BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Folding Mechanism of an Extremely Thermostable $(\beta\alpha)_8$-Barrel Enzyme: A High Kinetic Barrier Protects the Protein from Denaturation

Linn Carstensen,[†] Gabriel Zoldák,[‡] Franz-Xaver Schmid,[§] and Reinhard Sterner*[,†]

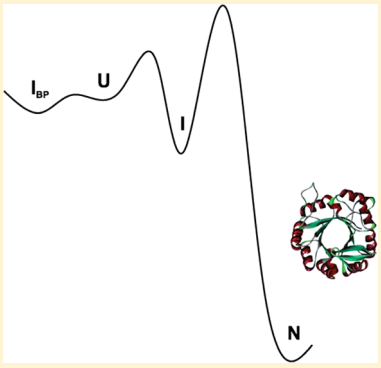
[†]Universität Regensburg, Institut für Biophysik und physikalische Biochemie, Universitätsstrasse 31, D-93053 Regensburg, Germany

[‡]Physik Department E22, Technische Universität München, D-85748 Garching, Germany

[§]Universität Bayreuth, Laboratorium für Biochemie, D-95440 Bayreuth, Germany

Supporting Information

ABSTRACT: HisF, the cyclase subunit of imidazole glycerol phosphate synthase (ImGPS) from *Thermotoga maritima*, is an extremely thermostable $(\beta\alpha)_8$-barrel protein. We elucidated the unfolding and refolding mechanism of HisF. Its unfolding transition is reversible and adequately described by the two-state model, but 6 weeks is necessary to reach equilibrium (at 25 °C). During refolding, initially a burst-phase off-pathway intermediate is formed. The subsequent productive folding occurs in two kinetic phases with time constants of ∼3 and ∼20 s. They reflect a sequential process via an on-pathway intermediate, as revealed by stopped-flow double-mixing experiments. The final step leads to native HisF, which associates with the glutaminase subunit HisH to form the functional ImGPS complex. The conversion of the on-pathway intermediate to the native protein results in a 10^6-fold increase of the time constant for unfolding from 89 ms to 35 h (at 4.0 M GdmCl) and thus establishes a high energy barrier to denaturation. We conclude that the extra stability of HisF is used for kinetic protection against unfolding. In its refolding mechanism, HisF resembles other $(\beta\alpha)_8$-barrel proteins.



Approximately 1200 different protein folds have been identified to date [Structural Classification of Proteins (SCOP) release 1.75, February 2009],[1] each of them being characterized by a distinct toplogical orientation of secondary structure elements. Whereas many folds are represented by only a few members of the protein database, others have been recruited extensively in the course of evolution. A prominent example is the $(\beta\alpha)_8$-barrel, which is among the most ancient, frequent, and versatile folds.[2−4] Approximately 10% of all proteins with known three-dimensional structure contain at least one $(\beta\alpha)_8$-barrel domain. With very few exceptions, all known $(\beta\alpha)_8$-barrels are enzymes, and SCOP distinguishes 33 superfamilies that catalyze more than 60 different reactions. They occur in five of the six Enzyme Commission (EC) classes, acting as oxidoreductases, transferases, lyases, hydrolases, and isomerases, and many of them are engaged in essential metabolic pathways.[5,6]

The canonical $(\beta\alpha)_8$-barrel fold is composed of eight units, each of which contains at least 25 residues. A single unit consists of a β-strand and an α-helix, which are linked by a $\beta\alpha$-loop, and the individual units are connected by $\alpha\beta$-loops. The eight strands form a central parallel β-sheet, which is surrounded by an outer layer of eight α-helices (Figure 1). Many $(\beta\alpha)_8$-barrel enzymes contain extensions to the canonical topology, at the N- or C-terminus or in loop segements. Remarkably, in all known $(\beta\alpha)_8$-barrels, the residues that are important for substrate specificity and catalytic activity are located at the C-terminal ends of the β-strands and in the subsequent $\beta\alpha$-loops. However, in spite of the common chain topology and the conserved location of the active site, the geometry of the central barrel can differ considerably between individual $(\beta\alpha)_8$-barrel proteins and the overall level of sequence identity is low. The question of whether all $(\beta\alpha)_8$-barrels are the products of divergent evolution from a common ancestor or whether the fold has developed independently several times by convergent evolution is therefore unresolved.[6−8] In any case, extensive sequence and structure comparisons suggest a common evolutionary origin for most $(\beta\alpha)_8$-barrels,[5,9] in line with the notion that conservation of structure exceeds conservation of amino acid sequence.

This observation raises the question of the extent to which the folding mechanism of $(\beta\alpha)_8$-barrel proteins has been conserved upon the vast diversification of their sequences in the course of evolution. Early insights into the folding of $(\beta\alpha)_8$-barrel proteins were obtained by fragmentation experiments performed with phosphoribosyl anthranilate isomerase (PRAI) and triosephosphate isomerase (TIM).[10−12] These studies provided evidence that both PRAI and TIM form stable folding intermediates consisting of four or six structured $(\beta\alpha)$ units, followed by the association of the unstructured segment to

Received: February 10, 2012
Revised: March 27, 2012
Published: March 28, 2012

   dx.doi.org/10.1021/bi300189f | *Biochemistry* 2012, 51, 3420−3432

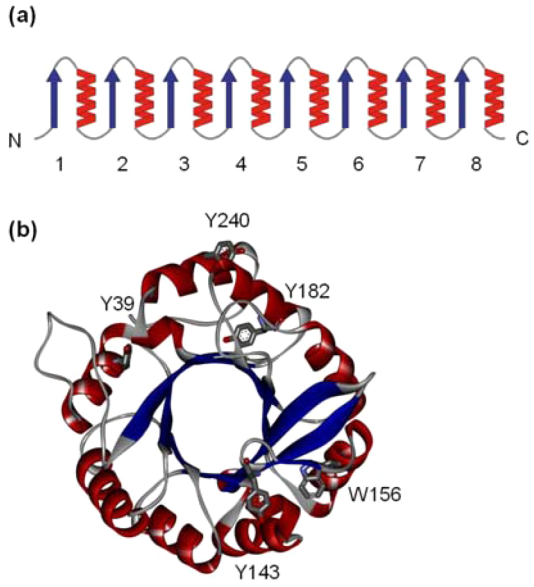


Figure 1. (a) Schematic depiction of the $(\beta\alpha)_8$-barrel fold and (b) ribbon diagram of the crystal structure of HisF (PDB entry 1THF).[23] α-Helices and β-strands are colored red and blue, respectively; loops are colored gray. Tyr and Trp residues are displayed as sticks. The figure was generated using Accelrys DS.

finally form the native $(\beta\alpha)_8$-barrel. Recent in-depth kinetic studies with the α-subunit of tryptophan synthase from *Escherichia coli* (αTS),[13,14] the indole-3-glycerol-phosphate synthase from *Sulfolobus solfataricus* (IGPS),[15−17] and a $(\beta\alpha)_8$-barrel protein of unknown function encoded by the *Bacillus subtilis iol*I gene (IOLI)[18] have provided more insights into the folding of $(\beta\alpha)_8$-barrels. All three proteins apparently form an off-pathway burst-phase intermediate, which must unfold to enter productive folding via one or more on-pathway intermediates. Although these similarities have suggested that the common topology is the main determinant for the folding mechanism of $(\beta\alpha)_8$-barrel proteins, the native secondary structural elements formed in the on-pathway intermediates seem to differ between αTS and IGPS. Moreover, whereas the slow cis−trans isomerization of peptide bonds involving proline residues results in several parallel folding pathways in the case of αTS, IGPS and IOLI contain only trans prolyl peptide bonds and show simpler folding mechanisms.[17]

We have studied the folding mechanism of the cyclase subunit HisF of the imidazole glycerol phosphate synthase from *Thermotoga maritima* (Figure 1b). HisF is an interesting protein for several reasons. First, it is active in its hyperthermophilic host above 80 °C and provides an excellent model for testing the hypothesis that high thermodynamic stability is related with a large kinetic barrier to unfolding.[19−22] Second, unlike other $(\beta\alpha)_8$-barrel proteins, HisF shows a striking 2-fold symmetry and contains phosphate binding sites in both the N-terminal and C-terminal half-barrel, which anchor its biphosphorylated substrate.[23] IGPS and αTS are evolutionarily related to HisF but lack such a clear 2-fold symmetry and contain only a single conserved C-terminal phosphate binding site,[24] which makes a comparison of their folding mechanisms particularly attractive. Finally, as HisF forms a 1:1 complex with the glutaminase subunit HisH of the imidazole glycerol phosphate synthase,[25,26] it offers an opportunity to study the relationship between protein folding and association.

MATERIALS AND METHODS

Mutagenesis, Cloning and Expression of Genes, and Purification of Recombinant Proteins. Point mutations were introduced into the *his*F gene by overlap extension polymerase chain reaction,[27] followed by cloning of the genes into plasmid pET24a(+) via the NdeI and BamHI restriction sites. Wild-type *his*F and its mutants were expressed in *E. coli* T7-Express cells (New England Biolabs). After induction with 0.5 mM IPTG, cells were grown for 4 h at 37 °C and harvested. The recombinant proteins with molecular masses of ∼27.7 kDa were purified as described previously.[28] According to sodium dodecyl sulfate−polyacrylamide gel electrophoresis, all HisF variants were >95% pure. Approximately 100 mg of protein per liter of culture was obtained and dialyzed against 50 mM Tris-HCl buffer (pH 7.5).

Equilibrium Unfolding Transitions. Samples with 4 μM protein were prepared in 50 mM Tris-HCl buffer (pH 7.5) containing various concentrations of GdmCl. GdmCl (ultrapure) was purchased from MP Biomedicals (Illkirch, France), and its concentration was determined by the refractive index of the solution.[29] The samples were incubated at 25 and 45 °C for 6 weeks and 19 days, respectively. The far-UV circular dichroism (CD) signal at 225 nm was monitored with a JASCO model J815 CD spectrophotometer, using a 5 mm cuvette and a bandwidth of 1 nm. Following excitation at 280 nm (bandwidth of 3 nm), the fluorescence emission signal at 320 nm (bandwidth of 5 nm) was monitored in a JASCO model FP-6500 spectrofluorimeter. The transitions were analyzed according to the two-state equilibrium model. The free energy change of unfolding in the absence of denaturant (ΔG_D°) was obtained assuming a linear dependency of ΔG_D on GdmCl concentration.[30]

Kinetics of Unfolding and Refolding. All experiments were performed in 50 mM Tris-HCl buffer (pH 7.5) at 25 or 45 °C. Unfolding was initiated by dilution of the native protein to various denaturing concentrations of GdmCl. Refolding was measured after dilution of the denatured (in 6.0 M GdmCl) protein to folding conditions.

Manual mixing experiments monitored by the far-UV CD signal were performed with a JASCO model J815 CD spectrophotometer. Following a 10-fold dilution, the CD signal of 4 μM of protein (final concentration) was recorded at 225 nm, using a 5 mm cuvette, a bandwidth of 1 nm, and an averaging time of 1 s. Manual mixing experiments followed by Trp/Tyr fluorescence were performed with a JASCO model FP-6500 spectrofluorometer using a 1 cm cuvette. The fluorescence signal of 4 μM protein (final concentration) was followed using an excitation wavelength of 280 nm (bandwidth of 3 nm) and an emission wavelength of 320 nm (bandwidth of 5 nm). Monoexponential or biexponential functions were fit to the data using GraFit6 from Erithacus (West Sussex, U.K.).

Stopped-flow fluorescence experiments were performed using a SX.20MV spectrometer from Applied Photophysics (Leatherhead, U.K.). After rapid mixing, the fluorescence signal of 2 μM protein (final concentration) was followed using a 320 nm emission cutoff filter after excitation at 280 nm (Tyr and Trp) or 290 nm (Trp). The path length of the observation cell was 2 mm, and the bandwidth was 10 nm. The kinetics was measured eight times under identical conditions, averaged, and analyzed using exponential functions.

The time-dependent formation of I from U was followed by stopped-flow fluorescence interrupted refolding experiments.

dx.doi.org/10.1021/bi300189f | *Biochemistry* 2012, 51, 3420−3432

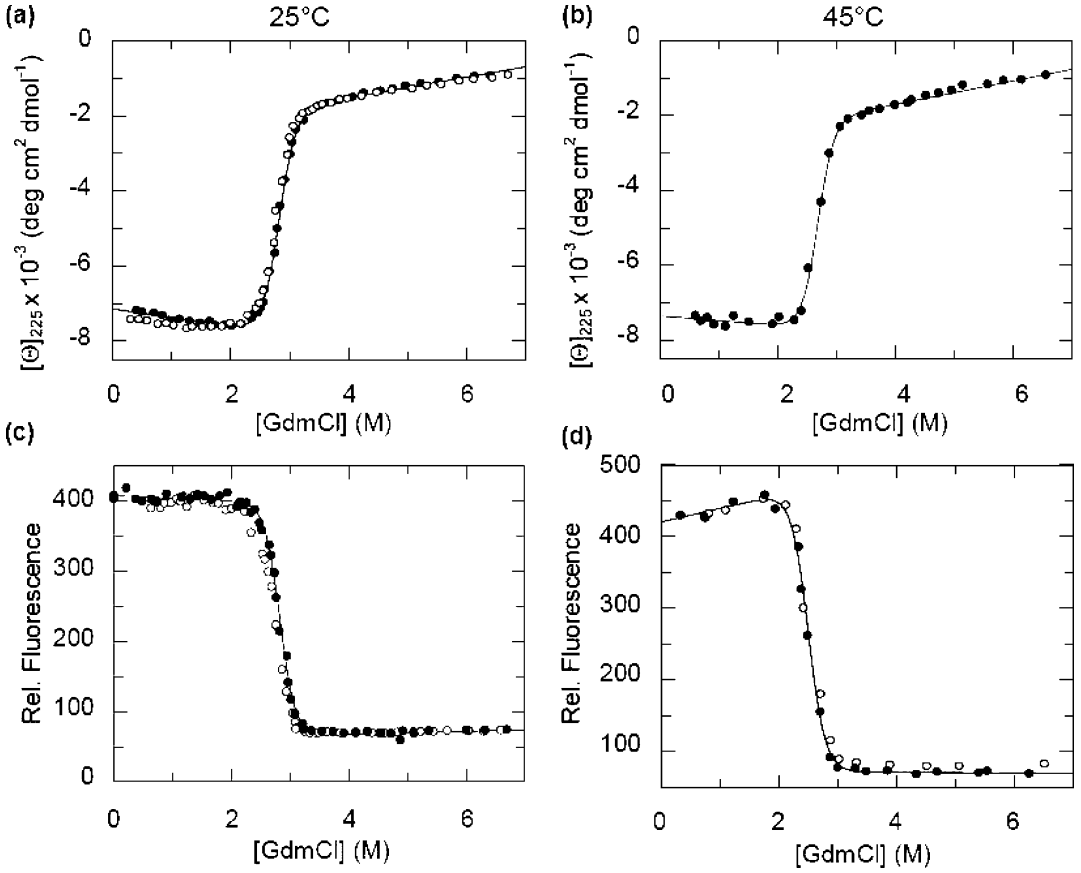


Figure 2. GdmCl-induced equilibrium unfolding of HisF at 25 °C (a and c) and 45 °C (b and d). The transitions were followed by the CD signal at 225 nm (a and b) and Trp/Tyr fluorescence (excitation at 280 nm; emission at 320 nm) (c and d) in 50 mM Tris-HCl buffer (pH 7.5). Filled symbols represent data from unfolding experiments, started with folded protein, and empty symbols represent data from refolding experiments, started with protein that was previously unfolded in 6.0 M GdmCl. The solid lines represent fits to the unfolding transitions on the basis of the two-state model; the parameters derived from the fit are listed in Table 1.

Unfolded HisF (132 μM in 6.0 M GdmCl) was diluted 6-fold to 1.0 M GdmCl and incubated for 10 s to allow for the conversion of U to I. The sample was then diluted 11-fold to final GdmCl concentrations between 1.9 and 7.0 M, which resulted in the rapid conversion of I to U. A monoexponential equation with a linear factor was fit to the unfolding traces to determine the rate constant for the I → U reaction at the respective denaturant concentration.

The kinetics of the sequential U → I → N reaction was monitored in further stopped-flow fluorescence interrupted refolding experiments. Unfolded HisF (132 μM in 6.0 M GdmCl) was diluted 6-fold to 1.0 M GdmCl and incubated for various times to allow for refolding to I and N. The sample was then diluted 11-fold to 2.3 or 6.9 M GdmCl, which resulted in the rapid unfolding of I and the slow unfolding of N. Monoexponential or double-exponential equations with a linear factor were fit to the kinetic traces. An equation describing the consecutive model was fit to the obtained amplitudes.[31,32]

The recovery of native HisF molecules during refolding was probed by double-jump stopped-flow fluorescence experiments combining refolding of HisF with binding to HisH. Unfolded HisF-W156Y (12 μM in 6.0 M GdmCl) was diluted 6-fold to 1.0 M GdmCl and refolded for variable times. The sample was then mixed with an equal volume of a solution containing the same concentration of HisH in buffer without GdmCl. A monoexponential equation was fit to the obtained amplitudes.

RESULTS

Equilibrium Transitions. The thermodynamic stability of HisF was characterized by GdmCl-induced equilibrium unfolding transitions at 25 and 45 °C. The loss of secondary structure was monitored by far-UV circular dichroism (CD) at 225 nm, and the loss of tertiary structure was followed by tryptophan and tyrosine fluorescence. HisF contains a single tryptophan residue (W156) located in α-helix 5 and four tyrosine residues (Y39, Y143, Y182, and Y240), which are distributed across the protein (Figure 1b). Their fluorescence served as an indicator for global folding.

HisF is a thermostable protein [T_M > 90 °C (data not shown)], and the folding equilibrium in the transition region is attained very slowly. After incubation at 45 °C for 10 days, the unfolding curve was still shifted to higher denaturant concentrations with respect to the refolding curve, indicating that the unfolding reaction had not yet reached equilibrium (Figure S1 of the Supporting Information). However, after 19 days, the unfolding and refolding curves coincided, indicating that equilibrium had been reached. Remarkably, unfolding remained fully reversible even after this excessively long period of incubation at intermediate denaturant concentrations. At 25 °C, unfolding was even slower, and an incubation time of 6 weeks was required to reach equilibrium.

The folding reaction of HisF is cooperative, and the transitions determined by CD and fluorescence superimpose well (Figure 2), suggesting that folding intermediates are not

Table 1. Thermodynamic Parameters for the GdmCl-Induced Unfolding of HisF and I_{BP}[a]

	25 °C			45 °C		
protein, signal	$\Delta G_D°$ (kJ mol^{-1})	m (kJ mol^{-1} M^{-1})	$[D]_{1/2}$ (M)	$\Delta G_D°$ (kJ mol^{-1})	m (kJ mol^{-1} M^{-1})	$[D]_{1/2}$ (M)
HisF, far-UV CD	53.5 ± 1.7	19.1 ± 0.6	2.80	52.7 ± 3.3	19.7 ± 1.2	2.68
HisF, fluorescence	55.6 ± 2.3	19.9 ± 0.9	2.79	47.9 ± 2.8	18.0 ± 1.1	2.66
I_{BP}, far-UV CD	15.2 ± 2.4	7.9 ± 0.9	1.90			

[a]Values are for 50 mM Tris-HCl buffer (pH 7.5) at the respective temperature. The thermodynamic parameters were calculated from the data shown in Figures 2 and 4.

populated at equilibrium. The data were analyzed on the basis of the two-state model, yielding the thermodynamic parameters listed in Table 1. At 25 and 45 °C, HisF unfolds with a similar cooperativity (m ~ 19 kJ mol^{-1} M^{-1}), which is close to the value of 24 kJ mol^{-1} M^{-1} that is expected for a protein of its size (253 residues).[33] The differences between the two temperatures with respect to the transition midpoint ($[D]_{1/2}$ ~ 2.8 M vs ~2.7 M) and the Gibbs free energy of unfolding ($\Delta G_D°$ ~ 54 kJ mol^{-1} vs ~50 kJ mol^{-1}) are small. Thus, the thermodynamic stability of HisF seems to be only weakly dependent on temperature.

Kinetics of Unfolding and Refolding. Unfolding and refolding kinetics of the secondary structure of HisF were measured at 25 and 45 °C by following the far-UV CD signal of the protein after manual mixing at various concentrations of GdmCl. Unfolding was a monoexponential reaction, as shown in Figure 3a for the reaction in 6.0 M GdmCl, and the amplitude accounted for the entire CD change as observed in the equilibrium unfolding transition. Refolding at low GdmCl concentrations was biexponential, as shown in Figure 3b for the reaction at 0.7 M GdmCl. The faster refolding phase vanished above 2 M GdmCl. The CD change observed in the refolding kinetics was significantly smaller than that expected from the equilibrium transitions (Figure 4a), indicating that the major part of the CD change occurred within the dead time of the manual mixing experiment. The kinetics of this CD change could also not be resolved after stopped-flow mixing, indicating that a burst-phase intermediate (I_{BP}) formed in <5 ms. This CD change is independent of protein concentration and thus probably not caused by rapid reversible aggregation.[34] The amplitude of the burst-phase refolding reaction decreased with an increasing concentration of GdmCl in a sigmoidal manner (Figure 4a). This cooperative unfolding process and the high magnitude of the relative CD signal of I_{BP} compared to N (71% at 1 M GdmCl and 36% at 2 M GdmCl) suggest that I_{BP} contains a high content of compact secondary structure. A tentative two-state analysis of the unfolding transition of I_{BP} yielded an apparent $\Delta G_D°$ of 15.2 kJ mol^{-1} and an m value of 7.9 kJ mol^{-1} M^{-1}. Thus, I_{BP} exhibits ~30% of the stability and ~40% of the compactness of the native state (Table 1).

Unfolding and refolding kinetics of the tertiary structure of HisF were measured at 25 and 45 °C by following the fluorescence changes of the single tryptophan and the four tyrosine residues (Figure 1b). Again, the unfolding kinetics were monophasic, and the refolding kinetics at low GdmCl concentrations were biphasic (Figure S2 of the Supporting Information). In both unfolding and refolding, the fluorescence amplitudes accounted for the entire signal change as expected from the equilibrium transitions (Figure 4b), which indicates that I_{BP} lacks significant tertiary structure.

The dye 1-anilino-8-naphthalenesulfonate (ANS) binds to hydrophobic regions of folding intermediates, which leads to an increase in its fluorescence intensity.[35,36] Stopped-flow refolding experiments in the presence of ANS yielded a strong fluorescence increase within the dead time of mixing (Figure S3a of the Supporting Information), indicating that I_{BP} contains hydrophobic surfaces that are capable of ANS binding. The ANS fluorescence then decreased in two exponential phases with rates that correspond to the refolding rates detected by CD and by Trp/Tyr fluorescence (Figure S3b of the Supporting Information). These results show that the multistep transition from I_{BP} to N leads to a continuous decrease in the exposed hydrophobic surface area of HisF.

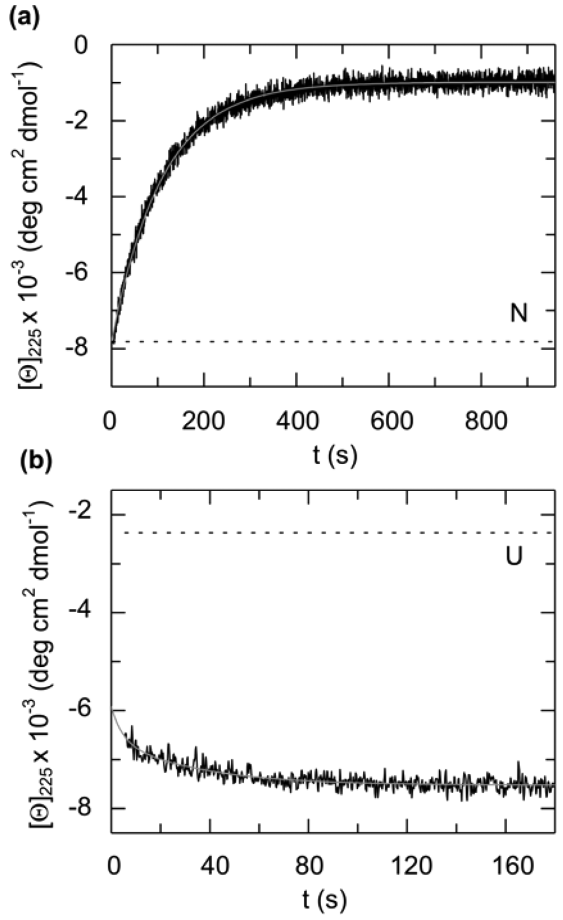


Figure 3. (a) Unfolding kinetics of HisF in 6.0 M GdmCl and (b) refolding kinetics in 0.7 M GdmCl as followed by far-UV CD. Following manual mixing, the traces of the CD signal at 225 nm were monitored in 50 mM Tris-HCl buffer (pH 7.5) at 25 °C. Gray solid lines represent fits of monoexponential and biexponential functions to the unfolding (τ = 110 s) and refolding curves (τ_1 = 3.1 s, and τ_2 = 34 s), respectively. Dashed lines indicate the initial ellipticities characterizing the native (N) state (a) and the unfolded (U) state (b).

The observed rate constants of unfolding and refolding at 25 and 45 °C are shown in panels a and b of Figure 5 as a function of GdmCl concentration in the form of chevron plots. The two

dx.doi.org/10.1021/bi300189f | Biochemistry 2012, 51, 3420−3432

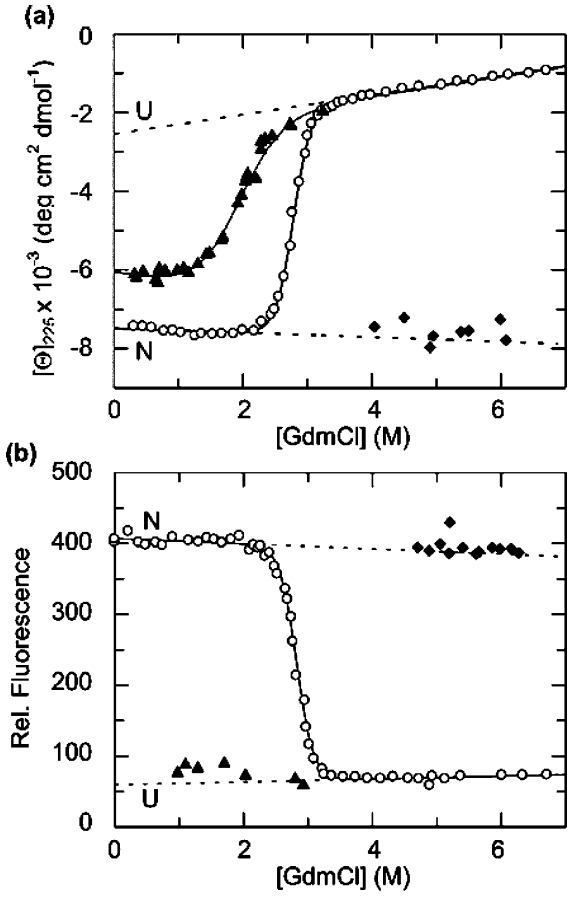


Figure 4. GdmCl-induced equilibrium unfolding transitions of HisF and initial values of refolding and unfolding kinetics. The equilibrium transitions (empty circles) were followed by (a) the CD signal at 225 nm and (b) Trp/Tyr fluorescence (excitation at 280 nm; emission at 320 nm) in 50 mM Tris-HCl buffer (pH 7.5) at 25 °C (data taken from Figure 2a,c) The dashed lines indicate the baselines for the pure N and U states. The initial values of the unfolding (filled diamonds) and refolding (filled triangles) kinetics were obtained by extrapolation to zero time (see Figure 3). The data show that a compact burst-phase refolding intermediate (I_{BP}) with a high content of secondary structure but no detectable tertiary structure is formed. The solid lines represent fits based on the two-state model to the data, yielding the thermodynamic parameters listed in Table 1.

refolding rate constants and the single unfolding rate constant as determined by CD and Trp/Tyr fluorescence superimpose well, indicating that, in these kinetic phases, secondary and tertiary structure formation and denaturation occur simultaneously.

The two refolding rates differ by ~10-fold at 25 °C and ~4-fold at 45 °C and are largely independent of denaturant concentration below 2 M GdmCl (Figure 5a,b). Denaturant-independent kinetics are often observed when folding is limited in rate by prolyl isomerization.[14,37] However, native HisF lacks cis prolines, and both refolding phases are insensitive to the cis/trans prolyl isomerase *Thermococcus* FKBP18 (data not shown). Decreased denaturant dependencies as in panels a and b of Figure 5 are usually called a "roll-over" and assumed to originate from the population of folding intermediates.[15] We suggest that the roll-over in the fast folding rate is caused by the burst-phase intermediate I_{BP}, which is populated at low GdmCl concentrations and shows a transition midpoint of 1.9 M GdmCl (see Figure 4a).

The roll-over observed for the slow folding rate might be caused by a folding intermediate I that is formed in the fast folding phase. In such a case, the two refolding phases observed between 0 and 2 M GdmCl (Figure 5a,b) should reflect two steps on a sequential pathway via a partially folded intermediate that is unstable above 2 M GdmCl. Such a sequential model is supported by the denaturant dependence of the refolding amplitudes (Figure 5c,d). These amplitudes change in a reciprocal fashion. The fast phase dominates at low GdmCl concentrations, but with an increasing GdmCl concentration, it loses amplitude and the amplitude of the slow phase increases in a reciprocal fashion, as expected for a sequential folding mechanism. At 45 °C, the decrease in the amplitude of the fast phase occurs at lower GdmCl concentrations than at 25 °C, suggesting that the intermediate is less stable at the higher temperature.

Above 2 M GdmCl, refolding and unfolding become monophasic at 25 and 45 °C, and the rates decrease and increase exponentially with denaturant concentration (Figure 5a,b). This apparent two-state refolding and unfolding behavior indicates that the putative folding intermediates are destabilized at high denaturant concentrations. The strong dependence of the observed refolding and unfolding rates on GdmCl concentration (corresponding to large m values) leads to extremely slow kinetics at intermediate denaturant concentrations, in line with the excessively long incubation time required to reach equilibrium in the unfolding transitions (Figure S1 of the Supporting Information). At 25 °C, the unfolding reaction exhibits a time constant ($\tau = 1/\lambda$) of 3.5 days at 4 M GdmCl, and between 4 and 2.5 M GdmCl, the kinetics were too slow for the determination of the involved rates.

Interrupted Refolding Assays. The refolding kinetics of HisF are biphasic at low denaturant concentrations (Figure 5), and the amplitude analysis (Figure 5c,d) suggested that the two phases might reflect the formation of an intermediate and its subsequent transformation to the native state. To examine whether such an intermediate is in fact formed as a transient species during refolding of HisF, we employed a stopped-flow double-mixing approach. In these experiments, refolding was interrupted after various times, and the concentrations of species present at the time of sampling were determined from the rates and amplitudes of their unfolding.

At 1.0 M GdmCl, the time constants of the two folding phases are 3.3 and 33 s [at 25 °C (Figure 5a)]. After refolding had been conducted for 10 s, a putative intermediate should thus be well populated and the concentration of fully folded molecules should still be very small. In the double-mixing experiment, refolding was therefore interrupted after 10 s, and the sample was transferred to unfolding conditions of 6.0 M GdmCl (Figure S4 of the Supporting Information), to monitor the rate and the amplitude of its unfolding reaction. Under these conditions, native HisF unfolds with a time constant of ~170 s (Figure 5a). The molecules formed in the fast refolding reaction, however, unfolded 3 orders of magnitude more rapidly with a time constant of 0.02 s. The product of the fast refolding reaction is thus not the native protein, and a parallel folding mechanism can be ruled out.

In further double-mixing experiments, unfolded HisF was also refolded in 1.0 M GdmCl for 10 s but then diluted to GdmCl concentrations between 1.9 and 7.5 M. In all cases, monoexponential unfolding kinetics were observed, and the rate constants were significantly higher (~10^5-fold at 5.0 M

dx.doi.org/10.1021/bi300189f | *Biochemistry* 2012, 51, 3420−3432

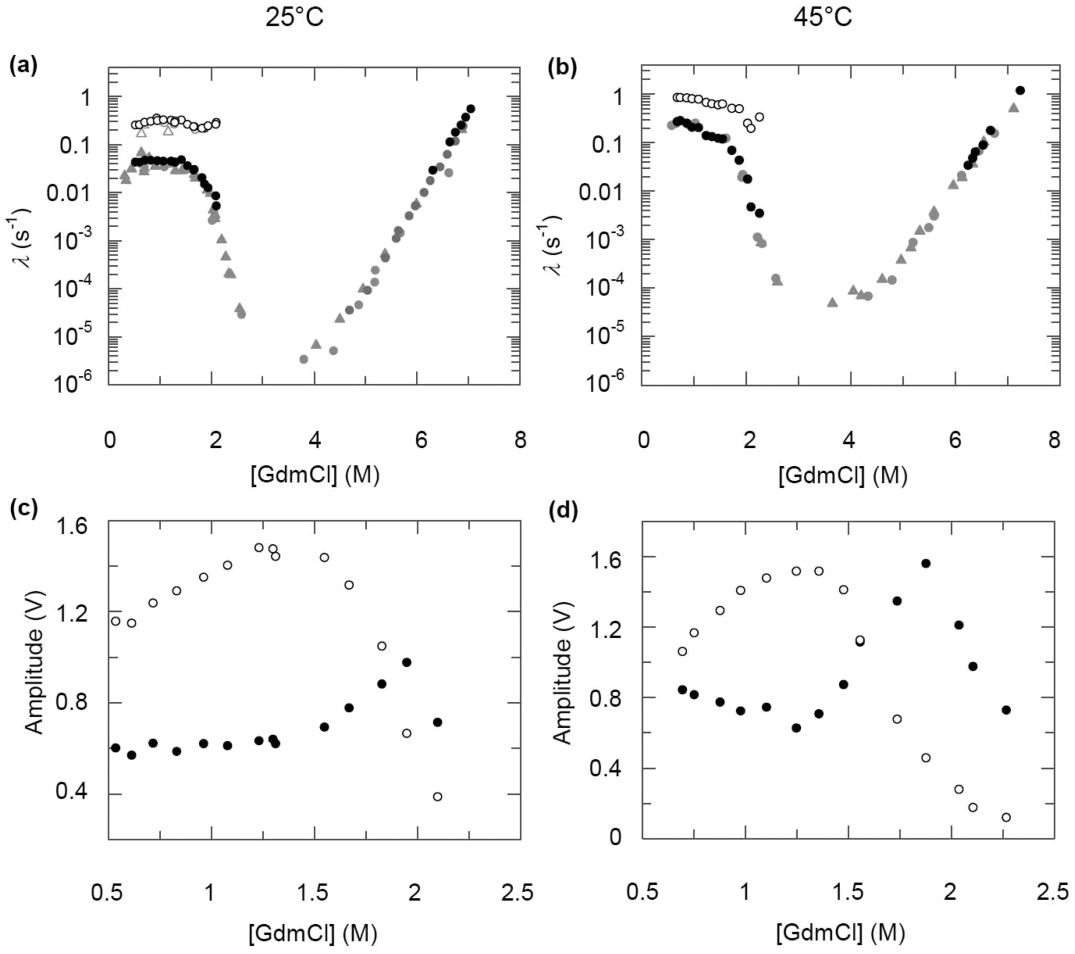


Figure 5. GdmCl dependence of the apparent rate constants (λ) of refolding and unfolding of HisF at (a) 25 °C and (b) 45 °C (chevron diagrams). Following manual mixing, the traces of the CD signal at 225 nm (gray triangles) and Trp/Tyr fluorescence (excitation at 280 nm; emission at 320 nm) (gray circles) were monitored in 50 mM Tris-HCl buffer (pH 7.5). Rate constants for the fast refolding phase are shown with empty symbols, and rate constants for the slow refolding and the single unfolding phases are shown with filled symbols. Reactions with λ values of >0.004 s^{-1} were also followed after stopped-flow mixing monitoring Trp/Tyr fluorescence (black circles). The observed amplitudes of the fast and slow refolding phases are shown (empty and filled black circles, respectively) at (c) 25 and (d) 45 °C.

GdmCl) than the rate constants of unfolding of the native protein, confirming that they reflect the unfolding of an intermediate (I) that formed during the fast folding phase. The measured unfolding rates of I connect smoothly with the previously determined fast refolding rates in the transition region around 2.1 M GdmCl (Figure 6a), indicating that the folding of I is reversible. The resulting chevron shows a minimum near 2.1 M GdmCl, in good agreement with the increase in the amplitude of unfolding of I (Figure 6b), which reflects its stability. Thus, I seems to be more stable than the burst-phase intermediate I_{BP} ($[D]_{1/2}$ = 1.9 M GdmCl) and much less stable than native HisF ($[D]_{1/2}$ = 2.8 M GdmCl). The unfolding rate of I increases between 2 and 4 M GdmCl but becomes denaturant-independent above 5.0 M GdmCl (Figure 6a). The reason for this behavior remains unclear. It might originate from a shift in the free energy of the transition state relative to the native state at high GdmCl concentrations (Hammond behavior)[38−40] or from the accumulation of an unfolding intermediate due to ground state effects.[41]

Interrupted refolding assays were also used to follow the time course of the accumulation and depletion of intermediate I. In these experiments, samples of HisF were allowed to refold in 1.0 M GdmCl for 1−200 s, and then they were transferred to 2.3 M GdmCl. At this denaturant concentration, native HisF molecules remain folded (Figure 2), but the intermediate I unfolds with a time constant of 2 s. The fluorescence amplitude of this unfolding reaction is proportional to the concentration of the intermediate present at the time when refolding was interrupted (Figure 7a) and thus traces the time course of the intermediate I during refolding (Figure 7b). The concentration of I first increased and reached a maximum at a refolding time of 10 s; then it decreased to almost zero after refolding had been conducted for ∼200 s. A fit of a double-exponential equation to these data yielded time constants of 2.9 and 40 s for the formation and depletion of I, respectively, which coincide with the time constants of the fast and the slow reaction, as determined in the conventional refolding kinetics (Figure 7b). This result demonstrates that these two phases reflect the formation and decay of intermediate I.

To examine whether the folding intermediate is directly transformed into the native protein, the unfolding assay in the second step of the double-mixing experiments was performed at 6.9 M GdmCl. At this high denaturant concentration, both the intermediate and the native state unfold, with time constants of 56 ms and 2.6 s, respectively. This large difference in rates allows one to trace the concentrations of I and N

dx.doi.org/10.1021/bi300189f | *Biochemistry* 2012, 51, 3420−3432

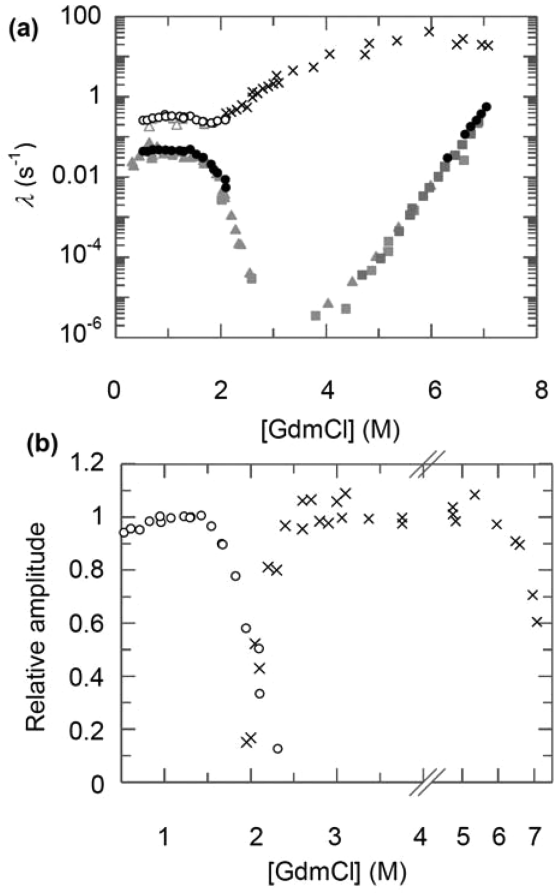


Figure 6. Analysis of the folding intermediate of HisF by interrupted refolding experiments. (a) Unfolding rate constants (λ) observed after refolding in 1.0 M GdmCl for 10 s in stopped-flow double-mixing experiments at 25 °C using Trp/Tyr fluorescence (excitation at 280 nm; emission at 320 nm) (crosses). The rate constants determined by single-mixing experiments (Figure 5a) are shown for comparison. (b) GdmCl dependence of the fluorescence amplitudes of the fast refolding phase (empty circles) derived from single-mixing experiments and of the unfolding phase (crosses) as observed in the double-mixing experiments. The amplitudes are constant between 0 and 1.4 M GdmCl for refolding and between 2.6 and 6.0 M GdmCl for unfolding. These amplitudes represent complete refolding from U to I and complete unfolding of I to U, respectively, and were therefore set to 1. Both amplitudes have a value of 0.5 at 2.1 M GdmCl, which is consistent with the transition midpoint ($[D]_{1/2}$) as estimated from the chevron diagram of I shown in panel a.

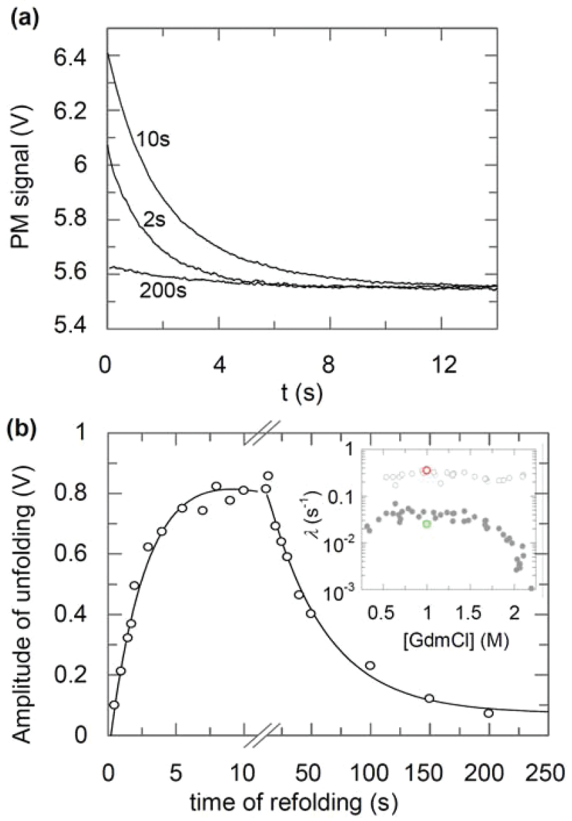


Figure 7. Time course of the folding intermediate (I) of HisF determined by stopped-flow double-mixing experiments. (a) Refolding of unfolded protein was initiated by dilution to 1 M GdmCl. After various refolding times (shown are 2, 10, and 200 s), the protein was diluted to 2.3 M GdmCl, and unfolding monitored by the decay of Trp/Tyr fluorescence (excitation at 280 nm; emission at >320 nm) was followed. A monoexponential function and a linear term, which represents the onset of the slow refolding reaction, were fit to the experimental curves. The time constant of unfolding of the intermediate ($\tau \sim 2$ s) was largely independent of the refolding time. (b) Relative amplitudes of unfolding in 2.3 M GdmCl as a function of refolding time in 1.0 M GdmCl. The solid line represents the fit of a double-exponential equation to the experimental points, yielding time constants of 2.9 s for the accumulation of the intermediate and 40 s for its depletion. In the inset, the observed rate constants for the accumulation (red circle) and depletion (green circle) of the intermediate are compared with the refolding rates determined by single-jump experiments (gray circles, data taken from Figure 5a). All experiments were conducted in 50 mM Tris-HCl buffer (pH 7.5) at 25 °C.

simultaneously and with high precision from the amplitudes of their unfolding reactions.

In a sequential three-state reaction, the formation of the native state should show a lag, because the concentration of the intermediate is zero at time zero. This hallmark of a sequential reaction is observed for the folding of HisF (Figure 8a). The lag in the time course for the native state is emphasized in the expanded view in Figure 8b. An equation describing a consecutive reaction was used for individual fits of the time courses of the intermediate and of the native state (Figure 8). The obtained rate constants were very similar to each other, showing that the intermediate is directly converted into the native form of HisF. Again, these rates coincided with those observed directly by fluorescence (Figure 8a). These findings complete the evidence that folding of HisF is sequential and that the intermediate is directly converted to the native protein.

Recovery of Protein Binding Competence during Refolding. The interrupted refolding experiments showed that molecules with the specific unfolding rate of the native state are formed in the slow folding phase. To confirm that correctly folded and functional HisF is recovered during this phase, refolding experiments were combined with a functional assay. In imidazole glycerol phosphate synthase, HisF (the cyclase subunit) and HisH (the glutaminase subunit) form a functional 1:1 complex.[25] The formation of this complex can be followed by a change in the fluorescence of the single Trp123 residue of HisH. This residue is solvent-exposed in the isolated protein but becomes buried upon binding to HisF (Figure 9a), which leads to a blue shift of its emission spectrum (Figure 9b). To avoid interference from the single Trp156 of HisF, the HisF-W156Y variant was used to follow the association of the

dx.doi.org/10.1021/bi300189f | Biochemistry 2012, 51, 3420−3432

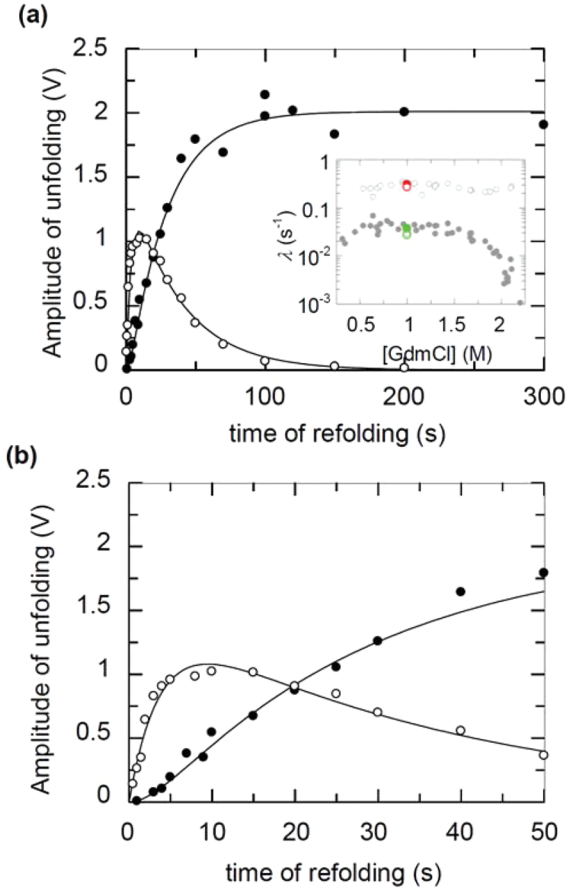


Figure 8. Time course of the unfolding intermediate and the native state of HisF determined by stopped-flow double-mixing experiments. Data were obtained as described in the legend of Figure 6, but unfolding was induced by dilution to 6.9 M GdmCl. A double-exponential function in combination with a linear term was fit to the observed fluorescence decays. The relative amplitudes of the fast [(empty circles) τ = 56 ms] and slow [(filled circles) τ = 2.6 s] phases, which characterize the unfolding of the intermediate and of the native state, respectively, are shown as a function of the duration of refolding in 1.0 M GdmCl (a) for the entire time range and (b) for the first 50 s of refolding. The solid lines represent a fit of an equation describing a consecutive process to the data.[31,32] In the inset of panel a, the obtained rate constants from the time course of I [(empty red circle) τ = 3.2 s, and (empty green circle) τ = 27 s] and from the time course of N [(filled red circle) τ = 3.7 s, and (filed green circle) τ = 36 s] are compared with the refolding rates determined by single-jump experiments (gray circles, data taken from Figure 5a).

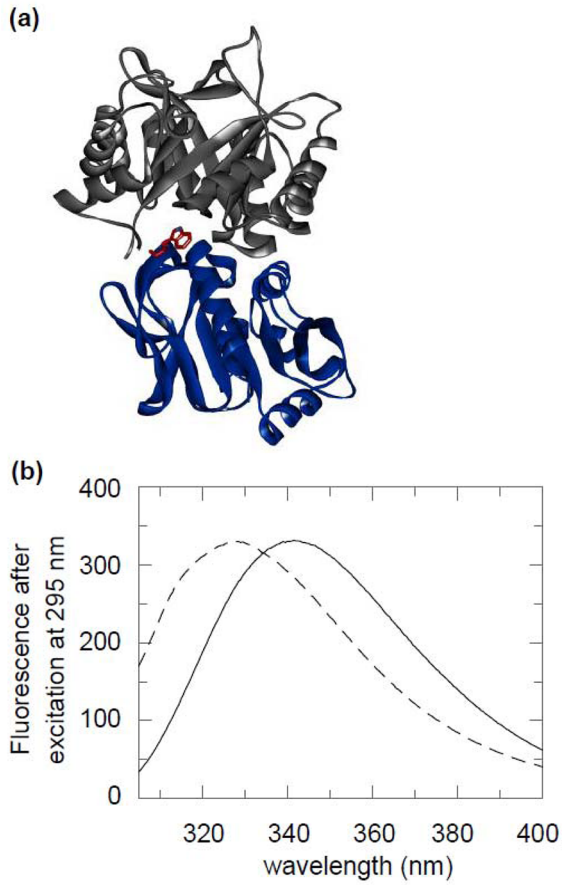


Figure 9. Stoichiometric complex of HisF and HisH that constitutes the imidazole glycerol phosphate synthase from *T. maritima*. (a) Ribbon diagram of the heterodimer (PDB entry 1GPW)[26] with HisH colored blue and HisF colored gray. The single Trp123 of HisH is displayed as red sticks. (b) The fluorescence emission of Trp123 from HisH (solid line, maximum at 341 nm) is blue-shifted upon binding to HisF-W156Y (dashed line, maximum at 327 nm). The fluorescence spectra of 2 μM protein were measured in 50 mM Tris-HCl buffer (pH 7.5) at 25 °C.

two proteins. The dissociation constant of the HisH−HisF-W156Y complex (K_D) was determined by equilibrium fluorescence titration to be ~5 nM (data not shown). The kinetics of association between 1 μM HisF and 1 μM HisH were measured in a stopped-flow apparatus. The apparent time constant of 0.07 s shows that the association is much more rapid than both the fast (50-fold) and slow (500-fold) refolding reaction of HisF. To determine the time course of the formation of binding-competent HisF upon refolding, double-mixing stopped-flow experiments were performed. Unfolded HisF-W156Y was first refolded in 1.0 M GdmCl for various times. Then the sample was mixed with HisH, and binding was followed by the decay of the fluorescence of Trp123 of HisH (Figure 10a). The observed amplitudes directly reflect the amount of binding-competent HisF present after various periods of refolding. A monoexponential equation was fit to the time-dependent increase in the amplitude of the binding reaction. The time constant of this reaction (τ = 40 s) is virtually identical with the time constant of the slow refolding phase (Figure 10b). These results show that the product of the rate-limiting slow folding step is native HisF. The fast folding phase (τ = 3.3 s) was not accompanied by a fluorescence decrease, suggesting that the folding intermediate I is not yet capable of binding to HisH. A lower limit for the dissociation constant (K_D) describing a putative binding of I to HisH can be estimated. After HisF had refolded for 2 s, half of the protein exists as the folding intermediate (Figure 7b), corresponding to 0.5 μM I in our binding assay. The maximal concentration of a potential complex between I and HisH can be estimated from the observation that the relative binding amplitude after refolding of HisF-W156Y for 2 s is <0.06 (Figure 10b), indicating that no more than 6% (<0.03 μM) of I could be involved in a complex with HisH. Consequently, the K_D for the binding of I to HisH must be larger than 17 μM, suggesting that the affinities of I and native HisF for HisH differ at least 1000-fold.

dx.doi.org/10.1021/bi300189f | *Biochemistry* 2012, 51, 3420−3432

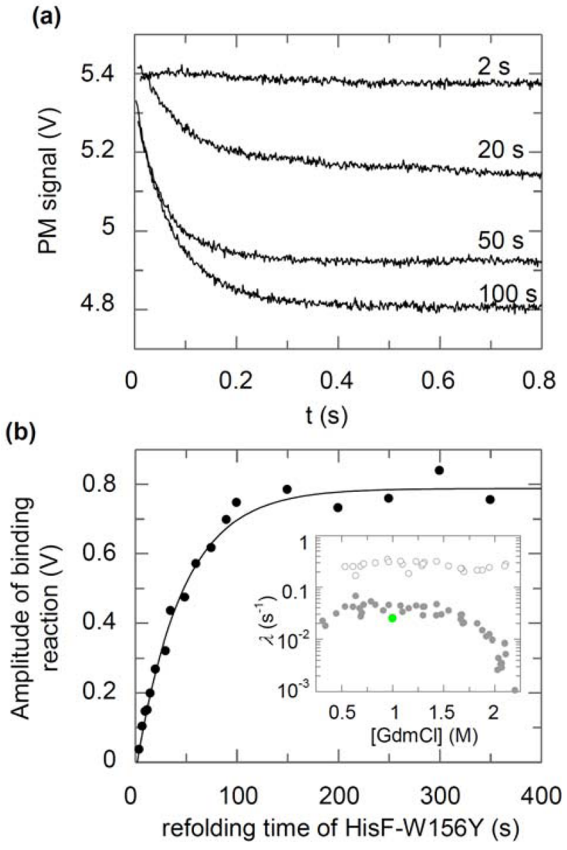


Figure 10. Kinetics of the formation of the HisF−HisH complex upon refolding of HisF determined by stopped-flow double-mixing experiments. (a) Refolding of unfolded HisF-W156Y was initiated by dilution to 1.0 M GdmCl. After various time intervals, HisF was mixed with HisH (final concentrations of 1.0 μM for each protein and 0.47 M GdmCl), and complex formation was followed by the decay of the fluorescence of W123 from HisH (excitation at 295 nm; emission at >320 nm). The experimental traces were fit with a monoexponential function. As expected, the time constant for complex formation (τ = 0.07 s) was independent of refolding time, whereas the binding amplitudes increased with time (shown are 2, 20, 50, and 100 s). (b) Amplitudes of the binding of HisF-W156Y to HisH as a function of refolding time in 1.0 M GdmCl. The solid line represents a fit of a single-exponential equation to the experimental points, yielding a time constant of 40 s. In the inset, this rate constant (green circle) is compared with the refolding rates determined by single-jump experiments (gray circles, data taken from Figure 5a). All experiments were conducted in 50 mM Tris-HCl buffer (pH 7.5) at 25 °C.

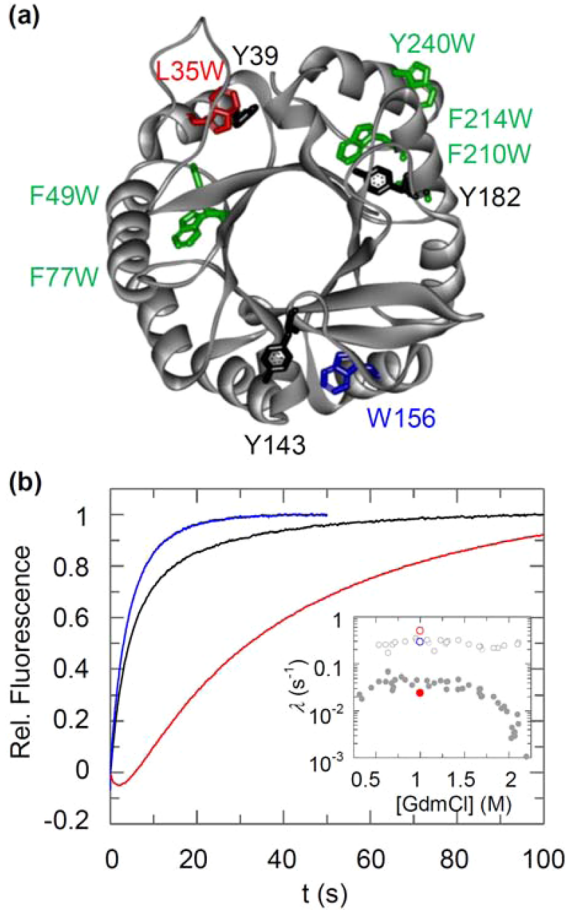


Figure 11. Monitoring refolding of HisF via individual Trp residues. (a) Ribbon diagram of the HisF structure (PDB entry 1THF)[23] with marked Tyr and Trp residues. The introduced Trp residues are colored red and green; native W156 is colored blue. (b) Refolding kinetics of wild-type W156-HisF and W35-HisF in 0.6 M GdmCl. Traces are shown for the Tyr/Trp fluorescence of wild-type W156-HisF (excitation at 280 nm; emission at >320 nm) (black), the Trp156 fluorescence of W156-HisF (excitation at 290 nm; emission at >320 nm) (blue), and the Trp35 fluorescence of W35-HisF (excitation at 290 nm; emission at >320 nm) (red). Monoexponential and biexponential functions were fit to the experimental curves. In the inset, the determined rate constants (λ) when monitoring the fluorescence of Trp156 [(empty blue circle) τ = 5.1 s] and Trp35 [(empty red circle) τ_1 = 2.2 s, and (filled red circle) τ_2 = 39 s] are compared with the refolding rates obtained by Tyr/Trp fluorescence [(empty gray circles) τ_1 = 4 s, and (filled gray circles) τ_2 = 25 s at 0.6 M GdmCl] taken from Figure 5a.

Following the Fluorescence of Individual Tryptophan Residues upon Refolding. To obtain the first evidence of the structure of the folding intermediate of HisF, we produced variants with Trp residues placed in different regions of the protein and analyzed their refolding reactions. Wild-type HisF contains only a single Trp (Trp156). It is located in α-helix 5 and monitors formation of tertiary structure in the central region of the protein chain. To follow formation of the tertiary structure in the N-terminal segment of HisF, we removed Trp from position 156 (by the W156Y substitution) and introduced a Trp into α-helix 1 at position 35 (by the L35W substitution). We call this variant "W35-HisF". Because HisF shows a pronounced 2-fold symmetry,[42] positions 35 and 156 in the N- and C-terminal half-barrels, respectively, are symmetry-related (Figure 11a). In the native state, Trp35 has a lower fluorescence quantum yield than Trp156 but the same emission maximum, and the spectrum becomes similarly red-shifted upon unfolding (Figure S5 of the Supporting Information). Equilibrium transitions at 45 °C showed that W35-HisF unfolds with cooperativity similar to that of wild-type HisF (W156-HisF) and is only slightly less stable (Figure S6 of the Supporting Information), suggesting that the folding mechanism is not changed by the shift of the reporter Trp from position 156 to position 35.

The kinetics of refolding of HisF, as monitored by the fluorescence of Trp156 in α-helix 5 or Trp35 in α-helix 1 (excitation at 290 nm; emission at >320 nm) at 0.6 M GdmCl, are compared in Figure 11b with the kinetics of refolding of wild-type HisF (W156-HisF) monitored by the fluorescence of both Trp156 and the four Tyr residues (excitation at 280 nm;

dx.doi.org/10.1021/bi300189f | Biochemistry 2012, 51, 3420−3432

emission at >320 nm). In accordance with Figure 5a, the Tyr/Trp fluorescence of W156-HisF monitors the formation of the intermediate (τ = 4 s) and its conversion to the native state (τ = 25 s). In contrast, when refolding of wild-type HisF was selectively followed by the fluorescence of Trp156, only the fast reaction with a time constant of 5.1 s was observed. Apparently, Trp156 in α-helix 5 is already in a native-like environment in the folding intermediate I and does not undergo further changes during the slow (rate-determining) transition from I to the native state N. This finding suggests that the central part of HisF adopts already a native-like tertiary structure in I, in which Trp156 is shielded from the solvent as efficiently as in N. When the refolding of W35-HisF was selectively followed by the fluorescence of Trp35, the signal was dominated by the slow reaction with a time constant of 39 s (Figure 11b). This result indicates that Trp35 in α-helix 1 is largely in an unfolded-like environment in the intermediate I and that the tertiary structure surrounding it is mainly formed upon transition from I to the native state N. Additionally, a fluorescence phase with a negative amplitude and a time constant of 2.2 s was observed. The latter agrees with the fast folding rate observed by CD and Tyr/Trp fluorescence (Figure 11b), but the relative amplitude is less than 5%. Taken together, our results suggest that the folding intermediate I of HisF contains native-like tertiary structure in its central part but is partly unfolded in its N-terminal region. To gain more insight into the tertiary structure content of I, we moved the single Trp residue also to positions 49, 77, 210, 214, and 240 (Figure 11a) and measured their refolding kinetics (excitation at 290 nm; emission at >320 nm). All these Trp residues detected both folding phases, suggesting that, at these positions, the folding intermediate is not yet nativelike (Table S1 of the Supporting Information).

DISCUSSION

Thermodynamic Stability. The GdmCl-induced equilibrium unfolding of HisF from the hyperthermophilic bacterium *T. maritima* is described well by the two-state mechanism. Remarkably, it took ~3 weeks at 45 °C and ~6 weeks at 25 °C to reach the folding equilibrium, which is reflected in excessively slow folding kinetics in the transition region. Importantly, HisF unfolds also very slowly in the physiological temperature range of *T. maritima*, which allows for the purification of the recombinant protein by heat incubation of the *E. coli* host cell extract.[28] Extremely slow unfolding kinetics have also been observed for other thermostable proteins,[19−22] and the properties of HisF further support the notion that a high kinetic barrier to unfolding provides a major mechanism for protecting proteins from hyperthermophiles against thermal denaturation.[43]

Folding Mechanism. In refolding, HisF forms a burst-phase intermediate I_{BP} with a significant amount of secondary structure. Then, in a sequential folding reaction, first a folding intermediate (I) is formed that is converted into the native state in the final, rate-limiting step of folding. Interrupted refolding experiments showed that the intermediate I unfolded much faster than the native protein and that folding is a sequential process with I as an obligatory on-pathway intermediate and a lag in the formation of the native protein. Both intermediates I_{BP} and I are less stable than N and not populated in the transition region. The on-pathway intermediate I cannot bind to HisH, which forms a functional complex with HisF. I_{BP} is formed within the dead time of refolding. The pronounced roll-over observed for the fast folding phase below 1.0 M GdmCl (Figure 5a), resulting in a positive m value of refolding, indicates that I_{BP} is off-pathway and that non-native interactions within I_{BP} have to be broken before I can form. However, the alternative of I_{BP} being on-pathway cannot be excluded. At >2 M GdmCl, both folding and unfolding of HisF can be described by a single phase (Figure 5), most probably because I_{BP} and I are destabilized by the denaturant and do not accumulate.[44] As a consequence, our data are in accordance with a linear four-species model (Scheme 1).

Scheme 1. Linear Four-Species Model for HisF Folding[a]

$$I_{BP} \rightleftharpoons U \rightleftharpoons I \rightleftharpoons N$$

[a]An off-pathway equilibrium between the unfolded state (U) and the burst-phase intermediate (I_{BP}) precedes the formation of the on-pathway intermediate (I) and the native state (N). At >2 M GdmCl, I_{BP} and I are too unstable to accumulate to detectable concentrations. Although it is improbable, our data cannot unambiguously exclude the possibility that I_{BP} is an on-pathway intermediate that exists in an equilibrium with I.

Structural Properties of the Burst-Phase Intermediate I_{BP} and the On-Pathway Intermediate I. Our experimental data provide information about the structural properties of the folding intermediates I_{BP} and I. The formation of the burst-phase intermediate I_{BP} in the first few milliseconds of refolding is accompanied by a large change in the far-UV CD signal in the dead time of stopped-flow mixing. The magnitude of this CD signal decreases in a cooperative manner with an increasing concentration of GdmCl, and the resulting "unfolding" curve suggests that I_{BP} is significantly less stable than the native protein. I_{BP} contains a high content of secondary structure but not a defined tertiary structure, as evidenced by a fluorescence that resembles that of the unfolded state U. In line with this finding, I_{BP} efficiently binds the dye ANS (Figure S3a of the Supporting Information), which is known to have a high affinity for solvent-exposed hydrophobic patches.[35,36] I_{BP} is probably a molten globule-like state[45] that forms rapidly when the unfolded protein is exposed to native conditions.

The on-pathway intermediate I binds weaker to ANS than I_{BP} (Figure S3b of the Supporting Information), suggesting that in I the hydrophobic side chains are shielded more efficiently from the solvent. The formation of I is accompanied by further changes in far-UV CD and Trp/Tyr fluorescence, which indicates that it contains both secondary and tertiary structure. Analysis of refolding by following the fluorescence of engineered Trp residues at various positions in HisF shows that tertiary structure formation differs at different positions. It seems more complete in the central region of the protein chain than in the N-terminal region (Figure 11). Also, the HisF folding intermediate cannot associate with its functional partner HisH (Figure 10), indicating that the binding interface is not yet formed. This finding is consistent with the crystal structure of the HisH−HisF complex demonstrating that the N-terminal segment of HisF is required for docking to HisH.[26]

Comparison of the Folding Mechanisms of HisF and Other $(\beta\alpha)_8$-Barrel Proteins. It is assumed that the shape of the folding landscape of a protein is determined by its chain topology and by individual sequence patterns.[46−50] To evaluate the relevance of these factors for the folding of $(\beta\alpha)_8$-barrels, we compared the folding properties of HisF with those of other $(\beta\alpha)_8$-barrel proteins, αTS,[13,14] IGPS,[15−17] and IOLI.[18] Unlike HisF, which displays two-state equilibrium unfolding, these

dx.doi.org/10.1021/bi300189f | Biochemistry 2012, 51, 3420−3432

three proteins form one or two equilibrium intermediates that contain ~50% of the far-UV CD signal of the native state. These proteins are much less stable than HisF from *T. maritima* and can be unfolded by the weak denaturant urea. Obviously, folded HisF contains additional stabilizing interactions, and therefore, >2.5 M GdmCl is required for its unfolding, conditions under which partially folded forms of the protein are destabilized already. Thus, the strong difference in stability between HisF and the other $(\beta\alpha)_8$-barrel proteins explains probably why equilibrium unfolding intermediates are not populated in the unfolding transition of HisF.

Despite these differences in equilibrium unfolding and despite the low overall level of sequence identity of <20% [as detected by a pairwise comparison with DALI-lite (http://www.ebi.ac.uk/Tools/dalilite/)], HisF resembles the three other analyzed $(\beta\alpha)_8$-barrel proteins in the kinetic folding mechanism. In all cases, a burst-phase intermediate I_{BP} is formed,[14,17,18] which is marginally stable but contains a high fraction of the far-UV CD signal of the native state (HisF, 71%; IGPS, 46%; αTS, 47%; IOLI, 37%). Global folding analysis or Gō-model simulation has suggested that the I_{BP} forms of αTS,[14,51] IGPS,[17] and IOLI[18] are off-pathway species, meaning that the unfolding of their non-native structural elements controls access to the productive on-pathway intermediates. Our kinetic data for HisF suggest that its I_{BP} is also an off-pathway intermediate. Substantial burst-phase reactions have also been observed for the $(\beta\alpha)_8$-barrel proteins fructose-1,6-bisphosphate aldolase[52] and PRAI,[53] as well as for the $(\beta\alpha)_5$-barrel proteins CheY[54] and Spo0F.[55] The latter were proposed to be evolutionarily related to $(\beta\alpha)_8$-barrels via a common $(\beta\alpha)_4$-half-barrel precursor.[56,57] The accumulation of non-productive burst-phase intermediates in the folding of $(\beta\alpha)_8$-proteins might be linked with the complexity of this structure and the concomitant slowness of its formation. Thus, rapid alternative folding paths that lead to the exclusion of solvent from side chains and from the main chain and to secondary structure formation compete with productive folding.[16,17,51,54,55]

At first glance, the chevron diagrams show different degrees of complexity. In particular, αTS folds along four parallel folding channels, which was explained by the cis−trans isomerization reactions of three proline residues.[58−60] HisF, IGPS, and IOLI contain only trans prolyl peptide bonds and show simpler folding mechanisms.[17,18] As an explicit similarity, all four proteins show roll-over phenomena at low denaturant concentrations and a strong decrease of the folding rates at moderate denaturant concentrations.

In addition to the off-pathway intermediate I_{BP}, one or two on-pathway kinetic intermediates are formed on the folding pathways of these $(\beta\alpha)_8$-barrel proteins.[14,15,18,61] How well conserved are the structures of these productive folding intermediates? PRAI, IGPS, and αTS catalyze three consecutive steps in tryptophan biosynthesis and contain a conserved phosphate binding site in the C-terminal half of their protein chain, and most probably, they diverged from a common ancestor by several gene duplications.[24,62] Early fragment sudies with PRAI suggested a common "6+2" model for the folding for these proteins, in which a prefolded $(\beta\alpha)_{1-6}$ module serves as a template for the association and folding of the remaining $(\beta\alpha)_{7-8}$ moiety.[10] Hydrogen exchange mass spectrometry (HX-MS) experiments with IGPS indicated that productive folding begins with the formation of an intermediate (termed I_a) that comprises the central $\beta_2(\beta\alpha)_{3-5}\beta_6$ segment. Subsequently, the adjacent α_1, α_6 and β_7 segments fold, yielding the intermediate I_b, and finally, the terminal segments α_0, α_7, β_8, and α_8 are structured to form the complete $(\beta\alpha)_8$-barrel (N).[16,17] These findings point to a "6+2" mechanism for the formation of the $(\beta\alpha)_8$ structure. For αTS, however, HX protection experiments rather suggest that the N-terminal half-barrel $(\beta\alpha)_{1-4}$ is formed in the rapidly formed folding intermediate.[63] This is reminiscent of the "4+4" folding model that was suggested for HisF, based on the finding that a catalytically active HisF−NC complex could be reconstituted in the joint refolding of the isolated N-terminal (HisF-N) and C-terminal (HisF-C) half-barrels.[42] Our refolding experiments with the individual Trp variants of HisF, however, suggest that its intermediate rather consists of a folded central region and unfolded termini (Figure 11), which resembles more the intermediate of IGPS than that of αTS.

How well conserved are the rates for the formation and for the unfolding of the $(\beta\alpha)_8$-barrel structure in these proteins? A direct comparison of the unfolding and refolding rates of the four $(\beta\alpha)_8$-barrel proteins HisF, αTS, IGPS, and IOLI is complicated by the fact that different denaturants, GdmCl and urea, were used. However, the rate-limiting steps of refolding of the four proteins at low denaturant concentrations are very similar. Despite the low level of sequence similarity and the strong differences in stability, they all show time constants of ~10 s, supporting the view that chain topology is an important determinant for refolding of $(\beta\alpha)_8$-barrels.

Strong differences are observed, however, for the unfolding rates of the four $(\beta\alpha)_8$-barrel proteins when compared at equal distances from the midpoint of the transition curve, that is at equal stability. Unfolding of HisF from the hyperthermophilic organism *T. maritima* is at least 1000-fold slower than the unfolding of αTS and IOLI, which are from mesophilic organisms. Remarkably, unfolding of HisF is also ~100-fold slower than unfolding of truncated IGPS, which originates also from a hyperthermophilic organism (*S. solfataricus*) but shows a lower thermodynamic stability.[15]

These findings suggest that for refolding, the proteins from hyperthermophiles use the same mechanism as their counterparts from mesophilic organisms. The stabilizing amino acid substitutions exert their influence late on the folding pathway, after the passage through the transition state, by selectively stabilizing the native state relative to the transition state. The increased equilibrium stability of these proteins is thus used entirely for a maximal increase in their kinetic stability.

■ ASSOCIATED CONTENT

Supporting Information

Figures showing GdmCl-induced equilibrium unfolding and refolding transitions of HisF (Figure S1), unfolding and refolding kinetics followed by fluorescence of Trp/Tyr residues (Figure S2) and bound ANS (Figure S3), unfolding of the intermediate in an interrupted refolding experiment (Figure S4), fluorescence emission spectra of W156-HisF and W35-HisF (Figure S5), and the equilibrium refolding transition of W35-HisF (Figure S6) and relative amplitudes and rate constants of refolding calculated from the fluorescence signal of different single Trp residues (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +49-941-943 3015. Fax: +49-941-943 2813. E-mail: reinhard.sterner@biologie.uni-regensburg.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Jochen Reinstein (Max Planck Institute for Medical Research, Heidelberg, Germany) for stopped-flow CD measurements.

ABBREVIATIONS

HisF, cyclase subunit of imidazole glycerol phosphate synthase (ImGPS) from *T. maritima*; HisH, glutaminase subunit of ImGPS; N, native state; U, unfolded state; I_{BP}, burst-phase intermediate; I, kinetic on-pathway intermediate; W156-HisF, wild-type HisF containing the single tryptophan residue at position 156; W35-HisF, HisF variant containing the W156Y and L35W amino acid exchanges; αTS, α-subunit of tryptophan synthase from *E. coli*; IGPS, indole-3-glycerol phosphate synthase from *S. solfataricus*; IOLI, protein of unknown function encoded by the *B. subtilis iol*I gene; PDB, Protein Data Bank.

REFERENCES

(1) Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol. 247*, 536−540.

(2) Caetano-Anollés, G., Kim, H. S., and Mittenthal, J. E. (2007) The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. *Proc. Natl. Acad. Sci. U.S.A. 104*, 9358−9363.

(3) Wierenga, R. K. (2001) The TIM-barrel fold: A versatile framework for efficient enzymes. *FEBS Lett. 492*, 193−198.

(4) Sterner, R., and Höcker, B. (2005) Catalytic versatility, stability, and evolution of the $(\beta\alpha)_8$-barrel enzyme fold. *Chem. Rev. 105*, 4038−4055.

(5) Nagano, N., Orengo, C. A., and Thornton, J. M. (2002) One fold with many functions: The evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol. 321*, 741−765.

(6) Pujadas, G., and Palau, J. (1999) TIM barrel fold: Structural, functional and evolutionary characteristics in natural and designed molecules. *Biologica (Bratislava, Slovakia) 54*, 231−253.

(7) Raine, A. R., Scrutton, N. S., and Mathews, F. S. (1994) On the evolution of alternate core packing in eightfold $(\beta\alpha)$-barrels. *Protein Sci. 3*, 1889−1892.

(8) Reardon, D., and Farber, G. (1995) The structure and evolution of $\alpha\beta$-barrel proteins. *FASEB J. 9*, 497−503.

(9) Copley, R. R., and Bork, P. (2000) Homology among $(\beta\alpha)_8$-barrels: Implications for the evolution of metabolic pathways. *J. Mol. Biol. 303*, 627−641.

(10) Eder, J., and Kirschner, K. (1992) Stable substructures of eightfold $(\beta\alpha)_8$-barrel proteins: Fragment complementation of phosphoribosylanthranilate isomerase. *Biochemistry 31*, 3617−3625.

(11) Soberón, X., Fuentes-Gallego, P., and Saab-Rincón, G. (2004) In vivo fragment complementation of a $(\beta\alpha)_8$-barrel protein: Generation of variability by recombination. *FEBS Lett. 560*, 167−172.

(12) Bertolaet, B. L., and Knowles, J. R. (1995) Complementation of fragments of triosephosphate isomerase defined by exon boundaries. *Biochemistry 34*, 5736−5743.

(13) Gualfetti, P. J., Bilsel, O., and Matthews, C. R. (1999) The progressive development of structure and stability during the equilibrium folding of the α-subunit of tryptophan synthase from *Escherichia coli*. *Protein Sci. 8*, 1623−1635.

(14) Bilsel, O., Zitzewitz, J. A., Bowers, K. E., and Matthews, C. R. (1999) Folding mechanism of the α-subunit of tryptophan synthase, an $\beta\alpha$-barrel protein: Global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry 38*, 1018−1029.

(15) Forsyth, W. R., and Matthews, C. R. (2002) Folding mechanism of indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*: A test of the conservation of folding mechanisms hypothesis in $(\beta\alpha)_8$-barrels. *J. Mol. Biol. 320*, 1119−1133.

(16) Gu, Z., Zitzewitz, J. A., and Matthews, C. R. (2007) Mapping the structure of folding cores in TIM barrel proteins by hydrogen exchange mass spectrometry: The roles of motif and sequence for the indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*. *J. Mol. Biol. 368*, 582−594.

(17) Gu, Z., Rao, M. K., Forsyth, W. R., Finke, J. M., and Matthews, C. R. (2007) Structural analysis of kinetic folding intermediates for a TIM barrel protein, indole-3-glycerol phosphate synthase, by hydrogen exchange mass spectrometry and Gō model simulation. *J. Mol. Biol. 374*, 528−546.

(18) Forsyth, W. R., Bilsel, O., Gu, Z., and Matthews, C. R. (2007) Topology and sequence in the folding of a TIM barrel protein: Global analysis highlights partitioning between transient off-pathway and stable on-pathway folding intermediates in the complex folding mechanism of a $(\beta\alpha)_8$-barrel of unknown function from *B. subtilis*. *J. Mol. Biol. 372*, 236−253.

(19) Mukaiyama, A., Takano, K., Haruki, M., Morikawa, M., and Kanaya, S. (2004) Kinetically robust monomeric protein from a hyperthermophile. *Biochemistry 43*, 13859−13866.

(20) Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M., Janicke, R., and Schmid, F. X. (1998) Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol. 5*, 229−235.

(21) Dams, T., and Jaenicke, R. (1999) Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochemistry 38*, 9169−9178.

(22) Deutschman, W., and Dahlquist, F. (2001) Thermodynamic basis for the increased thermostability of CheY from the hyperthermophile *Thermotoga maritima*. *Biochemistry 40*, 13107−13113.

(23) Lang, D., Thoma, R., Henn-Sax, M., Sterner, R., and Wilmanns, M. (2000) Structural evidence for evolution of the $\beta\alpha$-barrel scaffold by gene duplication and fusion. *Science 289*, 1546−1550.

(24) List, F., Sterner, R., and Wilmanns, M. (2011) Related $(\beta\alpha)_8$-barrel proteins in histidine and tryptophan biosynthesis: A paradigm to study enzyme evolution. *ChemBioChem 12*, 1487−1494.

(25) Beismann-Driemeyer, S., and Sterner, R. (2001) Imidazole glycerol phosphate synthase from *Thermotoga maritima*: Quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex. *J. Biol. Chem. 276*, 20387−20396.

(26) Douangamath, A., Walker, M., Beismann-Driemeyer, S., Vega-Fernandez, M. C., Sterner, R., and Wilmanns, M. (2002) Structural evidence for ammonia tunneling across the $(\beta\alpha)_8$-barrel of the imidazole glycerol phosphate synthase bienzyme complex. *Structure 10*, 185−193.

(27) Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*, 51−59.

(28) Thoma, R., Obmolova, G., Lang, D. A., Schwander, M., Jenö, P., Sterner, R., and Wilmanns, M. (1999) Efficient expression, purification and crystallisation of two hyperthermostable enzymes of histidine biosynthesis. *FEBS Lett. 454*, 1−6.

(29) Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol. 131*, 266−280.

(30) Santoro, M. M., and Bolen, D. W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry 27*, 8063−8068.

(31) McDaniel, D. H., and Smoot, C. R. (1956) Approximations in the kinetics of consecutive reactions. *J. Phys. Chem. 60*, 966−969.
(32) Casado, J., González, J., and Moreno, M. (1987) Consecutive reactions in chemical kinetics: A method for the treatment of experimental data. *React. Kinet. Catal. Lett. 33*, 357−362.
(33) Myers, J. K., Pace, C. N., and Martin Scholtz, J. (1995) Denaturant *m*-values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci. 4*, 2138−2148.
(34) Silow, M., and Oliveberg, M. (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6086.
(35) Stryer, L. (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J. Mol. Biol. 13*, 482−495.
(36) Jones, B. E., Jennings, P. A., Pierre, R. A., and Matthews, C. R. (1994) Development of nonpolar surfaces in the folding of *Escherichia coli* dihydrofolate reductase detected by 1-anilinonaphthalene-8-sulfonate binding. *Biochemistry 33*, 15250−15258.
(37) Lorenz, T., and Reinstein, J. (2008) The influence of proline isomerization and off-pathway intermediates on the folding mechanism of eukaryotic UMP/CMP kinase. *J. Mol. Biol. 381*, 443−455.
(38) Hammond, G. S. (1955) A correlation of reaction rates. *J. Am. Chem. Soc. 77*, 334−338.
(39) Oliveberg, M., Tan, Y.-J., Silow, M., and Fersht, A. R. (1998) The changing nature of the protein folding transition state: Implications for the shape of the free-energy profile for folding. *J. Mol. Biol. 277*, 933−943.
(40) Silow, M., and Oliveberg, M. (1997) High-energy channeling in protein folding. *Biochemistry 36*, 7633−7637.
(41) Zoldak, G., Carstensen, L., Scholz, C., and Schmid, F. X. (2009) Consequences of domain insertion on the stability and folding mechanism of a protein. *J. Mol. Biol. 386*, 1138−1152.
(42) Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A., and Sterner, R. (2001) Dissection of a $(\beta\alpha)_8$-barrel enzyme into two folded halves. *Nat. Struct. Biol. 8*, 32−36.
(43) Sterner, R., and Liebl, W. (2001) Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol. 36*, 39−106.
(44) Bachmann, A., and Kiefhaber, T. (2001) Apparent two-state Tendamistat folding is a sequential process along a defined route. *J. Mol. Biol. 306*, 375−386.
(45) Ptitsyn, O. B. (1995) Molten globule and protein folding. *Adv. Protein Chem. 47*, 83−229.
(46) Wensley, B. G., Gärtner, M., Choo, W. X., Batey, S., and Clarke, J. (2009) Different members of a simple three-helix bundle protein family have very different folding rate constants and fold by different mechanisms. *J. Mol. Biol. 390*, 1074−1085.
(47) Friel, C. T., Capaldi, A. P., and Radford, S. E. (2003) Structural analysis of the rate-limiting transition states in the folding of lm7 and lm9: Similarities and differences in the folding of homologous proteins. *J. Mol. Biol. 326*, 293−305.
(48) Olofsson, M., Hansson, S., Hedberg, L., Logan, D. T., and Oliveberg, M. (2007) Folding of S6 structures with divergent amino acid composition: Pathway flexibility within partly overlapping foldons. *J. Mol. Biol. 365*, 237−248.
(49) Lappalainen, I., Hurley, M. G., and Clarke, J. (2008) Plasticity within the obligatory folding nucleus of an immunoglobulin-like domain. *J. Mol. Biol. 375*, 547−559.
(50) Lam, A. R., Borreguero, J. M., Ding, F., Dokholyan, N. V., Buldyrev, S. V., Stanley, H. E., and Shakhnovich, E. (2007) Parallel folding pathways in the SH3 domain protein. *J. Mol. Biol. 373*, 1348−1360.
(51) Wu, Y., Vadrevu, R., Kathuria, S., Yang, X., and Matthews, C. R. (2007) A tightly packed hydrophobic cluster directs the formation of an off-pathway sub-millisecond folding intermediate in the α subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol. 366*, 1624−1638.
(52) Rudolph, R., Siebendritt, R., and Kiefhaber, T. (1992) Reversible unfolding and refolding behavior of a monomeric aldolase from *Staphylococcus aureus*. *Protein Sci. 1*, 654−666.
(53) Jasanoff, A., Davis, B., and Fersht, A. R. (1994) Detection of an intermediate in the folding of the $(\beta\alpha)_8$-barrel N-(5′-phosphoribosyl)-anthranilate isomerase from *Escherichia coli*. *Biochemistry 33*, 6350−6355.
(54) Kathuria, S. V., Day, I. J., Wallace, L. A., and Matthews, C. R. (2008) Kinetic traps in the folding of $\beta\alpha$-repeat proteins: CheY initially misfolds before accessing the native conformation. *J. Mol. Biol. 382*, 467−484.
(55) Hills, R. D., Jr., Kathuria, S. V., Wallace, L. A., Day, I. J., Brooks, C. L., III, and Matthews, C. R. (2010) Topological frustration in $\beta\alpha$-repeat proteins: Sequence diversity modulates the conserved folding mechanisms of $\alpha/\beta/\alpha$ sandwich proteins. *J. Mol. Biol. 398*, 332−350.
(56) Höcker, B., Schmidt, S., and Sterner, R. (2002) A common evolutionary origin of two elementary enzyme folds. *FEBS Lett. 510*, 133−135.
(57) Bharat, T. A. M., Eisenbeis, S., Zeth, K., and Höcker, B. (2008) A $\beta\alpha$-barrel built by the combination of fragments from different folds. *Proc. Natl. Acad. Sci. U.S.A. 105*, 9942−9947.
(58) Wu, Y., and Matthews, C. R. (2002) Parallel channels and rate-limiting steps in complex protein folding reactions: Prolyl isomerization and the α-subunit of Trp synthase, a TIM barrel protein. *J. Mol. Biol. 323*, 309−325.
(59) Wu, Y., and Matthews, C. R. (2002) A *cis*-prolyl peptide bond isomerization dominates the folding of the α-subunit of trp synthase, a TIM barrel protein. *J. Mol. Biol. 322*, 7−13.
(60) Wu, Y., and Matthews, C. R. (2003) Proline replacements and the simplification of the complex, parallel channel folding mechanism for the α-subunit of Trp synthase, a TIM barrel protein. *J. Mol. Biol. 330*, 1131−1144.
(61) Saab-Rincon, G., Froebe, C. L., and Matthews, C. R. (1993) Urea-induced unfolding of the α-subunit of tryptophan synthase: One-dimensional proton NMR evidence for residual structure near histidine-92 at high denaturant concentration. *Biochemistry 32*, 13981−13990.
(62) Wilmanns, M., Hyde, C. C., Davies, D. R., Kirschner, K., and Jansonius, J. N. (1991) Structural conservation in parallel $(\beta\alpha)$-barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry 30*, 9161−9169.
(63) Rojsajjakul, T., Wintrode, P., Vadrevu, R., Robert Matthews, C., and Smith, D. L. (2004) Multi-state unfolding of the α-subunit of tryptophan synthase, a TIM barrel protein: Insights into the secondary structure of the stable equilibrium intermediates by hydrogen exchange mass spectrometry. *J. Mol. Biol. 341*, 241−253.