# Exposing the distinctive modular behavior of β-strands and α-helices in folded proteins

Huabing Wang[a], Derek T. Logan[b], Jens Danielsson[a], and Mikael Oliveberg[a,1]

[a]Department of Biochemistry and Biophysics, Arrhenius Laboratories of Natural Sciences, Stockholm University, S-106 91 Stockholm, Sweden; and [b]Department of Chemistry, Division of Biochemistry & Structural Biology, Lund University, 22100 Lund, Sweden

Although folded proteins are commonly depicted as simplistic combinations of β-strands and α-helices, the actual properties and functions of these secondary-structure elements in their native contexts are just partly understood. The principal reason is that the behavior of individual β- and α-elements is obscured by the global folding cooperativity. In this study, we have circumvented this problem by designing frustrated variants of the mixed α/β-protein S6, which allow the structural behavior of individual β-strands and α-helices to be targeted selectively by stopped-flow kinetics, X-ray crystallography, and solution-state NMR. Essentially, our approach is based on provoking intramolecular "domain swap." The results show that the α- and β-elements have quite different characteristics: The swaps of β-strands proceed via global unfolding, whereas the α-helices are free to swap locally in the native basin. Moreover, the α-helices tend to hybridize and to promote protein association by gliding over to neighboring molecules. This difference in structural behavior follows directly from hydrogen-bonding restrictions and suggests that the protein secondary structure defines not only tertiary geometry, but also maintains control in function and structural evolution. Finally, our alternative approach to protein folding and native-state dynamics presents a generally applicable strategy for in silico design of protein models that are computationally testable in the microsecond–millisecond regime.

structural cooperativity | secondary structure | protein dynamics | protein design

One of the most intriguing properties of naturally occurring proteins is their high folding cooperativity. Generally, this cooperativity promotes a relatively slow (minutes to hours) and highly concerted unfolding process where only the fully native and fully denatured species populate (1). Although this two-state behavior plays a vital role in safeguarding native-state integrity and biological function, it presents a major challenge for the experimentalist who wants to map out the global dynamics: Structural fluctuations that are sufficiently large to report on the secondary-structure interplay are so sparsely populated that they generally defy detection (2). This intrinsic difficulty of isolating high-energy events has not only impeded our understanding of the folding cooperativity itself, but also prevented detailed insight into the more extensive conformational changes involved in biological function and protein-aggregation disease (2). Experimental insight into native-state dynamics has so far been provided by relaxation-dispersion NMR (3), fluorescence resonance energy transfer labeling (4), and various hydrogen-exchange techniques, monitoring the fluctuations of individual hydrogen bonds in the protein backbone (5). While these strategies have been instrumental in establishing breathing modes at the low free-energy level (6–8) and global exchange (8, 9), they fail to capture the rare events occurring at the high free-energy level. Due to their local foci, they tend also to be ambiguous about the degree of structural concertedness, i.e., it is not always clear to what extent the observed local fluctuations are indeed cooperatively coupled or report on more diffusely distributed motions (5). A notable exception is the pioneering hydrogen–deuterium

(HD)-exchange studies by Englander and coworkers where the concerted local unfolding of individual α-helices in cytochrome $c$ was inferred by GdmCl titration (10). In essence, these experiments reveal that the helices of cytochrome $c$ maintain their own subglobal cooperativity, resulting in a stepwise loss of structure on the native side of the unfolding barrier (10). When it comes to the structural behavior of β-strands, however, the observations are more confounding. For example, native-like intermediates with locally unfolded β-strands are, in some cases, implicated as precursors for pathological aggregation (11–13), while the very elongation and growth of fibrillar structures proceed in other cases by the incorporation of globally disordered monomers (14–16). Moreover, the formation of individual β-hairpins in peptides is observed to occur on the microsecond timescale (17), while the exchange of individual β-strands in the domain swap is notably slow by following the global unfolding transition (18). The clue to understanding these differences in structural and dynamic behaviors seems, thus, to rest in how the properties of the secondary-structure elements change upon transfer from solution into the tertiary context of folded topologies. In this study, we target this issue by employing an unconventional engineering strategy based on intramolecular domain swap where the domain is an individual β-strand or α-helix (Fig. 1). The results show that the structural behaviors of the β-strands and α-helices in the ribosomal protein S6 are, indeed, very distinct: The β-strands cannot be dislodged by local fluctuations but require global unfolding to swap, whereas the α-helices are free to swap locally in the native basin, to hybridize in new tertiary
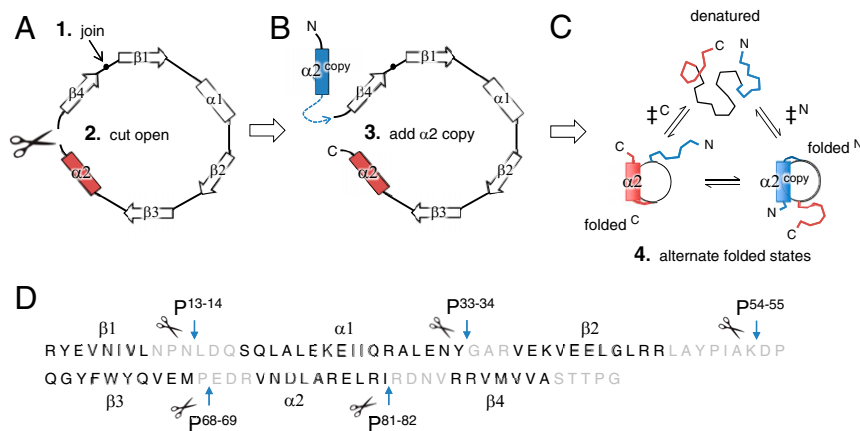
**Fig. 1.** Principle of construct design. (*A*) The protein is first circularized by joining the original N and C termini and then cut open next to the target secondary-structure element (here, α2). (*B*) An identical copy of the target element (α2$^{copy}$) is added to the N terminus, producing a primary sequence with α2 at either ends. (*C*) The construct is now free to fold by incorporating either α2 or α2$^{copy}$, leading to alternate folded states where the helical element is also allowed to swap locally. (*D*) Outline of the S6 pseudowild-type sequence showing permutant incisions and the secondary-structure elements analyzed in this study.

alignments, and to promote intermolecular association by "sliding" onto neighboring protein molecules. From these observations, we conclude that the advantage of having two types of secondary-structure elements lies not only in shape and complementarity, but also in evolvability through functional optimization of rigidity versus flexibility.

## Results

**Design of Swap Constructs.** The principal idea of this study is to access new levels of structural behavior by engineering proteins that can freely swap between given secondary-structure elements in their folded states. Our design of such swap constructs is outlined in Fig. 1, showing the example where α-helix 2 (α2) is chosen as a target element. First, wild-type S6 (PDB code 1ris) was circularized and cut open in a loop on the C-terminal side of α2. Next, a second copy of the target secondary-structure element was added to the new N terminus (α2$^{copy}$), yielding a construct with two identical secondary-structure elements at either end of the sequence. This construct can now fold by integrating either the original C-terminal secondary-structure element (folded$^C$) or the new N-terminal copy (folded$^N$). Of primary interest, the construct is also free to swap between the two identical secondary-structure elements in the native state (Fig. 1).

Regarding nomenclature, we denote the "double element" construct targeting α2 as swap$^{α2}$ with the alternate folded states swap$^{α2}$/folded$^C$ and swap$^{α2}$/folded$^N$. As independent references for the structures and folding behavior of the alternate swap$^{α2}$/folded$^C$ and swap$^{α2}$/folded$^N$, we used the "clean" circular permutants P$^{81-82}$ (folded$^C$) and P$^{68-69}$ (folded$^N$) with incisions between sequence positions 81 and 82, and 68 and 69, respectively (Fig. 2 and *SI Appendix*, Table S1). Analogous strategy and nomenclature are used for the constructs targeting the secondary-structure elements β1 and β2 of the S6 sheet (Fig. 3). The complete set of constructs and PDB codes to their X-ray structures are listed in *SI Appendix*, Table S1.

**β-Strand Swap Does Not Occur Locally but Proceeds via Global Unfolding.** To examine the properties of the S6 β-sheet, we opted for the constructs swap$^{β1}$ and swap$^{β2}$ (Fig. 3 *A* and *B* and *SI Appendix*, Table S1). The central β1 composes, here, the structural overlap between the two competing folding nuclei of the S6 folding process (1, 9, 19–21), whereas β2 is the dynamic (22) and structurally dispensable (23) strand at the sheet edge (Fig. 3B). Our constructs targeting the remaining strands β3 and β4 turned out to be unsuitable for swap analysis because the rate constants of the involved ground states are too poorly separated, cf. ref. 21 (*SI Appendix*, Fig. S2). To optimize the analysis of swap$^{β1}$, the
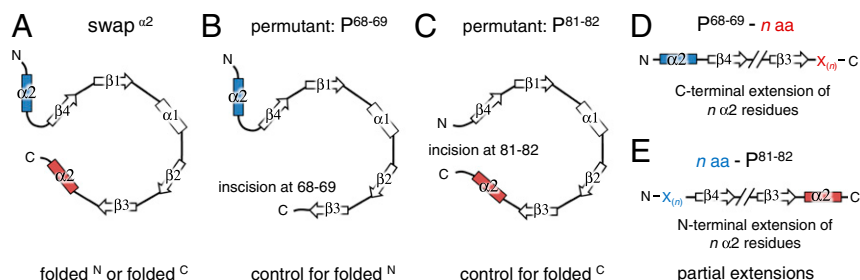


**Fig. 2.** Nomenclature as exemplified for the α2-constructs. The full set of constructs is listed in *SI Appendix*, Table S1. (*A*) Sequence outline of swap$^{α2}$ containing dual α2-copies. This construct is able to form the alternate states folded$^N$ and folded$^C$, according to Fig. 1C. (*B*) The circular permutant P$^{68-69}$ containing only the N-terminal α2-element. This variant is used as the control for the structure and folding pathway of folded$^N$. (*C*) Corresponding outline of the permutant P$^{81-82}$ forming the control for folded$^C$. (*D*) Construct with partial extension at the C terminus, in this case, *n* amino acids of α2. Besides from accounting for the effects of minor C-terminal modifications, this type of construct is used below for determining the effects of stepwise titration of the full α2-segment. (*E*) Corresponding outline of constructs with partial N-terminal extensions.

**Fig. 3.** Stopped-flow data showing the distinct global- and local-swap behaviors of β-strands and α-helices in the S6 structure. The sequences of the analyzed protein variants are outlined in *SI Appendix*, Table S1 and fitted parameters in *SI Appendix*, Tables S2 and S3. (*A, Left*) Constructs probing the swap behavior of β1. (*Middle*) The references folded$^C$ (6aa-P$^{13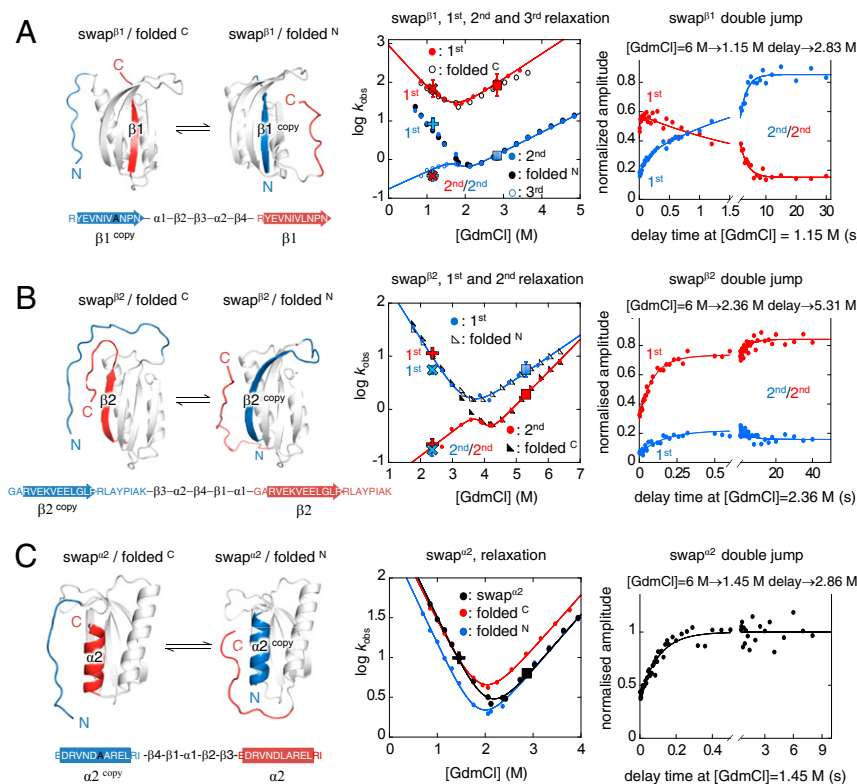-14}$) and folded$^N$ (P$^{97-3}$-5aa) show archetypical two-state behavior, manifested in v-shaped chevron plots with single folding relaxations. The swap-competent construct swap$^{β1}$ shows, instead, double folding relaxations that superimpose with both of the single chevrons of the references plus a slow phase that represents conversion between the alternate ground states swap$^{β1}$/folded$^C$ and swap$^{β1}$/folded$^N$ via the globally unfolded state. Large markers show the consistent relaxations obtained from the double-jump analysis on the *Right*. Fits are from the three-state relaxation in Scheme 1 (*SI Appendix*). (*B*) Corresponding data for swap$^{β2}$, folded$^C$ (8aa-P$^{54-55}$), and folded$^N$ (P$^{33-34}$-3aa). (*C*) The contrasting behavior of α2, which shows rapid swapping in the folded ground states according to the triangular mechanism in Fig. 1C. Constructs swap$^{α2}$, folded$^C$ (6aa-P$^{81-82}$), and folded$^N$ (P$^{68-69}$-6aa).

alternate ground states swap$^{β1}$/folded$^C$ and swap$^{β1}$/folded$^N$ were first poised to similar stabilities by the mutation L10A in the N-terminal β1 copy (*SI Appendix*, Tables S1–S3 and Figs. S1 and S2). This was to achieve balanced ground-state populations at equilibrium to maximize the resolution of the strand-swap events. Second, we equipped the references for folded$^C$ and folded$^N$ with short N/C-termini extensions to account for disordered-end effects (cf. Fig. 2 *D* and *E*). Generally, such short extensions are slightly destabilizing (2) and were, here, made up of five to six amino acids from either half of the β1 element (*SI Appendix*, Table S1). The resulting folding- and unfolding-kinetics of swap$^{β1}$ and the terminally modified references for folded$^C$ (6aa-P$^{13-14}$) and folded$^N$ (P$^{97-3}$-5aa) are shown in Fig. 3A. Most evidently, swap$^{β1}$ shows two relaxation phases that superimpose precisely with the chevron plots of either of the reference states (Fig. 3A). Additionally, the swap$^{β1}$ construct reveals a characteristic slow refolding phase in the timescale of global unfolding (Fig. 3A), analogous to that observed for strand-swap events in SOD1 (2). Taken together, these features suggest that the competing β1 elements in swap$^{β1}$ do not swap rapidly in the folded basin but need to await global unfolding according to (2)

$$\text{folded}^C \rightleftharpoons \text{denatured} \rightleftharpoons \text{folded}^N \quad \text{[Scheme 1]}$$
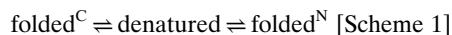
where the procedures for fitting are described in the *SI Appendix*. The swap$^{β1}$ relaxation phase that superimposes with the refolding

limb of the folded$^N$ reference (P$^{97-3}$-5aa) is linked to heterogeneity in the denatured ensemble due to generic proline isomerization (Fig. 3A). In essence, the fraction of denatured molecules with *cis* prolines in the linker to the C-terminal β1 segment is unable to fold rapidly via the folded$^C$ pathway and becomes naturally biased to the somewhat slower folded$^N$ alternative (*SI Appendix*, Fig. S3). Consistently, stopped-flow double-jump analysis yields refolding- and unfolding-rate constants that superimpose precisely with the chevron data of swap$^{β1}$ (Fig. 3A). Analysis of the edge strand β2 was undertaken in the same way, using the constructs swap$^{β2}$, folded$^C$ (8aa-P$^{54-55}$), and folded$^N$ (P$^{33-34}$-3aa) (*SI Appendix*, Table S1). The result is analogous to that of β1: The behavior of swap$^{β2}$ complies with the global swap mechanism according to Scheme 1 (Fig. 3B). A notable feature of the β2 swap event is that it involves the ~10-residue loop connecting β2 to β3 (*SI Appendix*, Table S1). Even if this loop is largely disordered in solution (9, 22), it folds over the sheet in the crystal lattice to form an external "minicore" at the monomer's packing interface (24). To see if the global-swap behavior of β2 is somehow coupled to such minicore interactions persisting in the solution state, we produced a truncated swap$^{β2}$ variant and reference states where the loop is replaced by a GAG linker (*SI Appendix*, Fig. S4). The control shows that the loop truncation has a negligible effect on the β2 swap behavior, which remains slaved to the global folding transition (*SI Appendix*, Fig. S4). It is thus evident that neither of the two S6 strands are free to detach and reinsert locally in the native basin

but are as slaved to global unfolding (2). Considering that β2 is partly dynamic in wild-type S6 (22) and can be cut away completely without impact on folding cooperativity (23), this a notable and somewhat surprising result.

**The Contrasting Behavior of α-Helices.** The question is now to what extent the S6 helices behave differently. Our focus is on α2 as the analysis of α-helix 1 is precluded by the reference constructs having unsuitable refolding kinetics (21) (*SI Appendix*, Fig. S2). Following the procedures of the strand-swap constructs, the competing states of swap$^{\alpha2}$ were first poised to matching stabilities by the mutation L75A in the N-terminal α2-copy of the swap construct (*SI Appendix*, Table S1 and Fig. S2). The results show that the kinetic behavior of this construct is markedly different from swap$^{\beta1}$ and swap$^{\beta2}$: The relaxation of swap$^{\alpha2}$ displays one phase only, and its chevron plot matches neither of the reference permutants (Fig. 3$C$). The single-phase relaxation is maintained in the double-jump controls, yielding the same single refolding- and unfolding-rate constants as the chevron analysis (Fig. 3$C$). This apparent two-state behavior of swap$^{\alpha2}$ suggests that the helix-swap events occur rapidly in the native state. In the simplest case, the linear Scheme 1 extends, then, to a triangular three-state relaxation with fast folded$^C$ ⇌ folded$^N$ interconversion in the native basin as outlined in Fig. 1$C$ and *SI Appendix*, Scheme S2. We also note that the data set as a whole resembles an ideal φ-value analysis where swap$^{\alpha2}$ is the wild type, and folded$^C$ (6aa-P$^{81-82}$) and folded$^N$ (P$^{68-69}$-6aa) are the mutant variants with φ = 0 and φ = 1, respectively (25). Before embarking on more detailed kinetic analysis, however, we characterized the folded ground states of the various α2-constructs by X-ray crystallography and solution-state NMR as described below.

**X-ray Structures Reveal Helix-Packing Plasticity.** Further clues to the swap$^{\alpha2}$ behavior is provided from X-ray data (Fig. 4 and *SI Appendix*, Figs. S5–S7). Starting with the folded$^N$ reference (L75), this construct shows an overall wild-type-like structure (*SI Appendix*, Fig. S5$D$). The only notable difference is that the strain release following the incision between residues 68 and 69 causes α2 to relax and to elongate by half a turn at the N-terminal end (*SI Appendix*, Fig. S5$D$). A similar structural arrangement is found for swap$^{\alpha2}$ (L75) where the presence of a complete α2-segment at the C terminus is not enough to overcome the high stability of the folded$^N$ state (*SI Appendix*, Fig. S6$B$). Upon destabilization of the folded$^N$ reference by the mutation L75A, however, α2 moves closer to the core, indicating that its docking is to some extent frustrated and susceptible to readjustments (*SI Appendix*, Fig. S5$B$). This frustration manifests itself fully in the likewise mutated swap$^{\alpha2}$ where both copies of α2 are structured but adjusted into nonnative positions: The C-terminal terminal helix slides out to bind the neighboring molecule in the crystal lattice, allowing the N-terminal copy to seal the volume left behind (Fig. 4$C$). Interestingly, this hybrid arrangement of the swap$^{\alpha2}$ helices leads to an expansion of the core around the original α2-position where the new N- and C-terminal copies occupy similar volumes (Fig. 4$A$). Elongating instead the sequence from the N-terminal end, the clean folded$^C$ reference (P$^{81-82}$) fails to produce useful diffraction data. Even so, an overall wild-type arrangement of the P$^{81-82}$ structure is verified by solution-state NMR (*SI Appendix*, Fig. S11). The ability to form ordered crystals is restored by a nine-residue extension of the N terminus in the reference construct 9aa-P$^{81-82}$, yielding a hybridized helical arrangement similar to that of swap$^{\alpha2}$ (Fig. 4$A$). Although these lattice-stabilized structures of 9aa-P$^{81-82}$ and swap$^{\alpha2}$ are not necessarily transferrable to solution conditions, they underpin the notion that the tertiary arrangement of α-helices — because of their self-contained cooperativity (26, 27) — is more flexible than that of globally fixed β-strands (28). Of particular
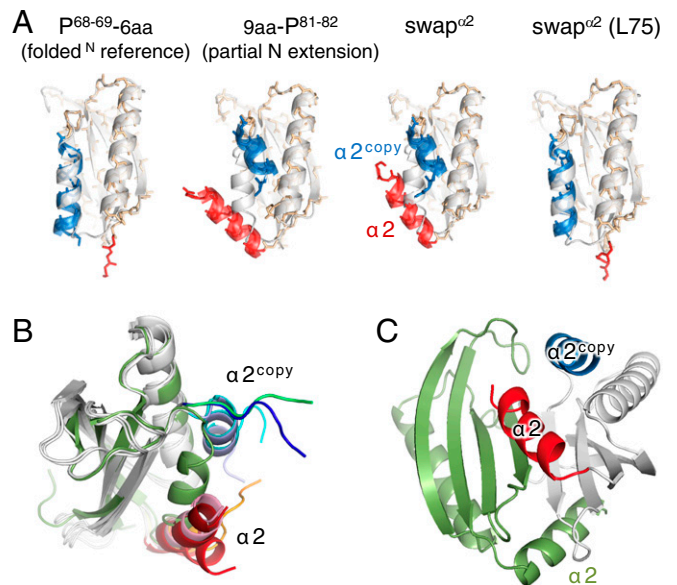


**Fig. 4.** X-ray structures showing helix hybridization and dimer-interface sliding. The C-terminal α2-element is in red, and the N-terminal α2$^{copy}$ is in blue. (*A*) The folded$^N$-reference (P$^{68-69}$-6aa) shows a wild-type (gray)-like structure with the 6aa α2-extension protruding disordered in the crystal lattice. The folded$^C$-titration construct 9aa-P$^{81-82}$ starts to undergo hybridization at an α2$^{copy}$-addition of 9aa (cf. *SI Appendix*, Fig. S5$F$). Similar hybridization is found for the full swap$^{\alpha2}$ construct. Reintroduction of the stabilizing L75 in the α2$^{copy}$ of swap$^{\alpha2}$, leads to complete expulsion of α2. (*B*) Overlay of representative α2-swap structures on S6 wild type (green). (*C*) Crystallographic protein—protein interface showing how α2 slides over to merge with the neighboring molecule.

interest is that the C-terminal α2 in several cases slides over to the neighboring S6 molecule and thereby contributes to extend the dimer-interface area (Fig. 4$C$). The self-contained behavior of α2 is, thus, not only revealed by intramolecular packing flexibility, but also by "induced" intermolecular fit. Similar sliding of helices in the crystal packing is consistently observed in four-helix bundles (29).

**NMR Shows Local Helix Exchange.** To determine to what extent helix hybridization occurs in the free swap$^{\alpha2}$ monomers, we employed solution-state NMR. The immediate observation is that the heteronuclear single quantum coherence (HSQC) spectra of swap$^{\alpha2}$, the folded$^N$, and folded$^C$ references (P$^{68-69}$-6aa and 6aa-P$^{81-82}$) show exchange broadening at protein concentrations >100 μM, which is not present in wild-type S6 (9, 22). Examination by elution chromatography and stopped-flow analysis shows further that these spectral components coincide with reversible (low-level) oligomerization of the folded species (*SI Appendix*, Fig. S9). It is reasonable to suspect that the oligomerization relates to the intermolecular interactions in the crystal packing, i.e., induced-interface fit by α2-sliding (Fig. 4$C$). To test this possibility, we perturbed the crystallographic dimer interface by the surface mutation A92K, which does not significantly affect the system's helix-swap behavior (*SI Appendix*, Fig. S10). The mutation diminishes the solution oligomerization and renders the HSQC spectra characteristic of monomeric S6 species (*SI Appendix*, Fig. S11). As an additional control, we verified the monomeric nature of the various A92K-containing swap$^{\alpha2}$ variants under NMR conditions with NMR relaxation analysis (*SI Appendix*). The new folded$^N$ and folded$^C$ references, i.e., P$^{68-69}$-6aa (A92K) and 6aa-P$^{81-82}$ (A92K), produce well-dispersed HSQC spectra typical for uniformly folded permutant structures (*SI Appendix*, Fig. S8$A$). These ideal spectral features are

overall maintained upon complete helix duplication in swap$^{\alpha 2}$ (A92K) but with the interesting difference that some cross peaks shift to further new positions (Fig. 5$A$). The obvious first step is, here, to assume that the swap$^{\alpha 2}$ (A92K) spectrum is the combination of folded$^N$ and folded$^C$ in rapid exchange (30), i.e., the helices swap dynamically with one another, generating a single cross peak that is the weighted average of the two competing folded states. However, this reductionist interpretation is challenged by several swap$^{\alpha 2}$ (K92) cross peaks being offset from the expected folded$^N$ and folded$^C$ average (Fig. 5$A$). The spectral divergence indicates that the helix duplication not just leads to competition for wild-type-like docking with the nonrecruited element protruding disordered into solvent but to a new helical arrangement of similar stability (Fig. 5$A$). Judged by the single set of swap$^{\alpha 2}$ (K92) cross peaks, this new helical arrangement seems further to be in rapid exchange with folded$^N$ and folded$^C$, leaving a balanced three-state equilibrium on the native side of the folding barrier(s) (Fig. 5$B$). Henceforth, we refer tentatively to the new swap$^{\alpha 2}$ (K92) species as the folded$^{hybrid}$ state since the present NMR data sheds no information of its actual helical arrangement: It can be either a true hybrid as observed in Fig. 4$A$ or a variant where one helix adapts fully native-like packing and the other perturbs the chemical shifts by dynamic sampling of the folded surface (*SI Appendix*).

**Clues from α2-Structural Titration: the Competing Ground States.** To shed mechanistic detail on the helix-swap events, we "titrated in" the competing α2-segments — one residue at the time — to gradually end up with the full swap$^{\alpha 2}$ protein (Fig. 6 and *SI Appendix*, Table S1). The effects on folding, stability, and ground-state structures were monitored both by chevron analysis (cf. Fig. 3) and by NMR (*SI Appendix*, Fig. S8 and S11). Starting with the clean reference folded$^N$ (P$^{68-69}$), appreciable effects on protein stability do not materialize before position V72 (i.e., construct P$^{68-69}$-4aa with a C-terminal addition of four residues) (Fig. 6$B$ and *SI Appendix*, Table S1). This slight destabilization is tentatively assigned to steric interference between the added residues and the folded helical element, possibly in combination with altered electrostatic interactions, including the C-terminal charge. The destabilization then maintains relatively constant between V72 and E78 (P$^{68-69}$-10aa) after which the addition of the final two to three residues promotes a sharp drop

to full swap$^{\alpha 2}$ stability (Fig. 6$B$ and *SI Appendix*, Table S1). The behavior suggests that, when titrated from the C-terminal end of the S6 sequence, a nearly complete copy of the α2-segment is required for the helix-swap events to commence. Starting conversely from the folded$^C$ reference (P$^{81-82}$), the extensions yield overall matching results. Mirroring the folded$^N$ titration, the destabilization cuts in relatively sharply between residues 81–79 while the recovery of stability upon adding residues 75–69 is more gradual (Fig. 6$A$). Regardless of how the titration is started, the decisive features for triggering the swap motions seem, thus, contained in the L75-I81 region. We also note that this region is likely to encompass the highest degree of helix–helix repulsion due the unmatched positive charges at positions R77 and R80 where the introduction of an addition positive terminal charge in the titration from the N-terminal end is expected to exacerbate this repulsion further. Corroborating the overall titration trend, the chemical-shift changes accompanying the helix titrations match the protein-stability profiles: The titration yields overall steeper response between positions 79 and 81 than between positions 69 and 75 (Fig. 6$A$ and *SI Appendix*, Fig. S12). Stability and NMR data thus yield consistent results in the form of a population-averaged transition of the initial and final species in rapid exchange.

**Clues from Structural Titration: the Competing Transition States.** An interesting outcome of the modular helical behavior in Figs. 5 and 6 is that it sheds lights on the ϕ-value character of the swap$^{\alpha 2}$ chevron plots (Fig. 3). To pin down the details behind this orderly response, we followed the effects of structural titration on the individual unfolding ($k_u$) and refolding ($k_f$) -rate constants (Fig. 6$C$). Starting from the C-terminal end (folded$^N$), $k_u$ remains relatively unaffected by the helix elongation, while $k_f$ undergoes a pronounced acceleration upon addition of the final L79-I81 residues. As such, the titration confirms that the C-terminal α2-element catalyzes global folding by providing full native stabilization already in the folding transition state, i.e., ϕ = 1. Starting instead from the N-terminal end (folded$^C$), the α2-titration shows only marginal impact on $k_f$ but a pronounced increase in $k_u$ (Fig. 6$B$). This selective $k_u$ effect suggests that the N-terminal α2-copy does not take part in the folding transition state of swap$^{\alpha 2}$, but structures first in the native basin, i.e., ϕ = 0. There are two implications of these results for the folding
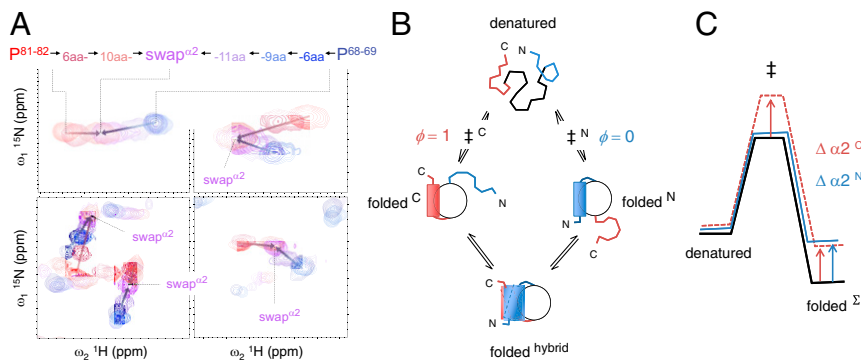


**Fig. 5.** Solution-state NMR analysis indicating α2-hybridization of free monomers. All constructs contain the dimer-splitting mutation A92K. (*A*) Chemical-shift changes upon structural titration from the clean circular permutants P$^{81-82}$ (folded$^C$) and P$^{68-69}$ (folded$^N$) to swap$^{\alpha 2}$. Step-by-step additions of native amino acids on the N-terminal end of P$^{81-82}$ result in a linear response in chemical shift toward the chemical shift of swap$^{\alpha 2}$. A corresponding response is seen upon addition of native amino acids on the P$^{68-69}$ C terminus. Even so, the kinks in the combined chemical-shift contours indicate that the swap$^{\alpha 2}$ state is not simply a linear combination of folded$^C$ and folded$^N$ but rather constitutes a third state. (*B*) Scheme showing cooperative folding of the alternate states folded$^C$ and folded$^N$, and rapid (>200 s$^{-1}$) α2 swapping in the native basin, involving the hybridized state folded$^{hybrid}$. The competing pathways for escaping and reentering the native basin are controlled by two transition states where the ‡$^C$ engages α2 (ϕ = 1) and ‡$^N$ lacks α2-influence (ϕ = 0). (*C*) Corresponding two-state projection of the free-energy profile, explaining the effects of *N*- and C-terminal α2-truncation on the swap$^{\alpha 2}$ background (cf. Fig. 3$C$) where folded$^\Sigma$ is the sum of folded$^C$, folded$^{hybrid}$, and folded$^N$.
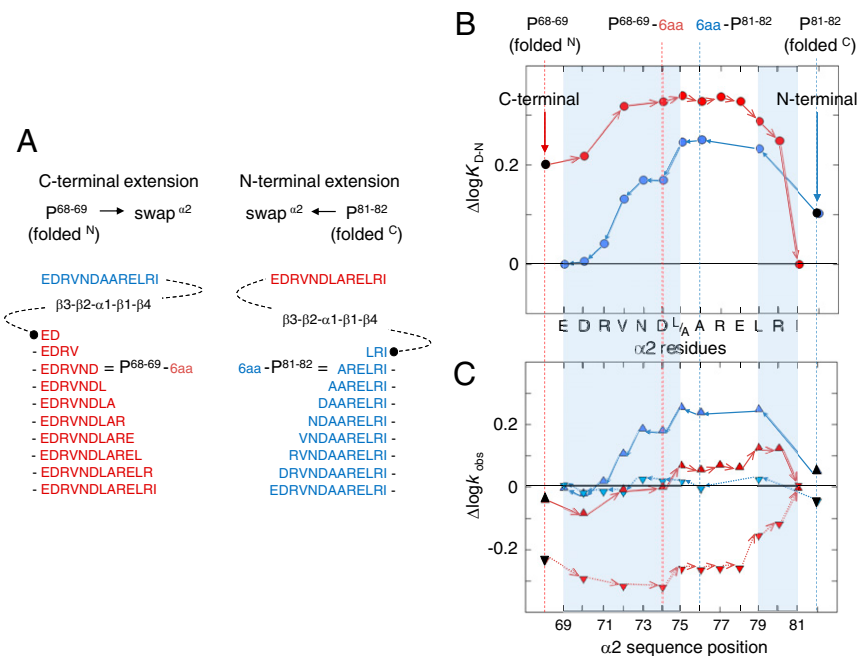
**Fig. 6.** Stepwise addition of an α2-copy. (A) Progressive extension of α2-residues to the C terminus of folded$^N$ (P$^{68-69}$) to finally end up with the full swap$^{α2}$ and the analogous procedure starting with folded$^C$ (P$^{81-82}$) (*SI Appendix*, Table S1). (B) Effects on protein stability relative the full swap$^{α2}$. The initial destabilizations show the penalty of adding extra C- or N-terminal residues to folded S6, and the final convergence to $\Delta \log K_{D-N} = 0$ shows the formation of the full swap$^{α2}$ species as detected by NMR in Fig. 5A. Notably, the addition of residues 69–75 yields a more gradual stability response than addition of residues 79–81, regardless of titration direction. (C) Corresponding changes of the unfolding- (▲) and refolding- (▼) rate constants. The kinetic profiles suggest that the C-terminal α2-element catalyzes global folding by providing full native stabilization already in the folding transition state (φ = 1), whereas the N-terminal α2-copy forms first in the native basin (φ = 0).

mechanism. First, the absence of the N-terminal α2-copy in the barrier crossing suggests that the preferred transition state of swap$^{α2}$ is similar to that of the clean folded$^C$ construct (P$^{81-82}$). We denote this dominant transition state ‡$^{φ=1}$ = ‡$^C$ (Fig. 7B). Likewise, there exists a parallel but less favorable transition state
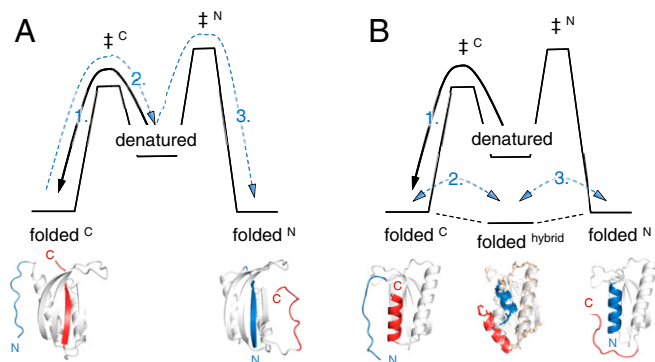


**Fig. 7.** Folding free-energy profiles illustrate the different swap behaviors of β-strands and α-helices, respectively. The arrows follow a typical refolding experiment as carried out in this study. (A) Proteins that first fold from denatured (D) to folded$^C$ with the C-terminal β-strand cannot swap to the N-terminal β-strand unless they return to D and refold to folded$^N$. The mechanistic reason for this global-swap behavior has been assigned to the need to establish β-sheet register early in the folding process, i.e., in the rate-limiting step D to ‡ (2). (B) Swapping of α-helices can occur much more flexibly in the native basin, and this also increases the susceptibility to hybridization. The mechanistic reason for this local-swap behavior is assigned to the self-contained cooperativity of the α-helical structure itself. Accordingly, the β-strands and α-helices seem to provide opposing properties in the optimization of structural rigidity versus conformational freedom in protein evolution.

where none of the α2-elements contribute, and this seems similar to that of the clean folded$^N$ construct (P$^{68-69}$), i.e., ‡$^{φ=0}$ = ‡$^N$ (Fig. 7B). The two pathways (‡$^{φ=1}$ and ‡$^{φ=0}$) are not just accessible to swap$^{α2}$ but also to the folded$^C$ reference. This minimal extension of the triangular scheme in Fig. 1C explains why the refolding rate constant of swap$^{α2}$ ($k_f^{swap\,α2}$) is identical to $k_f^{folded\,C}$ and not ($k_f^{folded\,C} + k_f^{folded\,N}$) (*SI Appendix*, Fig. S15). Evidence for such dual-pathway behavior is also found in the V72-L79 region of the folded$^N$ titration where $k_u$ and $k_f$ increase in concert as the faster ‡$^{φ=1}$ pathway gradually kicks in (Fig. 6C). Second, the conclusion that swap$^{α2}$ and folded$^C$ employ similar transition state (‡$^C$) turns out to have implications for the folded swap$^{α2}$ structure. The clue is in the response to α2-perturbation. Mutation of the folded$^C$ structure (P$^{81-82}$) has previously been shown to yield overall low φ-values in the α2-region, i.e., φ < 0.2, indicating that the native interactions of the helix are largely missing in the folding transition state (20, 21). In stark contrast, perturbations of the C-terminal α2 in swap$^{α2}$ yield, here, φ = 1 (Fig. 3C). Given that ‡$^C$ is still the same, these differences must stem from the helix mixing in the swap$^{α2}$ ground state (Figs. 5 and 6). One possibility is thus that the interactions of the C-terminal α2 in the swap$^{α2}$ structure (i.e., folded$^{hybrid}$) are weakened to level where they match those of the folded$^C$ transition state. Whether this correspondence is just by chance, system specific, or report on the inherent features of structural self-organization remains to be tested.

## Discussion

**The Different Behaviors of β-Strands and α-Helices in Folded Proteins.** The deceptively simple organization of protein structures into bundled α-helices and β-strands is often taken to infer how the protein folds, i.e., the primary chain first forms the secondary-structure elements, which, then, assemble into the native tertiary

topology. However, this view is not entirely true. Rather, the secondary structure and tertiary contacts are biased to form concomitantly since the principal role of the intramolecular hydrogen bonds is not in providing stability but in promoting specificity by compensating for lost water interactions upon hydrophobic collapse (31–33). For as yet poorly understood reasons, this conflict between satisfying hydrogen-bonding constraints and tertiary close packing is typically resolved at high-energy level, giving rise to substantial folding free-energy barriers and high folding cooperativity. Following microscopic reversibility, the native-state dynamics leading to global unfolding is the reverse, i.e., rupture of the tertiary integrity coupled to solvation of the backbone hydrogen bonds. Although this concerted two-state nature of structural transitions reduces complexity in the crowded interior of live cells (34, 35), it leaves us with difficulties in deconvoluting exactly what is going on (2). Some pressing examples are how secondary-structure properties direct structural evolution (36), the coupling between structural motions and protein function (37, 38), and to what extent proteins can locally unfold to trigger pathologic aggregation and misfolding disease (2, 39). We target, here, the problem by exposing in an alternative way the local and global behaviors of individual secondary-structure elements in the split β−α−β protein S6 (Fig. 1). Our approach is based on engineering sequence variants where the examined secondary-structure element is sterically allowed to swap place with an identical copy in the native state (Figs. 3 and 4). Taken together, the results from these secondary-structure duplications show that the β-strands and α-helices behave very differently. The β-strands are so firmly locked into their tertiary context that they cannot swap unless the protein globally ruptures and refolds, whereas the α-helices are free to swap dynamically on the native side of the unfolding barrier (Fig. 7).

**Comparison with Other Systems.** Although the results in Fig. 7 essentially confirm what is predicted from long-range (β-sheets) and local (α-helices) hydrogen-bond patterns alone, the very experimental verification of distinct secondary-structure behavior in the context of a folded structure sets the basis for a more detailed examination. First, we note that the partitioning into globally slaved and locally free motions of the β- and α-segments is fully discrete and without intermediate behavior.

Somewhat surprisingly, this partitioning includes also the edge strand β2, which is clearly dynamic as measured by HD-exchange (9) and NMR-relaxation data (22). The result indicates that the breathing and solvent accessibility of individual hydrogen bonds does not necessarily report on the concerted motions of the entire secondary-structure segment (2, 40). Mechanistically, one possibility is that the local breathing occurs just diffusely along the main chain without affecting the topology as a whole. Alternatively, the swap behavior is controlled by just a few globally fixed hydrogen bonds along the otherwise dynamic strand that, in this case, would be confined to the N-terminal end (9). In the case of helices, however, the agreement between our current data and the HD-exchange patterns appears more clear-cut: The individual helical elements of cytochrome c are found to undergo concerted unfolding in the native basin (10), consistent with the schematic outline in Fig. 7B. This difference in local dynamic behavior stems most likely from the intrinsic cooperativity of helical segments, in general (26, 27). Independent support for the global nature of β-strand transitions is provided by domain swap where two or more proteins "entangle" by exchanging secondary-structure elements with one another (41, 42). The domain swap follows mechanistically from the funneled landscapes of globular proteins (43), and is employed naturally in functional oligomerization (44), i.e., the domain-swapped oligomers have higher activity (45–47). Domain swap also tends to occur spontaneously in refolding of monomeric proteins if the protein concentration is high enough (48) or loop restraints are

relieved (49, 50). When the domain-swapped elements include β-strands, the dissociation into folded monomers is observed to require global unfolding (51–53). An illustrative example is provided by cyanovirin-N where the barrier for domain swap matches that for global unfolding of the free monomers (51). When the swapped element is instead a single α-helix, the domain swapping is reported to occur more rapidly by local unfolding in the native basin (54). Additional evidence for dynamic partitioning of β- and α-elements is provided from studies of a protein that can reversibly switch between two folded structures, i.e., protein $G_A$ (all-α) and protein $G_B$ (α/β) (55, 56). Following the constraints of the β-strands in Fig. 7A, the transition between the $G_A$ and the $G_B$ topologies does not take place locally but proceeds via global unfolding (55, 56). To this end, the conspicuous insert of a single β-strand into a partners sheet, which governs the functional polymerization of serpins, is observed to rely on concerted changes in secondary and tertiary structures (57) and, similarly, the swapping of individual β-strands in the amyotrophic lateral sclerosis-associated protein SOD1 requires global unfolding to occur (2). The advantage of this mechanism has, in both cases, been linked to the avoidance of sticky intermediates, which could conceivably populate if the swaps were based on hybridizing α-helices (Fig. 7). Taken together, this overall consistent separation of β-strand and α-helix behaviors suggests that our present observations are not unique for S6 but reflect some degree of generality.

**Notes on Structural Evolution.** Evolution of protein function is not just a matter of achieving suitable topology, but also more elusively about tuning the structural properties. The latter involves maintaining sufficient structural stability (58), balancing affinity and interaction promiscuity (59), tuning the quinary interplay, proteome self-organization, intracellular search (60, 61), and optimizing the dynamic motions (62–64). Our present focus is on the evolutionary role of the motions and properties of entire β- and α-elements, which has, so far, gained relatively sparse attention (2, 65, 66). Nevertheless, these secondary-structure elements constitute the basic units of insertions and deletions (indels) that, together with point mutations, are the most common events in protein evolution (36). Our simplified arguments are as follows. Large-scale analysis of existing protein structures shows that α-helical elements are overall more robust to mutations than β-strands, i.e., helices tolerate more sequence change without changing secondary structure (65). In view of the current data, this implies that α-helical elements more easily maintain their hydrogen-bonding integrity because they are generally freer to readjust in the folded structure upon shape perturbations. Such helix adjustments are, indeed, apparent upon superposition of S6 homologs and seem further to correlate with sequence divergence (*SI Appendix*, Figs S13 and S14). It is, thus, not surprising that proteins that encapsulate a bulky heme group are better suited to a malleable envelope of α-helices than tertiary restricted β-sheets. As exemplified by hemoglobin, helical topologies are also readily amenable to long-range allosteric crosstalk and conformational changes through coupled mutual helical motions (67, 68). Similar cofactor-dependent conformational heterogeneity is observed for the helical protein GAD65 (69). The extreme case of unified helical motions is possibly found in transmembrane proteins where the lipid confinement not only allows larger conformational changes than in an aqueous solvent, but also stabilizes the individual helical elements (70). By comparison, conformational changes within rigid β-sheets are intrinsically more problematic to achieve. A revealing example is the allosteric communication between the metal-binding sites in the homodimeric β-barrels of SOD1: To propagate the signal, this protein strains the sheet by twisting in a single β−β hydrogen bond into the core, introducing a "bridge" of dynamic frustration across the two subunits (6). Accordingly,

proteins based on mixed α- and β-motifs can be said to combine the best of two worlds and represent the most common motifs in proteomes across all organisms (71). Besides increasing the repertoire of steric shapes, rigid β-sheets covered on one or both sides by adjustable α-helical elements are overall more tolerant to mutations than corresponding all-β arrangements (65, 72). From the observations in this study, it is indicated that mixed α- and β-topologies also have the evolutionary advantage of mixed dynamical properties. First, the seamless merge of sheets enabled by the sliding motion of α2 in Fig. 4*C* not only produces a new snug interface, but also seems to create a new "foldon" in the form of one-helix-docking-against-two strands (9, 21). In essence, foldons are proposed to make up the minimal cooperative units (folding nuclei) and to propagate cooperativity within domains by being combined with structural overlap (1). As such, the induced S6 interface in Fig. 4*C* can be seen as a cooperative binding motif in its own right. Although the affinity provided by this induced "interface foldon" is just in the millimolar regime as judged by solution NMR (*SI Appendix*), it is easy to imagine how it can be evolutionarily fixed by additional point mutations or altered sequence connectivity. Second, the ability of identical helical elements to hybridize dynamically upon sequence duplication (Figs. 4 and 5), taken together with the finding that S6 still folds cooperatively upon complete truncation of the β2-strand (23), provides experimental evidence that structural evolution by insertion and deletion works in practice, and is even more flexible than envisioned from static structures alone. In particular, it is interesting that the stability change accompanying the insertion and hybridization of α2 is small and unlikely to compromise

in vivo foldability. From this structure-centric analysis, we conclude that the balance between global rigidity and controlled local motions is not only key to enzymatic efficiency and functional conformational changes (63, 73), but also can play a role in facilitation of evolution of the structures per se. Notably, the experimental observations behind this conjecture are easily put to test: The β-strands are rigid and unyielding because their hydrogen bonds are cooperatively slaved to the global unfolding transition, whereas the α-helical elements are free to dynamically swap and hybridize in the native basin.

## Materials and Methods

Molecular biology was performed as described in ref. 2, and proteins were produced as in ref. 21. Stopped-flow analysis, NMR analysis, and X-ray crystallography were performed as in ref. 2. All measurements were performed at 25 °C in 50 mM 2-(*N*-morpholino)ethanesulfonic acid (Sigma-Aldrich) buffer at pH6.3 unless otherwise stated. Detailed materials and methods are in the *SI Appendix*.

**Data Availability.** X-ray structures are deposited in the PDB databank (PDB codes 6I69, 6I6I, 6I6E, 6I6O, 6I6S, 6I6U, 6I6W, and 6I6Y). Raw data from kinetic, thermodynamic, and NMR analysis are available by contacting the corresponding author.

1. M. O. Lindberg, M. Oliveberg, Malleability of protein folding pathways: A simple reason for complex behaviour. *Curr. Opin. Struct. Biol.* **17**, 21–29 (2007).
2. H. Wang, L. Lang, D. T. Logan, J. Danielsson, M. Oliveberg, Tricking a protein to swap strands. *J. Am. Chem. Soc.* **138**, 15571–15579 (2016).
3. A. J. Baldwin, L. E. Kay, NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5**, 808–814 (2009).
4. E. Haas, The study of protein folding and dynamics by determination of intramolecular distance distributions and their fluctuations using ensemble and single-molecule FRET measurements. *ChemPhysChem* **6**, 858–870 (2005).
5. S. W. Englander, L. Mayne, Z. Y. Kan, W. Hu, Protein folding-how and why: By hydrogen exchange, fragment separation, and mass spectrometry. *Annu. Rev. Biophys.* **45**, 135–152 (2016).
6. J. Danielsson et al., Global structural motions from the strain of a single hydrogen bond. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3829–3834 (2013).
7. A. L. Serrano, O. Bilsel, F. Gai, Native state conformational heterogeneity of HP35 revealed by time-resolved FRET. *J. Phys. Chem. B* **116**, 10631–10638 (2012).
8. S. Bhatia, G. Krishnamoorthy, J. B. Udgaonkar, Site-specific time-resolved FRET reveals local variations in the unfolding mechanism in an apparently two-state protein unfolding transition. *Phys. Chem. Chem. Phys.* **20**, 3216–3232 (2018).
9. E. Haglund et al., The HD-exchange motions of ribosomal protein S6 are insensitive to reversal of the protein-folding pathway. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21619–21624 (2009).
10. Y. Bai, T. R. Sosnick, L. Mayne, S. W. Englander, Protein folding intermediates: Native-state hydrogen exchange. *Science* **269**, 192–197 (1995).
11. F. Chiti, C. M. Dobson, Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5**, 15–22 (2009).
12. M. G. Iadanza, M. P. Jackson, E. W. Hewitt, N. A. Ranson, S. E. Radford, A new era for understanding amyloid structures and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 755–773 (2018).
13. Y. S. Eisele et al., Targeting protein aggregation for the treatment of degenerative diseases. *Nat. Rev. Drug Discov.* **14**, 759–780 (2015).
14. L. Lang et al., SOD1 aggregation in ALS mice shows simplistic test tube behavior. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9878–9883 (2015).
15. M. Bacci, J. Vymětal, M. Mihajlovic, A. Caflisch, A. Vitalis, Amyloid β fibril elongation by monomers involves disorder at the tip. *J. Chem. Theory Comput.* **13**, 5117–5130 (2017).
16. R. Gaspar et al., Secondary nucleation of monomers on fibril surface dominates α-synuclein aggregation and provides autocatalytic amyloid amplification. *Q. Rev. Biophys.* **50**, e6 (2017).
17. V. Muñoz, P. A. Thompson, J. Hofrichter, W. A. Eaton, Folding dynamics and mechanism of beta-hairpin formation. *Nature* **390**, 196–199 (1997).
18. F. Rousseau, J. W. Schymkowitz, L. S. Itzhaki, The unfolding story of three-dimensional domain swapping. *Structure* **11**, 243–251 (2003).
19. M. Lindberg, J. Tångrot, M. Oliveberg, Complete change of the protein folding transition state upon circular permutation. *Nat. Struct. Biol.* **9**, 818–822 (2002).
20. M. O. Lindberg, E. Haglund, I. A. Hubner, E. I. Shakhnovich, M. Oliveberg, Identification of the minimal protein-folding nucleus through loop-entropy perturbations. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4083–4088 (2006).
21. E. Haglund, M. O. Lindberg, M. Oliveberg, Changes of protein folding pathways by circular permutation. Overlapping nuclei promote global cooperativity. *J. Biol. Chem.* **283**, 27904–27915 (2008).
22. A. Ohman, T. Oman, M. Oliveberg, Solution structures and backbone dynamics of the ribosomal protein S6 and its permutant P(54-55). *Protein Sci.* **19**, 183–189 (2010).
23. E. Haglund et al., Trimming down a protein structure to its bare foldons: Spatial organization of the cooperative unit. *J. Biol. Chem.* **287**, 2731–2738 (2012).
24. M. Lindahl et al., Crystal structure of the ribosomal protein S6 from Thermus thermophilus. *EMBO J.* **13**, 1249–1254 (1994).
25. A. Matouschek, J. T. Kellis Jr., L. Serrano, A. R. Fersht, Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122–126 (1989).
26. T. Bereau, M. Deserno, M. Bachmann, Structural basis of folding cooperativity in model proteins: Insights from a microcanonical perspective. *Biophys. J.* **100**, 2764–2772 (2011).
27. K. Ghosh, K. A. Dill, Theory for protein folding cooperativity: Helix bundles. *J. Am. Chem. Soc.* **131**, 2306–2312 (2009).
28. M. Manning, W. Colón, Structural basis of protein kinetic stability: Resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward beta-sheet structure. *Biochemistry* **43**, 11248–11254 (2004).
29. S. Geremia et al., Response of a designed metalloprotein to changes in metal ion coordination, exogenous ligands, and active site volume determined by X-ray crystallography. *J. Am. Chem. Soc.* **127**, 17266–17276 (2005).
30. A. G. Palmer 3rd, C. D. Kroenke, J. P. Loria, Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **339**, 204–238 (2001).
31. A. R. Fersht, Optimization of rates of protein folding: The nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10869–10873 (1995).
32. M. Oliveberg, P. G. Wolynes, The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* **38**, 245–288 (2005).
33. P. Kukic et al., Structural characterization of the early events in the nucleation-condensation mechanism in a protein folding process. *J. Am. Chem. Soc.* **139**, 6899–6910 (2017).
34. J. Danielsson et al., Thermodynamics of protein destabilization in live cells. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12402–12407 (2015).
35. X. Mu et al., Physicochemical code for quinary protein interactions in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4556–E4563 (2017).
36. N. V. Grishin, Fold change in evolution of protein structures. *J. Struct. Biol.* **134**, 167–185 (2001).
37. Z. Kurkcuoglu, A. Bakan, D. Kocaman, I. Bahar, P. Doruker, Coupling between catalytic loop motions and enzyme global dynamics. *PLoS Comput. Biol.* **8**, e1002705 (2012).
38. L. W. Yang, I. Bahar, Coupling between catalytic site and collective dynamics: A requirement for mechanochemical activity of enzymes. *Structure* **13**, 893–904 (2005).

39. A. Bille, S. A. E. Jónsson, M. Akke, A. Irbäck, Local unfolding and aggregation mechanisms of SOD1: A Monte Carlo exploration. *J. Phys. Chem. B* **117**, 9194–9202 (2013).

40. R. G. McAllister, L. Konermann, Challenges in the interpretation of protein h/d exchange data: A molecular dynamics simulation perspective. *Biochemistry* **54**, 2683–2692 (2015).

41. M. J. Bennett, S. Choe, D. Eisenberg, Domain swapping: Entangling alliances between proteins. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3127–3131 (1994).

42. M. J. Bennett, M. P. Schlunegger, D. Eisenberg, 3D domain swapping: A mechanism for oligomer assembly. *Protein Sci.* **4**, 2455–2468 (1995).

43. S. Yang *et al.*, Domain swapping is a consequence of minimal frustration. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13786–13791 (2004).

44. F. Rousseau, J. W. Schymkowitz, H. R. Wilkinson, L. S. Itzhaki, Intermediates control domain swapping during folding of p13suc1. *J. Biol. Chem.* **279**, 8368–8377 (2004).

45. M. Libonati, M. Bertoldi, S. Sorrentino, The activity on double-stranded RNA of aggregates of ribonuclease A higher than dimers increases as a function of the size of the aggregates. *Biochem. J.* **318**, 287–290 (1996).

46. M. Libonati, A. Floridi, Breakdown of double-stranded RNA by bull semen ribonuclease. *Eur. J. Biochem.* **8**, 81–87 (1969).

47. Y. Liu, G. Gotte, M. Libonati, D. Eisenberg, Structures of the two 3D domain-swapped RNase A trimers. *Protein Sci.* **11**, 371–380 (2002).

48. M. Silow, Y. J. Tan, A. R. Fersht, M. Oliveberg, Formation of short-lived protein aggregates directly from the coil in two-state folding. *Biochemistry* **38**, 13006–13012 (1999).

49. Y. W. Chen, K. Stott, M. F. Perutz, Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1257–1261 (1999).

50. S. Sambashivan, Y. Liu, M. R. Sawaya, M. Gingery, D. Eisenberg, Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure. *Nature* **437**, 266–269 (2005).

51. L. Liu, I. J. Byeon, I. Bahar, A. M. Gronenborn, Domain swapping proceeds via complete unfolding: A 19F- and 1H-NMR study of the cyanovirin-N protein. *J. Am. Chem. Soc.* **134**, 4229–4235 (2012).

52. F. Rousseau, J. W. Schymkowitz, H. R. Wilkinson, L. S. Itzhaki, Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5596–5601 (2001).

53. R. Jerala, E. Zerovnik, Accessing the global minimum conformation of stefin A dimer by annealing under partially denaturing conditions. *J. Mol. Biol.* **291**, 1079–1089 (1999).

54. X. Kang *et al.*, Foldon unfolding mediates the interconversion between M(pro)-C monomer and 3D domain-swapped dimer. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14900–14905 (2012).

55. A. Morrone *et al.*, The denatured state dictates the topology of two proteins with almost identical sequence but different native structure and function. *J. Biol. Chem.* **286**, 3863–3872 (2011).

56. R. Giri *et al.*, Folding pathways of proteins with increasing degree of sequence identities but different structure and function. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17772–17776 (2012).

57. L. Yang, J. A. Irving, W. Dai, M. I. Aguilar, S. P. Bottomley, Probing the folding pathway of a consensus serpin using single tryptophan mutants. *Sci. Rep.* **8**, 2121 (2018).

58. A. I. Gilson, A. Marshall-Christensen, J. M. Choi, E. I. Shakhnovich, The role of evolutionary selection in the dynamics of protein structure evolution. *Biophys. J.* **112**, 1350–1365 (2017).

59. T. Alhindi *et al.*, Protein interaction evolution from promiscuity to specificity with reduced flexibility in an increasingly complex network. *Sci. Rep.* **7**, 44948 (2017).

60. E. H. McConkey, Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3236–3240 (1982).

61. S. Ribeiro, S. Ebbinghaus, J. C. Marcos, Protein folding and quinary interactions: Creating cellular organisation through functional disorder. *FEBS Lett.* **592**, 3040–3053 (2018).

62. Y. Liu, I. Bahar, Sequence evolution correlates with structural dynamics. *Mol. Biol. Evol.* **29**, 2253–2263 (2012).

63. Z. Nevin Gerek, S. Kumar, S. Banu Ozkan, Structural dynamics flexibility informs function and evolution at a proteome scale. *Evol. Appl.* **6**, 423–433 (2013).

64. E. C. Campbell *et al.*, Laboratory evolution of protein conformational dynamics. *Curr. Opin. Struct. Biol.* **50**, 49–57 (2018).

65. G. Abrusán, J. A. Marsh, Alpha helices are more robust to mutations than beta strands. *PLoS Comput. Biol.* **12**, e1005242 (2016).

66. C. A. Barnes *et al.*, Remarkable rigidity of the single α-helical domain of myosin-VI as revealed by NMR spectroscopy. *J. Am. Chem. Soc.* **141**, 9004–9017 (2019).

67. Y. Yuan, M. F. Tam, V. Simplaceanu, C. Ho, New look at hemoglobin allostery. *Chem. Rev.* **115**, 1702–1724 (2015).

68. M. D. Vesper, B. L. de Groot, Collective dynamics underlying allosteric transitions in hemoglobin. *PLoS Comput. Biol.* **9**, e1003232 (2013).

69. I. Kass *et al.*, Cofactor-dependent conformational heterogeneity of GAD65 and its role in autoimmunity and neurotransmitter homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2524–E2529 (2014).

70. M. Mravic *et al.*, Packing of apolar side chains enables accurate design of highly stable membrane proteins. *Science* **363**, 1418–1423 (2019).

71. D. S. Moss, S. Jelaska, S. Pongor, Eds., *Essays in Bioinformatics*, (IOS Press, Amsterdam, 2005).

72. N. Goldman, J. L. Thorne, D. T. Jones, Assessing the impact of secondary structure and solvent accessibility on protein evolution. *Genetics* **149**, 445–458 (1998).

73. E. Campbell *et al.*, The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **12**, 944–950 (2016).