

Clusters of Branched Aliphatic Side Chains Serve As Cores of Stability in the Native State of the HisF TIM Barrel Protein

Basavanapura N. Gangadhara, Jennifer M. Laine, Sagar V. Kathuria, Francesca Massi and C. Robert Matthews

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605, USA

Correspondence to Francesca Massi and C. Robert Matthews: francesca.massi@umassmed.edu;

c.robert.matthews@umassmed.edu

<http://dx.doi.org/10.1016/j.jmb.2013.01.002>

Edited by S. Marqusee

Abstract

Imidazole-3-glycerol phosphate synthase is a heterodimeric allosteric enzyme that catalyzes consecutive reactions in imidazole biosynthesis through its HisF and HisH subunits. The unusually slow unfolding reaction of the isolated HisF TIM barrel domain from the thermophilic bacteria, *Thermotoga maritima*, enabled an NMR-based site-specific analysis of the main-chain hydrogen bonds that stabilize its native conformation. Very strong protection against exchange with solvent deuterium in the native state was found in a subset of buried positions in α -helices and pervasively in the underlying β -strands associated with a pair of large clusters of isoleucine, leucine and valine (ILV) side chains located in the $\alpha_7(\beta\alpha)_8(\beta\alpha)_{1-2}$ and $\alpha_2(\beta\alpha)_{3-6}\beta_7$ segments of the $(\beta\alpha)_8$ barrel. The most densely packed region of the large cluster, $\alpha_3(\beta\alpha)_{4-6}\beta_7$, correlates closely with the core of stability previously observed in computational, protein engineering and NMR dynamics studies, demonstrating a key role for this cluster in determining the thermodynamic and structural properties of the native state of HisF. When considered with the results of previous studies where ILV clusters were found to stabilize the hydrogen-bonded networks in folding intermediates for other TIM barrel proteins, it appears that clusters of branched aliphatic side chains can serve as cores of stability across the entire folding reaction coordinate of one of the most common motifs in biology.

© 2013 Elsevier Ltd. All rights reserved.

Introduction

The $(\beta\alpha)_8$ TIM barrel motif one of the most common folds in biology^{1,2} has been the subject of numerous studies of its structure,²⁻⁴ function,⁵⁻⁸ folding,⁹⁻¹² design¹³⁻²¹ and evolution.^{2,22-26} Evolution presumably recapitulated the $\beta\alpha$ module through gene duplication until its 8-fold manifestation resulted in a stable closed barrel with β_1 hydrogen bonded to β_8 and intervening α -helices forming a continuous amphipathic shell around the hydrophobic β -barrel.²⁴ The short loops/turns between the α -helices and the β -strands at one end of the barrel are crucial for stability.^{27,28} The longer loops between the β -strands and the subsequent α -helices at the opposite end of the barrel invariably form the

active site for a host of enzymes from all three superkingdoms.^{29,30}

Although the TIM barrel motif is recognized as a single structural domain, studies of its folding mechanism have revealed the presence of an early off-pathway intermediate and a pair of on-pathway intermediates for four barrels of very low sequence identity.^{9,31-33} Evidently, the conserved topology dictates the basic features of the folding free-energy surface, and the sequence modulates the energies of the various species and intervening transition-state ensembles (TSEs).³² Structural analyses of the intermediates for the α TS²⁸ and sIGPS³⁴ barrels employing hydrogen exchange (HX) methods revealed that the networks of main-chain amide hydrogens protected against

exchange with solvent were not conserved. Rather, the patterns of protection in the partially folded states varied with the location of large clusters of isoleucine, leucine and valine (ILV) side chains. We had previously hypothesized that clusters of branched aliphatic side chains, whose side-chain analogs preferentially partition into the vapor phase *versus* water,³⁵ would be especially resistant to the penetration of water and/or hydroxide required for HX.³⁶ We concluded that clusters of ILV side chains served as cores of stability in these partially folded states. Unfortunately, the time constraints of the HSQC (heteronuclear single-quantum coherence) NMR experiment required for site-specific exchange measurements, ~30 min for data collection, and the comparable lifetime of the native state to unfolding at neutral pH and room temperature precluded a test of the role of ILV clusters in the HX properties of the native state of the α TS TIM barrel.

The HisF TIM barrel subunit of the heterodimeric *Thermotoga maritima* imidazole-3-glycerol phosphate synthase (ImGPS) by contrast proves to be an especially favorable candidate to explore sequence-specific contributions to the HX properties of its native state. Lang *et al.* and Richter *et al.* have examined the evolution of the HisF barrel and concluded that the $(\beta\alpha)_8$ motif arose from a $(\beta\alpha)_4 + (\beta\alpha)_4$ gene duplication event.^{22,37} Subsequent studies revealed differential stability of the N- and C-terminal $(\beta\alpha)_4$ modules of the HisF barrel,³⁸ supporting this view and suggesting that differential HX patterns might persist in these modules in the native state. More recently, these workers have examined the kinetic folding mechanism of HisF and reported an exceedingly large kinetic stability,³³ equivalent to a very long native-state lifetime, that would enable a site-specific HX examination of the essential hydrogen-bonding network in the native thermodynamic state^{39,40} via NMR spectroscopy. Relevant to the present study, a follow-up protein engineering effort revealed a core of stability in the $(\beta\alpha)_{4-7}$ segment of the HisF TIM barrel.⁴¹

Similar to their roles in stabilizing the hydrogen-bonding networks in folding intermediates of TIM barrel proteins,^{28,34,36} we report that two large ILV clusters preferentially protect networks of main-chain amide hydrogen bonds (H-bonds) from exchange through native-state dynamics in HisF. The evolution of the sequence since the gene duplication event appears to have resulted in clusters that are not simply associated with either the N- and/or C-terminal halves of HisF. The larger ILV cluster, spanning $\alpha_2(\beta\alpha)_{3-6}\beta_7$, is closely associated with the $(\beta\alpha)_{4-7}$ segment and provides a rationale for its role as a core of stability in the native state of HisF.

Results

Equilibrium and kinetic folding mechanisms

As a prelude to the HX analysis of HisF, it was necessary to determine its equilibrium and kinetic folding properties under the conditions selected for the NMR experiment, 40 °C and pH 7.2. Guanidine hydrochloride (Gdn-HCl) induced equilibrium denaturation experiments monitored the disruption of secondary structure by far UV-CD spectroscopy. The reversibility was tested by comparing the transition curves generated from incubating the native state with increasing concentrations of denaturant with those generated by incubating the denatured state with decreasing concentrations of denaturant. Surprisingly, the CD-detected unfolding reaction required ~18 days at pH 7.2 and 40 °C to equilibrate (Fig. 1a). By contrast, the refolding reaction equilibrated within 24 h (Fig. 1b). Although all of the data are consistent with a cooperative two-state unfolding transition, the contrasting responses in the approach to equilibrium imply the presence of

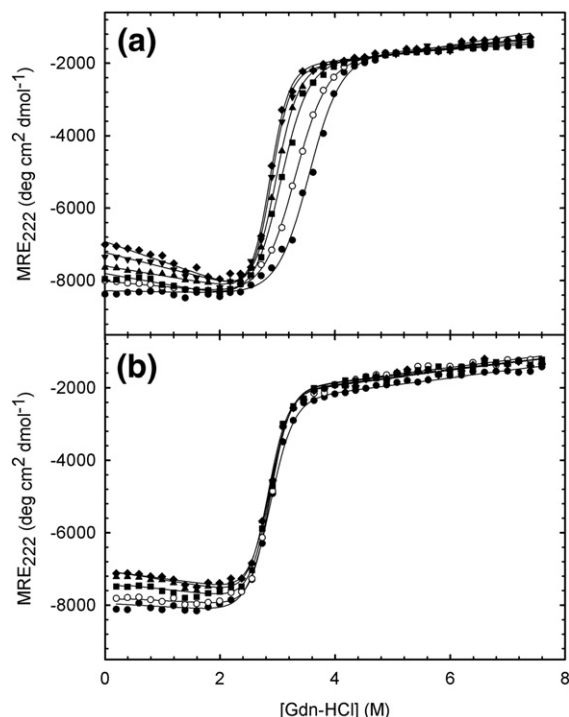


Fig. 1. Chemical denaturation curves of HisF monitored by mean residue ellipticity at 222 nm and plotted as a function of Gdn-HCl concentration in a buffer containing 10 mM KPi, 1 mM K₂EDTA and 0.5 mM DTT at pH 7.2 and 40 °C. (a) Unfolding equilibrium: 1 day (●), 2 days (○), 4 days (■), 10 days (▲), 13 days (▼) and 18 days (◆). (b) Refolding equilibrium: 1 day (●), 2 days (○), 4 days (■), 10 days (▲) and 18 days (◆). Continuous lines indicate fits to a two-state model.

an additional rapidly accessible and native-like state on the unfolding side of the major barrier that typically separates the native state from denatured states in folding reactions. After 18 days at pH 7.2 and 40 °C, the estimated apparent free energy of folding is 10.81 ± 0.16 kcal mol⁻¹ for the unfolding equilibrium reaction and 10.10 ± 0.33 kcal mol⁻¹ for the refolding equilibrium reaction (Table S1). The small but progressive increase in the denaturant dependence of the ellipticity in the native baseline region over the time course of the experiment (Fig. 1a and b) suggests that prolonged incubation in Gdn-HCl may lead to the slow degradation and/or aggregation of the protein.

The kinetic properties of the unfolding and refolding reactions of HisF were monitored by manual mixing and stopped-flow CD spectroscopy. The traces were fit to one or a sum of exponentials and a semi-log plot of the observed relaxation times is shown as a function of the denaturant concentration in the form of a chevron plot at 40 °C and pH 7.2 (Fig. 2). A satisfactory fit of the unfolding traces between 4.4 M and 5.6 M Gdn-HCl required two exponentials, revealing a major slow phase that accelerates exponentially with increasing concentrations of Gdn-HCl and a minor faster phase, <5%, in hundreds of seconds time range that is weakly dependent on the denaturant concentration. Above 5.6 M Gdn-HCl, only a single, denaturant-dependent

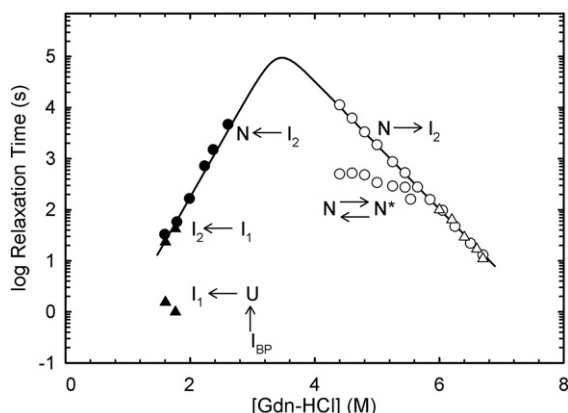
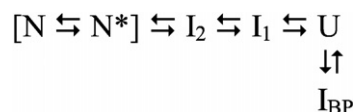


Fig. 2. Chevron plot of unfolding (open circles, open triangles) and refolding (closed circles, closed triangles) phases of HisF at pH 7.2 and 40 °C. Data were collected by manual mixing (●, ○) and stopped-flow CD (▲, △) spectroscopy and fit to exponentials as described in Materials and Methods. Assignments of the phases to specific steps in the folding mechanism shown in Scheme 1 are indicated. The rollover in the refolding leg of the chevron, ascribed to the $I_1 \rightarrow I_2$ reaction, is obscured by aggregation at 40 °C but can be observed at 20 °C (Fig. S1a). The extrapolated unfolding and refolding relaxation times in the absence of denaturant are ~116 years and 0.008 s, respectively. The buffer used was 10 mM KPi, 1 mM K₂EDTA and 0.5 mM DTT at pH 7.2 and 40 °C.

phase that smoothly converges with the major unfolding phase was observed. Refolding was more complex, beginning with a sub-millisecond burst-phase that accounts for 60% of the refolding amplitude at 1.6 M Gdn-HCl. The remainder of the native ellipticity involves a small amplitude phase, ~5%, in the 1- to 10 s time range just below 1.8 M Gdn-HCl and a major, ~35%, slow phase that accelerates exponentially between 2.6 M and ~1.6 M Gdn-HCl (Fig. 2). Unfortunately, aggregation below 1.6 M Gdn-HCl precluded reliable measurement of the refolding reaction under more strongly folding conditions. Repeating the kinetic folding studies at 20 °C, where aggregation does not occur, revealed a rollover at ~60 s in the refolding leg of the chevron below 1.6 M Gdn-HCl and an unfolding-like dependence of the faster refolding relaxation time on the denaturant concentration (Fig. S1a). Similar behavior in the αTS,⁹ IGPS³¹ and IOLI³² TIM barrels was attributed to the presence of an off-pathway intermediate. We also observed a progressive increase in the amplitude of the burst-phase reaction with decreasing final Gdn-HCl concentration (Fig. S1b), indicative of a marginally stable sub-millisecond intermediate. With the exception of the minor fast unfolding phase, the data are very similar to those reported by Carstensen *et al.* at pH 7.5 and both 25 °C and 45 °C.³³ Although these authors originally concluded that the folding mechanism involved an early off-pathway intermediate and a single subsequent on-pathway intermediate, our own analysis at 20 °C, and a subsequent study from the Sterner laboratory on an engineered variant,⁴¹ revealed an off-pathway burst-phase intermediate, I_{BP} , and a pair of on-pathway intermediates, I_1 and I_2 . Significantly, we attribute the fast unfolding phase to the conversion of the native state to a minor native-like species, N^* , in the native basin (Scheme 1).

The small amplitude of the minor unfolding phase and its modest denaturant dependence implies a subtle change in the secondary structure of N^* compared to the native state, N . The assignments of the steps in the mechanism to the observed refolding and unfolding phases are shown in Fig. 2 and Fig. S1a.

The denaturant dependence of the relaxation times for the major unfolding and refolding phases in HisF, an inverted V shape or chevron shape, is typical of reversible protein folding reactions. By linear extrapolations to the absence of denaturant, the chevron can be used to estimate the rate



Scheme 1. Proposed folding mechanism of HisF.

constants for the rate-limiting unfolding and refolding reactions in buffer.⁴² The inverse of the relaxation time, τ , equals the rate constant, k . Although the predicted refolding rate constant in the absence of denaturant, $k_f = 113 \text{ s}^{-1}$, at pH 7.2 and 40 °C is in the range observed previously for another TIM barrel,⁴³ the very small unfolding rate constant, $k_u = 2.92 \times 10^{-10} \text{ s}^{-1}$, is responsible for the substantial stability of the N state relative to the I_2 state; $\Delta G^\circ_{N/I_2} = -RT \ln(k_u/k_f) = 12.02 \text{ kcal mol}^{-1}$. Pertinent to the HX study, the lifetime of the N state in the absence of Gdn-HCl is estimated to be an astounding ~ 116 years at pH 7.2 and 40 °C. The denaturant dependence of the unfolding relaxation time is such that the lifetime of the N state exceeds 115 days up to 2.0 M Gdn-HCl at pH 7.2 and 40 °C. One might challenge the linear extrapolation required to obtain the lifetime of the native state in the absence of denaturant, and indeed, a previous study by Pace and Vanderburg suggests that nonlinearity below 1 M Gdn-HCl might lead to a reduction in the lifetime by less than 30-fold.⁴⁴ The resulting estimate of a lifetime of ~ 4 years would still far exceed the 10-day time course of the HX experiment, enabling insight into the protection patterns of the native TIM barrel fold for HisF.

HX-NMR experiments

Examination of the TROSY (transverse relaxation optimized spectroscopy) two-dimensional (2D) ^{15}N - ^1H correlation spectrum (Fig. S2a) revealed that 222 out of 240 cross-peaks were sufficiently well resolved at 40 °C to enable accurate measurements of their intensities for this study. Exchange of the amide hydrogen atoms for deuterium was monitored by recording the TROSY 2D ^{15}N - ^1H correlation spectrum as a function of time after dissolving lyophilized protein into $^2\text{H}_2\text{O}$ buffer at pH 6.8 (meter reading) at 40 °C (Fig. S2b and c). Note that the relationship between pH and pD, $\text{pD} = \text{pH} (\text{meter reading}) + 0.4$,⁴⁵ means that the actual acidity in solution was equivalent to pH 7.2. The precision of the HX measurements was enhanced by leaving the sample in the NMR probe for the entire time course of the experiment, 10 days.

The 90 Class I amide hydrogen atoms (NHs) exchanged within the time required to dissolve the sample, shim the magnet and collect a useful 2D spectrum, 30 min. The 79 Class II NHs underwent $>95\%$ exchange over 7 days, and almost all of the 53 Class III NHs underwent less than 65% exchange over the full 10 day time course of the experiment (see [Materials and Methods](#)). The exchange rate constants for the Class II NHs were obtained by fitting the fractional occupancy of hydrogen as a function of time to a single exponential function. Traces and fits at 0 M Gdn-HCl for a representative set of NHs are shown in

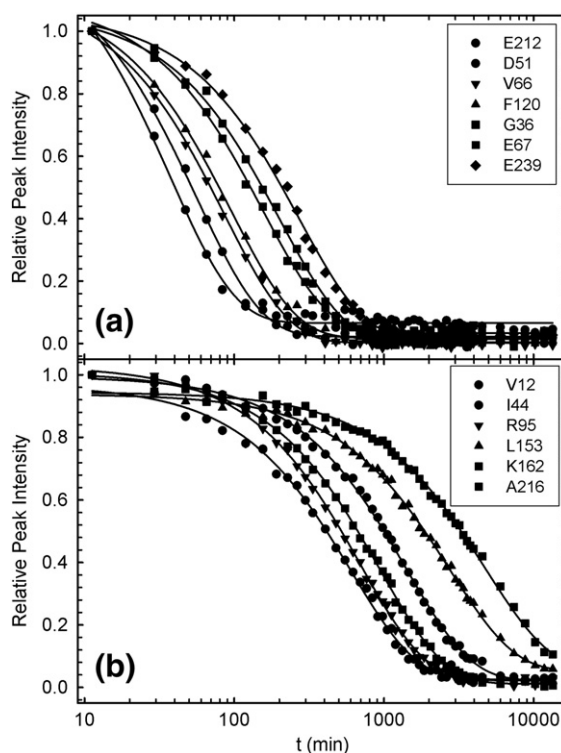


Fig. 3. Normalized intensity of ^{15}N - ^1H cross-peak versus log time plots for a representative set of NHs that exchange over the time frame. (a) 30–1000 min. (b) 30–10,000 min at 40 °C and pH 7.2 in deuterated 10 mM KPi, 50 mM KCl and 1 mM K_2EDTA . The presence of 50 mM KCl, included to mimic the buffer conditions used for the resonance assignments,⁴⁶ had no detectable effect on the rate constants for the major unfolding and refolding phases corresponding to the $[\text{N} \rightleftharpoons \text{N}^*] \rightleftharpoons \text{I}_2$ reaction (data not shown) (Scheme 1) and, therefore, the lifetime of the native state of HisF.

Fig. 3a and b for two different time regimes, 30–1000 min and 30–10,000 min, spanned by these experiments. From these measurements, protection factors (PFs) were calculated as the ratio of the exchange rate constant of the amino acid in an unstructured peptide, k_{int} , to the observed exchange rate constant for a particular amino acid in the protein, k_{obs} , $\text{PF} = k_{\text{int}}/k_{\text{obs}}$. The k_{int} was taken as the exchange rate constant for amino acids in a random-coil conformation calculated using the program Sphere†. PFs range from 4.12×10^3 for the more rapidly exchanging Class II NHs to 1.73×10^7 for slowly exchanging NHs in the absence of Gdn-HCl (Table S2). Although there was no evidence in the NMR spectra for the time-dependent changes observed in the native baseline region of the CD data (Fig. 1a and b), the fits of the hydrogen occupancies were only considered over the first 7 days of exchange. Less than 10% of the native ellipticity was lost in this time frame.

EX1 versus EX2 limit: characterizing the exchange regime

To interpret the PFs for HisF, one must determine whether exchange is limited by the opening rate constant for breaking the H-bond, the EX1 limit, or by the fraction of the open state required for exchange, the EX2 limit.⁴⁷ This determination can be made by comparing the k_{obs} values at two different pH values.^{47–50} In the EX1 limit, a linear dependence in the log–log plot is expected to intersect the y-axis at the same log unit as the x-axis and have a gradient of one. In the EX2 limit, one expects a linear dependence intersecting the y-axis at one log unit higher than the x-axis and with a gradient of one. The shift along the y-axis reflects the known acceleration of the exchange reaction with increasing pH above ~2.3.⁵¹ As shown in Fig. 4a, the raw data do not follow the expectation of either the EX1 or the EX2 regime. To probe the possible involvement of a change in stability of HisF with pH, we examined the effect of pH on the rate-limiting unfolding and refolding reactions. Although the refolding reaction is unperturbed, the unfolding reaction slows by ~5 fold when the pH is increased

from 6.2 to 7.2 (Fig. S3). Relative to the free energy of the I_2 state, the N state decreases, that is, it is stabilized, by 0.90 kcal mol^{−1} when the pH is increased from 6.2 to 7.2. The magnitude of the stabilization is precisely that would render the exchange process for the Class II NHs to obey the EX2 limit (Fig. 4b). Although HX cannot occur through the I_2 state over the time course of the experiment, N* is a potential conduit for HX in the native basin, and indeed, the slowing of the N→N* reaction at higher pH (Fig. S3) is consistent with this possibility.

In the EX2 limit, the PFs reflect the free-energy difference between the closed and open forms of the H-bond forming/breaking reactions, $\Delta G^\circ_{\text{HX}}$ ⁵² [$\Delta G^\circ_{\text{HX}} = -RT \ln(k_{\text{obs}}/k_{\text{int}})$]. The $\Delta G^\circ_{\text{HX}}$ values for the Class II NHs in the absence of denaturant, range from 5 to 10.5 kcal mol^{−1}, and along with values for k_{obs} and k_{int} , are shown in Table S2. Although the exchange behavior of the Class I and Class III NHs could not be measured, if both classes exchange by an EX2 mechanism, their $\Delta G^\circ_{\text{HX}}$ values would be less than 5 kcal mol^{−1} and greater than 10.5 kcal mol^{−1}, respectively. However, as will be discussed below (see Discussion, Mechanism of HX), the accuracy of these estimates for $\Delta G^\circ_{\text{HX}}$ is suspect and, therefore, should be considered as “apparent” values based upon the assumption of HX through unfolded segments of a protein.⁴⁰

HX protection patterns in HisF

The protection patterns for the Class I, Class II and Class III NHs are portrayed on ribbon diagrams and a 2D map for HisF in Figs. 5 and 6, respectively. The rapidly exchanging Class I NHs preferentially appear in the active-site loops at the C-termini of the β -strands, which for β_4 , β_5 and β_8 are either short α -helices (β_4 and β_8) or a β -hairpin (β_5). Class I NHs also tend to be found at the N-termini of the α -helices, contiguous with the active-site loops. For example, residues 57–65 form a particularly long stretch of Class I NHs in α_2 . Class II NHs are generally found in the C-terminal segments of the α -helices, at the C-termini of five of the eight β -strands and at the N-termini of β_1 , β_3 and β_8 . With the exception of the C-terminus of HisF, they also can be found in the short C-terminal loops and turns following the α -helices. These loops and turns have previously been implicated in stabilizing the TIM barrel.^{27,28} Non-exchanging Class III NHs are predominantly found in the β -strands, with two or more members of this class in all eight β -strands. With the exception of α_1 and α_2 , at least two Class III NHs are also found in the other six α -helices. Thus, the majority of the NHs in the β -barrel and segments of six of the α -helices in the external shell are highly resistant to HX in the native manifold of conformers of HisF.

The Class III level of protection for the NHs of E46 at the N-terminus of β_2 and E167 at the N-terminus of

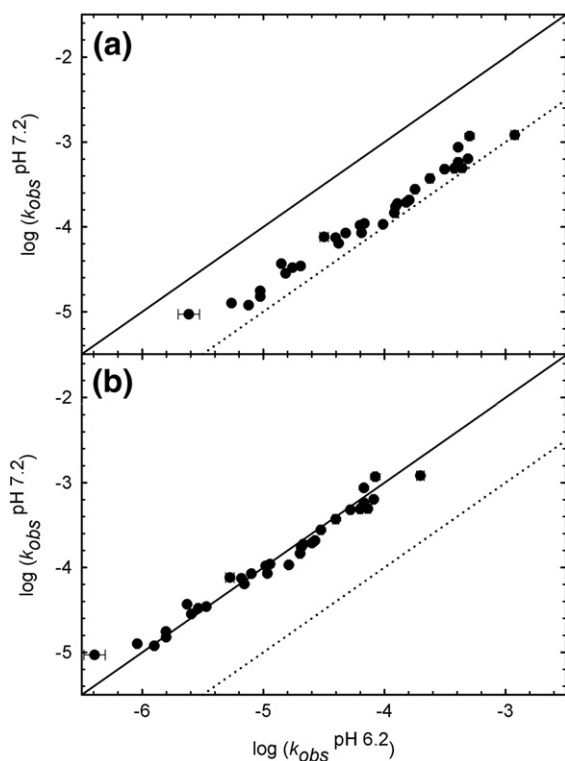


Fig. 4. (a) The log–log plot of the observed exchange rate constants measured at pH 7.2 and pH 6.2 at 40 °C for amide hydrogens of HisF, before correction for the effect of pH on stability. (b) The log–log plot of the observed exchange rate constants, after correction for the effect of pH on stability (Fig. S3). The continuous line represents the EX2 limit, and the dotted line represents EX1 limit.

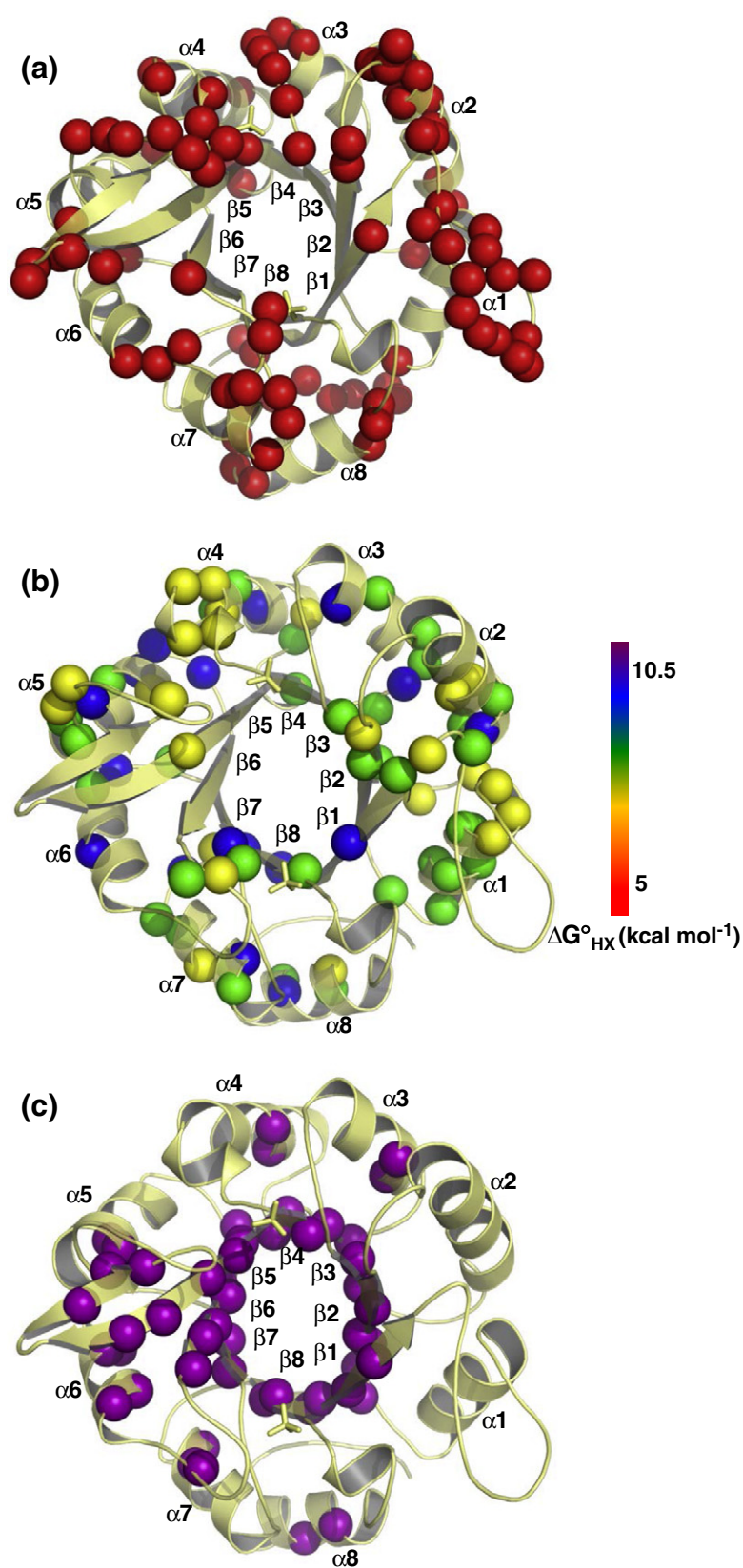


Fig. 5. A three-dimensional representation of the location of the three classes of protection for HisF at pH 7.2 and 40 °C. (a) Class I: red, $\Delta G^{\circ}_{\text{HX}} < 5 \text{ kcal mol}^{-1}$; (b) Class II: yellow, $\Delta G^{\circ}_{\text{HX}} \approx 5\text{--}7 \text{ kcal mol}^{-1}$; green, $\Delta G^{\circ}_{\text{HX}} \approx 7\text{--}9 \text{ kcal mol}^{-1}$; blue, $\Delta G^{\circ}_{\text{HX}} \approx 9\text{--}10.5 \text{ kcal mol}^{-1}$; (c) Class III: purple, $\Delta G^{\circ}_{\text{HX}} > 10.5 \text{ kcal mol}^{-1}$. The figures were generated using PyMOL.⁵³

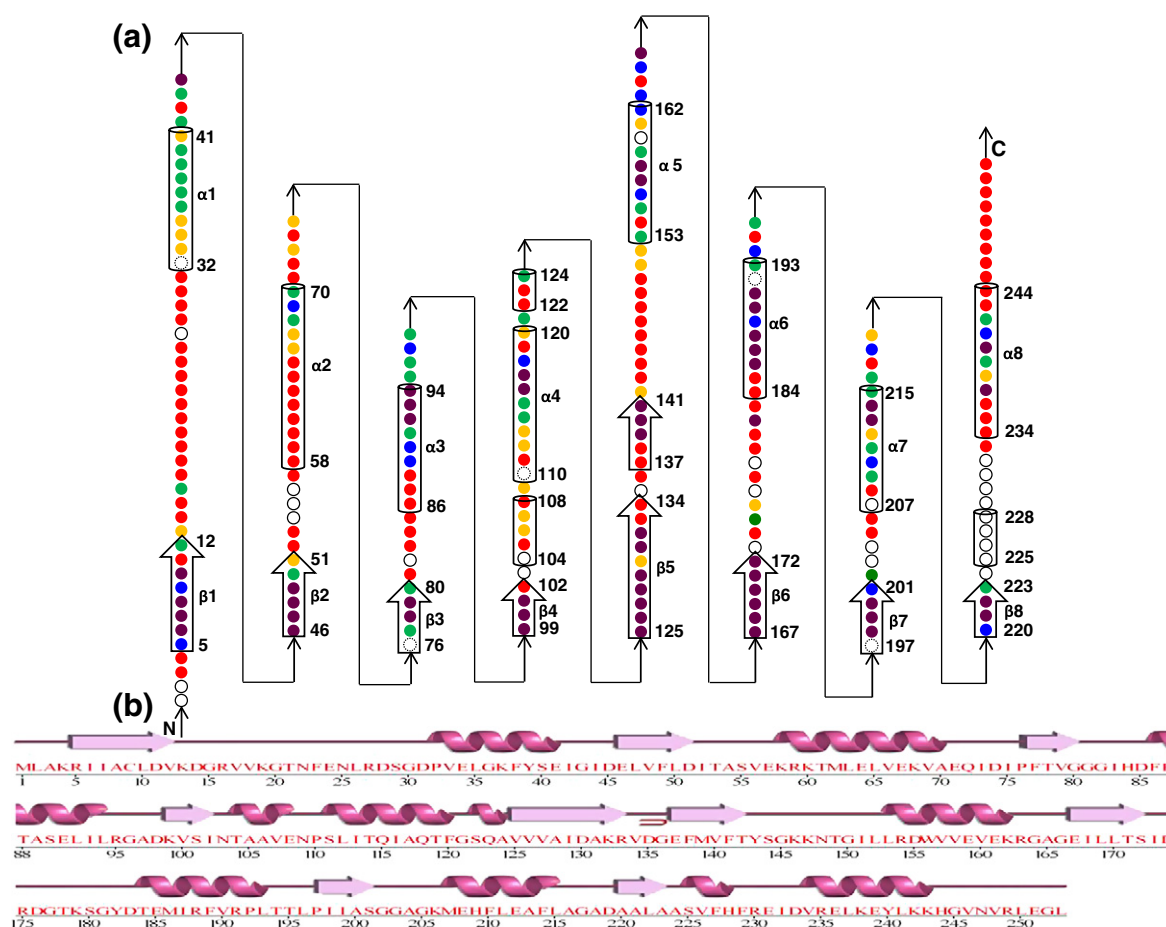


Fig. 6. (a) 2D representation of protection against exchange in HisF at pH 7.2 and 40 °C. Class I: red, $\Delta G^{\circ}_{\text{HX}} < 5$ kcal mol⁻¹; Class II: yellow, $\Delta G^{\circ}_{\text{HX}} \approx 5-7$ kcal mol⁻¹; green, $\Delta G^{\circ}_{\text{HX}} \approx 7-9$ kcal mol⁻¹; blue, $\Delta G^{\circ}_{\text{HX}} \approx 9-10.5$ kcal mol⁻¹; Class III: purple, $\Delta G^{\circ}_{\text{HX}} > 10.5$ kcal mol⁻¹. Open circles indicate either the absence of the NMR assignment or the inability to obtain accurate fits due to spectral overlap and (:::) indicates a proline residue. (b) The secondary structure of HisF mapped on its sequence.

β6 is of special note. The H-bonds to the carbonyl oxygens in I6 and V125 are at or beyond the expected length of H-bonds, 3.64 and 3.37 Å, respectively. The N-termini of the β-strands are twisted in a fashion that enables C=O of I6 to form a second H-bond with the NH of D45 and the C=O of V125 to form a second H-bond with the NH of G166 [insets (a) and (b) in Fig. 7]. Although the length of the E46 and E167 H-bonds and the necessity to share a common acceptor with other NHs implies weaker H-bonds and, possibly, weaker protection, the absence of exchange over a period of at least 10 days demonstrates that the local environments in the native state are remarkably resistant to the conformational dynamics required to expose the NHs to deuterium.

Structural dynamics of HX in the Class II NHs

Insights into the conformational changes accompanying the HX reactions at individual sites can be gained by performing HX in increasing concentra-

tions of denaturant.⁵⁴ The denaturant dependence of the PFs is correlated with the exposure of buried surface area in the opening reaction⁴⁷ and, thereby, is indicative of the magnitude of the conformational change. The lack of Gdn-HCl dependence of $\Delta G^{\circ}_{\text{HX}}$ for 62 of the 79 Class II NHs between 0 M and 2 M Gdn-HCl (see representative examples in Fig. S4) demonstrates that, within the native manifold, exchange occurs without a significant exposure of buried surface area.⁵⁵⁻⁵⁸ The remaining 17 Class II NHs could not be well resolved at increasing concentrations of Gdn-HCl. Unfortunately, attempts to collect data at higher denaturant concentrations and for longer time periods were precluded by the aggregation of HisF.

Discussion

The native-state HX-NMR experiment has most often been employed to highlight the protection

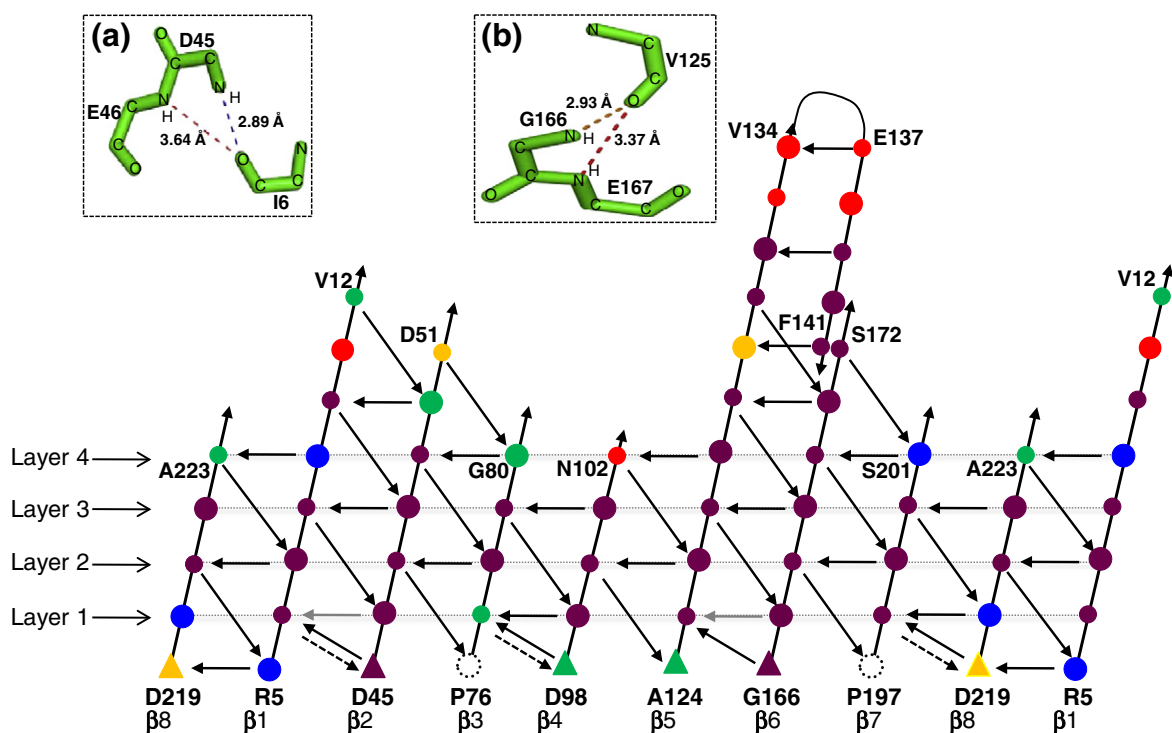


Fig. 7. 2D representation of the hydrogen-bonding pattern of the β -sheet in HisF. The arrows (\rightarrow) depict the H-bonds, pointing from the backbone amide (NH) to the main-chain carbonyl oxygen (C=O) with the exception of I6 in β 1, F77 in β 3 and I198 in β 7 where their respective main-chain amides are hydrogen bonded to the side-chain carboxyl of D45, D98 and D219, respectively (broken lines). The (\blacktriangle) represents amino acids that are not a part of β -strands. Small circles represent main-chain positions whose side chains point outside the barrel, and large circles represent main-chain positions whose side chains pointing inside the barrel. The NHs are color-coded as Class I: red, $\Delta G^\circ_{\text{HX}} < 5$ kcal mol $^{-1}$; Class II: yellow, $\Delta G^\circ_{\text{HX}} \approx 5\text{--}7$ kcal mol $^{-1}$; green, $\Delta G^\circ_{\text{HX}} \approx 7\text{--}9$ kcal mol $^{-1}$; blue, $\Delta G^\circ_{\text{HX}} \approx 9\text{--}10.5$ kcal mol $^{-1}$; Class III: purple, $\Delta G^\circ_{\text{HX}} > 10.5$ kcal mol $^{-1}$. The four layers of side chains within the β -strands are indicated by broken lines. (a) and (b) in the insets illustrate the pair of H-bond donors to the I6 and V125 carbonyl acceptors (gray arrows in the main figure).

patterns of amide hydrogens in rare high-energy partially folded states of globular proteins. In favorable cases, the patterns reveal substructures with increasing resistance to exchange and residue-specific information on the persistent patterns of secondary structure.^{28,55,58} The application of this experiment to the protection patterns in the actual native state of globular proteins is often precluded by the comparable lifetime of the native state and the minimum time required to obtain useful spectra, ~ 30 min. The resulting contribution of partially folded states to the protection pattern can obscure the protection offered by the native state. Fortunately, the very long lifetime of the hyperthermophilic HisF TIM barrel provides site-specific insights into the main-chain H-bonds that stabilize the native fold and, in the process, reveals networks that define cores of stability.

HX properties of HisF

The protection of main-chain NHs against exchange in the native basin of HisF at pH 7.2 and 40 °C falls into three classes:

1. Class I NHs exchange with solvent within the time required to initiate the exchange reaction by dilution of lyophilized protein into deuterated buffer and collect a usable 2D NMR spectrum, ~ 30 min. In general, the Class I NHs are found in loops connecting the C-termini of β -strands with their subsequent α -helices, that is, the loops that form the active site of HisF and the N-termini of α 2, α 3, α 6 and α 8 (Figs. 5a and 6). NMR relaxation experiments revealed dynamic behavior in this region of α 2 and attributed it to the allosteric communication between the active sites of the HisF and HisH subunits in ImGPS.⁵⁹ Because the Class I NHs exchange far more rapidly than the rate of unfolding of HisF from the native state, they are most likely exchanging by an EX2 mechanism and have apparent $\Delta G^\circ_{\text{HX}}$ values less than 5 kcal mol $^{-1}$.
2. Class II NHs exchange over the time period from ~ 30 min to 10 days via an EX2 mechanism. The Class II NHs are primarily found at the C-termini of the α -helices that dock on the

core of the β -barrel, in the short turns linking α -helices to the subsequent β -strands and, strikingly, at the N-termini of only β 1, β 3 and β 8 (Figs. 5b and 6). As noted above, the N-termini of β -strands are responsible for stabilizing the β -barrel, and indeed, the corresponding residues in the remaining five β -strands all fall into the Class III category. The apparent $\Delta G^{\circ}_{\text{HX}}$ values range from 5 to 10.5 kcal mol⁻¹ for the Class II NHs (Table S2). Exchange occurs at 62 distributed sites within the native manifold without a significant exposure of buried surface area. The pH-dependent stability correction that rationalizes the EX2 behavior (Fig. 4) implies that there is another state within the native manifold, likely the N* state in Scheme 1, that serves as a conduit for the exchange of the NHs for deuterium.

3. Class III NHs do not exchange significantly over 10 days at pH 7.2 and 40 °C. Class III NHs most often appear in the eight β -strands but are also found in six of the eight canonical helices, the short loops/turns preceding β 2 and β 6 and in the β -hairpin following β 5 (Figs. 5c and 6). With the assumption of an EX2 mechanism, the apparent $\Delta G^{\circ}_{\text{HX}}$ values for Class III NHs are greater than 10.5 kcal mol⁻¹. If a subset of the Class III NHs exchange through partially unfolded states on the native side of the barrier, they must do so through species whose free energies are at least 10.5 kcal mol⁻¹ above the native state.

Insights into the molecular underpinnings of the HisF TIM barrel

A closer examination of the protection patterns in the β -strands, α -helices, loops and turns in HisF (Fig. 6) hints at the essential structural features that define its native thermodynamic state. We take these structural features to be delineated by H-bonds that either do not break or are not accessible to deuterioxide over the 10-day time course of the HX-NMR experiment. Nearly 90% (47 out of 53) of the Class III NHs reside in the canonical β -strands and the segments of the hydrophobic faces of six of the α -helices that dock on the β -barrel. Most notably, the eight canonical β -strands contain 57% of the non-exchanging NHs and 68% of the β -strand NHs fall into the Class III category (Fig. 7).

The interior of the HisF barrel is well described by four layers of side chains, as originally recognized by Nagano *et al.*⁶⁰ Each layer consists of four side chains from either the odd-numbered or the even-numbered β -strands, alternating as the layers proceed from the N- to the C-termini of the β -strands (Fig. 7). Only the second and third layers each

contain eight Class III NHs from the eight canonical β -strands, I7 and A8 in β 1, L47 and V48 in β 2, T78 and V79 in β 3, V100 and S101 in β 4, V126 and V127 in β 5, I168 and L169 in β 6, I199 and A200 in β 7 and A221 and L222 in β 8, that do not exchange over the period of 10 days. Examination of the structure further shows that one or two Class III side chains in six of the α -helices, L92 and I93 in α 3, I116 in α 4, V157 in α 5, I187 and V190 in α 6, A213 and F214 in α 7 and L237 in α 8, are closely associated with the external side chains in layers 2 and 3 of the β -barrel (Fig. S5). This almost entirely aliphatic collection of side chains linking the β -strands (all of the eight side chains pointing out of the barrel are ILVA) and α -helices (eight of nine side chains pointing toward the barrel are ILVA) may well represent one of the essential features defining the ensemble of structures comprising the native thermodynamic state of the HisF barrel. Six of the eight side chains in layers 2 and 3 inside the β -barrel are also aliphatic. The very strong bias toward aliphatic side chains for these Class III NHs may be related to the preferential partitioning of side-chain analogs into the vapor phase *versus* water compared to their aromatic or polar counterparts (see below).

Because the free-energy difference between the native state and the I₂ intermediate (Scheme 1) is 12.02 kcal mol⁻¹, the intervening TSE must exceed that value. Access to the PFs between 10.5 and at least 12.02 kcal mol⁻¹ on the native side of the barrier, that is, those most crucial to the integrity of the native thermodynamic state, was, unfortunately, precluded by aggregation of the protein on prolonged exposure to 40 °C and 2.0 M Gdn-HCl where such states might be revealed. Note that the discrepancy between the free-energy difference between the N and U states estimated from the equilibrium unfolding analysis, 10.5 kcal mol⁻¹, and that from the kinetic analysis of scheme 1, 20.8 kcal mol⁻¹, reflects the presence of the intermediate states with native-like CD properties.

The BASiC hypothesis and HisF

The BASiC hypothesis proposes that large clusters of branched aliphatic side chains play crucial roles in stabilizing partially folded states in TIM barrel proteins.³⁶ A survey of 55 TIM barrel proteins posited that stabilizing clusters contain at least 10 ILV side chains each of which buries at least 10 Å² in an ILV self-contact and a total of at least 500 Å².⁶¹ Application of the BASiC hypothesis to the α TS²⁸ and siGPS^{34,62} TIM barrels found a very good correlation between HX protection in their partially folded states and large ILV clusters observed in their native conformations. The enhanced protection is ascribed to the preferential resistance of such clusters to the penetration of the water and/or hydroxide ion required for exchange of underlying

NHs and is consistent with the hydrophobicity scales of Wolfenden *et al.*⁶³ and Kyte and Doolittle.⁶⁴

HisF has five ILV clusters, ranging from 2 to 27 side chains in size (Fig. 8a). The side chains in a 22-residue cluster (Cluster 1) are associated with NHs in the $\alpha_7(\beta\alpha)_8(\beta\alpha)_{1-2}$, and the side chains in a 27-residue cluster (Cluster 2) are associated with NHs in the $\alpha_2(\beta\alpha)_{3-6}\beta_7$ region. A 6-residue cluster inside the barrel links β_1 , β_2 , β_5 , β_6 , β_7 and β_8 ; a 3-residue cluster links the loops after β_1 and β_2 , and the 2-residue cluster supports the loop after α_8 . The NHs

in Clusters 1 (Fig. 8b) and 2 (Fig. 8c) are dominated by Class II and Class III levels of protection against HX; however, the side chains in Cluster 2 are more efficiently packed with each other and a majority of the NHs display Class III levels of protection. The side chains in Cluster 1 primarily link the external loops and the C-terminus with the α -helices and β -strands in the $\alpha_8(\beta\alpha)_{1-2}$ segment, and those in Cluster 2 primarily fuse the α -helical shell to the β -barrel in the $\alpha_3(\beta\alpha)_{4-6}\beta_7$ segment (Fig. 8d). Consistent with this partitioning of the structure into two

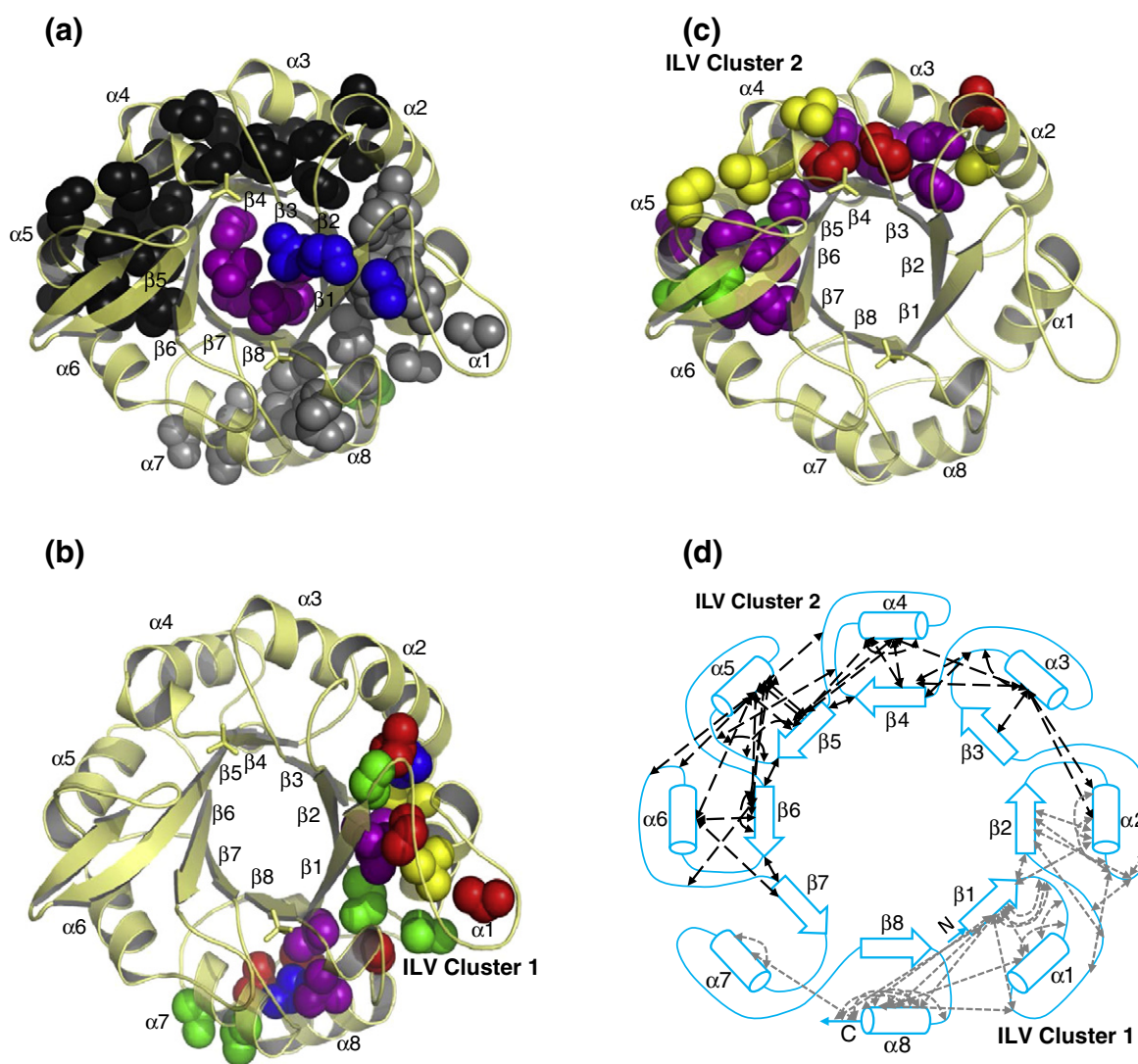


Fig. 8. HisF ILV clusters represented based on their residue contact and protection pattern. (a) HisF five ILV clusters: Cluster 1 (gray), Cluster 2 (black), Cluster 3 (green), Cluster 4 (purple) and Cluster 5 (blue). Clusters of ILV side chains in the crystal structure of HisF, Protein Data Bank entry 1THF,²² were identified using in-house software⁶¹ (<http://quipu.umassmed.edu/ccss/CCSS.cgi>). This software identifies networks of ILV side chains each of which bury at least 10 Å² and together bury more than 500 Å² surface areas using the CSU software.⁶⁵ (b) In the α_7 – α_2 regions, 22-residue Cluster 1 NHs are color-coded based on their $\Delta G^\circ_{\text{HX}}$ pattern: yellow, $\Delta G^\circ_{\text{HX}} \approx 5$ –7 kcal mol^{−1}; green, $\Delta G^\circ_{\text{HX}} \approx 7$ –9 kcal mol^{−1}; blue, $\Delta G^\circ_{\text{HX}} \approx 9$ –10.5 kcal mol^{−1} and purple, $\Delta G^\circ_{\text{HX}} > 10.5$ kcal mol^{−1}. (c) In the α_2 – β_7 regions, 27-residue Cluster 2 NHs are color-coded based on their $\Delta G^\circ_{\text{HX}}$ pattern. (d) Contact map of the two large ILV clusters illustrating the mutual contacts in the clusters spanning the α_2 – β_7 and the α_7 – α_2 segments of HisF. The figures were generated using PyMOL.⁵³

regions of protection against HX is the observation that two Class II NHs are found in the first core layer of the β -barrel, F77 in $\beta 3$ and A220 in $\beta 8$ (Fig. 7); the remainder are Class III. Although the N-terminus of the β -barrel generally stabilizes the fold, the evident dynamics at the N-termini of this pair of β -strands highlights the stronger protection for the intervening $(\beta\alpha)_{1-2}$ and $(\beta\alpha)_{4-7}$ regions of HisF.

Three orthogonal lines of evidence support the conclusion that the $\alpha_3(\beta\alpha)_{4-6}\beta_7$ region is a stability core for HisF. (1) The Rosetta design software package⁶⁶ was employed to predict the most stable 2-fold symmetric HisF TIM barrel,⁶⁷ motivated by the 4+4 gene duplication studies of Richter *et al.*³⁷ and an earlier conjecture of 2-fold symmetry of many TIM barrels.⁶⁸ Exploring the full sequence of HisF, the cut points for the $(\beta\alpha)_4$ modules were moved along the sequence, the modules duplicated, fused and docked and the energies were minimized. The most stable symmetric barrels invariably contained a pair of covalently connected $(\beta\alpha)_{4-7}$ segments that agrees closely with the strongly protected region spanned by the large ILV cluster (Fig. 8a and c). (2) A protein engineering effort by Carstensen *et al.* found substantially enhanced stability for the Sym2 construct, $\beta 5/\alpha 5/\beta 6/\alpha 6/\beta 7/\alpha 7/\beta 4/\alpha 4/\beta 5/\alpha 5/\beta 6/\alpha 6/\beta 7/\alpha 7/\beta 4/\alpha 4/$, that contains two $(\beta\alpha)_{4-7}$ elements, one overlapping the N- and C-termini.⁴¹ The X-ray structure of the engineered construct revealed two large and symmetrical ILV clusters, each of which spans the $(\beta\alpha)_{4-6}\beta_7$ elements (Fig. S6). (3) NMR and computational studies of the HisF domain in the heterodimeric ImGPS complex have found that binding of the substrate *N*-[(5'-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide perturbs the dynamics of the $\alpha_2(\beta\alpha)_{3-6}\beta_7$ and $\alpha_7(\beta\alpha)_8(\beta\alpha)_{1-2}$ sides of the barrel differently, suggesting a role for ILV clusters in enabling the allosteric mechanism of this enzyme.⁶⁹

We acknowledge that the BASiC hypothesis provides an estimate of the patterns of strong protection against HX in HisF. The protection patterns illustrated in Figs. 5 and 6 are not nearly as distinct as the large ILV clusters illustrated in Fig. 8. The hypothesis simply provides a chemical explanation for why large ILV clusters might preferentially (but not absolutely) protect against HX. In HisF, a 6 Å shell around the two large ILV clusters captures 90% of the strongly protected (Class II and Class III) amide hydrogens. Many of these NHs are associated with non-ILV side chains, showing that proximity to the clusters can bias their HX properties. We assume that the bias, in part, reflects the involvement of their aliphatic methylenes, linking polar or charged terminal atoms to the main chain, in the ILV clusters. It is also the case that about half of the amide hydrogens within this 6 Å shell are not strongly protected. While we do not understand the reasons for the over-prediction of strong protection, we attribute this

behavior to the structural fluctuations that are not apparent in a crystal structure. It is evident that the BASiC hypothesis provides a simple but approximate prediction of strong protection against HX in TIM barrel proteins. Refinement of the BASiC hypothesis will require further studies of the dynamics of the protein, for example, using long-timescale molecular dynamics simulations with explicit solvent.

All of these results validate the assumption that the HX protection pattern in the native state of HisF corresponds to cores of thermodynamic stability and support the application of the BASiC hypothesis to the native conformation of the HisF TIM barrel protein.

Mechanism of HX

The mechanism by which amide hydrogens exchange with solvent has been a subject of discussion and debate since the pioneering work of Berger and Linderstrøm-Lang,⁷⁰ Linderstrøm-Lang and Schellman⁷¹ and Hvidt and Nielsen.⁷² The local unfolding model proposes that exchange takes place in segments whose exchange properties resemble those of unfolded chains.^{71,73} By contrast, the broad category of solvent penetration models⁷⁴⁻⁷⁷ supposes that water and/or hydroxide diffuses into the body of the nominally folded protein and catalyzes exchange. In both models, the H-bond is assumed to be broken in the exchange-competent state, and if H-bond reformation is rapid compared to the pseudo-first-order exchange reaction, the exchange process falls into the EX2 limit. Englander and Kallenbach, in a comprehensive analysis of both models, argued persuasively for the local unfolding model.⁷⁸ They supposed that, among other issues, diffusion of water and especially anionic hydroxide into the low dielectric interior of a protein is very unlikely. Subsequently, Englander *et al.* acknowledged the possibility that the depth of the H-bond from the surface of a protein and/or the packing density might also play a role in limiting exchange in the native conformation.³⁹ Indeed, the pattern of protection in HisF shows that amides buried deep in the interior of the protein invariably have high PFs, whereas the surface amides range from completely unprotected to very strongly protected. Further, the most densely packed hydrophobic region (data not shown) corresponds to the $\alpha_3(\beta\alpha)_{4-6}\beta_7$ segment that closely corresponds to the most strongly protected region. In those situations, access to hydroxide would be expected to involve the coordinated disruption of structure and/or multiple packing interactions, for example, fraying, in the dynamic native manifold. Further consideration of this issue⁴⁰ led to the conclusion that the HX properties of an individual NH will reflect multiple aspects of its environment and are unlikely to yield reliable generalizations.

If fraying-like mechanisms are operative for HisF, how can one reconcile the small implied changes in structure with $\Delta G^\circ_{\text{HX}}$ values (at least 10 kcal mol⁻¹) approaching the free-energy difference between the N state and the I₂ state, 12.02 kcal mol⁻¹, across the major barrier? For comparison, the difference in buried surface area between the rate-limiting TSE and the native state can be calculated from the kinetic unfolding *m*-value, 1.7 kcal mol⁻¹ M⁻¹,⁷⁹ and is estimated to be 3200 Å². As noted previously,⁸⁰ the values of PFs and $\Delta G^\circ_{\text{HX}}$ extracted from the NMR data would be inappropriately high if the peptide models from which *k*_{int} values are derived do not accurately represent the exchange behavior of the frayed segments. Although the details of the putative fraying reactions are undoubtedly complex, the constraints on the local sequences required for the heterogeneous protection patterns observed in the α -helices and β -strands for HisF almost assuredly mean that the segments are not behaving as structureless peptides. As a consequence, the *k*_{int} values would be smaller than the values in the database employed to calculate the PFs and $\Delta G^\circ_{\text{HX}}$ values, and the end result would be a substantial over-estimation of both parameters. Thus, the $\Delta G^\circ_{\text{HX}}$ values reported in this study should be considered as apparent estimates of the free-energy difference between the closed and open states of the exchangeable H-bonds in the native state of HisF. This conjecture would also hold for other proteins for which the *k*_{obs} are independent of denaturant concentration up to the unfolding transition zone.^{55,58,81,82}

The possibility that the denaturant independence of *k*_{obs} is insensitive to a substantial exposure of buried surface elsewhere in HisF, as modeled by Wooll *et al.*, is very unlikely.⁸³ The *k*_{obs} for the 62 Class II NHs, distributed in every element of secondary structure and representing almost 30% of the NHs in HisF, do not reveal a significant denaturant dependence indicative of a large-scale conformational change. Also, the heterogeneous distribution of multiple classes of protection in each of the eight canonical helices is not consistent with this proposal. The inappropriate application of the simple HX model is more likely to provide an explanation for the apparently high values for ΔG_{HX} without a corresponding sensitivity to the denaturant concentration.

Insights into the evolution of the HisF barrel

The locations of the HX-resistant cores of stability in the native state of HisF do not mirror the 4+4 gene duplication event thought to have given rise to this TIM barrel protein.^{2,22} Sequence variation following the duplication event has caused one core to span three overlapping $\beta/\alpha/\beta$ modules, $\beta 4/\alpha 4/\beta 5/\alpha 5/\beta 6/\alpha 6/\beta 7$, and another core to stabilize the $\beta 1/\alpha 1/\beta 2$

module. $\beta/\alpha/\beta$ modules have previously been shown to be the minimal stable building blocks for the α TS TIM barrel.⁸⁴

It is interesting, however, that the locations of a different stabilizing feature of TIM barrel proteins, the main chain-side chain H-bond clamp, are consistent with the 4+4 model. These clamps are normally formed between the NH of the amino acid in the second position from the N-terminus in an odd-numbered β -strand and, most often, the carboxylic moiety of an aspartic acid immediately preceding the subsequent even-numbered β -strand.⁸⁵ The side chains associated with the NHs in the odd-numbered β -strand are large hydrophobes, often branched aliphatic side chains, that sequester the H-bond from solvent and significantly decrease the propensity of the NH to exchange with solvent. The Class II F77–D98 H-bond clamp is found between $\beta 3/\alpha 3/\beta 4$, and the Class III I198–D219 clamp is found in the analogous position between $\beta 7/\alpha 7/\beta 8$ (Fig. 7). Although the clamp is conserved, the large hydrophobe is not: F77 in $\beta 3$ shields the $\beta 3/\alpha 3/\beta 4$ clamp and I198 shields the $\beta 7/\alpha 7/\beta 8$ clamp. However, both clamps preserve a proline at the first position of $\beta 3$ and $\beta 7$, and the sequence preceding both D98 and D219, GAD, is conserved.⁸⁵ This tripeptide is consistent with previous observations of a conserved GXD sequence prior to even-numbered β -strands in TIM barrel proteins.²

Interestingly, a third clamp between I6 in $\beta 1$ and D45 prior to $\beta 2$ is also found in HisF but its analog between $\beta 5$ and $\beta 6$ is not (Fig. 7). The GID sequence prior to $\beta 2$ is replaced with GAG prior to $\beta 6$, eliminating the carboxylic acid H-bond acceptor required for the clamp with $\beta 5$. The sequence of the $\beta 1/\alpha 1/\beta 2$ module has been predicted to most closely resemble that for the earliest progenitor of the fundamental $\beta/\alpha/\beta$ module,³⁷ suggesting that this clamp might have been a key stabilizing element in that module. The loss of the clamp as these modules diverged in sequence presumably reflects the development of other stabilizing components in the sequence, for example, the large ILV cluster that spans the $\alpha 3(\beta \alpha)_{4-6}\beta 7$ region and, in particular, the $\beta 5/\alpha 5/\beta 6$ module missing the $\beta \alpha$ clamp. Although the side-chain clamp is missing between $\beta 5$ and $\beta 6$, the pair of main-chain H-bond donors from D45 and E46 in $\beta 2$ to the carbonyl oxygen of I6 in $\beta 1$ and a similar pair between G166 and E167 donors in $\beta 6$ and the V125 acceptor in $\beta 5$ (Fig. 7, insets) is conserved and in accordance with the 4+4 sequence duplication hypothesis for HisF.

Summary

The cores of stability in the native state for HisF, as assessed by protection of main-chain amide hydrogens against exchange with solvent, correlate well with the location of two large clusters of ILV side

chains. When considered with the exchange properties of two other TIM proteins, α TS²⁸ and sIGPS,³⁴ whose ILV clusters offer protection against HX in folding intermediates, it is evident that ILV clusters can serve as cores of stability all along the folding reaction coordinate. These clusters are not conserved in sequence or structure. Rather, the clusters move as the sequences evolve. What is conserved, however, is the existence of one or more ILV clusters that spans multiple adjacent $\beta/\alpha/\beta$ modules and plays a primary role in stabilizing folding intermediates and the native state of TIM barrel proteins.

Materials and Methods

Reagents

Ultrapure Gdn-HCl was purchased from MP Biomedicals, LLC (Solon, OH); deuterium oxide (99.9%), ¹⁵N ammonium chloride and potassium ethylenediaminetetraacetic acid (EDTA) were obtained from Sigma-Aldrich (St. Louis, MO). Monobasic and dibasic potassium phosphates were purchased from J. T. Baker Inc. (Phillipsburg, NJ) and Fisher Scientific (Fair Lawn, NJ), respectively. All other chemicals were reagent grade.

Protein expression and purification

HisF was expressed from the pET11c plasmid containing the HisF gene (a gift from Dr. Reinhard Sterner) in *Escherichia coli* BL21 (DE 3) cells and was purified according to a published protocol.⁸⁶ An overnight culture of freshly transformed cells in LB medium supplemented with 0.1 mg/ml ampicillin was used to inoculate a 10 l culture stock. The cells were harvested, lysed by sonication and centrifuged, and the supernatant was heat shocked at 75 °C for 15 min to precipitate host proteins. The HisF in the supernatant was dialyzed against 10 mM potassium phosphate (KPi) containing 2 mM K₂EDTA and 1 mM DTT buffer and purified by column chromatography with a DEAE-Sephacrose fast-flow column followed by a Sephacryl S-200 column. The purity of the HisF was confirmed by SDS-PAGE and by measuring the molecular mass with liquid chromatography–electrospray ionization mass spectrometry.

For NMR studies, the transformed *E. coli* BL21 (DE 3) cells were grown on minimal media supplemented with ¹⁵N ammonium chloride or ¹⁵N ammonium chloride and ¹³C glucose.⁴⁶ Cells were grown at 37 °C until an optical density at 600 nm reached 0.9, at which point the temperature was reduced to 30 °C, and the cells were allowed to grow for an additional 14 h. The purification followed the above procedures, and the purity of the HisF was confirmed by SDS-PAGE and by mass spectrometry.

Equilibrium measurements

Circular dichroism

All CD spectroscopy was performed on a JASCO-810 spectropolarimeter (Jasco Inc., Easton, MD) equipped with a water-cooled Peltier temperature control system. The CD

spectra were obtained using a 5 mm pathlength quartz cuvette, a scan rate of 50 nm min⁻¹ and a response time of 2 s. The buffer contained 10 mM KPi (pH 7.2), 1 mM K₂EDTA and 0.5 mM DTT. The Gdn-HCl induced unfolding was monitored from 215 to 250 nm, and the protein concentration was 5 μ M. Samples for the unfolding titrations were prepared by adding appropriate volumes of 0–8 M Gdn-HCl in standard buffer to a stock of native HisF in buffer. Samples for the refolding titrations were prepared by adding appropriate volumes of buffer to unfolded HisF in 7.5 M Gdn-HCl. The samples were incubated at 40 °C for up to 18 days, and aliquots were periodically withdrawn and their CD spectra were recorded to measure the progress of the unfolding and refolding reactions to equilibrium. The Gdn-HCl concentration was determined by refractive index on a Leica Mark II refractometer.⁸⁷ The data were fit to a two-state model as described previously.⁴³

Kinetic measurements

Circular dichroism

The slow unfolding and refolding kinetics of HisF were initiated by manual mixing and monitored with a JASCO-810 spectropolarimeter (Jasco Inc.). Data were collected at 222 nm in a 10 mm cuvette under continuous stirring with a solution volume of 2 ml (manual mixing dead time, ~6 s). The fast refolding and unfolding kinetics were monitored at 222 nm using an AVIV-202 stopped-flow CD spectrophotometer (dead time, ~5 ms). Unfolding experiments were initiated from the native state in standard folding buffer, and refolding experiments were initiated from protein incubated in 7.5 M Gdn-HCl overnight. The kinetic data were fit to one or more exponentials,^{43,88} using an in-house nonlinear least-squares fitting program, Savuka, as previously described.⁹ The logarithm of the observed relaxation time (s) was plotted as a function of final denaturant concentration to produce a chevron plot.⁴²

Native-state hydrogen–deuterium exchange

Purified HisF was dialyzed overnight into a buffer containing 10 mM KPi (pH 7.2), 50 mM KCl and 1 mM K₂EDTA. The sample was concentrated for the HX–NMR study with an Amicon Ultra-15 centrifugal filter unit with a 10 kDa membrane and then lyophilized. Exchange was initiated by dissolving the lyophilized protein in ²H₂O buffer, which had been prepared at the required pH and buffer conditions (all the components in the buffers were pre-deuterated by repeated cycles of dissolution and lyophilization). All pH values reported are corrected meter readings. Upon addition of ²H₂O buffer, the sample was immediately transferred to a 5 mm NMR tube (Wilmad LabGlass, Vineland, NJ) and placed in the spectrometer at 40 °C; the HisF concentration was ~0.4 mM. The time between the initiation of exchange, the transfer to the NMR tube, placement in the spectrometer, tuning and shimming and the beginning of data collection averaged 15 min. TROSY 2D ¹⁵N–¹H correlation spectra were recorded over a period of hours to days, and the sample remained in the spectrometer for the entire course of the exchange reactions. All NMR experiments were recorded on a Varian 600-MHz spectrometer, and the spectra were processed in

NMRPipe⁸⁹ and analyzed with Sparky.⁹⁰ The temperature was calibrated using a sample of 100% methanol. Three sets of independent experiments were collected to determine the experimental error.

The backbone resonance assignment of HisF is available from the BioMagResBank[†] under accession number BMRB-15741. Due to its spectral complexity and the different experimental conditions used in our experiments, buffer and temperature, we could not unambiguously assign all the backbone amide resonances based on the published assignment on HisF. For this reason, we collected three-dimensional sensitivity-enhanced gradient-selected [¹H,¹⁵N]-TROSY triple resonance experiments [HNCA, HN(CO)CA, HNCACB, HN(COCA)CB] at 14.1 T and at 40 °C using a uniformly labeled ¹³C/¹⁵N sample of HisF. Using these experiments, we were able to assign the backbone frequencies of 222 of the 240 residues of HisF.

Analysis of hydrogen–deuterium exchange data

Exchange rate constants for main-chain NHs were obtained by fitting the decay of ¹⁵N–¹H TROSY spectrum cross-peak intensities as a function of exchange time to a single exponential, $I = I_0 \exp(-k_{\text{obs}}t)$, with the initial intensity I_0 and the observed exchange rate k_{obs} as free variables in the fit. The exchange time was defined as the period of time from the dissolution of lyophilized HisF in ²H₂O buffer to the end of each TROSY 2D experiment. The uncertainties for the k_{obs} values were taken as standard errors of the fits. Under the demonstrated EX2 exchange mechanism for the Class II NHs, $\Delta G^\circ_{\text{HX}}$ was obtained from the equation: $\Delta G^\circ_{\text{HX}} = -RT \ln(k_{\text{obs}}/k_{\text{int}})$, where k_{int} is the intrinsic exchange rate calculated for amide protons in unstructured peptides. The k_{int} values were obtained using the program Sphere from model data.⁹¹ Residues that underwent greater than 70% exchange before the first time spectrum was collected were considered to have $\text{PF} < 4.12 \times 10^3$ and were classified as Class I. Residues that underwent less than 35% exchange at the end of 10 days were considered to have $\text{PF} > 1.73 \times 10^7$ and were classified as Class III. Further, residues that exhibited an incomplete exponential decay of proton occupancy and had a remaining amplitude >30% after 10 days were also classified as Class III.

Analysis of ILV clusters from the crystal structure of HisF

Clusters of ILV side chains in the crystal structure of HisF, 1THF,²² were identified using in-house software.⁶¹ This software identifies networks of ILV side chains each of which bury at least 10 Å² and together bury more than 500 Å² surface areas using the CSU software.⁶⁵

Acknowledgements

We thank Dr. Reinhard Sterner for generously providing the pET11c expression plasmid. Thanks also to Drs. Osman Bilsel, Jill Zitzewitz, Ramakrishna Vadrevu and Can Kayatekin and to Mr. Paul Nobrega

for helpful discussions. This work was supported by National Institutes of Health grants GM 23303 to C.R.M. and GM 98763 to F.M.

Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2013.01.002>

Received 25 February 2012;

Received in revised form 1 January 2013;

Accepted 3 January 2013

Available online 16 January 2013

Keywords:

ILV clusters;
NMR spectroscopy;
native-state hydrogen exchange;
protein folding

† www.fccc.edu/research/labs/roder/sphere/

‡ <http://www.bmrwisc.edu/>

Abbreviations used:

EDTA, ethylenediaminetetraacetic acid;
Gdn-HCl, guanidine hydrochloride;
H-bond, hydrogen bond; HX, hydrogen exchange;
ILV, isoleucine, leucine and valine; ImGPS, imidazole-3-glycerol phosphate synthase; PF, protection factor;
TSE, transition-state ensemble; 2D, two-dimensional.

References

1. Branden, C. I. (1991). The TIM barrel—the most frequently occurring folding motif in proteins. *Curr. Opin. Struct. Biol.* **1**, 978–983.
2. Nagano, N., Orengo, C. A. & 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.
3. Farber, G. K. & Petsko, G. A. (1990). The evolution of α/β barrel enzymes. *Trends Biochem. Sci.* **15**, 228–234.
4. Beaucamp, N., Hofmann, A., Kellerer, B. & Jaenicke, R. (1997). Dissection of the gene of the bifunctional PGK-TIM fusion protein from the hyperthermophilic bacterium *Thermotoga maritima*: design and characterization of the separate triosephosphate isomerase. *Protein Sci.* **6**, 2159–2165.
5. Nickbarg, E. B. & Knowles, J. R. (1988). Triosephosphate isomerase: energetics of the reaction catalyzed by the yeast enzyme expressed in *Escherichia coli*. *Biochemistry*, **27**, 5939–5947.
6. Wilmanns, M., Hyde, C. C., Davies, D. R., Kirschner, K. & Jansonius, J. N. (1991). Structural conservation in parallel β/α -barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry*, **30**, 9161–9169.

7. Knowles, J. R. (1991). Enzyme catalysis: not different, just better. *Nature*, **350**, 121–124.
8. Go, M. K., Amyes, T. L. & Richard, J. P. (2009). Hydron transfer catalyzed by triosephosphate isomerase. products of the direct and phosphite-activated isomerization of [1-¹³C]-glycolaldehyde in D₂O. *Biochemistry*, **48**, 5769–5778.
9. Bilsel, O., Zitzewitz, J. A., Bowers, K. E. & Matthews, C. R. (1999). Folding mechanism of the α -subunit of tryptophan synthase, an α/β barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry*, **38**, 1018–1029.
10. Rojsajakul, T., Wintrode, P., Vadrevu, R., Matthews, C. R. & Smith, D. L. (2004). Multi-state unfolding of the alpha 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.
11. Finke, J. M. & Onuchic, J. N. (2005). Equilibrium and kinetic folding pathways of a TIM barrel with a funneled energy landscape. *Biophys. J.* **89**, 488–505.
12. Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R. & Bilsel, O. (2008). Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc. Natl Acad. Sci. USA*, **105**, 13367–13372.
13. Tanaka, T., Kimura, H., Hayashi, M., Fujiyoshi, Y., Fukuhara, K. I. & Nakamura, H. (1994). Characteristics of a de novo designed protein. *Protein Sci.* **3**, 419–427.
14. Joseph-McCarthy, D., Petsko, G. A. & Karplus, M. (1995). Use of a minimum perturbation approach to predict TIM mutant structures. *Protein Eng.* **8**, 1103–1115.
15. Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R. J., Barrett, W. C., Reed, G. H. *et al.* (1996). The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the α -protons of carboxylic acids. *Biochemistry*, **35**, 16489–16501.
16. Babbitt, P. C. & Gerlt, J. A. (2001). New functions from old scaffolds: how nature reengineers enzymes for new functions. *Adv. Protein Chem.* **55**, 1–28.
17. Silverman, J. A., Balakrishnan, R. & Harbury, P. B. (2001). Reverse engineering the $(\beta/\alpha)_8$ barrel fold. *Proc. Natl Acad. Sci. USA*, **98**, 3092–3097.
18. Bonneau, R., Strauss, C. E. M., Rohl, C. A., Chivian, D., Bradley, P., Malmstrom, L. *et al.* (2002). De novo prediction of three-dimensional structures for major protein families. *J. Mol. Biol.* **322**, 65–78.
19. Shukla, A. & Guptasarma, P. (2004). Folding of β/α -unit scrambled forms of *S. cerevisiae* triosephosphate isomerase: evidence for autonomy of substructure formation and plasticity of hydrophobic and hydrogen bonding interactions in core of $(\beta/\alpha)_8$ -barrel. *Proteins*, **55**, 548–557.
20. Rothlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J. *et al.* (2008). Kemp elimination catalysts by computational enzyme design. *Nature*, **453**, 190–195.
21. Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Rothlisberger, D., Zanghellini, A. *et al.* (2008). De novo computational design of retro-aldol enzymes. *Science*, **319**, 1387–1391.
22. Lang, D., Thoma, R., Henn-Sax, M., Sterner, R. & Wilmanns, M. (2000). Structural evidence for evolution of the β/α barrel scaffold by gene duplication and fusion. *Science*, **289**, 1546–1550.
23. Gerlt, J. A. & Babbitt, P. C. (2001). Barrels in pieces? *Nat. Struct. Biol.* **8**, 5–7.
24. Höcker, B., Jürgens, C., Wilmanns, M. & Sterner, R. (2001). Stability, catalytic versatility and evolution of the $(\beta/\alpha)_8$ -barrel fold. *Curr. Opin. Biotechnol.* **12**, 376–381.
25. Setiyaputra, S., Mackay, J. P. & Patrick, W. M. (2011). The structure of a truncated phosphoribosylanthranilate isomerase suggests a unified model for evolution of the $(\beta\alpha)_8$ barrel fold. *J. Mol. Biol.* **408**, 291–303.
26. Reardon, D. & Farber, G. K. (1995). The structure and evolution of α/β barrel proteins. *FASEB J.* **9**, 497–503.
27. Niermann, T. & Kirschner, K. (1990). Improving the prediction of secondary structure of “TIM” barrel enzymes. *Protein Eng.* **4**, 137–147.
28. Vadrevu, R., Wu, Y. & Matthews, C. R. (2008). NMR analysis of partially folded states and persistent structure in the alpha subunit of tryptophan synthase: implications for the equilibrium folding mechanism of a 29-kDa TIM barrel protein. *J. Mol. Biol.* **377**, 294–306.
29. Anantharaman, V., Arvind, L. & Koonin, E. V. (2003). Emergence of diverse biochemical activities in evolutionarily conserved structural scaffolds of proteins. *Curr. Opin. Chem. Biol.* **7**, 12–20.
30. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. (2006). Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *J. Mol. Biol.* **361**, 1003–1034.
31. Forsyth, F. R. & 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.
32. Forsyth, W. R., Bilsel, O., Gu, Z. & 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.
33. Carstensen, L., Zoldák, G., Schmid, F.-X. & Sterner, R. (2012). Folding mechanism of an extremely thermostable $(\beta\alpha)_8$ -barrel enzyme: a high kinetic barrier protects the protein from denaturation. *Biochemistry*, **51**, 3420–3432.
34. Gu, Z., Zitzewitz, J. A. & 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 *S. solfataricus*. *J. Mol. Biol.* **368**, 582–594.
35. Radzicka, A. & Wolfenden, R. (1988). Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. *Biochemistry*, **27**, 1664–1670.
36. Wu, Y., Vadrevu, R., Kathuria, S., Yang, X. & 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.
37. Richter, M., Bosnali, M., Carstensen, L., Seitz, T., Durchschlag, H., Blanquart, S. *et al.* (2010). Computational and experimental evidence for the evolution of a $(\beta\alpha)_8$ -barrel protein from an ancestral quarter-barrel stabilised by disulfide bonds. *J. Mol. Biol.* **398**, 763–773.
 38. Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A. & Sterner, R. (2001). Dissection of a $(\beta\alpha)_8$ -barrel enzyme into two folded halves. *Nat. Struct. Biol.* **8**, 32–36.
 39. Englander, S. W., Mayne, L. & Krishna, M. M. G. (2008). Protein folding and misfolding: mechanism and principles. *Q. Rev. Biophys.* **40**, 287–326.
 40. Skinner, J. J., Lim, W. K., Bédard, S., Black, B. E. & Englander, S. W. (2012). Protein hydrogen exchange: testing current models. *Protein Sci.* **21**, 987–995.
 41. Carstensen, L., Sperl, J. M., Bocola, M., List, F., Schmid, F.-X. & Sterner, R. (2012). Conservation of the folding mechanism between designed primordial $(\beta\alpha)_8$ -barrel proteins and their modern descendant. *J. Am. Chem. Soc.* **134**, 12786–12791.
 42. Matthews, C. R. (1987). Effect of point mutations on the folding of globular proteins. *Methods Enzymol.* **154**, 498–511.
 43. Wu, Y. & Matthews, C. R. (2002). Parallel channels and rate-limiting steps in complex protein folding reactions: prolyl isomerization and the alpha subunit of trp synthase, a TIM barrel protein. *J. Mol. Biol.* **323**, 309–325.
 44. Pace, C. N. & Vanderburg, K. E. (1979). Determining globular protein stability: guanidine hydrochloride denaturation of myoglobin. *Biochemistry*, **18**, 288–292.
 45. Covington, A. K., Paabo, M., Robinson, R. A. & Bates, R. G. (1968). Use of the glass electrode in deuterium oxide and the relation between the standardized pD (p_{aD}) scale and the operational pH in heavy water. *Anal. Chem.* **40**, 700–706.
 46. Lipchock, J. M. & Loria, J. P. (2008). ^1H , ^{15}N and ^{13}C resonance assignment of imidazole glycerol phosphate (IGP) synthase protein HisF from *Thermotoga maritima*. *Biomol. NMR Assignments*, **2**, 219–221.
 47. Englander, S. W. (2000). Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 213–238.
 48. Kiefhaber, T. & Baldwin, R. L. (1995). Kinetics of hydrogen bond breakage in the process of unfolding of ribonuclease A measured by pulsed hydrogen exchange. *Proc. Natl Acad. Sci. USA*, **92**, 2657–2661.
 49. Yan, S., Kennedy, S. D. & Koide, S. (2002). Thermodynamic and kinetic exploration of the energy landscape of *Borrelia burgdorferi* OspA by native-state hydrogen exchange. *J. Mol. Biol.* **323**, 363–375.
 50. Huang, J. R., Craggs, T. D., Christodoulou, J. & Jackson, S. E. (2007). Stable intermediate states and high energy barriers in the unfolding of GFP. *J. Mol. Biol.* **370**, 356–371.
 51. Matthew, J. B. & Richards, F. M. (1983). The pH dependence of hydrogen exchange in protein. *J. Biol. Chem.* **258**, 3039–3044.
 52. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1994). Protein stability parameters measured by hydrogen exchange. *Proteins: Struct., Funct., Genet.* **20**, 4–14.
 53. Delano, W. L. (2002). The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
 54. Xu, Y., Mayne, L. & Englander, S. W. (1998). Evidence for an unfolding and refolding pathway in cytochrome c. *Nat. Struct. Biol.* **5**, 774–778.
 55. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.
 56. Chamberlain, A. K., Fischer, K. F., Reardon, D., Handel, T. M. & Marqusee, S. (1999). Folding of an isolated ribonuclease H core fragment. *Protein Sci.* **8**, 2251–2257.
 57. Feng, H., Vu, N. D. & Bai, Y. (2004). Detection and structure determination of an equilibrium unfolding intermediate of Rd-apocytochrome b_{562} : native fold with non-native hydrophobic interactions. *J. Mol. Biol.* **343**, 1477–1485.
 58. Chamberlain, A. K., Handel, T. M. & Marqusee, S. (1996). Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nat. Struct. Biol.* **3**, 782–787.
 59. Lipchock, J. L. & Loria, J. P. (2010). Nanometer propagation of millisecond motions in V-type allostery. *Structure*, **18**, 1596–1607.
 60. Nagano, N., Hutchinson, E. G. & Thornton, J. M. (1999). Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels. *Protein Sci.* **8**, 2072–2084.
 61. Kathuria, S. V., Day, I. J., Wallace, L. A. & 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.
 62. Gu, Z., Rao, M. K., Forsyth, W. R., Finke, J. M. & 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.
 63. Wolfenden, R., Andersson, L., Cullis, P. M. & Southgate, C. C. B. (1981). Affinities of amino acid side chains for solvent water. *Biochemistry*, **20**, 849–855.
 64. Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.
 65. Sobolev, V., Sorokine, A., Prilusky, J., Abola, E. E. & Edelman, M. (1999). Automated analysis of interatomic contacts in proteins. *Bioinformatics*, **15**, 327–332.
 66. Rohl, C., Strauss, C., Misura, K. & Baker, D. (2004). Protein structure prediction using Rosetta. *Methods Enzymol.* **383**, 66–93.
 67. Fortenberry, C., Bowman, E. A., Proffitt, W., Dorr, B., Combs, S. A., Harp, J. M. *et al.* (2011). Exploring symmetry as an avenue to the computational design of large protein domains. *J. Am. Chem. Soc.* **133**, 18026–18029.
 68. Wierenga, R. K. (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett.* **492**, 193–198.

69. Rivalta, I., Sultan, M. M., Lee, N.-S., Manley, G. A., Loria, J. P. & Batista, V. S. (2012). Allosteric pathways in imidazole glycerol phosphate synthase. *Proc. Natl Acad. Sci. USA*, **109**, E1428–E1436.
70. Berger, A. & Linderstrøm-Lang, K. (1957). Deuterium exchange of poly-DL-alanine in aqueous solution. *Arch. Biochem. Biophys.* **69**, 106–118.
71. Linderstrøm-Lang, K. U. & Schellman, J. A. (1959). Protein structure and enzyme activity. In (Boyer, P. D., Lardy, H. & Myrback, K., eds), pp. 443–510, Academic Press, New York, NY.
72. Hvidt, A. & Nielsen, S. O. (1966). Hydrogen exchange in proteins. *Adv. Protein Chem.* **21**, 287–386.
73. Skinner, J. J., Lim, W. K., Bédard, S., Black, B. E. & Englander, S. W. (2012). Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **21**, 996–1005.
74. Lumry, R. & Rosenberg, A. (1975). The mobile defect hypothesis of protein function. *Coll. Int. C. N. R. S. l'Eau. Syst. Biol.* **246**, 55–63.
75. Woodward, C. & Li, R. H. (1998). The slow-exchange core and protein folding. *T. I. B. S.* **23**, 379.
76. Richards, F. M. (1977). Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.
77. Wagner, G. & Wuthrich, K. (1979). Structural interpretation of the amide proton exchange in the basic pancreatic trypsin inhibitor and related proteins. *J. Mol. Biol.* **134**, 75–94.
78. Englander, S. W. & Kallenbach, N. R. (1983). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521–655.
79. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
80. Maity, H., Lim, W. K., Rumbley, J. N. & Englander, S. W. (2003). Protein hydrogen exchange mechanism: local fluctuations. *Protein Sci.* **12**, 153–160.
81. Vu, N. D., Feng, H. & Bai, Y. (2004). The folding pathway of barnase: the rate-limiting transition state and a hidden intermediate under native conditions. *Biochemistry*, **43**, 3346–3356.
82. Chu, R., Pei, W., Takei, J. & Bai, Y. (2002). Relationship between the native-state hydrogen exchange and folding pathways of a four-helix bundle protein. *Biochemistry*, **41**, 7998–8003.
83. Wooll, J. O., Wrabl, J. O. & Hilser, V. J. (2000). Ensemble modulation as an origin of denaturant-independent hydrogen exchange in proteins. *J. Mol. Biol.* **301**, 247–256.
84. Zitzewitz, J. A., Gualfetti, P. J., Perkons, I. A., Wasta, S. A. & Matthews, C. R. (1999). Identifying the structural boundaries of independent folding domains in the α subunit of tryptophan synthase, a β/α barrel protein. *Protein Sci.* **8**, 1200–1209.
85. Yang, X., Kathuria, S. V., Vadrevu, R. & Matthews, C. R. (2009). $\beta\alpha$ -Hairpin clamps brace $\beta\alpha\beta$ modules and can make substantive contributions to the stability of TIM barrel proteins. *PLoS One*, **4**, e7179.
86. Thoma, R., Obmolova, G., Lang, D. A., Schwander, M., Jenö, P., Sterner, R. *et al.* (1999). Efficient expression, purification and crystallisation of two hyperthermostable enzymes of histidine biosynthesis. *FEBS Lett.* **454**, 1–6.
87. Nozaki, Y. (1972). The preparation of guanidine hydrochloride. *Methods Enzymol.* **27**, 43–50.
88. Wu, Y. & 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.
89. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multi-dimensional spectral processing system based on unix pipes. *J. Biomol. NMR*, **6**, 277–293.
90. Kneller, D. G. & Kuntz, I. D. (1993). UCSF Sparky: an NMR display, annotation and assignment tool. *J. Cell. Biochem.* **53**, 254.
91. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct., Funct., Genet.* **17**, 75–86.