Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

3.8 Protein and Nucleic Acid Folding: Domain Swapping in Proteins

L Liu and AM Gronenborn, University of Pittsburgh, Pittsburgh, PA, USA

Abbreviations

3-D	three-dimensional
BS-RNase	bovine seminal ribonuclease
CA	capsid
CTD	C-terminal domain
cryo-EM	cryo-electron microscopy
CV-N	cyanovirin-N
DHasA	HasA dimer
DT	diphtheria toxin
Eps8	kinase substrate 8
Glx I	glyoxalase I
GNM	Gaussian network models
gp	glycoproteins
GRFT	Griffithsin
HIV	human immunodeficiency virus
hRNase	human pancreatic ribonuclease
MA	matrix
MD	molecular dynamics
NC	nucleocapsid
NMR	nuclear magnetic resonance
NTD	N-terminal domain
PDB	Protein Data Bank
PrP^C	prion protein
RNase A	ribonuclease A
RYMV	rice yellow mottle virus
trpR	trp repressor

Glossary

Closed/primary interface For domain-swapped proteins, the intermolecular interfaces in the oligomer that possess identical intramolecular counterparts in the monomer form are called the 'closed' or primary interface.

Domain swapping Domain swapping is a special type of oligomerization, during which two or more homo-polypeptide chains exchange identical units.

Hinge loop The hinge loop is the only region of the protein that adopts a different conformation in monomeric and domain-swapped structures.

Morpheeins Special cases of multimers are the so-called morpheeins, which are homo-oligomeric proteins that can switch their structure between functionally distinct alternate quaternary states.

Open/secondary interface For domain-swapped proteins, the newly created contact surfaces in the oligomer constitute the 'open' or secondary interface.

 doi:10.1016/B978-0-12-374920-8.00309-X

3.8.1 Introduction

It is generally accepted as a central truth in biochemistry that the amino acid sequence of a protein encodes all necessary information for the chain in a given environment to fold into a single, well-defined stable structure.[1] For most proteins, this structure is under physiological conditions, the native, functional state. In certain circumstances, however, proteins may be able to fold into distinctly different structures, and during the past few years, increasing numbers of alternative folds have been discovered. Lymphotactin[2] and Mad2 (the mitotic arrest deficiency 2 protein)[3] are extreme examples of this type.

The most common alternative structures comprise different multimeric assemblies of identical polypeptide chains. Multimers are endowed with structural and functional advantages, such as improved stability and control over the accessibility and specificity of active sites, explaining why oligomerization is favored during protein evolution.[4] Special cases of multimers are the so-called morpheeins, which are homo-oligomeric proteins that can switch their structure between functionally distinct alternate quaternary states. The prototypical example of a morpheein is the enzyme porphobilinogen synthase, which exists in an equilibrium between an octamer, a hexamer, and two dimer conformations.[5] Another special case of oligomerization has been described as 'three-dimensional (3-D) domain swapping'.[6] A 'domain-swapped' structure contains two or more polypeptide chains that exchange identical units. The exchanged portion may consist of a single secondary structure element or an entire globular domain. If exchange is reciprocal between two monomers, dimers are formed, and if more chains are involved, oligomers ensue.

Folding into the native state is driven by a combination of entropic and enthalpic forces that result in burial of hydrophobic residues in the interior and exposure of polar residues on the surface of the protein. This network of defined attractive and repulsive forces arranges the chain in well-defined, secondary structure elements. In multimers, each single polypeptide chain usually adopts the same conformation, although the assembly of individual chains in the oligomer can vary. Often, small changes in protein composition or environment can tip the balance from one arrangement to the next, with some proteins coexisting in more than one oligomeric state. A classic example of alternate oligomers is the Bence-Jones protein, characterized by X-ray diffraction more than 40 years ago. This protein exists in the crystal in three quaternary structures[7] that vary in their domain interactions.

Currently, the Protein Data Bank (PDB)[8] contains 26 939 homo-oligomeric protein structures. The most commonly found assembly patterns are side-by-side and 'head-to-tail', but domain-swapped structures are becoming a sizeable fraction. In this chapter, not all oligomeric structures are considered; we are concerned with proteins for which domain swapping has been described.

The term 3-D domain swapping, or simply domain swapping, was originally coined by Eisenberg and colleagues to describe the X-ray structure of a diphtheria toxin (DT) dimer in 1994.[9] However, already in 1962, a report was published describing the exchange of an N-terminal fragment for bovine pancreatic ribonuclease A (RNase A) upon dimerization.[10] The first protein X-ray structures that contained domain-swapped elements were determined in the early 1980s,[11–13] with increasingly more structures of domain-swapped multimers following suit (**Figure 1**).

This chapter introduces terms and features concerning domain swapping; reports on domain-swapped structures that are available in the PDB; summarizes ideas about putative mechanisms for this type of oligomerization; and describes a few examples in detail, for some of which domain swapping may be important for regulating function or triggering disease.

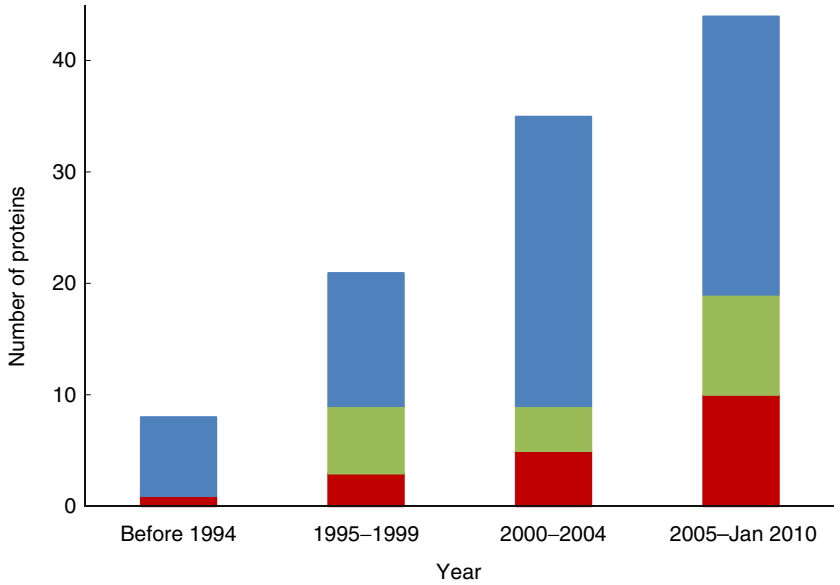

Figure 1 Growth in domain-swapped structures deposited in the PDB. Proteins with identical sequences for monomeric and oligomeric states are shown in red, proteins that share greater than 90% sequence identity between the monomer and oligomer are shown in green, and proteins for which swapped structures have been described without monomeric counterparts are shown in blue.

3.8.2 General Aspects

3.8.2.1 Terms and Definitions

True domain-swapped structures require that both monomeric and oligomeric states must be observed for a single protein.[6] However, this stringent designation is not always adhered to in the literature. Sometimes, structures are called domain-swapped even if no structure of the closed monomer has ever been observed or where only a homolog exhibits a closed monomer. In the first case, the protein is a 'candidate' for domain swapping, whereas in the second, the oligomers are classified as 'quasi-domain-swapped.'

In true domain-swapped structures, the exchanged subunit or domain in the oligomer is identical to the one in the corresponding monomer, exhibiting no differences in phi, psi backbone angles, except for the region that links the exchanging domains. This region is called the 'hinge loop' and often adopts an extended conformation in the domain-swapped oligomer while it folds back on itself in the monomer. Although called 'domain swapping,' the term 'domain' encompasses a variety of structural units: The largest may be an independently folded domain, whereas the smallest can be single secondary structure elements, such as a single β strand or an isolated α helix. The intermolecular interfaces in the oligomer that possess identical intramolecular counterparts in the monomer form are called the 'closed' or primary interface, whereas newly created contact surfaces constitute the 'open' or secondary interface. A schematic representation of different domain-swapping scenarios as well as the delineation of the different structural interfaces is provided in **Figure 2**.

This chapter considers mainly those proteins that contain swapped elements in their multimeric forms and for which a monomeric structure is seen for a mutant or close relative.

3.8.2.2 Data Set of Domain-Swapped Proteins

Currently, more than 100 domain-swapped structures are deposited in the PDB, with 38 examples for which both monomeric and oligomeric structures are available (**Tables 1** and **2**). These 38 proteins are nonrelated and exhibit less than 20% pairwise sequence identity. Among them, 19 cases exist with identical sequences for monomeric and oligomeric states; thus, they are examples of true domain swapping. The other 19 share greater than 90% sequence identity between the monomer and oligomer polypeptide, some involving single amino acid changes. Not surprisingly, structures for most domain-swapped oligomeric proteins have been determined by X-ray crystallography.

Analysis of the chain lengths, structural classification, or amino acid composition does not reveal any special properties associated with domain-swapped proteins. In the authors' data set, the shortest protein is the immunoglobulin binding domain B1 of streptococcal protein G (GB1),[55] which comprises only 56 residues, and the longest one is DT with 535 amino acids.[44] The ratio of all α proteins, all β proteins, and mixed α/β proteins for domain-swapped proteins is 2:2:5,

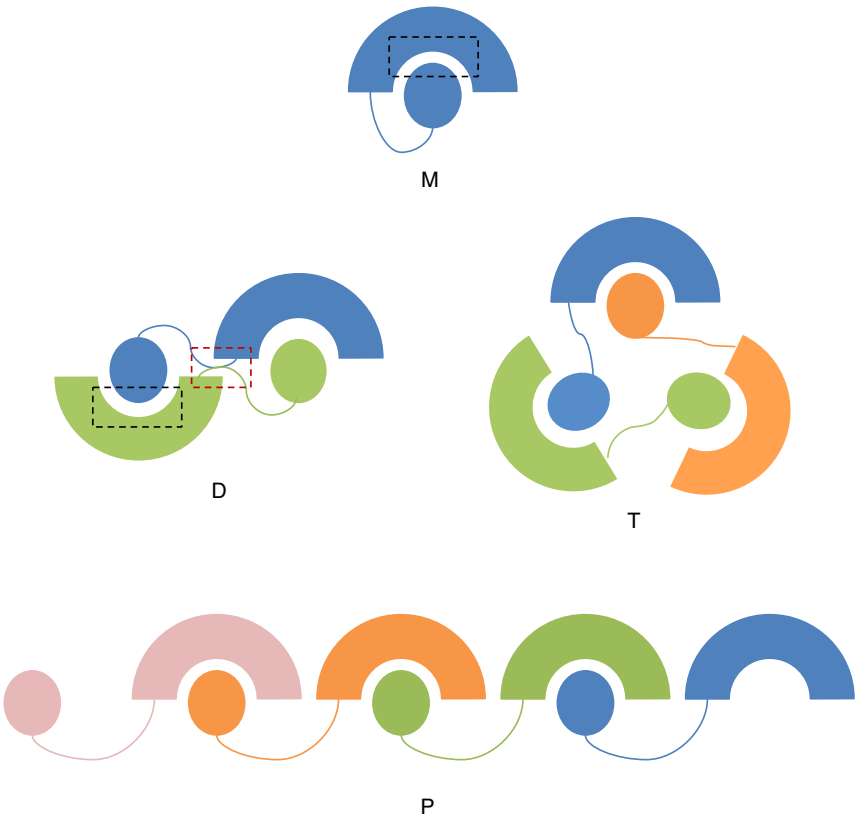

Figure 2 Schematic representation of domain-swapped structures and their pertinent features. D, dimer; M, monomer; P, daisy chain-type multimer; T, trimer. Closed and open interfaces are enclosed in black and red stippled rectangles, respectively.

Table 1 Proteins for which monomeric and swapped oligomeric structures are available for the identical polypeptide sequence

Protein	*PDB ID monomer*[a]	*PDB ID oligomer*	*Polypeptide length*[b]	*Hinge location*[c]	*Exchanged element(s)*	*References*
Syntaxin TLG1	2C5K	2C5J	95	65–69	Helix	14
VAMP-7	2VX8	2VX8	169	40–45[d]	Helix	15
spo0A	1QMP	1DZ3	130	107	Helix	16,17
Barnase	1BRN[e]	1YVS	110	37–41	Helices	18,19
FOXP2	2A07	2A07	93	538,544	Helices	20
Bcl2-L-1	1R2D	2B48	218	158–159	Helices	21,22
trpR	1P6Z[d,e]	1MI7	107	64–67,76–78	Helices	23,24
CD47	2JJS	2VSC	127	101–102	β strand	25
DAP-150	2HKQ	2HKN	97	37–40	β strand	26
LB1	1K50	1K50	63	52–56	β strand	27
cspB	1C9O	2HAX	66	37	β strands	28,29
CV-N	2EZM	3EZM	101	50–54	β strands	30,31
ATIII	1ATH	2ZNH	432	338–339,390–406[d]	β strands	32,33
RNase A						34–36
N-swap	5RSA	1A2W	124	19–20	Helix	
C-swap		1F0V		112	β strand	
ASP1	3BFB	3CYZ	119	13	β strand	37,38
yopH	1M0V	1K46	136	28–29	Mixed	39,40
Cystatin-A	1DVC	1N9J	98	48–50	Mixed	41,42
ptsH	1Y51	1Y50	88	54	Mixed	43
DT	1MDT	1DDT	535	379–386	Domain	9,44

[a]Some structures are not available as isolated monomers.
[b]Sequence information was obtained from the FASTA file in the PDB. Coordinate information may not be available for all residues in the PDB file.
[c]Hinge residues are numbered according to the monomer PDB file; these numbers may differ between monomer and dimer.
[d]The protein contains a cleaved peptide bond in the hinge region or has no coordinate information in the PDB file.
[e]No monomeric structure is available. The comparison is carried out for the monomer unit in a non-swapped dimer or oligomer.

Table 2 Proteins for which monomeric and swapped oligomeric structures are available for closely related polypetide sequences[a]

Protein	*PDB ID monomer*[b]	*PDB ID oligomer*	*Polypeptide length*[c]	*Mutation; extension*[d]	*Hinge location*[e]	*Exchanged element(s)*	*References*
TRX	2O7K	3DIE	107	1; 1	27–30	Mixed	45,46
CABP	1N65	1HT9	75	1; 1	42–45	Helices	47,48
CD2	1T6W	1CDC	99	3; 0	45–46	β strands	49,50
Rab27b	2ZET	2IF0	203	0; 3	43, 77	β strands	51,52
GRB2	1BM2	1FYR	117	1; 2	121–122	Mixed	53,54
GB1	1GB1	1Q10	56	4; 0	38–41	β strands	55,56
OBP	2HLV	1OBP	160	4; 0	121–122	Mixed	57,58
PrP^C	2W9E	1I4M	113	0; 5	190–197	Helix	59,60
HasA	1YBJ	2CN4	178	0; 5	48–50	Mixed	61,62
iNOS	1M8D[f]	1QOM	434	0; 6	104	Mixed	63,64
TNase	1SNC	1SND	149	6; 0	112–120[g]	Helix	65,66
GR	3BQD	3H52	255	7; 0	547–552	Mixed	67,68
Trk-A	1WWW	1WWA	101	0; 8	297	β strand	69,70
IL-10	1LK3	1ILK	160	6; 3	107–114	Helices	71,72
HDGF	1RI0	2NLU	110	0; 10	34–41	β strands	73,74
CA-CTD	2KOD[f]	2ONT	70	2; 12	177	Helix	75,76
EMMPRIN	3B5H[f]	3I84	184	>15	93–94	β-strand	77,78
RGS7	2D9J	2A72	139	>15	100	Helix	79,80
afaD	2IXQ	2AXW	142	>15	116–130	β strand	81,82

[a]The monomeric and swapped oligomeric structures for each pair are in the same entry in UniProt.[83]
[b]Some structures are not available as isolated monomers.
[c]Sequence information was obtained from the FASTA file in the PDB. Coordinate information may not be available for all residues in the PDB file.
[d]Sequence information was obtained from the FASTA file in the PDB. The polypeptide lengths in the pairs are different. Some such cases, such as HasA, are indeed bona fide examples of domain swapping.
[e]Hinge residues are numbered according to the monomer PDB file; these numbers may differ between monomer and dimer.
[f]No monomeric structure is available. The comparison is carried out for the monomer unit in a non-swapped dimer or oligomer.
[g]The protein contains a cleaved peptide bond in the hinge region or has no coordinate information in the PDB file.

which is identical to the ratio reported for all structures in Structural Classification of Proteins,[84] and there appears to exist no specific amino acid requirements for domain-swapped proteins, compared to overall protein space.

Similar findings hold when examining only the exchanged domains. They exhibit different sizes, ranging from a few residues to more than 100 amino acids. Single α helix or β strand can be swapped, bundles of α helices or β hairpins are found exchanged, and even mixed α helix and β strand elements can serve as the swapped domain, without any discernable sequence signature among them.[6] Although the exchanging unit can be located anywhere in the sequence, it is often found at one of the two termini. Human antithrombin III is an example in which the exchanged domain resides in the middle of the protein; this kind of exchange has also been termed 'hairpin insertion'.[32] An example in which almost one-half of the entire polypeptide chain is exchanged is cyanovirin-N (CV-N).[30]

The previous analysis reveals that proteins found in domain-swapped structures display the same diversity as any protein in the PDB. This suggests that almost any protein may be capable of undergoing domain swapping and that domain swapping is solely a specialized form of oligomer assembly.

3.8.2.3 Mechanistic Considerations

Comparison between the closed conformation of the monomeric polypeptide chain and the open conformation of the same chain in the domain-swapped dimer implies that the observed large conformational differences most likely require some kind of un/refolding. Intramolecular interactions involving hydrophobic contacts, hydrogen-bonding, electrostatic interactions, and even disulfide bridge interactions[60,85,86] at the closed interface in the monomer are exchanged to intermolecular interactions. Naturally, such breaking and reforming of contacts requires energy – the activation energy for 3-D domain swapping.[87] To overcome the activation barrier between the monomer and the dimer, changes in environment, particularly conditions that favor unfolding, may play a role.

For proteins capable of domain swapping, folding from the unfolded polypeptide chain can lead, in principle, to either the closed monomer or the domain-swapped dimer. Partitioning between the two products is determined by their free energy difference. This difference is naturally very small, given that all interactions within the two structures are extremely similar; only the hinge loop conformation is distinct. Therefore, any free energy difference needs to be traced to the hinge loop, which can either introduce or relieve strain during monomer-dimer interconversion.

3.8.2.3.1 The hinge loop

The hinge loop is the only region of the protein that adopts a different conformation in monomeric and domain-swapped structures. Therefore, sequences and secondary structures have received considerable attention in the search for local signals that could cause or influence domain swapping.

Several studies show that altering the length of the hinge loop can switch the domain swapping propensity of a protein. Intuitively, one would expect that long loops preferentially result in monomers and short ones in dimer structures: A short loop will make it difficult for the polypeptide to fold back on itself and in turn allow the swapped portion of the chain to find partners more easily. This is clearly the case in staphylococcal nuclease.[65] The only sequence difference between the monomer and the domain-swapped dimer is the loop length, with the monomer loop containing six more residues than the hinge in the dimer. Loop residue deletion has also been used in some designed proteins. An elegant example illustrating the importance of loop length is provided by two different three-helix bundles that were engineered in the Eisenberg laboratory.[88] Loop deletion in one of these caused the formation of a domain-swapped dimer, whereas loop deletion in the other resulted in fibril formation. On the other hand, Perutz and colleagues found that adding a stretch of polyglutamines into the active site loop of chymotrypsin inhibitor 2 caused domain swapping and higher order oligomer formation.[89] Indeed, in this case, oligomerization increased with increasing loop lengths. Therefore, a universal statement regarding the influence of hinge loop length cannot currently be made.

Not every amino acid in the hinge loop region has to change conformation. Sometimes, the alternative conformation is observed for only one or two residues. These could be the key hinge amino acids, and only their backbone phi and psi angles may have to change between monomer and dimer conformations. In the authors' data set, alanine and glycine are the most frequent amino acids in these key hinge positions, with their occurrence being much higher than commonly found. Glycine can adopt phi and psi angles in all four quadrants of the Ramachandran plot due to the lack of a side chain; therefore, it is possible to accommodate a glycine in any kind of turn, even quite sharp ones, which are sterically forbidden for other residues. For the cold shock protein cspB,[28] a flip in the backbone of G37 ($\Delta\varphi \approx 180^\circ$) is observed between monomer and domain-swapped dimer. Similarly, the small alanine residue is also more tolerant in terms of steric effects, and in the N-terminal swapped dimer of RNase A, only two adjacent alanines change their conformation compared to the monomer structure.

In the middle of hinge loop sequences, one also finds conserved prolines.[90] Because proline residues are thought to impart rigidity to the polypeptide backbone, Rousseau and colleagues suggested for the cyclin-dependent kinase regulatory subunit suc1[91] that the proline-caused strain in the hinge loop influences domain swapping. Indeed, replacement of the first proline in the hinge with an alanine stabilized the monomer form, whereas the same substitution of the second proline stabilized the dimer form. Rousseau et al. suggested that tension in the hinge loop in the monomer caused it to behave like a loaded molecular spring that is released when the alternative conformation is adopted in the dimer.[91] Unlike in suc1, mutation of the single proline in the hinge loop of CV-N to glycine substantially stabilized both states of the protein, with greater stabilization of the monomer compared to the dimer.[92] Furthermore, adding a second proline residue by mutating a neighboring amino acid caused the domain-swapped dimer to become the thermodynamically most stable state.[92] Similarly, the change of alanine in the hinge loop of the forkhead box P2 (FOXP2) to proline prevented

the formation of the swapped dimer.[20] This suggests that the addition or deletion of prolines creates no uniform outcome and that each protein may have its unique signature of hinge loop residues.

In addition to glycine, alanine, and proline, other amino acids in the hinge loops may also play a role in stabilizing particular secondary structure elements in the swapped domains. For example, a hinge loop could be a coil in the monomer form but become embedded into a long β strand or an α helix. This could stabilize the dimeric forms of these proteins, given the higher degree of secondary structure and the elimination of a flexible hinge region.

For a region in the protein to function as a hinge loop, it needs to be pliable enough to adopt different conformations. RNase illustrates this point. RNase A,[34] bovine seminal ribonuclease (BS-RNase),[93] and a human pancreatic ribonuclease (hRNase) chimera[94] share greater than 60% sequence identity, and all three proteins undergo domain swapping of their N-terminal helices, albeit with different relative orientations of the helix and different conformations in the three hinge loops. RNase A is also one of the rare examples that can swap either N- or C-terminal parts, with C-terminal strand exchange resulting in a domain-swapped dimer[35] or cyclic swapped trimer.[95]

Overall, the combined results obtained for hinge loop properties provide useful hints with respect to domain swapping. However, no clear, predictive rules have emerged.

3.8.2.3.2 Mutations promoting domain swapping outside of the hinge loop

There are several examples in which residue changes in other parts of the protein besides the hinge loop are associated with domain swapping. A prime example is GB1. Compared to wild-type monomeric GB1, the domain-swapped dimer comprises four mutations – L5V, F30V, Y33F, and A34F – none of which are located in the hinge region.[56] A theoretical analysis of the quadruple mutant and wild-type GB1 from Wodak's group[96] suggested different effects caused by each change: L5V introduces general destabilization due to unfavorable interactions with its surrounding residues; F30V induces local strain due to a clash with its own backbone; and A34F not only destabilizes the monomer conformation by forcing W43 to adopt a strained side chain conformation, and therefore disrupts the hydrophobic core of GB1, but also stabilizes the swapped dimer by tightly packing its side chains from both subunits against each other in the dimer core. The importance of the individual mutated residues (L5V/F30V/Y33F/A34F) in the integrity of the domain-swapped structure was also investigated by modeling and mutagenesis.[56] Inspection of the dimer structure suggested that the shorter mutant side chains of the L5V and F30V variants could easily be accommodated within the core, although possibly causing some destabilization of the structure. Indeed, each change is tolerated within the wild-type structure. The Y33F mutation represents a conservative change, and either side chain can substitute for the other in the respective cores. The position of F34 in the domain-swapped dimer appeared to be most crucial. This was verified experimentally: Reverting F34 in the amino acid sequence of the domain-swapped dimer mutant back to the wild-type alanine residue resulted in a monomeric protein with a very similar structure as that of wild-type GB1.[56]

In the T cell surface antigen CD2, the propensity for dimer formation could be modulated by mutations in the new interface that is created by domain swapping.[97] In addition, a R87A mutation that destabilizes the monomer simultaneously increased dimer formation. However, as with the majority of other proteins, the hinge residues in CD2 were still the most crucial amino acids with respect to domain swapping.[97]

In summary, residues distant from the hinge region can shift the relative stabilities of monomer and domain-swapped dimer and thereby modulate domain swapping properties. However, compared to the amino acids in the hinge loop region, they appear to play only a secondary role.

3.8.2.3.3 Stability and folding of the monomer

Despite substantial efforts, no compelling proposal for a generally applicable and unified molecular mechanism of domain swapping has emerged to date.[6,98–101]

Eisenberg and colleagues suggested a free energy diagram involving pathways for domain swapping based on their studies of DT (**Figure 3**).[87] In their scenario, the 'open monomer' conformation retains the native fold of other parts of the 'closed monomer,' and only interactions at the closed interface are disrupted during unfolding of the monomer. Such a partial unfolding scheme may be at play in multidomain proteins in which separate, independently folding domains are exchanged. However, the existence of a stable open monomer is unlikely for most domain-swapped proteins in which only a few secondary structural elements are exchanged. These isolated structural elements will be unstable, and therefore complete un/refolding is more likely to be at play in these cases.

In RNase A, more than one portion of the chain can exchange, creating different oligomers (**Figure 4**). Two different domain-swapped dimers and two domain-swapped trimers are formed in different relative proportions.[95] Among the two dimers, the C-terminal swapped dimer is the major form, suggesting that it is more stable. For the trimers, only the crystal structure of the cyclic C-terminal swapped form has been solved. Biochemical studies suggested that the second, uncharacterized trimer may be a linear trimer in which one RNase A molecule swaps its N-terminal helix with a neighboring RNase A molecule at one end and its C-terminal strand at the other end.[95] In this kind of trimer, both types of exchange occur simultaneously at very distant sites in the same protein molecule, supporting the notion that the closed monomers may fully unfold and refold to form these various forms of domain-swapped oligomers.

In the cyclin-dependent kinase regulatory subunit Cks1, exchange of the last β strand, β_4, is involved in dimer formation.[102] Nuclear magnetic resonance (NMR) studies indicate that β_4 in free monomeric Cks1 exhibits conformational heterogeneity.[103] This motion is abrogated by binding of Cdk2 to Cks1, resulting in a more homogeneous conformation of Cks1. Because Cdk2 binds to one face of the Cks1 β sheet, the flexibility of β_4 is reduced, preventing domain

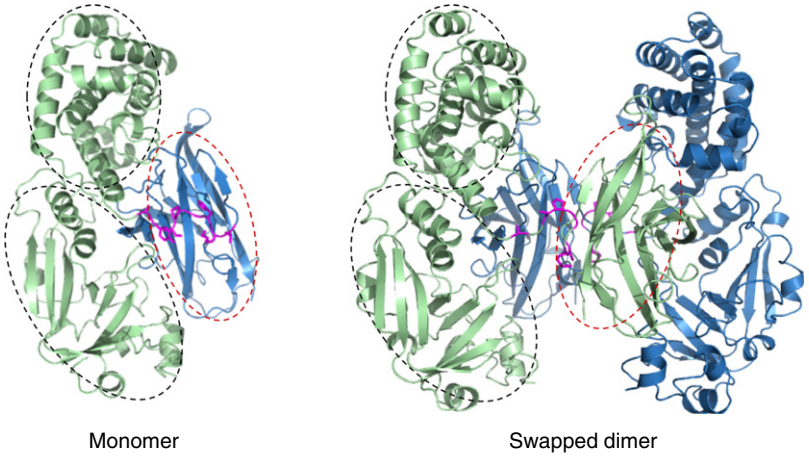

Figure 3 Structures of DT. In the monomer, the swapped elements are shown in blue. In the dimer, different polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and in magenta. Independently folding domains are encircled in black for non-swapped domains and in red for swapped domains.

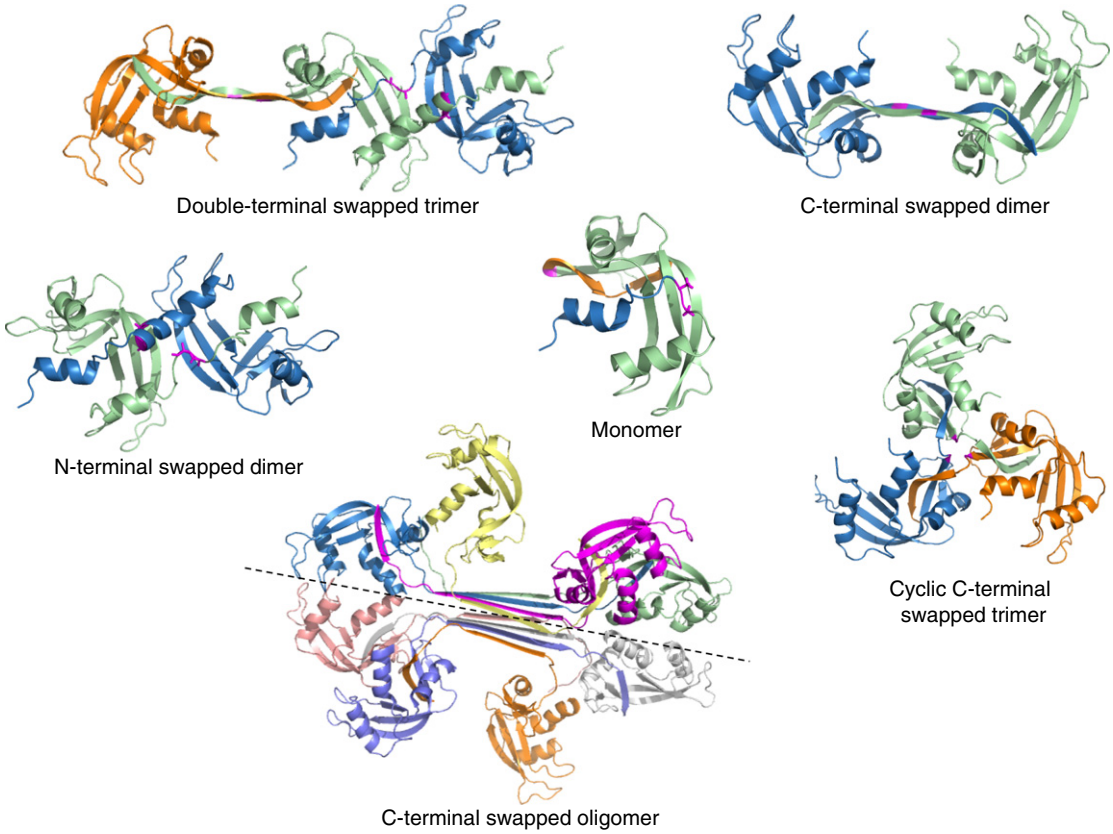

Figure 4 Structures of RNase A. In the monomer, the two secondary structure elements involved in exchange are shown in blue and orange. In the dimers and trimers, the individual polypeptide chains are shown in green, blue, and orange, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta.

swapping. Interestingly, the binding of Cdk2 increases the binding affinity of Cks1 for phosphopeptides that bind to the other face of the β sheet.[103] Therefore, configurational entropy influences not only ligand binding of Cks1 but also domain swapping.

3.8.2.4 Theoretical and Computational Explorations

A number of computational approaches for deciphering the basic events in protein folding and assembly are available that use reduced models and detailed atomistic simulations.

Several groups are applying these methodologies to domain swapping. Movement of the polypeptide chain by Brownian motion through a funneled energy landscape with structure formation dominated by native stability[104] is the most elegant and widely accepted protein folding concept. This concept has also been applied to protein associations in domain-swapped multimers. In particular, Onuchic et al.[105] have used a symmetrized Go-type potential to simulate domain swapping in molecular dynamics (MD) simulations. For the epidermal growth factor receptor kinase substrate 8 (Eps8) SH3 dimer, they discovered a frustrated hinge region and suggested the following most favorable path for domain swapping: native monomers → partially folded monomers → unfolded monomers → open-end domain-swapped dimers → domain-swapped dimers. Onuchic et al. suggested that the overall monomeric topology, rather than local signals in the hinge region, determines where in the polypeptide chain domain swapping will occur.[105] Although plausible, it appears at odds with some experimental results. For instance, in GB1 and LB1 (protein L B1 domain), proteins with identical monomeric topologies, different domain-swapped dimers are observed, clearly at odds with expectations if topology plays the dominant role. Proteins with intrinsic symmetry of the sequence and/or structure are 'highly frustrated' in the language of these authors, and in their simulations multimode domain swapping was observed and necessitated the inclusion of inter- or intramolecular disulfide bonds.[106] Two proteins that fall into the 'highly frustrated' category are the human prion protein (PrP^C) and CV-N. However, at least for CV-N, the presence of disulfide bonds is not necessary for domain swapping because several homologs of CV-N with varying numbers of disulfide bonds appear to lack domain swapping,[107,108] and no differences in disulfides were noted for the monomers or domain-swapped dimers.

Coarse-grained MD simulations for several known domain-swapped proteins were also performed by Ding et al.,[109] who found that starting from monomeric conformations, domain-swapped dimers sometimes formed. Based on native contact changes and topology maps, a web server for predicting the hinge region of domain-swapped proteins was created.[109] Testing the predictive value with the current set of 38 proteins resulted in correct predictions for only approximately one-third of the proteins in this set.

Analyzing large-scale domain motions of DT via Gaussian network models (GNM), Kundu and Jernigan[110] uncovered the major hinge in this protein based on the observed slower modes in GNM. The direction of the motion of the swapped domain about the hinge was predicted using the anisotropic network model.[110] However, it appears that DT is a special case among the domain-swapped proteins, given its multiple domain structure and the fact that a true folded domain undergoes the exchange and not single secondary structural elements.

We performed GNM analysis on the monomeric conformations of all 38 domain-swapped proteins in order to uncover any motions that may induce domain swapping. Initially, the domain-swapped structures were not used and were simply employed as controls in this analysis. For each protein, hinge residues were defined by comparing backbone dihedral angles for the experimentally determined monomer and dimer structures (dihedral angle changes $> 60^\circ$ at the open interface). The motional behavior for all residues via the first slow modes from GNM were examined. GNM did not successfully distinguish hinge residues for our diverse set of domain-swapped proteins. Investigating the behavior of every residue, we found that the hinge residues are neither the most mobile nor the most rigid ones in some proteins. For that matter, taking the picture of a hinge literally, the actual hinge usually stays fixed, with the two objects that are connected by the hinge changing their relative positions. This would translate to relative rigidity of hinge residues and mobility at the edge of the hinge. On the other hand, hinge residues are often located in loops that are naturally more mobile than the cores of proteins, thereby allowing conformational changes to occur more easily.

A quite different mechanism of domain swapping that involves a progressive and reversible transformation between monomer and dimer has been proposed by Wodak's group.[111] This process starts from either end of the polypeptide chain, and intramolecular contacts are traded for equivalent intermolecular ones, with the total number of native contacts remaining essentially constant. In this manner, increasingly more of the monomer chains are substituted for each other until a stable state is reached. Exchange initiated at one end, such as the C terminus, and did not involve unfolding. Conformational changes within the individual monomers and the binding between them were tightly coupled, and the total number of native contacts was maximized. In this process, a large number of hinge conformations and association modes are sampled by the intermediates, suggesting that the exchange reaction is nonspecific and the amino acid sequence plays only a minor role. However, to date, there is no experimental evidence for such a mechanism, and it remains highly speculative.

3.8.3 Instructive Examples and Biological Implications

Is domain swapping an *in vitro* curiosity or does it serve a biological function? A number of results suggest that this type of oligomerization could be exploited in biology. One possible role for domain swapping could be to regulate protein function by modulating the populations of active molecules or the availability of functional sites. In addition, domain swapping could play a role in the allosteric regulation and signal transduction. Furthermore, in protein oligomerization scenarios, possible cytotoxic aggregation could be inhibited by domain-swapped dimerization. Finally, domain swapping is an efficient means for supramolecular structural organization of oligomers, such as seen in viral capsid structures. Therefore, although domain swapping may be involved in misfolding, aggregation, and amyloid formation of many proteins,[99,112] this may not be the only function it serves.

Next, several notable examples of domain-swapped proteins are discussed in more detail. These are not stringent examples as defined previously, and for the associated proteins, a stably folded monomeric structure may not be available.

3.8.3.1 RNase A

RNase A is the classic example of a protein engaged in domain swapping. Dimerization involving exchange of the N terminus

was proposed in 1962 prior to any structural information by Crestfield, Stein, and Moore to explain its behavior under acidic conditions.[10] The first X-ray structure for a domain-swapped RNase A dimer was solved in the late 1990s by Eisenberg,[34] and the Eisenberg laboratory subsequently identified more domain-swapped dimers, trimers, and multimers (**Figure 4**).[35,95] Because of its versatility, RNase A is frequently portrayed as the prototypical domain-swapped protein, and with its different oligomeric states, it well illustrates the remarkable options of domain swapping modes.

Different folding conditions result in different types of RNase A oligomerization. Dimers are found at pH 6.5 and 37 °C, close to the physiological conditions. However, the dissociation constant for the dimer under these conditions is about 2 mM, approximately 20-fold greater than the concentration of RNase A in the bovine pancreas. Polyethylene glycol 10 000 stabilizes the RNase A minor trimer under crystallization conditions at pH 3.5.[95] Interestingly, RNase A oligomers exhibit higher enzyme activity on double-strand RNA than the monomer,[113] and this is easily explained by the spatial arrangement of amino acids from different subunits that create the active site. Indeed, catalytic histidines are contributed by the N-terminal α helix and the C-terminal β strand, respectively.[114]

In one of the trimer forms of RNase A, both N- and C-terminal units are exchanged, resulting in a linear arrangement.[95] In the other trimer that only exhibits swapping of the C-terminal strand, a cyclic structure is formed. Therefore, for proteins that can swap two different domains, a variety of assembled oligomeric structures can be formed, and models for such trimers, tetramers, and other oligomers have been proposed for RNase A.[115]

Although wild-type RNase A does not form fibrils, a variant with a polyglutamine insertion in its hinge loop (RNase A Q_{10}) forms amyloids *in vitro*.[116] A model for the RNase A Q_{10} fibrils was proposed in which the Q_{10}-containing hinge loop residues form β strands that arrange into two β sheets. The individual domains in this model keep their native fold and are involved in 'runaway' domain swapping.[116] In addition to the linear-type arrangements, simultaneous exchange of two different domains allows the formation of branched aggregates, possibly explaining the observation of some nonfibrillar aggregates.

3.8.3.2 B1 Domain

GB1 is a small, 56-residue, stable, single domain protein. It comprises a four-stranded β sheet with a single α helix packed on top of it.[55] This protein exhibits astounding structural variability. A number of surprising structural variants were obtained in a large mutagenesis study involving a library of randomized hydrophobic core residues. Among the alternative structures was a domain-swapped dimer in which one hairpin was exchanged between the subunits.[56] The dimeric structure comprises an eight-stranded β sheet made from four adjacent hairpins, resulting in two extensive new interfaces (**Figure 5**). The two α helices are antiparallel and cross at their C termini. Half of the dimer, composed of the first β hairpin and the α helix from one polypeptide chain and the second β hairpin from the other chain, is essentially identical to the monomer structure. The dimer dissociates into partially folded,

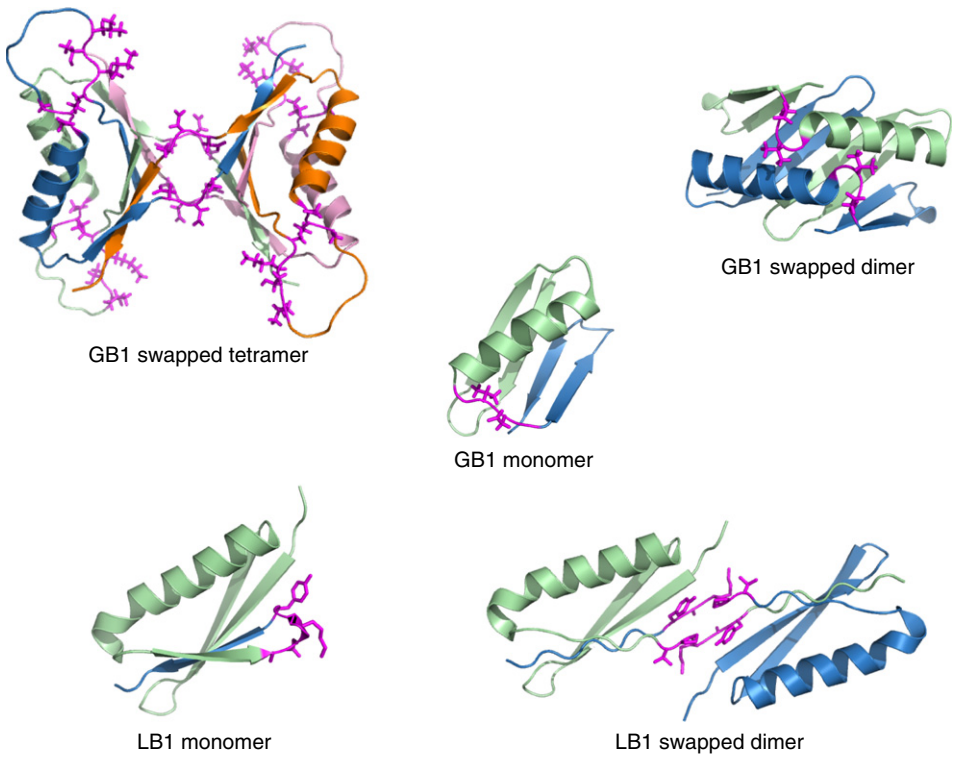

Figure 5 Structures of B1 domains. In monomers, exchanged elements are shown in blue. In the dimers, individual polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta.

monomeric species at low micromolar protein concentrations. The monomer is not a native, stable structure but, rather, a partially folded protein with extensive motions on the micro- to millisecond timescale. Despite these conformational fluctuations, the overall architecture of the monomer resembles that of wild-type GB1. Thus, for this variant, dimerization via domain swapping stabilizes the molten, monomeric hydrophobic core.[117]

Structural comparison between the domain-swapped dimer and the wild-type monomer suggested that the F34 side chain was the pivot for the monomer-dimer switch. Indeed, changing this residue back to the wild-type alanine resulted in a wild-type-like monomer structure. Interestingly, changing A34 to phenylalanine in the wild-type sequence did not induce domain swapping but resulted in a side-by-side dimer.[118]

GB1 variants are also capable of fibril formation, especially those sequences that are prone to domain swapping. Mutants that fold into the stable, wild-type GB1 structure or variants that exist as a highly destabilized, fluctuating ensemble of random, folded, and partially folded structures under the same experimental conditions do not easily fibrilize. A left-handed helical ribbon model for the fibril was built, based on experimental disulfide cross-linking results, containing the swapped dimer structure as the smallest unit.[119]

An additional amino acid change in the domain-swapped dimer core caused a further dramatic change in structure: A symmetric tetramer ensued with intermolecular strand exchange involving all four units.[120] Three β strands and the α helix were retained in the tetramer, although their intra- and intermolecular interactions were radically different, with strand β_2 of the first hairpin missing. The β_3-β_4 hairpin was changed to a side-by-side arrangement of strands β_3 and β_4 from one subunit, running antiparallel to β_3 and β_4 of another one. This topological change was accompanied by a shift in register. In addition to strand exchange of the domain swapping kind, a new interface between surface elements of the individual chains was formed.

LB1 exhibits the same fold as the GB1 monomer;[121] however, a quite different domain-swapped structure was found for its mutants (**Figure 5**). Substitution of a glycine by alanine in the turn of the second β hairpin caused exchange of the C-terminal β strand between the subunits, with the wild-type hairpin straightening and creating the intermolecular β-sheet interface. These long β strands are kinked, causing both B1 units to be rotated around the hinge region. Exchange of valine to alanine in the hydrophobic core also resulted in this type of domain-swapped structure.[27] Interestingly, in the X-ray structure, the asymmetric unit contains two wild-type-like monomers and a domain-swapped dimer. Novel intermolecular hydrophobic contacts as well as intermolecular hydrogen bonds between the exchanged β strands contribute to the stability of the domain swap.[27]

The previously described different oligomeric B1 structures are illuminating examples for structural evolutionary paths from monomers to multimers.

3.8.3.3 Llama VHH

Certain immunoglobulin isotypes of old (camels and dromedaries) or new (llamas and vicuna) world camelids comprise a single domain antibody, referred to as VHH, instead of the common four-domain Fab fragment. This VHH domain is generated via DNA recombination between dedicated VHH germline gene segments and D and J minigenes. All VH and VHH domains posses two cysteines that form a conserved disulfide bridge. The llama VHH sequence contains changes in four conserved amino acids, and its CDR1 (complementarity determining region 1) is two residues longer (nine amino acids) than the typical length of most VHHs. Its CDR2 contains seven residues, as seen in many camelid VHHs; however, its CDR3 is only six residues long, which is shorter than the average size of camel or llama CDR3s.

The X-ray structure of a llama VHH domain (VHH-R9) revealed the protein as a domain-swapped dimer (**Figure 6**).[123] The first seven residues in the first β strand are missing, and CDR3 and the last strand exhibit domain swapping. An antiparallel β structure connects the two symmetry-related molecules, creating a dumbbell shape. Thus, one incomplete VHH monomer together with the last β strand from the other monomer forms the classical VHH structure. Because the VHH fold is conserved and no open interface is created in the domain-swapped VHH dimer, it is believed that CDR3 strain release and favorable interactions along the extended hinge may account for dimer stabilization.

The crystal contains an extended β sheet throughout the lattice, formed by trigonal structures that are arranged around the cubic threefold axis from N- and C-terminal segments of the chain. In this extended β-sheet structure, six VHH molecules are linked by a dense and large network of interstrand hydrogen bonds. In this manner, higher order oligomeric structures are stabilized. These extensive polymeric contacts are only possible because of the truncated N terminus.

3.8.3.4 Lectins

Several lectin structures were found to exhibit domain-swapped multimers. The first example was CV-N, originally isolated from an aqueous extract of the cyanobacterium *Nostoc ellipsosporum*. It inactivates human immunodeficiency virus (HIV) and other enveloped viruses. The original solution structure of the 101-amino acid protein was a monomeric one,[30] whereas in the subsequently determined X-ray structures, domain-swapped dimers were found (**Figure 7**).[31,124] Either form predominates, depending on experimental conditions, and can be isolated for biophysical, structural, and functional studies.[92] The monomer structure exhibits a compact, bilobal fold with pseudo-symmetry. Interestingly, the amino acid sequence repeats of CV-N do not constitute the structural repeats. Rather, the two symmetrically related domains are formed by portions from both sequence repeats. Each domain comprises a triple-stranded β sheet with a β hairpin packed on top, connected by a helical linker. This linker becomes the hinge in the domain-swapped dimer. In the dimer structure, the two pseudo-monomer halves contain the same interactions as were seen in the monomer. Biophysical studies revealed that the domain-swapped dimer is a kinetically trapped folding intermediate that converts into the slightly more stable monomeric form at physiological ($>30\ ^\circ C$) temperatures. At low temperature (room temperature or lower),

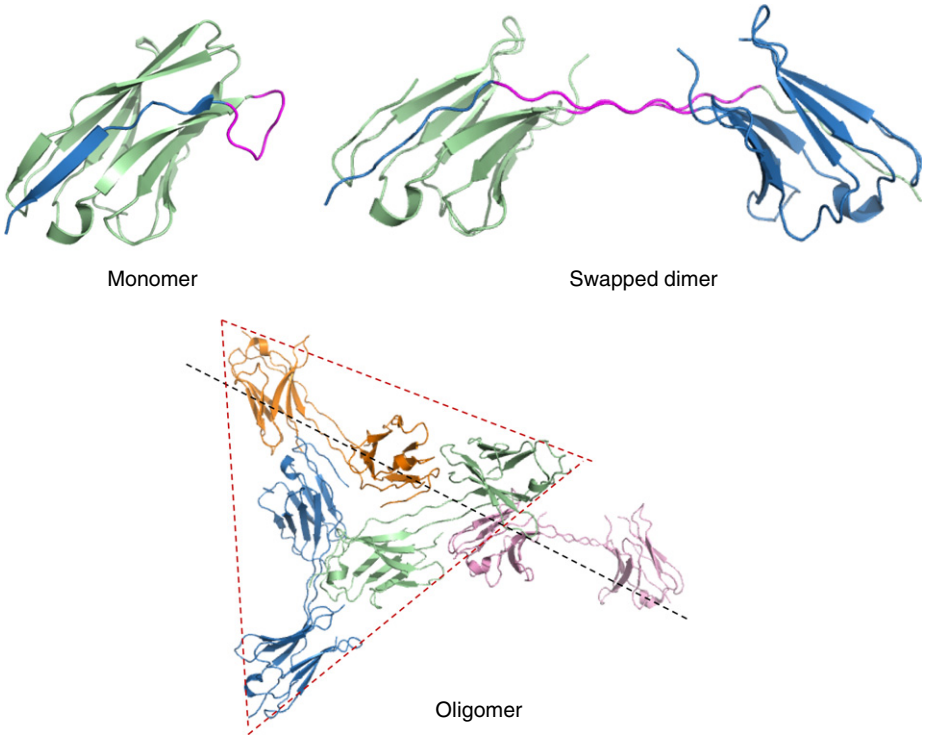

Figure 6 Structures of llama VHH. In the monomer, exchanged elements are shown in blue. In the dimer, individual polypeptide chains are shown in green and blue, respectively. The monomer structure is derived from a homolog of the swapped dimer that shares 74% sequence identity.[122] Hinge residues are shown in magenta.

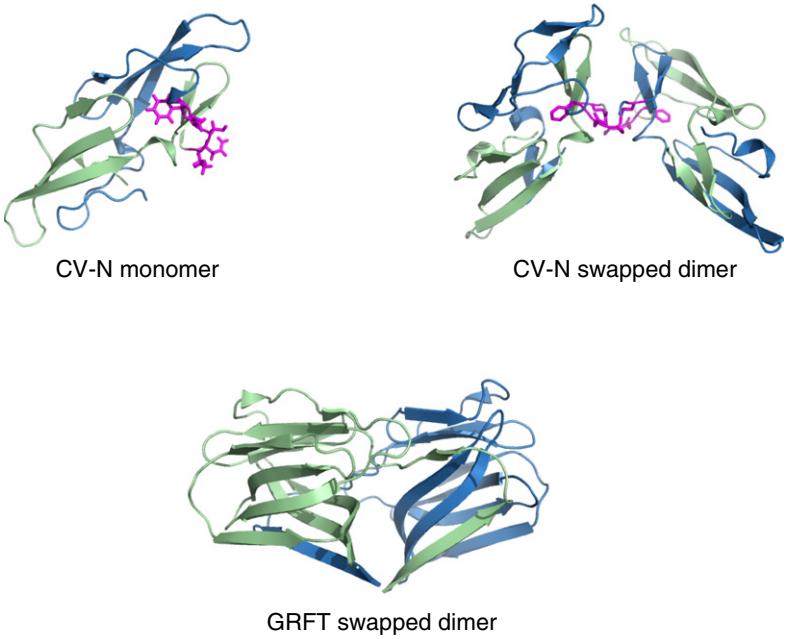

Figure 7 Structures of lectins. In the monomer, exchanged elements are shown in blue. In the dimers, individual polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta.

the dimer lifetime is sufficiently long for structural characterization in solution.[92] The solution dimer structure is virtually identical in interdomain packing and overall folding to the structures in the trigonal and tetragonal crystals;[31,124] only the relative domain-domain orientations of the pseudomonomeric halves are distinct. In the solution dimer, hinge residues exhibit motions on the microsecond timescale, suggesting the possibility of reorientation around this hinge

with different orientations being trapped in alternate crystal lattices.

The coexistence of both monomeric and domain-swapped dimeric CV-N in solution indicates that the free energies of folding for both quaternary states must be comparable and the kinetic barrier between the states has to be significant. Mutation of residues in the hinge region allows for modulation of the relative stabilities. For instance, substituting the single proline in the hinge region by glycine resulted in a greater than 5 kcal mol^{-1} stabilization of the monomeric P51G variant compared to wild type. The S52P mutant exists predominantly as a dimer due to destabilization of the monomer, and for the ΔQ50 deletion mutant, no folded monomer was found.[125] Interestingly, a mutant that was created to eliminate the sugar binding site in one of the domains, CVN^{mutDB}, resulted in a protein that is monomeric in both solution and the crystal.[107] For this variant, no domain swapping was observed under all conditions, although in the crystal structures, two monomers were found in the asymmetric unit. Intriguingly, the protein-protein interface between the two monomers in the crystal resembles the open interface in one of the domain-swapped dimer structures of wild-type CV-N.[107]

The second antiviral lectin that exhibited domain swapping was Griffithsin (GRFT).[126] GRFT is a 121-amino acid protein of the red alga *Griffithsia* sp. It exhibits antiviral activity against HIV-1 and severe acute respiratory syndrome virus by binding to various viral glycoproteins (gp), such as gp120, gp41, and gp160, in a monosaccharide-dependent manner.[126,127] The structure of GRFT closely resembles jacalin lectins and comprises three repeats of a four-stranded antiparallel β sheet. In the swapped dimer, the first two β strands of one chain complete the β prism of the other chain (**Figure 7**). To date, GRFT is the only example of a jacalin-fold protein for which a domain-swapped structure has been observed. GRFT is also the only member in its fold family that contains three carbohydrate binding sites. Other jacalins usually have a single one. The prism structure of GRFT is encoded by its triple sequence repeat. The three sugar binding sites reside in the loops of the β hairpins formed by the second and third strand of each β sheet.[126]

Another lectin, *Microcystis viridis* lectin, was also suggested to show a domain-swapped structure. However, because no monomeric structure is available, it is difficult to ascertain that indeed a domain swapping has occurred.[128]

Although CV-N and GRFT undergo domain swapping, the extent of the exchanged sequence is quite different. In CV-N, half of the molecule is involved in the swap, whereas in GRFT only the first 2 β strands out of 12 are swapped. In addition, for CV-N, both monomeric and dimeric structures have been extensively characterized, whereas for GRFT only the dimeric structure is available.

Regarding their anti-HIV activities, the previously discussed lectins interact with oligosaccharides on viral envelope glycoproteins. The GRFT dimer contains six sugar binding sites, whereas CV-N exhibits two (monomer) or four (dimer). Both proteins are highly potent and inhibit HIV-1 at nanomolar concentrations.[126,129] The binding sites on CV-N interact with the terminal epitopes (D1 and D3 arms) of the large, branched oligosaccharides. For GRFT, a similar binding mode has been proposed.[126,129]

3.8.3.5 Fab 2G12

Among all anti-HIV antibodies, very few neutralizing ones have been described. One of these is the human antibody 2G12, which neutralizes a broad range of HIV-1 isolates. It binds to a cluster of high-mannose sugars on the 'silent' face of the gp120 envelope glycoprotein. In the crystal structure of 2G12, two Fabs assembled into an interlocked dimer via domain swapping of their VH domains (**Figure 8**).[130] Biochemical, biophysical, and mutagenesis data revealed that the productive form that recognizes glycosylated gp120 is indeed the swapped dimer. In the crystal structure of this antibody, the arrangement of the combining sites creates an extended surface for multivalent interaction with the carbohydrates. Such dimeric assembly has not been observed in any of the hundreds of Fab structures in the PDB. The VH domain exchange in 2G12 is associated with a twist of the variable regions relative to the constant region, although the individual variable (V_H and V_L) and constant domains (C_{H1} and C_L) are very similar to those in other Fab molecules. The swapped dimer lacks the highly conserved 'ball-and-socket joint' between V_H and C_{H1} that allows the variable domains to adjust their position relative to the constant domains, even though the required residues are present. The V_H domains within the dimer are related by a noncrystallographic twofold axis ($\sim 180^\circ$), with the two Fabs arranged side by side. This conformation places the combining sites approximately 35 Å apart but facing in the same direction. Analysis of the Fab 2G12 structure suggests three causes for domain swapping in this case. First, the conserved interaction between adjacent glutamines in the V_H and V_L domains is missing at the closed V_H/V_L interface. These glutamines usually hydrogen bond to each other at the base of antibody combining sites. In 2G12, however, this position is occupied by a rarely observed arginine. Second, the elbow region connecting the V_H and C_{H1} domains that forms the hinge loop comprises an unusual sequence. Therefore, the V_H domain pivots around P113, stabilizing a hydrophobic contact between proline and valine that aids in domain swapping. Third, a favorable open

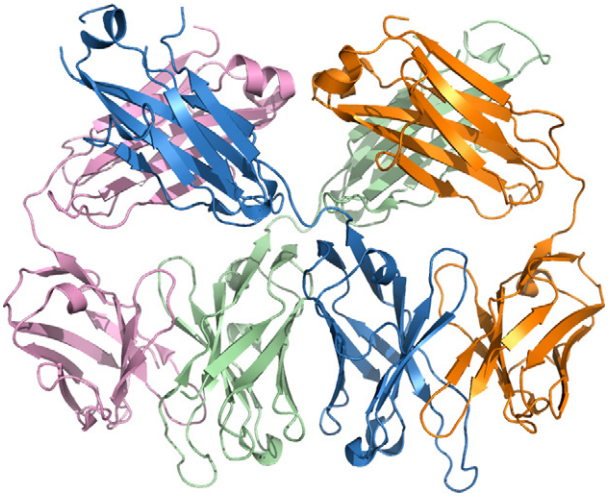

Figure 8 Structure of the 2G12 Fab. Heavy chains are shown in green and blue and light chains in pink and orange, respectively.

interface is created between V_H/V_H domains through an extensive hydrogen bonding and salt bridge network, comprising a total of 10 hydrogen bonds and 136 van der Waals interactions, as well as π-stacking interactions between several aromatic residues that create a substantial buried area.[130]

The neutralizing properties of the 2G12 antibody are intimately connected to domain exchange of the V_H domains in the two Fab regions. The extensive multivalent binding surface created from two conventional combining sites therefore is able to recognize high-mannose sugars. In the 2G12 V_H/V_H interface, many conserved germline-encoded amino acids are found, including three uncommon mutations that stabilize this interaction. Although in principle, any IgG molecule could potentially recognize two oligomannose chains in a bivalent manner, this would require an energetically unfavorable near parallel orientation of the two Fab arms. In contrast, the 2G12 domain-exchanged structure is well suited for recognizing two oligomannose chains that are 35 Å apart through a virtually continuous protein surface that matches the geometrical spacing of the carbohydrate cluster on gp120.

3.8.3.6 Viral Capsid Protein

The gag polyprotein is encoded by all retroviruses and directs the formation and release of immature viral particles. During virus budding, the gag precursor protein is cleaved into three products: matrix (MA), capsid (CA), and nucleocapsid (NC). Simultaneously with gag processing, termed maturation, the morphology of the viral particle changes dramatically. Different retroviruses exhibit morphologically distinct mature virions, characterized by the shape of their CA core structure. In HIV-1 and other lentiviruses, the mature capsid is conical, whereas it is spherical or polyhedral in HTLV-1 and MLV.

Limited high-resolution structural information is available for the arrangement of CA proteins within the shell of immature or mature retroviral particles. However, cryo-electron microscopy (cryo-EM) studies of *in vitro* assemblies and high-resolution studies of isolated CA fragments provide important clues. In general, the overall protein structure is highly conserved among retroviral CA proteins, comprising two independently folded domains – the N-terminal domain (NTD) and the C-terminal domain (CTD). Both domains of CA are predominantly α-helical, with the NTD comprising seven α helices and an amino-terminal β hairpin, whereas the CTD is composed of four short α helices and a single-turn 3_{10} helix. Capsid assembly involves dimerization of the CTD, followed by assembly into polymeric forms.

The nature of the CTD dimerization interface has been the subject of considerable interest and debate. In early crystallographic studies, face-to-face dimerization was observed for the CTD. In the crystal dimer, the protein packs through helices 9 and 9′ (notation for the entire CA structure), and changes in this interface can completely abolish CA dimerization in solution. Although mutations of residues in the crystal dimer interface have pronounced effects on viral assembly, they do not completely block it, and interface residues are not conserved in different retroviruses. Currently, several X-ray crystal structures and one NMR structure of the HIV-1 capsid CTD are available, and five possible arrangements of the CTD dimer have been observed.[76,131–134] All of these dimers exhibit different packing arrangements for helix 9 from each monomer, comprising distinct crossing angles. Interestingly, the dimer arrangement found in the solution NMR structure fits well into the cryo-EM density map of assembled capsid.[131] These observations indicate that conformational changes in the capsid protein are possible during assembly.

Surprisingly, a deletion variant of the HIV-1 CA-CTD resulted in a domain-swapped structure in the crystal (**Figure 9**).[75] This dimer, with an entirely new dimer interface, was found after the solution NMR structure of an evolutionary and structurally related SCAN domain had been determined.[135] The Δ177 CA-CTD deletion mutant exhibits essentially the same architecture as the SCAN dimer. The swapped unit comprises the N-terminal strand, a turn, and helix 1. The hinge for the domain swapping is located between helices 8 and 9, and the deleted A177 resided in this region. In addition to domain swapping, a second structural change occurs in helix 9, compared to the previously determined X-ray structures. In the original X-ray structures, helix 9 was kinked and resides at the crystallographic dimer interface, whereas in the domain-swapped dimer structure, helix 9 is a regular, straight helix.

In the full-length CA protein, the Δ177 mutation can assemble into core-like structures *in vitro*. However, there is little experimental evidence to support the domain-swapped model

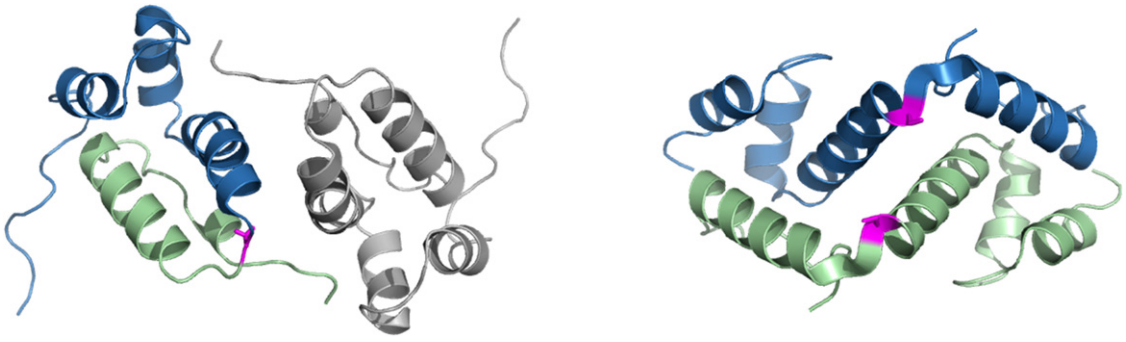

Figure 9 Structures of capsid CTD. In one of the chains in the dimer, elements that are exchanged in the domain-swapped counterpart are shown in green. In the swapped dimer, individual polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta. In the domain-swapped dimer, the location of the deletion is shown in magenta.

over the conventional side-by-side model of CA-CTD dimerization in the mature virions, especially given the excellent fit of the solution NMR CA-CTD structure into the electron density of the cryo-EM maps of assembled capsid tubes.[131] Thus, it is more likely that formation of the swapped dimer is predominantly caused by shortening of the hinge and that native capsid assembly involves a side-by-side dimer formation for the CTD.

In the rice yellow mottle virus (RYMV) from the *Sobemoviridae* family, the virus particle is a nearly spherical icosahedder, with small protrusions and depressions. The capsid consists of 180 copies of the 238-amino acid comprising coat protein, exhibiting $T=3$ quasi-equivalency. Each asymmetric unit of the capsid contains three subunits (A–C), and the subunit structure is similar to the common eight-stranded jelly-roll β-sandwich fold of most icosahedral viruses. Subunit A is arranged into pentamers, and alternating subunits B and C form hexamers. The N-terminal 49 residues are not seen in subunits A and B, but 23 of them (residues 27–49) are visible in subunit C. They form an additional β strand, βA, that is tucked between a neighboring B subunit and another C subunit. As a result, the exchange of these βA strands via domain swapping creates long-range interactions between the subunits and most likely increases the stability of the RYMV capsid through intermolecular contacts.[136]

3.8.3.7 Prion Protein

Mammalian prion proteins contain approximately 210 residues with one conserved disulfide bond. The monomer structure of the human PrP^C has been determined in solution by NMR, and the domain-swapped dimer structure was solved by crystallography (**Figure 10**).[60,137] The monomer is made of three α helices and a small two-stranded β sheet. The disulfide bond connects helices 2 and 3. In the domain-swapped dimer, the C-terminal helix is exchanged. Intriguingly, the nature of the disulfide bond between helices 2 and 3 is different in the monomer and dimer: In monomeric PrP^C, it is intramolecular, whereas in the domain-swapped dimer it is intermolecular. For this reason, conversion from monomer to dimer cannot occur without reduction and reoxidation of the disulfide.

For the prion protein, it is possible to seed the conversion of 'normal' recombinant protein into amyloid fibrils *in vitro* through a specialized redox process.[138] This conversion process is proposed to proceed by runaway domain swapping, in which the hinge loop and adjacent regions of successive subunits form a continuous β sheet in the center of the fibril. In this model, the domain-swapped PrP^C globular domains decorate the exterior of the β sheet, providing the characteristic templating features of prions. In the growing fibril, there are always two free helices with one cysteine each. They can potentially catalyze sulfhydryl-disulfide exchange on the fibril surface, opening a new PrP^C monomer and readying it for addition to the growing fibril. However, whether the prion protein ever encounters a reducing environment *in vivo* is uncertain, and it appears that the intramolecular disulfide bond is required for infectivity.

3.8.3.8 Cystatin

Cystatins belong to a large superfamily of cysteine protease inhibitors. They generally contain 100–120 amino acids, and their fold is composed of a five-stranded β sheet and an α helix that is packed orthogonally onto the β strands.[42] The substrate binding sites of cystatins are highly conserved and located in β hairpin loops that connect strands 2 and 3 (site I) and 4 and 5 (site II), respectively.[139] Similar stable domain-swapped dimers have been seen for several members of the cystatin family at physiological concentrations.[42,140,141] In the swapped dimer, the first two β strands and the helix are exchanged, generating two antiparallel β strands around the hinge region (**Figure 11**). In the monomer structure, this region comprises binding site I. As in most domain-swapped proteins, mutations in key positions promote or prevent domain swapping. For instance, in chicken cystatin, the I66Q core mutation increases the kinetics of dimerization from immeasurably slow to easily measurable, whereas other mutations at or near the hinge region (V55D or I102K) block dimerization.[42]

Cystatin can also form amyloids, with cystatin amyloidosis involved in a number of diseases.[142] Crystal structures of cystatin tetramers and octamers have been solved,[143] as well as a larger oligomeric form.[140] In the latter crystal, two domain-swapped

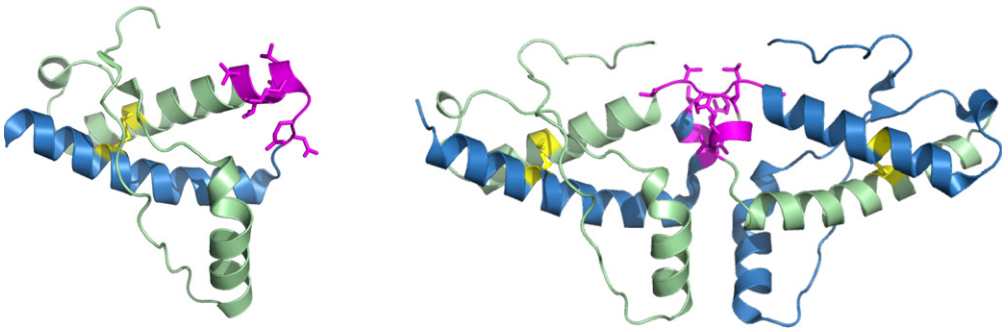

Figure 10 Structures of PrP^C. In the monomer, exchanged elements are shown in blue. In the dimer, individual polypeptide chains are shown in green and blue, respectively. Hinge residues and disulfide bonds are shown with their side chains in stick representation and colored magenta and yellow, respectively.

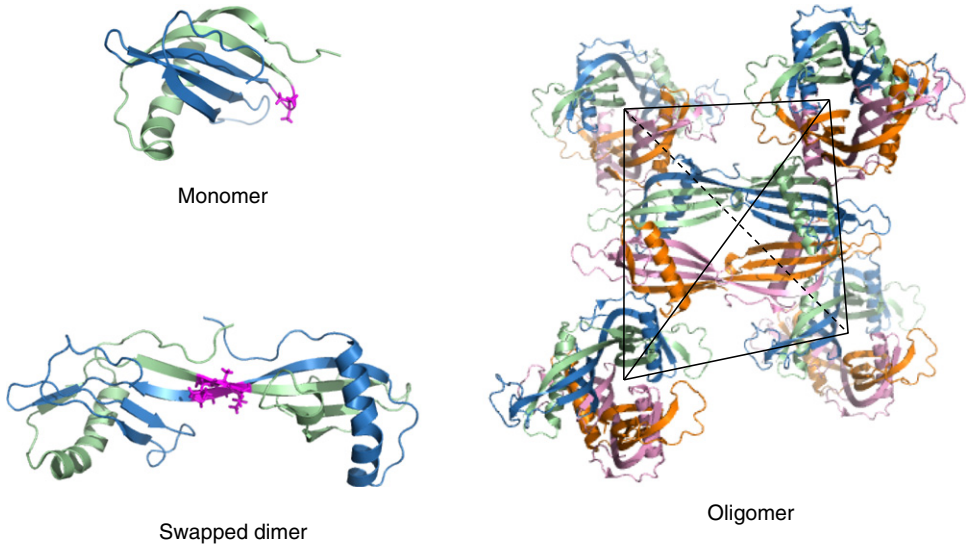

Figure 11 Structures of cystatin. In the monomer, exchanged elements are shown in blue. In the dimer, individual polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta.

dimers interact with each other and form a tetramer, with tetramer packing into higher order structures in which the β sheets of all dimers are perpendicular to a common axis. A similar model for fibrils involving the stacking of domain-swapped dimers was described previously for the GB1 amyloid fibrils.[119]

3.8.3.9 Glyoxalase I

Glyoxalase I (Glx I) catalyzes the reaction between methylglyoxal and glutathione, forming *S*-lactoylglutathione as the product. Human Glx I contains 183 amino acids, and in its monomeric form, it folds into two domains and an isolated N-terminal helix. Each domain is composed of a four-stranded β sheet flanked by two α helices. The active form of the enzyme is a dimer (**Figure 12**), with the N-terminal domain of one monomer and the C-terminal domain of the other monomer arranged side by side in an eight-stranded β sheet.[144] The active site is located at the dimer interface at the center of the β sheet, comprising residues from the two different polypeptide chains.

In the dimer, the N-terminal helix is the exchanged element and packs against the body of the protein formed by other polypeptides. Sequence alignment of Glx I family members reveals that this N-terminal helix is missing in several members, indicating that it may not be strictly necessary for creating the active enzyme. It may, however, contribute to the stability of the Glx I dimer. It has also been suggested that in the dimer, the C-terminal domain is involved in domain swapping.[144] As with other such cases, given that no closed monomer structure is available to date, this cannot be verified. In the crystallographic dimer, one polypeptide chain does not comprise a complete active site and therefore may be inactive. However, in *Pseudomonas putida* Glx I (55% sequence identity to human Glx I), both the monomer and the dimer are active.[145] In addition, *P. putida* Glx I exhibits a monomer-dimer equilibrium in solution, with a less active monomer and a more active dimer. This monomer-dimer equilibrium could be regulated by glutathione[145] and is an example of how ligand binding can modulate a protein's quaternary state and function.

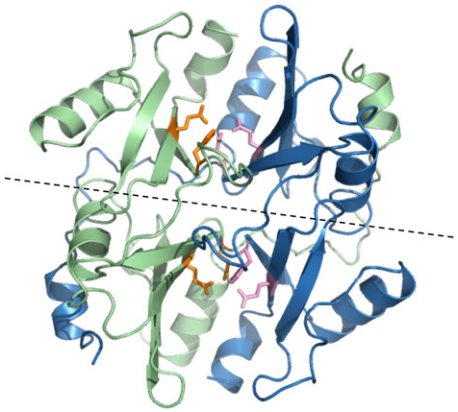

Figure 12 Structure of glyoxalase I. Individual polypeptide chains are shown in green and blue, respectively, and active-site residues from different chains are shown with their side chains in stick representation and colored orange and magenta, respectively.

3.8.3.10 Forkhead Domain of FOXP2

The protein forkhead box P2 (FOXP2) belongs to a newly defined subfamily, the P branch of the forkhead box transcription factors. Previous studies indicated that disease-causing mutations are located in the forkhead domains of FOXP proteins.[146–149] This domain is the signature feature of FOX family members and is responsible for DNA binding. The forkhead domain of FOXP2 contains 93 amino acids and exhibits two monomers and two domain-swapped dimers in the same crystal asymmetric unit, with two bound double-stranded DNA segments (**Figure 13**).[20] The monomeric form is composed mainly of three stacked α helices and a three-stranded antiparallel β sheet, which caps the structure at one

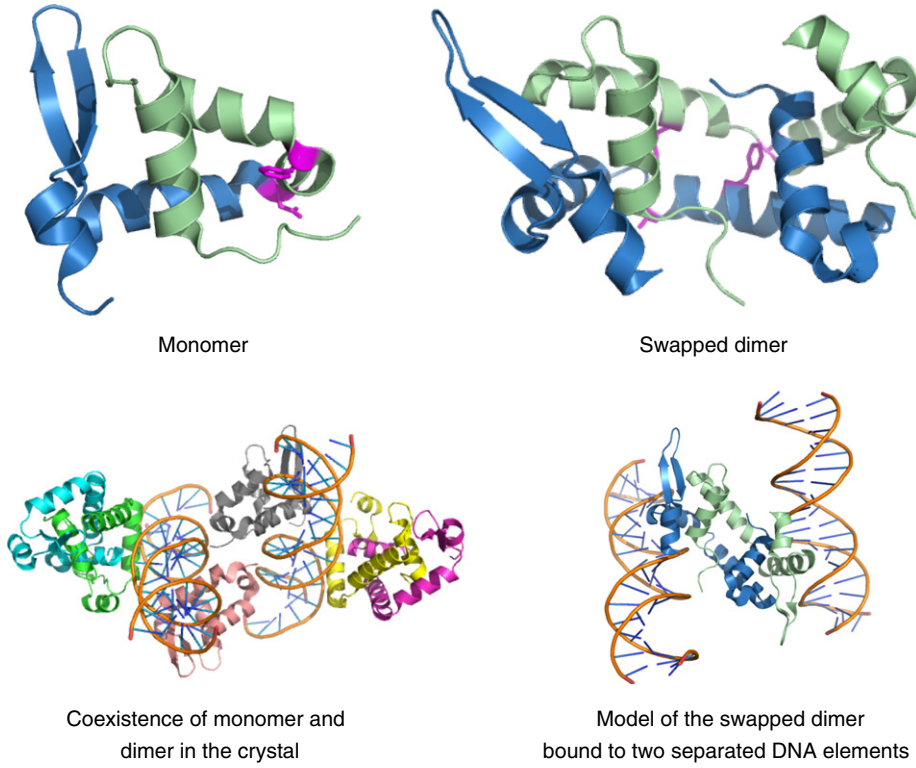

Figure 13 Structures of forkhead domain of FOXP2. In the monomer, exchanged elements are shown in blue. In the dimer, different polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta.

end. In the domain-swapped structure, the third helix and the second and third strands are exchanged. Interestingly, the third helix is the DNA recognition motif in both the monomeric and the domain-swapped dimeric structure.

The hinge region of the domain-swapped dimer contains an alanine that is conserved in the FOXP subfamily but replaced by a proline in other POX family members. Mutation of alanine to proline in the forkhead domain of FOXP2 abrogates the formation of a swapped dimer, indicating its pivotal role in the monomer-dimer switch. Moreover, some disease-causing mutations also map to the dimer interface. A model for the domain-swapped FOXP2 dimer bound to two separated DNA recognition elements suggests that for FOXP2, domain swapping may help in the assembly of high-order DNA-protein complexes.[20]

3.8.3.11 TrpR

The trp repressor (trpR) binds the operator region of the trp operon and prevents the initiation of transcription. Its smallest functional unit is a dimer,[150] although tetrameric and higher order species are also observed.[151,152] Under extreme conditions (30% isopropanol), an infinite crystalline 3-D supramolecular array was found,[24] possibly involving domain swapping.

Whether or not the trpR dimer constitutes a true domain-swapped case is unclear because a highly intertwined structure was observed and no monomeric homolog is available. Each single polypeptide chain in trpR dimer is composed of six α helices (A–F) in a relative closed conformation (**Figure 14(a)**).[23] In the crystalwide assembly, helices C–E rearrange and form a very long, single helix, resulting in polypeptide termini that are separated by a large distance (**Figure 14(b)**). Two such dimers come together and create a substructure (half of the tetramer) very similar to that seen in the dimer (**Figure 14(c)**). Therefore, the dimer-like structure is formed by segments of four different polypeptide chains, two providing intertwined N-terminal regions and two providing C-terminal regions,[24] with the polymer constituting a branched aggregate rather than a daisy chain-type linear one. Given that the polymeric form was crystallized in the presence of a significant amount of alcohol, its relevance to any physiological state is unclear. Indeed, it is well-known that alcohols increase the helical content of flexible peptides and/or destabilize tertiary structures significantly.[24,153,154]

3.8.3.12 Serpins

A number of human diseases are associated with 'serpinopathies,' which are deposition disorders of a large and diverse family of protease inhibitors, the serpin.[155] Serpins are relatively large proteins, with approximately 330–500 amino acids per chain. Their structures are composed of three β sheets and eight or nine α helices. During approximately the past 25 years, a number of studies have revealed that inhibitory serpins can undergo dramatic conformational changes. The reactive center loop (RCL) region of these proteins can insert itself as an additional β strand (β_4) into the central β sheet, either after peptide bond cleavage or spontaneously.

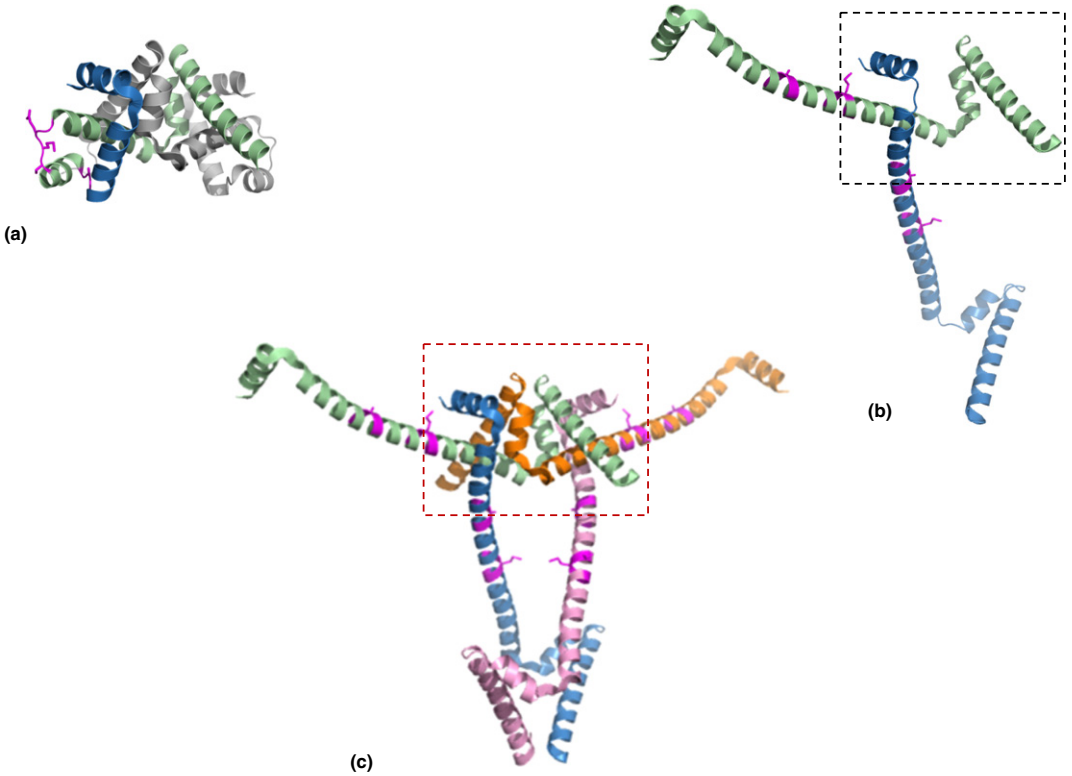

Figure 14 Structures of trpR. (a) trpR dimer. Each chain exhibits a relatively closed conformation, and exchanged elements are shown in blue. (b) The closed conformation is formed by two different polypeptide chains in the crystal lattice, shown in green and blue, respectively. (c) The trpR oligomer in the crystal lattice. Hinge residues are shown with their side chains in stick representation and colored magenta in all panels.

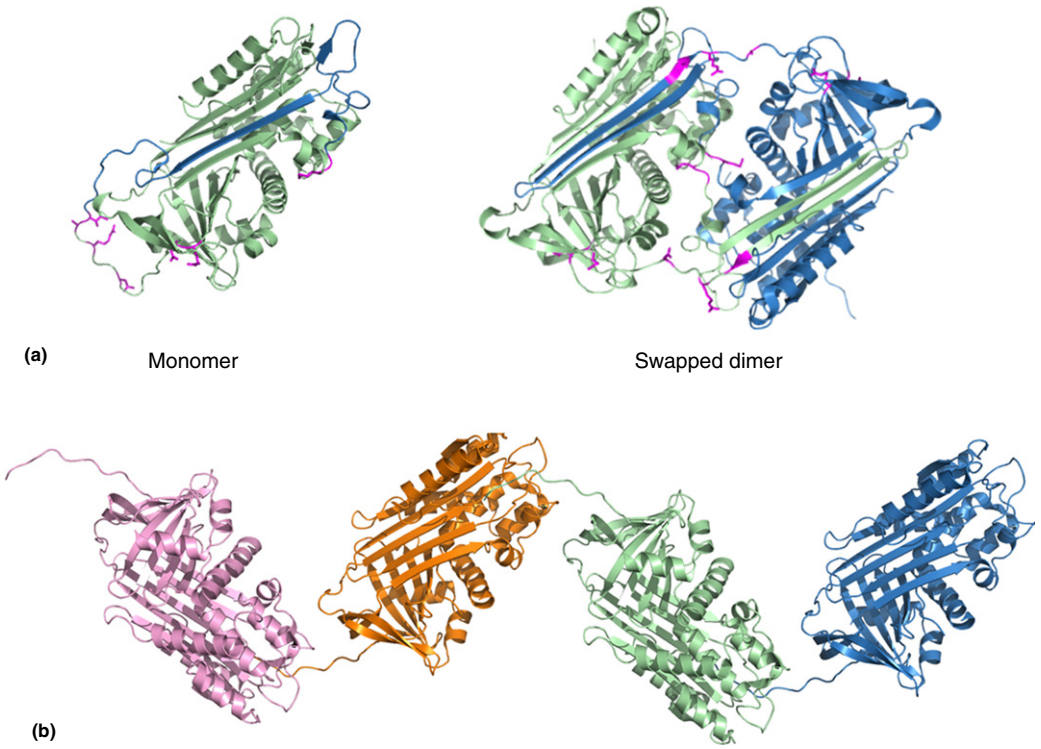

Figure 15 Structures of serpins. (a) Antithrombin. In the monomer, exchanged elements are shown in blue. In the dimer, individual polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta. (b) α_1-Antitrypsin daisy chain-type multimer.

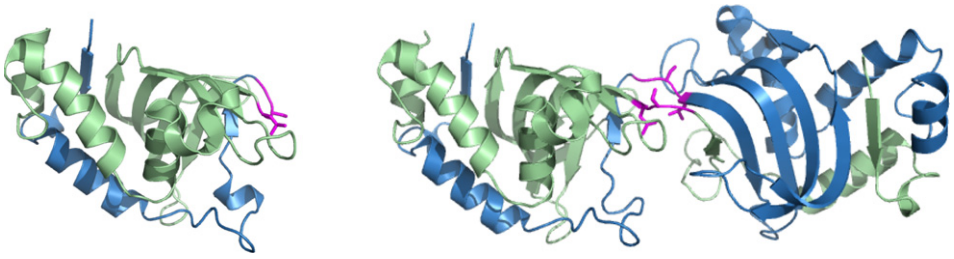

Figure 16 Structures of hemophore HasA. In the monomer, exchanged elements are shown in blue. In the dimer, individual polypeptide chains are shown in green and blue, respectively. Hinge residues are shown with their side chains in stick representation and colored magenta.

A mutant antithrombin shows a surprising twist to this story: In its domain-swapped dimer structure, one monomer donates a long β hairpin to another monomer, inserting it into the middle of the principal β sheet. In the polymer and the cleaved serpin, these two β strands constitute the RCL, forming the β_4 strand, as well as its neighboring β_5 strand (**Figure 15(a)**).[33] Based on this unusual structure, Yamasaki et al. proposed that such β hairpin insertion across several molecules may create a highly stable polymer.[33] In fact, the crystal structure of a linear polymer of α_1-antitrypsin, another human serpin, supports this notion (**Figure 15(b)**).[156] Such oligomerization of serpin polymers explains their rapid propagation and extreme stability. In addition, serpins constitute excellent examples for irreversible β sheet expansion in protein deposition diseases.

3.8.3.13 Hemophore HasA

The HasA protein is used by pathogenic bacteria to acquire the host's iron from either free heme or heme-containing hemoglobin. The secreted protein binds heme in a 1:1 stoichiometry and transfers it its receptor HasR, which in turn releases it into the bacterium.[157] *Serratia marcescens* HasA is a 19-kDa protein whose monomer structure is composed of seven β strands arranged into a central six-stranded antiparallel β sheet with a high degree of curvature. Four α helices are packed on one side against this sheet, whereas a loop connecting strands β_5 and β_6 covers its hydrophobic other side. Two small 3_{10} helices connect strand β_7 to helix α_2. The heme is held in a relatively solvent-exposed environment by two loops, one of which connects helix α_1 to strand β_2 and the other connects strands β_4 and β_5 at the interface of the predominately α and β regions of the molecule.[158] In addition to the monomeric variant, a HasA dimer (DHasA) is also secreted *in vivo*, and its crystal structure revealed a domain-swapped dimer. The pseudo-monomeric portion of the structure is very similar to the isolated monomer structure (**Figure 16**). However, it appears to be less flexible, and the region around the heme binding pocket is almost inaccessible in the dimer.[61] DHasA is not capable of transferring the heme to HasR because the appropriate productive HasA-HasR complex cannot be formed. Indeed, the regions involved in the interaction between HasA and HasR[159] become partially inaccessible by dimerization in DHasA. However, the heme can be transferred from DHasA to a HasA monomer.[61] Therefore, DHasA was proposed to constitute a heme reservoir, keeping it available for use by the Has system.[61]

3.8.4 Conclusions

During the past several decades, increasingly more domain-swapped protein structures have become available, and at least in some cases, there is evidence in support of the dimer or multimer constituting biologically important species. Indeed, irrespective of whether domain swapping is a specific mechanism for regulating function *in vivo*, it is becoming clear that it is not solely an *in vitro* artifact.

Despite considerable efforts by numerous groups, no unifying molecular mechanism of domain swapping has emerged: Each protein seemingly behaves in a distinctive and individual manner, and a general explanation for how proteins exchange domains remains elusive. What seems to emerge is that domain swapping is closely associated with the unfolding/folding process of proteins. For some proteins, distinct intermediates – in which some hydrophobic part of the monomeric protein becomes exposed and, thereby, is available for interaction with a 'like' molecule – may play a role, whereas for others, complete unfolding may occur. The fact that high protein concentration and additives (always present during crystallization) promote domain swapping suggests a switch in solute-solvent interaction. For example, exposed hydrophobic regions may no longer undergo unfavorable interactions with the aqueous solvent but, rather, favorable ones with another polypeptide chain. In this manner, an oligomeric structure can be trapped in either a crystal or an aggregate. Such behavior may also occur *in vivo* under conditions in which monomer promoting factors are missing or high local protein concentrations are induced through compartmentalization or the action of protein-protein interaction modules.

A more thorough understanding of the underlying features associated with domain swapping is certainly desirable. On the one hand, domain swapping seems to be a means by which stable multimers can be generated under evolutionary pressure, and it provides ways to improve protein stability. On the other hand, the fact that increasingly more proteins that exhibit disease-related aggregation can also form domain-swapped structures suggests a possible involvement in protein deposition diseases. Therefore, it may be possible to suppress

aggregation by modulating domain swapping – an unexplored avenue in drug discovery.

Acknowledgments

Work in our laboratory is supported by the National Institutes of Health (GM082251 and GM080642).

References

[1] Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230.
[2] Tuinstra, R. L.; Peterson, F. C.; Kutlesa, S.; Elgin, E. S.; Kron, M. A.; Volkman, B. F. Interconversion between two unrelated protein folds in the lymphotactin native state. *Proc. Nat. Acad. Sci. USA* **2008**, *105*, 5057–5062.
[3] Luo, X.; Tang, Z.; Xia, G.; Wassmann, K.; Matsumoto, T.; Rizo, J.; Yu, H. The Mad2 spindle checkpoint protein has two distinct natively folded states. *Nature Struct. Mol. Biol.* **2004**, *11*, 338–345.
[4] Marianayagam, N. J.; Sunde, M.; Matthews, J. M. The power of two: Protein dimerization in biology. *Trends Biochem. Sci.* **2004**, *29*, 618–625.
[5] Lawrence, S. H.; Ramirez, U. D.; Tang, L.; Fazliyez, F.; Kundrat, L.; Markham, G. D.; Jaffe, E. K. Shape shifting leads to small-molecule allosteric drug discovery. *Chem. Biol.* **2008**, *15*, 586–596.
[6] Liu, Y.; Eisenberg, D. 3D domain swapping: As domains continue to swap. *Protein Sci.* **2002**, *11*, 1285–1299.
[7] Huang, D. B.; Ainsworth, C. F.; Stevens, F. J.; Schiffer, M. Three quaternary structures for a single protein. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7017–7021.
[8] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucl. Acids Res.* **2000**, *28*, 235–242.
[9] Bennett, M. J.; Choe, S.; Eisenberg, D. Refined structure of dimeric diphtheria toxin at 2.0 Å resolution. *Protein Sci.* **1994**, *3*, 1444–1463.
[10] Crestfield, A. M.; Stein, W. H.; Moore, S. On the aggregation of bovine pancreatic ribonuclease. *Arch. Biochem. Biophys.* **1962**, *1*, 217–222.
[11] Remington, S.; Wiegand, G.; Huber, R. Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. *J. Mol. Biol.* **1982**, *158*, 111–152.
[12] Anderson, W. F.; Ohlendorf, D. H.; Takeda, Y.; Matthews, B. W. Structure of the cro repressor from bacteriophage lambda and its interaction with DNA. *Nature* **1981**, *290*, 754–758.
[13] Fita, I.; Rossmann, M. G. The NADPH binding site on beef liver catalase. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1604–1608.
[14] Fridmann-Sirkis, Y.; Kent, H. M.; Lewis, M. J.; Evans, P. R.; Pelham, H. R. Structural analysis of the interaction between the SNARE Tlg1 and Vps51. *Traffic* **2006**, *7*, 182–190.
[15] Pryor, P. R.; Jackson, L.; Gray, S. R.; Edeling, M. A.; Thompson, A.; Sanderson, C. M.; Evans, P. R.; Owen, D. J.; Luzio, J. P. Molecular basis for the sorting of the SNARE VAMP7 into endocytic clathrin-coated vesicles by the ArfGAP Hrb. *Cell* **2008**, *134*, 817–827.
[16] Lewis, R. J.; Brannigan, J. A.; Muchova, K.; Barak, I.; Wilkinson, A. J. Phosphorylated aspartate in the structure of a response regulator protein. *J. Mol. Biol.* **1999**, *294*, 9–15.
[17] Lewis, R. J.; Muchova, K.; Brannigan, J. A.; Barak, I.; Leonard, G.; Wilkinson, A. J. Domain swapping in the sporulation response regulator Spo0A. *J. Mol. Biol.* **2000**, *297*, 757–770.
[18] Buckle, A. M.; Fersht, A. R. Subsite binding in an RNase: Structure of a barnase-tetranucleotide complex at 1.76-Å resolution. *Biochemistry* **1994**, *33*, 1644–1653.
[19] Zegers, I.; Deswarte, J.; Wyns, L. Trimeric domain-swapped barnase. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 818–822.
[20] Stroud, J. C.; Wu, Y.; Bates, D. L.; Han, A.; Nowick, K.; Paabo, S.; Tong, H.; Chen, L. Structure of the forkhead domain of FOXP2 bound to DNA. *Structure* **2006**, *14*, 159–166.
[21] Manion, M. K.; O'Neill, J. W.; Giedt, C. D.; Kim, K. M.; Zhang, K. Y.; Hockenbery, D. M. Bcl-XL mutations suppress cellular sensitivity to antimycin A. *J. Biol. Chem.* **2004**, *279*, 2159–2165.
[22] O'Neill, J. W.; Manion, M. K.; Maguire, B.; Hockenbery, D. M. BCL-XL dimerization by three-dimensional domain swapping. *J. Mol. Biol.* **2006**, *356*, 367–381.
[23] Benoff, B.; Lawson, C.; Berman, H.; Carey, J. Long-range effects on structure in a temperature-sensitive mutant of trp repressor. To be published.
[24] Lawson, C. L.; Benoff, B.; Berger, T.; Berman, H. M.; Carey, J. *E. coli* trp repressor forms a domain-swapped array in aqueous alcohol. *Structure* **2004**, *12*, 1099–1108.
[25] Hatherley, D.; Graham, S. C.; Turner, J.; Harlos, K.; Stuart, D. I.; Barclay, A. N. Paired receptor specificity explained by structures of signal regulatory proteins alone and complexed with CD47. *Mol. Cell* **2008**, *31*, 266–277.
[26] Honnappa, S.; Okhrimenko, O.; Jaussi, R.; Jawhari, H.; Jelesarov, I.; Winkler, F. K.; Steinmetz, M. O. Key interaction modes of dynamic +TIP networks. *Mol. Cell* **2006**, *23*, 663–671.
[27] O'Neill, J. W.; Kim, D. E.; Johnsen, K.; Baker, D.; Zhang, K. Y. Single-site mutations induce 3D domain swapping in the B1 domain of protein L from *Peptostreptococcus magnus*. *Structure* **2001**, *9*, 1017–1027.
[28] Max, K. E.; Zeeb, M.; Bienert, R.; Balbach, J.; Heinemann, U. Common mode of DNA binding to cold shock domains: Crystal structure of hexathymidine bound to the domain-swapped form of a major cold shock protein from *Bacillus caldolyticus*. *FEBS J.* **2007**, *274*, 1265–1279.
[29] Mueller, U.; Perl, D.; Schmid, F. X.; Heinemann, U. Thermal stability and atomic-resolution crystal structure of the Bacillus caldolyticus cold shock protein. *J. Mol. Biol.* **2000**, *297*, 975–988.
[30] Bewley, C. A.; Gustafson, K. R.; Boyd, M. R.; Covell, D. G.; Bax, A.; Clore, G. M.; Gronenborn, A. M. Solution structure of cyanovirin-N, a potent HIV-inactivating protein. *Nature Struct. Mol. Biol.* **1998**, *5*, 571–578.
[31] Yang, F.; Bewley, C. A.; Louis, J. M.; Gustafson, K. R.; Boyd, M. R.; Gronenborn, A. M.; Clore, G. M.; Wlodawer, A. Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. *J. Mol. Biol.* **1999**, *288*, 403–412.
[32] Schreuder, H. A.; de, B. B.; Dijkema, R.; Mulders, J.; Theunissen, H. J.; Grootenhuis, P. D.; Hol, W. G. The intact and cleaved human antithrombin III complex as a model for serpin-proteinase interactions. *Nature Struct. Mol. Biol.* **1994**, *1*, 48–54.
[33] Yamasaki, M.; Li, W.; Johnson, D. J.; Huntington, J. A. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* **2008**, *455*, 1255–1258.
[34] Liu, Y.; Hart, P. J.; Schlunegger, M. P.; Eisenberg, D. The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-Å resolution. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3437–3442.
[35] Liu, Y.; Gotte, G.; Libonati, M.; Eisenberg, D. A domain-swapped RNase A dimer with implications for amyloid formation. *Nature Struct. Mol. Biol.* **2001**, *8*, 211–214.
[36] Wlodawer, A.; Borkakoti, N.; Moss, D. S.; Howlin, B. Comparison of two independently refined models of ribonuclease-A. *Acta Crystallogr. B* **1986**, *42*, 379–387.
[37] Pesenti, M. E.; Spinelli, S.; Bezirard, V.; Briand, L.; Pernollet, J. C.; Tegoni, M.; Cambillau, C. Structural basis of the honey bee PBP pheromone and pH-induced conformational change. *J. Mol. Biol.* **2008**, *380*, 158–169.
[38] Pesenti, M. E.; Spinelli, S.; Bezirard, V.; Briand, L.; Pernollet, J. C.; Campanacci, V.; Tegoni, M.; Cambillau, C. Queen bee pheromone binding protein pH-induced domain swapping favors pheromone release. *J. Mol. Biol.* **2009**, *390*, 981–990.
[39] Khandelwal, P.; Keliikuli, K.; Smith, C. L.; Saper, M. A.; Zuiderweg, E. R. Solution structure and phosphopeptide binding to the N-terminal domain of *Yersinia* YopH: Comparison with a crystal structure. *Biochemistry* **2002**, *41*, 11425–11437.
[40] Smith, C. L.; Khandelwal, P.; Keliikuli, K.; Zuiderweg, E. R.; Saper, M. A. Structure of the type III secretion and substrate-binding domain of *Yersinia* YopH phosphatase. *Mol. Microbiol.* **2001**, *42*, 967–979.
[41] Martin, J. R.; Craven, C. J.; Jerala, R.; Kroon-Zitko, L.; Zerovnik, E.; Turk, V.; Waltho, J. P. The three-dimensional solution structure of human stefin A. *J. Mol. Biol.* **1995**, *246*, 331–343.
[42] Staniforth, R. A.; Giannini, S.; Higgins, L. D.; Conroy, M. J.; Hounslow, A. M.; Jerala, R.; Craven, C. J.; Waltho, J. P. Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. *EMBO J.* **2001**, *20*, 4774–4781.
[43] Sridharan, S.; Razvi, A.; Scholtz, J. M.; Sacchettini, J. C. The HPr proteins from the thermophile *Bacillus stearothermophilus* can form domain-swapped dimers. *J. Mol. Biol.* **2005**, *346*, 919–931.
[44] Bennett, M. J.; Eisenberg, D. Refined structure of monomeric diphtheria toxin at 2.3 Å resolution. *Protein Sci.* **1994**, *3*, 1464–1475.
[45] Roos, G.; Garcia-Pino, A.; Van, B. K.; Brosens, E.; Wahni, K.; Vandenbussche, G.; Wyns, L.; Loris, R.; Messens, J. The conserved active

site proline determines the reducing power of *Staphylococcus aureus* thioredoxin. *J. Mol. Biol.* **2007**, *368*, 800–811.
[46] Garcia-Pino, A.; Martinez-Rodriguez, S.; Wahni, K.; Wyns, L.; Loris, R.; Messens, J. Coupling of domain swapping to kinetic stability in a thioredoxin mutant. *J. Mol. Biol.* **2009**, *385*, 1590–1599.
[47] Hakansson, M.; Svensson, A.; Fast, J.; Linse, S. An extended hydrophobic core induces EF-hand swapping. *Protein Sci.* **2001**, *10*, 927–933.
[48] Jimenez, B.; Poggi, L.; Piccioli, M. Monitoring the early steps of unfolding of dicalcium and mono-Ce^{3+}-substituted forms of P43M calbindin D9k. *Biochemistry* **2003**, *42*, 13066–13073.
[49] Murray, A. J.; Lewis, S. J.; Barclay, A. N.; Brady, R. L. One sequence, two folds: A metastable structure of CD2. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7337–7341.
[50] Yang, W.; Wilkins, A. L.; Ye, Y.; Liu, Z. R.; Li, S. Y.; Urbauer, J. L.; Hellinga, H. W.; Kearney, A.; van der Merwe, P. A.; Yang, J. J. Design of a calcium-binding protein with desired structure in a cell adhesion molecule. *J. Am. Chem. Soc.* **2005**, *127*, 2085–2093.
[51] Kukimoto-Niino, M.; Sakamoto, A.; Kanno, E.; Hanawa-Suetsugu, K.; Terada, T.; Shirouzu, M.; Fukuda, M.; Yokoyama, S. Structural basis for the exclusive specificity of Slac2-a/melanophilin for the Rab27 GTPases. *Structure* **2008**, *16*, 1478–1490.
[52] Chavas, L. M.; Torii, S.; Kamikubo, H.; Kawasaki, M.; Ihara, K.; Kato, R.; Kataoka, M.; Izumi, T.; Wakatsuki, S. Structure of the small GTPase Rab27b shows an unexpected swapped dimer. *Acta Crystallogr. D Biol. Crystallogr.* **2007**, *63*, 769–779.
[53] Ettmayer, P.; France, D.; Gounarides, J.; Jarosinski, M.; Martin, M. S.; Rondeau, J. M.; Sabio, M.; Topiol, S.; Weidmann, B.; Zurini, M.; Bair, K. W. Structural and conformational requirements for high-affinity binding to the SH2 domain of Grb2(1). *J. Med. Chem.* **1999**, *42*, 971–980.
[54] Schiering, N.; Casale, E.; Caccia, P.; Giordano, P.; Battistini, C. Dimer formation through domain swapping in the crystal structure of the Grb2-SH2-Ac-pYVNV complex. *Biochemistry* **2000**, *39*, 13376–13382.
[55] Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **1991**, *253*, 657–661.
[56] Byeon, I. J.; Louis, J. M.; Gronenborn, A. M. A protein contortionist: Core mutations of GB1 that induce dimerization and domain swapping. *J. Mol. Biol.* **2003**, *333*, 141–152.
[57] Ramoni, R.; Spinelli, S.; Grolli, S.; Conti, V.; Merli, E.; Cambillau, C.; Tegoni, M. Deswapping bovine odorant binding protein. *Biochim. Biophys. Acta* **2008**, *1784*, 651–657.
[58] Tegoni, M.; Ramoni, R.; Bignetti, E.; Spinelli, S.; Cambillau, C. Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nature Struct. Mol. Biol.* **1996**, *3*, 863–867.
[59] Antonyuk, S. V.; Trevitt, C. R.; Strange, R. W.; Jackson, G. S.; Sangar, D.; Batchelor, M.; Cooper, S.; Fraser, C.; Jones, S.; Georgiou, T.; Khalili-Shirazi, A.; Clarke, A. R.; Hasnain, S. S.; Collinge, J. Crystal structure of human prion protein bound to a therapeutic antibody. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 2554–2558.
[60] Knaus, K. J.; Morillas, M.; Swietnicki, W.; Malone, M.; Surewicz, W. K.; Yee, V. C. Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nature Struct. Mol. Biol.* **2001**, *8*, 770–774.
[61] Czjzek, M.; Letoffe, S.; Wandersman, C.; Delepierre, M.; Lecroisey, A.; Izadi-Pruneyre, N. The crystal structure of the secreted dimeric form of the hemophore HasA reveals a domain swapping with an exchanged heme ligand. *J. Mol. Biol.* **2007**, *365*, 1176–1186.
[62] Wolff, N.; Izadi-Pruneyre, N.; Couprie, J.; Habeck, M.; Linge, J.; Rieping, W.; Wandersman, C.; Nilges, M.; Delepierre, M.; Lecroisey, A. Comparative analysis of structural and dynamic properties of the loaded and unloaded hemophore HasA: Functional implications. *J. Mol. Biol.* **2008**, *376*, 517–525.
[63] Crane, B. R.; Rosenfeld, R. J.; Arvai, A. S.; Ghosh, D. K.; Ghosh, S.; Tainer, J. A.; Stuehr, D. J.; Getzoff, E. D. N-terminal domain swapping and metal ion binding in nitric oxide synthase dimerization. *EMBO J.* **1999**, *18*, 6271–6281.
[64] Rosenfeld, R. J.; Garcin, E. D.; Panda, K.; Andersson, G.; Aberg, A.; Wallace, A. V.; Morris, G. M.; Olson, A. J.; Stuehr, D. J.; Tainer, J. A.; Getzoff, E. D. Conformational changes in nitric oxide synthases induced by chlorzoxazone and nitroindazoles: Crystallographic and computational analyses of inhibitor potency. *Biochemistry* **2002**, *41*, 13915–13925.
[65] Green, S. M.; Gittis, A. G.; Meeker, A. K.; Lattman, E. E. One-step evolution of a dimer from a monomeric protein. *Nature Struct. Mol. Biol.* **1995**, *2*, 746–751.
[66] Loll, P. J.; Lattman, E. E. The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+}, and the inhibitor pdTp, refined at 1.65 Å. *Proteins* **1989**, *5*, 183–201.
[67] Schoch, G. A.; D'Arcy, B.; Stihle, M.; Burger, D.; Bar, D.; Benz, J.; Thoma, R.; Ruf, A. Molecular switch in the glucocorticoid receptor: Active and passive antagonist conformations. *J. Mol. Biol.* **2010**, *395*, 568–577.
[68] Suino-Powell, K.; Xu, Y.; Zhang, C.; Tao, Y. G.; Tolbert, W. D.; Simons, Jr. S.S; Xu, H. E. Doubling the size of the glucocorticoid receptor ligand binding pocket by deacylcortivazol. *Mol. Cell Biol.* **2008**, *28*, 1915–1923.
[69] Wiesmann, C.; Ultsch, M. H.; Bass, S. H.; de Vos, A. M. Crystal structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor. *Nature* **1999**, *401*, 184–188.
[70] Ultsch, M. H.; Wiesmann, C.; Simmons, L. C.; Henrich, J.; Yang, M.; Reilly, D.; Bass, S. H.; de Vos, A. M. Crystal structures of the neurotrophin-binding domain of TrkA, TrkB and TrkC. *J. Mol. Biol.* **1999**, *290*, 149–159.
[71] Josephson, K.; Jones, B. C.; Walter, L. J.; DiGiacomo, R.; Indelicato, S. R.; Walter, M. R. Noncompetitive antibody neutralization of IL-10 revealed by protein engineering and X-ray crystallography. *Structure* **2002**, *10*, 981–987.
[72] Zdanov, A.; Schalk-Hihi, C.; Gustchina, A.; Tsang, M.; Weatherbee, J.; Wlodawer, A. Crystal structure of interleukin-10 reveals the functional dimer with an unexpected topological similarity to interferon gamma. *Structure* **1995**, *3*, 591–601.
[73] Sue, S. C.; Chen, J. Y.; Lee, S. C.; Wu, W. G.; Huang, T. H. Solution structure and heparin interaction of human hepatoma-derived growth factor. *J. Mol. Biol.* **2004**, *343*, 1365–1377.
[74] Sue, S. C.; Lee, W. T.; Tien, S. C.; Lee, S. C.; Yu, J. G.; Wu, W. J.; Wu, W. G.; Huang, T. H. PWWP module of human hepatoma-derived growth factor forms a domain-swapped dimer with much higher affinity for heparin. *J. Mol. Biol.* **2007**, *367*, 456–472.
[75] Ivanov, D.; Tsodikov, O. V.; Kasanov, J.; Ellenberger, T.; Wagner, G.; Collins, T. Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4353–4358.
[76] Byeon, I. J.; Meng, X.; Jung, J.; Zhao, G.; Yang, R.; Ahn, J.; Shi, J.; Concel, J.; Aiken, C.; Zhang, P.; Gronenborn, A. M. Structural convergence between cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* **2009**, *139*, 780–790.
[77] Luo, J.; Teplyakov, A.; Obmolova, G.; Malia, T.; Wu, S. J.; Beil, E.; Baker, A.; Swencki-Underwood, B.; Zhao, Y.; Sprenkle, J.; Dixon, K.; Sweet, R.; Gilliland, G. L. Structure of the EMMPRIN N-terminal domain 1: Dimerization via beta-strand swapping. *Proteins* **2009**, *77*, 1009–1014.
[78] Yu, X. L.; Hu, T.; Du, J. M.; Ding, J. P.; Yang, X. M.; Zhang, J.; Yang, B.; Shen, X.; Zhang, Z.; Zhong, W. D.; Wen, N.; Jiang, H.; Zhu, P.; Chen, Z. N. Crystal structure of HAb18G/CD147: Implications for immunoglobulin superfamily homophilic adhesion. *J. Biol. Chem.* **2008**, *283*, 18056–18065.
[79] Zhang, H. P.; Nagashima, T.; Hayashi, F.; Yokoyama, S. Solution structure of the RGS domain of Regulator of G-protein signaling 7. To be published.
[80] Soundararajan, M.; Willard, F. S.; Kimple, A. J.; Turnbull, A. P.; Ball, L. J.; Schoch, G. A.; Gileadi, C.; Fedorov, O. Y.; Dowler, E. F.; Higman, V. A.; Hutsell, S. Q.; Sundstrom, M.; Doyle, D. A.; Siderovski, D. P. Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6457–6462.
[81] Cota, E.; Jones, C.; Simpson, P.; Altroff, H.; Anderson, K. L.; du Merle, L.; Guignot, J.; Servin, A.; Le Bouguénec, C.; Mardon, H.; Matthews, S. The solution structure of the invasive tip complex from Afa/Dr fibrils. *Mol. Microbiol.* **2006**, *62*, 356–366.
[82] Jedrzejczak, R.; Dauter, Z.; Dauter, M.; Piatek, R.; Zalewska, B.; Mroz, M.; Bury, K.; Nowicki, B.; Kur, J. Structure of DraD invasin from uropathogenic Escherichia coli: A dimer with swapped beta-tails. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 157–164.
[83] The UniProt Consortium. The Universal Protein Resource (UniProt) in 2010. *Nucl. Acids Res.* **2010**, *38*, D142–D148.
[84] Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **1995**, *247*, 536–540.
[85] Diederichs, K.; Jacques, S.; Boone, T.; Karplus, P. A. Low-resolution structure of recombinant human granulocyte-macrophage colony stimulating factor. *J. Mol. Biol.* **1991**, *221*, 55–60.
[86] Milburn, M. V.; Hassell, A. M.; Lambert, M. H.; Jordan, S. R.; Proudfoot, A. E.; Graber, P.; Wells, T. N. A novel dimer configuration revealed by the crystal structure at 2.4 Å resolution of human interleukin-5. *Nature* **1993**, *363*, 172–176.
[87] Bennett, M. J.; Schlunegger, M. P.; Eisenberg, D. 3D domain swapping: A mechanism for oligomer assembly. *Protein Sci.* **1995**, *4*, 2455–2468.

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

[89] Chen, Y. W.; Stott, K.; Perutz, M. F. Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1257–1261.

[90] Bergdoll, M.; Remy, M. H.; Cagnon, C.; Masson, J. M.; Dumas, P. Proline-dependent oligomerization with arm exchange. *Structure* **1997**, *5*, 391–401.

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

[92] Barrientos, L. G.; Louis, J. M.; Botos, I.; Mori, T.; Han, Z.; O'Keefe, B. R.; Boyd, M. R.; Wlodawer, A.; Gronenborn, A. M. The domain-swapped dimer of cyanovirin-N is in a metastable folded state: Reconciliation of X-ray and NMR structures. *Structure* **2002**, *10*, 673–686.

[93] Mazzarella, L.; Capasso, S.; Demasi, D.; Dilorenzo, G.; Mattia, C. A.; Zagari, A. Bovine seminal ribonuclease – Structure at 1.9-Angstrom resolution. *Acta Crystallogr. D Biol. Crystallogr.* **1993**, *49*, 389–402.

[94] Canals, A.; Pous, J.; Guasch, A.; Benito, A.; Ribo, M.; Vilanova, M.; Coll, M. The structure of an engineered domain-swapped ribonuclease dimer and its implications for the evolution of proteins toward oligomerization. *Structure* **2001**, *9*, 967–976.

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

[96] Sirota, F. L.; Hery-Huynh, S.; Maurer-Stroh, S.; Wodak, S. J. Role of the amino acid sequence in domain swapping of the B1 domain of protein G. *Proteins* **2008**, *72*, 88–104.

[97] Murray, A. J.; Head, J. G.; Barker, J. J.; Brady, R. L. Engineering an intertwined form of CD2 for stability and assembly. *Nature Struct. Mol. Biol.* **1998**, *5*, 778–782.

[98] Rousseau, F.; Schymkowitz, J.; Itzhaki, L. S. Implications of 3D Domain Swapping for Protein Folding, Misfolding and Function. In *Protein Dimerization (and Oligomerization) in Biology*; Matthews, J., Ed.; Landes Bioscience: Austin, TX, 2010.

[99] Bennett, M. J.; Sawaya, M. R.; Eisenberg, D. Deposition diseases and 3D domain swapping. *Structure* **2006**, *14*, 811–824.

[100] Gronenborn, A. M. Protein acrobatics in pairs – Dimerization via domain swapping. *Curr. Opin. Struct. Biol.* **2009**, *19*, 39–49.

[101] Newcomer, M. E. Protein folding and three-dimensional domain swapping: A strained relationship? *Curr. Opin. Struct. Biol.* **2002**, *12*, 48–53.

[102] Parge, H. E.; Arvai, A. S.; Murtari, D. J.; Reed, S. I.; Tainer, J. A. Human CksHs2 atomic structure: A role for its hexameric assembly in cell cycle control. *Science* **1993**, *262*, 387–395.

[103] Seeliger, M. A.; Spichty, M.; Kelly, S. E.; Bycroft, M.; Freund, S. M.; Karplus, M.; Itzhaki, L. S. Role of conformational heterogeneity in domain swapping and adapter function of the Cks proteins. *J. Biol. Chem.* **2005**, *280*, 30448–30459.

[104] Wolynes, P.; Luthey-Schulten, Z.; Onuchic, J. Fast-folding experiments and the topography of protein folding energy landscapes. *Chem. Biol.* **1996**, *3*, 425–432.

[105] Yang, S.; Cho, S. S.; Levy, Y.; Cheung, M. S.; Levine, H.; Wolynes, P. G.; Onuchic, J. N. Domain swapping is a consequence of minimal frustration. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13786–13791.

[106] Cho, S. S.; Levy, Y.; Onuchic, J. N.; Wolynes, P. G. Overcoming residual frustration in domain-swapping: The roles of disulfide bonds in dimerization and aggregation. *Phys. Biol.* **2005**, *2*, S44–S55.

[107] Matei, E.; Furey, W.; Gronenborn, A. M. Solution and crystal structures of a sugar binding site mutant of cyanovirin-N: No evidence of domain swapping. *Structure* **2008**, *16*, 1183–1194.

[108] Koharudin, L. M.; Viscomi, A. R.; Jee, J. G.; Ottonello, S.; Gronenborn, A. M. The evolutionarily conserved family of cyanovirin-N homologs: Structures and carbohydrate specificity. *Structure* **2008**, *16*, 570–584.

[109] Ding, F.; Prutzman, K. C.; Campbell, S. L.; Dokholyan, N. V. Topological determinants of protein domain swapping. *Structure* **2006**, *14*, 5–14.

[110] Kundu, S.; Jernigan, R. L. Molecular mechanism of domain swapping in proteins: An analysis of slower motions. *Biophys. J.* **2004**, *86*, 3846–3854.

[111] Malevanets, A.; Sirota, F. L.; Wodak, S. J. Mechanism and energy landscape of domain swapping in the B1 domain of protein G. *J. Mol. Biol.* **2008**, *382*, 223–235.

[112] Fink, A. L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **1998**, *3*, R9–R23.

[113] Gotte, G.; Bertoldi, M.; Libonati, M. Structural versatility of bovine ribonuclease A: Distinct conformers of trimeric and tetrameric aggregates of the enzyme. *Eur. J. Biochem.* **1999**, *265*, 680–687.

[114] Piccoli, R.; Di Donato, A.; D'Alessio, G. Co-operativity in seminal ribonuclease function: Kinetic studies. *Biochem. J.* **1988**, *253*, 329–336.

[115] Nenci, A.; Gotte, G.; Bertoldi, M.; Libonati, M. Structural properties of trimers and tetramers of ribonuclease A. *Protein Sci.* **2001**, *10*, 2017–2027.

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

[117] Byeon, I. J.; Louis, J. M.; Gronenborn, A. M. A captured folding intermediate involved in dimerization and domain-swapping of GB1. *J. Mol. Biol.* **2004**, *340*, 615–625.

[118] Jee, J.; Byeon, I. J.; Louis, J. M.; Gronenborn, A. M. The point mutation A34F causes dimerization of GB1. *Proteins* **2008**, *71*, 1420–1431.

[119] Louis, J. M.; Byeon, I. J.; Baxa, U.; Gronenborn, A. M. The GB1 amyloid fibril: Recruitment of the peripheral beta-strands of the domain swapped dimer into the polymeric interface. *J. Mol. Biol.* **2005**, *348*, 687–698.

[120] Kirsten, F. M.; Dyda, F.; Dobrodumov, A.; Gronenborn, A. M. Core mutations switch monomeric protein GB1 into an intertwined tetramer. *Nature Struct. Mol. Biol.* **2002**, *9*, 877–885.

[121] Wikstrom, M.; Drakenberg, T.; Forsen, S.; Sjobring, U.; Bjorck, L. Three-dimensional solution structure of an immunoglobulin light chain-binding domain of protein L: Comparison with the IgG-binding domains of protein G. *Biochemistry* **1994**, *33*, 14011–14017.

[122] Dolk, E.; van der Vaart, M.; Hulsik, D. L.; Vriend, G.; de Haard, H.; Spinelli, S.; Cambillau, C.; Frenken, L.; Verrips, T. Isolation of llama antibody fragments for prevention of dandruff by phage display in shampoo. *Appl. Environ. Microbiol.* **2005**, *71*, 442–450.

[123] Spinelli, S.; Desmyter, A.; Frenken, L.; Verrips, T.; Tegoni, M.; Cambillau, C. Domain swapping of a llama VHH domain builds a crystal-wide beta-sheet structure. *FEBS Lett.* **2004**, *564*, 35–40.

[124] Botos, I.; O'Keefe, B. R.; Shenoy, S. R.; Cartner, L. K.; Ratner, D. M.; Seeberger, P. H.; Boyd, M. R.; Wlodawer, A. Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. *J. Biol. Chem.* **2002**, *277*, 34336–34342.

[125] Barrientos, L. G.; Lasala, F.; Delgado, R.; Sanchez, A.; Gronenborn, A. M. Flipping the switch from monomeric to dimeric CV-N has little effect on antiviral activity. *Structure* **2004**, *12*, 1799–1807.

[126] Ziolkowska, N. E.; O'Keefe, B. R.; Mori, T.; Zhu, C.; Giomarelli, B.; Vojdani, F.; Palmer, K. E.; McMahon, J. B.; Wlodawer, A. Domain-swapped structure of the potent antiviral protein griffithsin and its mode of carbohydrate binding. *Structure* **2006**, *14*, 1127–1135.

[127] Mori, T.; O'Keefe, B. R.; Sowder, R. C.; Bringans, S.; Gardella, R.; Berg, S.; Cochran, P.; Turpin, J. A.; Buckheit, Jr. R. W.; McMahon, J. B.; Boyd, M. R. Isolation and characterization of griffithsin, a novel HIV-inactivating protein, from the red alga Griffithsia sp. *J. Biol. Chem.* **2005**, *280*, 9345–9353.

[128] Williams, Jr. D. C.; Lee, J. Y.; Cai, M.; Bewley, C. A.; Clore, G. M. Crystal structures of the HIV-1 inhibitory cyanobacterial protein MVL free and bound to Man3GlcNAc2: Structural basis for specificity and high-affinity binding to the core pentasaccharide from n-linked oligomannoside. *J. Biol. Chem.* **2005**, *280*, 29269–29276.

[129] Ziolkowska, N. E.; Wlodawer, A. Structural studies of algal lectins with anti-HIV activity. *Acta Biochim. Polonica* **2006**, *53*, 617–626.

[130] 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.; Kelly, J. W.; Rudd, P. M.; Dwek, R. A.; Katinger, H.; Burton, D. R.; Wilson, I. A. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* **2003**, *300*, 2065–2071.

[131] Gamble, T. R.; Yoo, S.; Vajdos, F. F.; von Schwedler, U. K.; Worthylake, D. K.; Wang, H.; McCutcheon, J. P.; Sundquist, W. I.; Hill, C. P. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **1997**, *278*, 849–853.

[132] Ganser-Pornillos, B. K.; Cheng, A.; Yeager, M. Structure of full-length HIV-1 CA: A model for the mature capsid lattice. *Cell* **2007**, *131*, 70–79.

[133] Ternois, F.; Sticht, J.; Duquerroy, S.; Krausslich, H. G.; Rey, F. A. The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nature Struct. Mol. Biol.* **2005**, *12*, 678–682.

[134] Worthylake, D. K.; Wang, H.; Yoo, S.; Sundquist, W. I.; Hill, C. P. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55*, 85–92.

[135] Ivanov, D.; Stone, J. R.; Maki, J. L.; Collins, T.; Wagner, G. Mammalian SCAN domain dimer is a domain-swapped homolog of the HIV capsid C-terminal domain. *Mol. Cell* **2005**, *17*, 137–143.
[136] Qu, C.; Liljas, L.; Opalka, N.; Brugidou, C.; Yeager, M.; Beachy, R. N.; Fauquet, C. M.; Johnson, J. E.; Lin, T. 3D domain swapping modulates the stability of members of an icosahedral virus group. *Structure* **2000**, *8*, 1095–1103.
[137] Zahn, R.; Liu, A.; Luhrs, T.; Riek, R.; von Schroetter, C.; Lopez, G. F.; Billeter, M.; Calzolai, L.; Wider, G.; Wuthrich, K. NMR solution structure of the human prion protein. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 145–150.
[138] Lee, S.; Eisenberg, D. Seeded conversion of recombinant prion protein to a disulfide-bonded oligomer by a reduction-oxidation process. *Nature Struct. Mol. Biol.* **2003**, *10*, 725–730.
[139] Nakai, S.; Li-Chan, E. C.; Dou, J. Pattern similarity study of functional sites in protein sequences: Lysozymes and cystatins. *BMC Biochemistry* **2005**, *6*, 9.
[140] Janowski, R.; Kozak, M.; Abrahamson, M.; Grubb, A.; Jaskolski, M. 3D domain-swapped human cystatin C with amyloidlike intermolecular beta-sheets. *Proteins* **2005**, *61*, 570–578.
[141] Jenko, K. S.; Guncar, G.; Stern, I.; Morgan, G.; Rabzelj, S.; Kenig, M.; Staniforth, R. A.; Waltho, J. P.; Zerovnik, E.; Turk, D. Essential role of proline isomerization in stefin B tetramer formation. *J. Mol. Biol.* **2007**, *366*, 1569–1579.
[142] Sanders, A.; Jeremy, C. C.; Higgins, L. D.; Giannini, S.; Conroy, M. J.; Hounslow, A. M.; Waltho, J. P.; Staniforth, R. A. Cystatin forms a tetramer through structural rearrangement of domain-swapped dimers prior to amyloidogenesis. *J. Mol. Biol.* **2004**, *336*, 165–178.
[143] Janowski, R.; Abrahamson, M.; Grubb, A.; Jaskolski, M. Domain swapping in N-truncated human cystatin C. *J. Mol. Biol.* **2004**, *341*, 151–160.
[144] Cameron, A. D.; Olin, B.; Ridderstrom, M.; Mannervik, B.; Jones, T. A. Crystal structure of human glyoxalase I: Evidence for gene duplication and 3D domain swapping. *EMBO J.* **1997**, *16*, 3386–3395.
[145] Saint-Jean, A. P.; Phillips, K. R.; Creighton, D. J.; Stone, M. J. Active monomeric and dimeric forms of *Pseudomonas putida* glyoxalase I: Evidence for 3D domain swapping. *Biochemistry* **1998**, *37*, 10345–10353.
[146] Bennett, C. L.; Christie, J.; Ramsdell, F.; Brunkow, M. E.; Ferguson, P. J.; Whitesell, L.; Kelly, T. E.; Saulsbury, F. T.; Chance, P. F.; Ochs, H. D. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nature Genet.* **2001**, *27*, 20–21.
[147] Brunkow, M. E.; Jeffery, E. W.; Hjerrild, K. A.; Paeper, B.; Clark, L. B.; Yasayko, S. A.; Wilkinson, J. E.; Galas, D.; Ziegler, S. F.; Ramsdell, F. Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nature Genet.* **2001**, *27*, 68–73.
[148] Lai, C. S.; Fisher, S. E.; Hurst, J. A.; Vargha-Khadem, F.; Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **2001**, *413*, 519–523.
[149] Wildin, R. S.; Ramsdell, F.; Peake, J.; Faravelli, F.; Casanova, J. L.; Buist, N.; Levy-Lahad, E.; Mazzella, M.; Goulet, O.; Perroni, L.; Bricarelli, F. D.; Byrne, G.; McEuen, M.; Proll, S.; Appleby, M.; Brunkow, M. E. X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nature Genet.* **2001**, *27*, 18–20.
[150] Joachimiak, A.; Kelley, R. L.; Gunsalus, R. P.; Yanofsky, C.; Sigler, P. B. Purification and characterization of trp aporepressor. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 668–672.
[151] Fernando, T.; Royer, C. Role of protein-protein interactions in the regulation of transcription by trp repressor investigated by fluorescence spectroscopy. *Biochemistry* **1992**, *31*, 3429–3441.
[152] Martin, K. S.; Royer, C. A.; Howard, K. P.; Carey, J.; Liu, Y. C.; Matthews, K.; Heyduk, E.; Lee, J. C. Electrostatic forces contribute to interactions between trp repressor dimers. *Biophys. J.* **1994**, *66*, 1167–1173.
[153] Babu, K. R.; Moradian, A.; Douglas, D. J. The methanol-induced conformational transitions of beta-lactoglobulin, cytochrome c, and ubiquitin at low pH: A study by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 317–328.
[154] Buck, M. Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* **1998**, *31*, 297–355.
[155] Law, R. H.; Zhang, Q.; McGowan, S.; Buckle, A. M.; Silverman, G. A.; Wong, W.; Rosado, C. J.; Langendorf, C. G.; Pike, R. N.; Bird, P. I.; Whisstock, J. C. An overview of the serpin superfamily. *Genome Biol.* **2006**, *7*, 216.
[156] Huntington, J. A.; Pannu, N. S.; Hazes, B.; Read, R. J.; Lomas, D. A.; Carrell, R. W.A 2.6 Å structure of a serpin polymer and implications for conformational disease. *J. Mol. Biol.* **1999**, *293*, 449–455.
[157] Izadi-Pruneyre, N.; Huche, F.; Lukat-Rodgers, G. S.; Lecroisey, A.; Gilli, R.; Rodgers, K. R.; Wandersman, C.; Delepelaire, P. The heme transfer from the soluble HasA hemophore to its membrane-bound receptor HasR is driven by protein-protein interaction from a high to a lower affinity binding site. *J. Biol. Chem.* **2006**, *281*, 25541–25550.
[158] Arnoux, P.; Haser, R.; Izadi, N.; Lecroisey, A.; Delepierre, M.; Wandersman, C.; Czjzek, M. The crystal structure of HasA, a hemophore secreted by *Serratia marcescens*. *Nature Struct. Mol. Biol.* **1999**, *6*, 516–520.
[159] Letoffe, S.; Debarbieux, L.; Izadi, N.; Delepelaire, P.; Wandersman, C. Ligand delivery by haem carrier proteins: The binding of *Serratia marcescens* haemophore to its outer membrane receptor is mediated by two distinct peptide regions. *Mol. Microbiol.* **2003**, *50*, 77–88.