Investigations of the thermostability of rubredoxin models using molecular dynamics simulations

ELIZABETH A. BRADLEY,¹ DAVID E. STEWART,² MICHAEL W.W. ADAMS,¹
AND JOHN E. WAMPLER¹

¹ Department of Biochemistry and the Center for Metalloenzyme Studies, University of Georgia, Athens, Georgia 30602

² University Computing and Networking Services, University of Georgia, Athens, Georgia 30602

(RECEIVED July 17, 1992; REVISED MANUSCRIPT RECEIVED December 17, 1992)

Abstract

The effects of differences in amino acid sequence on the temperature stability of the three-dimensional structure of the small β -sheet protein, rubredoxin (Rd), was revealed when a set of homology models was subjected to molecular dynamics simulations at relatively high temperatures. Models of Rd from the hyperthermophile, *Pyrococcus furiosus* (Pf), an organism that grows optimally at 100 °C, were compared to three mesophilic Rds of known X-ray crystal structure. Simulations covering the limits of known Rd thermostabilities were carried out at temperatures of 300 K, 343 K, 373 K, and 413 K. They suggest that Rd stability is correlated with structural dynamics. Because the dynamic behavior of three Pf Rd models was consistently different from the dynamic behavior of the three mesophilic Rd structures, detailed analysis of the temperature-dependent dynamic behavior was carried out. The major differences between the models of the protein from the hyperthermophile and the others were: (1) an obvious temperature-dependent transition in the mesophilic structures not seen with the Pf Rd models, (2) consistent AMBER energy for the Pf Rd due to differences in nonbonded interaction terms, (3) less variation in the average conformations for the Pf Rd models with temperature, and (4) the presence of more extensive secondary structure for the Pf Rd models. These unsolvated dynamics simulations support a simple, general hypothesis to explain the hyperthermostability of Pf Rd. Its structure simplifies the conformational space to give a single minimum accessible over an extreme range of temperatures, whereas the mesophilic proteins sample a more complex conformational space with two or more minima over the same temperature range.

Keywords: hyperthermostability; molecular modeling; molecular dynamics; rubredoxin

The tools of computational chemistry allow investigation of the structural forces within proteins as well as simulation of their dynamic behavior. A number of computational studies employing molecular mechanics force fields have made use of small peptides and proteins in order to study protein folding, unfolding, or relative stabilities. The first was by Levitt and coworkers who studied the folding of bovine pancreatic trypsin inhibitor (Warshel & Levitt, 1975; Levitt, 1976) and carp myogen (Warshel & Levitt, 1976) using molecular mechanics but with a considerably simplified force field and representation of the rotational freedom of bonds. The simulation of calbindin D_{9k} in aqueous solution was recently reported (Ahlström

et al., 1989) as were two other studies (Brooks et al., 1988; Tirado-Rives & Jorgensen, 1991) that employed molecular dynamics simulations to probe unfolding of small helical peptides. Brooks et al. (1988), using the simplified molecular mechanics approach described by Levitt (1976) along with a Brownian dynamics component, simulated the movements of a 15-residue polyvaline helix. With this model they observed correlated, concerted transitions from helix to coil with unwinding rates at 298 K limited to about 10⁷/s. A more rigorous computational approach was used by Tirado-Rives and Jorgensen (1991) who modeled an altered 15 residue S-peptide fragment from the N-terminal of ribonuclease A using short helical peptides like those from the experimental studies of Finkelstein et al. (1991). The altered peptide had a relatively high α -helical content (~45% at 276 K) and a melting temperature below 20 °C. Molecular dynamics simulations of the helical peptide in model water (Jorgensen

Reprint requests to: John E. Wampler, Department of Biochemistry, University of Georgia, Athens, Georgia 30602.

Abbreviations: Cp, *Clostridium pasteurianum*; Dg, *Desulfovibrio gigas*; Dv, *D. vulgaris*; Pf, *Pyrococcus furiosus*; Rd, rubredoxin; RMS, root mean square.

et al., 1983) showed that the helical content was lower and more extended during higher temperature (358 K) simulations as compared to lower temperature (298 K) simulations. This structural difference, while similar to that observed by Brooks et al. (1988), occurred after much shorter simulation times.

Free energy perturbation calculations (see van Gunsteren & Weiner, 1989) have shown that computational methods using molecular mechanics force fields hold promise for understanding thermostability. In several cases, such calculations have been applied to evaluate the effect of mutations on the stability of protein structures. For example, Akke and Forsen (1990) used a simple force field of coulombic energies to evaluate the energetics of three different mutations in calbindin D_{9k}. Although the accuracy of the results was not impressive, the order of the predicted effect (and its sign) correlated well with experimental values. Similarly, the effects of two mutations in T4 lysozyme were simulated by Dang et al. (1989) and Tidor and Karplus (1991) using more complete force fields. The results in both cases were of the same magnitude and sign as experimental measurements. Recently, Sneddon and Tobias (1992) investigated mutations in the ribonuclease T1 core using this method, but did not include experimental results for comparison.

This report is the first attempt to use computational tools to investigate the relative dynamic behavior of a protein, in this case Rd, using sequences and structures from organisms covering an extreme range of optimal growth temperatures (~70 °C). Results from a preliminary study were recently reported (Wampler et al., 1992), but the hyperthermostable model that was generated was based on preliminary sequence data (see Blake et al., 1991). As previously discussed (Wampler et al., 1992, 1993) Rd was chosen for this study for four reasons: (1) its simple β -sheet structure, (2) its small size (approximately 50 residues), (3) the availability of X-ray crystal structures of several mesophilic Rds (Watenpaugh et al., 1979, 1980; Frey et al., 1987; Stenkamp et al., 1990; Adman et al., 1991; Dauter et al., 1992), and (4) the availability of the sequence of the hyperthermostable form, Pf Rd, which is stable at 95 °C for at least 24 h (Blake et al., 1991). Thermostability data, although sketchy, indicate that the mesophilic Rds are considerably less stable (Lovenberg & Sobel, 1965; Papavassiliou & Hatchikian, 1985) than Pf Rd (Blake et al., 1991). In the accompanying paper (Wampler et al., 1993) the three-dimensional structure of Pf Rd was determined by homology modeling using the structures of the mesophilic Rds from Cp, Dg, and Dv. This was one part of a tripartite effort also involving X-ray crystallography (Day et al., 1992) and NMR spectroscopy (Blake et al., 1992) to structurally characterize this protein. Herein we focus on the differences in dynamic behavior of mesophilic and hyperthermostable Rd structures that result from their differences in amino acid sequences and conformational interactions.

Results

General comparisons

The potential of the specific amino acid sequence of Pf Rd to stabilize it relative to the other mesophilic homologs from Cp, Dg, and Dv was investigated by molecular dynamics simulations performed at four temperatures (300 K, 343 K, 373 K, 413 K). Two temperatures, 300 K and 373 K, were chosen to be near the natural environmental temperature for the mesophilic and hyperthermophilic proteins respectively. One temperature, 343 K, was chosen to be near to or above the transition temperatures for reversible thermal denaturation of the mesophilic Rds but below the transition temperature for Pf Rd, and 413 K was chosen to be above the reversible denaturation temperature of Pf Rd (Klump et al., in prep.).

The individual comparisons of each parent protein model with the Pf Rd modeled from it (Figs. 1, 2) show that in all cases the Pf model has the lowest AMBER energy (although at the highest temperatures the AMBER energy of Cp Rd approaches that of Pf), indicating that the Pf Rd models are more stable than the mesophilic Rd models. The average AMBER energies of the three Pf models are similar at all temperatures (Fig. 1), and their trajectories (Fig. 2) show a similar range of variation. The radii of gyration of the Pf models are fairly constant with temperature (Fig. 3). In contrast, the simulations of the three mesophilic protein models give higher energies (Figs. 1, 2), show more variation in size (Fig. 3), and undergo a temperature-dependent structural transition between 343 K and 373 K (Fig. 1). Variation of the mean structures with temperature (Figs. 4, 5) and of the α -carbon atoms about the mean dynamics structures (calculated with the AMBER 4.0 molecular dynamics analysis package; see Materials and methods) over the course of the simulations (Fig. 6) clearly indicates that the Pf models are less flexible over the entire temperature range than their mesophilic counterparts. Specifically, there is a considerable reduction in the variability of the first 13 residues and residues 31–45 (Pf numbering). Flexibility is inherent in the loop region of residues 15–26 (Pf numbering) for all structures, but the Pf models display less-specific variability in this region when compared to their respective parent structures.

A more detailed analysis of the level of structural variability between the mesophilic and hyperthermophilic structures was obtained from all-atom RMS comparisons between molecular dynamics average structures (see Materials and methods for a description of calculation). For each Rd simulation (300 K, 343 K, 373 K, and 413 K), an average structure was obtained. These average structures were fit in pairs (300 K:343 K, 343 K:373 K, 373 K:413 K) in order to determine overall all-atom RMS deviations between structures representing concurrent temperature steps. Each residue's RMS value (per atom) was examined

Fig. 1. Average total AMBER energies from molecular dynamics simulations of rubredoxin (Rd) models. Averages were taken for 100 sampled structures in each case (10 ps simulation time). Error bars in each case represent ± 1 standard deviation. In each panel the mesophilic protein results are given by the dashed line and those for the *Pyrococcus furiosus* (Pf) Rd modeled from it are given by the solid line. **A:** *Clostridium pasteurianum* (Cp) and Pf/Cp. **B:** *Desulfovibrio gigas* (Dg) and Pf/Dg. **C:** *Desulfovibrio vulgaris* (Dv) and Pf/Dv.

using an arbitrary criterion (at least two atoms per residue with an RMS difference greater than 2.0 Å indicates a residue with a large position change), and it was determined that 68% of the residues in the mesophilic structures are involved in large fluctuations in position with each temperature step. In contrast, only 27% of the residues in the hyperthermophilic structures undergo similar

Fig. 2. Total AMBER energy trajectories for 10-ps simulations at four temperatures. In each case the broad lines are from the mesophilic structures: Cp (solid line), Dg (dark pattern), Dv (light pattern). The thin lines are for the Pf models: Pf/Cp (solid line), Pf/Dg (pattern), Pf/Dv (dotted line). Simulation temperatures are indicated by the Y-axis labels.

Fig. 3. Average radii of gyration from molecular dynamics simulations of Rd models. Averages were taken for 100 sampled structures in each case (10 ps simulation time). Error bars in each case represent ± 1 standard deviation. In each panel the mesophilic protein results are given by the dashed line and those for the Pf Rd modeled from it are given by the solid line. A: Cp and Pf/Cp. B: Dg and Pf/Dg. C: Dv and Pf/Dv.

fluctuation. A number of residues whose positions vary the most between 300 K and 343 K in the mesophilic structures are much less variable in the hyperthermophilic structures. These include (mesophilic numbering) residues 45–49, which comprise a helical region in the hyperthermophilic structures (see above), β -sheet residue 50, specific residues of the iron cage loops (9 and 43), and several nonpolar (44 and 45) and charged (17 and 36) residues. In addition, there are several residues that are highly variable in both Dg and Dv structures but which do not meet the criteria of fluctuation described above for the mesophilic Cp and the hyperthermophilic models. These are iron cage loop residues 10 and 42, and nonpolar residues 12, 33, and 38, as well as charged residue 36. Between 343 K and 373 K, the positions of some aromatic residues (4, 13, 30, and 37) vary in the mesophilic structures of Dg and Dv along with residues 7, 8, 41, and 42, which are

crucial to both of the iron cage loops. Residue 50, which contributes to the extended β -sheet region, continues to have large fluctuations in its positions in all mesophilic structures as indicated above. Between 373 K and 413 K, most of the flexibility of the mesophilic structures persists (except for the aromatic residues 4, 30, and 37) while the fluctuations in corresponding hyperthermophilic residues are still much less than those observed in the mesophilic structures in all temperature ranges.

When kinetic energy is added to the Pf Rd structures in the form of higher simulation temperatures, the changes in total AMBER are mostly the results of changes in the nonbonded van der Waals and electrostatic energy terms (Table 1). However, the patterns of specific electrostatic and hydrogen-bonding interactions (discussed below) remain relatively constant. The increase in total AMBER energy, which is nearly linear with temperature, occurs via numerous, fairly uniform, small differences rather than at the expense of some distinct and identifiable change in specific interactions. To a large extent the same can be said for the simulations of the mesophilic models except that in each of those cases heating from 343 K to 373 K does not increase the AMBER energy and in one simulation (Cp at 343 K), a transition to a lower energy state begins after about 3.5 ps of simulation time (Fig. 2).

The anomalous behavior observed during the Cp Rd simulation at 343 K was analyzed further in order to investigate the nature of the transition. There are two interpretations of this conformational change. One is that all three mesophilic models undergo a transition between 343 K to 373 K, but that the transition for Cp is actually sampled during the simulation at 343 K. The second interpretation is that Cp undergoes a unique transition at 343 K not sampled by the other X-ray models. To examine this question and the nature of the transition seen in the Cp simulation at 343 K, three types of analyses were employed. First, trajectories of each energy component of the Cp simulation at 343 K were examined and fit to models of one and two energy states using DANFIP (Wampler, 1990) analysis. Second, the RMS deviations in atomic positions of the average structures from all other simulations at 343 K and 373 K were calculated relative to average Cp structures determined separately for both of the 343 K energy states: one representing the first portion of the transition and the other representing the second portion of the transition. Finally, details of the secondary structure of the average structures for the Pf, Dv, and Dg Rd models at 343 K and 373 K were compared with both of the 343 K average Cp structures.

DANFIP analysis (Wampler, 1990) of the energy trajectory of the Cp simulation at 343 K shows that it can be modeled as two separate energy states. Sixty-six and one half percent of the time the structure is varying about a conformation with average AMBER energy of -894 ± 21 kcal/mol. For the remaining time, the structure varies about a conformation with average AMBER energy

Fig. 4. Stereo plots of the average main-chain traces for Pf Rd simulations. Average structures were generated as described in the text from the 100 structures sampled from each 10-ps simulation. The main-chain traces (connected α -carbon atoms) are shown with temperature indicated by line type. **A:** Pf/Cp. **B:** Pf/Dg. **C:** Pf/Dv. Solid line, 300 K; dashed line, 343 K; shaded pattern, 373 K; dotted line, 413 K.

of -818 ± 11 kcal/mol. When examining the fits of the transition to one- and two-state models, it was discovered that no single energy component transition from the 343 K Cp trajectories matches the same transition that is observed in the total energy trajectory. Interestingly, the only component that gives a two-state fit that is better (reduced chi square) than one state is the nonbonded electrostatic energy component, but the fit gives only 30% of the low-energy component ($-1,714$ kcal/mol) and 70%

of the high-energy component ($-1,665$ kcal/mol). Thus, the higher energy state predominates in this fit and the trajectory does not match the time course of the total energy trajectory. The nonbonded van der Waals term, on the other hand, seems to show transitions, but the fit to two components versus one does not improve chi squared. When the two energies are combined (nonbonded electrostatic and nonbonded van der Waals), the result is a trajectory that does indeed match the total energy and can

Fig. 5. Stereo plots of the average main-chain traces for mesophilic Rd simulations. Average structures were generated as described in the text from the 100 structures sampled from each 10-ps simulation. The main-chain traces (connected α -carbon atoms) are shown with temperature indicated by line type. **A:** Cp Rd. **B:** Dg Rd. **C:** Dv Rd. Solid line, 300 K; dashed line, 343 K; shaded pattern, 373 K; dotted line, 413 K.

be easily resolved using the DANFIP procedure into two components: 66.6% of a lower energy state ($-1,989$ kcal/mol, SD = 22 kcal/mol) and 33.4% of a higher energy state ($-1,928$ kcal/mol, SD = 15 kcal/mol). Thus, the combined nonbonded energies (electrostatic and van der Waals) explain both the shape, distribution, and en-

ergy differences seen in the total energy trajectory for Cp at 343 K.

The RMS differences in the average structures for the simulations at 343 K and 373 K are compared with the components of the Cp simulation at 343 K in Table 2. Average structures were calculated using the molecular

Fig. 6. Root mean square (RMS) deviations from the mean structure of the C α atoms of each residue. Panels are identified for each structure simulated. Dotted line, 300 K; short dashed line, 343 K; long dashed line, 373 K; solid line, 413 K.

dynamics analysis routine of AMBER 4.0 for the full trajectories (100 structures) of all models except Cp at 343 K. The Cp 343 K trajectory was divided into two groups of structures (the first 30 = Group A average; the last 40 = Group B average). At 373 K (for Dv and Dg) and during the last half of the 343 K simulation (for Cp), the average structures are more like the Pf structures at 343 K and 373 K than are the 343 K averages (average A in the case of Cp). For both Dg and Dv, the average structure of the 373 K simulation is considerably different from the 343 K structure. The change in average conformation with temperature is much smaller for the Pf models (also illustrated by Figs. 4, 5). It is therefore apparent that all three mesophilic structures undergo a conformational change between 343 K and 373 K, resulting in 373 K structures that are more like Pf structures than are the 343 K struc-

tures. Remarkably, in the case of Dv Rd, the 373 K structure resembles the Pf structures at that temperature more than it does its own 343 K structure.

Finally, secondary structure comparisons (detailed below in the next section) failed to reveal any specific change associated with the 343 K–373 K transitions, except with the Cp Rd 343 K trajectory where there is a loss of helix (residues 45–49) and concomitant formation of a β -bridge (residues 14 and 18) on the transition to lower energy. These differences are not consistently observed in comparing 343 K and 373 K structures of the other mesophilic Rds, although the 45–49 helix does show consistent temperature-dependent melting in these structures.

Detailed comparisons: Electrostatics

Details from the energy calculations (Table 1) show that in all cases the difference in total AMBER energy between the Pf Rd model and its parent resides in the nonbonded interaction terms (both electrostatic and van der Waals). No other components of the AMBER energy show a consistent or significantly large difference. On average, there is no consistent pattern of change in the van der Waals term, but the electrostatic term of each Pf model is consistently lower than that of its parent structure. Figure 7 shows the average electrostatic energy differences per residue between the parent structures and the Pf structures modeled from them. Extension of the bars above the X-axis indicates a stabilizing electrostatic contribution in the Pf Rd structure relative to the mesophilic structures, whereas extensions of the bars below the X-axis indicate a stronger stabilizing electrostatic interaction in the mesophilic proteins. Because residue 1 of the Pf Rd sequences aligns with residue 2 of the mesophilic structures, the residue numbers in this figure correspond to mesophilic Rds and the interaction energy of the first mesophilic residue is lost upon mutation to Pf Rd. This gives a negative bar for the first residue position in the figure.

Examination of the plots of Figure 7 reveals 9 residues in Pf Rd (Lys 2, Lys 6, Ile 7, Glu 14, Lys 28, Asp 34, Glu 47, Glu 49, Lys 50, and Glu 52) that make a consistently strong, differential, stabilizing electrostatic contribution. The predominant interactions of these residues are indicated in Table 3. Most of these interactions are not found uniformly in the simulations of the mesophilic Rds. In particular, it is likely that interactions of Ala 1, Glu 14, Lys 28, Glu 30, and Lys 52 act to tie the main-chain strands together. Similarly, the unique interaction of Lys 6 with Glu 47 and Glu 49 may serve to bridge regions of the protein that are far too separated to be stabilized by backbone interactions. Of these interactions, however, none exhibit major changes on heating except the small Glu 30 interaction in the Pf models from Cp and Dg.

Destabilizing interactions or loss of stabilizing electrostatic interactions are exhibited by four residues of the mesophilic structures (Fig. 7) at positions 1, 12, 17, and

Table 1. Average AMBER energy components of the calculations^a

Energy component	Temperature (K)	Rubredoxin structure dynamics simulated					
		Cp	Pf/Cp	Dv	Pf/Dv	Dg	Pf/Dg
TOTAL	300	-938	-972	-832	-990	-827	-976
	343	-867	-933	-714	-933	-778	-925
	373	-905	-925	-752	-912	-782	-895
	413	-834	-858	-640	-855	-755	-847
ANGLE	300	215	212	210	210	204	227
	343	232	233	227	234	225	247
	373	259	250	241	248	238	260
	413	264	261	256	264	259	276
DIHEDRAL	300	219	216	225	231	224	221
	343	235	239	240	238	229	233
	373	249	245	246	251	252	246
	413	254	260	254	251	253	252
N-B VDW	300	-278	-282	-275	-254	-272	-293
	343	-294	-271	-252	-255	-257	-290
	373	-285	-272	-268	-264	-267	-294
	413	-296	-272	-254	-270	-257	-281
N-B EEL	300	-1,714	-1,739	-1,624	-1,812	-1,618	-1,769
	343	-1,674	-1,749	-1,561	-1,788	-1,592	-1,770
	373	-1,760	-1,773	-1,607	-1,781	-1,635	-1,767
	413	-1,690	-1,734	-1,532	-1,741	-1,635	-1,741
HBOND ^b	300	-16	-16	-18	-15	-14	-17
	343	-17	-14	-17	-16	-13	-16
	373	-18	-16	-16	-14	-16	-16
	413	-17	-15	-17	-10	-16	-16
N-B 1,4 VDW	300	107	109	106	108	109	104
	343	110	108	110	109	105	109
	373	109	112	109	109	111	109
	413	114	106	110	110	112	110
N-B 1,4 EEL	300	529	527	544	541	540	553
	343	540	521	538	546	552	561
	373	542	530	544	539	535	566
	413	536	535	542	542	528	554

^a Energies in kilocalories/mole. Nonbonded energies for atoms linked 1,4 are calculated separately in AMBER. TOTAL, total AMBER energy; ANGLE, bond angle; DIHEDRAL, dihedral angle; N-B, nonbonded; VDW, van der Waals; EEL, electrostatic; Cp, *Clostridium pasteurianum*; Dg, *Desulfovibrio gigas*; Dv, *Desulfovibrio vulgaris*; Pf, *Pyrococcus furiosus*.

^b HBOND refers to the 10–12 potential energy term of the AMBER energy, which replaces the usual 6–12 van der Waals potential energy term when a nonbonded interaction involves a hydrogen bond. The electrostatic energy of the hydrogen bond is found within the total N-B EEL term.

25. The large destabilizing loss at position 1 is, of course, because N-terminal methionine in the mesophiles is eliminated on conversion to Pf Rd (Blake et al., 1991). Like the stabilizing interactions, the destabilizing interactions do not show any dramatic or specific temperature-dependent changes.

Detailed comparisons: Secondary structure

The electrostatic energy term in AMBER calculations includes the major contributions of salt bridges and hydrogen bonds. A smaller component, the hydrogen bond

10–12 potential, is listed separately in Table 1. Figure 8 shows the average main-chain hydrogen bond counts during the various simulations. Most of the Pf Rd structures during all of the simulations contain a greater number of main-chain hydrogen bonds than the mesophilic structures. Average hydrogen bond lengths show no significant change with time, temperature, or structure (data not shown).

In the Rd structures, main-chain hydrogen bonding is involved primarily in stabilizing β -sheets and, to a lesser extent, a small α -helix and some turns. Results from analysis using the secondary structure prediction program, DSSP (Kabsch and Sander, 1983), confirm that, com-

Table 2. Root mean square (RMS) differences in main-chain atom positions of average structures from molecular dynamics simulations of rubredoxin in Ångströms

Protein	Temperature (K)	Protein and temperature (K)							
		Cp 343—A	Cp 343—B	Dg 343	Dv 343	Pf(Cp) 343	Pf(Dg) 343	Pf(Cp) 373	Pf(Dg) 373
Cp	343—A		<u>1.55^a</u>	2.51	3.15	1.80	2.09	1.76	1.84
	343—B			3.13	4.13	2.48	1.89	1.62	1.61
Dg	343				3.29	2.47	2.73	2.47	2.68
Dg	373	2.55	3.01	<u>2.11</u>	3.64	2.24	2.41	2.32	2.41
Dv	343					2.97	3.92	3.62	3.93
Dv	373	2.18	2.81	2.99	<u>2.82</u>	2.03	2.39	2.34	2.37
Pf(Cp)	343						1.94	1.61	1.88
Pf(Cp)	373					<u>1.61</u>	1.23		1.23
Pf(Dg)	343							1.23	0.98
Pf(Dg)	373						0.98	1.23	

^a Underlined values indicate self-comparison of models.

pared to the mesophilic structures, there are higher numbers of hydrogen bonds in the Pf Rd models. The data also show that the β -sheet region (residues 2, 3, 4, 5, 10, 11, 12, 13, 48, 49, 50, and 51 in Pf Rd) is extensive, although highly dynamic with different bridging interactions of the sheet region present at different temperatures and times. In the Dg and Dv simulations, the β -sheet re-

gion involves less of the third, C-terminal strand, but with Cp Rd all three strands are as involved in the β -sheet, as they are in the Pf Rd. The difference in extent of β -sheet between the three mesophilic structures can probably be attributed, therefore, to the shortened C-terminal end of Dg and Dv (Figs. 4, 5). Also important from DSSP secondary structure analysis is the fact that residues Ala 1 and Glu 14 in the Pf models are not involved in the sheet region because their side-chain interaction precludes this. Inclusion of solvent might alleviate this distortion removing the direct side-chain interaction and allowing participation of both residues in β -sheet formation.

There are also several β -bridges that occur in the Pf Rd models in the temperature range studied. Bridges between residues 13 and 17 as well as between 18 and 23 are present in the Pf Rd models at most temperatures, whereas bridges between residues 36 and 45 and between 38 and 43 are more dynamic. A few of these interactions are seen in the mesophilic simulations (14–18 in Cp at 343 K and 413 K and in Dg at 373 K; 38–45 in Cp at 373 K; 19–24 in Dg at all temperatures except 413 K; and 38–45 and 39–44 in Dg at 343 K), but the β -sheet region is obviously much more extensive in the Pf Rd models and more stable with temperature (Fig. 4 vs. Fig. 5).

In addition to the β -sheet region, there is a short helical segment near the end of the Pf Rd models involving residues 44–48 (sequence = PKSEF), which runs perpendicular to the β -sheet and actually flows into it at its C-terminal end (see Fig. 9A). According to DSSP, this 5-residue helix appears as a less extensive turn in most of the mesophilic structures, particularly at higher temperatures. In Cp (corresponding sequence = GKDEF), the helical conformation is present at 343 K and 373 K, while in Dv (PKSEF), it is present at 300 K, and, in Dg (SKDAF), it is only seen in the minimized structure. In

Table 3. Strong electrostatic interactions between residues in Pf rubredoxin and in mesophilic rubredoxins (Cp, Dg, and Dv)

Pf rubredoxin ^{a,b}	Cp ^{b,c}	Dg ^{b,c}	Dv ^{b,c}
A1:E14	M1:E16	M1:K2	N/O ^d
K2:D13&D15	K3:D14&E16	N/O	K2:E12 K3:D14
A1:E30 (melts)	M1:D29	M1:E31	M1:D31
A1:D53	M1:E53 K2:E53&E54	M1:K51	N/O
K6:E47&E49	N/O	N/O	N/O
I7:Fe	V8&D47:Fe	C6&T7&V8:Fe	N/O
E48:Fe	L41&V44:Fe	N/O	G43&A44:Fe
E14:K28	N/O	N/O	N/O
K28:E30&E31	D29:K31 K31:D32	K29:E31 K29:E32	K25:D32
E30&E31:K50	N/O	E31:K51	N/O
D34:K45	D35:K46	D35:K46	N/O
K50:E52	N/O	N/O	N/O

^a This work.

^b Residue numbers refer to respective sequences.

^c From X-ray crystal structures.

^d N/O, not observed.

Fig. 8. Average numbers of main-chain hydrogen bonds from molecular dynamics simulations of Rd models. Averages were taken for 100 sampled structures in each case (10 ps simulation time). Error bars in each case represent ± 1 standard deviation. In each panel the mesophilic protein results are given by the dashed line and those for the Pf Rd modeled from it are given by the solid line. **A:** Cp and Pf/Cp. **B:** Dg and Pf/Dg. **C:** Dv and Pf/Dv.

contrast, this helix is found in most of the Pf Rd models and is present in one or more at all temperatures. In the hyperthermophilic models, this helical region may play a key stabilizing role because it is connected to both sides

Fig. 7. Average electrostatic interaction differences between the mesophilic Rd simulations and those for the Pf models. In each case all interactions of each residue in each protein for the 100 structures (10 ps simulation time) were averaged. Plotted are the global differences (per residue), which represent the sum of the interactions/residue for the three mesophilic simulations, minus the sum of the interactions/residue for the three different Pf simulations. Sequence numbering for the Pf structures was shifted 1 residue for proper alignment (see text). A bar above the X-axis indicates a more favorable interaction in the Pf structures; a bar below the X-axis indicates a more favorable interaction in the mesophilic structures. The results for each temperature are plotted (see Y-axis labels).

Fig. 9. **A:** Chain trace linking the α -carbons for the average structure of Pf/Dg Rd at 343 K showing the helix (residue 44–49) and locations of β -bridges (3:12, 4:49, 13:17, 18:23, 36:45, 38:43). **B:** The “molecular belt” is illustrated with the residues involved (bold line), the hydrogen bonds (dashed lines), and the chain trace of the rest of the protein (as in A). **C:** The nonpolar residues (4, 11, 51; 7, 40; 37, 44) on the surface of Pf/Dg Rd (at 343 K) are shown (bold line), and hydrogen bonds are shown with the dashed lines. The rest of the structure is shown with the light patterned line.

of the large β -sheet region via β -bridges and interior side-chain interactions. In fact, on one side of the β -sheet region, the end of the helix is directly attached to the last strand, whereas on the other, the helix is linked through a network of hydrogen bonds to the β -sheet (Asp 18 is linked to Asn 21, which in turn is linked to the side chain of Trp 36 and the carbonyl of Val 37), while three adja-

cent bridges (18–23, 36–45, 38–43) complete the network. Thus, the entire protein is restricted in its middle by “a molecular belt” of hydrogen and ionic bonds (Fig. 9B). In the mesophilic structures, the Trp 37 side-chain interaction with Asp 19 is present below 373 K, but the residue 22 (Asp in Cp, Asn in Dv, and Ser in Dg) and Val 38 linkage is not seen at any temperature.

Detailed comparisons: Nonpolar residue surface exposure

Over half of the amino acid residues in Pf Rd are nonpolar or aromatic (Wampler et al., 1993). Many of these nonpolar residues appear to be largely exposed to solvent as measured by DSSP (Kabsch & Sander, 1983; data not presented), and according to visual inspection although without explicit solvent, any conclusions made about solvent accessibility maybe entirely speculative. The nonpolar side chains of residues Ile 7, Ile 11, Ile 40, and Leu 51 appear to be extremely exposed to solvent in all average models at all temperatures. Other residues that appear to be moderately exposed are: Val 4, Ile 23, Leu 32, Trp 36, Val 37, and Pro 44. What is interesting about these nonpolar solvent-exposed residues is that they are poised around secondary structural elements. It is plausible to speculate that if the residues are exposed to the extent previously indicated, van der Waals contact between specific residues could significantly contribute to thermostability, especially if contact alters the thermodynamic equilibrium of folding. Val 4, Ile 11, and Leu 51 project perpendicularly out of the surface of the three-stranded β -sheet and are proximal to one another (Fig. 9C). Residues Ile 23, Thr 27, Leu 32, and Trp 36 are involved in tight interior packing interactions as determined from visual inspection of the space-filled structures and serve to overlay the previously mentioned hydrogen bond region of Asp 18–Asn 21–Trp 36 and Asn 21–Val 37 while residing below the short helical region of residues 44–48. Val 37 and Pro 44 perpendicularly project from the side surface of one of the loops that comprises the iron cage (residues 38–42). Ile 7 and Ile 40 perpendicularly project out of the top surface of the iron cage loops (residues 38–42 and 5–9) (Fig. 9C).

An additional difference in the composition of the proteins that might contribute to their differential thermostabilities is the percentage of nonpolar isoleucine residues that each structure contains. Pf Rd contains more isoleucine residues (4 vs. 2, 2, and 0 for mesophilic structures) and fewer valine residues (2 vs. 5, 3, and 5 for mesophilic structures) than the mesophilic proteins (Wampler et al., 1993), and most of these isoleucine residues are highly exposed to solvent while being proximal to other highly exposed nonpolar residues. According to the hydrophobicity scale presented by Creighton (1984), isoleucine is more hydrophobic than valine (2.5 vs. 1.5 kcal/mol, respectively) so it is probable that isoleucine would undergo stronger hydrophobic interactions.

Discussion

Molecular dynamics simulations using the modeled structures and the minimized X-ray structures give a consistent indication of stability for Pf Rd relative to the other pro-

teins. When the dynamics are simulated at 300 K, 343 K, 373 K, and 413 K, all three Pf Rd models exhibit lower AMBER energies (Fig. 1) and their radii of gyration are less variable with temperature than those of the mesophilic proteins (Fig. 3). Also in support of this conclusion are the smaller deviations of the Pf models about their mean structures during MD (Fig. 6) and the retention of more similar average structures with temperature (Fig. 4 vs. Fig. 5). The lack of increase in AMBER energy and large changes in radii of gyration seen in the mesophilic simulations at 343 K and 373 K indicate a conformational change not seen with the Pf Rd models. The trajectory for Cp at 343 K (Fig. 2) appears to sample this transition.

As Figures 4 and 5 indicate, the Pf Rd models remain virtually unchanged with increasing temperature, whereas the mesophilic structures change dramatically. Of these, the Cp Rd structure changes the least between 300 K and 413 K (55% of the residues exhibiting high variability). Indeed, Cp Rd has the closest sequence homology to Pf Rd (Blake et al., 1991; Wampler et al., 1993) and shows consistently lower AMBER energies at all temperatures compared to Dv Rd and Dg Rd. The Cp structure behaves more like the simulated Pf Rd modeled from it, whereas the other mesophilic structures behave less like their corresponding hyperthermophilic models.

The changes that occur in converting each of the three X-ray structures to the models of Pf Rd make possible significantly increased nonbonded interactions and prevent the temperature-induced transition. Analysis of the Cp Rd trajectory at 343 K indicates that the transition is composed of roughly equal parts of change in electrostatic and van der Waals interactions. Examination of the details of the electrostatic interactions before and after this transition shows that there are no major differences in interacting pairs with temperature (Fig. 7). More detailed examinations of all of the individual interactions (not shown) did not reveal any hidden information masked by averaging. Electrostatic interactions in Pf Rd such as those between Ala 1, Glu 14, Lys 28, Glu 30, and Lys 52 and that of Lys 6 with Glu 47 and Glu 49 appear to stabilize the structure at high temperatures since most of these interactions are not found during the mesophilic simulations.

Apparently, the differences in sequence between the mesophilic and the hyperthermophilic protein allow slightly more secondary structure interactions in the latter case with the energy of this difference hidden in the electrostatic term. This difference is also quantified by the number of main-chain hydrogen bonds (Fig. 8). The electrostatic interactions that comprise the molecular belt mentioned above are particularly important because they also include the helical region of residues 44–48 in the Pf Rd models, which might stabilize the iron cage loops (see Fig. 9C) and are differentially maintained between mesophilic and hyperthermophilic protein models. This region, along with differences in numbers of β -bridges, accounts

for the increased hydrogen bond count observed for the Pf Rd models.

The absence of electrostatic interactions with solvent in these simulations increases the apparent interaction energies between the residues even with the use of a distance-dependent dielectric constant (Weiner et al., 1984). Because this is particularly true for surface residue interactions, some of the strong electrostatic interactions that are observed in unsolvated simulations will be lost when the proteins are solvated. Indeed, several of the electrostatic interactions observed in the Pf Rd models but not seen in the mesophilic models (see Table 3) are not present in the X-ray crystal structure of Pf Rd (Day et al., 1992). Instead several of these residues have direct interaction with the solvent. For example, the close proximity of Lys 6, Glu 47, and Glu 49, which allows their strong electrostatic interaction in the Pf models, are not as strong in the X-ray structure with only Lys 6 and Glu 49 close enough (2.014 Å) to make a hydrogen bond. Similarly, the hydrogen-bonding atoms of Glu 14, Lys 50, Glu 30, and Trp 3 surround a water molecule in the Pf X-ray structure, whereas they interact directly in the Pf Rd models. If several of these strong electrostatic interactions are lost on solvation as suggested by the X-ray structures, conclusions made about the contribution of localized electrostatic interactions to overall protein stability may be overstated. However, since ionic interactions are entropically driven (Creighton, 1984), higher temperatures may result in greater ion pairing. Thus, the lack of temperature effect on the electrostatic interactions (Table 3) seen in our simulations may change with solvation so that enhanced stability with temperature may indeed be partly attributed to increased electrostatic interactions.

It has long been thought that the hydrophobic effect is a major factor in determining protein stability (Privalov & Gill, 1988) and large differences in thermostability have been achieved with single hydrophobic residue mutations (Lim et al., 1992). In fact, direct correlation between the fraction of total nonpolar buried residues and thermostability has been shown (Privalov & Makhataadze, 1990). Enzymes with larger proportions of the surface areas of their nonpolar residues buried, such as metmyoglobin, have higher melting temperatures as determined by scanning microcalorimetry (Privalov et al., 1989). In addition, packing interactions in hydrophobic cores are thought to stabilize myoglobin conformation (Adachi et al., 1992). Because the nonpolar core is highly conserved in all of the Rd structures, despite possible stability differences, it may be unlikely that hydrophobic interior packing is a contributor to enhanced thermostability of Rd from Pf. Because explicit solvent was not included in the simulations, it is not possible, however, to substantiate conclusions about hydrophobic packing at this time.

However, an additional role for hydrophobic interactions is suggested by the variation between structures obtained using three different approaches, namely molecular

modeling (Wampler et al., 1993), X-ray crystallography (Day et al., 1992), and NMR (Blake et al., 1992). All show that there are a number of exposed and/or partially exposed hydrophobic cluster regions involving Val 4, Ile 7, Ile 11, Ile 23, Leu 32, Val 37, Ile 40, Pro 44, and Leu 51. In the simulated models and in the NMR structure at 45 °C, these residues are more fully exposed to solvent than in the X-ray structures, and they are situated relative to one another in such a way as to suggest that they could strongly interact. For example, such an interaction between Val 4, Ile 11, and Leu 51 would stabilize the three-stranded β -sheet at elevated temperatures thereby stabilizing a major portion of the structure. It is interesting that Cp Rd, whose simulations resemble the Pf simulations the most (Figs. 4, 5), has Thr, Ile, and Val in these positions, whereas Dg and Dv have Val, Glu, and Gln or Ala in these positions, respectively. Additionally, a hydrophobic interaction between Ile 40 and Pro 44 in Pf Rd would stabilize one of the iron cage loops. In the Pf Rd models, the hydrophobic interactions of Ile 23, Thr 27, Leu 32, and Thr 36 could stabilize the helical region (residues 44–48), especially because this cluster rests below the helix. These hydrophobic interactions should be strong in solvents such that the entropic cost of ordered water around them would be reduced (Creighton, 1984). Evidence for this is found in studies of β -spiral structures (Urry, 1990), wherein compression of an extended structure due to packing of nonpolar residues has been shown to occur above certain transition temperatures as nonpolar groups aggregate. Because the temperature at which the transition occurs is highly dependent on pH and hydrophobicity of the residues within the repeating segments (Urry, 1990), close packing of solvent-exposed hydrophobic clusters in the Pf Rd structure would be expected to occur at specific temperatures. This should be particularly true with clusters comprised of very hydrophobic residues such as isoleucine, leucine, and valine. Moreover, if there is an antagonism between the structure of water required for polar hydration and that required for apolar hydration (Urry, 1990), the existence of the Glu 12 residue in the mesophilic structures (from Dg and Dv) in place of Ile 11 and Ile 12 in the Pf and Cp structures, respectively, might shift the folding equilibrium in such a way that a significant portion of the difference in thermal stability could be accounted for. Unfortunately, this issue and the overall extent of hydrophobic cluster packing interactions are not addressed by these unsolvated simulations and are presented as suggestions for enhanced structure thermostability since this study failed to indicate startling specific structural forces (electrostatics, secondary structure) that can be assumed to be major contributors to protein thermostability. It is possible that these specific forces were not apparent due to the lack of specific solvent, especially because Daggett and Levitt (1992) have shown that short helical peptides fail to unfold in molecular dynamics simulations in vacuo at elevated temperatures.

Although we have not assessed the relative strength in Pf Rd of the hydrophobic interactions in these clusters, it seems possible that they would withstand high temperatures due to a solvent-exposed hydrophobic effect. For example, a correlation between buried nonpolar surface area and heat capacity has been suggested (Privalov & Makhataдзе, 1990) and the packing of solvent-exposed nonpolar residues has been shown to dramatically change the conformation of a helical structure beyond a specific transition temperature (Urry, 1990). Because specific hydrophobic interactions have been experimentally observed, it is possible that hydrophobic interactions such as those suggested may significantly contribute to protein thermostability.

Although we have explicitly shown that there is a greater number of secondary structure interactions in the Pf modeled structures and an increase in AMBER electrostatic interactions compared to the mesophilic structures, the presence of enhanced secondary structure or side-chain interaction may not be the sole basis for structural thermostability, especially because these increased interactions are indifferent to elevated temperatures. Rather, these secondary structure interactions may have enhanced thermostability due to other forces that withstand extreme temperatures and in so doing permit the protein structure to remain organized. One possible added effect untested by this study is that of packing interactions within solvent-exposed hydrophobic clusters.

Overview

The Pf Rd models are less variable with temperature than their mesophilic counterparts and absorb heat during simulation in a uniform way. Surprisingly, the mesophilic models undergo a transition between 343 K and 373 K that makes them *more* like Pf Rd at high temperatures, not less. This suggests that the temperature stability in Pf Rd is due to differences that cause it to occupy a common region of conformational space over a broad temperature range. The mesophilic proteins seem to sample this conformation only at higher temperatures. If this is the case, there must be compensating changes in the structure or surface character of Pf Rd to allow it to carry out its biological function in a conformation that is characteristic of inactive mesophilic proteins. Thus, the conformational space sampled by the hyperthermostable protein has a single deep, broad minimum where the average structure is fairly constant over a broad temperature range, while the mesophilic proteins sample a more complex conformational space with two or more minima and the distribution of species between them is temperature dependent.

Compared to the simulations of Tirado-Rives and Jorgensen (1991), the data reported here involve much longer peptide chains (52 vs. 15 residues), no explicit solvation, and much shorter simulation times. One advantage of not including explicit water is that shorter simulation times

can be used to cover a broader area of conformational space. In some cases (see Wendoloski et al., 1987) the results with and without solvation appear to be very similar except for the longer time required by the solvated simulation. In other cases (see Stewart & Wampler, 1991), the solvated and unsolvated trajectories do not seem to be so directly related. Because there is good evidence from scanning differential microcalorimetry (Privalov & Gill, 1988) that internal interactions may well determine differences in stability, nonsolvated simulations are justified as a first step. However, extension of the simulation periods as well as the investigation of solvation effects are part of our ongoing studies.

Examination of the details of interactions from this study further supports the widely held opinion that thermostability arises from multiple effects. A number of reviews covering different aspects of this problem have recently been published (Adams & Kelly, 1992). Perhaps the current view is best summarized by the concluding remarks of Straume et al. (1992, p. 133): "The calculations presented here, in conjunction with the existing compositional and structural data for thermophilic proteins, indicate that the increased thermostability of those proteins cannot be accounted for by an enhancement of any particular interaction but in terms of the combined and cumulative effects of changes in most, if not all, interactions."

Materials and methods

Computations and molecular modeling were carried out using a Silicon Graphics 4D/20 Iris graphics workstation, the computational resources of the Biological Sciences Sequence/Structure Computational Facility and those provided by University Computing and Networking Services, both at the University of Georgia. The molecular dynamics calculations were carried out using a CYBER 205 high-performance computer and a Silicon Graphics 4D/35 workstation. The software utilized has been previously described (Stewart et al., 1987; Stewart, 1989; Wampler et al., 1990). Molecular mechanics and dynamics calculations were performed using the AMBER 3.0 suite of programs (Singh et al., 1986). The crystal structures used in this study were obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977; Abola et al., 1987) and by homology modeling as previously described (Wampler et al., 1993). The Rd metal center parameters used in this study were those used in our previous homology modeling (Stewart et al., 1987; Stewart, 1989; Wampler et al., 1992, 1993) and molecular dynamics simulations (Stewart and Wampler, 1991). Molecular modeling and general manipulations of the structures were carried out using the program SYBYL (Tripos Associates, Inc., St. Louis, Missouri), and data analysis and presentation were carried out using SPECOS SA (Wampler, 1991). Analyses of the components of the distribution of states were carried out

using the DANFIP procedure as described by Wampler (1990).

Minimizations

Prior to simulation all Rd structures/models were minimized using AMBER 3.0 as discussed in the accompanying paper (Wampler et al., 1993). For both minimization and dynamics simulations the following selections and options of the AMBER program were used: united atom force field, distance-dependent dielectric constant, 9.5 Å cutoff for nonbonded calculations, and conjugate gradient minimization until the RMS gradient in the energy was less than 0.001 kcal/mol·Å.

Molecular dynamics

The AMBER molecular dynamics calculations were carried out using the fully minimized structures in discrete intervals of 0.001 ps with the following precession: warming, equilibration, stabilization, and data collection. Each structure was warmed from 0 K to the final temperatures (300 K, 343 K, 373 K, 413 K) in increments of 30 K (0.2 ps) with final temperature increments as necessary. For example, warming to 343 K involved heating in 30 K intervals to 330 K with a final heating interval of 13 K. Warming was achieved using the coupled heat bath option of AMBER (Berendsen et al., 1984) so that the temperature of the system increased sequentially according to the reference temperature of the heat bath. Following equilibration (2.0 ps total in 0.2-ps intervals) and stabilization (4.0 ps total in 0.5-ps intervals) at the final temperatures, data collection was done in 0.5 ps intervals with structures collected every 0.1 ps for a total period of 10.0 ps. The AMBER options used for the molecular dynamics calculations were as follows: 9.5 Å cutoff for nonbonded interactions, distance-dependent dielectric with a value of 1.0000 for electrostatic interactions, united-atom force field, and the SHAKE option (van Gunsteren & Berendsen, 1977) for all bonds with 0.0005 Å tolerance. The force field parameters and other selections have been previously described (Wampler et al., 1990).

Static and dynamics analysis

Static analysis of the heated structures was performed using the Static Analysis module of AMBER 3.0. Energy decomposition and calculation of the radius of gyration, molar volume, dihedral torsion angles, and hydrogen-bonding patterns were performed. Average heated structures were obtained using the Molecular Dynamics Analysis module of AMBER 4.0, but average dynamics measurements were obtained on a per structure basis (one structure per 10^{-13} s) using the Static Analysis module. Average heated structures for each simulation were obtained after translation and rotation of each coordinate set (collected at 0.1-ps intervals) in order to make the best fit with respect to the first set of coordinates.

Acknowledgments

This work was funded in part by grants from the following: National Institutes of Health (GM41482), the North Atlantic Treaty Organization (0404/88), the National Science Foundation (DIR-9014281 and BCS-9011583), and the Office of Naval Research (N00014-90-J-1894). Special thanks are due to Dr. Walter McRae, Assistant Vice President for Computing, for support funds and computer time on the high performance machines, to Doug C. Rees for kindly providing the Pf Rd X-ray crystal coordinates, and to Michael F. Summers for providing the NMR solution-state coordinates.

References

- Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F., & Weng, J. (1987). Protein Data Bank. In *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F.H., Bergerhoff, G., & Sievers, R., Eds.), pp. 107–132. Commission of the International Union of Crystallography, Bonn/Cambridge/Chester.
- Adachi, S., Sunohara, N., Ishimori, K., & Morishima, I. (1992). Structure and ligand binding properties of leucine 29 (B10) mutants of human myoglobin. *J. Biol. Chem.* 267, 12614–12621.
- Adams, M.W.W. & Kelly, R.M., Eds. (1992). *Biocatalysis at Extreme Temperatures: Enzyme Systems Near and Above 100°C*. American Chemical Society, Washington, D.C.
- Adman, E.T., Sieker, L.C., & Jensen, L.H. (1991). Structure of rubredoxin from *Desulfovibrio vulgaris* at 1.5 Å resolution. *J. Mol. Biol.* 217, 337–352.
- Ahlström, P., Teleman, O., Kördel, J., Forsén, S., & Jönsson, B. (1989). A molecular dynamics simulation of bovine calbindin D_{9k}. Molecular structure and dynamics. *Biochemistry* 28, 3205–3211.
- Akke, M. & Forsen, S. (1990). Protein stability and electrostatic interactions between solvent exposed charged side chains. *Proteins Struct. Funct. Genet.* 8, 23–29.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., & Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81, 3684–3690.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112, 535–542.
- Blake, P.R., Park, J.-B., Bryant, F.O., Shigetoshi, A., Magnuson, J.K., Eculston, E., Howard, J.B., Summers, M.F., & Adams, M.W.W. (1991). Determinants of protein hyperthermostability: 1. Purification, amino acid sequence and secondary structure by NMR of the rubredoxin from the hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Biochemistry* 30, 10885–10895.
- Blake, P.R., Park, J.-B., Zhou, Z.H., Hare, D.R., Adams, M.W.W., & Summers, M.F. (1992). Solution-state structure by NMR of zinc-substituted rubredoxin from the marine hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Sci.* 1, 1508–1521.
- Brooks, C.L., III, Karplus, M., & Pettitt, B.M. (1988). Proteins: A theoretical perspective of dynamics, structure and thermodynamics. *Adv. Chem. Phys.* 71, 128–129.
- Creighton, T.E. (1984). Physical forces that determine the properties of proteins. In *Proteins: Structures and Molecular Principles*, pp. 133–157. W.H. Freeman and Company, New York.
- Daggett, V. & Levitt, M. (1992). Molecular dynamics simulations of helix denaturation. *J. Mol. Biol.* 223, 1121–1138.
- Dang, L.X., Merz, K.M., & Kollman, P.A. (1989). Free energy calculations of protein stability: Thr-157 → Val-157 mutation of T4 lysozyme. *J. Am. Chem. Soc.* 111, 8505–8508.
- Dauter, Z., Sieker, L.C., & Wilson, K.S. (1992). Refinement of rubredoxin from *Desulfovibrio vulgaris* at 1.0 Å with and without restraints. *Acta Crystallogr.* B48, 42–59.
- Day, M.W., Hsu, B.T., Joshua-Tor, L., Park, J.-B., Zhou, Z.H., Adams, M.W.W., & Rees, D.C. (1992). X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hy-

- perthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Sci.* 1, 1494–1507.
- Finkelstein, A.V., Badretidinov, A.Y., & Ptitsyn, O.B. (1991). Physical reasons for secondary structure stability: α -Helices in short peptides. *Proteins Struct. Funct. Genet.* 10, 287–299.
- 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 Ångströms resolution. *J. Mol. Biol.* 197, 525–541.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., & Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.
- Levitt, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* 104, 59–107.
- Lim, W.K.L., Brouillete, C., & Hardman, J.K. (1992). Thermal stabilities of mutant *Escherichia coli* tryptophan synthase α -subunits. *Arch. Biochem. Biophys.* 292, 34–41.
- Lovenberg, W. & Sobel, B.E. (1965). Rubredoxin: A new electron transfer protein from *Clostridium pasteurianum*. *Proc. Natl. Acad. Sci. USA* 54, 193–199.
- Papavassiliou, P. & Hatchikian, E.C. (1985). Isolation and characterization of a rubredoxin and a two (4Fe-4S) ferredoxin from *Thermodesulfovibrio commune*. *Biochim. Biophys. Acta* 810, 1–11.
- Privalov, P.L. & Gill, S.T. (1988). Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39, 191–234.
- Privalov, P.L. & Makhatadze, G.I. (1990). Heat capacity of proteins II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: Protein unfolding effects. *J. Mol. Biol.* 213, 385–391.
- Privalov, P.L., Tiktopulo, E.I., Venyaminov, S.Y., Griko, Y.V., Makhatadze, G.I., & Khechinashvili, N.N. (1989). Heat capacity and conformation of proteins in the denatured state. *J. Mol. Biol.* 205, 737–750.
- Singh, U.C., Wiener, P.K., Caldwell, J.W., & Kollman, P.A. (1986). AMBER (USCF), Version 3.0. Department of Pharmaceutical Chemistry, University of California, California.
- Sneddon, S.F. & Tobias, D.J. (1992). The role of packing interactions in stabilizing folded proteins. *Biochemistry* 31, 2842–2846.
- Stenkamp, R.E., Sieker, L.C., & Jensen, L.H. (1990). The structure of rubredoxin from *Desulfovibrio desulfuricans* strain 27774 at 1.5 Å resolution. *Proteins Struct. Funct. Genet.* 8, 352–364.
- Stewart, D.E. (1989). The structure, interactions, and dynamics of electron transport proteins from *Desulfovibrio*: A molecular modeling and computational study. Ph.D. Dissertation, University of Georgia, Athens, Georgia.
- Stewart, D.E. & Wampler, J.E. (1991). Molecular dynamics simulations of the cytochrome c_3 -rubredoxin complex from *Desulfovibrio vulgaris*. *Proteins Struct. Funct. Genet.* 11, 142–152.
- Stewart, D.E., Weiner, P.K., & Wampler, J.E. (1987). Prediction of the structure of proteins using related structures, energy minimization and computer graphics. *J. Mol. Graph.* 5, 137–144.
- Straume, M., Murphy, K.P., & Freire, E. (1992). Thermodynamic strategies for protein design: Increased temperature stability. In *Biocatalysis at Extreme Temperatures: Enzyme Systems Near and Above 100 °C* (Adams, M.W.W. & Kelly R.M., Eds.), pp. 122–135. American Chemical Society, Washington, D.C.
- Tidor, B. & Karplus, M. (1991). Simulation analysis of the stability mutant R96H of T4 lysozyme. *Biochemistry* 30, 3217–3228.
- Tirado-Rives, J. & Jorgensen, W.L. (1991). Molecular dynamics simulations of the unfolding of an α -helix analogue of ribonuclease A S-peptide in water. *Biochemistry* 30, 3864–3871.
- Urry, D.W. (1990). Protein folding and assembly. An hydration-mediated free energy driving force. In *Protein Folding: Deciphering the Second Half of the Genetic Code* (Gierasch, L.M. & King, J., Eds.), pp. 63–71. American Association for the Advancement of Science, Washington, D.C.
- van Gunsteren, W.F. & Berendsen, H.J.C. (1977). Algorithms for macromolecular dynamics and constraint dynamics. *J. Mol. Phys.* 34, 1311–1327.
- van Gunsteren, W.F. & Weiner, P.K., Eds. (1989). *Computer Simulation of Biomolecular Systems*. ESCOM Science Publishers B.V., Leiden, The Netherlands.
- Wampler, J.E. (1990). Analysis of the probability distribution of small random samples by non-linear fitting of integrated probabilities. *Anal. Biochem.* 186, 209–219.
- Wampler, J.E. (1991). Specos SA, A computer program for managing, graphing and manipulating laboratory data. *Anal. Instrum.* 19, 203–230.
- Wampler, J.E., Bradley, E.A., Adams, M.W.W., & Stewart, D.E. (1992). Computational approaches to modeling and analyzing thermostability in proteins. In *Biocatalysis at Extreme Temperatures: Enzyme Systems Near and Above 100 °C* (Adams, M.W.W. & Kelly, R.M., Eds.), pp. 153–173. American Chemical Society, Washington, D.C.
- Wampler, J.E., Bradley, E.A., Stewart, D.E., & Adams, M.W.W. (1993). Modeling the structure of *Pyrococcus furiosus* rubredoxin by homology to other X-ray structures. *Protein Sci.* 2, 640–649.
- Wampler, J.E., Stewart, D.E., & Gallion, S.L. (1990). Molecular dynamics simulations of proteins and protein-protein complexes. In *Computer Simulation Studies in Condensed Matter Physics II* (Landau, D.P., Mon, K.K., & Schütler, H.-B., Eds.), pp. 68–84. Springer-Verlag, New York.
- Warshel, A. & Levitt, M. (1975). Computer simulation of protein folding. *Nature* 253, 694–698.
- Warshel, A. & Levitt, M. (1976). Folding and stability of a helical protein: Carp myogen. *J. Mol. Biol.* 106, 421–437.
- Watenpugh, K.D., Sieker, L.C., & Jensen, L.H. (1979). The structure of rubredoxin at 1.2 Å resolution. *J. Mol. Biol.* 131, 509–522.
- Watenpugh, K.D., Sieker, L.C., & Jensen, L.H. (1980). Crystallographic refinement of rubredoxin at 1.2 Ångströms resolution. *J. Mol. Biol.* 138, 615–633.
- Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., & Weiner, P.L. (1984). A new force field for molecular mechanical simulations of nucleic acids and proteins. *J. Am. Chem. Soc.* 106, 765–784.
- Wendoloski, J.J., Matthew, J.B., Weber, P.C., & Salemme, F.R. (1987). Molecular dynamics of a cytochrome c -cytochrome b_5 electron transfer complex. *Science* 238, 794–797.