per water molecule, and the solvent-solvent energy per water molecule.

^bTotal energy includes the bonding terms.

extensive for RdDv than for RdPf. Comparison with the 300 K plot (Fig. 4, note that the cutoff for the 373 K and the other high-temperature simulations is 1.5, rather than the value 1.3 used at 300 K) reflects the greater distortion of the structure at 373 K. The

Fig. 7. Simulation results at 373 K. **A:** Backbone RMSD from crystal structure. **B:** Radius of gyration as a function of time. Dark lines correspond to RdPf and light lines to RdDv.

Fig. 8. Contact plots for the 373 K simulations. Dot means that a particular contact is present at that time (interatomic distance less than 1.5 times that in the crystal structure); see Table 4 for definition of contacts. **A:** RdDv. **B:** RdPf.

unfolding “pathway” for RdDv can be deduced from the contact plot. The order of disappearance of contacts is (excluding contacts that disappear already at 300 K): #29 (β 2-L2 hydrophobic), #24 (L1-L2 hydrophobic), #37 (L1 backbone hydrogen bond), #40 (tertiary in L1), #27 (L1-L2 tertiary at the rim of hydrophobic core), #18 (a side-chain-backbone hydrogen bond between β 2-L1), #34 and 33 in the hydrophobic core, #16 (β 2-L1 hydrogen bond), and #25 (hydrophobic between L1 and L2). The β -sheet is very stable in this simulation. Unfolding starts by the separation of the L1 and L2 loops and exposure of the hydrophobic core. This is accompanied by penetration of water molecules into the hydrophobic core. In the last coordinate frame of the simulation, two water molecules are seen to hydrogen bond to the hydroxyl group of Tyr 13, one as a donor and one as an acceptor. These molecules also hydrogen bond to other water molecules closer to the rim of the pocket. Similar observations have been made in an extensive analysis of solvation during unfolding simulations of barnase (Cafilisch & Karplus, 1994a; 1995). The increase in R_g is paralleled by an increase in accessible surface area, which, at the end of the simulation, is $3,587 \text{ \AA}^2$, compared to $3,255 \text{ \AA}^2$ for the crystal structure.

473 K Simulations, set 1

473 K is a typical temperature for studies of unfolding by molecular dynamics (Li & Daggett, 1994) because unfolding at this temperature is fast enough (on the order of a few hundred ps) to be observed in the relatively short simulation times accessible with present day computers. The disadvantage of 473 K for the present study is that both proteins are thermodynamically unstable. Nevertheless, it is of interest to compare the nature of the folding transition for the two proteins to see if there are any significant differences. This could not be done at 373 K because RdPf did not unfold. Also, one might expect that the rate of unfolding is correlated with thermodynamic stability, although this is not necessarily true (see the Discussion).

The RMSD from the crystal structure is shown in Figure 9A. For the 473 K simulations, the RMSD is measured with respect to the starting simulation structure. In the first 100 ps of the simulation, RdDv exhibits a higher RMSD than RdPf, but RdPf appears to undergo a transition around that time; from 100 to 200 ps, the behavior is similar. The RMSD values after 200 ps are about 2.5 Å (backbone) and 3–3.5 Å (all atoms) for both proteins. The backbone RMSD between the two proteins is 3.09 Å. In the Rg for RdDv, there is an initial increase and a temporary decrease at about 100 ps. The Rg of RdPf exhibits a steady increase throughout the trajectory (Fig. 9B).

Contact plots are presented in Figure 10. In the RdPf simulation, the order of disappearance of contacts is: 27 and 24 (L1-L2 tertiary), 37 (L1 hydrogen bond), 33 (β 2-L2 side chain-backbone hydrogen bond), and 40 (L1 tertiary). All of them break early in

Fig. 9. First set of simulations at 473 K. **A:** Backbone RMSD. **B:** Radius of gyration. See Figure 7 for explanation.

Fig. 10. Contact plot for the first set of 473 K simulations. **A:** RdDv. **B:** RdPf. See Figure 8 for explanation.

the 373 K simulation of RdDv, which indicates that the proteins unfold in the same way. For RdDv, the order of disappearance of contacts is: #28 (β 2-L2 hydrophobic), #18 (β 2-L1 side-chain-backbone hydrogen bond), #35 (β 2- β 3 hydrophobic), #32 (β 2-L2 hydrophobic), #17 (β 2-L1 tertiary), #20 and 19 (β 3-L2 hydrogen bonds), #9 (β 1- β 3 hydrogen bond), and #27 (L1-L2 tertiary). The contacts that disappear early here are not identical to those in the 373 K simulations of RdDv, but they belong to the same regions, i.e., the loop region and the hydrophobic core. The first step in unzipping of the β -sheet, the break of contacts #2 and 7 (β 1- β 2 backbone hydrogen bonds), occurs at about 135 ps.

473 K Simulations, set 2

For RdPf, the RMSD increases slowly in the first 150 ps, then exhibits a more rapid increase up to 180 ps and a drop to the final value of about 2 Å (backbone) and 2.5 Å (all atoms) (Fig. 11A). The RdDv run shows similar behavior to 160 ps and after that has an increased RMSD. At 180 ps, the protein collided with the boundary and was resolved in a larger sphere (see Materials and meth-

ods). The simulation was continued, and the RMSD dropped somewhat before increasing again to the final value of about 4 Å (backbone) and 5 Å (all atoms). The R_g is shown in Figure 11B.

For RdPf, the contacts that disappear early are: #25, 27, 24, 40, and 37, as in the previous simulations, and #17 and 18, which are β 2-L1 tertiary contacts in the hydrophobic core. The unfolding pathway is very similar to that found in the previous simulation at 473 K. For RdDv, the order is: #37, 25, 17, 27, 9, 35, 29 (β 2-L2 hydrophobic), 18, 16 (β 2-L1 hydrogen bond), 40, 13 (β 1- β 3 hydrophobic on surface), 32, and 11 (β 1- β 3 hydrogen bond). Most of these contacts also disappeared early in previous simulations.

In the RdDv simulation, the β -sheet unfolds significantly for the first time. The β 1- β 3 interactions disappear from 105 to 310 ps and the β 1- β 2 interactions start breaking up at 280 ps. The order of disappearance of contacts in the β -sheet is: #9 (backbone hydrogen bond at the edge of β 1- β 3), #13 (β 1- β 3 surface hydrophobic), #11 (backbone hydrogen bond at the other edge of β 1- β 3), #10 and 8 (the remaining β 1- β 3 backbone hydrogen bonds), #6 and 3 (β 1- β 2 backbone hydrogen bonds), #12 (β 1- β 3 buried tertiary contact), #2 and 7 (β 1- β 2 tertiary involving Pro 15), and #4 (β 1- β 2 hydrogen bond). The reason that contacts #2 and 7 appear to break after #6 and 3 is in part due to the convention used for the presence of a contact (distance $\leq 1.5d$, where d is the distance in the crystal). The distance between Pro 15 and Lys 2 is relatively large in the crystal and therefore a longer distance is required to consider the contact broken. Broken protein-protein hydrogen bonds are replaced by protein-water hydrogen bonds. The observed behavior is consistent with the findings in unfolding simulations of barnase (Caflisch & Karplus, 1995).

The contacts that remain until the end are mostly local contacts in the polypeptide chain: #5 (β 1- β 2 hydrogen bond), #19 and 20 (β 3-L2 local hydrogen bond), #23 (L1-L2 local hydrogen bond), #31 and 34 (β 2-L2 hydrophobic), #38 and 39 (L1 local hydrogen bonds), #42 and 45 (L2 local hydrogen bonds), and #47 (L2 local hydrophobic), and the iron-sulfur cluster bonds, which are treated as covalent. Snapshots of RdDv at different times during the simulation are shown in Figure 12; the view is the same as in Figure 2.

500 K Simulations

These simulations lasted 260 ps. They exhibit different behavior from the previous simulations in that RdPf unfolds more rapidly than RdDv. The RMSD of RdPf increases steadily to a final value of about 5 Å (backbone), 6 Å (all atoms), whereas that of RdDv stabilizes at about 3 Å (backbone), 4 Å (all atoms) (Fig. 13A). The R_g plots are shown in Figure 13B.

Most contacts that disappear early in the RdPf simulation are similar to those that disappear early in other simulations. In addition, #36 (β 2- β 3 hydrophobic core) and #44 (L2 local hydrogen bond) also break early. The β -sheet starts unfolding at about 100 ps. Contact #1, which is unique in RdPf (a hydrogen bond between the side chain of Glu 25 and the indole of Trp 4), disappears first, along with adjacent #7 (15N-2O hydrogen bond). They are followed by #2, 3, and 6, which are all β 1- β 2 hydrogen bonds. In this simulation, the β 1- β 3 interactions remain till the end. Snapshots of the backbone during the simulation are shown in Figure 14.

In the RdDv simulation, there is limited unfolding. The β -sheet is much more stable. The contacts that disappear early are similar

Fig. 11. Second set of simulations at 473 K. **A:** Backbone RMSD. **B:** Radius of gyration. See Figure 7 for explanation.

Fig. 12. Backbone trace of structures along the unfolding trajectory of RdDv in the second 473 K simulation. Only CA, C, and N atoms are shown. Orientation is the same as in Figure 2.

to those observed in other simulations: #33, 18, 13, 11, 29, 32, 27, and also #15 (β 2-L1 local hydrogen bond). Again the loop region is much more labile than the β -sheet.

Discussion

Molecular dynamics simulations in solution can be used to obtain information on the dynamic behavior of proteins. Simulations at room temperature result in stable trajectories with small deviations from the experimental structure, whereas simulations at elevated temperatures result in extensive unfolding, in accord with the thermal denaturation of proteins observed experimentally. Thus, molecular dynamics is a natural tool to investigate differences between mesophilic and thermophilic or hyperthermophilic proteins. Room-temperature simulations can be used to determine the flexibility and high-temperature simulations can provide information on the rate and mechanism of unfolding.

In this paper, simulations at different temperatures have been performed for a hyperthermophilic rubredoxin from *P. furiosus* and one of its mesophilic homologues from *D. vulgaris*. At room temperature, both proteins are stable over a period of 400 ps and exhibit similar, moderate deviations from the crystal structures that are of the same order as those found for other proteins. The RMS fluctuations of RdPf are found to be smaller than those of RdDv, but the difference is not large, i.e., for residues 2–52, the values for RdPf and RdDv are 0.76 Å and 0.88 Å, respectively. Also, the principal components analysis supports the idea of a somewhat greater flexibility for the mesophilic protein at room temperature.

Fig. 13. Simulation at 500 K. **A:** Backbone RMSD. **B:** Radius of gyration. See Figure 7 for explanation.

Fig. 14. Backbone trace of structures along the unfolding trajectory of RdPf in the 500 K simulation. Orientation is the same as in Figure 2A.

This is the first time such a difference in flexibility has been observed in a simulation. For rubredoxin, a protein involved in oxidation–reduction reactions, only a small degree of flexibility appears necessary for function because the oxidized and reduced structures are very similar (Yelle et al., 1995). It will be of interest to determine whether the same effect exists for larger proteins that have enzymatic functions. It was found also that the intraprotein energy is lower in RdPf than RdDv, even on a per atom basis. This may be a factor in increasing the stability. The protein–solvent interaction energies are essentially identical on an atom basis.

As the temperature was raised, both proteins showed unfolding behavior. At 373 K, where RdPf is still marginally stable and RdDv is unstable, slow unfolding of the latter was observed. RdPf remained with its backbone RMSD at 2 Å from the crystal structure, whereas RdDv reached a deviation of 3 Å within the same time period. At 473 K and 500 K, both proteins unfold. In two sets of simulations at 473 K, RdDv unfolds before RdPf, as one would expect if the rate of unfolding correlated with decreased thermodynamic stability. In the 500 K simulations, however, RdPf unfolds faster. This is likely due not to the slightly higher simulation temperature, but to the stochastic element of the simulations. Simulations of chymotrypsin inhibitor 2 trajectories performed under the same conditions have been found to differ dramatically in the time when unfolding starts (Lazaridis & Karplus, 1997). Ideally, a large number of simulations are needed to obtain definitive results.

The high-temperature simulations provide information concerning the unfolding process, i.e., the sequence by which the protein structural elements disappear. The exact sequence of events is not

the same in the various simulations, but a consistent picture of unfolding emerges. Interactions involving the loop region and the hydrophobic core are relatively unstable and tend to disappear earlier. This can be seen by calculating the RMSDs for the separate regions. For the most unfolded structures, the last frame from 473 K-set 2 for RdDv and the last frame from 500 K from RdPf, the backbone RMSD for the β -sheet (residues 1–7, 10–15, 48–51) is 2.72 for RdDv and 2.35 Å for RdPf, whereas the RMSD for the remainder of the protein is 4.34 for RdDv and 5.72 Å for RdPf. This means that the β -sheet is kinetically more stable and suggests that it is more stable thermodynamically. Within the β -sheet, the β 1- β 2 interaction is more stable, as expected from the more extensive contacts, than the β 1- β 3 interaction. In one simulation, however (500 K for RdPf), the β 1- β 2 interactions disappeared first. Figure 15 shows the 48 contacts of RdDv divided into three classes depending on their kinetic stability. Contacts that disappear early are shown as thin dashed lines, those of intermediate stability as medium dashed lines, and those that are most persistent as thick dashed lines.

From examination of the structure of rubredoxin, it has been suggested that the N terminus of rubredoxin may be the first part of the protein to unfold because it is the region where major differences are observed between the hyperthermophilic and the mesophilic proteins (Blake et al., 1991, 1992a). All simulations show that the loop region (residues 16–47) is the most labile, especially at the rim of the hydrophobic core (residues 22 and 37). Moreover, the additional hydrogen bond that extends the β -sheet in RdPf (Glu 15–Ala 2) does not appear to be particularly strong, because it breaks in the room-temperature simulation. This is in accord with site-directed mutagenesis results, which showed that this extension of the β -sheet does not contribute to stability (Richie et al., 1996; Eidsness et al., 1997). Most labile are the contacts between loop residues and those between loop residues and the β -sheet. This is consistent with the experimental results showing that interactions between the β -sheet and the loop residues are important for kinetic stability (Eidsness et al., 1997).

Because flexibility is often involved in discussions of the stability and activity of thermophilic proteins (Jaenicke & Zavodszky, 1990), some general comments about the connection between flexibility, stability, and activity are appropriate. First, there is no single measure of flexibility. Protein dynamics is characterized by a large number of modes of motion that occur on vastly different time scales. Some of them may be functionally significant and others may be irrelevant (Frauenfelder et al., 1991). Thus, one cannot talk of flexibility in general, but must specify the type of flexibility. A protein can be "rigid" on a nanosecond time scale and "flexible" on a millisecond time scale. Second, flexibility cannot be a "cause" for thermodynamic instability. Flexibility per se implies increased conformational entropy of the native state and is therefore, favorable to thermodynamic stability. Rigidity offers no thermodynamic advantage to a protein, unless it leads to very favorable interactions or is required at room temperature if the protein is to be stable at higher temperatures. The latter has been suggested (Zuber, 1988; Wrba et al., 1990), but there is no evidence for this hypothesis. In the present rubredoxin simulations, there is some evidence that, on the picosecond time scale, the hyperthermophilic protein is less flexible than the mesophilic protein at room temperature. This is in accord with the lower van der Waals energy of the former.

There is no fundamental reason for correlation between stability and flexibility. In simple systems, flexibility is indicative of weak interactions, which, in turn, imply low stability. But in a complex system like a protein, this correspondence may not always hold. Flexibility depends on the shape of the energy landscape around the native state; thermodynamic stability depends on the difference in free energy between the folded and unfolded states, and kinetic stability (the rate of unfolding) depends on the activation barrier between folded state and the transition region for unfolding. Based on linear free energy relations (Fersht, 1985), the rate of unfolding often correlates with thermodynamic stability because a more stable protein is likely to have a higher barrier for unfolding, but this is not always the case. A change in the sequence of a protein can

Fig. 15. Stereo plot of RdDv (wall-eyed) showing the native contacts according to their stability in the unfolding trajectories. Only the main chain and the side chains involved in these contacts are shown. Thin dashed lines indicate the contacts that disappear early; medium dashed lines, the contacts of intermediate stability; and thick dashed lines, the contacts that tend to persist long in the simulations. Orientation is the same as in Figure 2A.

alter the stability and flexibility in different ways. For example, the creation of a cavity in the protein interior usually destabilizes the protein (Eriksson et al., 1992) and at the same time probably makes it more flexible (Beeser et al., 1997). In other cases, this correlation should not be expected. For example, the introduction of a salt bridge in the protein interior tends to make the protein more rigid, but at the same time can destabilize it thermodynamically (Hendsch & Tidor, 1994; Waldburger et al., 1995). A thermostable disulfide-bridged mutant of T4 lysozyme exhibited similar temperature factors in the crystal as the wild type (Pjura et al., 1990). Also, there are many modes of motion that have nothing to do with stability. For example, hinge bending motions between domains have no reason to be correlated with the stability of the domains themselves.

One argument for the correlation of stability and flexibility is based on hydrogen exchange data (Wrba et al., 1990). However, hydrogen exchange is mediated by both local and global fluctuations (Wagner & Wüthrich, 1978; Delepierre et al., 1983; Clarke et al., 1993). Slow hydrogen exchange is often determined by global unfolding of the protein, which occurs on a long time scale. Correlation of hydrogen exchange with the free energy of unfolding is therefore to be expected and can be used to measure the free energy change (Bai et al., 1994). By contrast, ^{13}C NMR relaxation experiments have shown that no correlation exists between thermostability and mobility on the nanosecond time scale (Wüthrich et al., 1980). The increased resistance of thermophilic proteins to proteolysis (Daniel et al., 1982; Fontana, 1988) does not necessarily imply flexibility on a short time scale because global or subglobal unfolding may be required for proteolysis to occur. Such global or subglobal unfolding is probably irrelevant to enzyme catalysis.

Rigidity is often cited as a cause of inactivity of thermophilic proteins at room temperature (Zuber, 1988; Wrba et al., 1990). Some correlation between dynamics and function has been observed in mesophilic proteins. For example, it has been found that ribonuclease A (Rasmussen et al., 1992) and bacteriorhodopsin (Ferand et al., 1993) become inactive at temperatures below 200 K and that this loss of activity is associated with a reduction in the atomic fluctuations; below the transition temperature, the motions are essentially harmonic (Smith et al., 1990). In the present simulations at room temperature, we found slightly higher fluctuations for the mesophilic rubredoxin. However, the differences appear to be too small to lead to an important decrease in activity (which for rubredoxin is not precisely known) of the hyperthermostable protein at room temperature; there are no measurements concerning the activity of RdPf at room temperature. Of course, it is possible that other thermophilic proteins would show more striking differences. One point that is not clear is whether the room-temperature inactivity (or low activity) of thermophiles deviates from a simple Arrhenius dependence of the reaction rate or whether a "freezing" phenomenon similar to the glass transition observed in many proteins (Frauenfelder et al., 1991) is involved.

Alternative mechanisms for the lack of activity of thermophilic enzymes at room temperature cannot be excluded. It has been suggested that thermophilic microorganisms and their enzymes have not evolved as extensively as mesophilic or psychrophilic ones (Adams, 1993). They may be less efficient for structural rather than dynamic reasons. Small shifts in the position of residues in the active site can produce large changes in activity, as has been shown for certain proteins (Joseph-McCarthy et al., 1995). The average structure of the enzyme may be slightly different at different temperatures. Near-UV CD and fluorescence studies of

thermophilic proteins showed temperature-dependent changes, suggesting structural reorganization in the protein (Wrba et al., 1990; Kotik & Zuber, 1992). Another reason for the low activity of thermophilic enzymes at room temperature may be the temperature dependence of pK_a values of basic residues (Danson et al., 1996).

Correlations between stability and activity have been observed in some cases (Matsumura et al., 1986), but not in others (Serrano et al., 1993). In fact, in some cases there is an inverse correlation (Gassner et al., 1996), i.e., a reduction in activity is accompanied by a reduction in stability. Thus, changes in activity are not necessarily coupled to changes in stability, unless active site residues are involved (Meiering et al., 1992). If that is the case, combination of the features that make thermophilic proteins stable with those that make meso/psychrophilic enzymes efficient could lead to enzymes with markedly improved performance. The reason why mesophilic proteins are marginally stable and consequently have a limited lifetime in the cell may be due to the general fact that the lifetime of most proteins is relatively short and to the need for a rapid response to regulatory signals (Hochachka & Somero, 1984), rather than a need for flexibility. Moreover, there is no evolutionary pressure to create superstable proteins.

More experimental and theoretical studies are needed to understand the differences between mesophilic and thermophilic proteins. In most cases, a very subtle set of structural and dynamic changes seems to be involved. These have so far evaded a clear characterization. The present paper has shown that, at least for rubredoxin, the hyperthermophilic protein has smaller room-temperature fluctuations and is likely to be kinetically more stable. However, multiple longer simulations in solution are needed to confirm this result. As in general analyses of protein thermodynamics (Lazaridis et al., 1995; Makhatazde & Privalov, 1995), simplified models may be useful.

Materials and methods

Energy function

The empirical energy function used here is the CHARMM polar hydrogen model (Brooks et al., 1983; Neria et al., 1996) in which polar hydrogen atoms are treated explicitly and nonpolar hydrogens (CH_3 , CH_2 , CH , SH) are incorporated into the heavy atom to which they are bonded. For the iron-sulfur cluster, parameters developed by Ichiye and coworkers were used (Shenoy & Ichiye, 1993; Yelle et al., 1995). The Fe-S interactions are treated as covalent bonds and so cannot be broken during the simulation; this is a good approximation under normal denaturation conditions. The parameters for the oxidized form of rubredoxin (Shenoy & Ichiye, 1993; Yelle et al., 1995), which is the one employed in most experimental studies, was used in the present simulations. The modified TIP3P model (Jorgensen et al., 1983; Neria et al., 1996) was used for water.

The nonbonded interactions were spherically truncated at 7.75 Å (Brooks et al., 1983). The electrostatic interactions were shifted to zero at 7.75 Å and the van der Waals interactions were switched off from 7.25 to 7.75 Å. The nonbonded list was generated with an 8-Å cutoff and updated every 10 steps for the 300 K and 373 K simulations and every 5 steps for the 473 K and 500 K simulations.

System preparation

The starting points of the simulations were the crystal structures of the two proteins: oxidized rubredoxin from *P. furiosus* (PDB code

Table 4. Native contacts used for analysis of the simulations

	RdPf		RdDv		Type ^a	Class ^b
1	Trp 4 NE1	Glu 15 OE2		— ^c	s-s	β 1- β 2
2	Ala 2 N	Glu 15 OE1	Lys 2 N	Pro 15 CG	m-s	
3	Trp 4 N	Tyr 13 O	Tyr 4 N	Tyr 13 O	m-m	
4	Cys 6 N	Tyr 11 O	Cys 6 N	Tyr 11 O	m-m	
5	Gly 10 N	Cys 6 O	Gly 10 N	Cys 6 O	m-m	
6	Tyr 13 N	Trp 4 O	Tyr 13 N	Tyr 4 O	m-m	
7	Glu 15 N	Ala 2 O	Pro 15 CD	Lys 2 O	m-m	
8	Val 5 N	Glu 50 O	Val 5 N	Glu 50 O	m-m	β 1- β 3
9	Lys 7 N	Glu 48 O	Thr 7 N	Glu 48 O	m-m	
10	Glu 50 N	Val 5 O	Glu 50 N	Val 5 O	m-m	
11	Leu 52 N	Lys 3 O	Ala 52 N	Lys 3 O	m-m	
12	Trp 4 O	Phe 49 CD2	Tyr 4 O	Phe 49 CD2	m-s	
13	Val 5 CG1	Leu 52 CD2	Val 5 CG1	Ala 52 CB	s-s	
14	Lys 7 NZ	Glu 50 OE1	Thr 7 CG2	Glu 50 OE2	s-s	
15	Gly 18 N	Asp 14 O	Gly 18 N	Asp 14 O	m-m	β 2-L1
16	Gly 27 N	Glu 15 O	Gly 27 N	Pro 15 O	m-m	
17	Tyr 13 OH	Ile 24 CG2	Tyr 13 OH	Val 24 CG1	s-s	
18	Tyr 13 OH	Thr 28 O	Tyr 13 OH	Thr 28 O	m-s	
19	Glu 48 N	Pro 45 O	Glu 48 N	Pro 45 O	m-m	β 3-L2
20	Phe 49 N	Lys 46 O	Phe 49 N	Lys 46 O	m-m	
21	Phe 30 CE1	Phe 49 CG	Phe 30 CE2	Phe 49 CG	s-s	
22	Trp 37 CZ3	Phe 49 CZ	Trp 37 CZ3	Phe 49 CZ	s-s	
23	Glu 32 N	Lys 29 O	Asp 32 N	Ser 29 O	m-m	L1-L2
24	Ile 24 CG2	Leu 33 CD2	Val 24 CG2	Leu 33 CD2	s-s	
25	Ile 24 CG1	Trp 37 CD1	Val 24 CG1	Trp 37 CD1	s-s	
26	Asp 19 OD2	Trp 37 NE1	Asp 19 OD2	Trp 37 NE1	s-s	
27	Asn 22 ND2	Pro 40 CB	Asn 22 ND2	Pro 40 CB	s-s	
28	Tyr 11 CZ	Pro 40 CG	Tyr 11 CZ	Pro 40 CG	s-s	β 2-L2
29	Tyr 11 CD1	Ile 41 CG1	Tyr 11 CD1	Val 41 CG2	s-s	
30	Tyr 13 CD1	Trp 37 CZ2	Tyr 13 CD1	Trp 37 CZ2	s-s	
31	Tyr 13 CD2	Phe 30 CD1	Tyr 13 CD2	Phe 30 CD2	s-s	
32	Tyr 13 CZ	Leu 33 CD2	Tyr 13 CZ	Leu 33 CD2	s-s	
33	Glu 15 OE2	Phe 30 N	Pro 15 CG	Phe 30 N	m-s	
34	Trp 4 CE2	Phe 30 CB	Tyr 4 CD2	Phe 30 CB	s-s	β 1-L2
35	Tyr 11 CD2	Phe 49 CE2	Tyr 11 CD2	Phe 49 CE2	s-s	
36	Tyr 13 CB	Phe 49 CE2	Tyr 13 CB	Phe 49 CE2	s-s	
37	Asp 19 N	Ile 24 O	Asp 19 N	Val 24 O	m-m	L1
38	Asn 22 N	Asp 19 O	Asn 22 N	Asp 19 O	m-m	
39	Gly 23 N	Pro 20 O	Gly 23 N	Pro 20 O	m-m	
40	Ala 17 O	Pro 26 CG	Glu 17 O	Pro 26 CG	m-s	
41	Asp 19 OD1	Asn 22 ND2	Asp 19 OD2	Asn 22 ND2	s-s	
42	Leu 33 N	Phe 30 O	Leu 33 N	Phe 30 O	m-m	L2
43	Trp 37 N	Pro 34 O	Trp 37 N	Pro 34 O	m-m	
44	Cys 39 N	Ala 44 O	Cys 39 N	Ala 44 O	m-m	
45	Gly 43 N	Cys 39 O	Gly 43 N	Cys 39 O	m-m	
46	Leu 33 CD1	Trp 37 CE3	Leu 33 CD1	Trp 37 CE3	s-s	
47	Trp 37 CZ2	Pro 40 CD	Trp 37 CZ2	Pro 40 CD	s-s	
48	Phe 30 O	Lys 46 NZ	Ala 35 O	Lys 46 NZ	m-s	
49	Leu 33 O	Lys 46 NZ	Trp 37 O	Lys 46 NZ	m-s	

^am, main chain; s, side chain.^b β 1, residues 1–8; β 2, residues 9–15; β 3, residues 48–52; L1, residues 16–29; L2, residues 30–47.^cContact #1 does not exist in RdDv (Tyr 4 is too far from Pro 15).

1CAA) and oxidized rubredoxin from *D. vulgaris* (PDB code 7RXN). Hydrogen atoms were built with the HBUILD algorithm in CHARMM (Brünger & Karplus, 1988). Both structures were minimized gently (200 steepest descent steps) in vacuum with their crystal waters (61 for RdPf and 180 for RdDv) to relieve any local strain. A sphere of TIP3P water molecules of 23 Å radius was then introduced, centered on the C α of Tyr 13. With this size sphere, all atoms in the native structure had a distance of at least 6 Å from the surface of the sphere (except for the C terminus of RdPf, which had a closest approach of 3.25 Å from the boundary). The C terminus is involved in contacts in the crystal, but has been shown by NMR to be disordered in solution (Blake et al., 1992b). The average water density in the sphere was 0.0327 Å⁻³ (0.978 g/cm³). All water molecules with their oxygen within 2.8 Å of any protein heavy atom or crystal water oxygen were deleted. The solvent was thermalized with the protein rigid for 1,000 steps (1 ps). The same sphere was overlaid four times with different orientations to fill any holes in the water structure, and each time all water within 2.8 Å from any protein heavy atom or other water molecule were deleted. The solvent was thermalized for another ps. The final number of water molecules (including crystal waters) was 1,390 for RdPf and 1,418 for RdDv. The total number of atoms was 4,658 for RdPf (488 protein atoms) and 4,714 for RdDv (460 protein atoms).

Molecular dynamics

The stochastic boundary method was used (Brooks & Karplus, 1983) and all calculations were performed with the CHARMM program (Brooks et al., 1983).³ Molecular dynamics with the Verlet algorithm was performed in the inner 21-Å sphere (the "reaction region") and Langevin dynamics was performed between 21 and 23 Å (the "buffer region"). The friction constant for the Langevin dynamics was 62 ps⁻¹ and was applied to the water oxygen atoms. All protein atoms are in the reaction region and are treated by molecular dynamics. The list of solvent molecules belonging to the buffer region was updated every 50 steps; there are approximately 400 water molecules in the buffer region.

SHAKE (van Gunsteren & Berendsen, 1977) was used for all bonds involving hydrogen (including the HH distance in TIP3P water) with a tolerance of 10⁻⁶ Å. For historical reasons, a time step of 1 fs was used for the 300 K and 373 K simulations. After various tests and other simulations at 600 K (Caflisch & Karplus, 1994a, 1995) showed that a 2-fs time step yielded meaningful results, the additional simulations at 473 K and 500 K were run with a 2-fs time step. The Langevin region provides contact with a constant temperature heat bath, so that the dynamics corresponds to a canonical ensemble. Coordinates were saved every 0.1 ps for analysis. Table 1 gives an overview of the simulations. The 300 K simulations serve as a control and provide information on the dynamics of the two proteins at room temperature. At 373 K, RdPf is expected to be stable and RdDv unstable. At 473 K and 500 K, both proteins are unstable and unfolding kinetics is fast enough to observe in short simulations. Such high temperatures are used commonly in molecular dynamics simulations to study protein unfolding (Caflisch & Karplus, 1994a; Li & Daggett, 1994).

³Version 22 of CHARMM was used. This and more recent versions of the program can be obtained under a licence agreement from the CHARMM Development Project. The e-mail address is marci@tammy.harvard.edu.

Analysis of the simulations

To analyze the simulations, the RMSD from the crystal structure and the Rg were monitored as a function of time. Also, a number of native contacts (49 in RdPf and 48 in RdDv, see Table 4), including hydrogen bonds, hydrophobic and other tertiary contacts, were examined. They were selected from a set of pairwise atom-atom interactions in the crystal structure with a distance of less than about 4 Å and provide information on rearrangements of the structural elements of the protein. This approach is based on that used previously to characterize the multiplicity of pathways in chymotrypsin inhibitor 2 (Lazaridis & Karplus, 1997). A contact is said to exist if the distance between the two atoms defining the contact is less than that in the crystal structure times a cutoff value. The cutoff value was chosen to be 1.5 for the unfolding (high-temperature) trajectories and slightly less, 1.3, for the 300 K trajectories, to reveal smaller fluctuations in the native contacts.

The room-temperature simulations were also subjected to principal components analysis. This is done by quasiharmonic analysis of the trajectories (Levy et al., 1984; Brooks et al., 1995) to obtain the quasiharmonic normal modes and subsequent projection of the trajectories on the two lowest frequency modes (Caves et al., 1998). For the room-temperature simulations, the RMS fluctuations relative to the average structure were calculated over the whole trajectory and over its four 100-ps segments. For this calculation, the coordinate frames from the trajectories were first superposed on the crystal structure to remove any effects of overall translation and rotation. The calculated RMS fluctuations were compared to those derived from the crystallographic temperature (*B*) factors through the equation $\text{RMSF}(i) = \sqrt{3B(i)/8\pi^2}$. This neglects the effects of static disorder on the *B* factors.

Hydrogen bonds were evaluated with the following criterion: the hydrogen-acceptor distance must be less than 2.5 Å and the D-H...A and H...A-AA angles must be greater than 90°.

Note added in proof

A recent paper (Hiller et al., 1997) uses hydrogen exchange measurements to infer that the "apparent melting temperature" of RdPf approaches 200 °C. It is interesting to note that unfolding occurs in the present simulations at 473 K (200 °C).

Acknowledgments

We thank Prof. D.C. Rees for providing the coordinates of the hyperthermophilic rubredoxin prior to their release from the PDB, Prof. T. Ichiye for providing the parameters for the iron-sulfur cluster, and Dr. A. Caflisch and Dr. L.S. Caves for useful advice on simulation methodologies. This work was supported in part by a grant from the National Institutes of Health.

References

- Adams MWW. 1993. Enzymes and proteins from organisms that grow near and above 100 °C. *Annu Rev Microbiol* 47:627-658.
- Adams MWW, Kelly RM. 1992. *Biocatalysis at extreme temperatures: Enzyme systems near and above 100 °C*. Washington: ACS.
- Adams MWW, Kelly RM. 1995. Enzymes from microorganisms in extreme environments. *Chem & Eng News* 73:32-42.
- Adman EJ, Sieker LC, Jensen LH. 1991. Structure of rubredoxin from *Desulfovibrio vulgaris* at 1.5 Å resolution. *J Mol Biol* 217:337-352.
- Alber T. 1989. Mutational effects on protein stability. *Annu Rev Biochem* 58:765-798.

- Alber T, Bell J, Dao-Pin S, Nicholson H, Wozniak J, Cook S, Matthews B. 1988. Replacements of Pro⁸⁶ in phage T4 lysozyme extend an α -helix but do not alter protein stability. *Science* 239:633–635.
- Argos P, Rossmann MG, Gran UM, Zuber H, Frank G, Tratschin JD. 1979. Thermal stability and protein structure. *Biochemistry* 18:5698–5703.
- Bai Y, Milne J, Mayne L, Englander S. 1994. Protein stability parameters measured by hydrogen exchange. *Proteins Struct Funct Genet* 20:4–14.
- Baumann H, Knapp S, Lundback T, Ladenstein R, Hard T. 1994. Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. *Struct Biol* 1:808–819.
- Beeser SA, Goldenberg DP, Oas TG. 1997. Enhanced protein flexibility caused by a destabilizing amino acid replacement of BPTI. *J Mol Biol* 269:154–164.
- Biesecker G, Harris JI, Thierry JC, Walker JE, Wonacott AJ. 1977. Sequence and structure of D-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *Nature* 266:328–333.
- Blake PR, Day MW, Hsu BT, Joshua-Tor L, Park JB, Hare DR, Adams MWW, Rees DC, Summers MF. 1992b. Comparison of the X-ray structure of native rubredoxin from *Pyrococcus furiosus* with the NMR structure of the zinc-substituted protein. *Protein Sci* 1:1522–1525.
- Blake PR, Park JB, Bryant FO, Aono S, Magnuson JK, Eccleston E, Howard JB, Summers MF, Adams MWW. 1991. Determinants of protein hyperthermostability: Purification and amino acid sequence of rubredoxin from the hyperthermophilic Archaeobacterium *Pyrococcus furiosus* and secondary structure of the zinc adduct by NMR. *Biochemistry* 30:10885–10895.
- Blake PR, Park JB, Zhou ZH, Hare DR, Adams MWW, Summers MF. 1992a. Solution-state structure by NMR of zinc-substituted rubredoxin from the marine hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Sci* 1:1508–1521.
- Bradley EA, Stewart DE, Adams MWW, Wampler JE. 1993. Investigations of the thermostability of rubredoxin models using molecular dynamics simulations. *Protein Sci* 2:650–665.
- Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: A program for macromolecular energy minimization and dynamics calculations. *J Comput Chem* 4:187–217.
- Brooks BR, Janecz D, Karplus M. 1995. Harmonic analysis of large systems I. Methodology. *J Comput Chem* 16:1522–1542.
- Brooks CL III, Karplus M. 1983. Deformable stochastic boundaries in molecular dynamics. *J Chem Phys* 79:6312–6325.
- Brooks CL III, Karplus M, Pettitt BM. 1988. Proteins: A theoretical perspective of dynamics, structure, and thermodynamics. *Adv Chem Phys* 71.
- Brünger AT, Karplus M. 1988. Polar hydrogen positions in proteins: Empirical energy placement and neutron diffraction comparison. *Proteins Struct Funct Genet* 4:148–156.
- 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* 91:1746–50.
- Caflisch A, Karplus M. 1994b. Molecular dynamics studies of protein and peptide folding and unfolding. In: Merz K, LeGrand S, eds. *The protein folding problem and tertiary structure prediction*. Boston: Birkhauser. pp 193–230.
- Caflisch A, Karplus M. 1995. Acid and thermal denaturation of barnase investigated by molecular dynamics simulations. *J Mol Biol* 252:672–708.
- Cavagnero S, Zhou ZH, Adams MWW, Chan SI. 1995. Response of rubredoxin from *Pyrococcus furiosus* to environmental changes: Implications for the origin of hyperthermostability. *Biochemistry* 34:9865–9873.
- Cavagnero S, Zhou ZH, Adams WW, Chan SI. 1997. Unfolding kinetics as a signature of protein stability. *Biochemistry* (in press).
- Caves LSD, Evansek JD, Karplus M. 1998. Locally accessible conformations of proteins: Multiple molecular dynamics simulations of crambin. *Protein Sci* (in press).
- Chan MK, Mukund S, Kletzin A, Adams MWW, Rees DC. 1995. Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* 267:1463–1469.
- Chothia C, Lesk AM. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J* 5:823–826.
- Clarke J, Hounslow AM, Bycroft M, Fersht AR. 1993. Local breathing and global unfolding in hydrogen exchange of barnase and its relationship to protein folding pathways. *Proc Natl Acad Sci USA* 90:9837–9841.
- Daggett V, Levitt M. 1993. Protein unfolding pathways explored through molecular dynamics simulations. *J Mol Biol* 232:600–619.
- Daniel RM, Cowan DA, Morgan HW, Curran MP. 1982. A correlation between protein thermostability and resistance to proteolysis. *Biochem J* 207:641–644.
- Danson M, Hough D, Russell R, Taylor G, Pearl L. 1996. Enzyme thermostability and thermoactivity. *Protein Eng* 9:629–630.
- Dauter Z, Sieker LC, Wilson KS. 1992. Refinement of rubredoxin from *Desulfovibrio vulgaris* at 1.0 Å with and without restraints. *Acta Crystallogr B* 48:42–59.
- Davies GJ, Gamblin SJ, Littlechild JA, Watson HC. 1993. The structure of a thermally stable 3-phosphoglycerate kinase and a comparison with its mesophilic equivalent. *Proteins Struct Funct Genet* 15:283–289.
- Day MW, Hsu BT, Joshua-Tor L, Park JB, Zhou ZH, Adams MWW, Rees DC. 1992. X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Sci* 1:1494–1507.
- Delepierre M, Dobson CM, Selvarajan S, Wedin RE, Poulsen FM. 1983. Correlation of hydrogen exchange behavior and thermal stability of lysozyme. *J Mol Biol* 168:687–692.
- Edmondson SP, Qiu L, Shriver JW. 1995. Solution structure of the DNA-binding protein sac7d from the hyperthermophile *Sulfolobus acidocaldarius*. *Biochemistry* 34:13289–13304.
- Eidsness MK, Richie KA, Burden AE, Kurtz DMJ, Scott RA. 1997. Dissecting contributions to the thermostability of *Pyrococcus furiosus* rubredoxin: β -sheet chimeras. *Biochemistry* 36:10406–10413.
- Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. 1992. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* 255:178–183.
- Ferrand M, Dianoux A, Petry W, Zaccai G. 1993. Thermal motion and function of bacteriorhodopsin in purple membranes: Effects of temperature and hydration studied by neutron scattering. *Proc Natl Acad Sci USA* 90:9668–9672.
- Fersht A. 1985. *Enzyme structure and mechanism*, 2nd ed. New York: Freeman.
- Fontana A. 1988. Structure and stability of thermophilic enzymes. Studies on thermolysin. *Biophys Chem* 29:181–193.
- Frauenfelder H, Sligar SG, Wolynes PG. 1991. The energy landscapes and motions of proteins. *Science* 254:1598–1603.
- Frey M, Sieker L, Payan F, Haser R, Bruschi M, Pepe G, LeGall J. 1987. Rubredoxin from *Desulfovibrio gigas*. A molecular model of the oxidized form at 1.4 Å resolution. *J Mol Biol* 197:525–541.
- Frommel C, Sander C. 1989. Thermitase, a thermostable subtilisin: Comparison of predicted and experimental structures and the molecular cause of thermostability. *Proteins Struct Funct Genet* 5:22–37.
- Gassner NC, Baase WA, Matthews BW. 1996. A test of the “jigsaw puzzle” model for protein folding by multiple methionine substitutions within the core of T4 lysozyme. *Proc Natl Acad Sci USA* 93:12155–12158.
- Hendsch ZS, Tidor B. 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci* 3:211–226.
- Hennig M, Darimont B, Sterner R, Kirschner K, Jansonius JN. 1995. 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: Possible determinants of protein stability. *Structure* 3:1295–1306.
- Hiller R, Zhou ZH, Adams MWW, Englander SW. 1997. Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200 °C. *Proc Natl Acad Sci USA* 94:11329–11332.
- Hochachka PW, Somero GN. 1984. *Biochemical adaptation*. Princeton: Princeton University Press.
- Jaenicke R. 1981. Enzymes under extremes of physical conditions. *Annu Rev Biophys Bioeng* 10:1–67.
- Jaenicke R, Zavadzsky P. 1990. Proteins under extreme physical conditions. *FEBS Lett* 268:344–349.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
- Joseph-McCarthy D, Petsko G, Karplus M. 1995. Use of a minimum perturbation approach to predict TIM mutant structures. *Protein Eng* 8:1103–1115.
- Jung D, Kang N, Jhon M. 1997. Site-directed mutation study on hyperthermostability of rubredoxin from *Pyrococcus furiosus* using molecular dynamics simulations in solution. *J Phys Chem A* 101:466–471.
- Klump HH, Adams MWW, Robb FT. 1994. Life in the pressure cooker: The thermal unfolding of proteins from hyperthermophiles. *Pure & Appl Chem* 66:485–489.
- Kollman P. 1996. Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Acc Chem Res* 29:461–469.
- Korndorfer I, Steipe B, Huber R, Tomschy A, Jaenicke R. 1995. The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J Mol Biol* 246:511–521.
- Kotik M, Zuber H. 1992. Evidence for temperature-dependent conformational changes in the L-lactate dehydrogenase from *Bacillus stearothermophilus*. *Biochemistry* 31:7787–7795.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Kuriyan J, Petsko GA, Levy RM, Karplus M. 1986. Effect of anisotropy and anharmonicity on protein crystallographic refinement. *J Mol Biol* 190:227–254.

- Lazaridis T, Archontis G, Karplus M. 1995. Enthalpic contributions to protein stability: Insights from atom-based calculations and statistical mechanics. *Adv Protein Chem* 47:231–306.
- Lazaridis T, Karplus M. 1997. Multiple unfolding simulations reconcile the “new view” of protein folding with the old. *Science* (in press).
- Levy RM, Srinivasan AR, Olson WK, McCammon JA. 1984. Quasi-harmonic method for studying very low frequency modes in proteins. *Biopolymers* 23:1099–1112.
- Li A, Daggett V. 1994. Characterization of the transition state of protein unfolding by use of molecular dynamics. Chymotrypsin inhibitor 2. *Proc Natl Acad Sci USA* 91:10430–10434.
- Lovenberg W, Sobel BE. 1965. Rubredoxin: A new electron transfer protein from *Clostridium pasteurianum*. *Proc Natl Acad Sci USA* 54:193–199.
- Makhatadze GI, Kim KS, Woodward C, Privalov PL. 1993. Thermodynamics of BPTI folding. *Protein Sci* 2:2028–36.
- Makhatadze GI, Privalov PL. 1995. Energetics of protein structure. *Adv Protein Chem* 47:307–425.
- Mark AE, van Gunsteren WF. 1992. Simulation of the thermal denaturation of hen egg white lysozyme: Trapping the molten globule state. *Biochemistry* 31:7745–7748.
- Matsumura M, Yasumura S, Aiba S. 1986. Cumulative effect of intragenic amino-acid replacements on the thermostability of a protein. *Nature* 323:356–358.
- Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu Rev Biochem* 62:139–60.
- Matthews BW, Jansonius JN, Colman PM, Schenborn BP, Dupourque D. 1972. Three dimensional structure of thermolysin. *Nature New Biol* 238:37–41.
- Matthews BW, Nicholson H, Becktel WJ. 1987. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci USA* 84:6663–6667.
- McCammon JA, Harvey SC. 1987. *Dynamics of proteins and nucleic acids*. Cambridge: Cambridge University Press.
- Meiering EM, Serrano L, Fersht AR. 1992. Effect of active site residues in barnase on activity and stability. *J Mol Biol* 225:585–589.
- Menendez-Arias L, Argos P. 1989. Engineering protein thermal stability. Sequence statistics point to residue substitutions in α -helices. *J Mol Biol* 206:397–406.
- Neria E, Fischer S, Karplus M. 1996. Simulation of activation free energies in molecular systems. *J Chem Phys* 105:1902–1921.
- Papavassiliou P, Hatchikian EC. 1985. Isolation and characterization of a rubredoxin and a two-(4Fe-4S) ferredoxin from *Thermodesulfobacterium commune*. *Biochim Biophys Acta* 810:1–11.
- Perutz MF, Raidt H. 1975. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature* 255:256–259.
- Phillips G. 1990. Comparison of the dynamics of myoglobin in different crystal forms. *Biophys J* 57:381–383.
- Pjura PE, Matsumura M, Wozniak JA, Matthews BW. 1990. Structure of a thermostable disulfide-bridged mutant of phage T4 lysozyme shows that an engineered cross-link in a flexible region does not increase the rigidity of the folded protein. *Biochemistry* 29:2592–2598.
- Rasmussen B, Stock A, Ringe D, Petsko G. 1992. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* 357:423–424.
- Rees DC, Adams MWW. 1995. Hyperthermophiles: Taking the heat and loving it. *Structure* 3:251–254.
- Richie KA, Teng Q, Elkin CJ, Kurtz DMJ. 1996. 2D ^1H and 3D ^1H - ^{15}N NMR of zinc-rubredoxins: Contributions of the β -sheet to thermostability. *Protein Sci* 5:883–894.
- Russell RJM, Hough DW, Danson MJ, Taylor GL. 1994. The crystal structure of citrate synthase from the thermophilic Archaeon, *Thermoplasma acidophilum*. *Structure* 2:1157–1167.
- Serrano L, Day AG, Fersht AR. 1993. Stepwise mutation of barnase to binase. A procedure for engineering increased stability of proteins and an experimental analysis of the evolution of protein stability. *J Mol Biol* 233:305–312.
- Serrano L, Kellis JT Jr, Cann P, Matouschek A, Fersht AR. 1992. The folding of an enzyme II. Substructure of barnase and the contribution of different interactions to protein stability. *J Mol Biol* 224:783–804.
- Shenoy VS, Ichiye T. 1993. Influence of protein flexibility on the redox potential of rubredoxin: Energy minimization studies. *Proteins Struct Funct Genet* 17:152–160.
- Sieker LC, Stenkamp RE, LeGall J. 1994. Rubredoxin in crystalline state. *Methods Enzymol* 243:203–216.
- Smith J, Kuczera K, Karplus M. 1990. Dynamics of myoglobin: Comparison of simulation results with neutron scattering spectra. *Proc Natl Acad Sci USA* 87:1601–1605.
- Spassov VZ, Karshikoff AD, Ladenstein R. 1995. The optimization of protein-solvent interactions: Thermostability and the role of hydrophobic and electrostatic interactions. *Protein Sci* 4:1516–1527.
- Starich MR, Sandman K, Reeve JN, Summers MF. 1996. NMR structure of HMFB from the hyperthermophile *Methanothermobacter fervidus* confirms that this archaeal protein is a histone. *J Mol Biol* 255:187–203.
- Stenkamp RE, Sieker LC, Jensen LH. 1990. The structure of rubredoxin from *Desulfovibrio desulfuricans* strain 27774 at 1.5 Å resolution. *Proteins Struct Funct Genet* 8:352–364.
- Stetter KO. 1982. Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105 °C. *Nature* 300:258–260.
- Swartz P, Ichiye T. 1996. Temperature dependence of the redox potential of rubredoxin from *Pyrococcus furiosus*: A molecular dynamics study. *Biochemistry* 35:13772–13779.
- Tanner JJ, Hecht RM, Krause KL. 1996. Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry* 35:2597–2609.
- van Gunsteren HJC, Berendsen HJC. 1977. Algorithms for macromolecular dynamics and constraint dynamics. *Mol Phys* 34:1311–1327.
- Varley PG, Pain RH. 1991. Relation between stability, dynamics and enzyme activity in 3-phosphoglycerate kinases from yeast and *Thermus thermophilus*. *J Mol Biol* 220:531–538.
- Vogt G, Argos P. 1997. Protein thermal stability, hydrogen bonds or internal packing. *Folding & Design* 2:540–546.
- Volki P, Markiewicz P, Stetter KO, Miller JH. 1994. The sequence of a subtilisin-type protease (aerolysin) from the hyperthermophilic archaeum *Pyrobaculum aerophilum* reveals sites important to thermostability. *Protein Sci* 3:1329–1340.
- Wagner G, Wüthrich K. 1978. Dynamic model of globular protein conformations based on NMR studies in solution. *Nature* 275:247–248.
- Waldburger CD, Schildback JF, Sauer RT. 1995. Are buried salt bridges important for protein stability and conformational specificity? *Nature Struct Biol* 2:122–128.
- Wampler JE, Bradley EA, Stewart DE, Adams MWW. 1993. Modeling the structure of *Pyrococcus furiosus* rubredoxin by homology to other X-ray structures. *Protein Sci* 2:640–649.
- Watanabe K, Chishiro K, Kitamura K, Suzuki Y. 1991. Pro residues responsible for the thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase. *J Biol Chem* 266:24287–24294.
- Watenpugh KD, Sieker LC, Jensen LH. 1979. The structure of rubredoxin at 1.2 Å resolution. *J Mol Biol* 131:509–522.
- Watenpugh KD, Sieker LC, Jensen LH. 1980. Crystallographic refinement of rubredoxin at 1.2 Å resolution. *J Mol Biol* 138:615–633.
- Wrba A, Schweiger A, Schultes V, Jaenicke R, Zavodszky P. 1990. Extremely thermostable D-glyceraldehyde-3-phosphate dehydrogenase from the eubacterium *Thermotoga maritima*. *Biochemistry* 29:7584–7592.
- Wüthrich K, Wagner G, Richarz R, Braun W. 1980. Correlations between internal mobility and stability of globular proteins. *Biophys J* 32:549–558.
- Yamada T, Akutsu N, Miyazaki K, Kakinuma K, Yoshida M, Oshima T. 1990. Purification, catalytic properties, and thermal stability of Threo-Ds-3-Isopropylmalate dehydrogenase coded by leuB gene from an extreme thermophile, *Thermus thermophilus* strain HB8. *J Biochem* 108:449–456.
- Yelle RB, Park NS, Ichiye T. 1995. Molecular dynamics simulations of rubredoxin from *Clostridium pasteurianum*: Changes in structure and electrostatic potential during redox reactions. *Proteins Struct Funct Genet* 22:154–167.
- Yip KSP, Stillman TJ, Britton KL, Artymiuk PJ, Bocker PJ, Sedelnikova SE, Engel PC, Pasquo A, Chiavaluce P, Consalvi V, Scandurra R, Rice DW. 1995. The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* 3:1147–1158.
- Zuber H. 1988. Temperature adaptation of lactate dehydrogenase: Structural, functional, and genetic aspects. *Biophys Chem* 29:171–179.