Structure *14*, 5–14, January 2006 ©2006 Elsevier Ltd All rights reserved DOI 10.1016/j.str.2005.09.008

Topological Determinants of Protein Domain Swapping

Feng Ding,[1] Kirk C. Prutzman,[1] Sharon L. Campbell,[1] and Nikolay V. Dokholyan[1,*]

[1] Department of Biochemistry and Biophysics
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599

Summary

Protein domain swapping has been repeatedly observed in a variety of proteins and is believed to result from destabilization due to mutations or changes in environment. Based on results from our studies and others, we propose that structures of the domain-swapped proteins are mainly determined by their native topologies. We performed molecular dynamics simulations of seven different proteins, known to undergo domain swapping experimentally, under mildly denaturing conditions and found in all cases that the domain-swapped structures can be recapitulated by using protein topology in a simple protein model. Our studies further indicated that, in many cases, domain swapping occurs at positions around which the protein tends to unfold prior to complete unfolding. This, in turn, enabled prediction of protein structural elements that are responsible for domain swapping. In particular, two distinct domain-swapped dimer conformations of the focal adhesion targeting domain of focal adhesion kinase were predicted computationally and were supported experimentally by data obtained from NMR analyses.

Introduction

Genetic alterations or changes in cellular environment such as salt concentration or pH value often destabilize proteins, and they can lead to structural alterations or misfolding. One of the consequences of protein destabilization is protein aggregation, which has been associated with a number of human diseases, such as Huntington's and Parkinson's diseases. Protein aggregation can result in the formation of stable, often hyperstable, oligomeric structures. This process can be thought of as a natural compensation for protein destabilization and rescue from degradation. The protein domain-swapping phenomenon, in which a fragment of a protein exchanges with a corresponding fragment of another like protein, was described by Eisenberg and coauthors in 1994 (Bennett et al., 1994) and has been repeatedly observed since in a variety of proteins. It has been proposed that domain swapping is similar to the aggregation phenomenon, i.e., certain proteins oligomerize to compensate for protein instability. Furthermore, naturally occurring domain-swapped proteins may evolve from monomeric proteins that became destabilized due to mutations. There are a number of key experimental observations that support the destabilization-compensation hypothesis (O'Neill et al., 2001; Byeon et al., 2003; Frank et al., 2002; Zegers et al., 1999; Kishan et al., 2001; Ultsch et al., 1999; Liu and Eisenberg, 2002).

First, reengineered proteins tend to oligomerize (O'Neill et al., 2001; Byeon et al., 2003; Frank et al., 2002; Green et al., 1995; Schymkowitz et al., 2001; Rousseau et al., 2001; Murray et al., 1998). For example, O'Neill et al. (2001) observed that a single amino acid substitution (V49A) in the B1 domain of protein L resulted in perturbation of the hydrophobic core and dimerization of PpI^{V49A}. Byeon et al. (2003) observed that the immunoglobulin binding domain (B1) of streptococcal protein G (GB1) oligomerized and, upon mutagenesis, induced destabilization of its hydrophobic core (L5V/F30V/A34F or L5V/F30V/Y33F/A34F). Moreover, a single core mutation (A26F) resulted in GB1 tetramerization (Frank et al., 2002). Second, subjecting proteins to destabilizing environments in some cases can result in oligomerization. For example, barnase, a monomeric protein under physiological pH, forms a domain-swapped trimer at pH 4.5 (Zegers et al., 1999). The Eps8 SH3 domain forms a domain-swapped dimer when separated from the rest of the protein (Kishan et al., 2001). Third, truncated proteins often exhibit domain swapping. For example, truncation of Domain 5 from TrkA, TrkB, and TrkC proteins results in domain swapping (Ultsch et al., 1999). Other examples are described by Liu and Eisenberg (2002).

Here, we propose that native topology is a general determinant of protein domain-swapped structures, consistent with previous work on domain-swapped dimers of SH3 domains (Ding et al., 2002b; Yang et al., 2004). To test our hypothesis, we studied the dimerization of seven distinct proteins shown experimentally to form domain-swapped dimers. Our results indicate that domain swapping occurs at a hinge region, around which proteins tend to locally unfold prior to complete unfolding. To aid in identification of protein hinge regions that promote formation of domain-swapped dimers, we have constructed a theoretical predictor. The predictor correctly identified hinge regions for all of the proteins under study (A web-based server for hinge region prediction is available at: http://dokhlab.unc.edu/tools/h-predictor/).

The focal adhesion targeting domain (FAT) of focal adhesion kinase (FAK) can form a domain-swapped dimer in which helix 1 exchanges with a second molecule of the FAT domain (Arold et al., 2002), and we were able to computationally reproduce the domain-swapped dimer structure observed by X-ray crystallography. Additionally, an alternative domain-swapped dimer structure was observed in simulations, in which the N-terminal region of the FAT domain containing helix 1, loop 1, and helix 2, rather than helix 1 alone, formed the dimer interface. Although this alternative domain-swapped dimeric form of the FAT domain has not yet been confirmed experimentally, NMR amide resonances corresponding to residues A978 and S979 within loop 2 are broadened and undetectable, indicating that these resonances may be in conformational exchange on the NMR timescale (see Figure 3H) (Prutzman et al., 2004). Moreover, loop 2 contains two conserved prolines that may induce conformational strain and create a second hinge region.

*Correspondence: dokh@med.unc.edu

Table 1. The List of Seven Proteins that Form Domain-Swapped Dimers in Experiments

Protein	PDB ID	Number of Residues per Monomer	Hinge Region		
			Experiment (Dimer)	Simulation (Monomer)	Prediction (Monomer)
De novo three-helix bundle	M:N/A	48 (1–48)	33–34	32–35	32–33
	D:1G6U (Ogihara et al., 2001)	48 (1–48)			
Ribonuclease A	M:1A5P (Pearson et al., 1998)	124 (1–124)	15–22, 112–115	15–22	19–21
	D:1A2W (Liu et al., 1998)	124 (1–124)			
SH3 domain	M:1NLO (Feng et al., 1996)	56 (9–64)	34–39 (Src loop)	37–41 (Src loop)	36–41
	D:1AOJ (Kishan et al., 1997)	61 (5–65)			
Staphylococcal nuclease	M:1EYD (Chen et al., 2000)	135 (6–141)	112–120$\Delta^{114–119}$	115–119	115–116
	D:1SND (Green et al., 1995)	135 (7–141$\Delta^{114–119}$)			
VHH domain	M:1QD0 (Spinelli et al., 2000)	128 (1–128)	95–100 (β8-β9 loop)	107–113 (β8-β9 loop)	106–109
	D:1SJV (Spinelli et al., 2004)	107 (8–110, +25A–25E)			
Cyanovirin-N	M:2EZM (Bewley et al., 1998)	101 (1–101)	50–53	50–53	49–51
	D:3EZM (Yang et al., 1999)	101 (1–101)			
FAT domain	M:1PV3 (Prutzman et al., 2004)	134 (920–1053)	943–948 (H1-H2)	943–948 (H1-H2); 974–978 (H2-H3)	943–948 (H1-H2); 974–978 (H2-H3, second minimum)
	D:1K04 (Arold et al., 2002)	142 (908–1049)			

The second column contains the PDB codes for the monomeric (M) and dimeric (D) states of these proteins. "Δ" indicates the deletion mutation in the staphylococcal nuclease. In the case of the SH3, VHH, and FAT domains, the monomeric and dimeric states of the proteins are determined for different, but homologous, proteins. For these proteins, due to a discrepancy in the numbering of the protein residues in the monomeric and dimeric forms, we show the region of homology in the brackets (columns 4 and 5). Columns 4, 5, and 6 are the hinge regions for the seven proteins observed in experiments, determined in simulations, and predicted by using a simplified physical model of a monomer protein. The index in column 4 is taken from the dimer PDB file, and the indices in columns 5 and 6 are taken from the monomer PDB files.

Taken together, these data and observations provide support for the existence of an alternative domain-swapped dimer form of the FAT domain that differs from the one observed by X-ray crystallography.

Results and Discussion

To test our hypothesis that protein topology is a determinant in domain swapping, we performed molecular dynamics simulations on seven proteins with different folds (Table 1 and Figure 1): de novo-designed three-helix bundle (Ogihara et al., 2001), ribonuclease (RNase) A (Liu et al., 1998, 2001), the SH3 domain (Kishan et al., 1997), staphylococcal nuclease (SNase) (Green et al., 1995), camelid heavy-chain antibody VH domain (VHH) (Spinelli et al., 2004), cyanovirin-N (CVN) (Boyd et al., 1997; Mori et al., 1997; Bewley et al., 1998; Yang et al., 1999), and the focal adhesion targeting (FAT) domain of focal adhesion kinase (Arold et al., 2002; Hoellerer et al., 2003).

From equilibrium simulations of the Src SH3 domain (Ding et al., 2002b) at various temperatures we found that the Src SH3 domain aggregated most likely near the transition temperatures, T_F, where the Src SH3 domain coexists in native and unfolded states. Therefore, to find the optimal temperature for domain swapping of the proteins under study, we first performed a quasi-equilibrium molecular dynamics simulation of single monomeric proteins to estimate their T_F (Ding et al., 2005). The temperature for each protein was then gradually increased until it fully unfolded. The T_F was identified by monitoring sharp potential energy, E, transitions and the radius of gyration (R_G) of the protein. The temperature of the system was maintained by using a Berendsen's thermostat (Berendsen et al., 1984). To study the formation of domain-swapped dimers, we performed simulations on two identical proteins in a cubic box with a side length of 150 Å (with periodic boundary conditions) at a concentration of 1 mM. For each protein, we performed ten separate simulations near T_F with different initial velocities and relative orientations for each protein. Consistent with our previous study (Ding et al., 2002b), domain-swapped dimers formed only in some simulations, while other simulations led to amyloidogenic aggregates. Only the domain-swapped dimers were selected. All-atom structures of the domain-swapped dimers were generated from the coarse-grained structures after the three-step process described in Experimental Procedures. The computationally determined domain-swapped dimers were compared with the experimental structures, as shown in Figure 1. The similarity between experimental and computationally determined structures of the domain-swapped dimers is striking. Since amino acid interactions in our simulations were determined by the native protein topology, the agreement observed between experimentally determined structures and our simulations suggests that domain swapping is determined by protein topology.

Domain-Swapped Dimers of Seven Different Proteins

De Novo-Designed Three-Helix Bundle

Ogihara et al. (2001) designed a domain-swapping peptide by deleting the loops of a Ser coil peptide. The designed peptide can form both a domain-swapped dimer and fibrous oligomers. The dimer is formed by two peptides exchanging their N-terminal helices. As the structure of this protein in its monomeric state has not been determined, we reconstructed a monomeric conformation from the dimer structure, as described in Experimental Procedures. The monomer and dimer structures are shown in Figure 1A.

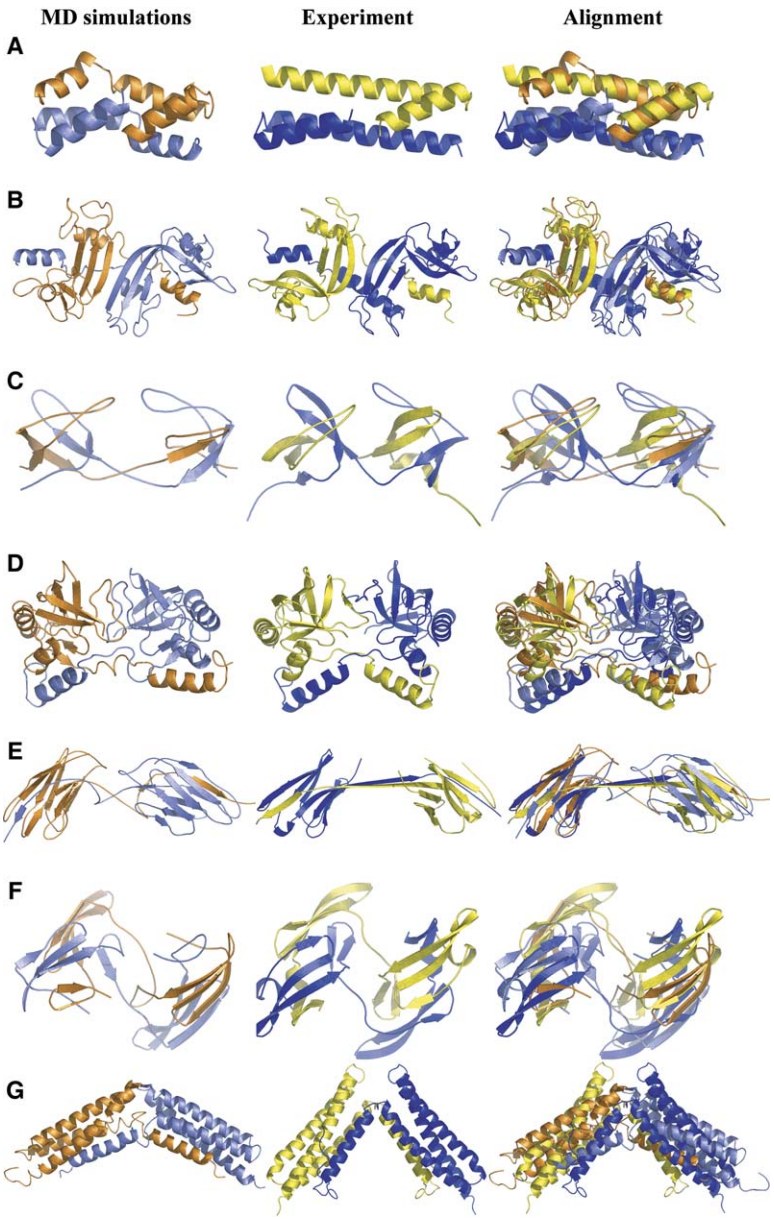

Figure 1. Seven Domain-Swapped Proteins

(A–G) The first and the second columns are the domain-swapped dimeric structures obtained in DMD simulations and experiments, respectively. The third column is the alignment of computational and experimental structures. (A) De novo three-helix bundle, (B) Ribonuclease A, (C) the SH3 domain, (D) staphylococcal nuclease, (E) the VHH domain, (F) Cyanovirin-N, and (G) the FAT domain.

Ribonuclease A

RNase A (PDB: 1a5p) is an enzyme that catalyzes the breakdown of RNA into smaller components. RNase A can adopt two predominant forms of domain-swapped dimers depending on whether the N-terminal α helix (Liu et al., 1998) (Figure 1B) or the C-terminal β strand (Liu et al., 2001) exchange. However, in simulations we only observed the N-terminal α helix-swapped dimers (PDB: 1a2w). Although we believe that protein topology is an important determinant, other mechanisms such as structural strain (Dehouck et al., 2003) also affect the formation of domain-swapped structures.

The SH3 Domain

SH3 domains are found in numerous signal transduction proteins and play an important role in mediating protein-protein interactions. Intriguingly, the Eps8 SH3 domain (Kishan et al., 1997) forms a domain-swapped dimer. Previously, we simulated the aggregation process of the Src SH3 domain (PDB: 1nlo), a homolog of the Eps8 SH3 (PDB: 1aoj) domain, and observed the formation of a domain-swapped dimer (Figure 1C and Figure 1B in Ding et al. [2002b]) similar to the experimentally determined Eps8 SH3 dimer.

Staphylococcal Nuclease

SNase is an enzyme that hydrolyzes the phosphodiesterase bonds in DNA or RNA. SNase (PDB: 1eyd) shows 135 ordered residues in the crystal structure (Chen et al., 2000), consisting of an N-terminal β sheet domain and a C-terminal cluster of α helices. SNase forms a domain-swapped dimer structure (Green et al., 1995) upon deletion of a 6 residue surface loop (PDB: 1snd) (Figure 1D). The dimer is formed by exchange of the C-terminal helices. Maity and Eftink (2004) demonstrated that denaturing conditions induced a thermodynamic intermediate SNase state in which the C-terminal helix unfolds, consistent with the dimer structure reported by

Green et al. (1995). The dimer structure (Figure 1D) from our simulations further confirms these experimental findings.

VHH Domain

Among mammals, camelids have a unique immunological system, since they produce functional antibodies devoid of light chains and CH1 domains. Camelids use the single VH domain from their heavy chain (VHH) to bind antigens. Spinelli et al. (2004) reported a llama VHH domain (PDB: 1sjv) that builds a crystal-wide β sheet structure by forming a domain-swapped dimer (Figure 1E). The authors postulated that strain in the C-terminal strand of monomeric VHH leads to a domain-swapped dimer in which the C-terminal strands are exchanged. A homologous monomeric VHH domain (PDB: 1qd0) was used in our simulations. Since our model only employs a structure-based interaction potential, the observed domain-swapped dimer formed in our simulation suggests that swapping of the C-terminal strands is an intrinsic property of the monomer fold.

Cyanovirin-N

CVN is a monomeric 11 kDa cyanobacterial protein that inactivates diverse strains of human immunodeficiency virus through high-affinity interactions with the surface envelope glycoprotein gp120 to disrupt cell fusion (Boyd et al., 1997; Mori et al., 1997). Monomeric CVN (PDB: 2ezm) (Bewley et al., 1998) is an elongated, largely β sheet protein that displays internal 2-fold pseudo symmetry. The two sequence repeats (residues 150 and 51,101) share 32% sequence identity and superimpose with a backbone atomic root mean square difference of 1.3 Å. Interestingly, in the crystal, CVN forms a domain-swapped dimer (PDB: 3ezm) by complementary exchange of the two subdomains (residues 150 and 51,101) (Yang et al., 1999). The dimer structure from our simulations shows good agreement with the crystal structure (Figure 1F).

The Focal Adhesion Targeting Domain of Focal Adhesion Kinase

FAK is a 125 kDa protein that colocalizes with integrins at focal adhesions upon cell adhesion to the extracellular matrix (Schlaepfer et al., 1999; Schaller, 2001; Parsons, 2003), and it is involved in regulating a number of biological processes such as cell motility and cell survival. The FAT domain is a C-terminal domain of FAK that adopts an amphipathic, antiparallel four-helix bundle fold (Arold et al., 2002; Prutzman et al., 2004; Hayashi et al., 2002). The FAT domain is critical for targeting FAK to focal adhesions where FAK is activated (Shen and Schaller, 1999; Cooley et al., 2000). Interestingly, in simulations we found that the FAT domain (avian, PDB: 1pv3) forms two types of domain-swapped dimers. In these dimers, (1) exchange occurs between the first N-terminal helix of each monomer (Figure 1G) and (2) two helices in the N terminus are exchanged. The first type of FAT domain dimer (human, PDB: 1k04) has been observed by X-ray crystallography (Arold et al., 2002).

We have computationally identified dimer structures for seven different proteins that are known to undergo domain swapping. These structures are in agreement with the experimental structures. In the case of RNase A, we do not observe the second type of the domain-swapped dimer (Liu et al., 2001), in which the C-terminal β strands are swapped. Given our results, we believe that, while protein topology is an important determinant of the domain-swapped structures, other factors such as conformational strain (Dehouck et al., 2003) also affect the formation of domain-swapped structures. Moreover, using only structural information of the native monomers, our model cannot reproduce non-native packing of domain-swapped dimers, as often revealed in the experimental structures.

Interestingly, simulations on the FAT domain suggest the existence of an alternative domain-swapped dimer. However, this second domain-swapped dimeric form of FAT domain has not been previously observed experimentally. NMR data presented here support the existence of this domain-swapped FAT domain structure (discussed below).

The Hinge Region of Domain Swapping Can Be Predicted from the Monomeric Protein Topology

One possible reason why some proteins oligomerize through domain swapping may be due to heterogeneity of interactions between various protein fragments with the rest of the protein. For example, in the case of the Eps8 SH3 domain, the largest, least-structured fragment of the protein, the Src loop, unfurls and exchanges with the Src loop from a second molecule (Kishan et al., 2001). Simulations of the Src SH3 domain, a homologous protein of the Eps8 SH3 domain, reproduced a domain-swapped dimer, in which fragments of the Src loops are exchanged (Ding et al., 2002b). Dehouck et al. (2003) argued that "frustrations" in protein sequences are the "hot-spots" for domain swapping, i.e., amino acid fragments that are not optimal with regard to the stability of the native structure are more prone to be exchanged with the same fragments of identical proteins. Here, we propose to characterize the structural determinants of protein domain swapping. In most experimentally observed domain-swapped oligomers, the swapped domains correspond to one or several secondary structural elements from either the N or C termini (Liu and Eisenberg, 2002). Only in some rare instances are the swapped domains positioned in the middle of the protein. The domain-swapped oligomeric structures are, therefore, mainly determined by the location and the properties of the hinge region. Indeed, in simulations we observed that hinge regions in all seven proteins are the "hot-spots," around which proteins tend to locally unfold prior to complete protein unfolding. Our simulations also suggest that the tendency of the hinge region to unfold locally is mainly determined by the overall structure of the monomer.

We further propose that domain-swapped proteins are formed by their "intermediate" states, which resemble the structure of a single peptide in the domain-swapped oligomers. In simulations that rely on a priori knowledge of protein structure, we observed these states in all seven proteins at timescales that support dimerization. We have thus developed a simple predictor based on protein structure that identifies hinge regions. We virtually created "intermediate" protein states by disrupting the native state of a monomer around residue i into two subdomains, each of which retain their native structures (Figure 2). We simply removed from the native contact map all of the interactions between residues j and k, where $j < i$ and $k > i$. The hypothetical intermediate

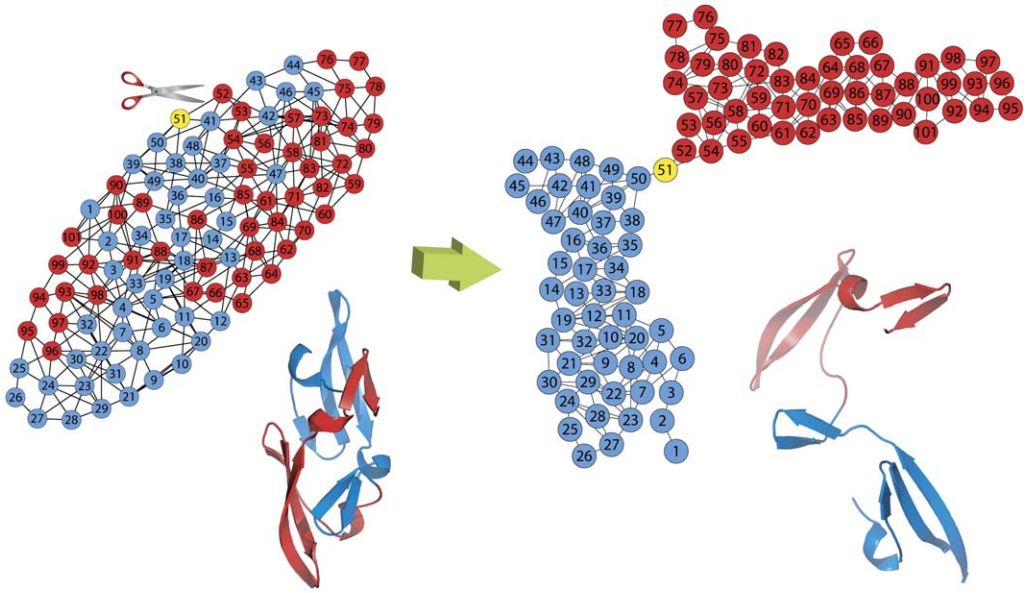

Figure 2. Scheme Associated with the Simplified Physical Model for Predicting the Hot-Spot Hinge Regions

For each protein, we defined a graph that consists of nodes representing protein residues and edges connecting those pairs of residues that are in contact. The change in entropy was calculated by changes in the properties of the amino acid interaction networks upon a hypothetical disruption of the monomer into two subdomains near the hinge region. The two subdomains are colored blue and red. In this scheme, we disrupted the interactions between the two subdomains of CVN separated by hinge region residue 51, which is colored yellow.

state corresponds to the conformational ensemble consistent with the resulting contact map. To identify the most probable intermediate state, we evaluated for each disruption the energetic and structural changes from the native state. Here, these changes were computed solely based on the contact maps. The intermediate states, which feature smaller enthalpic increases, ΔE (see Experimental Procedures), but larger conformational entropy increases, ΔS, upon disruption of the native state, are expected to appear with a higher probability.

To estimate the change of entropy upon subdomain partitioning of a protein, a protein graph (Dokholyan et al., 2002) that represents a network of protein interactions was defined (see Experimental Procedures; Figure 2). The minimal path between two residues, *i* and *j*, is defined as a minimal number of nodes on the protein graph needed in order to transverse from *j* to *i*. The average minimal path *L* is the minimal path connecting nodes *i* and *j* on a graph averaged over all pairs of nodes *i* and *j*. Thus, *L* is a measure of typical separation between any two nodes on a graph. This quantity has been proven to be a sensitive measure in differentiating protein conformations with respect to the free energy barrier of the protein folding transition (Dokholyan et al., 2002). Assuming that a protein state with minimal path *L* is analogous to an ideal polymer with chain length *L*, the conformational entropy is proportional to *L* since the entropy of the effective polymer is $k_B \ln(z^L) = k_B L \ln(z)$, where *z* is the coordination number of a monomer in the polymer. Thus, to estimate the change of entropy, ΔS, upon disruption of the protein at residue *i*, the change in the average minimal path, $\Delta L(i)$, was determined.

Since the most probable intermediate states featured small changes in enthalpy, $\Delta E(i)$, and a large change in entropy, $\Delta S(i) \sim \Delta L(i)$, we chose the ratio $\Delta E/\Delta S(i)$ as a predictor of a protein hinge region. Physically, $\Delta E/\Delta S$ is the effective equilibrium temperature between the two states: native state and the corresponding intermediate state with residue *i* as the center of the hinge region. The smaller the value of $\Delta E/\Delta S(i)$, the smaller the temperature at which the intermediate is observed, and, therefore, the higher the probability to find the intermediate state with a hinge region around residue *i*. Hence, ΔE and ΔL can be estimated to construct a predictor, $\Delta E/\Delta S(i) \sim \Delta E/\Delta L(i)$.

The results show (Table 1 and Figures 3A–3G) that the global minimum of $\Delta E/\Delta L$ corresponds to residues involved in the experimentally determined hinge regions. The exceptions are proteins featuring more than one type of domain-swapped dimer. For example, RNase A (Figure 3B) has a second domain-swapping (Liu et al., 2001) hinge region (as indicated by a blue arrow in Figure 3B). The corresponding $\Delta E/\Delta L$ value for this alternative hinge region is even higher than that of the second minimum near residue 90. In addition, comparable local minima are also observed (such as Figures 3A and 3E). The complexity in interpreting $\Delta E/\Delta L$ minima may result from our use of a simple model that does not take into account other factors besides topology, such as the heterogeneity in pairwise interactions and conformational strain in the native structure (Dehouck et al., 2003). In order to construct an elaborate predictor of the domain-swapping hinge region, one needs to include features other than the topology. Nevertheless, an agreement between the predicted and observed hinge regions supports our postulation that native protein topology is an important determinant of protein domain swapping. Given these results, we postulate that regions with global minima in $\Delta E/\Delta L$ can provide experimentally testable leads for identifying domain-swapping structures,

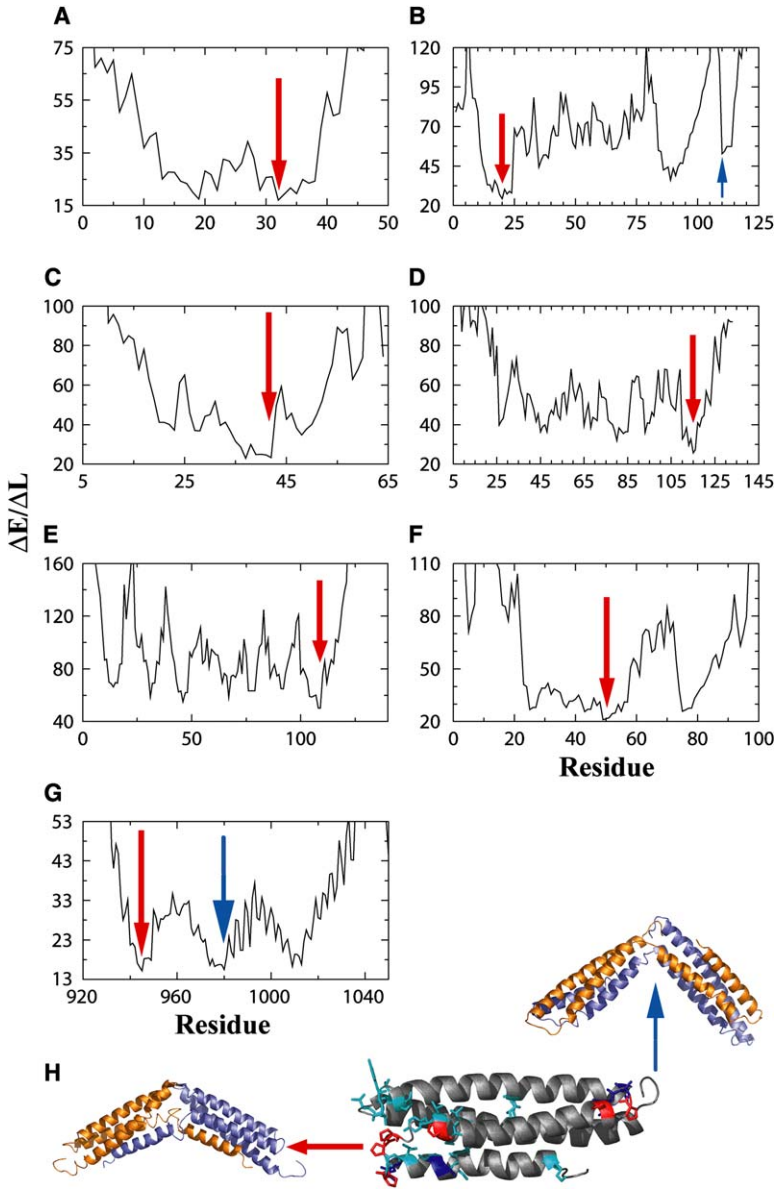

Figure 3. Identifying the Hot-Spot Hinge Regions of Domain Swapping

(A–G) The predictor $\Delta E/\Delta L(i)$ as a function of residue index i for seven proteins: (A) De novo three-helix bundle, (B) RNase A, (C) the SH3 domain, (D) staphylococcal nuclease, (E) the VHH domain, (F) Cyanovirin-N, and (G) the FAT domain. The global minimum of $\Delta E/\Delta L(i)$ corresponds to residues belonging to the predicted hinge regions (red arrow). In some cases, there are competing, but not global, minima that do not correspond to a domain-swapped structure, except for RNase A and the FAT domain. In the case of RNase A, there are two experimentally observed hinge regions, corresponding to N (red arrow) and C termini (blue arrow). The predictor $\Delta E/\Delta L$ for the second hot-spot region does not reach its global minimum, suggesting that the C-terminal domain-swapped dimer may appear with lower probability than the N-terminal one. The FAT domain case is similar to the RNase case.

(H) Ribbon diagram of the avian FAT domain solution structure (PDB: 1PV3). Backbone HN peaks in the (^{1}H-^{15}N)-HSQC spectra that are either unobservable (blue) or significantly less intense when the temperature is lowered from 37°C to 27°C (cyan), indicative of conformational exchange, are indicated. Proline residues (red) are predicted to create strain in loop 1 (red arrow) and loop 2 (blue arrow). The left and right structures of two domain-swapped FAT domains obtained in simulations are shown.

and that comparable local minima may be candidates for putative hinge regions, for example, in the case of multiple domain-swapped dimers (Figures 3B and 3H).

In addition, the formation of domain-swapped dimers is not predicated on the existence of a thermodynamically stable intermediate protein state. For example, transient unfolding around the hinge region at high concentrations may promote domain swapping of two-state proteins. The proposed predictor is not a measure of the protein's propensity for domain-swapping, but rather a structural propensity that a hinge region may result in domain swapping.

Two Domain-Swapped Dimers of the FAT Domain

In two cases, RNase A and the FAT domain, we are able to predict and conduct molecular dynamics simulations that support the existence of two types of domain-swapped dimers. The FAT domain has been shown previously to dimerize as a result of helix 1 domain swapping (Arold et al., 2002; Hoellerer et al., 2003). While there is currently no known biological function for the domain-swapped dimer form of the FAT domain, prevailing evidence suggests that the structural plasticity of helix 1 modulates paxillin binding and Y925 (Y926 in avian FAK) phosphorylation, which is important for FAK localization and FAK-mediated signaling and cell adhesion function (Arold et al., 2002; Prutzman et al., 2004). Consistent with observations that the FAT domain forms a domain-swapped dimer (Arold et al., 2002; Prutzman et al., 2004) (Figure 1G), we have determined that the FAT domain is in equilibrium between a monomer and dimeric form by gel filtration chromatography and proposed that the FAT domain contains a strained hinge region between helix 1 and helix 2 due to the presence of three conserved prolines with abnormal φ-ψ angles in the monomer structure (Arold et al., 2002; Prutzman et al., 2004). Moreover, analytical ultracentifugation results show that the FAT domain undergoes self-association (K_d = 0.5 mM) at

pH 7.5, thereby providing additional support for FAT domain dimerization in solution. We thus hypothesized that a strained hinge region exists in loop 1 of the FAT domain, which promotes the formation of a domain-swapped dimer. This premise was based on several observations (Prutzman et al., 2004). A PROCHECK (Laskowski et al., 1996) analysis of the 20 reported solution structures revealed that residues 943–951 in and around loop 1 sample aberrant φ-ψ angles, as determined by low average G factor values. The G factor provides a measure of how "normal" or, alternatively, how "unusual" a given stereochemical property is, with a low G factor indicating that the property (φ-ψ torsion angles in this case) corresponds to a low probability conformation. Low G factors were observed for two or more residues within this putative hinge region in each structure of the ensemble. A PROCHECK structural analysis of the 20 avian FAT domain solution structures also showed that the same residues had high circular variance values. The circular variance provides a measure of the spread of a set of dihedral angles, and it is a measure of how tightly or loosely a given residue's φ or ψ torsion angles cluster together across the entire family or ensemble. Similar φ and ψ angle violations in hinge residues 946–950 (corresponding to residues 947–951 in the avian FAT domain) were observed in the X-ray crystal structure of the mouse FAT domain monomer (PDB: 1k40) reported by Hayashi et al. (2002). In addition, the monomeric form of the human FAT domain reported by Arold et al. (2002) showed low G factor values for residues 939–950 (residues 940–951 in the avian FAT domain). In this study, the monomer crystallized with three molecules in the unit cell. Interestingly, there was variability in the G factor values for residues in this putative hinge region among the three molecules, suggesting that this region adopts multiple conformations. In contrast, only a slight φ-ψ angle violation was observed in P944 (945 in the avian FAT domain) in the structure of the domain-swapped dimer (Arold et al., 2002). Further, an NMR analysis showed that residues in and around loop 1 display temperature- and field-dependent line broadening indicative of conformational exchange (Prutzman et al., 2004). The observed φ-ψ angle violations, high circular variance, and temperature-dependent amide peak broadening in the putative hinge region of the FAT domain support the existence of conformational strain in this region.

We thus postulated that the proline-rich region in loop 1 caused conformational strain resulting in enhanced conformational dynamics and partitioning of helix 1 away from the four-helix bundle (Arold et al., 2002; Prutzman et al., 2004), consistent with our detection of an intermediate form of the FAT domain in which helix 1 partitions from the helix bundle-core and becomes more conformationally mobile and less structured. This intermediate was detected by using a novel method employing NMR-derived amide protection factors and discrete molecular dynamic simulations (Dixon et al., 2004) and is further supported by gel filtration, NMR (Figure 3H), and structural analyses (Prutzman et al., 2004).

An additional domain-swapped dimer was predicted in the present study, suggesting that a second hinge region may exist in the FAT domain. However, in this alternative domain-swapped dimer form of the FAT domain, both helix 1 and helix 2, rather than only helix 1, partition away from the four-helix bundle to form a dimer with a second molecule (Figure 1G). In support of the existence of conformational strain in loop 2 (residues 973–980 in the avian FAT domain), this loop contains two conserved prolines (P974, P977 in the avian FAT domain), and residues A978 and S979 exhibit NMR exchange broadening (Figure 3H). Thus, in addition to those in loop 1, our NMR data suggest that residues in loop 2 may undergo conformational exchange between two or more conformations on the microsecond–millisecond timescale. We also observed that residues 974–979 (avian FAT domain) in loop 2 had φ-ψ angle violations in both the NMR solution structures as well as the X-ray crystal structures, as determined by low G factor values from a PROCHECK analysis. However, it is possible that the observed violations in loop 2 may be due to their location in a loop, as similar low G factor values were observed in loop 3. Additionally, loop 2 residues 977–980 (avian FAT domain) in the family of NMR solution structures had high circular variance values and may sample multiple conformations. As these observations may also reflect a lack of restraints in structure calculations, further studies are underway to determine whether this second domain-swapped dimer form of the FAT domain exists in solution and if so, whether this form of FAT domain plays a role in FAK biology. While the data presented in this study (NMR, $\Delta E/\Delta L$ model predictor, φ-ψ angles) provide support for the role of proline-induced strain in the promotion of FAT domain swapping, we recognize that additional or distinct processes may facilitate domain swapping. For example, we are intrigued by the presence of a methionine cluster that lies close to the putative hinge region between helix 1 and helix 2, which may promote entropic strain in FAT domain. Moreover, as helix 1 is the shortest helix in FAT domain, there may be fewer interactions with helices 2–4. In fact, our computational studies suggest that helix 1 makes fewer stabilizing interactions within the helix interface in the monomer form of the FAT domain relative to the other helices.

Biological Implications

Often, domain-swapped dimers form in living cells, and the functional relevance of dimeric states or intermediate states leading to dimerization has been established for several systems (Barrientos and Gronenborn, 2002; Calarese et al., 2003). Thus, dimerization may function to either downregulate or upregulate (Calarese et al., 2003) protein function by altering the activity of the dimerizing proteins (Rousseau et al., 2003). For example, the human antibody 2G12 domain-swapped VH dimer neutralizes a broad range of human immunodeficiency virus type 1 (HIV-1) isolates by binding an unusually dense cluster of carbohydrate moieties on the "silent" face of the gp120 envelope glycoprotein (Calarese et al., 2003). In this case, dimerization may alter cellular function by increasing the concentration of active molecules at a specific location within the cell, without affecting the activity of the dimerizing proteins. Intriguingly, Glyoxalase I from *Pseudomonas aeruginosa* can exist both in an active domain-swapped dimeric state and as a metastable and less active monomer (Saint-Jean et al., 1998). Glutathione can modulate the relative populations of dimeric and monomeric glyoxalase I in vitro.

Although no evidence exists for such a mechanism in vivo, this finding demonstrates that functional regulation by domain swapping can, in principle, be achieved.

Dimerization may also play a role in preventing destabilized proteins from forming cytotoxic aggregates. For example, Ding et al. (2002b) found that the Src SH3 domain forms two types of dimeric states upon destabilization: a closed and an open state. The former is a domain-swapped dimer (Figure 1C); the latter is an aggregation-prone state capable of forming larger amyloid aggregates (Ding et al., 2002b). The competitive formation of the domain-swapped dimers decreases the probability of formation of amyloid fibrils.

Domain swapping has also been proposed as a mechanism for protein misfolding, aggregation, and amyloid formation (Liu et al., 2001; Fink, 1998). For example, a mutant of cystatin C forms domain-swapped dimers that may represent precursors to amyloid formation (Sanders et al., 2004; Staniforth et al., 2001; Janowski et al., 2001), and thus may function in amyloid disease. Although it is not clear how domain swapping gives rise to the final assembly of cross-β structures characteristic of amyloid fibrils, this phenomenon may play a role in the early stages of fibril formation.

Domain swapping has also been found to play a role in the structural organization of viruses. For example, the viral capsid of rice yellow mottle virus (RYMV) contains an icosahedral caspid structure and was recently found to be composed of domain-swapped dimers. Interestingly, RYMV has higher thermostability than other viruses from the same family that lack the domain-swapped architecture (Qu et al., 2000). Finally, domain swapping has been proposed to contribute to structural diversification and the emergence of oligomers during evolution (Schlunegger et al., 1997).

It is clear that the biological role of domain-swapped dimers cannot be disregarded. Our simple predictor of domain swapping, $\Delta E/\Delta L$, may thus aid in efforts to alter the stability of proteins prone to domain swapping to develop novel reagents that modulate a wide array of biological processes.

Experimental Procedures

Simulations of Coarse-Grained Protein Models

All proteins were modeled as "beads-on-a-string," in which each amino acid is represented by its C_α and C_β atoms (C_α atom only for Gly). The neighboring atoms along the peptide chain are covalently constrained according to the corresponding distance distributions obtained from the Protein Data Bank (PDB) (Ding et al., 2002a), forcing our coarse-grained models to mimic the peptide backbone flexibility and reduce unphysical degrees of freedom. To study the formation of the domain-swapped dimer, we introduced structure-based amino acid interactions. These include intra- and interchain amino acid interactions, Gō interactions, and nonspecific backbone hydrogen bonds (Ding et al., 2002b). In the Gō interaction model, the native contacts are favored over the non-native contacts by assigning attractive interaction potentials to the former. The interaction strengths of the structure-specific Gō potential and nonspecific backbone hydrogen bonds were assigned as ε and 3ε, respectively. Recently, Yang et al. (2004) applied the same type of interaction potentials to study domain-swapped dimer formation of the Eps8 SH3 domain. Although a different computational approach from ours was used in their study (Yang et al., 2004), the final structure appeared to be the same as that predicted by Ding et al. (2002b), which is in agreement with the experimental structure (Kishan et al., 1997). A rapid discrete molecular dynamics (DMD) algorithm (Dokholyan et al., 1998, 2003; Ding et al., 2002a) was used to perform simulations on the model proteins.

Many proteins that form domain-swapped dimers contain disulfide bonds in their folded monomeric states. In order to model the covalent feature of the disulfide linkage, we introduced permanent constraints between these native disulfide cysteine pairs to prevent disruption during the DMD simulations. In addition, structures of monomeric forms for many domain-swapped proteins, such as the de novo-designed three-helix bundle, are not available (Ogihara et al., 2001). The monomer form of the de novo-designed three-helix bundle (Ogihara et al., 2001) was modeled by reconstructing the monomer contact map out of the domain-swapped dimer. We assume a native contact between i and j if an intrachain contact is formed between i_1 and j_2, where the subscript denotes different peptides in the domain-swapped dimer. DMD simulations were then performed with the generated contact map to reconstruct the native state of the monomeric protein by annealing a stretched peptide to a low temperature.

Reconstruction of All-Atom Representations from the Coarse-Grained Dimer Structures

The dimeric structures obtained from the DMD simulations are in coarse-grained representations, and secondary structures in these structures are often not well-defined. To obtain atomic resolution structures, an all-atom representation of each domain-swapped protein dimer was reconstructed from typical snapshots taken from simulations of coarse-grained protein models as described above.

Step I: Reconstruction of the Backbone Heavy Atoms N and C′

Based on the typical dimeric structures in the coarse-grained representation, backbone N and C′ were added beside C_α and C_β atoms by keeping each amino acid as a D-amino acid. To reduce the extra rotational freedom of the N and C′ atoms along the C_α and C_β axes, a strong constraint—that the neighboring amino acids must form a planar peptide bond—was used. Harmonic-like potentials were introduced between neighboring backbone atoms: C'_i-N_{i+1}, $C_{\alpha,i}$-N_{i+1}, C'_i-$C_{\alpha,i+1}$, where i represents the residue index. The distances between these pairs correspond to the average distances determined from PDB structures. The geometry of D-amino acids was maintained by introducing strong distance constraints between: $C_{\alpha,i}$-$C_{\beta,i}$, N_i-$C_{\alpha,i}$, $C_{\alpha,i}$-C'_i, and N_i-$C_{\beta,i}$, C_i-C'_i, $C_{\beta,i}$-C'_i. Thus, the handedness of each amino acid is intact after the initial introduction of N and C′ atoms. To keep the structure intact, the C_α atoms were immobilized by setting their mass to infinity. Short molecular dynamic simulations were then performed to relax the system.

Step II: Refinement of Secondary Structure Elements α Helices and β Strands

Nonspecific backbone-backbone hydrogen bonds were introduced as described by Ding et al. (2003). The same Gō interactions as those used to obtain the domain-swapped dimers were assigned between C_β atoms. The interaction strengths of the backbone-backbone hydrogen bonds and side chain-side chain interactions were 5ε and ε, respectively. Short molecular dynamics simulations were performed at low temperatures, i.e., $0.6\ \varepsilon/k_B$, to insure that the overall dimeric structure does not change while the number of hydrogen bonds is maximized.

Step III: Reconstruction of an All-Atom Representation of Protein Dimers

The side chain atoms and backbone O and H atoms for each amino acid were added based on the coordinates of N, C_α, C_β, and C′ atoms, and Monte Carlo-based simulated annealing was employed to search the rotamer space for optimal arrangement of the side chains. The Dunbrack and Cohen (1997) backbone-dependent rotamer library was used to select rotamers according to their natural occurrences. The scoring function for rotamer optimization included van der Waals, solvation, and hydrogen bond interactions. The van der Waals radii and interaction strengths were adopted from the Cedar force field (Hermans et al., 1984). The EEF1 solvation model (Lazaridis and Karplus, 1999) was applied, and a statistical potential for hydrogen bonds, as proposed by Kortemme et al. (2003), was used.

The Hinge Region Predictor

Calculation of Enthalpy Change, ΔE

The enthalpy change $\Delta E(i)$ was estimated from the change in the number of native contacts, ΔN^{NC}; native hydrogen bonds, ΔN^{HBond}

(with strength ϵ^{HB} = 3.0); and native disulfide bonds, ΔN^{SS} (with strength ϵ^{SS} = 100.0): $\Delta E = \Delta N^{NC} + \epsilon^{HB}\Delta N^{Hbond} + \epsilon^{SS}\Delta N^{SS}$.

Construction of Protein Graphs

Graphs corresponding to protein conformations were constructed in which the nodes represent amino acids and the edges represent pairs of amino acids that are geometrically located within the interaction distance from each other. The interaction distance between two amino acids was calculated between their C_β (C_α for Gly) atoms. If the distance was smaller than 7.5 Å, an edge was assigned between the respective amino acids.

Expression and Purification of the FAT Domain

The FAT domain, containing residues 920–1053 of avian focal adhesion kinase plus a 12 amino acid N-terminal linker, was expressed and purified as previously described (Prutzman et al., 2004).

NMR Spectroscopy

Purified proteins were exchanged into NMR buffer (25 mM Tris-d_{11} Maleate-d_2 [pH 6.0], 150 mM NaCl, 0.1% NaN_3, PPACK [1 μM], pefabloc [0.5 mg/ml], and 10% D_2O) by Centricon filtration (MW cutoff 5000). 2D ^1H-^{15}N HSQC experiments were conducted at 27ºC and 37ºC on a Varian INOVA 600 MHz spectrometer as previously described (Prutzman et al., 2004). NMR spectra were processed with NMRPipe (Delaglio et al., 1995) and analyzed with NMRView (Johnson and Blevins, 1994).

Analytical Ultracentrifugation

Sedimentation equilibrium experiments were performed in a Beckman Optima XL-I analytical ultracentrifuge with an An50 Ti rotor. Samples of the FAT domain at concentrations of 0.100 mM, 0.067 mM, and 0.033 mM were equilibrated against 10 mM potassium phosphate (pH 6.5), 150 mM NaCl, and 0.01% NaN3 and were centrifuged at 20,000 rpm at 20ºC. Radial scans at 280 nm were taken every 2 hr. Scans at 12 hr, 14 hr, and 16 hr were identical, indicating that equilibrium conditions were reached. The data were fit, and a monomer-dimer dissociation constant (K_d) was calculated by using Beckman XL-A/XL-I Analysis Software, Version 4.0.

Acknowledgments

We thank J. Hermans, S. Khare, B. Kuhlman, and J. Richardson for stimulating discussions; S. Sharma for kindly setting up the webserver; A. Tripathy for technical assistance; M. King and Dr. Schaller for expressing and purifying the ^{15}N-labeled avian FAT domain. This work is supported in part by Muscular Dystrophy Association grant MDA3720, Research Grant No. 5-FY03-155 from the March of Dimes Birth Defect Foundation, the UNC/IBM Junior Investigator Award (to N.V.D.), and National Institutes of Health PO1 HL451000-10 (to S.L.C.).

Received: June 27, 2005
Revised: September 6, 2005
Accepted: September 7, 2005
Published: January 10, 2006

References

Arold, S.T., Hoellerer, M.K., and Noble, M.E.M. (2002). The structural basis of localization and signaling by the focal adhesion targeting domain. Structure *10*, 319–327.

Barrientos, L.G., and Gronenborn, A.M. (2002). The domain-swapped dimer of cyanovirin-N contains two sets of oligosaccharide binding sites in solution. Biochem. Biophys. Res. Commun. *298*, 598–602.

Bennett, M.J., Choe, S., and Eisenberg, D. (1994). Domain swapping: entangling alliances between proteins. Proc. Natl. Acad. Sci. USA *91*, 3127–3131.

Berendsen, H.J.C., Postma, J.P.M., Vangunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular-dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Bewley, C.A., Gustafson, K.R., Boyd, M.R., Covell, D.G., Bax, A., Clore, G.M., and Gronenborn, A.M. (1998). Solution structure of cyanovirin-N, a potent HIV-inactivating protein. Nat. Struct. Biol. *5*, 571–578.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., Okeefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., et al. (1997). Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. Antimicrob. Agents Chemother. *41*, 1521–1530.

Byeon, I.J., Louis, J.M., and Gronenborn, A.M. (2003). A protein contortionist: core mutations of GB1 that induce dimerization and domain swapping. J. Mol. Biol. *333*, 141–152.

Calarese, D.A., Scanlan, C.N., Zwick, M.B., Deechongkit, S., Mimura, Y., Kunert, R., Zhu, P., Wormald, M.R., Stanfield, R.L., Roux, K.H., et al. (2003). Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. Science *300*, 2065–2071.

Chen, J.M., Lu, Z.Q., Sakon, J., and Stites, W.E. (2000). Increasing the thermostability of staphylococcal nuclease: implications for the origin of protein thermostability. J. Mol. Biol. *303*, 125–130.

Cooley, M.A., Broome, J.M., Ohngemach, C., Romer, L.H., and Schaller, M.D. (2000). Paxillin binding is not the sole determinant of focal adhesion localization or dominant-negative activity of focal adhesion kinase/focal adhesion kinase-related nonkinase. Mol. Biol. Cell *11*, 3247–3263.

Dehouck, Y., Biot, C., Gilis, D., Kwasigroch, J.M., and Rooman, M. (2003). Sequence-structure signals of 3D domain swapping in proteins. J. Mol. Biol. *330*, 1215–1225.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

Ding, F., Dokholyan, N.V., Buldyrev, S.V., Stanley, H.E., and Shakhnovich, E.I. (2002a). Direct molecular dynamics observation of protein folding transition state ensemble. Biophys. J. *83*, 3525–3532.

Ding, F., Dokholyan, N.V., Buldyrev, S.V., Stanley, H.E., and Shakhnovich, E.I. (2002b). Molecular dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. J. Mol. Biol. *324*, 851–857.

Ding, F., Borreguero, J.M., Buldyrey, S.V., Stanley, H.E., and Dokholyan, N.V. (2003). Mechanism for the α-helix to β-hairpin transition. Proteins *53*, 220–228.

Ding, F., Jha, R.K., and Dokholyan, N.V. (2005). Scaling behavior and structure of denatured proteins. Structure *13*, 1047–1054.

Dixon, R.D.S., Chen, Y., Ding, F., Khare, S.D., Campbell, S.L., and Dokholyan, N.V. (2004). New insights into FAK signaling and localization based on detection of a FAT domain folding intermediate. Structure *12*, 2161–2171.

Dokholyan, N.V., Buldyrev, S.V., Stanley, H.E., and Shakhnovich, E.I. (1998). Discrete molecular dynamics studies of the folding of a protein-like model. Fold. Des. *3*, 577–587.

Dokholyan, N.V., Li, L., Ding, F., and Shakhnovich, E.I. (2002). Topological determinants of protein folding. Proc. Natl. Acad. Sci. USA *99*, 8637–8641.

Dokholyan, N.V., Borreguero, J.M., Buldyrev, S.V., Ding, F., Stanley, H.E., and Shakhnovich, E.I. (2003). Identifying importance of amino acids for protein folding from crystal structures. Macromol. Crystallogr. D *374*, 618–640.

Dunbrack, R.L., Jr., and Cohen, F.E. (1997). Bayesian statistical analysis of protein side-chain rotamer preferences. Protein Sci. *6*, 1661–1681.

Feng, S., Kapoor, T.M., Shirai, F., Combs, A.P., and Schreiber, S.L. (1996). Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis. Chem. Biol. *3*, 661–670.

Fink, A.L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. Fold. Des. *3*, R9–R23.

Frank, M.K., Dyda, F., Dobrodumov, A., and Gronenborn, A.M. (2002). Core mutations switch monomeric protein GB1 into an intertwined tetramer. Nat. Struct. Biol. *9*, 877–885.

Green, S.M., Gittis, A.G., Meeker, A.K., and Lattman, E.E. (1995). One-step evolution of a dimer from a monomeric protein. Nat. Struct. Biol. *2*, 746–751.

Hayashi, I., Vuori, K., and Liddington, R.C. (2002). The focal adhesion targeting (FAT) region of focal adhesion kinase is a four-helix bundle that binds paxillin. Nat. Struct. Biol. *9*, 101–106.

Hermans, J., Berendsen, H.J.C., Vangunsteren, W.F., and Postma, J.P.M. (1984). A consistent empirical potential for water-protein interactions. Biopolymers *23*, 1513–1518.

Hoellerer, M.K., Noble, M.E.M., Labesse, G., Campbell, I.D., Werner, J.M., and Arold, S.T. (2003). Molecular recognition of paxillin LD motifs by the focal adhesion targeting domain. Structure *11*, 1207–1217.

Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., and Jaskolski, M. (2001). Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. Nat. Struct. Biol. *8*, 316–320.

Johnson, B., and Blevins, R.A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

Kishan, K.V.R., Scita, G., Wong, W.T., DiFiore, P.P., and Newcomer, M.E. (1997). The SH3 domain of Eps8 exists as a novel intertwined dimer. Nat. Struct. Biol. *4*, 739–743.

Kishan, K.V.R., Newcomer, M.E., Rhodes, T.H., and Guilliot, S.D. (2001). Effect of pH and salt bridges on structural assembly: molecular structures of the monomer and intertwined dimer of the Eps8 SH3 domain. Protein Sci. *10*, 1046–1055.

Kortemme, T., Morozov, A.V., and Baker, D. (2003). An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. J. Mol. Biol. *326*, 1239–1259.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Lazaridis, T., and Karplus, M. (1999). Effective energy function for proteins in solution. Proteins *35*, 133–152.

Liu, Y., and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. Protein Sci. *11*, 1285–1299.

Liu, Y.S., Hart, P.J., Schlunegger, M.P., and Eisenberg, D. (1998). The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-angstrom resolution. Proc. Natl. Acad. Sci. USA *95*, 3437–3442.

Liu, Y.S., Gotte, G., Libonati, M., and Eisenberg, D. (2001). A domain-swapped RNase A dimer with implications for amyloid formation. Nat. Struct. Biol. *8*, 211–214.

Maity, H., and Eftink, M.R. (2004). Perchlorate-induced conformational transition of staphylococcal nuclease: evidence for an equilibrium unfolding intermediate. Arch. Biochem. Biophys. *431*, 119–123.

Mori, T., Shoemaker, R.H., Gulakowski, R.J., Krepps, B.L., McMahon, J.B., Gustafson, K.R., Pannell, L.K., and Boyd, M.R. (1997). Analysis of sequence requirements for biological activity of cyanovirin-N, a potent HIV (human immunodeficiency virus)-inactivating protein. Biochem. Biophys. Res. Commun. *238*, 218–222.

Murray, A.J., Head, J.G., Barker, J.J., and Brady, R.L. (1998). Engineering an intertwined form of CD2 for stability and assembly. Nat. Struct. Biol. *5*, 778–782.

O'Neill, J.W., Kim, D.E., Johnsen, K., Baker, D., and Zhang, K.Y.J. (2001). Single-site mutations induce 3D domain swapping in the B1 domain of protein L from *Peptostreptococcus magnus*. Structure *9*, 1017–1027.

Ogihara, N.L., Ghirlanda, G., Bryson, J.W., Gingery, M., DeGrado, W.F., and Eisenberg, D. (2001). Design of three-dimensional domain-swapped dimers and fibrous oligomers. Proc. Natl. Acad. Sci. USA *98*, 1404–1409.

Parsons, J.T. (2003). Focal adhesion kinase: the first ten years. J. Cell Sci. *116*, 1409–1416.

Pearson, M.A., Karplus, P.A., Dodge, R.W., Laity, J.H., and Scheraga, H.A. (1998). Crystal structures of two mutants that have implications for the folding of bovine pancreatic ribonuclease A. Protein Sci. *7*, 1255–1258.

Prutzman, K.C., Gao, G., King, M.L., Iyer, V.V., Mueller, G.A., Schaller, M.D., and Campbell, S.L. (2004). The focal adhesion targeting domain of focal adhesion kinase contains a hinge region that modulates tyrosine 926 phosphorylation. Structure *88*, 881–891.

Qu, C., Liljas, L., Opalka, N., Brugidou, C., Yeager, M., Beachy, R.N., Fauquet, C.M., Johnson, J.E., and Lin, T. (2000). 3D domain swapping modulates the stability of members of an icosahedral virus group. Struct. Fold. Des. *8*, 1095–1103.

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

Rousseau, F., Schymkowitz, J.W., and Itzhaki, L.S. (2003). The unfolding story of three-dimensional domain swapping. Structure (Camb.) *11*, 243–251.

Saint-Jean, A.P., Phillips, K.R., Creighton, D.J., and Stone, M.J. (1998). Active monomeric and dimeric forms of *Pseudomonas putida* glyoxalase I: evidence for 3D domain swapping. Biochemistry *37*, 10345–10353.

Sanders, A., Craven, C.J., Higgins, L.D., Giannini, S., Conroy, M.J., Hounslow, A.M., Waltho, J.P., and Staniforth, R.A. (2004). Cystatin forms a tetramer through structural rearrangement of domain-swapped dimers prior to amyloidogenesis. J. Mol. Biol. *336*, 165–178.

Schaller, M.D. (2001). Biochemical signals and biological responses elicited by the focal adhesion kinase. Biochim. Biophys. Acta *1540*, 1–21.

Schlaepfer, D.D., Hauck, C.R., and Sieg, D.J. (1999). Signaling through focal adhesion kinase. Prog. Biophys. Mol. Biol. *71*, 435–478.

Schlunegger, M.P., Bennett, M.J., and Eisenberg, D. (1997). Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly. Adv. Protein Chem. *50*, 61–122.

Schymkowitz, J.W., Rousseau, F., Wilkinson, H.R., Friedler, A., and Itzhaki, L.S. (2001). Observation of signal transduction in three-dimensional domain swapping. Nat. Struct. Biol. *8*, 888–892.

Shen, Y., and Schaller, M.D. (1999). Focal adhesion targeting: the critical determinant of FAK regulation and substrate phosphorylation. Mol. Biol. Cell *10*, 2507–2518.

Spinelli, S., Frenken, L.G., Hermans, P., Verrips, T., Brown, K., Tegoni, M., and Cambillau, C. (2000). Camelid heavy-chain variable domains provide efficient combining sites to haptens. Biochemistry *39*, 1217–1222.

Spinelli, S., Desmyter, A., Frenken, L., Verrips, T., Tegoni, M., and Cambillau, C. (2004). Domain swapping of a llama VHH domain builds a crystal-wide β-sheet structure. FEBS Lett. *564*, 35–40.

Staniforth, R.A., Giannini, S., Higgins, L.D., Conroy, M.J., Hounslow, A.M., Jerala, R., Craven, C.J., and Waltho, J.P. (2001). Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. EMBO J. *20*, 4774–4781.

Ultsch, M.H., Wiesmann, C., Simmons, L.C., Henrich, J., Yang, M., Reilly, D., Bass, S.H., and de Vos, A.M. (1999). Crystal structures of the neurotrophin-binding domain of TrkA, TrkB and TrkC. J. Mol. Biol. *290*, 149–159.

Yang, F., Bewley, C.A., Louis, J.M., Gustafson, K.R., Boyd, M.R., Gronenborn, A.M., Clore, G.M., and Wlodawer, A. (1999). Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. J. Mol. Biol. *288*, 403–412.

Yang, S., Cho, S.S., Levy, Y., Cheung, M.S., Levine, H., Wolynes, P.G., and Onuchic, J.N. (2004). From the cover: domain swapping is a consequence of minimal frustration. Proc. Natl. Acad. Sci. USA *101*, 13786–13791.

Zegers, I., Deswarte, J., and Wyns, L. (1999). Trimeric domain-swapped barnase. Proc. Natl. Acad. Sci. USA *96*, 818–822.