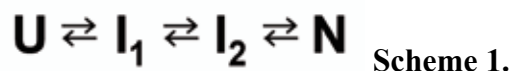


Supporting Text

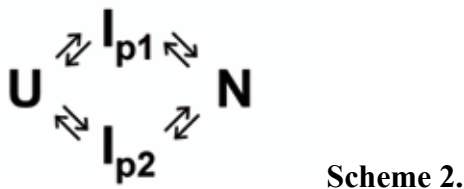
Sequential Folding Scheme of SMN. We observed three kinetic phases in the folding of single-chain monellin (SMN) using circular dichroism (CD), small-angle x-ray scattering (SAXS), Trp, and 1-anilino-naphthalene-8-sulfonate (ANS) fluorescence. The kinetic CD results suggested the presence of the burst phase that completes within 300 μ s and showed the first and second phases with time constants of ≈ 14 ms and ≈ 1.3 s, respectively (Fig. 1B). Both Trp-4 and ANS fluorescence data showed the first and second phases (Fig. 2A and B). The kinetic radius of gyration (R_g) trace indicated the major change in the burst phase and the minor change in the first phase (Fig. 3A). To explain the observed three phases, several kinetic models can be proposed. We will examine each of the possible models in turn below.

The first model is the sequential scheme with two folding intermediates (Scheme 1):



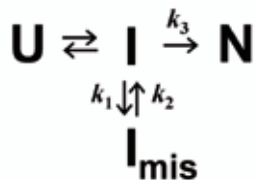
In this scheme, the formation of I_1 , I_2 , and the native state (N) corresponds to the burst, first, and second phases, respectively. As explained in the main text, this scheme can consistently explain all of the observations made in the current study.

The second possibility is the parallel folding scheme (Scheme 2), in which two folding intermediates, I_{p1} and I_{p2} , are formed in the burst phase, and the independent conversion of I_{p1} and I_{p2} to N creates the first and second kinetic phases, respectively.



If Scheme 2 is correct, we can firstly deduce that the CD spectra for both I_{p1} and I_{p2} are similar to that of U due to the similarity between the kinetic CD spectrum at 300 μ s and the static spectrum for the initial unfolded state (Fig. 1A). We can further deduce that more than half of the protein should be in the native state after the first kinetic phase, because the CD amplitude for the first phase is 57% (Fig. 2B). However, this is contradictory to the results of the kinetic titration experiments (Fig. 2C), in which more than 90% of SMN binds with ANS at 100 ms after the pH jump, demonstrating the absence of the native component after the first phase. The triangle schemes are also unlikely for the same reason.

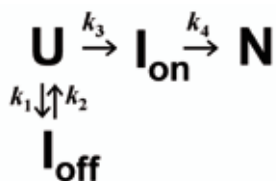
The third possibility is the dead-end model (Scheme 3), in which a misfolded component (I_{mis}) is in kinetic equilibrium with an on-pathway intermediate (I).



Scheme 3

To fit the observed three phases to Scheme 3, we have to assign the burst, first, and second kinetic phases to the formation of I from U, I_{mis} from I, and N from I_{mis} through I. From the apparent rate constant for the first phase that is well separated from the second phase, k_1 can be roughly estimated as $\approx 1/14 \text{ ms}^{-1} = 70 \text{ s}^{-1}$. The kinetic titration experiment demonstrated that I_{mis} populates more than 90% at 100 ms (Fig. 2C), giving the upper limit of k_2 and k_3 , that is, $k_2 + k_3 \leq 0.1 \cdot k_1$. In the time domain where the steady-state approximation for I is valid, the apparent formation rate of N can be written as $k_2 \cdot k_3 / (k_1 + k_2 + k_3)$. Under the restrictions that $k_1 \approx 70 \text{ s}^{-1}$ and $k_2 + k_3 \leq 0.1 \cdot k_1$, the largest possible value for $k_2 \cdot k_3 / (k_1 + k_2 + k_3)$ is $\approx 0.16 \text{ s}^{-1}$. This is much smaller than the observed rate constant for the appearance of N ($\approx 1/1.3 \text{ s}^{-1} = 0.77 \text{ s}^{-1}$). Thus, it is difficult to interpret the observed rate constants based on the dead-end model.

The fourth possibility is another dead-end model (Scheme 4), in which an off-pathway component (I_{off}) is in kinetic equilibrium with the unfolded state (U) and on-pathway intermediate (I_{on}), which are formed during the folding process from U to N.



Scheme 4

We found that Scheme 4 can also explain the current observations, if we assume that k_1 , k_2 , k_3 , and k_4 are 30,000, 1,500, 1,500, and 0.77 s^{-1} , respectively. This is understandable, because the above combination of the rate constants shows that Scheme 4 reproduces the stepwise folding, in which the burst ($< 300 \mu\text{s}$), first (70 s^{-1}), and second (0.77 s^{-1}) kinetic phases correspond to the formations of I_{off} from U, I_{on} from I_{off} through U, and N from I_{on} , respectively. Thus, we cannot rule out Scheme 4 as an alternative possibility of the current observations. We point out, however, that the structural properties of I_{off} and I_{on} in Scheme 4 correspond to those of I_1 and I_2 in Scheme 1, respectively. Therefore, the conformational landscape of SMN based on Scheme 4 should be very similar to that based on Scheme 1 presented in Fig. 4. Although we consider that the connection of the conformations in the order of U, I_1 , U, I_2 , and N in Fig. 4 is unnatural, further experiments are necessary to distinguish the possibilities.

Thus, although there remains a possibility that the dead-end model (Scheme 4) might describe the folding of SMN, the sequential folding scheme (Scheme 1) is the simplest and most consistent model to explain the current observations on the folding of SMN.