doi:10.1016/j.jmb.2007.12.020 *J. Mol. Biol.* (2008) **376**, 1142–1154

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Foldon Substructure of Staphylococcal Nuclease

Sabrina Bédard*, Leland C. Mayne, Ronald W. Peterson, A. Joshua Wand and S. Walter Englander

Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6059, USA

Received 16 August 2007; received in revised form 6 December 2007; accepted 10 December 2007
Available online
15 December 2007

To search for submolecular foldon units, the spontaneous reversible unfolding and refolding of staphylococcal nuclease under native conditions was studied by a kinetic native-state hydrogen exchange (HX) method. As for other proteins, it appears that staphylococcal nuclease is designed as an assembly of well-integrated foldon units that may define steps in its folding pathway and may regulate some other functional properties. The HX results identify 34 amide hydrogens that exchange with solvent hydrogens under native conditions by way of large transient unfolding reactions. The HX data for each hydrogen measure the equilibrium stability (ΔG_{HX}) and the kinetic unfolding and refolding rates (k_{op} and k_{cl}) of the unfolding reaction that exposes it to exchange. These parameters separate the 34 identified residues into three distinct HX groupings. Two correspond to clearly defined structural units in the native protein, termed the blue and red foldons. The remaining HX grouping contains residues, not well separated by their HX parameters alone, that represent two other distinct structural units in the native protein, termed the green and yellow foldons. Among these four sets, a last unfolding foldon (blue) unfolds with a rate constant of 6×10^{-6} s^{-1} and free energy equal to the protein's global stability (10.0 kcal/mol). It represents part of the β-barrel, including mutually H-bonding residues in the β4 and β5 strands, a part of the β3 strand that H-bonds to β5, and residues at the N-terminus of the α2 helix that is capped by β5. A second foldon (green), which unfolds and refolds more rapidly and at slightly lower free energy, includes residues that define the rest of the native α2 helix and its C-terminal cap. A third foldon (yellow) defines the mutually H-bonded β1–β2–β3 meander, completing the native β-barrel, plus an adjacent part of the α1 helix. A final foldon (red) includes residues on remaining segments that are distant in sequence but nearly adjacent in the native protein. Although the structure of the partially unfolded forms closely mimics the native organization, four residues indicate the presence of some nonnative misfolding interactions. Because the unfolding parameters of many other residues are not determined, it seems likely that the concerted foldon units are more extensive than is shown by the 34 residues actually observed.

Keywords: staphylococcal nuclease; protein folding; foldons; hydrogen exchange; kinetic native-state hydrogen exchange

Edited by C. R. Matthews

**Corresponding author.* E-mail address: bedard@mail.med.upenn.edu.

Abbreviations used: SNase, staphylococcal nuclease; N, I, and U, native, intermediate, and unfolded states; foldon, cooperative folding/unfolding unit; PUF, partially unfolded form (with some foldons formed and others not formed); HX, hydrogen exchange; NHX, native-state hydrogen exchange; ΔG_{HX}, unfolding free energy measured by hydrogen exchange; *m* value, $d(\Delta G_{HX})/d[GdmCl]$; EX1, monomolecular exchange where hydrogen exchange rate is equal to the determining structural opening rate; EX2, bimolecular exchange where hydrogen exchange rate is proportional to catalyst concentration and the equilibrium constant of the determining structural unfolding; GdmCl, guanidinium chloride; HSQC, heteronuclear single-quantum coherence; 3D, three-dimensional.

Introduction

Studies of the protein-folding process have serendipitously revealed a new dimension of protein structure.[1–3] It appears that proteins are made up of a small number of cooperative structural units, called foldons, that can be seen to experience repeated unfolding and refolding reactions even under native conditions, albeit at an exceedingly low level. Submolecular foldon units have now been demonstrated in many proteins by site-resolved hydrogen exchange (HX),[2,4–22] by a related thiol reactivity method,[23] by NMR relaxation dispersion,[24] and by theoretical analysis.[25,26] These results support the generality of the foldon phenomenon. Foldon structure and interactions are interesting in respect to protein stability, cooperativity, dynamics, and design, possibly even for protein evolution, and they have clear functional implications. The stepwise formation and assembly of native-like foldon units appear to account for the steps in folding pathways that carry unfolded polypeptides to their native folded state.[4] Having reached the native state, reversible foldon unfolding occurs repeatedly and can control functionally important structure change equilibria and site exposure rates.[27,28]

One wants to understand the determinants of these fundamental units of protein structure and folding and their behavior. The first implication for the existence of foldons and their possible functional significance came from the HX pulse-labeling method, which was developed to study the structure of kinetic folding intermediates.[13,29–31] This experiment, repeated for many proteins, has usually found intermediates that contain partial native-like structures. However, the experiment is limited to the study of well-populated kinetic intermediates. Most folding intermediates occur after the rate-limiting step and are invisible to all kinetically based observations.

Further progress came with the realization that native protein molecules repeatedly unfold and refold, in whole and in part, and continually reexplore all of the structural forms in their high free-energy landscape.[1,2] In principle, folding intermediates might then be studied over long time periods, whether or not they accumulate during kinetic folding. The problem is that this low-level unfolding–refolding behavior is invisible to most methods which are dominated by signals from the overwhelmingly populated native state. Fortunately, the opposite is true for HX measurements. HX rates measured for stably protected hydrogen atoms receive no contribution from the predominant native state, but are wholly determined by the cycling of protein molecules through their higher free-energy states. The HX study of transient unfolding reactions can, in favorable cases, detect the major components of the high free-energy landscape, determine their structure, measure their equilibrium stability and their kinetic unfolding and refolding rates, and evaluate their significance for functional properties, including protein-folding pathways.[4]

The experiment most used for these purposes, called equilibrium native-state hydrogen exchange (NHX),[1,2,32] attempts to amplify the equilibrium population of high free-energy partially unfolded forms (PUFs) so that they come to dominate the HX behavior that one measures. The different PUFs may then be distinguished by their different free-energy levels and identified by the sets of amide hydrogen atoms that they protect and expose. A more discriminating separation can be provided by a related approach called kinetic NHX,[12] which attempts to distinguish different unfolded forms by their unfolding rates rather than their population levels. This method exploits high pH conditions that tend to drive HX that is mediated by large unfolding reactions to the so-called EX1 limit. Here the measured HX rate of individual hydrogen atoms becomes equal to the rate of the transient structural unfolding reaction that exposes them to exchange.[31,33–35] The experiment, when successful, can measure both equilibrium stability and kinetic opening rates at many sites and thus provides a two-parameter separation of different unfolding reactions.

The present work extends these investigations to the much studied folding model staphylococcal nuclease (SNase), a member of the large OB-fold family of proteins.[36] SNase is a 149-residue mixed α/β protein with three α-helices ($\alpha1$–$\alpha3$), a major five-stranded β-barrel ($\beta1$–$\beta5$), and three minor β-strands.[37] In an earlier search for foldon units, Wrabl studied SNase by the equilibrium NHX method and found that many amide hydrogens naturally exchange by way of large unfolding reactions that approach the global unfolding in free energy.[38] As is often found, the various nearly global hydrogens displayed a spread of free energies that made it difficult to distinguish separate foldon units. This article describes a search for foldon units in SNase by use of the kinetic NHX method.

Results

Equilibrium stability from denaturant melting

The search for foldon units by native-state HX generally requires that protein stability be rather high, perhaps ≥ 9 kcal/mol, in order to provide a large dynamic range within which subglobal unfolding reactions might be separated and characterized.[2] We used a stabilized double mutant of SNase, P117G/H124L.[39] Equilibrium melting experiments indicate that each mutation increases SNase stability about equally [$\Delta C_m = 0.68$ M guanidinium chloride (GdmCl) at pH 8 and 20 °C], and the double mutant adds the separate effects. Structure changes due to these mutations are very local,[39,40] and the protein retains full activity.[39]

Most biophysical studies of SNase have been performed below neutral pH. However, HX experiments

designed to reach limiting EX1 behavior require a high pH condition where the chemical exchange of exposed amide hydrogen atoms becomes faster than the rate for reclosing of the structural unfolding reactions that expose the hydrogen atoms to exchange ($k_{ch} > k_{cl}$; Eq. (2)). We measured the equilibrium unfolding of SNase from pH 6.5 to pH 10 by fluorescence and far-ultraviolet CD (Fig. 1a). When analyzed by the linear extrapolation method,[41] the two spectroscopic probes show that stability is constant from pH 6.5 to pH 9.5 and decreases at higher pH (Fig. 1b). However, this analysis underestimates the true stability because SNase melting is not two-state (see HX analysis in later discussion).

Figure 1c reveals another issue. Fluorescence and CD_{222} amplitudes decrease above pH 8. The single fluorescence probe, Trp140, and the α3 helix are near the protein C-terminus. Heteronuclear single-quantum coherence (HSQC) spectra show that cross-peaks moving in from the C-terminus (residues 149–143) and from the N-terminus (residues 2–7) broaden and disappear when pH is raised above 6 (data not shown). These pH-dependent effects change the amplitude of measured equilibrium melting curves, but not their midpoint and slope or the extrapolated global stability until pH 10 when the onset of instability becomes apparent.

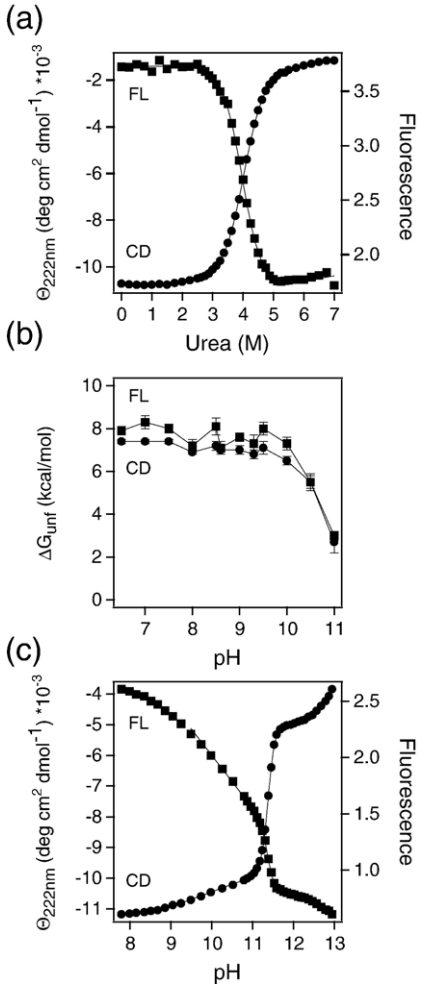

Fig. 1. Equilibrium properties of SNase P117G/H124L measured by fluorescence (■) and CD_{222} (●). (a) Urea unfolding at pH 8. (b) Stability as a function of pH obtained from urea melts by the linear extrapolation method, which shows pH dependence but underestimates true stability because SNase melting is not two-state. (c) SNase denaturation at high pH.

Kinetic NHX in pulse-labeling mode

The kinetic NHX approach attempts to distinguish foldon units by virtue of their different unfolding rates as they repeatedly unfold and refold under native conditions. At relatively high pH, HX rates can become equal to the rate of structural unfolding reactions (EX1 mode), which occurs when the unprotected chemical exchange rate is made faster than the rate for structural refolding. For example, at pH 10 and 20 °C, the average unprotected amide HX rate is about 10^4 s^{-1}. Only large unfoldings that refold more slowly than 100 μs can then produce EX1 behavior (Eq. (4)).

An initial search for EX1 behavior was performed in the conventional pulse-labeling mode[12,42] by mixing deuterium-exchanged native protein into H_2O at high pH for short labeling times (75 ms–3 h). At pH 8 and pH 10, the exchange of most protons was either too slow to measure or remained in the EX2 mode. At pH 11.5, SNase unfolds nonreversibly with a time constant of 30 s. The hydrogen atoms that would otherwise exchange more slowly than 30 s are then seen to exchange in an apparent EX1 manner, all with the same 30-s time constant. These results do not provide the kind of detailed structural and kinetic information necessary to identify and characterize SNase foldons.

Kinetic NHX in real-time mode

The very slowly exchanging hydrogens of SNase can be directly measured in real time, from hours to months, at pH conditions where the protein is stable. We first assigned NMR cross-peaks at pH 6.5–10 using three-dimensional (3D) NMR methods. D-to-H exchange was then measured in H_2O solutions at nine pH values from 7 to 10 by recording consecutive HSQC spectra in time. Accurate exchange rates and their dependence on pH were obtained for 75 individual amides and 1 tryptophan indole proton. Figure 2a illustrates HX data measured at pH 9.0 for some individual residues. Each HX time course is accurately fit by a monoexponential or a biexponential decay when cross-peaks overlap (e.g., 24/25). When necessary, overlapping crosspeaks were identified by triple resonance NMR after the faster hydrogen was largely eliminated by H-to-D exchange.

The HX rates of 41 of the 75 amide hydrogens measured remain in the EX2 mode over the pH range studied, continuing to increase by a factor of 10 per pH unit due to catalysis by OH^-. Examples are shown in the first line of Fig. 2b. These hydrogens

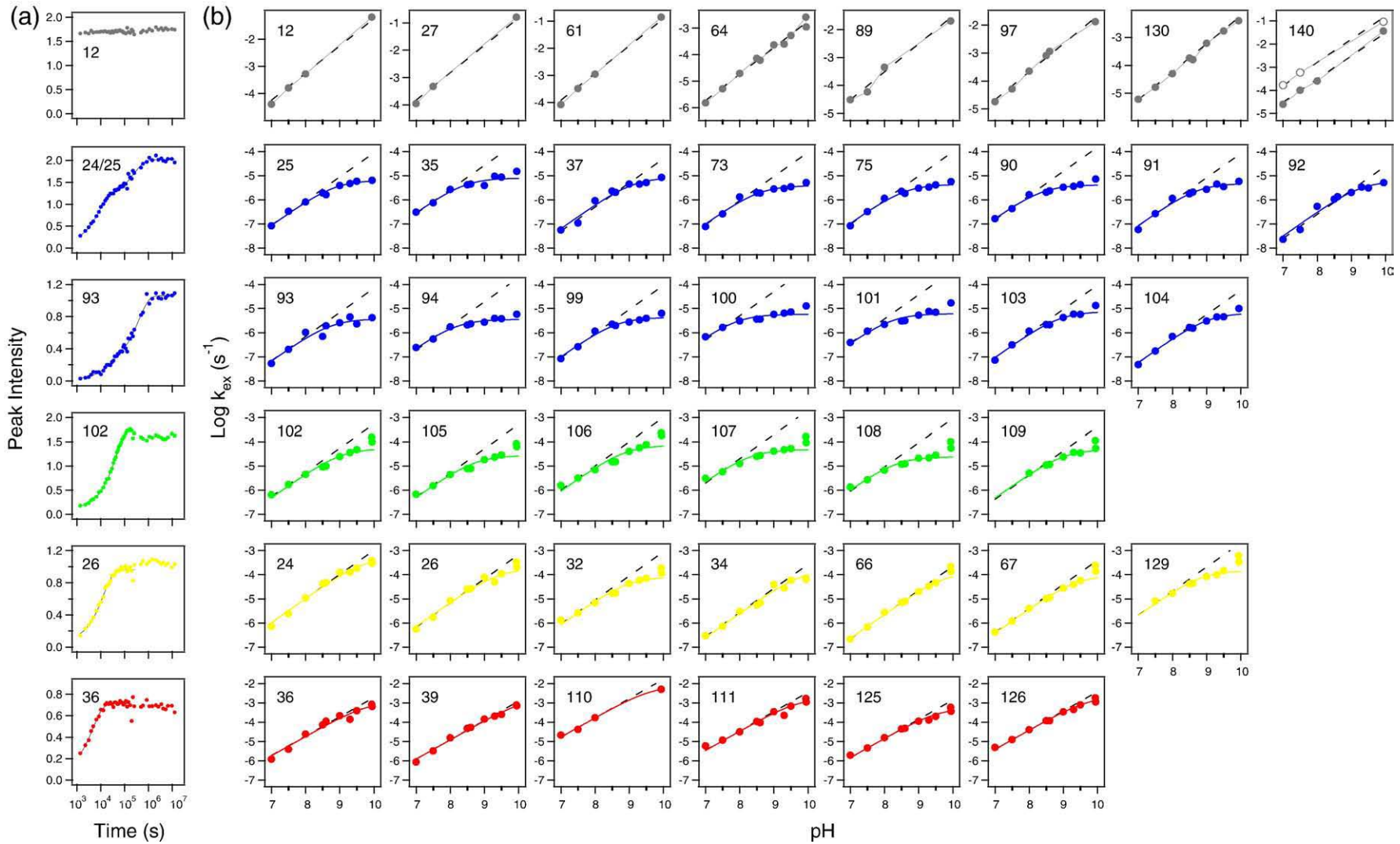

Fig. 2. Kinetic NHX results. (a) Illustrative D-to-H exchange data measured at pH 9.0. Time-dependent crosspeak amplitudes, fit to one or two exponentials, are normalized to a reference amplitude taken as the average of 10 fully protonated crosspeaks in the same HSQC spectrum and need not equal unity. (b) HX rates *versus* pH. Dashed lines show the unit slope of EX2 HX. Continuous lines are fit by Eq. (2) for an EX2-to-EX1 transition. Residues placed in each foldon unit (see Fig. 4) are in color. The participation in foldon unfolding reactions cannot be determined for sites that remain in EX2 mode (examples shown in gray) or were not measured. The double data points at pH 10 (pulse and continuous HX modes), which were systematically high due to decreased stability (see Fig. 1), were used to fit the red curves due to the paucity of other data points, but not the blue, green, and yellow data. The Trp140 panel shows both the amide (●) and indole (○) NH.

exchange by way of opening reactions, probably local fluctuational events,[43] that are able to reclose more rapidly than the chemical exchange rate ($k_{cl} > k_{ch}$), even at pH 10 where $k_{ch} \sim 10^4$ s^{-1}. The same residues undoubtedly also participate in larger foldon unfolding reactions, but this is obscured by their failure to adopt EX1 behavior.

In fortunate contrast, the HX behavior of the other 34 amide hydrogen atoms in Fig. 2b is seen to transit from EX2 to EX1 behavior (Eqs. (2)–(4)). EX1 behavior (Eq. (4)) distinguishes large unfolding reactions because only large unfoldings are able to have such slow reclosing rates (listed in Table 1). Local fluctuational openings are able to reclose very fast, faster than microseconds,[44] and so will maintain EX2 behavior even up to pH 12. Solvent-penetration mechanisms would not show EX1 behavior. The other possible explanation for the rollover in rate at high pH would be that exchange remains in the EX2 mode and the slowed HX rate is due to an increase in stability. This possibility is dispelled by the observation that stability decreases above pH 9.5 (Fig. 1b). A decrease in stability can cause HX rates to go faster, but not slower. Finally, many of these same residues were measured in the equilibrium NHX experiment of Wrabl and were

Table 1. Foldon parameters

Residue	k_{op} (s^{-1})[a]	k_{cl} (s^{-1})[a]	ΔG_{HX} (kcal/mol)[b]	Rollover pH[c]	m (kcal/mol M^{-1})[d]
Blue foldon	*(10^{-6})*[e]				
25	7.3±0.5	260±30	10.1±0.1	9.0±0.1	−3.8±0.2
35	8.4±0.4	350±30	10.2±0.1	8.5±0.1	n/a
37	8.3±0.5	220±20	10.0±0.1	9.2±0.1	−4.0±0.2
73	3.9±0.1	60±3	9.6±0.1	8.6±0.1	−4.6±0.3
75	4.6±0.1	81±4	9.7±0.1	8.6±0.1	n/a
90	4.3±0.1	132±6	10.0±0.1	8.4±0.1	−4.5±0.2
91	4.5±0.2	270±20	10.4±0.1	8.8±0.1	−4.6±0.1
92	6.4±0.5	460±40	10.5±0.1	9.4±0.1	n/a
93	3.3±0.2	140±10	10.2±0.1	8.7±0.1	−4.0±0.9
94	3.6±0.1	160±9	10.2±0.1	8.2±0.1	−5.2±0.1
99	4.4±0.2	101±8	9.9±0.1	8.7±0.1	−4.1±0.2
100	6.4±0.2	190±10	10.0±0.1	8.0±0.1	−4.7±0.3
101	6.2±0.2	100±7	9.7±0.1	8.3±0.1	−4.7±0.3
103	7.7±0.3	190±10	9.9±0.1	8.9±0.1	n/a
104	6.4±0.2	120±6	9.7±0.1	9.1±0.1	−4.3±0.2
Average	5.7±1.7	180±120	10.0±0.3	8.7±0.4	−4.4±0.4
Green foldon	*(10^{-5})*[e]				
102	5.8±0.2	800±50	9.6±0.1	9.1±0.1	n/a
105	2.9±0.1	430±30	9.6±0.1	8.8±0.1	n/a
106	12.1±0.6	3300±200	10.0±0.1	9.3±0.1	−4.7±0.2
107	5.3±0.1	720±30	9.6±0.1	8.5±0.1	−5.2±0.3
108	2.6±0.09	96±6	8.8±0.1	8.5±0.1	−4.9±0.2
109	4.6±0.2	510±60	9.4±0.1	9.0±0.5	−4.1±0.3
Average	5.5±3.5	1000±1200	9.5±0.4	8.9±0.3	−4.7±0.5
Yellow foldon	*(10^{-4})*[e]				
24	4.7±0.8	2400±500	9.0±0.2	9.6±0.1	−4.0±0.2
26	2.3±0.1	1800±100	9.3±0.1	9.6±0.1	−3.9±0.1
32	1.0±0.02	600±20	9.1±0.1	9.1±0.1	n/a
34	1.3±0.3	4000±1000	10.0±0.3	9.7±0.1	−4.3±0.1
66	1.3±0.2	1200±200	9.3±0.2	9.7±0.1	−4.1±0.3
67	1.0±0.1	510±50	9.0±0.1	9.4±0.1	−3.9±0.2
129	1.9±0.1	610±60	8.7±0.1	9.1±0.1	n/a
Average	1.9±1.3	1600±1300	9.2±0.4	9.5±0.3	−4.0±0.2
Red foldon	*(10^{-3})*[e]				
36	2.6±0.7	7000±2000	8.6±0.3	10.3±0.1	n/a
39	3.6±0.5	3400±500	8.0±0.2	10.5±0.1	−2.6±0.2
110	8±3	4000±2000	7.6±0.5	9.7±0.2	n/a
111	1.6±0.2	940±140	7.7±0.2	9.6±0.1	n/a
125	0.4±0.02	330±20	8.0±0.1	9.4±0.1	n/a
126	2.8±0.3	4400±500	8.3±0.1	9.9±0.1	−5.2±0.5
Average	3.2±2.6	3300±2400	8.0±0.4	9.9±0.4	−4±2

n/a = Not available.

[a] Obtained by fitting the kinetic NHX data in Fig. 2 using Eq. (2) (20 °C in H_2O with 50 mM buffer and 0.1 M KCl at pH 7.0–9.5). k_{op} is the rate for foldon unfolding starting from the predominant native state. k_{cl} is the rate for reprotecting the hydrogen atoms exposed in the partially unfolded form.

[b] $\Delta G_{HX} = -RT \ln(K_{op}) = -RT \ln(k_{op}/k_{cl})$.

[c] pH at which $k_{cl} = k_{ch}$.

[d] Calculated from equilibrium NHX data for SNase PHS at 25 °C in D_2O with 50 mM sodium acetate and 100 mM NaCl at pH_{read} 5.0 (personal communication and Wrabl[38]).

[e] Multiplication factor for all k_{op} values within a foldon.

seen to exchange with high dependence on denaturant (high m values) characteristic of large unfolding reactions (Table 1).[38]

Data analysis

The colored curves in Fig. 2b fit the HX data to Eq. (2), which describes the transition from EX2 to EX1 exchange at high pH. EX2 exchange is seen at the lower pH values, where HX rate increases in proportion to catalyst OH^- concentration (unit slope shown by dashed line). EX1 behavior appears at higher pH, where the HX rate is limited at the opening rate for the transient unfolding reaction that exposes each hydrogen to exchange (Eq. (4)).

Figure 3 and Table 1 show the ΔG_{HX} (unfolding free energy measured by HX) and k_{op} values obtained in this way for each residue. The results exhibit a general correlation between stability (ΔG_{HX}) and unfolding rate (k_{op}), as can be expected. More specifically, one wants to know how the measured HX parameters relate to dynamic structural unfolding events in the native protein. One suggestion is that the HX data might reflect some progressive fraying of the protecting structure. Comparison with the SNase structure (Fig. 4) provides no reasonable basis for this scenario.

Comparison of the results in Fig. 3 with the SNase structure makes obvious a quite different relationship. The HX data distribute into three distinct groupings. We applied a cluster analysis[46] to the HX data. The analysis computes a metric that compares the distance between data points to the distance between possible groupings, and assigns a quality score to each possible grouping. The metric peaks at the three visually apparent clusters circled in Fig. 3. The different HX groupings overlap significantly in

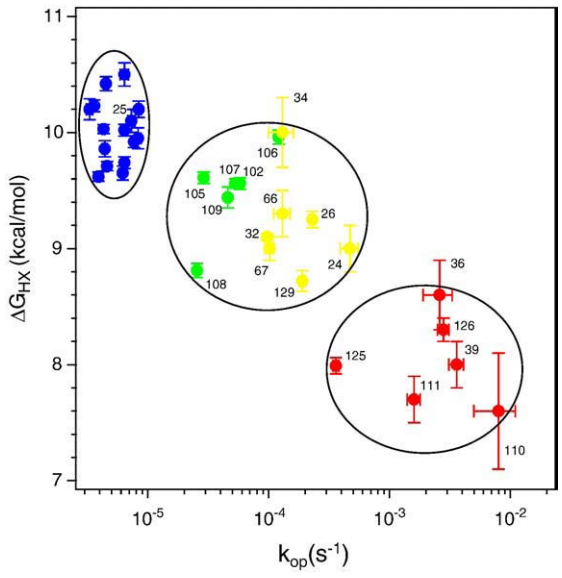

Fig. 3. Two-dimensional scatter plot separation of foldon units. ΔG_{HX} and k_{op} are the values measured from the fitted data in Fig. 2. The groupings shown are indicated visually, by formal cluster analysis, and by the placement of the amino acid residues in the native protein.

ΔG_{HX}, but much less in the k_{op} dimension. Their separation would be less obvious along either single dimension, but it becomes obvious in the two-dimensional display provided by the kinetic NHX experiment.

Independently, the identification of the blue and red sets by their HX parameters is strikingly validated by the obvious structural grouping of these amino acids in the native protein (Fig. 4), which provides a third independent dimension for foldon discrimination. Residues in the middle HX group fall into two clearly distinct structural groupings on opposite sides of the protein, indicated in green and yellow in Figs. 3 and 4. The HX data in Fig. 3 echo this separation in the k_{op} dimension, but do not alone make it clear. We assume this distribution in the following.

In summary, the hydrogens in Fig. 3 exchange by way of large unfolding reactions (EX1; m values); they fall into HX groupings defined by common free energy and rates of unfolding; and each HX group matches a coherent structural grouping in the native protein. Unfolding parameters obtained according to Eqs. (2)–(4) are listed in Table 1.

SNase foldon units

The pH dependence of the slowest exchanging hydrogens is coded in blue in Fig. 2. They display very similar HX parameters with high equilibrium stability and slow opening rate (Fig. 3, Table 1). Many were previously found to have a large m value.[38] They are grouped in the native protein (Fig. 4). They include the main-chain amide hydrogen atoms on both sides of the β5 strand other than the strand termini, their naturally paired H-bonded amides in the neighboring β3 and β4 strands, and α2 amides that intersect β5. We refer to the concerted unfolding/refolding unit defined in this way as the blue foldon. The actual foldon unit may be larger than is detected. For example, the end residues of β5 continue to exchange more rapidly and in EX2 mode, as expected for fraying reactions, obscuring their possible participation in a larger unfolding.

The residues coded in green and yellow in Fig. 3 are not cleanly separated by their HX properties, but they do have distinct placement within the native structure (Fig. 4). The green residues represent a native-like structural unit, the short doubly capped α2 helix (Fig. 4). The stability (ΔG_{HX}) of the green foldon is only modestly lower than the blue group, which explains why the equilibrium NHX experiment of Wrabl could not definitively separate them.[38] The average opening rate of the green foldon is 10 times faster than that of the blue foldon, and the reclosing rate is 5 times faster.

The HX deviation toward EX1 behavior of hydrogens in the yellow set is less impressive but seems clearly present (Fig. 2). Most of the grouped yellow residues occupy segments that are in direct contact in the native protein (Fig. 4). They represent the β2–β3 hairpin and the contiguous part of β1, which

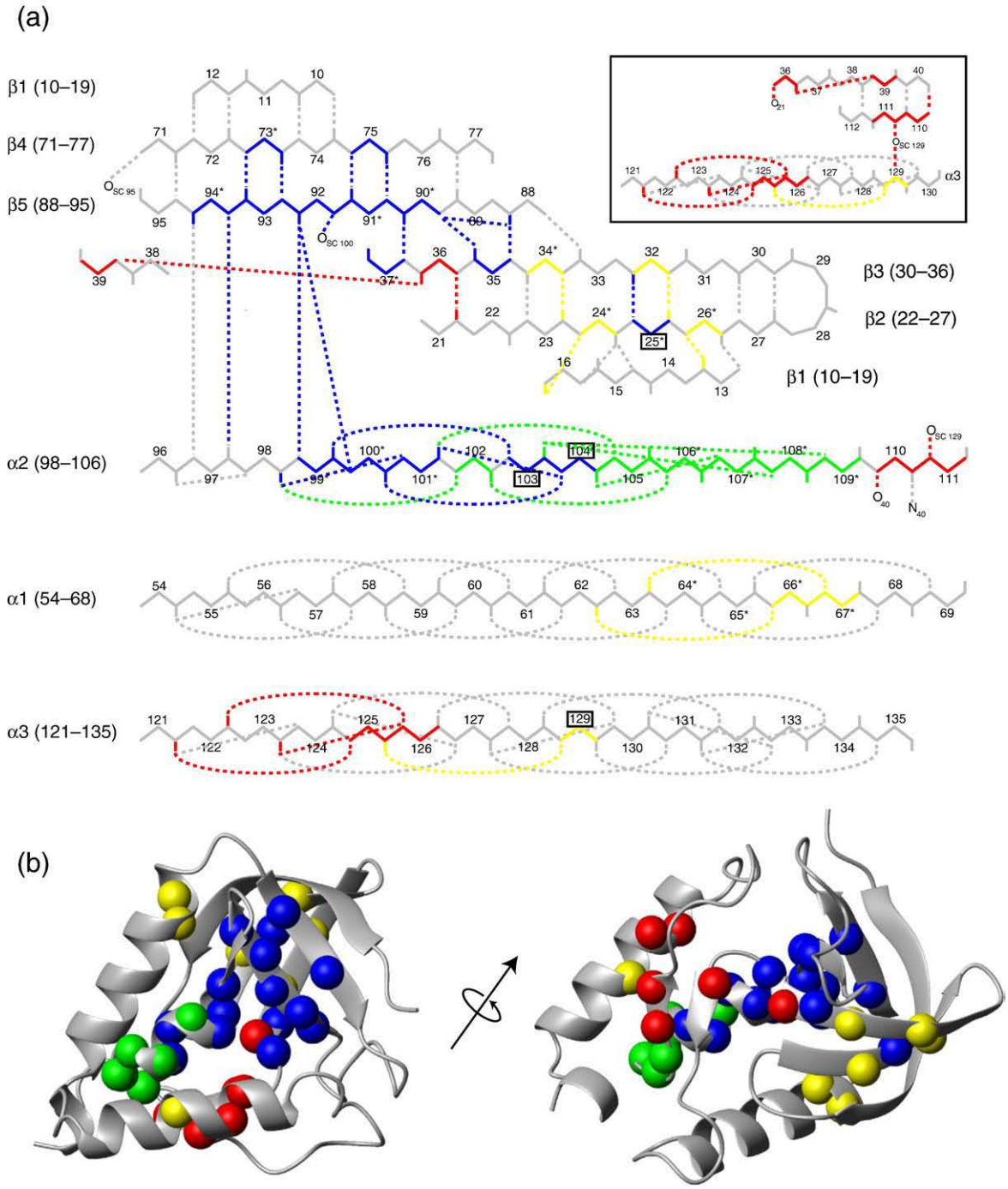

Fig. 4. Placement in the native structure of the amides measured in each color-coded foldon. (a) Foldon positions within the native SNase H-bonding pattern. Each residue carbonyl is indicated by a bar. The apparently misprotected residues (25, 129, 103, and 104) that assume nonnative protection within PUFs are boxed. Asterisks mark residues observed in the equilibrium NHX experiments of Wrabl to exchange by way of large unfoldings (high m values).[38] Inset: The red foldon H-bonding pattern. (b) Detected residues that indicate foldon positions in the SNase structure. Views were generated with MolMol[45] using Protein Data Bank entry 1SNP.

form a mutually H-bonded three-stranded β-meander sequence, and also the abutting C-terminus of α1 (Fig. 4b). Five of these residues were observed by Wrabl and found to exchange by way of a large unfolding reaction (large m value) as were also residues 64, 65, and maybe 69 (all in α1), suggesting a longer length of the α1 helix within the yellow foldon.[38]

The HX results in Fig. 2 marginally detect a final set, called red, that appears to adopt EX1 behavior. In apparent confirmation, the red residues are closely related in the native protein, where they form a small β-sheet plus a part of α3 (Fig. 4a, inset). The ΔG_{HX} values are well determined (EX2 region) and reasonably consistent (spread of 1 kcal/mol) as are the k_{op} values, except for one outlier. The rather fast

reclosing rate places the EX1 rollover point at high pH (where $k_{ch}=k_{cl}$, so that $k_{ex}=1/2k_{op}$; see Eq. (2)). This makes EX1 discrimination increasingly difficult, since fewer data points could be obtained above the rollover pH due to the onset of pH-dependent destabilization (Fig. 1). Therefore, in fitting the data for the red group, we were forced to use also the artifactually high pH 10 data point, which tends to minimize the true EX1 deviation. This factor makes it even more likely that more residues than are directly identified may be involved.

It is interesting that sites involved in function (affected by thymidine 3′,5′-diphosphate disodium salt and Ca^{2+} binding) tend to occur in the least stable red foldon, as has similarly been found for cytochrome *c*.[27,28]

Misfolding

Comparison of the HX data with the SNase structure pinpoints two structural anomalies that suggest nonnative interactions in the PUFs. Misfolding in PUFs has often been seen before.[16,22,47–50]

Residue 25 is exposed to HX with the rate and equilibrium parameters of the blue foldon, even though it is placed across the β2–β3 hairpin in the yellow foldon. No other β2–β3 residue is protected in this way, not even residue 32 which is the paired partner of residue 25 in the native protein. Similarly, residues 103 and 104 have blue foldon HX parameters, although they are placed within the green helix foldon. These results suggest a PUF in which the yellow and green foldons have previously unfolded in lower free-energy, more probable steps, but an energy-minimizing structural rearrangement reprotects these hydrogen atoms in some nonnative interactions in the still closed blue unit.

In another apparent misfolding, residue 129 exchanges with yellow foldon parameters even though it is structurally associated with the red foldon. It appears to find a protecting energy-minimizing interaction within the still folded yellow foldon when its parent red unit unfolds.

Discussion

Protein foldons

HX results for 12 proteins have now found structural elements, called foldons, that engage in large unfolding and refolding reactions under native conditions[51] and resemble coherent units of the native protein. This interesting behavior might have been expected. One has long understood that secondary structural elements tend to act as cooperative folding units.[52,53] When built into 3D proteins, their cooperative unit behavior may well be modified, but it seems unlikely that the cooperative property will be wholly lost.

Cooperativity considerations suggest that concerted foldon units may tend to involve entire native-like secondary structural elements. The cooperativity relationship captured in the Zimm–Bragg[52] and Lifson–Roig[53] formulations applies most obviously to helices. Some known foldons do encompass entire helical lengths,[1,16,19,54] but others seem not to do so.[9] Not surprisingly, entire Ω-loops also act as concerted units[6,12] because they are internally packed, self-contained structures.[55] β-Structures tend to break up into smaller separately cooperative units as seen here and before.[9,21]

In addition to the strikingly native-like configuration of known foldons, one often finds evidence for some nonnative interactions.[16,21,22,47–50] This, too, is not surprising. In PUFs, some of the interactions that delimit and stabilize native structural features are absent, and normally buried regions are exposed. It can be expected that these loosely structured forms can respond by adopting new energy-minimizing interactions.[16]

SNase foldons

The ability to distinguish different foldons in the present work was promoted by the fact that many SNase protons naturally exchange by way of large unfolding reactions and each hydrogen provides two different exposure parameters, ΔG_{HX} and k_{op}.

The many residues that define the blue SNase foldon (Fig. 2) by their common HX thermodynamic and kinetic parameters (Fig. 3, Table 1) specify a coherent native-like structural unit (Fig. 4) that includes the β4 and β5 strands and the immediately contiguous part of β3 and α2. Previous results have shown that the slowest exchanging hydrogen atoms of many proteins are exposed to exchange by the transient global unfolding reaction.[56] In this case, the HX parameters of the blue foldon, measured under reversible native conditions (Table 1), represent the last step in the SNase unfolding pathway (average, 6×10^{-6} s^{-1}; total spread, <3×) and the first step in folding. The measured ΔG_{HX} (average ΔG_{HX}, 10 kcal/mol; total spread, 0.9 kcal/mol) is larger than the stability measured by standard methods as in Fig. 1 (ΔG_{unf}=8–8.5 kcal/mol). This not uncommon circumstance is due to the fact that the global unfolding measured for SNase in high denaturants is not two-state.[18] The blue foldon refolding rate is equal to the rate for the first step measured for SNase folding by stopped-flow methods (to be described elsewhere).

The green foldon accounts for all or part of a short helix. The yellow foldon accounts for the sequential β1–β2–β3 meander, interconnected by multiple H-bonding interactions and short turns, and a contiguous length of α1. This separation of the β-barrel into two different foldons appears to be structurally reasonable. The yellow meander is largely orthogonal to the blue β-strands and is more surface exposed, and the connection between these two units may be tenuous. The lowest free-energy red foldon includes some β-strand ends and at least a part of α3. Credibility is suggested by the

fact that these residues occur together in the native protein (Fig. 4, inset).

Whether or not entire secondary structural lengths participate in the several SNase foldons, as has been seen in some other proteins, cannot be determined from the present data. As often occurs, many other residues are either not measured or continue to exchange by way of local fluctuations, obscuring their possible participation in foldon units. Accordingly, it seems likely that the foldons entrain more complete secondary structural segments than is definitively observed. It is hard to picture how a major midregion of well-structured helices or β-strands could concertedly unfold without also entraining their end residues in the same unfolding.

Problems in foldon detection

The HX detection of foldon units is made difficult by several factors. The fact that the exchange of many hydrogens is often dominated by small local fluctuations obscures their participation in larger unfolding reactions and makes the recognition of large unfoldings difficult. Even when the presence of unfolding reactions can be demonstrated, the separation of different foldons may not be obvious because foldon hydrogens commonly exhibit a spread of ΔG_{HX} values, on the order of 1 kcal/mol, as in Fig. 3. If the spread between foldons is not greater than the spread within, their discrimination is difficult. The spread may be due to fraying mechanisms that cause end hydrogens to exchange faster than others exposed to exchange by the same unfolding.[57] Independently, HX may be partially blocked in incompletely unfolded forms,[43] so that the k_{ch} value used in Eqs. (2) and (3) is incorrect, contributing to an artifactual spread in calculated ΔG_{HX} values. Finally, foldon identification can be confused by the presence of protecting nonnative interactions in unfolded forms.

Misfolding in partially unfolded states

Four SNase hydrogens exchange with parameters that identify them with HX groupings that do not match their position in the native protein, indicating the presence of some nonnative "misfolding" interactions in the PUFs. Insight into the source of these interactions comes from prior studies of TIM barrel proteins. Gu *et al.* have explored the role of large hydrophobic clusters of Ile, Leu, and Val residues in HX behavior.[58,59] They propose a significant role for ILV clusters in defining the structure of partially folded TIM barrels.[60] SNase has a large contiguous cluster of 13 such residues. Three of these account for the major misfolding observed here. Leu25(Y), Leu103(G), and V104(G) exhibit HX parameters of the blue foldon even though they are placed in other foldons in the native protein (as indicated). It appears that, when their parent foldons unfold, these residues continue to be held in their contiguous ILV cluster in an energy-minimizing arrangement that both shields their exposed hydrophobic side chains and is able to satisfy their polar main-chain amides (see Supplementary Data). This would confer blue foldon HX properties on their amide hydrogen atoms. A similar residual structure, indicated by mutational studies of the large hydrophobic side chains of SNase (m^+ and m^- mutants),[61] may reflect analogous interactions in a similar compact unfolded form.

Implications

The present results reveal foldon units in the SNase protein, define (part of) their structure, and measure their equilibrium stability and their kinetic unfolding and refolding rates. As for other proteins,[1,6,8–10,12–16,19–26,49,54,62–67] the foldons in SNase are found to resemble elements of the native protein. These results support the emerging view that protein molecules can be considered as an assemblage of separate but well-integrated foldon units. A related implication is that partially folded–partially unfolded but decisively native-like forms dominate the high free-energy space of proteins under native conditions. The intriguing possibility that the SNase foldons may illuminate the steps in its kinetic folding pathway[4,68] will be considered elsewhere.

Materials and Methods

Protein preparation

The plasmid pTSN2cc containing the SNase P117G/H124L gene (provided by John L. Markley, University of Wisconsin-Madison) was transformed into *Escherichia coli* BL21(DE3)/pLysS cells (for SNase from the V8 strain of *Staphylococcus aureus*). Cells were grown in minimal M9 media either with or without $^{15}NH_2SO_4$ and [^{13}C]glucose (37 °C, with ampicillin and chloramphenicol). IPTG was added (1 mM) at an optical density of 0.8 (600 nm), growth was continued for 3 h, and SNase was then purified in accordance with Royer *et al.*, except for the following details.[69] Buffers contained 20 mM borate at pH 9. Cell lysate was loaded directly onto a CM column, washed, and eluted with a salt gradient. Dialysis was performed against water only. Yields were 60 mg/L for unlabeled SNase, 40–50 mg/L for [^{15}N]SNase, and 20 mg/L for [^{15}N,^{13}C]SNase.

All experiments were performed at 20 °C in 0.1 M KCl, unless otherwise indicated.

Global unfolding

Equilibrium melting was measured as a function of urea and GdmCl concentration (urea shown) from pH 6.5 to pH 10 by fluorescence and CD using an AVIV model 202 CD spectrometer. The buffers (50 mM) were K_2HPO_4 for pH 6.5–7.5, Tris for pH 8.0–8.6, Ches for pH 9.0–9.5, 3-(Cyclohexylamino)-1-propanesulfonic acid for pH 10.0–10.5, and 4-(Cyclohexylamino)-1-butanesulfonic acid for pH 11.0. For pH-dependent melting (no urea), the starting buffer was 10 mM glycine (pH 8.0), and the final solution was KOH at pH 13.0, both containing 5 μM SNase. For each pH increment, a small volume of the second was

added to the first, pH was measured, and CD at 222 nm and fluorescence were recorded. Melting data were fitted by the Santoro–Bolen equation.[70]

NMR assignment

For the assignment of main-chain (^1H_N, ^{15}N, and $^{13}C\alpha$) and $^{13}C\beta$ crosspeaks, CBCA(CO)NH[71,72] and HNCACB[72,73] spectra of uniformly ^{15}N- and ^{13}C-labeled proteins were collected at pH 5.3, pH 8.0, and pH 10.0 in 0.1 M NaCl and buffer (with $CaCl_2$ and excess thymidine 3′,5′-diphosphate disodium salt at pH 5.3) using a 750-MHz Varian INOVA spectrometer and a 500-MHz instrument equipped with a cryoprobe. For backbone amide assignments at intermediate pH values, [^{15}N]HSQC spectra were collected at 0.5 pH unit intervals and mapped to known assignments to follow the movement of crosspeaks as pH was increased. Overlapping crosspeaks were separated by additional CBCA(CO)NH and HNCACB experiments coupled with H-to-D exchange. Assignments will be deposited in BioMagResBank†.

Hydrogen exchange

Fully deuterated SNase was prepared by dissolving lyophilized protein in D_2O, holding it unfolded for 10 min at a pD_r of 12, refolding at a pD_r of 5.3, and lyophilizing. The process was repeated. For HX experiments, pH was corrected to take into account the amount of D_2O [$pH_{corr} = pH_r + 0.4$(fraction D_2O)].

For pulse-quench experiments, lyophilized deuterated protein was dissolved in D_2O buffer (10 mM acetic acid), filtered, and titrated to a pD_r of 5.3. D-to-H exchange was performed by mixing with high pH pulse buffer to obtain the desired exchange pH [1:5 ratio into H_2O with 50 mM bicine (pH 8), 3-(Cyclohexylamino)-1-propanesulfonic acid (pH 10), or 4-(Cyclohexylamino)-1-butanesulfonic acid (pH 11.5)]. At measured times (75 ms–3 h), exchange was halted by mixing with quench buffer to bring the pH to 5.1 (ratio 3:1 into H_2O, 50 mM sodium acetate, and 20 mM $CaCl_2$). Mixing used stopped flow for exchange times of <3 s and manual mixing for longer times. Samples were then deep frozen pending NMR measurement.

For real-time HX, lyophilized deuterated protein was dissolved initially in D_2O (5 mM or 10 mM buffer), filtered, and titrated to the pD_{corr} of the experiment. D-to-H exchange was initiated by passage into H_2O through a spin column (gel filtration; G25 Sephadex) and prewashed in 50 mM buffer and 10% D_2O. Serial HSQC spectra were collected until exchange was complete (minutes to months). Buffers were the same as for equilibrium melting, except for bicine at pH 8.0. Protein concentration was 1–3 mM.

For the analysis of real-time D-to-H exchange measured by sequential 15-min HSQC spectra, crosspeak intensities analyzed using Felix 2.3 were normalized using the average of a subset of 10 cross-peaks that were already fully exchanged at zero time (each pH used a different subset). For NMR analysis of pulse-quench samples, 50-min HSQC spectra were collected. Spectra were normalized to a common basis by collecting a one-dimensional spectrum for each sample and by calculating the average area of the same five peaks for all samples. The exponential time course for the exchange of each amide proton and other data fitting used Igor Pro software (WaveMetrics, Inc.).

HX theory

The D-to-H exchange reaction proceeds in two steps (Eq. (1)):[33,74]

$$\text{Closed (D)} \underset{k_{cl}}{\overset{k_{op}}{\leftrightarrow}} \text{Opened (D)} \xrightarrow{k_{ch}} \text{Opened (H)} [\leftrightarrow \text{Closed (H)}] \quad (1)$$

A transient structural opening reaction (local fluctuation, subglobal unfolding, global unfolding) separates protecting H-bonds and exposes the proton to attack by solvent catalyst OH^- dominates above pH 4). In a chemical step, the HX catalyst is then able to remove the proton, which is rapidly replaced by a solvent proton in a non-rate-limiting way. The steady-state HX rate constant is given by Eq. (2):[33,75]

$$k_{ex} = (k_{op}k_{ch})/(k_{cl} + k_{ch}) \quad (2)$$

where $k_{ch} = k_{int}[OH^-]$ and k_{int} is the known intrinsic chemical HX rate calibrated for the equivalent residue amide at the experimental condition.[76–78]

In the transiently open condition, a kinetic competition between exchange and reclosing ensues. If reclosing is faster ($k_{cl} > k_{ch}$), the structural opening reaction appears as a preequilibrium step prior to the rate-limiting chemical exchange, and the HX rate is given by Eq. (3). In this so-called EX2 (bimolecular exchange) limit, the measured exchange rate reveals the fraction of time open and leads to the equilibrium constant of the opening reaction, K_{op}:[33]

$$k_{ex} = (k_{op}k_{ch})/k_{cl} = K_{op}k_{ch} \quad (3)$$

From the Boltzmann relationship ($\Delta G_{HX} = -RT \ln K_{op}$), one can then calculate the apparent free energy of the structural opening reaction that exposes the hydrogen to exchange.

If the chemical exchange rate is made faster, for example at high pH (high concentration of OH^- HX catalyst), the measured HX rate can deviate from EX2 behavior. As the chemical rate approaches and then supersedes the reclosing rate, the measured HX rate asymptotically approaches the pH-independent EX1 (monomolecular exchange) limit, where $k_{cl} < k_{ch}$. Equation (2) then reduces to Eq. (4), and the opening rate can be obtained:[33]

$$k_{ex} = k_{op} \quad (4)$$

These equations hold under steady-state conditions where the protecting structure is stable ($k_{cl} > k_{op}$; $K_{op} > 1$) and there is no significant population in the opened state. Where these conditions are not met, more general equations are required.[31,33] One assumes that EX1 behavior will only occur when exposure to exchange is controlled by a sizeable structural unfolding, so that the reclosing rate can be moderately slow. For example, an average amide k_{ch} reaches 10^4 s^{-1} at pH 10 and 20 °C. The reclosing of a small structural fluctuation is likely to be much faster, precluding EX1 behavior.

The kinetic native-state HX strategy attempts to measure k_{ex} over a pH range where EX2 exchange dominates and then transitions to EX1 exchange at increasing pH. When the exchange of a set of hydrogens becomes dominated by the cooperative unfolding of a protein segment, the measurement of HX at an amino-acid-

† http://www.bmrb.wisc.edu

resolved level can identify the unfolding segments in terms of the amino acid residues that participate. Quantitative HX data then allow the calculation of the kinetic (k_{op} and k_{cl}) and equilibrium (ΔG_{HX}) parameters of the unfolding reaction, according to Eqs. (2)–(4).

Acknowledgements

This work was supported by National Institutes of Health research grants GM031847 and GM075105 (S.W.E.) and GM035940 (A.J.W.), and by scholarships from the Natural Sciences and Engineering Research Council of Canada and the Fonds Québécois de la Recherche sur la Nature et les Technologies (S.B.). We thank Kim Sharp for help with the cluster analysis, J.O. Wrabl for his GdmCl-dependent SNase HX data, and C.R. Matthews and S. Kathuria for pointing out the connection between their Ile/Leu/Val studies and this work, and for information on ILV clustering in SNase.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.12.020

References

1. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.
2. Bai, Y. & Englander, S. W. (1996). Future directions in folding: the multi-state nature of protein structure. *Proteins: Struct. Funct. Genet.* **24**, 145–151.
3. Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. (1997). Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Sci.* **6**, 1101–1109.
4. Englander, S. W. (2000). Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 213–238.
5. Krishna, M. M. G., Lin, Y., Mayne, L. & Englander, S. W. (2003). Intimate view of a kinetic protein folding intermediate: residue-resolved structure, interactions, stability, folding and unfolding rates, homogeneity. *J. Mol. Biol.* **334**, 501–513.
6. Krishna, M. M. G., Lin, Y., Rumbley, J. N. & Walter Englander, S. (2003). Cooperative omega loops in cytochrome *c*: role in folding and function. *J. Mol. Biol.* **331**, 29–36.
7. Krishna, M. M. G., Maity, H., Rumbley, J. N. & Englander, S. W. (2007). Branching in the folding pathway of cytochrome *c*. *Protein Sci.* **16**, 1946–1956.
8. Cecconi, C., Shank, E. A., Bustamante, C. & Marqusee, S. (2005). Direct observation of the three-state folding of a single protein molecule. *Science*, **309**, 2057–2060.
9. 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.
10. Chamberlain, A. K. & Marqusee, S. (2000). Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms. *Adv. Protein Chem.* **53**, 283–328.
11. Bai, Y., Englander, J. J., Mayne, L., Milne, J. S. & Englander, S. W. (1995). Thermodynamic parameters from hydrogen exchange measurements. *Methods Enzymol.* **259**, 344–356.
12. Hoang, L., Bédard, S., Krishna, M. M. G., Lin, Y. & Englander, S. W. (2002). Cytochrome *c* folding pathway: kinetic native-state hydrogen exchange. *Proc. Natl Acad. Sci. USA*, **99**, 12173–12178.
13. Roder, H., Elove, G. A. & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome *c* by hydrogen-exchange labeling and proton NMR. *Nature*, **335**, 700–704.
14. 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.
15. Feng, H., Takei, J., Lipsitz, R., Tjandra, N. & Bai, Y. (2003). Specific non-native hydrophobic interactions in a hidden folding intermediate: implications for protein folding. *Biochemistry*, **42**, 12461–12465.
16. Feng, H., Zhou, Z. & Bai, Y. (2005). A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc. Natl Acad. Sci. USA*, **102**, 5026–5031.
17. Rojsajjakul, 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.
18. Mayne, L. & Englander, S. W. (2000). Two-state vs. multistate protein unfolding studies by optical melting and hydrogen exchange. *Protein Sci.* **9**, 1873–1877.
19. Fuentes, E. J. & Wand, A. J. (1998). Local stability and dynamics of apocytochrome *b*562 examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry*, **37**, 9877–9883.
20. Fuentes, E. J. & Wand, A. J. (1998). Local dynamics and stability of apocytochrome *b*562 examined by hydrogen exchange. *Biochemistry*, **37**, 3687–3698.
21. 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.
22. Bollen, Y. J. M., Kamphuis, M. B. & van Mierlo, C. P. M. (2006). The folding energy landscape of apoflavodoxin is rugged: hydrogen exchange reveals nonproductive misfolded intermediates. *Proc. Natl Acad. Sci. USA*, **103**, 4095–4100.
23. Silverman, J. A. & Harbury, P. B. (2002). The equilibrium unfolding pathway of a (b/a)8 barrel. *J. Mol. Biol.* **324**, 1031–1040.
24. Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Nardo, A. A. D., Davidson, A. R., Dobson, C. M. & Kay, L. E. (2004). Low-populated folding intermediates of fyn sh3 characterized by relaxation dispersion NMR. *Nature*, **430**, 586–590.
25. Weinkam, P., Zong, C. & Wolynes, P. G. (2005). A funneled energy landscape for cytochrome *c* directly predicts the sequential folding route inferred from hydrogen exchange experiments. *Proc. Natl Acad. Sci. USA*, **102**, 12401–12406.
26. Pletneva, E. V., Gray, H. B. & Winkler, J. R. (2005). Snapshots of cytochrome *c* folding. *Proc. Natl Acad. Sci. USA*, **102**, 18397–18402.
27. Hoang, L., Maity, H., Krishna, M. M. G., Lin, Y. & Englander, S. W. (2003). Folding units govern the

cytochrome *c* alkaline transition. *J. Mol. Biol.* **331**, 37–43.
28. Maity, H., Rumbley, J. N. & Englander, S. W. (2006). Functional role of a protein foldon—an ω-loop foldon controls the alkaline transition in ferricytochrome *c*. *Proteins: Struct. Funct. Bioinform.* **63**, 349–355.
29. Englander, S. W. & Mayne, L. (1992). Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 243–265.
30. Udgaonkar, J. B. & Baldwin, R. L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature*, **335**, 694–699.
31. Krishna, M. M. G., Hoang, L., Lin, Y. & Englander, S. W. (2004). Hydrogen exchange methods to study protein folding. *Methods*, **34**, 51–64.
32. Englander, S. W. (1998). Native-state HX. *Trends Biochem. Sci.* **23**, 378.
33. Hvidt, A. & Nielsen, S. O. (1966). Hydrogen exchange in proteins. *Adv. Protein Chem.* **21**, 287–386.
34. Englander, S. W. & Kallenbach, N. R. (1984). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521–655.
35. Arrington, C. B. & Robertson, A. D. (2000). Kinetics and thermodynamics of conformational equilibria in native proteins by hydrogen exchange. *Methods Enzymol.* **323**, 104–124.
36. Theobald, D., Mitton-Fry, R. & Wuttke, D. (2003). Nucleic acid recognition by ob-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 115–133.
37. Hynes, T. R. & Fox, R. O. (1991). The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution. *Proteins: Struct. Funct. Genet.* **10**, 92–105.
38. Wrabl, J. O. (1999). *Investigations of Denatured State Structure and m-Value Effects in Staphylococcal Nuclease*, Ph.D. Dissertation, The Johns Hopkins University, Baltimore, Maryland.
39. Truckses, D. M., Somoza, J. R., Prehoda, K. E., Miller, S. C. & Markley, J. L. (1996). Coupling between *trans*/*cis* proline isomerization and protein stability in staphylococcal nuclease. *Protein Sci.* **5**, 1907–1916.
40. Hynes, T. R., Hodel, A. & Fox, R. O. (1994). Engineering alternative b-turn types in staphylococcal nuclease. *Biochemistry*, **33**, 5021–5030.
41. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
42. Arrington, C. B. & Robertson, A. D. (2000). Microsecond to minute dynamics revealed by ex1-type hydrogen exchange at nearly every backbone hydrogen bond in a native protein. *J. Mol. Biol.* **296**, 1307–1317.
43. Maity, H., Lim, W. K., Rumbley, J. N. & Englander, S. W. (2003). Protein hydrogen exchange mechanism: local fluctuations. *Protein Sci.* **12**, 153–160.
44. Hernandez, G., Jenney, F. E., Adams, M. W. W. & LeMaster, D. M. (2000). Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proc. Natl Acad. Sci. USA*, **97**, 3166–3170.
45. Koradi, R., Billeter, M. & Wuethrich, K. (1996). Molmol: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, **14**, 29–32.
46. Kurita, T. (1991). An efficient agglomerative clustering algorithm using a heap. *Pattern Recognit.* **24**, 205–209.
47. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Biol.* **9**, 209–216.
48. Nishimura, C., Dyson, H. J. & Wright, P. E. (2006). Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin. *J. Mol. Biol.* **355**, 139–156.
49. Krishna, M. M. G., Lin, Y. & Englander, S. W. (2004). Protein misfolding: optional barriers, misfolded intermediates, and pathway heterogeneity. *J. Mol. Biol.* **343**, 1095–1109.
50. Wu, Y. & Matthews, C. R. (2003). Proline replacements and the simplification of the complex, parallel channel folding mechanism for the alpha subunit of trp synthase, a TIM barrel protein. *J. Mol. Biol.* **330**, 1131–1144.
51. Englander, S. W., Mayne, L. & Krishna, M. M. G. (2008). Protein folding and misfolding: mechanisms and principles from hydrogen exchange. *Q. Rev. Biophys.* In press.
52. Zimm, G. H. & Bragg, J. K. (1959). Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526–535.
53. Lifson, S. & Roig, A. (1961). On the theory of the helix–coil transition in polypeptides. *J. Chem. Phys.* **34**, 1963–1974.
54. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. (2003). Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. *Pharm. Res.* **20**, 1325–1336.
55. Leszczynski, J. F. & Rose, G. D. (1986). Loops in globular proteins: a novel category of secondary structure. *Science*, **234**, 849–855.
56. Huyghues-Despointes, B. M. P., Pace, C. N., Englander, S. W. & Scholtz, J. M. (2001). Measuring the conformational stability of a protein by hydrogen exchange. *Methods Mol. Biol.* **168**, 69–92.
57. Krishna, M. M. G., Lin, Y., Mayne, L. & Englander, W. S. (2003). Intimate view of a kinetic protein folding intermediate: residue-resolved structure, interactions, stability, folding and unfolding rates, homogeneity. *J. Mol. Biol.* **334**, 501–513.
58. 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 *Sulfolobus solfataricus*. *J. Mol. Biol.* **368**, 582–594.
59. 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.
60. Wu, Y., Vadrevu, R., Kathuria, S., Yang, X. Y. & Matthews, C. R. (2007). A tightly packed hydrophobic cluster directs the formation of an off-pathway submillisecond folding intermediate in the alpha subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol.* **366**, 1624–1638.
61. Shortle, D., Stites, W. E. & Meeker, A. K. (1990). Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. *Biochemistry*, **29**, 8033–8041.
62. Xu, Y., Mayne, L. & Englander, S. W. (1998). Evidence for an unfolding and refolding pathway in cytochrome *c*. *Nat. Struct. Biol.* **5**, 774–778.
63. Milne, J. S., Xu, Y., Mayne, L. C. & Englander, S. W. (1999). Experimental study of the protein folding landscape: unfolding reactions in cytochrome *c*. *J. Mol. Biol.* **290**, 811–822.
64. Maity, H., Maity, M. & Englander, S. W. (2004). How cytochrome *c* folds, and why: submolecular foldon

units and their stepwise sequential stabilization. *J. Mol. Biol.* **343**, 223–233.
65. Maity, H., Maity, M., Krishna, M. M. G., Mayne, L. & Englander, S. W. (2005). Protein folding: the stepwise assembly of foldon units. *Proc. Natl Acad. Sci. USA*, **102**, 4741–4746.
66. Krishna, M. M. G., Maity, H., Rumbley, J. N., Lin, Y. & Englander, S. W. (2006). Order of steps in the cytochrome *c* folding pathway: evidence for a sequential stabilization mechanism. *J. Mol. Biol.* **359**, 1411–1420.
67. Yan, S., Gawlak, G., Smith, J., Silver, L., Koide, A. & Koide, S. (2004). Conformational heterogeneity of an equilibrium folding intermediate quantified and mapped by scanning mutagenesis. *J. Mol. Biol.* **338**, 811–825.
68. Krishna, M. M. G. & Englander, S. W. (2007). A unified mechanism for protein folding: predetermined pathways with optional errors. *Protein Sci.* **16**, 449–464.
69. Royer, C. A., Hinck, A. P., Loh, S. N., Prehoda, K. E., Peng, X., Jonas, J. & Markley, J. L. (1993). Effect of amino acid substitutions on the pressure denaturation of staphylococcal nuclease as monitored by fluorescence and nuclear magnetic resonance spectroscopy. *Biochemistry*, **32**, 5222–5232.
70. Santoro, M. M. & 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.
71. Grzesiek, S. & Bax, A. (1992). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291–6293.
72. Muhandiram, D. R. & Kay, L. E. (1994). Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson. Ser. B*, **103**, 203–216.
73. Wittekind, M. & Mueller, L. (1993). Hncacb, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. Ser. B*, **101**, 201–205.
74. Linderstrom-Lang, K. (1955). Deuterium exchange between peptides and water.*Chem. Soc. (London) Spec. Publ.* **2**, 1–20; discussion, 21–4.
75. Linderstrøm-Lang, K. (1958). Deuterium exchange and protein structure. In *Symposium on Protein Structure* (Neuberger, A., ed.), pp. 23–24, Methuen, London.
76. Molday, R. S., Englander, S. W. & Kallen, R. G. (1972). Primary structure effects on peptide group hydrogen exchange. *Biochemistry*, **11**, 150–158.
77. Connelly, G. P., Bai, Y., Jeng, M. F. & Englander, S. W. (1993). Isotope effects in peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 87–92.
78. 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.