Folding and Association of Thermophilic Dimeric and Trimeric DsrEFH Proteins: Tm0979 and Mth1491

ARTICLE *in* BIOCHEMISTRY · MAY 2009

Impact Factor: 3.02 · DOI: 10.1021/bi801784d · Source: PubMed

9 AUTHORS, INCLUDING:

Céline Galvagnion
University of Cambridge
16 PUBLICATIONS **210** CITATIONS

SEE PROFILE

Aron Broom
University of Waterloo
8 PUBLICATIONS **84** CITATIONS

SEE PROFILE

Kenrick A Vassall
University of Guelph
18 PUBLICATIONS **202** CITATIONS

SEE PROFILE

Folding and Association of Thermophilic Dimeric and Trimeric DsrEFH Proteins: Tm0979 and Mth1491[†]

Céline Galvagnion,[‡,§] Martin T. J. Smith,[‡] Aron Broom,[‡] Kenrick A. Vassall,[‡] Gabriela Meglei,[‡,∥] Joseph A. Gaspar,[‡] Peter B. Stathopulos,[‡,⊥] Bo Cheyne,[‡,#] and Elizabeth M. Meiering*[,‡]

Guelph−Waterloo Centre for Graduate Work in Chemistry and Biochemistry and Department of Chemistry, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, Commissariat à l'Energie Atomique, Institut de Biologie et Technologies de Saclay, Service de Bioénergetique Biologie Structurale et Mécanismes, Laboratoire des protéines membranaires, Gif-sur-Yvette F-91191, France, Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada, Division of Signaling Biology and Department of Medical Biophysics, Ontario Cancer Institute and University of Toronto, Toronto M5G 1L7, Ontario, Canada, and Department of Civil Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

Received September 18, 2008; Revised Manuscript Received December 12, 2008

ABSTRACT: Although the majority of natural proteins exist as protein−protein complexes, the molecular basis for the formation and regulation of such interactions and the evolution of protein interfaces remain poorly understood. We have investigated these phenomena by characterizing the thermal and chemical denaturation of thermophilic DsrEFH proteins that have a common subunit fold but distinct quaternary structures: homodimeric Tm0979 and homotrimeric Mth1491. Tm0979 forms a moderate affinity dimer, and a monomeric intermediate is readily populated at equilibrium and during folding kinetics. In contrast, the Mth1491 trimer has extremely high stability, so that a monomeric form is not measurably populated at equilibrium, although it may be during folding kinetics. A common mechanism for evolution of quaternary structures may be facile formation of a relatively stable monomeric species, with stabilizing intermolecular interactions centering on alternative environments for a β-strand at the edge of the monomer, augmented by malleable hydrophobic interactions. The exceptional trimer stability arises from a remarkably slow unfolding rate constant, 6.5×10^{-13} s^{-1}, which is a common characteristic of highly stable thermophilic and/or oligomeric proteins. The folding characteristics of Tm0979 and Mth1491 have interesting implications for assembly and regulation of homo- and heterooligomeric proteins *in vivo*.

The large majority of proteins in nature exist as protein−protein complexes, and complex formation is central to protein function and regulation (*1, 2*). Current understanding of the formation of protein complexes is very limited, however, at a quantitative predictive level and at an experimental mechanistic level (*3, 4*). In recent years, the field of protein folding has advanced from mainly studies of simple monomeric proteins to analyses of protein complexes, ususally homomeric and especially homodimeric proteins (*5, 6*). Only a handful of trimers have been characterized in detail (*7*−*17*). Analyses of the folding of related oligomeric proteins can provide important insights into the molecular basis of protein assembly, regulation, and evolution. Two interesting examples are Cro monomers and dimers (*18*) and monomeric, dimeric, and tetrameric mutants of immunoglobulin-binding domain B1 of streptococcal protein G (*19*). For both sets of proteins, the mechanisms by which alternative quaternary structures arise appear to involve the formation of fluctuating monomeric species, which are further stabilized by the formation of variable intermolecular interactions. A common feature in these and many other natural and disease-related protein complexes is variable intermolecular interactions involving β-strands at the edge of a monomer (*20, 21*).

Here we report quantitative analysis of the folding of two thermophilic members of the DsrEFH protein superfamily: homodimeric DsrH protein, Tm0979 from *Thermotoga maritima,* and homotrimeric DsrE/F protein, Mth1491 from *Methanobacterium thermoautotrophicum.* Interestingly, different quaternary structures are often observed for different members of

[†] This research was funded by the Natural Sciences and Engineering Research Council of Canada.

* Corresponding author: e-mail, meiering@uwaterloo.ca; phone, 519-888-4567ext 32254; fax, 519-746-0435.

[‡] Guelph−Waterloo Centre for Graduate Work in Chemistry and Biochemistry and Department of Chemistry, University of Waterloo.

[§] Commissariat à l'Energie Atomique, Institut de Biologie et Technologies de Saclay, Service de Bioénergétique Biologie Structurale et Mécanismes et URA CNRS 2096, Laboratoire des Protéines membranaires, Gif-sur-Yvette, F-91191, France.

[∥] Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8.

[⊥] Division of Signaling Biology and Department of Medical Biophysics, Ontario Cancer Institute and University of Toronto, Toronto M5G 1L7, Ontario, Canada.

[#] Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

[1] Abbreviations: K_d, dissociation constant; ΔG_U, change in Gibbs free energy; m_U, dependence of ΔG on denaturant concentration; GdmCl, guanidinium chloride; CD, circular dichroism; T_m^{app}, apparent midpoint of a thermal unfolding transition; ΔH_{vH}, van't Hoff enthalpy of denaturation at T_m; ΔC_p, change in heat capacity of unfolding at T_m; ΔG_U, Gibbs free energy of unfolding in water; K_U, equilibrium constant of unfolding in water; k_{uH_2O}, rate of unfolding in water; k_{fH_2O}, rate of refolding in water; $m^{\ddagger}_u$, the dependence of the unfolding rate constant, k_u, on GdmCl concentration; $m^{\ddagger}_f$, the dependence of the refolding rate constant, k_f, on GdmCl concentration; SSE, sum of squared errors.

10.1021/bi801784d CCC: $40.75
Published on Web 03/16/2009

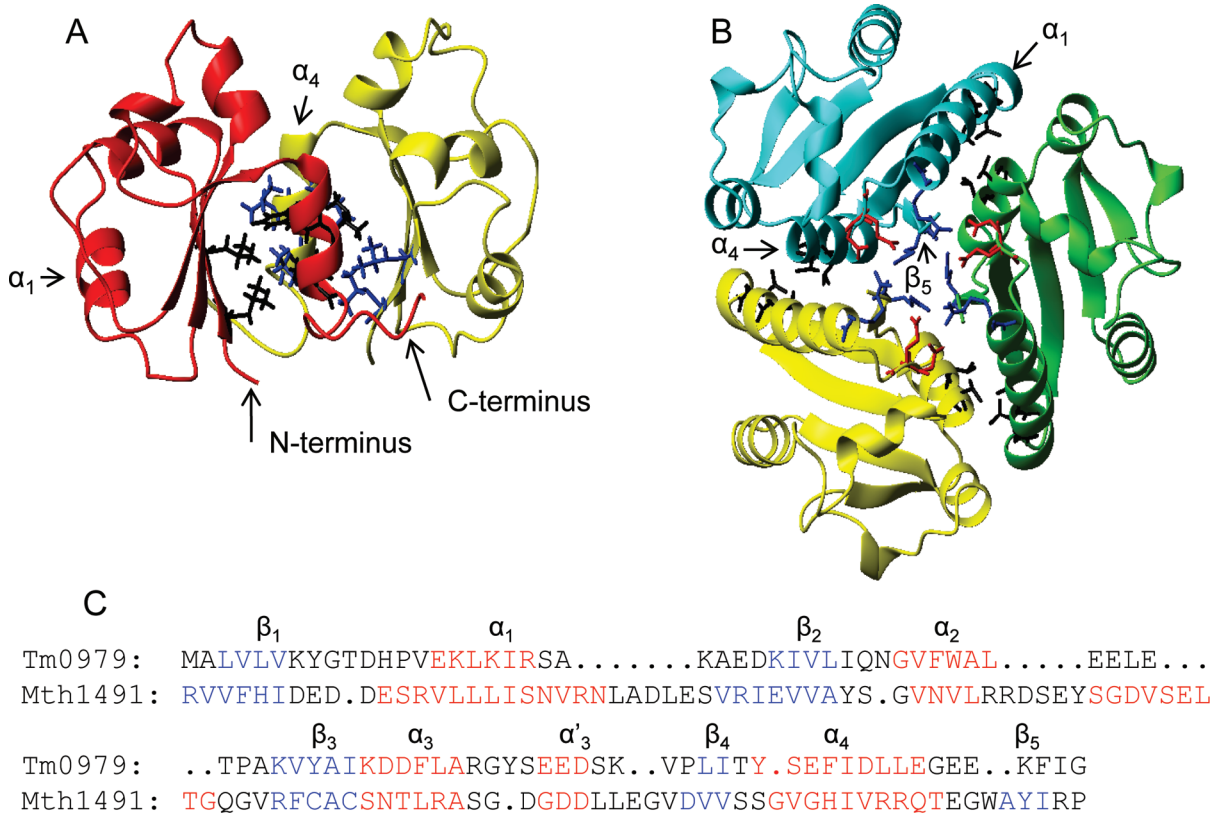

FIGURE 1: Structures of Tm0979 dimer and Mth1491 trimer in ribbon representation (A, B) and primary amino acid sequence alignment (C). (A) The two Tm0979 subunits are represented as red and yellow ribbons with side chains shown in blue and black, respectively. Subunit interfaces include hydrophobic interactions between the fourth α-helices (Phe75, Ile76, Leu79) of both subunits and between the fourth α-helix of one subunit (Ile76, Leu79) and first *β*-strand (Leu3, Leu5) of the other. (B) The three Mth1491 subunits are represented as cyan, yellow, and green ribbons. Subunit interfaces include hydrophobic interactions between subunits shown in black (Leu28, Leu32 from one subunit and Val97, Val101 from the other) as well as ionic interactions among acidic (Asp11, Glu12) and basic (Arg17, Arg112) residues shown in red and blue, respectively. Ribbon diagrams were generated using MolMol (*74*) and PDB codes 1x9a and 1l1s for Tm0979 and Mth1491, respectively. (C) Primary sequence alignment is based on Dali structural alignment. Residues in *β*-strands and α-helices are colored blue and red, respectively. The chain length for Tm0979 and Mth1491 is 87 and 112 amino acids, respectively. Sequence identity is 17%.

a given superfamily; however, little is known about the molecular basis for these variations (*4, 22*). Tm0979 and Mth1491 have very different primary sequences but share a common subunit fold consisting of four *β*/α units that form a parallel *β*-sheet with strand order 1234 and the first three helices packing on the opposite face of the sheet to the fourth helix (Figure 1). Mth1491 contains an additional C-terminal *β*-strand, located at the edge of the subunit parallel to the first strand and at the center of the trimeric structure; in contrast, the C-terminus of Tm0979 is somewhat shorter and poorly structured on the protein surface. Tm0979 can form a monomeric or a dimeric structure of moderate (micromolar) affinity (*23, 24*); no other stability or subunit affinity data are available for DsrEFH proteins. Other family members exhibit additional quaternary structure variations: homohexameric or heterohexameric, which are dimers of homo- or heterotrimers with central C-terminal *β*-strands (*25*−*27*). The functions of DsrEFH proteins are not well characterized but are involved in various aspects of sulfur metabolism (*23, 25, 28*−*32*).

We report here that the Tm0979 dimer dissociates relatively easily to form a monomer that is moderately stable and is populated readily during equilibrium and kinetic folding experiments. In contrast, Mth1491 forms a remarkably stable trimer, so that a monomer is not measurably populated at equilibrium but may be populated during folding kinetics. The results have specific and general implications for understanding and modeling the assembly and regulation of homo- and heterooligomeric proteins. The findings are also discussed in terms of evolution of quaternary structure via monomeric species and variations in interactions of edge *β*-strands.

MATERIALS AND METHODS

Protein Expression and Purification. Tm0979 and Mth1491 were prepared using pET15b/BL21(Gold*λ*DE3) expression systems, as described previously (*23, 28*), with the following modifications. For Mth1491, 1 mM ethylenediaminetetraacetic acid (EDTA), 10 mM dithiothreitol (DTT), and 10% (v/v) glycerol were added after elution of the protein from the Ni affinity column to minimize aggregation. Purified proteins were concentrated and exchanged by ultrafiltration (YM3 membrane, Millipore) into phosphate buffer (450 mM NaCl, 25 mM NaH_2PO_4/Na_2HPO_4, pH 6.5) for Tm0979 or into citrate buffer (450 mM NaCl, 1 mM EDTA, 10 mM DTT, 10% (v/v) glycerol, 20 mM $Na_2C_6H_6O_7$/$Na_3C_6H_5O_7$, pH 6.0) for Mth1491. All further experiments were conducted in these buffers unless otherwise indicated. Protein concentrations were determined from solution absorbance at 280 nm, A_{280}, using experimentally determined extinction coefficients and calculated molecular weights (*M*), His tag included, for the monomeric subunits of 11167.53 M^{-1} cm^{-1}

Table 1: Equations for Equilibrium Unfolding of Dimeric and Trimeric Proteins[a]

Dimer-2-state model		Dimer-3-state monomer intermediate model		Dimer-3-state dimer intermediate model	
$N_2 \xleftarrow{K_U} 2U$	Scheme I	$N_2 \xleftarrow{K_{U1}} 2I \xleftarrow{K_{U2}} 2U$	Scheme II	$N_2 \xleftarrow{K_{U1}} I_2 \xleftarrow{K_{U2}} 2U$	Scheme III
$K_U = \frac{[U]^2}{[N_2]}$	2.1	$K_{U1} = \frac{[I]^2}{[N_2]}, K_{U2} = \frac{[U]}{[I]}$	3.1a,b	$K_{U1} = \frac{[I_2]}{[N_2]}, K_{U2} = \frac{[U]^2}{[I_2]}$	4.1a,b
$f_{N_2} + f_U = 1$	2.2	$f_{N_2} + f_I + f_U = 1$	3.2	$f_{N_2} + f_{I_2} + f_U = 1$ [a]	4.2
$f_{N_2} = \frac{[N_2]}{Pt}, f_U = \frac{[U]}{2Pt}$	2.3a,b	$f_{N_2} = \frac{[N_2]}{Pt}, f_I = \frac{[I]}{2Pt}, f_U = \frac{[U]}{2Pt}$	3.3a,b,c	$f_{N_2} = \frac{[N_2]}{Pt}, f_{I_2} = \frac{[I_2]}{Pt}, f_U = \frac{[U]}{2Pt}$	4.3a,b,c
$4Ptf_U^2 + K_U f_U - K_U = 0$	2.4	$4Ptf_I^2 + K_{U1} f_I (1 + K_{U2}) - K_{U1} = 0$	3.4	$\left(\frac{1 + K_{U1}}{K_{U1}}\right)^2 f_{I_2}^2 - f_{I_2}\left(\frac{2(1 + K_{U1})}{K_{U1}} + \frac{K_{U2}}{4Pt}\right) + 1 = 0$	4.4
$f_U = \frac{-K_U + \sqrt{K_U^2 + 16PtK_U}}{8Pt}$	2.5	$f_I = \frac{-K_{U1}(1 + K_{U2}) + \sqrt{K_{U1}^2(1 + K_{U2}) + 16PtK_{U1}}}{8Pt}$	3.5	$f_{I_2} = \frac{\frac{2(1 + K_{U1})}{K_{U1}} + \frac{K_{U2}}{4Pt} + \sqrt{\frac{(1 + K_{U1})K_{U2}}{K_{U1}4Pt} + \left(\frac{K_{U2}}{4Pt}\right)^2}}{2\left(\frac{1 + K_{U1}}{K_{U1}}\right)^2}$	4.5
$Y_{obs} = Y_{N_2} + f_U(Y_U - Y_{N_2})$	2.6	$Y_{obs} = Y_{N_2} + f_I(Y_I - Y_{N_2} + K_{U2}(Y_U - Y_{N_2}))$	3.6	$Y_{obs} = Y_U + f_{I_2}(Y_{I_2} - Y_U + \frac{1}{K_{U1}}(Y_{N_2} - Y_U))$	4.6
Trimer-2-state model		**Trimer-3-state monomer intermediate model**		**Trimer-3-state trimer intermediate model**	
$N_3 \xleftarrow{K_U} 3U$	Scheme IV	$N_3 \xleftarrow{K_{U1}} 3I \xleftarrow{K_{U2}} 3U$	Scheme V	$N_3 \xleftarrow{K_{U1}} I_3 \xleftarrow{K_{U2}} 3U$	Scheme VI
$K_U = \frac{[U]^3}{[N_3]}$	5.1	$K_{U1} = \frac{[I]^3}{[N_3]}, K_{U2} = \frac{[U]}{[I]}$	6.1a,b	$K_{U1} = \frac{[I_3]}{[N_3]}, K_{U2} = \frac{[U]^3}{[I_3]}$	7.1a,b
$f_{N_3} + f_U = 1$	5.2	$f_{N_3} + f_I + f_U = 1$	6.2	$f_{N_3} + f_{I_3} + f_U = 1$	7.2
$f_{N_3} = \frac{[N_3]}{Pt}, f_U = \frac{[U]}{3Pt}$	5.3a,b	$f_{N_3} = \frac{[N_3]}{Pt}, f_I = \frac{[I]}{3Pt}, f_U = \frac{[U]}{3Pt}$	6.3a,b,c	$f_{N_3} = \frac{[N_3]}{Pt}, f_{I_3} = \frac{[I_3]}{Pt}, f_U = \frac{[U]}{3Pt}$	7.3a,b,c
$27Pt^2 f_U^3 + K_U f_U - K_U = 0$	5.4	$27Pt^2 f_I^3 + K_{U1} f_I (1 + K_{U2}) - K_{U1} = 0$	6.4	$\left(\frac{1+K_{U1}}{K_{U1}}\right)^3 f_{I_3}^3 - 3f_{I_3}^2\left(\frac{1+K_{U1}}{K_{U1}}\right)^2 + f_{I_3}\left(\frac{3(1+K_{U1})}{K_{U1}} + \frac{K_{U2}}{27Pt^2}\right) - 1 = 0$	7.4
$f_U = \sqrt[3]{\frac{K_U}{54Pt^2} + \frac{K_U}{54Pt^2}\sqrt{1 + \frac{4K_U}{729Pt^2}}} + \sqrt[3]{\frac{K_U}{54Pt^2} - \frac{K_U}{54Pt^2}\sqrt{1 + \frac{4K_U}{729Pt^2}}}$	5.5[b]	$f_I = \sqrt[3]{\frac{K_{U1}}{54Pt^2} + \frac{K_{U1}}{54Pt^2}\sqrt{1 + \frac{4K_{U1}(1 + K_{U2})^3}{729Pt^2}}} + \sqrt[3]{\frac{K_{U1}}{54Pt^2} - \frac{K_{U1}}{54Pt^2}\sqrt{1 + \frac{4K_{U1}(1 + K_{U2})^3}{729Pt^2}}}$	6.5[b]	$f_{I_3} = \sqrt[3]{\frac{-q + \sqrt{q^2 + \frac{4p^3}{27}}}{2}} + \sqrt[3]{\frac{-q - \sqrt{q^2 + \frac{4p^3}{27}}}{2}} + \frac{K_{U1}}{(1 + K_{U1})}$	7.5[b,c]
$Y_{obs} = Y_{N_3} + f_U(Y_U - Y_{N_3})$	5.6	$Y_{obs} = Y_{N_3} + f_I(Y_I - Y_{N_3} + K_{U2}(Y_U - Y_{N_3}))$	6.6	$Y_{obs} = Y_U + f_{I_3}(Y_{I_3} - Y_U + \frac{1}{K_{U1}}(Y_{N_3} - Y_U))$	7.6

[a] N, I, and U refer to the native, intermediate and unfolded subunits, with the subscript indicating the number of associated subunits (e.g., N_2 is native dimer, etc). K_U is the equilibrium constant of unfolding for a dimer (trimer) two-state mechanism. For the dimer (trimer) 3-state with monomer intermediate mechanism: K_{U1} is the equilibrium constant for dissociation of N_2 (N_3) to 2 (3) I, and K_{U2} is the equilibrium constant of unfolding from I to U. For the dimer (trimer) 3-state with dimer (trimer) intermediate mechanism: K_{U1} is the equilibrium constant for N_2 (N_3) unfolding to I_2 (I_3), and K_{U2} is the equilibrium constant for dissociation and unfolding of I_2 (I_3) to 2 (3) U. For each state of the protein, i, $[i]$, f_i, and Y_i are the concentration, mol fraction and corresponding optical signal, respectively. Y_{obs} is the total observed signal. Pt is the total protein concentration expressed in dimer or trimer equivalents. [b] Analytical solution of the cubic Equations 5−7.4 were obtained using Cardano's methods. [c] $p = (K_{U2}K_{U1}^3)/[27Pt^2(1 + K_{U1})^3]$; $q = [K_{U1}/(1 + K_{U1})]^4(K_{U2}/27Pt^2)$.

and 12040 g mol^{-1}, respectively, for Tm0979 and 12058 M^{-1} cm^{-1} and 14720 g mol^{-1} for Mth1491. The extinction coefficients were determined using the method of Gill and von Hippel (*33*).

Size Exclusion Chromatography. SEC was performed on a Biocad Sprint system with a prepacked Superose 12 10/300 GL (GE Healthcare) analytical gel filtration column equilibrated in the corresponding buffer for each protein (see above). All sample injection volumes were 25 μL. For each buffer, highly linear calibration curves of elution time versus log M were obtained using as standards: bovine serum albumin (66 kDa), ovalbumin (43.0 kDa), myoglobin (16.9 kDa), and cytochrome c (12.6 kDa). The apparent M of the eluted protein was determined from comparison with the standard curve, and the concentration of the protein during elution was calculated from the A_{280} at half-height of the elution peak (see above). For Tm0979, the dimer dissociation constant was determined using the method of Manning and co-workers (*34*). Briefly, percent of protein as dimer was calculated as a function of protein concentrations according to

$$\%\ \text{dimer} = 100(2^{(V_m - V)/(V_m - V_d)} - 1) \quad (1)$$

where V_m, V_d, and V are the elution times for the monomer, dimer, and mixture of species, respectively. The value of the dissociation constant, K_d,[1] was then determined using a linear fit of log(% dimer/(0.04(100 − % dimer)^2)) versus log[dimer], when % dimer/(0.04(100 − % dimer)^2) = 1, [dimer] = K_d.

Equilibrium GdmCl Curve Analysis. Tm0979 (Mth1491) GdmCl curves were fit to dimer (trimer) two-state and three-state models with monomer, I, or dimer (trimer) I_2 (I_3) intermediate using MATLAB 2007b, according to the numerical expressions (eqs 2.1−7.6), given in Table 1.

First, the curves measured at different protein concentrations were scaled to align the native protein and the unfolded protein signals. The observed change in signal (Y_{obs}) upon protein unfolding is related to the fraction (f_i) and signal (Y_i) of each species, i, involved in the unfolding process as given by

$$Y_{obs} = \sum_i f_i Y_i \quad (8)$$

Y_i were taken to vary linearly with GdmCl concentration according to

$$Y_i = Y_{i_0} + s_{i_0}[\text{GdmCl}] \quad (9)$$

where Y_{i_0} is the signal of the species, i, in 0 M GdmCl, and s_{i_0} is the dependence of Y_i on GdmCl concentration. The

Table 2: Equilibrium Unfolding Parameters for Tm0979 Dimer and Mth1491 Trimer[a]

	dimer three-state with monomer intermediate $N_2 \leftrightarrow 2I \leftrightarrow 2U$													
	dimer two-state $N_2 \leftrightarrow 2U$		$N_2 \leftrightarrow 2I$		$I \leftrightarrow U$		$N_2 \leftrightarrow 2U$							
Tm0979 dimer	ΔG_U (kcal (mol of dimer)$^{-1}$)	m_U (kcal (mol of dimer)$^{-1}$ M^{-1})	ΔG_{U1} (kcal (mol of dimer)$^{-1}$)	m_{U1} (kcal (mol of dimer)$^{-1}$ M^{-1})	ΔG_{U2} (kcal (mol of monomer)$^{-1}$)	m_{U2} (kcal (mol of monomer)$^{-1}$ M^{-1})	ΔG_{Utotal} (kcal (mol of dimer)$^{-1}$)	m_{Utotal} (kcal (mol of dimer)$^{-1}$ M^{-1})	Y_i (%)	s_i				
Fl	n/a[b]	n/a	7.2[c]	0.6 ± 0.2	6.2 ± 0.2	2.0 ± 0.1	19.5 ± 0.2	4.7 ± 0.2	86	(4.32 ± 1.10) × 10^6				
CD	n/a	n/a	7.2[c]	0.7 ± 0.2	5.7 ± 0.8	1.8 ± 0.2	18.6 ± 0.2	4.3 ± 0.3	250	0				
kinetics	n/a	n/a	n/a	n/a	6.8[d] ± 2.3	2.4[d] ± 0.1	n/a	n/a	n/a	n/a				

	trimer three-state with trimer intermediate $N_3 \leftrightarrow I_3 \leftrightarrow 3U$													
	trimer two-state $N_3 \leftrightarrow 3U$		$N_3 \leftrightarrow I_3$		$I_3 \leftrightarrow 3U$		$N_3 \leftrightarrow 3U$							
Mth1491 trimer	ΔG_U (kcal (mol of trimer)$^{-1}$)	m_U (kcal (mol of trimer)$^{-1}$ M^{-1})	ΔG_{U1} (kcal (mol of trimer)$^{-1}$)	m_{U1} (kcal (mol of trimer)$^{-1}$ M^{-1})	ΔG_{U2} (kcal (mol of trimer)$^{-1}$)	m_{U2} (kcal (mol of trimer)$^{-1}$ M^{-1})	ΔG_{Utotal} (kcal (mol of trimer)$^{-1}$)	m_{Utotal} (kcal (mol of trimer)$^{-1}$ M^{-1})	Y_i (%)	s_i				
Fl	44.9 ± 0.7	13.3 ± 0.3	18.0 ± 2.1	7.5 ± 0.9	27.1 ± 2.3	6.0 ± 0.9	45.1 ± 3.1	13.5 ± 1.3	55	0				
CD	38.6 ± 1.9	10.5 ± 0.8	19.1 ± 4.4	7.9 ± 1.7	21.9 ± 5.2	3.7 ± 2.0	27.0 ± 5.1	11.6 ± 2.6	76[c]	0				
CD	44[c]	12.8 ± 0.05	16.1 ± 3.1	6.8 ± 1.4	27[c]	5.6 ± 0.1	43.1 ± 3.1	12.4 ± 3.4	76[c]	0				

[a] ΔG_U represents the Gibbs free energy change in the unfolding direction for the various steps in the models, which are taken to be linearly dependent upon denaturant concentration according to the corresponding m values, which are proportional to the changes in solvent-accessible surface area for the transitions. For the dimer (trimer) two-state models, the free energy of unfolding of N_2 (N_3) to U is given by $\Delta G_U = 2G_U - G_{N2}$ ($3G_U - G_{N3}$). For the dimer three-state with monomer intermediate model, $\Delta G_{U1} = 2G_I - G_{N2}$ represents the free energy of dissociation of N_2 to monomer intermediate, I, and $\Delta G_{U2} = G_U - G_I$ represents the free energy of unfolding of I to U, so that $\Delta G_{Utotal} = \Delta G_{U1} + 2\Delta G_{U2}$, and $m_{Utotal} = m_{U1} + 2m_{U2}$. For trimer three-state unfolding with trimeric intermediate, $\Delta G_{U1} = G_{I3} - G_{N3}$ represents the free energy of trimer intermediate (I_3) formation from N_3 and $\Delta G_{U2} = 3G_U - G_{I3}$ represents the free energy of unfolding of I_3 to U, so that $\Delta G_{Utotal} = \Delta G_{U1} + \Delta G_{U2}$, and $m_{Utotal} = m_{U1} + m_{U2}$. Y_i and s_i are the signal and the change in signal with denaturant concentration, respectively, of the intermediate (I or I_3). Y_i expressed as a percentage of the total amplitude of the unfolding transition. [b] n/a: not applicable. [c] Values were fixed as described in Materials and Methods. [d] Equilibrium constants obtained from kinetic constants using eqs 10, 16.1, and 16.2.

folded (native) slope and intercepts (s_{f_0} and Y_{f_0}, respectively) were determined by linear regression of the data at low GdmCl concentration, i.e., before the transition region. The unfolded (denatured) slope and baseline (s_{u_0} and Y_{u_0}, respectively) were determined by linear regression of the data at high GdmCl concentration, i.e., after the transition region. Expressions for f_i, defined as functions of equilibrium constants, K, in Table 1 were related to Gibbs free energy values by

$$\Delta G_U = -RT \ln K_U \tag{10}$$

where R is the gas constant and T is the absolute temperature in Kelvin. Each ΔG_U was taken to vary linearly with GdmCl concentration according to m_U:

$$\Delta G_U{}^{GdmCl} = \Delta G_U + m_U[\text{GdmCl}] \tag{11}$$

For the two-state models, the slope and baseline of each species were fixed; ΔG_U and m_U were shared and allowed to be varied. For the three-state models, the thermodynamic constants (ΔG_U and m_U) for the two steps were first shared and allowed to be varied. For Tm0979, for the dimer 3-state monomer intermediate model, ΔG_{U2} was then fixed to 7.2, corresponding to a K_d of 5×10^{-6} M obtained by SEC. For Mth1491, the fitted constants differ slightly when determined by fitting the fluorescence or the CD data. Therefore, for the trimer three-state intermediate model, ΔG_{U2} was determined by fitting the fluorescence renaturation. The CD curves were then also fit by fixing the monomer stability, ΔG_{U2}, to the value determined by fluorescence, which is better defined. Also, the CD signal of the trimer intermediate was very poorly determined when allowed to vary, and so its value was also fixed to various values between those of the native and unfolded states; the reported value (Table 2) corresponds to the lowest SSE value obtained.

Unfolding and Refolding Kinetics. For Tm0979 and Mth1491, unfolding rates were measured by manual mixing experiments, using the same fluorescence and CD set up as for equilibrium curves. Unfolding was performed by diluting buffered native protein stock solution 10-fold into different final concentrations of buffered GdmCl. Final GdmCl concentrations were determined by refractive index (*35*) after completion of kinetics. To minimize temperature artifacts, both protein and GdmCl solutions were preequilibrated at 25.0 °C prior to initiation of unfolding.

For Tm0979, refolding kinetics were measured by manual mixing fluorescence and CD and by stopped-flow fluorescence. The inherently very low sensitivity of stopped-flow CD coupled with the intrinsically very low CD signal for Tm0979 precluded measurement of refolding kinetics by stopped-flow CD. Stock unfolded Tm0979 was prepared by incubating the protein in 6 M buffered GdmCl for 15 min and then diluting the stock into different final concentrations of GdmCl. Manual mixing measurements were made using the same spectral parameters as for equilibrium measurements. Fluorescence stopped-flow measurements were performed using a SFM4/Q instrument (Molecular Kinetics, Pullman, WA) interfaced to the FL3-22, thermostated to 25 °C using a circulating water bath.

For Mth1491, refolding kinetics were measured by manual mixing and stopped-flow fluorescence and CD, since this protein has much higher intrinsic CD signal than Tm0979.

Stock unfolded Mth1491 was prepared by incubating the protein in 5.8−6 M buffered GdmCl for at least 1 h and then diluting the stock into different final concentrations of GdmCl. Manual mixing experiments were performed in analogous fashion to Tm0979. CD stopped-flow measurements were made using a SFM-20 apparatus (Molecular Kinetics, Pullman, WA) interfaced to a Jasco 815. All solutions for Mth1491 experiments were degassed prior to using and were kept under reducing conditions by adding DTT and EDTA to final concentrations of 10 and 1 mM, respectively.

Kinetic Data Analysis. Monomer folding and unfolding kinetics, as well as dimer and trimer unfolding kinetics, are generally first order (eq 12), while two-state dimer and trimer folding are second and third order (eqs 13.1 and 13.2), respectively:

$$Y_{obs} = Y_0 + (Y_\infty - Y_0)(1 - e^{-k_{app}t}) \quad (12)$$

$$Y_{obs} = Y_0 + (Y_\infty - Y_0)\left(\frac{2Pt k_f t}{1 + 2Pt k_f t}\right) \quad (13.1)$$

$$Y_{obs} = Y_0 + (Y_\infty - Y_0)\left(1 - \frac{1}{\sqrt{1 + 18Pt^2 k_f t}}\right) \quad (13.2)$$

where t is time, Y_{obs} is the observed signal, Y_0 and Y_∞ are the signals at $t = 0$ and $t =$ infinity, respectively, k_{app} is the apparent folding rate constant, and $k_{app} = k_u + k_f$, where k_u is the unfolding rate constant and k_f is the folding rate constant. When obtained by fitting to the correct model, k_u and k_f are generally linearly proportional to the concentration of GdmCl:

$$\ln k_u = \ln k_{uH_2O} + m^{\ddagger}_u[\text{GdmCl}]/RT \quad (14.1)$$

$$\ln k_f = \ln k_{fH_2O} - m^{\ddagger}_f[\text{GdmCl}]/RT \quad (14.2)$$

where $m^{\ddagger}_u$ and $m^{\ddagger}_f$ are the dependencies of the unfolding and folding rate constants in water (k_{uH_2O} and k_{fH_2O}), respectively, on GdmCl concentration.

For a two-state monomer transition, the natural logarithm of the rate constants can be plotted as a Chevron plot according to

$$\ln(k_{obs}) = \ln\left(k_{fH_2O}\exp\left(-\frac{m_f[\text{GdmCl}]}{RT}\right) + k_{uH_2O}\exp\left(\frac{m_u[\text{GdmCl}]}{RT}\right)\right) \quad (15)$$

Furthermore, equilibrium constants can be calculated from the kinetic constants using the equations:

$$K_U = \frac{k_u}{k_f} \quad (16.1)$$

$$m_U = m^{\ddagger}_u + m^{\ddagger}_f \quad (16.2)$$

Tm0979 and Mth1491 refolding kinetic traces were fit to equations corresponding to a first-, second-, or third-order transition using eqs 12, 13.1, and 13.2, respectively, while unfolding kinetics were fit to a first-order transition using eq 12. First-order kinetics were fit using Biokine 32 (Biologic), whereas second- and third-order kinetics were fit using OriginPro7.5 (OriginLab).

RESULTS

Size Analysis. The oligomerization of Tm0979 and Mth1491 was investigated by SEC as a function of protein concentration (Figure 2). For Tm0979, a single elution peak was observed at all protein concentrations, with the elution time at higher protein concentrations approaching that expected for the dimeric form of the protein and shifting toward that expected for the monomeric form at lower protein concentrations (Figure 2A). The peak shapes are indicative of rapid equilibrium between the dimeric and monomeric forms of the protein, with the elution time being determined by the relative proportions of the two species (*34*).This allows the SEC data to be fit to a dimer dissociation model, which gives a K_d value of $(5 \pm 2) \times 10^{-6}$ M as the average of two independent determinations (Figure 2A). Analysis of the monomer−dimer equilibrium using dynamic light scattering as a function of protein concentration gave a similar value for K_d. In contrast, for Mth1491, although a single peak was observed at all protein concentrations (Figure 2B), the elution times did not vary and were consistent with the protein remaining trimeric. An upper limit of $K_d < 4 \times 10^{-19}$ M^2 can be estimated, assuming less than 10% dissociation of the trimer to monomers at the lowest protein concentration (3 nM trimer).

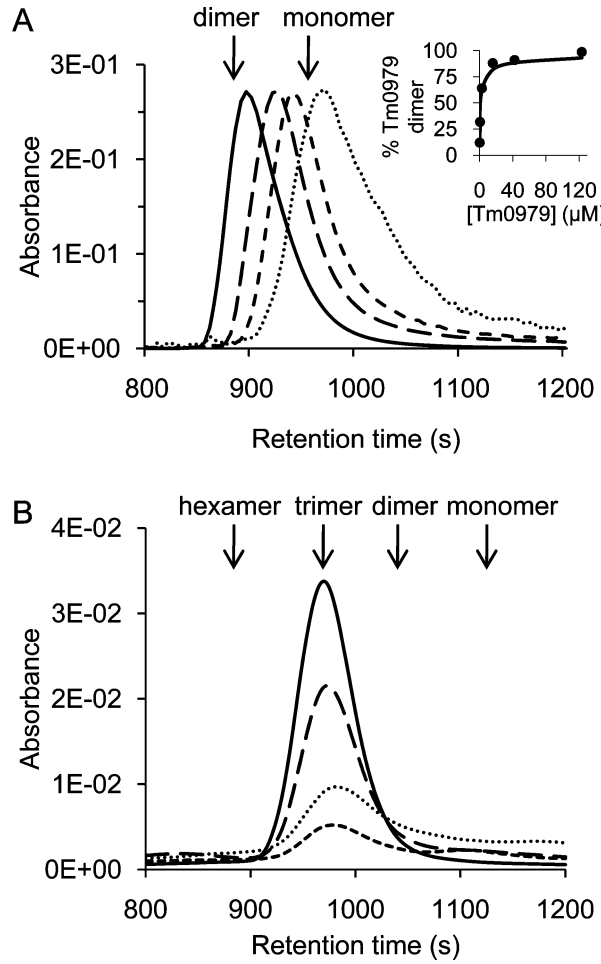

FIGURE 2: Determination of oligomeric state for Tm0979 and Mth1491 by size exclusion chromatography. (A) Elution profile for Tm0979 at total dimer concentrations of 21.1 μM (solid), 1.46 μM (long dash), 0.55 μM (short dash), and 0.11 μM (dotted). Expected retention times for Tm0979 dimer and monomer are indicated by arrows. Inset: % dimer versus total dimer concentration (eq 1). The line shows that the data are well described by a K_d value of 5 μM, determined as described in Materials and Methods. (B) Elution profile of Mth1491 measured at total trimer concentrations of 1.4 μM (solid), 0.12 μM (long dash), 9 nM (dotted dash), and 8 nM (short dash).

Calorimetry. Since both Tm0979 and Mth1491 are from thermophilic organisms, it was expected that they would be relatively stable toward thermal denaturation. DSC measurements showed that both proteins remain folded up to relatively high temperatures, but Tm0979 thermal unfolding is highly reversible, whereas Mth1491 unfolding is highly irreversible (Figure S1 of the Supporting Information). Various solution conditions were tested for Mth1491; however, at increased temperature (~76 °C) the data were characterized by an exotherm due to protein aggregation. Although the irreversible aggregation prohibits thermodynamic analysis of the data, it is likely that Mth1491 is sufficiently thermostable to remain folded under biological conditions, since the optimal growth temperature of *M. thermoautotrophicum* is 65 °C (*36*). In contrast, Tm0979 exhibits an endotherm at ~85 °C, with minimal aggregation. Thus, this protein also remains folded *in vivo*, as the optimal growth temperature of *T. maritima* is 80 °C (*37*).

DSC data for Tm0979 were also obtained as a function of protein concentration, and the apparent melting temperature did not vary, suggesting dimer dissociation largely occurs below this temperature. The data were not well fit by either a monomer two-state or a dimer two-state model (Table S1 of Supporting Information), indicating more complex behavior, such as monomer formation. ITC experiments revealed that dimer association is weakened as temperature is increased but is still significant at 75 °C (Figure S1C,D of the Supporting Information). Fits of the DSC data to a dimer three-state model with monomer intermediate corresponded well to the data but did not give well-defined thermodynamic parameters, particularly for the dimer dissociation step. Since the van't Hoff to calorimetric enthalpy ratios for the dimer two-state fits are somewhat higher than 1, aggregation may also be occurring to some extent. Thus, the fitted values may provide an estimate of thermodynamic properties but must be considered with caution.

Reversibility and Equilibrium of GdmCl Denaturation. The folding energetics and mechanisms for Tm0979 and Mth1491 were investigated further using chemical denaturation by GdmCl monitored by fluorescence and CD. A prerequisite for quantitative thermodynamic analysis is that chemically induced unfolding be reversible and at equilibrium. GdmCl denaturation of both proteins is highly reversible; however, the time to reach equilibrium is much shorter for Tm0979 (within 1 day for denaturation and renaturation) than for Mth1491 (more than 23 days for denaturation and 4 days for renaturation). Due to the shorter time required to reach equilibrium, further curves for Mth1491 were obtained by renaturation.

Equilibrium GdmCl Curve Analysis. (A) Tm0979 Populates a Monomer Intermediate at Equilibrium. Tm0979 GdmCl curves (Figure 3) were fit to various dimer folding models (Table 1, Schemes I−III) by global fitting of the curves obtained at protein concentrations ranging from 0.5 to 25 μM dimer. The two-state model (Table 1, Scheme I) could not account for the data, since the fitted lines for different protein concentrations were much more spread out than the experimental curves (not shown). Similarly, the dimer three-state model with dimeric intermediate (Table 1, Scheme III) gave very poor fits. The very minimal change in the curves with Tm0979 concentration (Figure 3) suggests

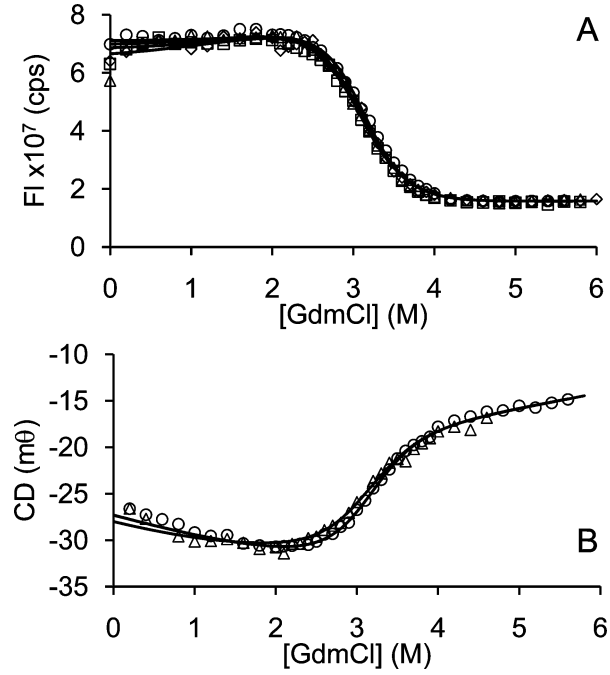

FIGURE 3: Tm0979 denaturation curves as a function of protein concentration fit to a dimer three-state model with monomer intermediate. (A) Tm0979 denaturation curves measured by fluorescence at dimer concentrations of 0.5 μM (◇), 2.5 μM (□), 5 μM (△), and 25 μM (○). (B) Tm0979 denaturation curves measured by circular dichroism at dimer concentrations of 5 μM (△) and 25 μM (○). The fitted traces correspond to the fitted constants given in Table 2, with corresponding equations given in Table 1.

unfolding of a monomeric species, as might be expected to become populated in GdmCl based on the moderate affinity K_d value of 5×10^{-6} M obtained for Tm0979 by SEC (Figure 2A). Initial fits to a dimer three-state model with monomer intermediate (Table 1, Scheme II) that allowed the thermodynamics constants of both steps to all vary did not converge. However, fixing the free energy for dimer dissociation, ΔG_{U1}, to 7.2 kcal mol^{-1}, corresponding to the K_d obtained by SEC (eq 10), and allowing the other parameters to vary resulted in good fits (Figure 3, Table 2). The fitted values obtained by fluorescence and CD are in agreement within error, providing further support for the applicability of the three-state monomer intermediate mechanism. The total free energy for unfolding of native Tm0979 dimer to unfolded monomers, ΔG_{Utotal}, is 19.0 kcal mol^{-1}, indicating that Tm0979 is relatively stable compared to other dimers (*6*). Relatively high stability is commonly observed for thermophilic proteins (*38*−*40*). Also, the Tm0979 monomer intermediate has significant stability, ΔG_{U2} of ~6 kcal mol^{-1}, comparable to various other naturally monomeric proteins (*5*). It should be noted, however, that at low concentrations favoring monomer formation Tm0979 tends to precipitate, probably as a consequence of exposure of the interface region.

(B) Mth1491 Populates Low Levels of Trimer Intermediate at Equilibrium. Mth1491 GdmCl curves monitored by fluorescence and CD, at protein concentrations ranging from 0.1 to 3 μM trimer, were globally fit to trimer folding models (Table 1, Schemes IV–VI). As shown in Figure 4, the trimer two-state model fits the data quite well. Also, the fitted values obtained for the fluorescence and CD data (Table 2) are in reasonable agreement. The CD values appear to be slightly

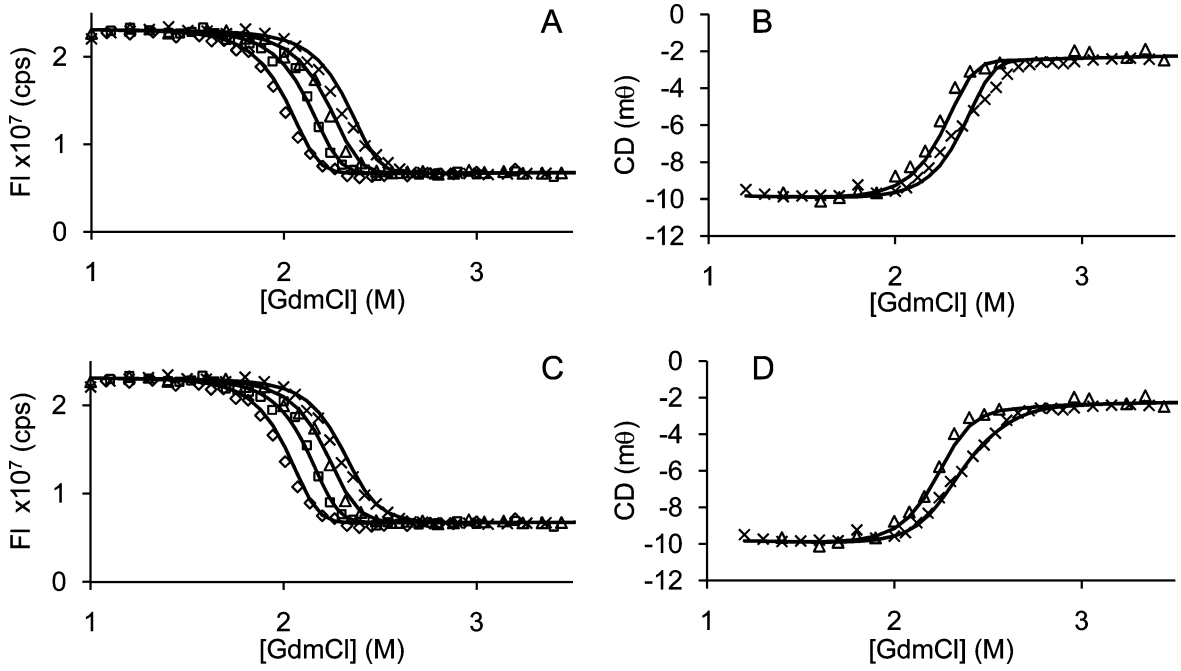

FIGURE 4: Mth1491 equilibrium renaturation. Renaturation curves measured by fluorescence (A, C) and circular dichroism (B, D) fit to trimer two-state model (A, B) or three-state model with trimer intermediate (C, D). Total trimer concentrations are 0.1 μM (◇), 0.3 μM (□), 1 μM (△), and 3 μM (×). Lines correspond to best fit values given in Table 2, with corresponding equations given in Table 1.

lower, however, and there is a small systematic deviation between the experimental data and the lines of best fit for GdmCl concentrations higher than 2.2 M, where the experimental data have a higher fluorescence signal than predicted by the fit and this is more pronounced at increased protein concentration. This behavior suggests increased formation of a trimer intermediate at increased protein concentration; analogous behavior has been reported for native protein dimers that form dimer intermediates (*41*−*44*). The Mth1491 curves were therefore also fit to trimer three-state models with trimer or monomer intermediate (Table 1, Schemes VI and V, respectively). The fit for the trimer three-state model with monomer intermediate was comparable to the trimer two-state fit for the last part of the transition; i.e., there was still systematic deviation between the data and the fits (not shown); however, the trimer three-state with trimer intermediate model accounted for the data very well throughout the transitions for all protein concentrations (Figure 4C,D).

From these analyses, the values obtained from the fluorescence and CD data agree within experimental error, supporting the applicability of three-state trimer intermediate model. The average of the fluorescence and CD values for trimer dissociation and monomer unfolding, ΔG_{U1} and ΔG_{U2}, are 17 and 27 kcal (mol of trimer)$^{-1}$, respectively, corresponding to an overall Gibbs free energy of unfolding, ΔG_{Utotal}, of 44 kcal (mol of trimer)$^{-1}$. This value is very similar to the ΔG_U value of 45 kcal (mol of trimer)$^{-1}$ obtained from the fluorescence two-state fits, which is reasonable given that the population of the trimer intermediate is very low in the fluorescence curves except at the highest protein concentration of 3 μM trimer, where its population reaches a maximum level of only ~14% (Figure 8; see Discussion). The difference is higher for the two-state CD values, likely due to the intermediate being significantly populated at the higher protein concentrations required for the CD measurements; increased population of intermediate would cause the transition to be broadened resulting in decreased m_U and ΔG_U values. The fluorescence and CD average values of m_{U1} and m_{U2} are 7.2 and 5.8 kcal (mol of trimer)$^{-1}$ M^{-1}, respectively, corresponding to an overall m_{Utotal} of 13.0 kcal (mol of trimer)$^{-1}$ M^{-1}, which is again very close to that obtained for the two-state fit of the fluorescence data. These fitted m values are comparable to the expected value calculated using an empirical equation derived from observed m values for many other proteins (*45*), 11.5 kcal (mol of trimer)$^{-1}$ M^{-1}. The ΔG_{Utotal} value of 44 kcal (mol of trimer)$^{-1}$ is relatively high, corresponding to a very low K_d (using $\Delta G_{Utotal} = -RT \ln K_d$) of 5×10^{-33} M^2, for which protein is half-dissociated at 6 pM total trimer concentration This K_d is very low compared to the K_d deduced from the ΔG_U of other trimeric proteins (Table 3), probably owing to the particularly high stability of thermophilic proteins (see Discussion).

Kinetics of Unfolding and Folding. In order to investigate the relationship between kinetic and equilibrium folding mechanisms, the kinetics of unfolding and folding for Tm0979 and Mth1491 were measured by fluorescence and CD at different denaturant and protein concentrations.

(A) Tm0979 Kinetics Are Dominated by Rate-Limiting Monomer Transitions. Tm0979 unfolding kinetics by fluorescence and CD are well fit by a single exponential equation, as is generally observed for protein unfolding reactions under highly denaturing conditions (Figure 5A,D). The amplitudes for the unfolding kinetics correspond to the amplitudes expected based on equilibrium curve measurements which reflect unfolding of monomer intermediate (Figure 5C,F). The unfolding rates measured by fluorescence and CD are also very similar (Figure 7A). This suggests that the observed kinetics are dominated by the unfolding of the monomer intermediate, with rapid dimer dissociation occurring in the dead time. Similar behavior has been reported for other dimers (*46*−*48*). The natural logarithm of the k_u varies linearly with GdmCl concentration (eq 14.1, Figure 7A). The magnitude of $m^{\ddagger}_u$ is smaller than expected for a dimer of

Table 3: Thermodynamic and Kinetic Parameters Characteristic of Chemically Induced Unfolding of Trimeric Proteins

			interface characteristics[b]											
protein	PDB code	chain length	interface area per monomer	% buried	% polar	equilibrium folding mechanism	denaturant	ΔG_U (kcal mol^{-1})	m_U (kcal (mol of trimer)$^{-1}$ M^{-1})	kinetic mechanism[h]	k_{uH_2O} (s^{-1})	$m^{\ddagger}_u$ (kcal (mol of trimer)$^{-1}$ M^{-1})	k_{fH_2O} (M^{-2} s^{-1})	$m^{\ddagger}_f$ (kcal (mol of trimer)$^{-1}$ M^{-1})
α-Proteins														
MHC class II TM domain (*11*)	nd[a]	25	nd	nd	nd	3U ↔ N_3		9.4[c]	nd	nd	nd	nd	nd	nd
three-stranded coiled coil (*12*)	nd	nd	nd	nd	nd	2U ↔ I_2	urea	8.7[d]	nd	three-state	1.5×10^{-2}	nd	7.8×10^{4}	nd
						$I_2 + I \leftrightarrow N_3$		7.9[d]			1.1		6.5×10^{5}	
						3U ↔ N_3[c]		**16.6**						
designed three-stranded coiled coil (*10*)	1COS	29	1117	37	12	2U ↔ I_2	GdmCl			nd	nd	nd	nd	nd
						$I_2 + I \leftrightarrow N_3$								
						3U ↔ N_3		**18.4**	**2.7**					
HIV gp41 (*16*) N34(L6)C28	1AIK	34	2096	37	9	3U ↔ N_3	urea	27.7	2.4	two-state	1.1×10^{-5}	0.5	1.3×10^{15}	2
SIV gp41 (*15, 16*) N34(L6)C28	2SIV	36	6421	59	18	3U ↔ N_3	urea	18.9	3.6	two-state	5.7×10^{-4}	1.4	1.1×10^{10}	2.1
Lpp-56 (*9*)	1KFM	56	1586	35		3U ↔ N_3	GdmCl	18.9	3.6	two-state	3.8×10^{-10}	2.7	6.8×10^{5}	0.5
α/β-Proteins														
bacteriophage T4 fibritin (*13*)	1RFO	27	1091	44	22	3U ↔ N_3	GdmCl	21.3	2.5	complex[e]				
Mth1491	1L1S	111	1473	25	30	3U ↔ I_3	GdmCl	27	5.8	nd	nd	nd	nd	nd
						$I_3 \leftrightarrow N_3$		17	7.2					
						3U ↔ N_3		**44**	**13**					
adenylate kinase (*7*)	1NKS	194	1706	17	34	3U ↔ N_3	GdmCl	29.5	5.5	nd	nd	nd	nd	nd
apo γ-carbonic anhydrase (*17*)	1QRG	206	2631	28	4	U ↔ I	urea	4	nd	nd	nd	nd	nd	nd
						3I ↔ N_3		13.2						
						3U ↔ N_3		**25.2**						
holo γ-carbonic anhydrase (*17*)						3U ↔ N_3	urea	70						
bacteriophage T4 gp45e (*72*)	1B77	228	890	8	24	3U ↔ N_3	n/a	9.1[c]	nd	nd	nd	nd	nd	nd
Hsc70 (*73*)	2QW9	384	nd	nd	nd	2U ↔ I_2	n/a	6.3[f]	nd	nd	nd	nd	nd	nd
						$I_2 + I \leftrightarrow N_3$		6.2[f]						
						3 ↔ N_3		**12.6**						
β-Proteins														
bovine factor B (*8*)	nd	nd	nd	nd	nd	3U ↔ N_3	n/a	**13.1**[g]	nd	nd	nd	nd	nd	nd

[a] nd: not determined. [b] Calculation for interface characteristics made using Getarea (*66*) (http://pauli.utmb.edu/cgi-bin/get_a_form.tcl). Interface area per monomer is calculated as (3(surface area of monomer) − (surface area of trimer))/3, % buried is the percent of surface area buried within the interface per monomer and is calculated as (interface area per monomer) × 100/(surface area of monomer), % polar is the percentage of polar residues in the interface and is calculated as (3(polar surface area of monomer) − (polar surface area of trimer)) × 100/(3(interface area per monomer)). [c] Text in bold typeface refers to the sum of the steps in the mechanism. [d] Determined from unfolding and refolding rates using eqs 10 and 16.1. [e] Kinetic mechanism involves monomer, dimer, and trimer intermediate. [f] Calculated using eq 10 and K_d determined from sedimentation equilibrium data. [g] ΔG_U calculated from K_d using eq 10. [h] The kinetic mechanism is shown in the unfolding direction.

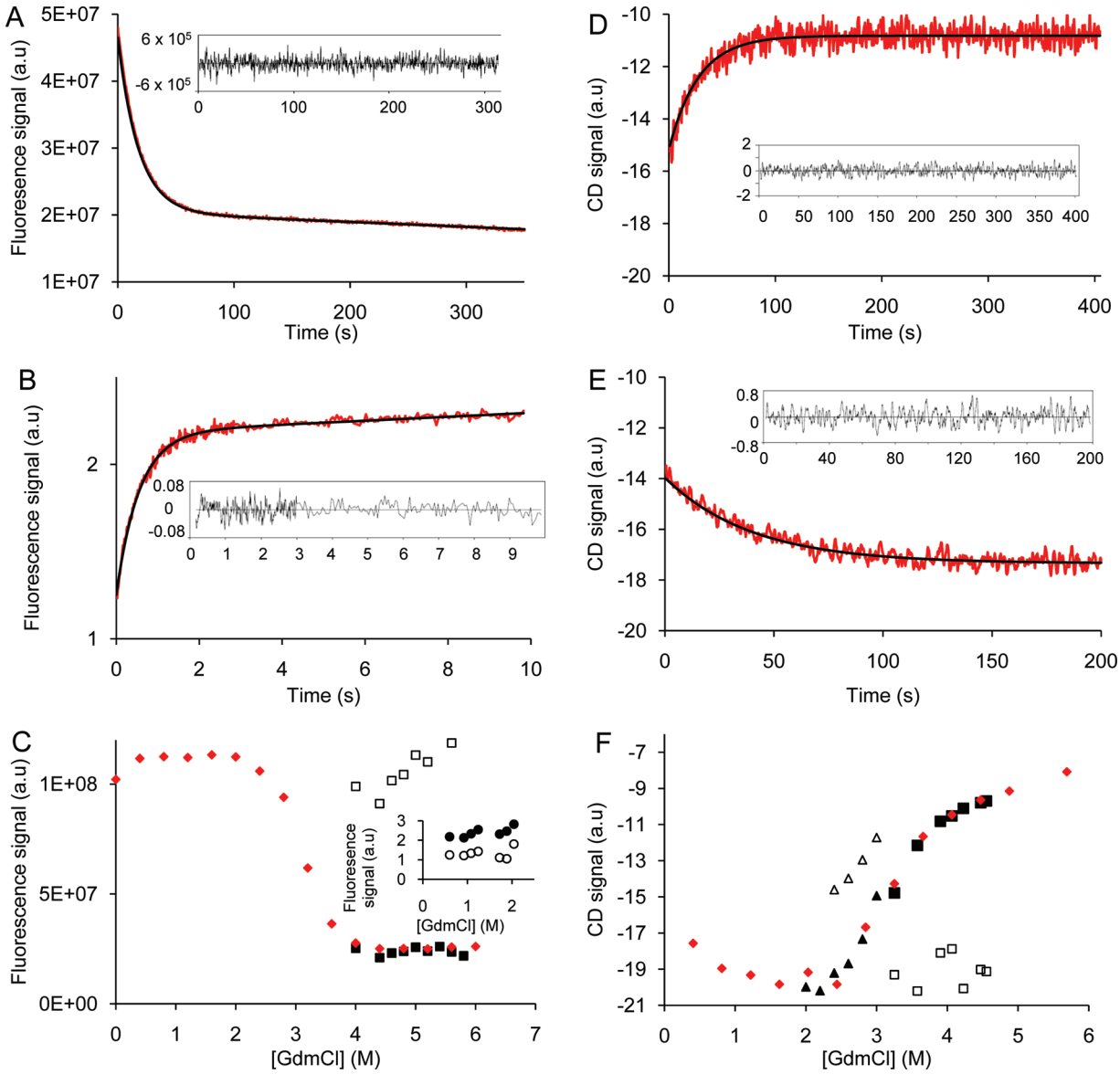

FIGURE 5: Unfolding and folding kinetics of Tm0979 monitored by fluorescence (A−C) and CD (D−F). Manual mixing unfolding kinetics monitored by (A) fluorescence in 4.6 M GdmCl, $k_u = 5.5 \times 10^{-2}$ s^{-1}, and (D) CD in 3.9 M GdmCl, $k_u = 3.6 \times 10^{-2}$ s^{-1}. Refolding kinetics monitored by (B) stopped-flow fluorescence in 0.60 M GdmCl, $k_f = 1.8$ s^{-1}, and (E) manual mixing CD in 2.8 M GdmCl, $k_f = 2.5 \times 10^{-2}$ s^{-1}. Insets in panels A, B, D, and E show random residuals for fitting data to single exponential functions with a drift (fluorescence) or without drift (CD). The protein concentrations for fluorescence and CD experiments are 1.5 and 25 μM (dimer equivalents), respectively. (C) and (F) show the kinetic data relative to equilibrium curve fluorescence and CD signals (◆), respectively. Unfolding kinetic initial (□) and final (■) signals and refolding initial (△) and final (▲) signals are shown in the main panels, and stopped-flow initial (○) and final (●) signals are shown as insets.

this size and close to that expected for a single subunit of Tm0979 (*5*), supporting the interpretation that Tm0979 monomer unfolding is the observed step. Further evidence for this comes from comparison of kinetic data with equilibrium stabilities, described below.

Tm0979 folding kinetics measured by fluorescence and CD at different protein concentrations are well fit by a single exponential equation (Figure 5B,E) and have very similar rate constants (Figure 7A); these observations suggest that the observed transition corresponds to monomer refolding. Systematic deviations were observed upon fitting the data to a second-order equation for simultaneous folding and association of monomers to form the native dimer. Moreover, if dimer association is the observed step, the kinetics should occur faster at higher protein concentration. Thus, the observed kinetics are inconsistent with a second-order folding process and appear to represent monomer folding.

Since both folding and unfolding kinetics of Tm0979 have characteristics of monomer transitions, the unfolding and refolding rate constants were fit to a monomer kinetic model using eq 15. The values obtained for the fitted constants are as follows: k_{uH_2O} of 6.5×10^{-4} s^{-1}; $m^{\ddagger}_u$ of 0.61 kcal (mol of monomer)$^{-1}$ M^{-1}; k_{fH_2O} of 60 s^{-1}; and $m^{\ddagger}_f$ of 1.8 kcal (mol of monomer)$^{-1}$ M^{-1}. These values were used in eqs 10, 16.1, and 16.2 to calculate equilibrium values, $\Delta G_U = 6.8$ kcal (mol of monomer)$^{-1}$ and $m_U = 2.4$ kcal (mol of monomer)$^{-1}$ M^{-1}, which are very similar to the values of ΔG_{U2} and m_{U2} obtained for monomer unfolding from equilibrium curve measurements (Table 2). Together, the above observations provide strong evidence that the kinetic measurements for

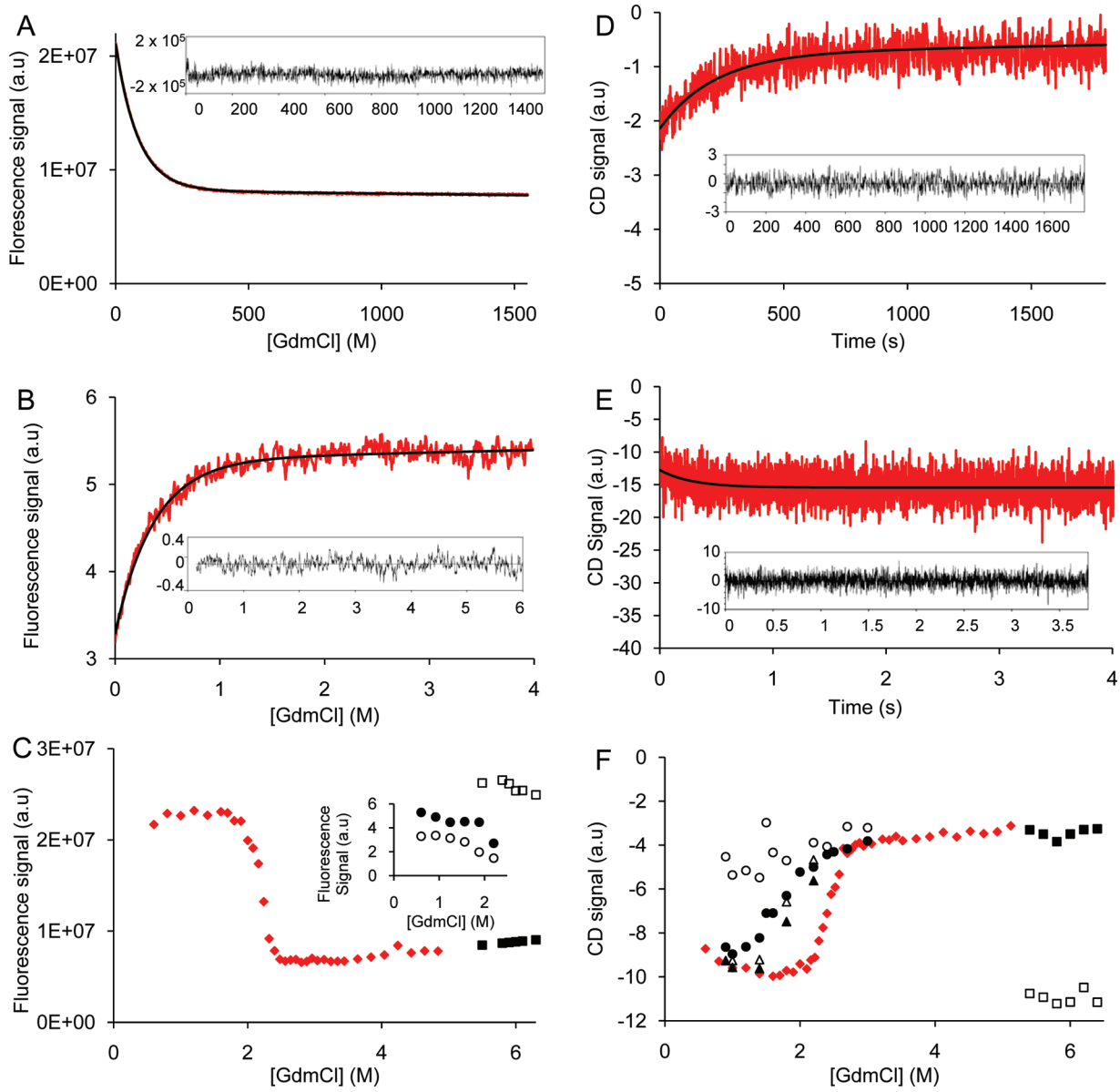

FIGURE 6: Unfolding and folding kinetic traces for Mth1491 monitored by fluorescence (A−C) and CD (D−F). Manual mixing unfolding kinetics monitored by (A) fluorescence in 6.1 M GdmCl, $k_u = 1.2 \times 10^{-2}$ s^{-1}, and (D) CD in 5.8 M GdmCl, $k_u = 4.5 \times 10^{-3}$ s^{-1}. Stopped-flow refolding kinetics monitored by (B) fluorescence in 0.60 M GdmCl, $k_f = 2.7$ s^{-1}, and (E) CD in 0.9 M GdmCl, $k_f = 3.5$ s^{-1}. Insets in panels A, B, D, and E show random residuals for fitting data to single exponential functions. The protein concentration for fluorescence and CD experiments is 1.0 and 3.0 μM, respectively (trimer equivalents). (C) and (F) show the kinetic data relative to equilibrium curve fluorescence and CD signals (◆), respectively. Manual mixing unfolding kinetic initial (□) and final (■) signals and manual mixing refolding initial (△) and final (▲) signals are shown in the main panels, and stopped-flow initial (○) and final (●) signals are shown as inset in (C) and in main panel in (F). The stopped-flow CD signals could be shown in this way because the signal was constant with time, whereas the stopped-flow fluorescence offset drifted over the course of the experiment.

Tm0979 correspond to rate-limiting folding and unfolding of the monomer intermediate that is also observed in equilibrium measurements.

(B) Mth1491 Kinetics Are Complex. Mth1491 unfolding kinetics monitored by fluorescence and CD are also well fit by a single exponential equation (Figure 6A,D), and the amplitudes are as expected based on equilibrium measurements for unfolding of native trimer to unfolded monomers (Figure 6C,F). The unfolding rate constants vary linearly with GdmCl concentration (Figure 7B); a linear fit of the data gives k_{uH_2O} of 6.5×10^{-13} s^{-1} and $m^{\ddagger}_u$ of 2.3 kcal (mol of trimer)$^{-1}$ M^{-1}. Comparing with values obtained for other proteins (*5*), this value of $m^{\ddagger}_u$ is considerably larger than would be expected for a single subunit of Mth1491. These observations indicate that the observed unfolding kinetics for Mth1491 correspond to the native trimer unfolding.

Mth1491 refolding kinetics are more complex and clearly multiphasic (Figures 6 and 7). The multiphasic nature is evident from the time dependence of renaturation curves, which require days to reach equilibrium in the transition region, where there is an initial change in optical signal followed by additional changes that occur over hours to days. The initial changes were monitored by stopped-flow fluorescence and CD (Figures 6B,C,E,F and 7B). The stopped-flow data are well fit by a single exponential equation (Figure 6B,E), the observed rates are very similar for both optical probes (Figure 7B), and the fitted apparent rate constants vary linearly with GdmCl

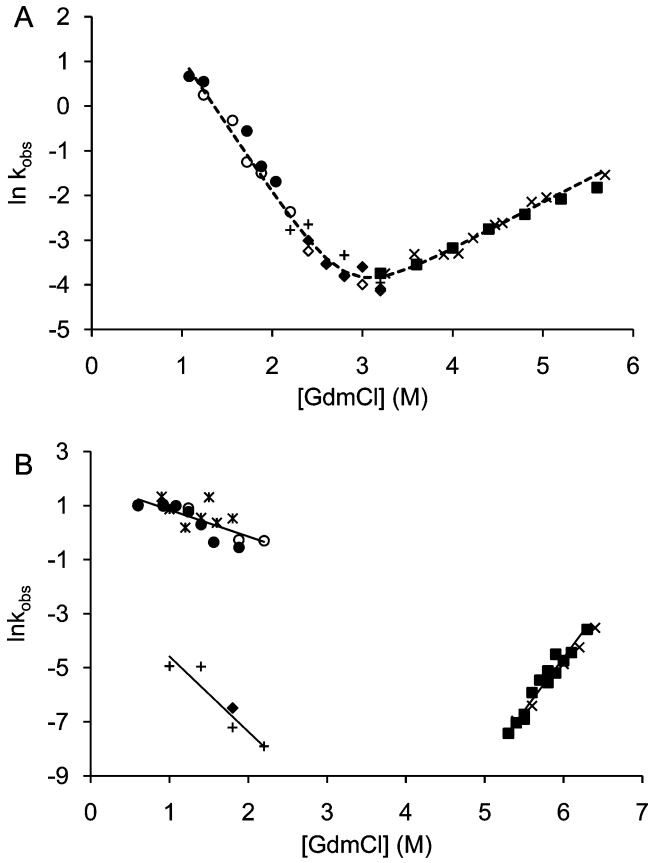

FIGURE 7: GdmCl dependence of the refolding and unfolding rate constants for Tm0979 and Mth1491. Rate constants were obtained by fluorescence/CD manual mixing for unfolding (■/×), manual mixing for refolding (◆◇/+), and stopped flow for refolding (●○/*). (A) Tm0979 unfolding rates were measured at 1.5 μM (■) and refolding rates were measured at 1.5 μM (◆●) and 3 μM (◇○) protein (dimer equivalents). (B) Mth1491 unfolding rates were measured at 1 μM (■) and 3 μM (×), and refolding rates were measured at 1 μM (●), 2 μM (○), and 3 μM (+*◆) protein (trimer equivalents). Values of k_{obs} are for fitting to single exponential equations. The dashed line in panel A corresponds to fitting the data to a two-state monomer unfolding transition, as described in Results and Materials and Methods. The fitted values are as follows: k_{fH_2O}, 60 s^{-1}; $m^{\ddagger}_f$, 1.8 kcal (mol of monomer)$^{-1}$ M^{-1}; k_{uH_2O}, 6.5 × 10^{-4} s^{-1}; $m^{\ddagger}_u$, 0.61 kcal (mol of monomer)$^{-1}$ M^{-1}. The solid lines in panel B were obtained by fitting the 1.5 μM data to a straight line using linear regression, which gives the following: k_{fH_2O}, 6.2 s^{-1}; $m^{\ddagger}_f$, 1.1 kcal (mol of monomer)$^{-1}$ M^{-1}; k_{uH_2O}, 6.5 × 10^{-13} s^{-1}; $m^{\ddagger}_u$, 2.3 kcal (mol of monomer)$^{-1}$ M^{-1}.

concentration and do not increase with increasing protein concentration (Figure 7B). These results are contrary to the expected behavior for a third- or second-order association reaction and strongly suggest that the observed kinetics correspond to a process that does not involve a change in molecularity, i.e., monomer folding (U → I). Comparing with monomeric proteins (*5*), the $m^{\ddagger}_f$ for Mth1491 is quite close to what may be expected for a protein the size of the Mth1491 monomer.

At least two additional slower kinetic phases are apparent from manual mixing refolding experiments and comparison with equilibrium curves (Figures 6F and 7B). An intermediate time scale phase is observed by both CD and fluorescence, and this phase has a much larger denaturant dependence than the fast phase (Figure 7B). This suggests larger scale conformational changes associated with larger species. Note that at low denaturant the full signal is regained very quickly by both CD and fluorescence, suggesting a switch from rate-limiting monomer folding at low denaturant to processes involving higher order species at high denaturant. It is not possible at this time to interpret the kinetic phases for Mth1491 in detail, as this would require considerable additional study (such as double jump experiments). However, we note that complex folding behavior has been reported for a number of other trimeric proteins (see Discussion).

DISCUSSION

Observed Folding Mechanisms for Tm0979 and Mth1491. Protein superfamilies often include members with different quaternary structure (*4, 22*), and as sequence similarity decreases among proteins that adopt the same fold, different subunit associations are commonly observed (*49*). The dsr family is interesting because it includes many different quaternary structures: monomeric, homodimeric, homotrimeric, and homo- and heterohexameric. Changes in quaternary structure may occur due to happenstance or be associated with the evolution of new functions and regulation of function (discussed further below).

In this study, we have characterized the equilibrium and kinetic folding properties of two members of the DsrEFH-like superfamily, homodimeric Tm0979 and homotrimeric Mth1491. These proteins have similar subunit folds (Figure 1) but very different sequences (17% identity), quaternary structures (Figure 1), and folding characteristics. Previous NMR and preliminary SEC and light scattering data had indicated that Tm0979 can form a micromolar affinity dimer (*23*) or a monomer (*24*) in solution. In order to investigate this further, we analyzed the characteristics of the dimer solution structure by interface analysis programs, DiMoVo (*50*), PISA (*51*), and NOXclass (*52*). All three programs confirm, with high confidence limits, that the dimeric structure is physically reasonable, with interface characteristics that are typical of and consistent with the observed affinity. The monomeric Tm0979 structure was determined using a slightly different construct, in a different buffer, and the secondary structural elements were less well defined than for the dimer structure; no additional biochemical data were reported (*24*). Since we also find that Tm0979 can populate a monomer at decreased protein concentration, the difference in the results is not very pronounced and may reasonably arise from the somewhat different experimental conditions. For Mth1491, previous SEC data also found evidence for trimer formation, but no other details were reported (*28*). To our knowledge, the only other reported folding studies on proteins with the same fold that form both dimers and trimers are for designed coiled-coil systems (*9, 10, 12, 53*). The coiled-coil proteins have simple helical interfaces; hence no comparable natural proteins with differing quaternary structures and complex interfaces have been characterized.

(A) Distinct Equilibrium Folding Mechanisms. A central difference in the equilibrium folding mechanisms of Tm0979 and Mth1491 is the population of a monomeric state, which occurs readily for Tm0979 but not for Mth1491. All the experimental data for Tm0979 (SEC, DSC, GdmCl curves, Figures 2A and 3 and Figure S1 of the Supporting Information) support a three-state folding mechanism with prominent formation of monomer intermediate. In contrast, no monomer

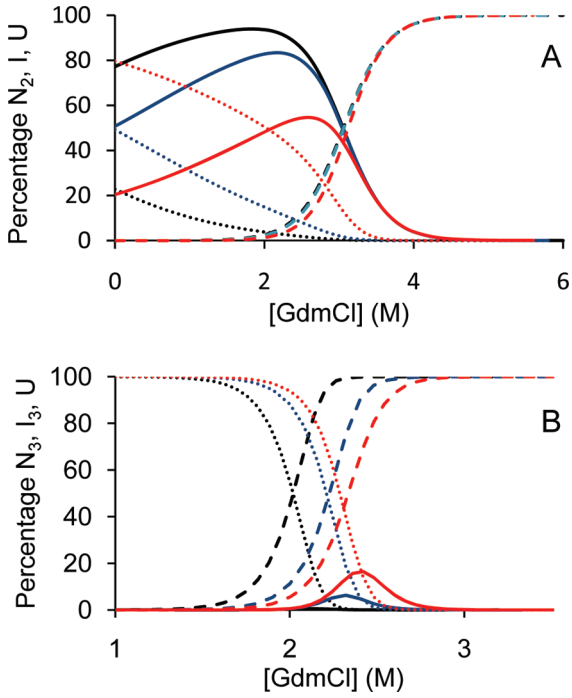

FIGURE 8: Fractional population of native, intermediate, and unfolded species in GdmCl curves for Tm0979 and Mth1491 as a function of protein concentration. (A) The fraction of protein existing as native dimer (dotted line), monomer intermediate (solid line), and unfolded monomer (dashed line) as a function of GdmCl concentration for total dimer concentrations of 0.5 μM (black), 2.5 μM (blue), and 25 μM (red). (B) The fraction of protein existing as native trimer (dotted line), trimer intermediate (solid line), and unfolded monomer (dashed line) as a function of GdmCl concentration for total trimer concentrations of 0.1 μM (black), 1 μM (blue), and 3 μM (red). The fractions of each species were calculated using the parameters for the three-state models given in Table 3 (native dimer with monomer intermediate for Tm0979 and native trimer with trimer intermediate for Mth1491) and the corresponding equations in Table 2.

intermediate is detected for Mth1491 by SEC or in GdmCl curves (Figures 2B and 4), although folding kinetics may suggest monomer formation (Figures 6C,D and 7B). The folding behavior of the proteins can be further illuminated by using the equilibrium parameters to calculate the populations of different species of Tm0979 and Mth1491 under different conditions (Table 2, Figure 8). The monomer for Tm0979 is very highly populated, particularly at low protein concentrations, so that the dominant observed transition in GdmCl curves is monomer unfolding, which is protein concentration independent. The high population of monomer occurs because the K_d is relatively large, favoring dissociation at low concentrations of protein or denaturant, while the monomer stability is relatively high (Table 2). In contrast, for Mth1491 the native trimer is the dominant species well into the folding transition, with some formation of a trimer intermediate at high protein concentrations (Figure 8). This occurs because trimer association is extremely strong (Figure 2, Table 2), so that at the relatively high GdmCl concentrations where the native trimer starts to unfold to trimer intermediate, the conditions may be too destabilizing for a monomer intermediate to be highly populated. In fact, the trimer intermediate has relatively high stability, comparable to various other native trimers (*vide infra* and Table 3).

(B) Distinct Kinetic Folding Mechanisms. The kinetic mechanisms for Tm0979 and Mth1491 also have very marked differences, which can be understood as a consequence of trimer formation being inherently more difficult than dimer formation, due to the higher order and increased size and complexity of the trimer. For Tm0979, the kinetic mechanism is consistent with the equilibrium mechanism: dimer three-state with formation of monomer intermediate. The observed kinetics are dominated by relatively slow unfolding and folding of the monomer intermediate, while dissociation and association of monomers is not apparent, probably because these are relatively rapid processes (Figure 7A). Similar observations have been reported for the folding of other dimeric proteins (*46*−*48*), but a wide range of folding rates and mechanisms have been reported for dimeric proteins in general (*6*) and prediction of specific folding behavior for a given dimer is not yet possible.

The kinetics for Mth1491 are more complex than for Tm0979. The observed unfolding process appears to correspond to native trimer unfolding, while the observed refolding kinetics are multiphasic and include a fast phase that may correspond to monomer folding, as well as slower phases that may involve higher order species (see Results and Figure 7). Assignment of the kinetic phases for Mth1491 is not possible based on the available data, and we note that various sequential or parallel kinetic schemes may apply. Additional states are frequently observed in kinetic compared to equilibrium experiments because species that do not have sufficient relative stability to be populated at equilibrium may nevertheless accumulate transiently during kinetic transitions. For homotrimeric proteins, kinetics have been reported to be two-state, e.g., for simple coiled-coil interfaces (*9, 16*), but they tend to be more complex for more complex interfaces, with various unimolecular and bimolecular steps resulting in population of not only monomeric but also sometimes dimeric and trimeric intermediates (*13, 54, 55*).

Characteristics and Determinants of Stability and Binding. Inherently linked with the differences in the folding mechanisms of Tm0979 and Mth1491 are the very different subunit affinities. SEC experiments revealed that Tm0979 forms a moderate affinity dimer with K_d of 5 μM, whereas Mth1491 monomer association is much stronger, with a K_d of 5.4×10^{-33} M^2, as determined by GdmCl curve analysis (Table 2), considerably smaller than has been reported for other trimers (Table 3). There is some uncertainty in this low K_d, and corresponding high ΔG_U, due to the extrapolation from higher denaturant concentration. The values are highly dependent on the m_U of unfolding, which is somewhat higher than expected for a protein of this size (*45*). However, it is nevertheless clear that Mth1491 is a very high affinity trimer.

(A) High Equilibrium and Kinetic Stability of Proteins from Thermophilic Organisms. The relatively high ΔG_U values for Mth1491 and Tm0979 (Table 2) are most likely related to their origin from thermophilic organisms. The stability measurements here were made at 25 °C; however, both proteins must function biologically at considerably higher temperature, since the optimal growth temperature for *M. thermoautotrophicum* is 65 °C (*36*), while that of *T. maritima* is 80 °C (*37*). Thermophilic proteins have been found, in general, to have higher stability against thermal and chemical denaturation than their mesophilic counterparts; most often this is simply due to an increase in ΔG_U for the thermophilic

proteins (*39, 40, 56*). Thus, thermophilic proteins have comparable, moderate stability under biological conditions as their mesophilic relatives and higher stability at lower temperatures. Very high stability and strong association have been reported for dihydrofolate reductase (DHFR) from *T. maritima* (*57*): at 15 °C this dimeric protein has ΔG_U of 33.9 kcal (mol of dimer)$^{-1}$, corresponding to K_d of 1.88×10^{-26} M. Thus, the DHFR dimer is half-dissociated at a protein concentration equal to K_d, which is much lower than the ~6 pM for half-dissociation of the Mth1491 trimer. As another example, dimeric ORF56 from the hyperthermophile *Sulfolobus islandicus* has ΔG_U of 20.3 kcal (mol of dimer)$^{-1}$ corresponding to K_d of 2.15×10^{-15} M at 25 °C (*58*).

The increased equilibrium stability for thermophilic proteins implies changes in kinetics, i.e., increased folding rates and/or decreased unfolding rates. Folding is quite fast for both Mth1491 and Tm0979, being complete within seconds in the absence of denaturant (Figure 7), but is not particularly fast compared with mesophilic proteins. Unfolding is fairly slow for Tm0979, with k_{uH_2O} of 7.0×10^{-4} s^{-1}, and it is remarkably slow for Mth1491, with k_{uH_2O} of 6.5×10^{-13} s^{-1}. The relative rates for Mth1491 and Tm0979 are also consistent with the times required for denaturation and renaturation curves to reach equilibrium; this was particularly long (23 days) for denaturation of Mth1491. Thus, for the two DsrEFH proteins increasing stability is associated with decreasing unfolding rate.

Although there is relatively little kinetic folding data available for thermophilic proteins, increased stability of some other thermophilic, highly stable proteins has also been attributed to extremely slow unfolding rates. In particular, k_{uH_2O} for dimeric DHFR from *T. maritima* is 4.6×10^{-12} s^{-1}, 10^8 fold slower compared to k_{uH_2O} for monomeric DHFR from *Escherichia coli* (*57*). Also, monomeric pyrrolidone carboxyl peptidase from the hyperthermophile *Pyrococcus furiosus* has k_{uH_2O} of 1.6×10^{-15} s^{-1}, 10^7-fold slower than for the homologous enzyme from *Bacillus amyloliquefaciens* (*59*). Similar, though less pronounced, slowed unfolding rates were also associated with increased stability from mesophilic to thermophilic and hyperthermophilic histone dimers and for several other monomeric proteins from thermophiles compared to mesophiles (*60*). It is reasonable that disruption of the larger, more complex, and more stable interface of the Mth1491 trimer compared to the Tm0979 dimer should occur more slowly. Similarly, the homotrimeric nonthermophilic coiled-coil Lpp-56 from *E. coli* also unfolds very slowly, with rate constants of 10^{-9}−10^{-11} s^{-1}; however, other trimers unfold much faster (Table 3). The folding rates and mechanisms for these trimers are also variable and stabilities are moderate (Table 3). Thus, as has been found for dimers (*6*), it is emerging that in general the folding behavior of trimeric proteins is very diverse.

(B) Stability and Interface Characteristics. An important question is what molecular mechanism(s) may underly the very high stability and slow unfolding of Mth1491. The high stability of thermophilic proteins has been attributed to many different mechanisms: two common ones are oligomerization and increased salt bridges (*38, 40, 61*). The increased number of subunits and a network of salt bridges in the interface for homotrimeric Mth1491 (*28*) may be related to its increased stability relative to dimeric Tm0979 and other proteins. Binding affinity tends to increase also with interface size (*62, 63*) and percentage of monomer buried (*64, 65*). Consistent with this, total interface areas for Mth1491 and Tm0979 are ~4400 and ~2200 Å, respectively, while percentage buried values are ~25% and 20% (calculated using Getarea (*66*); see also Table 3). From both an equilibrium and kinetic standpoint, disruption of the larger interface is likely to require more energy. More detailed molecular interpretations are currently difficult. Although there is a great deal of research into prediction of subunit affinities, as illustrated in results of CAPRI, accurate predictions remain elusive (*3*). The data obtained for Tm0979 and Mth1491 will be valuable for testing and further refining prediction algorithms. Additional clues regarding the relative stabilities may be gleaned from consideration of protein biological function.

Biological Significance of Folding Mechanisms. Recent structural analyses for heterohexameric proteins from the DsrEFH superfamily, TusBCD from *E. coli* (*26*) and DsrEFH from *Allochromatium vinosum* (*25*), strongly suggest that Tm0979 is involved in formation of an analogous, functional DsrEFH structure in *T. maritima.*

Furthermore, only DsrE/F proteins have been reported to form homotrimers (*28*) and homohexamers (*27*), while DsrH proteins have been observed as monomer, homodimer, and heterohexamers (*23*−*26*). C-Terminal β-strands from each subunit form the central core of all the trimer-based structures, including homotrimer and homo- and heterohexamer. Based on our biochemical data, it appears that Tm0979 cannot form a stable homotrimer or homohexamer. This is probably a consequence of Tm0979's shorter secondary structure elements compared to DsrE/F proteins, in particular shorter first and fourth helices and lack of a fifth strand, which are involved in forming the trimer interface (Figure 1). Although the Dsr E, F, and H proteins have related sequences (*32*), they are quite divergent, and the DsrH proteins (such as Tm0979) tend to be slightly shorter and have been classified into a different Pfam than the DsrE/F proteins (such as Mth1491). To date, activity has only been demonstrated for the heterohexameric complexes (*29*−*32*), while Mth1491 has been proposed to have a related activity distinct from DsrEFH, whose genes are nearby in the *M. thermoautotrophicum* genome (*29*).

The preceding considerations suggest that the homodimeric form of Tm0979 may be involved in assembly and regulation. Formation of homodimer may be important for protecting Tm0979 monomers with exposed heterooligomerization interfaces against undesired aggregation. This is supported by the observation that Tm0979 aggregates *in vitro* at low protein concentration where the protein is monomeric. The relatively weak Tm0979 homodimer association and rapid association/dissociation rates could facilitate assembly and disassembly of higher order complexes and regulation of function. Many proteins that assemble into complexes may behave in an analogous fashion (*67*). This could be particularly common for proteins that form heterocomplexes of homologous subunits; for example, homo- and heterooligomer formation has been observed for chaperones such as Tim9/Tim10 (*68*) and for S100 EF-hand signaling proteins (*69*). Also, formation of both homodimers and homotrimers is involved in regulation of human glucocorticoid-induced receptor ligand (*70*) and phospholipase A2 (*71*).

A major mechanism in the evolution of protein function and regulation is the formation or disruption of interactions between identical or related subunits. In depth analyses of the folding and assembly of related protein complexes can shine a light on both evolution and the general principles of protein−protein interactions (*4, 18, 19*). An interesting emerging mechanism involves β-strands at the subunit edges mediating protein−protein interactions. For example, in the DNA binding Cro proteins a dramatic evolutionary switch from a monomeric helical protein to a dimeric mixed α/β protein is based on conversion of an α-helix at the edge of the subunit to β-sheet, which forms the dimer interface (*18*). Also, simple mutations in the hydrophobic core of monomeric immunoglobulin-binding domain B1 of streptococcal protein G cause remarkable changes in quaternary structure, generating various dimers as well as tetramer (*19*). Here also edge β-strands mediate protein−protein interactions, stabilized by additional hydrophobic interactions. For the Dsr proteins, a monomeric ancestor appears likely, given the reasonably high stability of the Tm0979 monomer. Homotrimer followed by homohexamer (a dimer of trimers) (*27, 28*) could logically arise from such a monomer (*4*). The C-terminus of monomeric and dimeric Tm0979 structures is poorly ordered (*23*), but this region forms a well-ordered β-strand in the center of the trimer-based structures (*25*−*28*). The first and fourth helices have hydrophobic faces that form adjacent parts of the subunit interfaces in the trimer. However, the hydrophobic face of the fourth helix also forms the center of the dimer interface for Tm0979. Thus, it is easy to envision a switch between the monomer, dimer, and trimer complexes through mutations in the common interface elements. Again, a key aspect of these protein−protein interactions is variable formation and interactions of an edge β-strand.

Recent studies have highlighted that interchain β-strand interactions occur widely in protein quaternary structures, interactions between proteins, and protein aggregation and that they play a central role in many biological processes and in diseases (*20*). Quantitative analyses of the energetics of folding and assembly for related proteins with different quaternary structures provide important insights into the molecular basis for protein−protein interactions and the evolution of structure and function. They also provide valuable data for testing models of protein folding, assembly, and evolution, in order to reach the ultimate goal of accurate and detailed prediction of protein complexes.

ACKNOWLEDGMENT

We are very grateful to Rod Merrill and Gerry Prentice for access to and assistance with stopped-flow CD measurements. We also thank Cheryl Arrowsmith for the expression construct of Mth1491, Adelinda Yee for help with protein expression, and the members of our group for helpful discussions of this work. Jessica Rumfeldt and R. William Lewis provided valuable assistance with data analysis.

SUPPORTING INFORMATION AVAILABLE

DSC fitted parameters for Tm0979, calorimetry analyses of Tm0979 and Mth1491, and simulation of trimer two-state and three-state folding models. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Goodsell, D. S., and Olson, A. J. (2000) Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct. 29*, 105–153.
2. Marianayagam, N. J., Sunde, M., and Matthews, J. M. (2004) The power of two: protein dimerization in biology. *Trends Biochem. Sci. 29*, 618–625.
3. Lensink, M. F., Mendez, R., and Wodak, S. J. (2007) Docking and scoring protein complexes: CAPRI 3rd Edition. *Proteins 69*, 704–718.
4. Levy, E. D., Boeri Erba, E., Robinson, C. V., and Teichmann, S. A. (2008) Assembly reflects evolution of protein complexes. *Nature 453*, 1262–1265.
5. Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding Des. 3*, R81−R91.
6. Rumfeldt, J. A., Galvagnion, C., Vassall, K. A., and Meiering, E. M. (2008) Conformational stability and folding mechanisms of dimeric proteins. *Prog. Biophys. Mol. Biol.* .
7. Backmann, J., Schafer, G., Wyns, L., and Bonisch, H. (1998) Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon Sulfolobus acidocaldarius. *J. Mol. Biol. 284*, 817–833.
8. Belogrudov, G. I., Schirf, V., and Demeler, B. (2006) Reversible self-association of recombinant bovine factor B. *Biochim. Biophys. Acta 1764*, 1741–1749.
9. Bjelic, S., Karshikoff, A., and Jelesarov, I. (2006) Stability and folding/unfolding kinetics of the homotrimeric coiled coil Lpp-56. *Biochemistry 45*, 8931–8939.
10. Boice, J. A., Dieckmann, G. R., DeGrado, W. F., and Fairman, R. (1996) Thermodynamic analysis of a designed three-stranded coiled coil. *Biochemistry 35*, 14480–14485.
11. Dixon, A. M., Stanley, B. J., Matthews, E. E., Dawson, J. P., and Engelman, D. M. (2006) Invariant chain transmembrane domain trimerization: a step in MHC class II assembly. *Biochemistry 45*, 5228–5234.
12. Durr, E., and Bosshard, H. R. (2000) Folding of a three-stranded coiled coil. *Protein Sci. 9*, 1410–1415.
13. Guthe, S., Kapinos, L., Moglich, A., Meier, S., Grzesiek, S., and Kiefhaber, T. (2004) Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. *J. Mol. Biol. 337*, 905–915.
14. Hlodan, R., and Pain, R. H. (1995) The folding and assembly pathway of tumour necrosis factor TNF alpha, a globular trimeric protein. *Eur. J. Biochem. 231*, 381–387.
15. Jelesarov, I., and Lu, M. (2001) Thermodynamics of trimer-of-hairpins formation by the SIV gp41 envelope protein. *J. Mol. Biol. 307*, 637–656.
16. Marti, D. N., Bjelic, S., Lu, M., Bosshard, H. R., and Jelesarov, I. (2004) Fast folding of the HIV-1 and SIV gp41 six-helix bundles. *J. Mol. Biol. 336*, 1–8.
17. Simler, B. R., Doyle, B. L., and Matthews, C. R. (2004) Zinc binding drives the folding and association of the homo-trimeric gamma-carbonic anhydrase from Methanosarcina thermophila. *Protein Eng., Des. Sel. 17*, 285–291.
18. Roessler, C. G., Hall, B. M., Anderson, W. J., Ingram, W. M., Roberts, S. A., Montfort, W. R., and Cordes, M. H. (2008) Transitive homology-guided structural studies lead to discovery of Cro proteins with 40% sequence identity but different folds. *Proc. Natl. Acad. Sci. U.S.A. 105*, 2343–2348.
19. Jee, J., Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2008) The point mutation A34F causes dimerization of GB1. *Proteins 71*, 1420–1431.
20. Dou, Y., Baisnee, P. F., Pollastri, G., Pecout, Y., Nowick, J., and Baldi, P. (2004) ICBS: a database of interactions between protein chains mediated by beta-sheet formation. *Bioinformatics 20*, 2767–2777.
21. Kinch, L. N., and Grishin, N. V. (2002) Evolution of protein structures and functions. *Curr. Opin. Struct. Biol. 12*, 400–408.
22. Ponstingl, H., Kabir, T., Gorse, D., and Thornton, J. M. (2005) Morphological aspects of oligomeric protein structures. *Prog. Biophys. Mol. Biol. 89*, 9–35.
23. Gaspar, J. A., Liu, C., Vassall, K. A., Meglei, G., Stephen, R., Stathopulos, P. B., Pineda-Lucena, A., Wu, B., Yee, A., Arrowsmith, C. H., and Meiering, E. M. (2005) A novel member of the YchN-like fold: solution structure of the hypothetical protein Tm0979 from Thermotoga maritima. *Protein Sci. 14*, 216–223.

24. Peti, W., Herrmann, T., Zagnitko, O., Grzechnik, S. K., and Wuthrich, K. (2005) NMR structure of the conserved hypothetical protein TM0979 from Thermotoga maritima. *Proteins 59*, 387–390.
25. Dahl, C., Schulte, A., and Shin, D. H. (2007) Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of DsrEFH from Allochromatium vinosum. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 63*, 890–892.
26. Numata, T., Fukai, S., Ikeuchi, Y., Suzuki, T., and Nureki, O. (2006) Structural basis for sulfur relay to RNA mediated by heterohexameric TusBCD complex. *Structure 14*, 357–366.
27. Shin, D. H., Yokota, H., Kim, R., and Kim, S. H. (2002) Crystal structure of a conserved hypothetical protein from Escherichia coli. *J. Struct. Funct. Genomics 2*, 53–66.
28. Christendat, D., Saridakis, V., Kim, Y., Kumar, P. A., Xu, X., Semesi, A., Joachimiak, A., Arrowsmith, C. H., and Edwards, A. M. (2002) The crystal structure of hypothetical protein MTH1491 from Methanobacterium thermoautotrophicum. *Protein Sci. 11*, 1409–1414.
29. Dahl, C., Engels, S., Pott-Sperling, A. S., Schulte, A., Sander, J., Lubbe, Y., Deuster, O., and Brune, D. C. (2005) Novel genes of the dsr gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium Allochromatium vinosum. *J. Bacteriol. 187*, 1392–1404.
30. Ikeuchi, Y., Shigi, N., Kato, J., Nishimura, A., and Suzuki, T. (2006) Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thiouridine biosynthesis at tRNA wobble positions. *Mol. Cell 21*, 97–108.
31. Lauhon, C. T. (2006) Orchestrating sulfur incorporation into RNA. *Nat. Chem. Biol. 2*, 182–183.
32. Pott, A. S., and Dahl, C. (1998) Sirohaem sulfite reductase and other proteins encoded by genes at the dsr locus of Chromatium vinosum are involved in the oxidation of intracellular sulfur. *Microbiology 144* (Part 7), 1881–1894.
33. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem. 182*, 319–326.
34. Manning, L. R., Jenkins, W. T., Hess, J. R., Vandegriff, K., Winslow, R. M., and Manning, J. M. (1996) Subunit dissociations in natural and recombinant hemoglobins. *Protein Sci. 5*, 775–781.
35. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol. 131*, 266–280.
36. Smith, D. R., Doucette-Stamm, L. A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Reeve, J. N., et al. (1997) Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics. *J. Bacteriol. 179*, 7135–7155.
37. Adams, M. W. (1994) Biochemical diversity among sulfur-dependent, hyperthermophilic microorganisms. *FEMS Microbiol. Rev. 15*, 261–277.
38. Kumar, S., Tsai, C. J., and Nussinov, R. (2001) Thermodynamic differences among homologous thermophilic and mesophilic proteins. *Biochemistry 40*, 14152–14165.
39. Ladenstein, R., and Antranikian, G. (1998) Proteins from hyperthermophiles: stability and enzymatic catalysis close to the boiling point of water. *Adv. Biochem. Eng. Biotechnol. 61*, 37–85.
40. Razvi, A., and Scholtz, J. M. (2006) Lessons in stability from thermophilic proteins. *Protein Sci. 15*, 1569–1578.
41. Bose, K., and Clark, A. C. (2001) Dimeric procaspase-3 unfolds via a four-state equilibrium process. *Biochemistry 40*, 14236–14242.
42. Deu, E., and Kirsch, J. F. (2007) The unfolding pathway for Apo Escherichia coli aspartate aminotransferase is dependent on the choice of denaturant. *Biochemistry 46*, 5810–5818.
43. Doyle, S. M., Braswell, E. H., and Teschke, C. M. (2000) SecA folds via a dimeric intermediate. *Biochemistry 39*, 11667–11676.
44. Galani, D., Fersht, A. R., and Perrett, S. (2002) Folding of the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J. Mol. Biol. 315*, 213–227.
45. Geierhaas, C. D., Nickson, A. A., Lindorff-Larsen, K., Clarke, J., and Vendruscolo, M. (2007) BPPred: a Web-based computational tool for predicting biophysical parameters of proteins. *Protein Sci. 16*, 125–134.
46. Malecki, J., and Wasylewski, Z. (1997) Stability and kinetics of unfolding and refolding of cAMP receptor protein from Escherichia coli. *Eur. J. Biochem. 243*, 660–669.
47. Rietveld, A. W., and Ferreira, S. T. (1998) Kinetics and energetics of subunit dissociation/unfolding of TIM: the importance of oligomerization for conformational persistence and chemical stability of proteins. *Biochemistry 37*, 933–937.
48. Svensson, A. K., Bilsel, O., Kondrashkina, E., Zitzewitz, J. A., and Matthews, C. R. (2006) Mapping the folding free energy surface for metal-free human Cu, Zn superoxide dismutase. *J. Mol. Biol. 364*, 1084–1102.
49. Kim, W. K., Henschel, A., Winter, C., and Schroeder, M. (2006) The many faces of protein-protein interactions: A compendium of interface geometry. *PLoS Comput. Biol. 2*, e124.
50. Bernauer, J., Bahadur, R. P., Rodier, F., Janin, J., and Poupon, A. (2008) DiMoVo: a Voronoi tessellation-based method for discriminating crystallographic and biological protein-protein interactions. *Bioinformatics 24*, 652–658.
51. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774–797.
52. Zhu, H., Domingues, F. S., Sommer, I., and Lengauer, T. (2006) NOXclass: prediction of protein-protein interaction types. *BMC Bioinf. 7*, 27.
53. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science 262*, 1401–1407.
54. Forrer, P., Chang, C., Ott, D., Wlodawer, A., and Pluckthun, A. (2004) Kinetic stability and crystal structure of the viral capsid protein SHP. *J. Mol. Biol. 344*, 179–193.
55. Simkovsky, R., and King, J. (2006) An elongated spine of buried core residues necessary for in vivo folding of the parallel beta-helix of P22 tailspike adhesin. *Proc. Natl. Acad. Sci. U.S.A. 103*, 3575–3580.
56. Powers, S. L., Robinson, C. R., and Robinson, A. S. (2007) Denaturation of an extremely stable hyperthermophilic protein occurs via a dimeric intermediate. *Extremophiles 11*, 179–189.
57. Dams, T., and Jaenicke, R. (1999) Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium Thermotoga maritima. *Biochemistry 38*, 9169–9178.
58. Zeeb, M., Lipps, G., Lilie, H., and Balbach, J. (2004) Folding and association of an extremely stable dimeric protein from Sulfolobus islandicus. *J. Mol. Biol. 336*, 227–240.
59. Ogasahara, K., Nakamura, M., Nakura, S., Tsunasawa, S., Kato, I., Yoshimoto, T., and Yutani, K. (1998) The unusually slow unfolding rate causes the high stability of pyrrolidone carboxyl peptidase from a hyperthermophile, Pyrococcus furiosus: equilibrium and kinetic studies of guanidine hydrochloride-induced unfolding and refolding. *Biochemistry 37*, 17537–17544.
60. Topping, T. B., and Gloss, L. M. (2004) Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archael histones: the importance of folding intermediates. *J. Mol. Biol. 342*, 247–260.
61. Szilagyi, A., and Zavodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure 8*, 493–504.
62. Ali, M. H., and Imperiali, B. (2005) Protein oligomerization: how and why. *Bioorg. Med. Chem. 13*, 5013–5020.
63. Janin, J., Miller, S., and Chothia, C. (1988) Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol. 204*, 155–164.
64. Jones, S., and Thornton, J. M. (1995) Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol. 63*, 31–65.
65. Jones, S., and Thornton, J. M. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A. 93*, 13–20.
66. Fraczkiewicz, R., and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem. 19*, 319–333.
67. Nooren, I. M., and Thornton, J. M. (2003) Structural characterisation and functional significance of transient protein-protein interactions. *J. Mol. Biol. 325*, 991–1018.
68. Vial, S., Lu, H., Allen, S., Savory, P., Thornton, D., Sheehan, J., and Tokatlidis, K. (2002) Assembly of Tim9 and Tim10 into a functional chaperone. *J. Biol. Chem. 277*, 36100–36108.
69. Santamaria-Kisiel, L., Rintala-Dempsey, A. C., and Shaw, G. S. (2006) Calcium-dependent and -independent interactions of the S100 protein family. *Biochem. J. 396*, 201–214.
70. Zhou, Z., Song, X., Berezov, A., Zhang, G., Li, Y., Zhang, H., Murali, R., Li, B., and Greene, M. I. (2008) Human glucocorticoid-induced TNF receptor ligand regulates its signaling activity through multiple oligomerization states. *Proc. Natl. Acad. Sci. U.S.A. 105*, 5465–5470.

71. Singh, G., Gourinath, S., Saravanan, K., Sharma, S., Bhanumathi, S., Betzel, C., Srinivasan, A., and Singh, T. P. (2005) Sequence-induced trimerization of phospholipase A2: structure of a trimeric isoform of PLA2 from common krait (Bungarus caeruleus) at 2.5 Å resolution. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 61*, 8–13.
72. Alley, S. C., Shier, V. K., Abel-Santos, E., Sexton, D. J., Soumillion, P., and Benkovic, S. J. (1999) Sliding clamp of the bacteriophage T4 polymerase has open and closed subunit interfaces in solution. *Biochemistry 38*, 7696–7709.
73. Benaroudj, N., Batelier, G., Triniolles, F., and Ladjimi, M. M. (1995) Self-association of the molecular chaperone HSC70. *Biochemistry 34*, 15282–15290.
74. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51–55, 29−32.

BI801784D