The folding energy landscape of apoflavodoxin is rugged: Hydrogen exchange reveals nonproductive misfolded intermediates

Yves J. M. Bollen*, Monique B. Kamphuis†§, and Carlo P. M. van Mierlo*‡

*Department of Structural Biology, Vrije Universiteit, 1081 HV, Amsterdam, The Netherlands; and †Laboratory of Biochemistry, Wageningen University, 6703 HA, Wageningen, The Netherlands

Edited by Alan R. Fersht, University of Cambridge, Cambridge, United Kingdom, and approved January 23, 2006 (received for review October 19, 2005)

Many native proteins occasionally form partially unfolded forms (PUFs), which can be detected by hydrogen/deuterium exchange and NMR spectroscopy. Knowledge about these metastable states is required to better understand the onset of folding-related diseases. So far, not much is known about where PUFs reside within the energy landscape for protein folding. Here, four PUFs of the relatively large apoflavodoxin (179 aa) are identified. Remarkably, at least three of them are partially misfolded conformations. The misfolding involves side-chain contacts as well as the protein backbone. The rates at which the PUFs interconvert with native protein have been determined. Comparison of these rates with stopped-flow data positions the PUFs in apoflavodoxin’s complex folding energy landscape. PUF1 and PUF2 are unfolding excursions that start from native apoflavodoxin but do not continue to the unfolded state. PUF3 and PUF4 could be similar excursions, but their rates of formation suggest that they are on a dead-end folding route that starts from unfolded apoflavodoxin and does not continue all of the way to native protein. All PUFs detected thus are off the protein’s productive folding route.

Hydrogen/deuterium exchange | partially unfolded forms | protein folding

Protein folding can be described by using a free energy landscape model (1, 2). In this model an unfolded protein molecule descends along a funnel describing its free energy until it reaches the state with the lowest free energy, which is the native state. The landscape often contains local minima, which host folding intermediates. Whereas the native state of many proteins is well characterized, little is known about the metastable intermediate states and their positions along the folding funnel. Knowledge about these metastable states is necessary to improve the understanding of protein folding and folding-related diseases, because rarely populated partially folded forms of proteins are known to initiate the formation of amyloids (3).

Detection by nuclear magnetic resonance (NMR) spectroscopy of native state hydrogen/deuterium exchange (H/D exchange), i.e., in the presence of small amounts of a denaturant, allows the characterization of partially unfolded forms (PUFs) of a protein (4). These PUFs are interpreted to reside in high-energy minima in the folding energy landscape (5, 6). PUFs are often undetectable by other techniques, because, for example, they reside behind the highest-energy transition state for folding and because they do not populate significantly at equilibrium. Analysis of H/D exchange data with models that statistically sample the conformational space accessible to a protein can give insight into the high-energy conformations that a protein can adopt (7–9). However, little evidence exists that links PUFs to the folding intermediates that populate during equilibrium unfolding, kinetic unfolding, or kinetic refolding of proteins. This missing link makes it difficult to position PUFs within the energy landscape for protein folding (10).

Recently, both the denaturant-induced equilibrium unfolding and the kinetic folding of Azotobacter vinelandii flavodoxin have been characterized (11–15). Apoflavodoxin kinetic folding is described by (13)

\[
I_1 \Rightarrow U \Rightarrow I_2 \Rightarrow N,
\]

where U is unfolded protein, N is native apoflavodoxin, and I₁ and I₂ are folding intermediates. In the presence of FMN, first apoflavodoxin folds to its native state according to Eq. 1 and then FMN binds to form holoflavodoxin (15). The rate constants shown in Eq. 1 are in s⁻¹ and are obtained by extrapolation to 0 M GuHCl. The rate constants that describe departure from I₂, either to N or to U, are not absolute. The kinetic data obtained only allow the determination of their ratio (13). The molten globule-like intermediate I₁, which also populates to significant extents at GuHCl concentrations ranging from 1 to 3 M, acts as a trap. It has to unfold before native apoflavodoxin can be formed. Intermediate I₂ is highly unstable and thus is not observed during denaturant-induced equilibrium unfolding of apoflavodoxin. The formation of intermediates I₁ and I₂ during folding is linked to apoflavodoxin’s α-β parallel topology (14).

Here, H/D exchange measurements are used to characterize PUFs of the 179-residue A. vinelandii apoflavodoxin. A detailed picture, which includes the position of apoflavodoxin’s PUFs, is obtained of the folding energy landscape accessible to this relatively large protein. All PUFs turn out to be off the protein’s productive folding route.

Results

Backbone amide proton H/D exchange rates (kex) are determined for apoflavodoxin at deuterated guanidinium chloride (GuDCl) concentrations that range from 0 to 750 mM using NMR spectroscopy and converted into free energy differences for local opening of the protein structure, ΔGop (see Table 2, which is published as supporting information on the PNAS website). This conversion is valid when exchange takes place according to the EX2 mechanism, which is the case for apoflavodoxin at pD 5.7 to 5.9 (where pD is the pH meter reading of a deuterated solution), the pD range used in this study (see below).

Identification of Cooperatively (Un)Folding Clusters Within Apoflavodoxin. Some apoflavodoxin residues display a linear dependence of ΔGop on GuDCl concentration over the complete...
on the PNAS web site). Other residues have mental Procedures either from the globally unfolded state (type 1 residues) or from residues identified by similarity of the corresponding energy isotherm of the cluster involved. The isotherms for the four identified isotherms at GuDCl concentrations from a fit of a two-process model (see Eq. 4096).

Fig. 2. Native apoflavodoxin and its partially unfolded forms as detected by native-state H/D exchange data of amino acid residues of apoflavodoxin. (a) Typical examples of $\Delta G_{\text{op}}$ data that have a linear or curved dependency on GuDCl concentration (filled and open circles, respectively). Cluster 1 within apoflavodoxin (yellow) is represented by V141 (filled circles) and A140 (open circles), Cluster 2 (green) by A150 (filled circles) and K145 (open circles), Cluster 3 (blue) by F6 (filled circles) and V17 (open circles), and Cluster 4 (red) by L93 (filled circles) and F49 (open circles). (b) Clusters of residues identified by similarity of the corresponding $\Delta G_{\text{op}}(0)$ and $m$ values, respectively (errors are standard fitting errors). Only those residues for which the GuDCl-dependency of their $\Delta G_{\text{op}}$ data is linear in the GuDCl-range studied are taken into account. Four clusters are identified as follows: Cluster 1 (yellow; V141), Cluster 2 (green; F146, V147, G148, L149, and A150), Cluster 3 (blue; G4, L5, F6, I21, and K22), and Cluster 4 (red; I51, L52, G53, V91, A92, L93, F94, W167, H168, A169, Q170, I171, and A172). The average $\Delta G_{\text{op}}(0)$ and $m$ values of a particular cluster define the linear GuDCl-dependence of the free energy isotherm of the cluster involved. The isotherms for the four identified clusters are shown in a as solid lines. The curvature isotherms in $a$, which result from a fit of a two-process model (see Eq. S7 in Supporting Experimental Procedures) to the curving data, are dashed and join one of the linear isotherms at GuDCl concentrations $\cong 0$ M.

At low GuDCl concentration, the amides of these residues exchange through local opening of the native conformation (type 3 residues). Their amides exchange either from the globally unfolded state (type 1 residues) or from a PUF of the protein (type 2 residues; see Supporting Experimental Procedures, which is published as supporting information on the PNAS web site). Other residues have $\Delta G_{\text{op}}$ values that have a linear dependency on GuDCl concentration (Fig. 1a). At low GuDCl concentration, the amides of these residues exchange through local opening of the native conformation (type 3 residues).

Residues for which the $\Delta G_{\text{op}}$ values depend linearly on GuDCl concentration. For each residue that displays a linear dependence of $\Delta G_{\text{op}}$ on GuDCl concentration, the $\Delta G_{\text{op}}$ value at 0 M GuDCl [i.e., $\Delta G_{\text{op}}(0)$] and the corresponding slope ($m$) value are shown in Fig. 1b. Residues are clustered and colored based on similarities of their $\Delta G_{\text{op}}(0)$ and $m$ values, respectively. Residues in a particular cluster exchange their amides from the same (partially or globally) unfolded form of the protein (see Supporting Experimental Procedures). Four clusters are thus identified. Each cluster is characterized by a typical average $\Delta G_{\text{op}}(0)$ and average $m$ value (described below in Features of the four cooperatively unfolding clusters identified). These averaged $\Delta G_{\text{op}}(0)$ and $m$ values are all smaller than the stability against global unfolding of apoflavodoxin ($\Delta G_{N-U}$) and the corresponding GuHCl-dependence ($m_{N-U}$), respectively. This finding implies that none of the amides of the identified clusters exchanges through global protein unfolding, the characteristic of type 1 residues. The residues in Fig. 1b are thus all type 2 residues, and each of the four clusters unfolds through a specific subglobal process.

Residues for which the $\Delta G_{\text{op}}$ values have a curving dependency on GuDCl concentration. At low GuDCl concentrations, the $\Delta G_{\text{op}}$ values of the type 3 residues of apoflavodoxin change little with increasing GuDCl concentration, because local opening processes are the dominant mechanism for exchange of type 3 residues. These residues mainly reside in loops and at the edges of elements of secondary structure of apoflavodoxin, where it is reasonable to envisage local unfolding events. In contrast, type 2 residues (Fig. 1b and Table 2) are in the core of secondary-structure elements of apoflavodoxin.

At higher GuDCl concentrations, exchange of type 3 amide protons occurs through subglobal unfolding of the protein. As a consequence, the corresponding $\Delta G_{\text{op}}$ values curve toward the linear isotherms that characterize the unfolding of the four identified clusters (Fig. 1a). Type 3 residues can be assigned to belong to one of the subglobally unfolding clusters identified in Fig. 1b on the basis of the quality of the fit of a two-process model (see Eq. S7 in Supporting Experimental Procedures) to their curving $\Delta G_{\text{op}}$ data (assignments are listed in Table 2) (16, 17). In this model, amides can exchange from a (sub)globally unfolded conformation or a locally opened native-like conformation.

Features of the four cooperatively unfolding clusters identified. Cluster 1 within apoflavodoxin (yellow in Figs. 1 and 2a) is characterized...
by $\Delta G_{op}(0) = 4.87 \pm 0.03 \text{ kcal/mol}$ and $m = -2.3 \pm 0.1 \text{ kcal/mol.M}^{-1}$ (errors indicated are standard errors). Only one of its residues, V141, has up to 0.75 M a linear dependence of its $\Delta G_{op}$ on GuDCl concentration, characteristic for type 2 residues. Neighboring A140 has a somewhat lower $\Delta G_{op}(0)$ than V141, and its isotherm curves smoothly toward the one of V141. A140 and also V142 are type 3 residues belonging to Cluster 1. The three residues discussed form one of the two strands of a small $\beta$ sheet that is part of a loop typical for long-chain flavodoxins (Fig. 2a).

This loop connects $\beta$-strands 5A and 5B.

Cluster 2 consists of the other strand of the small $\beta$-sheet discussed and also comprises $\beta$-strand 5B. It is colored green in Figs. 1 and 2a and is characterized by $\Delta G_{op}(0) = 6.14 \pm 0.05 \text{ kcal/mol and } m = -3.1 \pm 0.1 \text{ kcal/mol.M}^{-1}$.

Cluster 3 (blue in Figs. 1 and 2a) is characterized by $\Delta G_{op}(0) = 7.95 \pm 0.09 \text{ kcal/mol and } m = -5.7 \pm 0.2 \text{ kcal/mol.M}^{-1}$. It mainly involves a large part of the first 24 N-terminal residues of apoflavodoxin. Of the residues belonging to Cluster 3, the $\Delta G_{op}(0)$ of G4 is significantly higher than the $\Delta G_{op}(0)$ values of the other residues (Fig. 1b). Because the $m$ value associated with G4 is identical within error to the average $m$ value of the residues belonging to Cluster 3, G4 is considered to be a member of this cluster as well. G4 seems to have residual protection against exchange when Cluster 3 is unfolded; its $\Delta G_{op}(0)$ and $m$ value are not used for the determination of the average $\Delta G_{op}(0)$ in Clusters 3.

Identification of a fifth cluster.

The residues not yet categorized (i.e., L110, G111, Y114, S115, F116, and F117) reside in $\alpha$-helix 4 of apoflavodoxin (purple in Fig. 2a) and have curve $\Delta G_{op}$ isotherms. Interestingly, these curves (except the one for L110) lie at high GuDCl concentrations above the isotherms of the four already-identified clusters within apoflavodoxin are shown by dashed lines.

Table 1. Thermodynamic parameters of the apoflavodoxin folding states identified

<table>
<thead>
<tr>
<th>Species</th>
<th>$\Delta G_{N-U}(0)$</th>
<th>$m$ value</th>
<th>$\alpha$ value</th>
<th>$k_{N-X}$</th>
<th>$k_{X-N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6.69 ± 0.01</td>
<td>-4.38 ± 0.02</td>
<td>0.29</td>
<td>7.4 × 10^{-5}</td>
<td>0.733</td>
</tr>
<tr>
<td>I1</td>
<td>10.58 ± 0.06</td>
<td>-6.22 ± 0.02</td>
<td>0.08-0.9</td>
<td>3.2</td>
<td>$\approx 1 \times 10^{3}$</td>
</tr>
<tr>
<td>I2</td>
<td>7.95 ± 0.09</td>
<td>-5.7 ± 0.2</td>
<td>0.09</td>
<td>(7 ± 4) × 10^{-5}</td>
<td>6 ± 4</td>
</tr>
<tr>
<td>U</td>
<td>8.67 ± 0.05</td>
<td>-4.6 ± 0.1</td>
<td>0.27</td>
<td>(4 ± 1) × 10^{-5}</td>
<td>13 ± 4</td>
</tr>
</tbody>
</table>

Listed are the free energy $[\Delta G_{N-U}(0), \text{ kcal/mol}]$ associated with the formation of a particular state starting from native apoflavodoxin in absence of denaturant, the denaturant dependence of $\Delta G_{N-U}$ (i.e., the $m$ value, kcal/mol.M^{-1}), the $\alpha$ value (i.e., the normalized $m$ value), the rate constant $k_{N-X}$ (s^{-1}) at which the species refolds from native apoflavodoxin, and the rate constant $k_{X-N}$ (s^{-1}) at which the species refolds to native apoflavodoxin. Data regarding N, I1, I2, and U are taken from refs. 13 and 15; the relative free energy of the unfolded state is corrected for prolyl peptide bond isomerization, as described in Results.
PUFs of Apoflavodoxin. Native-state H/D exchange data identify five clusters of residues within apoflavodoxin. Four of these clusters unfold subglobally in a cooperative manner. The resulting conformations are thus PUFs of the protein (4). It has been suggested that PUFs are snapshots of productive protein folding, forming sequentially in the order of their stability difference with respect to the native state (6, 22). This order is thought to be caused by the cooperative nature of protein folding (6). Assuming that such sequential unfolding occurs for native apoflavodoxin, conformations of its PUFs are identified (Fig. 2b).

The slope of the GuDCl-dependence of $\Delta G_{\text{N-PUF}}$, which is the free energy difference between N and a particular PUF (Figs. 1 and 3), is called $m_{\text{N-PUF}}$. The value of $m_{\text{N-PUF}}$ informs about the PUF's denaturation accessibility relative to the denaturant accessibility of native protein, because $m$ values scale to changes in accessible surface area (23). The denaturant accessibility of a PUF is usually expressed as the $\alpha$ factor, which is the normalized $m$ value of the specific PUF:

$$\alpha = (m_{\text{PUF-U}}/m_{\text{N-U}}) = (m_{\text{N-U}} - m_{\text{N-PUF}})/m_{\text{N-U}},$$

where $m_{\text{N-U}}$ is the $m$ value associated with global unfolding of apoflavodoxin (13, 15) and $m_{\text{PUF-U}}$ is the one associated with unfolding of a specific PUF to unfolded apoflavodoxin. When $\alpha = 0$, the corresponding PUF is as denaturant-accessible as fully unfolded apoflavodoxin, and when $\alpha = 1$ it is as denaturant-accessible as native apoflavodoxin. The PUFs $\alpha$ values are listed in Fig. 2b.

The thermodynamic parameters associated with the identified folding states of apoflavodoxin are listed in Table 1.

Folding and Unfolding Rates of the Clusters Identified Within Apoflavodoxin. The rate constant for the conversion of native protein to a specific PUF can be determined. Use is made of pD-dependent native-state H/D exchange rates of type 2 residues of the cluster involved (24-27). At elevated pD values, exchange enters the EX1 regime (see Fig. 4a). The fit of the equation shown in Fig. 4a to the curving pD-dependent $k_{\text{ex}}$ data leads to the determination of $k_{\text{op}}$ and $k_{\text{cl}}$ of the cluster involved. In the case of type 2 residues, $k_{\text{op}}$ equals the rate constant associated with unfolding of native apoflavodoxin to a specific PUF, $k_{\text{N-PUF}}$, and $k_{\text{cl}}$ equals the rate constant associated with refolding of the PUF in question (26). Comparison of the latter rate constants
with those obtained from stopped-flow folding data of a protein can give clues about the position a PUF has within the energy landscape for folding.

H/D exchange rates of backbone amide protons of apoflavodoxin are determined in the absence of GuDCl at pD values of 6.28, 6.80, and 8.23 (a similar change of pH from 6 to 8 marginally destabilizes native apoflavodoxin because the thermal midpoint of its unfolding decreases by only 2.5 K). Unfortunately, the amide protons of the type 2 residues of both Clusters 1 and 2 exchange too rapidly to be tracked at pD 6.80 and 8.23. Their exchange apparently increases significantly with increasing pD because they can be tracked at pD 6.28, and EX2 behavior is likely. Consequently, the closing rates of these residues, which equal the refolding rates of PUF1 and PUF2, respectively, need to be fast compared with their \( k_{\text{int}} \) values, which are at pD 6.28 between 0.1 and 1 s\(^{-1}\).

In case of Cluster 3, pD-dependent amide proton exchange could be determined for four of its five type 2 residues, which leads to: \( k_{\text{op}} = (7 \pm 4) \times 10^{-5} \text{ s}^{-1} \) and \( k_{\text{cl}} = 6 \pm 4 \text{ s}^{-1} \) (Fig. 4b). The \( k_{\text{op}} \) value equals the rate constant at which PUF3 is formed from N, whereas PUF3 refolds at rate constant \( k_{\text{cl}} \) (26). Analogously, in case of Cluster 4 global fitting of the pD-dependent H/D exchange data of all of its 13 type 2 residues leads to: \( k_{\text{op}} = (4 \pm 1) \times 10^{-5} \text{ s}^{-1} \) and \( k_{\text{cl}} = 13 \pm 4 \text{ s}^{-1} \) (Fig. 4b). Cluster 5 only contains type 3 residues. Therefore, the rate constant for global folding and unfolding of apoflavodoxin cannot be determined directly by H/D exchange.

The GuDCl-dependent exchange data, recorded at pD 5.7–5.9, are analyzed assuming that exchange takes place according to the EX2 mechanism. The pD-dependent exchange data validate this assumption. At pD 5.7–5.9, the \( k_{\text{int}} \) values are all <1.7 s\(^{-1}\). At these \( k_{\text{int}} \) values the residues shown in Fig. 4b exchange according to the EX2 mechanism. This finding also holds for all other apoflavodoxin residues except for F25 (data not shown).

Discussion

Position of the PUFs Identified Within the Energy Landscape for Apoflavodoxin Folding. A two-dimensional model of the energy landscape for apoflavodoxin folding has been constructed by using the reported denaturant-induced equilibrium and stopped-flow apoflavodoxin (un)folding results (Fig. 5a) (13). Comparison of the rate constants for interconversion between native apoflavodoxin and its PUFs with the rate constants for apoflavodoxin (un)folding extracted from stopped-flow data allows the positioning of the PUFs within this folding energy landscape.

Positioning PUF1 and PUF2. The rate constants at which two apoflavodoxin species interconvert can be calculated from Eq. 1. Unfolded apoflavodoxin (U) forms native apoflavodoxin (N) with a rate constant \( k_{\text{U,N}} \) of 7.3 \( \times 10^8 \text{ s}^{-1} \), and N visits U with an unfolding rate constant \( k_{\text{N,U}} \) of 0.013 s\(^{-1}\), respectively (see Supporting Experimental Procedures). Any folding intermediate that is positioned on the route between U and N must form native apoflavodoxin with a rate constant that is equal to or larger than \( k_{\text{U,N}} \). Such a high refolding rate implies that backbone amide protons that exchange from U, I2, or from any other potential apoflavodoxin folding intermediate that lies on the productive folding route between N and U are expected to display EX2 exchange behavior within the pD range studied (see Fig. 4a).

As discussed, EX2 behavior seems likely for the residues of Clusters 1 and 2 of apoflavodoxin. Both PUF1 and PUF2 thus could be on the productive folding route that links U with N, but PUF1 and PUF2 are too accessible for GuDCl to coincide with the kinetic apoflavodoxin folding intermediate I2 (Table 1).

Fig. 5b shows a free energy landscape for apoflavodoxin folding with PUF1 and PUF2 positioned on the productive folding route, taking into account their free energies and denaturant accessibilities. Because both PUFs are significantly more denaturant accessible than the rate-limiting transition state for folding (its \( \alpha \)-value is 0.73 (13)), they reside on the unfolded side of this transition state, which is marked \( \uparrow \). Both PUFs are significantly more stable than U. As a result, the transition from PUF1 via \( \uparrow \) to N would be the rate-limiting step in folding from U to N. This folding would be much slower (0.029 s\(^{-1}\); see legend of Fig. 5b) than experimentally observed (7.3 \( \times 10^8 \text{ s}^{-1}\)). Such slow folding would cause EX1 exchange behavior for amides that exchange from PUF1 and PUF2 at all pD values used here, which is not observed. Consequently, PUF1 and PUF2 cannot reside on the productive folding route of apoflavodoxin.

PUF1 and PUF2 thus must be excursions from the native state that are separate from the (un)folding route that links native and globally unfolded apoflavodoxin. This result is schematically shown in Fig. 5c.

Positioning PUF3 and PUF4. None of the backbone amide protons of residues that belong to Clusters 3 and 4 of apoflavodoxin display pure EX2 H/D exchange behavior over the entire pD range studied. PUF3 and PUF4 of apoflavodoxin thus are not on the productive folding route between U and N.

PUF3 and PUF4 could be excursions from the native state separate from those leading to PUF1 and PUF2. Remarkably, however, \( k_{\text{N,PUF3}} \) and \( k_{\text{N,PUF4}} \) are much smaller than \( k_{\text{N,U}} \), whereas \( k_{\text{U,PUF3}} \) and \( k_{\text{N,PUF4}} \) are within error identical to \( k_{\text{N,U}} \) (Table 1). An explanation for this observation is that PUF3, PUF4, and I4 reside in a free energy region that is separated from unfolded apoflavodoxin by the same high barrier (Fig. 5c). PUF3 and PUF4 do not correspond to I1, because they are too unstable, too accessible for GuDCl, and fold too rapidly (Table 1). Apparently, both PUF3 and PUF4, as well as I1, need to unfold before productive folding occurs.

PUFs Are Not Snapshots of Apoflavodoxin’s Productive Folding. PUFs were suggested to be snapshots of productive protein folding and to fold one after the other in a strictly hierarchical order determined by their individual free energy values, where the folding of one PUF facilitates the formation of the next one (5, 6, 22). However, apoflavodoxin’s pD-dependent exchange data clearly show that its PUFs do not reside on apoflavodoxin’s productive folding route. Furthermore, note that in Fig. 4a the lines that represent unfolding of native apoflavodoxin to PUF2 (i.e., the green line) and to PUF3 (i.e., the blue line) cross one another at 0.7 M GuDCl. This result implies that above 0.7 M GuDCl PUF3 is visited more often than PUF2, which conflicts with the idea that the stabilities of PUFs represent a sequential (un)folding order. Apoflavodoxin PUFs thus are not snapshots of apoflavodoxin’s productive folding.

Apoflavodoxin PUFs Are Partially Misfolded. The observation that the PUFs of apoflavodoxin do not develop from one another in a sequential manner means that the representations of the conformations of these PUFs in Fig. 2b are speculative. For example, in the case of PUF2, it is assumed in Fig. 2b that both Cluster 2 and Cluster 1 are unfolded. However, when PUF1 and PUF2 would be separate excursions from the native state, Cluster 1 needs not necessarily be unfolded in PUF2. In addition, the folded parts of a PUF need not necessarily be native-like. In fact, the H/D exchange data presented here show that the PUFs of apoflavodoxin are partially misfolded for the following reasons. (i) In PUF1, one of the two strands of a small \( \beta \)-sheet is unfolded, whereas the second strand remains folded. The residues in the remaining folded strand can only be shielded from water by making nonnative hydrogen bonds with other folded parts of the protein. (ii) In PUF2, both the long loop typical for long-chain flavodoxins and \( \beta \)-strand 5B are unfolded, whereas the backbone amide of a single residue in this part of the protein (i.e., W128) remains protected against exchange. In native
apoflavodoxin, the backbone amide of W128 is hydrogen-bonded to the carbonyl group of F146. F146 is water accessible in PUF2. Consequently, in PUF2, W128 must make a nonnative hydrogen bond. (iii) β-Strand 2 in PUF3, a strand that is at the periphery of apoflavodoxin, apparently is natively folded, whereas the neighboring β-strand 1 is unfolded. The backbone amides of β-strand 2 can only be protected against exchange in PUF3 by making nonnative hydrogen bonds.

In conclusion, three PUFs of apoflavodoxin represent partially misfolded conformations. The misfolding does not only involve side-chain contacts, as was recently observed for the PUFs of a four-helix bundle protein (29), but also involves the backbone of secondary structure elements of apoflavodoxin. The apoflavodoxin data thus show that the original idea (5, 6, 22) that PUFs represent a hierarchical folding order in which the parts of the PUFs that are folded are all native-like no longer holds.

It is striking that all four PUFs identified reside on folding and unfolding routes that differ from the one that links native and globally unfolded apoflavodoxin and that at least three of these PUFs are partially misfolded. Both observations probably are related because partially misfolded intermediates need to unfold to continue the route to the native state.

Rugged Folding Energy Landscape of Apoflavodoxin. The schematic energy landscape presented in Fig. 5c is consistent with the experimental apoflavodoxin equilibrium and kinetic folding data (13) as well as with the H/D exchange data presented here. However, the two-dimensional model is most likely an oversimplification of reality. In Fig. 5d an alternative representation of apoflavodoxin’s folding energy landscape is shown in which the known minima and the likely folding trajectories are presented. Productive apoflavodoxin folding proceeds along the solid lines that link U with N via the high-energy intermediate I2. No direct folding from I1 to N is observed (13). This kinetic mechanism for folding is shared by proteins with the flavodoxin-like topology (14). PUF3 and PUF4 have free energy values and denaturant accessibilities that position these species in a region between U and I1 in Fig. 5d. PUF3 and PUF4 thus could be PUFs of apoflavodoxin’s relatively stable, long-living folding intermediate I1. This hypothesis is supported by k_{N→PUF3} and k_{N→PUF4} being identical within error to k_{N→I1} and by k_{PUF3→N} and k_{PUF4→N} being slightly larger than k_{I1→N} (Table 1). However, although unlikely, PUF3 and PUF4 could be rare unfolding excursions that start directly from native apoflavodoxin (Fig. 5d).

Conclusion
A detailed view has been obtained of the folding energy landscape of A. vinelandii apoflavodoxin. The energy landscape is mapped by denaturant-induced equilibrium unfolding, stopped-flow folding and unfolding, and H/D exchange data under both EX2 and EX1 conditions. The PUFs of apoflavodoxin detected with H/D exchange are partially misfolded and all turn out to be off the protein’s productive folding route. Apoflavodoxin folding involves a search of many local minima in its complex folding-energy landscape. The search includes a population of a stable intermediate. Apparently, this search is characteristic for the folding of the many proteins that share the flavodoxin-like topology.

Methods
Monomeric uniformly 15N-labeled C69A A. vinelandii apoflavodoxin was obtained as described in refs. 11, 13, and 30. Lyophilized apoflavodoxin was dissolved in D2O containing 100 mM sodium pyrophosphate and various GuDCl concentrations ranging from 0 to 750 mM and KCl in various amounts to obtain a final salt concentration (i.e., GuDCl + KCl) of 750 mM. The final concentrations of protein are ±1 mM. Samples were immediately transferred into an AMX500 NMR spectrometer (Bruker, Rheinstetten, Germany) that operates at a proton frequency of 500.13 MHz. Subsequently, a series of gradient-enhanced 1H-15N Heteronuclear single-quantum coherence spectra (31, 32) were recorded to detect H/D exchange. The dead-time was ~5 min, and the temperature in the probe was 25°C. The pD was measured after exchange was complete. The spectra were processed, and a single-exponential function was fitted to the time-dependent peak intensities to obtain k_{ex}. Further details are provided in Supporting Experimental Procedures, together with the theory behind H/D exchange and Table 2.

We thank J. P. Mackay, H. Lill, and I. E. Sánchez for constructive comments.