The Foldon Substructure of Staphylococcal Nuclease

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

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 ($\Delta G_{\text{HX}}$) and the kinetic unfolding and refolding rates ($k_{\text{op}}$ and $k_{\text{cl}}$) 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 $6 \times 10^{-6}$ s$^{-1}$ and free energy equal to the protein’s global stability (10.0 kcal/mol). It represents part of the $\beta$-barrel, including mutually H-bonding residues in the $\beta$4 and $\beta$5 strands, a part of the $\beta$3 strand that H-bonds to $\beta$5, and residues at the N-terminus of the $\alpha$2 helix that is capped by $\beta$5. 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 $\alpha$2 helix and its C-terminal cap. A third foldon (yellow) defines the mutually H-bonded $\beta$1–$\beta$2–$\beta$3 meander, completing the native $\beta$-barrel, plus an adjacent part of the $\alpha$1 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.

© 2007 Elsevier Ltd. All rights reserved.

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

*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; $\Delta G_{\text{HX}}$, unfolding free energy measured by hydrogen exchange; $m$ value, d($\Delta G_{\text{HX}}$)/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.

0022-2836/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
Introduction

Studies of the protein-folding process have serendipitously revealed a new dimension of protein structure. 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 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), by a related thiol reactivity method, by NMR relaxation dispersion, and by theoretical analysis. 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. Having reached the native state, reversible foldon unfolding occurs repeatedly and can control functionally important structure change equilibria and site exposure rates.

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. 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 reexpose all of the structural forms in their high free-energy landscape. 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.

The experiment most used for these purposes, called equilibrium native-state hydrogen exchange (NHX), 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, 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. 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. SNase is a 149-residue mixed α/β protein with three α-helices (α1–α3), a major five-stranded β-barrel (β1–β5), and three minor β-strands. 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. 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. We used a stabilized double mutant of SNase, P117G/H124L. Equilibrium melting experiments indicate that each mutation increases SNase stability about equally [ΔCm = 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, and the protein retains full activity.

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, 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 amplitudes decrease above pH 8. The single fluorescence probe, Trp140, and the \( \alpha_3 \) 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.

**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 \( \mu \)s 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\( _2 \)O 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\( _2 \)O 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, 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, 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>
<thead>
<tr>
<th>Residue</th>
<th>$k_{op}$ (s$^{-1}$)$^a$</th>
<th>$k_{cl}$ (s$^{-1}$)$^a$</th>
<th>$\Delta G_{HX}$ (kcal/mol)$^b$</th>
<th>Rollover pH$^c$</th>
<th>$m$ (kcal/mol M$^{-1}$)$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue foldon ($10^{-5}$)$^e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>7.3±0.5</td>
<td>260±30</td>
<td>10.1±0.1</td>
<td>9.0±0.1</td>
<td>−3.8±0.2</td>
</tr>
<tr>
<td>29</td>
<td>8.4±0.4</td>
<td>350±30</td>
<td>10.2±0.1</td>
<td>8.5±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>37</td>
<td>8.3±0.5</td>
<td>220±20</td>
<td>10.0±0.1</td>
<td>9.2±0.1</td>
<td>−4.0±0.2</td>
</tr>
<tr>
<td>73</td>
<td>3.9±0.1</td>
<td>60±3</td>
<td>9.6±0.1</td>
<td>8.6±0.1</td>
<td>−4.6±0.3</td>
</tr>
<tr>
<td>75</td>
<td>4.6±0.1</td>
<td>81±4</td>
<td>9.7±0.1</td>
<td>8.6±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>90</td>
<td>4.3±0.1</td>
<td>132±6</td>
<td>10.0±0.1</td>
<td>8.4±0.1</td>
<td>−4.3±0.2</td>
</tr>
<tr>
<td>91</td>
<td>4.5±0.2</td>
<td>270±20</td>
<td>10.4±0.1</td>
<td>8.8±0.1</td>
<td>−4.6±0.1</td>
</tr>
<tr>
<td>92</td>
<td>6.4±0.5</td>
<td>460±40</td>
<td>10.5±0.1</td>
<td>9.4±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>93</td>
<td>3.3±0.2</td>
<td>140±10</td>
<td>10.2±0.1</td>
<td>8.7±0.1</td>
<td>−4.0±0.9</td>
</tr>
<tr>
<td>94</td>
<td>3.6±0.1</td>
<td>160±9</td>
<td>10.2±0.1</td>
<td>8.2±0.1</td>
<td>−5.2±0.1</td>
</tr>
<tr>
<td>99</td>
<td>4.4±0.2</td>
<td>101±8</td>
<td>9.9±0.1</td>
<td>8.7±0.1</td>
<td>−4.1±0.2</td>
</tr>
<tr>
<td>100</td>
<td>6.4±0.1</td>
<td>190±10</td>
<td>10.0±0.1</td>
<td>8.0±0.1</td>
<td>−4.7±0.3</td>
</tr>
<tr>
<td>101</td>
<td>6.2±0.2</td>
<td>100±7</td>
<td>9.7±0.1</td>
<td>8.3±0.1</td>
<td>−4.7±0.3</td>
</tr>
<tr>
<td>102</td>
<td>7.7±0.3</td>
<td>190±10</td>
<td>9.9±0.1</td>
<td>8.9±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>104</td>
<td>6.4±0.2</td>
<td>120±6</td>
<td>9.7±0.1</td>
<td>9.1±0.1</td>
<td>−4.3±0.2</td>
</tr>
<tr>
<td>Average</td>
<td>5.7±1.7</td>
<td>180±120</td>
<td>10.0±0.3</td>
<td>8.7±0.4</td>
<td>−4.4±0.4</td>
</tr>
<tr>
<td>Green foldon ($10^{-5}$)$^e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>5.8±0.2</td>
<td>800±50</td>
<td>9.6±0.1</td>
<td>9.1±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>105</td>
<td>2.9±0.1</td>
<td>430±30</td>
<td>9.6±0.1</td>
<td>8.8±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>106</td>
<td>12.1±0.6</td>
<td>3300±200</td>
<td>10.0±0.1</td>
<td>9.3±0.1</td>
<td>−4.7±0.2</td>
</tr>
<tr>
<td>107</td>
<td>5.3±0.1</td>
<td>720±30</td>
<td>9.6±0.1</td>
<td>8.5±0.1</td>
<td>−5.2±0.3</td>
</tr>
<tr>
<td>108</td>
<td>2.6±0.09</td>
<td>96±6</td>
<td>8.8±0.1</td>
<td>8.5±0.1</td>
<td>−4.9±0.2</td>
</tr>
<tr>
<td>109</td>
<td>4.6±0.2</td>
<td>510±60</td>
<td>9.4±0.1</td>
<td>9.0±0.5</td>
<td>−4.1±0.3</td>
</tr>
<tr>
<td>Average</td>
<td>5.5±3.5</td>
<td>1000±1200</td>
<td>9.5±0.4</td>
<td>8.9±0.3</td>
<td>−4.7±0.5</td>
</tr>
<tr>
<td>Yellow foldon ($10^{-4}$)$^e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4.7±0.8</td>
<td>2400±500</td>
<td>9.0±0.2</td>
<td>9.6±0.1</td>
<td>−4.0±0.2</td>
</tr>
<tr>
<td>26</td>
<td>2.3±0.1</td>
<td>1800±100</td>
<td>9.3±0.1</td>
<td>9.6±0.1</td>
<td>−3.9±0.1</td>
</tr>
<tr>
<td>32</td>
<td>1.0±0.02</td>
<td>600±20</td>
<td>9.1±0.1</td>
<td>9.1±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>34</td>
<td>1.3±0.3</td>
<td>4000±1000</td>
<td>10.0±0.3</td>
<td>9.7±0.1</td>
<td>−4.3±0.1</td>
</tr>
<tr>
<td>66</td>
<td>1.3±0.2</td>
<td>1200±200</td>
<td>9.3±0.2</td>
<td>9.7±0.1</td>
<td>−4.1±0.3</td>
</tr>
<tr>
<td>67</td>
<td>1.0±0.1</td>
<td>510±50</td>
<td>9.0±0.1</td>
<td>9.4±0.1</td>
<td>−3.9±0.2</td>
</tr>
<tr>
<td>129</td>
<td>1.9±0.1</td>
<td>610±60</td>
<td>8.7±0.1</td>
<td>9.1±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>Average</td>
<td>1.9±1.3</td>
<td>1600±1300</td>
<td>9.2±0.4</td>
<td>9.5±0.3</td>
<td>−4.0±0.2</td>
</tr>
<tr>
<td>Red foldon ($10^{-3}$)$^e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>2.6±0.7</td>
<td>7000±2000</td>
<td>8.6±0.3</td>
<td>10.3±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>39</td>
<td>3.6±0.5</td>
<td>3400±500</td>
<td>8.0±0.2</td>
<td>10.5±0.1</td>
<td>−2.6±0.2</td>
</tr>
<tr>
<td>110</td>
<td>5.3±0.4</td>
<td>4000±2000</td>
<td>7.6±0.5</td>
<td>9.7±0.2</td>
<td>n/a</td>
</tr>
<tr>
<td>111</td>
<td>1.6±0.2</td>
<td>940±140</td>
<td>7.7±0.2</td>
<td>9.6±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>125</td>
<td>0.4±0.02</td>
<td>330±20</td>
<td>8.0±0.1</td>
<td>9.4±0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>126</td>
<td>2.8±0.3</td>
<td>4400±500</td>
<td>8.3±0.1</td>
<td>9.9±0.1</td>
<td>−5.2±0.5</td>
</tr>
<tr>
<td>Average</td>
<td>3.2±2.6</td>
<td>3300±2400</td>
<td>8.0±0.4</td>
<td>9.9±0.4</td>
<td>−4±2</td>
</tr>
</tbody>
</table>

n/a = Not available.

*a* Obtained by fitting the kinetic NHX data in Fig. 2 using Eq. (2) (20 °C in H$_2$O 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* $\Delta G_{HX} = \Delta G_{op} = RT \ln(k_{op}) = RT \ln(k_{op}/k_{cl})$.

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

*d* Calculated from equilibrium NHX data for SNase PHS at 25 °C in D$_2$O with 50 mM sodium acetate and 100 mM NaCl at pHread 5.0 (personal communication and Wrabl38).

*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 \( \Delta G_{HX} \) (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 (\( \Delta G_{HX} \)) 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 \( \Delta G_{HX} \), 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 \( \beta 5 \) strand other than the strand termini, their naturally paired H-bonded amides in the neighboring \( \beta 3 \) and \( \beta 4 \) strands, and \( \alpha 2 \) amides that intersect \( \beta 5 \). 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 \( \beta 5 \) 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 \( \alpha 2 \) helix (Fig. 4). The stability (\( \Delta G_{HX} \)) 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 \( \beta 2-\beta 3 \) hairpin and the contiguous part of \( \beta 1 \), which

![Fig. 3. Two-dimensional scatter plot separation of foldon units. \( \Delta G_{HX} \) 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.](image-url)
form a mutually H-bonded three-stranded $\beta$-meander sequence, and also the abutting C-terminus of $\alpha_1$ (Fig. 4b). Five of these residues were observed by Wrabl and found to exchange by way of a large unfolding reaction (large $m$ value). $\alpha_1$ residues 64, 65, and maybe 69 (all in $\alpha_1$), suggesting a longer length of the $\alpha_1$ helix within the yellow foldon. 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 $\beta$-sheet plus a part of $\alpha_3$ (Fig. 4a, inset). The $\Delta G_{1HX}$ 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 $k_{d}=k_{cl}$, so that $k_{d}=1/2k_{op}$; 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 artificially 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 large unfolding and refolding reactions under native structural elements, called foldons, that engage in Protein foldons

**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 ($ΔG_{mol}$=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 finitively observed. It is hard to picture how a major complete secondary structural segments than is de-

ingly, it seems likely that the foldons entrain more 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, 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 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 E. 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 [15N]SNase and [13C]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. Buffers contained 20 mM borate at pH 9. Cell lystate 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 [15N]SNase, and 20 mg/L for [15N,13C]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 (80 mM) were K2HPO4 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 (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. Buffers contained 20 mM borate at pH 9. Cell lystate 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 [15N]SNase, and 20 mg/L for [15N,13C]SNase.

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

For the assignment of main-chain (1Hα, 15N, and 13Cα) and 13Cβ crosspeaks, CBCA(CO)NH71,72 and HNCACB72,73 spectra of uniformly 15N- and 13C-labeled proteins were collected at pH 5.3, pH 8.0, and pH 10.0 in 0.1 M NaCl and buffer (with CaCl₂ and excess (3′-deoxy-3′-phosphate 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, 15N]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 D₂O, holding it unfolded for 10 min at a pD₀ of 12, refolding at a pD₀ of 5.3, and lyophilizing. The process was repeated. For HX experiments, pH was corrected to take into account the amount of D₂O [pHcorr = pHr + 0.4(fraction D₂O)].

For pulse-quench experiments, lyophilized deuterated protein was dissolved in D₂O buffer (10 mM acetic acid) and buffer (with CaCl₂ and excess (3′-deoxy-3′-phosphate 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, 15N]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†.

HX theory

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

Closed (D)→Opened (D)→Opened (H)→Closed (H)

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):53,75

\[ k_{\text{ex}} = (k_{\text{op}}k_{\text{ch}})/(k_{\text{cl}} + k_{\text{ch}}) \]

where \( k_{\text{ch}} = \text{intrinsic rate constant} \) and \( k_{\text{ch}} \) is the known intrinsic chemical HX rate calibrated for the equivalent residue amide at the experimental condition.6–8

In the transiently open condition, a kinetic competition between exchange and reclosing ensues. If reclosing is faster (\( k_{\text{ch}} > k_{\text{cl}} \)), the structural opening reaction appears as a pre-equilibrium 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, \( K_{\text{op}} \).

\[ k_{\text{ex}} = (k_{\text{op}}k_{\text{ch}})/k_{\text{cl}} = K_{\text{op}}k_{\text{ch}} \]

From the Boltzmann relationship (\( \Delta G_{\text{HX}} = -RT \ln K_{\text{op}} \)), 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 \( k_{\text{ex}} > k_{\text{cl}} \). Equation (2) then reduces to Eq. (4), and the opening rate can be obtained:33

\[ k_{\text{ex}} = k_{\text{op}} \]

These equations hold under steady-state conditions where the protecting structure is stable 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 \( k_{\text{ex}} \) reaches 10⁶ 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 \( k_{\text{ex}} \) 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-
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 ($k_{on}$ and $k_{off}$) and equilibrium ($\Delta G^{\text{HX}}$) 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


units and their stepwise sequential stabilization.  


*J. Mol. Biol.* **359**, 1411–1420.


*Biochemistry,* **32**, 5222–5232.

*Biochemistry,* **27**, 8063–8068.


*Biochemistry,* **11**, 150–158.
