Published in final edited form as:

Curr Opin Struct Biol. 2009 February ; 19(1): 39–49. doi:10.1016/j.sbi.2008.12.002.

Protein acrobatics in pairs – Dimerization via Domain Swapping

Angela M. Gronenborn

Department of Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA

Introduction

Ever since Anfinsen's seminal experiments about 50 years ago, it is widely accepted that the amino acid sequence of a protein determines its unique three-dimensional structure. While generally true for single chain proteins, in multimers, two or more individual chains can be assembled in different ways. Therefore, making a distinction between tertiary and quaternary protein structure may be prudent when evoking Anfinsen's postulate. The folded, native structure of a protein is considered the lowest energy state, with folding being driven by a combination of entropic and enthalpic forces that result in the burial of hydrophobic residues and a specific distribution of polar residues mainly on the surface. This balanced interplay of forces creates a network of defined attractive and repulsive forces that arrange the chain in well-defined, secondary structure elements. The predominance of such cooperative folding on smooth, free energy landscapes for most small, naturally occurring single chain proteins has been explained as the product of natural selection acting at the level of an individual amino acid sequence.

When initially established, the protein structure database [1] primarily contained single chain, monomeric proteins; however, over the past several years, the number of multimeric proteins represented in the PDB has increased rapidly and continues to grow. Indeed, multimeric proteins are prevalent in all organisms, and oligomerization is generally believed to be favored during protein evolution [2], since multimers are endowed with structural and functional advantages, such as improved stability and control over the accessibility and specificity of active sites.

For multimers, the specification of a unique oligomeric state may not always be unambiguous. 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 three quaternary structures [3] that vary in their domain interactions. Alternate quaternary assemblies of functionally distinct homo-oligomeric proteins have recently been termed 'morphoeins' [4], with the enzyme porphobilinogen synthase (PBGs) as the prototypical example. PBGS exists in an equilibrium between an octamer, a hexamer, and two dimer conformations [5]

Angela M. Gronenborn, Department of Structural Biology, University of Pittsburgh School of Medicine, 1051 Biomedical Science Tower 3, 3501 Fifth Ave, Pittsburgh, PA 15261, Phone: (412) 648-9959, Fax: (412) 648-9008, Email: E-mail: amg100@pitt.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A special case of oligomer assembly occurs via domain swapping. The term three-dimensional (3D) domain-swapping, or simply domain swapping, was coined by Eisenberg [6] for an oligomerization mechanism in which two or more polypeptide chains 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, or, if more chains are involved, oligomers ensue (Figure 1).

Not surprisingly, most domain-swapped structures have been determined by X-ray crystallography. Currently, more than 60 examples of domain swapped structures are available in the PDB [1]. According to the original definition, both monomeric and oligomeric structures must be observed for an identical protein [7], with both states found either in crystals or one in a crystal and the other in solution. This purist's designation, however, has been relaxed over the last decade. 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. Originally, in the former case, the protein is a 'candidate' for domain swapping, while in the latter, the oligomers are classified as 'quasi-domain-swapped'. In the present review, 'quasi-domain-swapped' structures are termed domain swapped when the amino acid sequences of monomer and multimer are very similar, *i.e.* the proteins are mutants or close homologs.

The collection of domain swapped protein structures described up to 2002 has been summarized in reviews by Newcomer [8] and Liu and Eisenberg [7]. A more recent review focuses on proteins that display 3D domain swapping as well as fibril formation and discusses the possible involvement of domain swapping in protein deposition diseases [9]. Therefore, I only briefly summarized notations and terminology and the reader is referred to the above reviews for more in depth descriptions of the basic features of domain swapping.

Terms and descriptors

The subunit or domain structure in a domain swapped oligomer and the corresponding monomer is identical, exhibiting no differences in phi, psi backbone angles, except for the region that links the exchanging domains. This part of the polypeptide chain is called the "hinge loop" and often folds back on itself in the monomer, while adopting an extended conformation in the domain swapped oligomer. Although called "domain" swapping, the term 'domain' can refer to a variety of structural units: the largest may truly constitute independently folding domains, while the smallest can comprise single secondary structure elements, such as a sole β strand or an isolated α helix. In principle, the exchanging unit can be located anywhere in the sequence, although it is often located at one of the two termini. In rare cases, half of the entire structure may be involved in the domain swap, with CV-N as the prototypical example of this type. Inter-molecular interactions in dimers that possess an identical intra-molecular counterpart in the monomer form the 'closed' or primary interface while newly created contacts reside in the 'open' or secondary interface. A schematic representation of different domain swapping scenarios as well as delineation of the different structural interfaces is provided in Figure 1.

Mechanism of Domain Swapping

A number of studies have looked for distinguishing features of domain swapping proteins compared to all protein space. However, no obvious sequence similarities or resemblance in secondary structure arrangements have been found [7]. Figure 2 presents selected examples of domain swapped structures of different secondary structure and fold types. In search of local signals that may cause swapping, attention has focused on the hinge loop region, since this is the only part of these proteins that adopts a different conformation in monomeric and domain swapped forms. Frequently, proline residues are found in the hinge loop, suggesting their possible importance [10,11]. Rousseau and colleagues have suggested that these prolines create

strain in the hinge loop of suc1 and thereby influence domain swapping. Specifically, replacement of the first proline in the hinge with an alanine stabilizes the monomer form, whereas the same substitution of the second proline stabilizes the dimer form. These authors suggest that tension in the hinge loop in the monomer causes it to behave like a loaded molecular spring which is released when the alternative conformation is adopted in the dimer [11]. Unlike suc1, mutation of the single proline in the hinge loop of CVN, to glycine, substantially stabilizes both states of the protein, with greater stabilization of the monomer compared to the dimer [12]. In contrast, adding a second proline residue by mutating a neighboring amino acid causes the domain swapped dimer to become the thermodynamically most stable state [12]. In general, there are many examples of domain swapped structures that are devoid of prolines in the hinge loop, rendering the presence of a proline likely, but not necessary. Overall, it is doubtful that sequence features alone determine whether a protein will undergo domain swapping.

The propensity of a protein for domain swapping can be modulated by altering the length of the hinge loop. Intuitively, one may assume that lengthening the loop would aid monomer formation and shortening would cause dimers, as a short loop in the monomer would make it difficult for the polypeptide to fold back on itself and would allow the swapped portion of the chain to find partners more easily. However, this concept may be over-simplified, as lengthening the hinge has been shown to aid domain swapping in some cases. 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, which increased as the lengths of the insertions increased [13]. In other cases, dimer formation is promoted by shortening the hinge loop. Indeed, loop deletion is observed in a number of natural proteins and has been used in some designed ones. An example of the importance of loop length in protein behavior is provided by two different three helix bundles engineered by the Eisenberg lab [14]. Loop deletion in one caused the formation of a domain swapped dimer whereas loop deletion in the other resulted in fibril formation. Thus, the influence of loop length on protein architecture appears to be unique to each protein.

Although substantial efforts have been made to identify a unifying molecular mechanism for domain swapping, no compelling proposal has emerged to date.

Theoretical and computational explorations

Computational approaches for deciphering the elementary events of protein folding and assembly are beginning to be worked out, using reduced models and detailed atomistic simulations. Several groups are applying these methodologies to understand domain swapping. An elegant concept for understanding the folding of small, monomeric proteins involves movement of the polypeptide chain by Brownian motion through a funneled energy landscape that guides structure formation dominated by native stability [15]. This energy landscape concept, developed for gaining insights into folding/unfolding of monomeric proteins, has also been applied to protein associations in domain swapped multimers. In particular, Onuchic and Wolynes have used a symmetrized Go-type potential to simulate the mechanism of domain swapping using MD simulations with the replica exchange method. They discovered a frustrated hinge region in the Eps8 SH3 dimer 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. Based on these results, the authors propose that the overall monomeric topology, rather than local signals in the hinge region, determines where in the polypeptide chain domain swapping will occur [16]. This notion is at odds with the different domain swapped dimers observed for GB1 and LB1 (see below), proteins with identical monomeric topologies. The authors also suggest that partially folded intermediates can serve as “templates” for self-assembling aggregates. They called proteins with intrinsic symmetry of the sequence and/or structure highly frustrated and observed multi-

mode domain-swapping in their simulations, necessitating the inclusion of inter- or intra-molecular disulfide bonds [17]. Two proteins that fall into the ‘highly frustrated’ category are the human prion protein and CV-N. However, at least for CV-N, disulfide bonds do not appear to be a prerequisite for domain swapping experimentally; several homologs of CV-N with varying numbers of disulfide bonds do not exhibit domain swapping [18].

Molecular dynamics simulations have also been applied to study domain swapping and an illustrative example was carried out by Esposito and Daggett on RNase A [19]. Their simulations were conducted with parameters that mimic experimental conditions known to favor either the formation of N- or C-terminal domain swapped dimers and show that the same protein can have different pathways for domain swapping. Under mildly unfolding conditions, a partial or complete opening of the N-terminal helix was observed, whereas dislocation of the C-terminal strand, away from the core of the structure, only occurred when more extreme, low pH conditions were employed.

The mechanism of domain swapping has also been investigated by analyzing large-scale domain motions via Gaussian network models (GNM). In a study of diphtheria toxin, Kundu and Jernigan uncovered major hinges in this protein based on slower modes observed by GNM and determined the direction of the motion of the swapped domain about the hinge using the anisotropic network model (ANM) [20]. A principal hinge for the domain swapping transition was found in both the monomer and the ‘open’ chain in the dimer. The location of the hinge appears to be robust in their models, with the caveat that computations with the GNM and ANM models are based on contact maps that are virtually identical in the monomer and dimer. As a result, very similar motions are expected to dominate both structures.

A quite different mechanism of domain swapping involves the progressive and reversible transformation between monomer and dimer, starting from either end of the polypeptide chain. In this process, intra-molecular contacts are traded for equivalent inter-molecular ones, with the total number of native contacts remaining essentially constant. In this manner an increasing portion of the chains between monomers becomes substituted, until a stable conformational state is reached. This mechanism is proposed by the Wodak group, who recently computed the complete free energy profiles for GB1 using all atom models and classical force-fields [21]. Swapping started from the C-terminus of the protein and did not involve unfolding. All conformational changes within the individual monomers and the binding between them were tightly coupled. Throughout this process, the total number of native contacts was maximized and solvent exposure minimized. A large number of hinge conformations and nearly isoenergetic association modes were sampled by the intermediates, suggesting that the exchange reaction is a nonspecific process akin to encounter complex formation. As a result, amino acid sequence only played a marginal role.

Something old – something new: A few examples

RNase A

The classic example of a protein engaged in domain swapping is bovine pancreatic ribonuclease A (RNase A), for which dimerization, involving N-terminal exchange was proposed in 1962 by Crestfield, Stein, and Moore to explain its behavior under acidic conditions [22]. The first X-ray structures for a domain swapped dimer were solved in the late nineties [23] and the Eisenberg laboratory subsequently identified domain swapped dimers, trimers, and multimers [24,25]. Either the N-terminal helix or the C-terminal strand can domain swap and RNase A frequently is considered the prototype of a domain swapping protein. The swapped unit and its oligomeric state can vary significantly and RNase A beautifully illustrates the amazing possibilities of domain swapping modes. In the trimer both N- and C-terminal units are swapped and a circular arrangement ensues [25] (see Figure 2K, L, M). Given that RNase can swap two

different domains, a variety of assembled oligomeric structures can be formed [25,26]. Models for such trimers, tetramers, and other oligomers have been proposed [27]. Domain swapping and oligomerization creates two catalytic sites and yield a spatial arrangement of amino acids from different subunits that is characteristic of the active site of RNase, with the catalytic histidines contributed by the N-terminal α helix and the C-terminal β strand, respectively.

B1 domains

The immunoglobulin binding domain B1 of streptococcal protein G (GB1) is a small, 56 residue, stable, single domain protein composed of a four-stranded β -sheet, on top of which a single α helix is packed [28]. This protein exhibits surprising structural variability, as illustrated by the discovery of several unusual structural variants in an extensive study of mutants derived from a library of randomized hydrophobic core residues. One variant was a domain swapped dimer in which exchange of one hairpin between the subunits occurred [29]. In this dimer structure, four adjacent hairpins are put together using alternate monomer chains, resulting in a final eight-stranded β -sheet with two extensive new interfaces (Figure 2Q). The helices are arranged in an anti-parallel manner, crossing 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. At low micromolar protein concentrations, the dimer dissociates into partially folded, monomeric species that exhibit extensive conformational motions on the micro- to millisecond timescale for a substantial portion of the polypeptide chain. Despite these fluctuations, the overall architecture of the monomer resembles that of wildtype GB1, although neither stable H-bonds nor specific salt-bridges are locked-in. Thus, destabilization or opening up of the hydrophobic core in monomeric GB1 can be compensated for by extending the core into a larger structure *via* domain swapping and dimerization [30].

Comparison of the domain-swapped dimer structure with the wildtype monomer suggested that Phe34 was a pivotal side chain for the monomer-dimer switch. Reverting this residue back to the wildtype Ala yielded a wildtype-like monomer structure. This led to the speculation that the equivalent change, A34F, in the wildtype protein would also induce domain-swapping. Intriguingly, this was not the case and a different side-by-side dimer, without domain swapping, was found [31].

GB1 variants that can exist as domain swapped dimer structures can also form fibrils. No fibril formation occurs for mutants that fold into the stable, wildtype GB1 structure or for variants that exist as a highly destabilized, fluctuating ensemble of random, folded and partially folded structures under the same experimental conditions. A simple model for the GB1 fibril, based on experimental disulfide cross-linking results, has been constructed using a swapped dimer structure as the smallest unit. In this model, the β 3 edge strands are in close proximity and a continuous left-handed helical ribbon wraps around the helix axis [32].

Mutation of an additional single core amino acid in the domain swapped dimer caused a surprising and dramatic change in structure, namely a symmetric tetramer with inter-molecular strand-exchange involving all four units [33]. Four of the five secondary structure elements that are present in the monomeric wildtype GB1 are retained in the tetramer (three β strands and the α helix; strand β 2 of the first hairpin is missing), although their intra- and intermolecular interactions are radically altered. A notable difference, compared to the wildtype arrangement, is the loss of the β 3- β 4 hairpin interaction within the monomeric unit. This interaction is substituted by 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 goes hand in hand with a shift in register between interacting residues in β 3 and β 4. In addition to strand-exchange of the domain-swapping kind, a new interface is formed between surface elements of the individual chains.

For the B1 domain of protein L (LB1), a protein that adopts the same fold as the GB1 domain monomer [34], however, a quite different domain swapped structure was found (Figure 2P). In LB1, substitution of a glycine in the turn of the second β -hairpin by alanine caused exchange of the C-terminal β -strand between the subunits, with the wildtype hairpin turn straightening out and forming the inter-molecular β -sheet interface. At the mutation site, the long β -strands are kinked, causing both B1 units to be rotated around the hinge region. Another mutant, in which a valine in the hydrophobic core was changed to alanine, was also found to form this domain swapped structure [34]. In this case, the asymmetric unit contained two wildtype-like monomers and a domain swapped dimer. In the dimer structure, the void in the hydrophobic core, created by the loss of the two methyl groups, is compensated for by rotation of a leucine side chain that resides on the neighboring β strand belonging to the second monomer. In addition, novel intermolecular hydrophobic contacts between a Phe in one monomer and a Tyr side chains on the other are formed. These hydrophobic interactions as well as inter-molecular H bonds between the exchanged β strands contribute to the stability of the domain swap [34].

The different oligomeric B1 structures provide particularly intriguing examples of how oligomeric proteins can evolve from monomeric forms. In addition, the tetrameric GB1 mutant constitutes a rare case where core mutations, rather than surface changes caused the assembly into a stable oligomer.

Cyanovirin-N

Cyanovirin-N (CV-N) is a 101 amino acid cyanobacterial lectin that was originally isolated from an aqueous extract of the cyanobacterium *Nostoc ellipsosporum*. CV-N inactivates HIV and is a general virucidal agent against other enveloped viruses. The original solution structure was a monomer [35] whereas the subsequently determined X-ray structures [36,37] were domain swapped dimers (Figure 2I). Depending on experimental conditions, either form predominates and can be isolated for biophysical, structural, and functional studies [12]. The monomer structure exhibits a compact, bilobal fold with pseudo-symmetry. Interestingly, the amino acid sequence repeats of CV-N do not constitute the repeated structural domains. Rather, the two symmetrically related domains are formed by strand exchange across the two sequence repeats. Each domain comprises a triple-stranded β -sheet with a β -hairpin packed on top. A helical linker is located in the middle of the sequence and in the domain swapped dimer this linker acts as the hinge. In the dimer, the monomers open up, and the two halves of the dimer essentially contain the same interactions as the monomer. It was shown that the domain-swapped dimer is a kinetically trapped folding intermediate at high protein concentrations that convert into the slightly more stable monomer form at physiological ($>30^{\circ}\text{C}$) temperature. At room temperature or below, however, the dimer lifetime is sufficiently long for structural characterization in solution [12] and the solution dimer structure is virtually identical in inter-domain packing and overall folding to the structures in the trigonal and tetragonal crystals [36,37]. The sole difference between all swapped dimer structures is the relative domain-domain orientations of the pseudo-monomeric halves. In solution, residues in the hinge of the domain swapped dimeric species exhibit conformational mobility on the microsecond time scale, suggesting that reorientation around this hinge is easily possible and that different orientations can be trapped in alternate crystal lattices.

The fact that both monomeric and domain-swapped dimeric CV-N coexist in solution under identical conditions indicates that the free energies of folding for both quaternary states must be comparable and the kinetic barrier between the monomer and dimer has to be significant. This can be altered by mutation of residues in the hinge region. For instance, changing the single proline in the sequence that resides in the hinge region to glycine resulted in a substantial stabilization of this monomeric P51G mutant by >5 kcal/mol compared to wild-type. A S52P mutant yielded predominantly dimeric protein due to destabilization of the monomer, and a

deletion mutant, $\Delta Q52$, existed solely as a domain swapped dimer [38]. 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 in the crystal [39]. No domain-swapping was observed for this variant under all conditions, although two monomers are found in the asymmetric unit. Interestingly, the protein-protein interface between the two monomers is reminiscent of the open interface in one of the domain swapped dimer structures of wildtype CV-N [39].

GRFT

A second antiviral lectin for which domain swapping was observed is Griffithsin (GRFT) [40]. The structure of GRFT (Figure 2J) closely resembles jacalin lectins and consists of triangular prisms with three repeats of a four-stranded antiparallel β -sheet. The first two β -strands of one chain complete the β -prism of the other chain. Thus far, GRFT is the only example of a jacalin-fold protein for which a domain swapped structure has been observed.

Fab 2G12

The human antibody 2G12 neutralizes a broad range of HIV-1 isolates by binding to a cluster of high-mannose sugars on the “silent” face of the gp120 envelope glycoprotein. The crystal structure of Fab 2G12 revealed two Fabs assembled into an interlocked dimer via a three-dimensional swap of their V_H domains (Figure 2F) [41]. Biochemical, biophysical, and mutagenesis data suggested that the dimer is the productive binding state that recognizes glycosylated gp120. The arrangement of the combining sites in this antibody creates an extended surface for multivalent interaction with sugar ligands. This dimeric assembly has not been observed in hundreds of Fab structures in the PDB. Although the variable (V_H and V_L) and constant regions (C_{H1} and C_L) are structurally similar to those in other Fab molecules, the V_H domain swap in 2G12 is associated with a twist of the variable regions relative to the constant region. Even though the required residues are present, the dimer lacks a highly conserved ‘ball-and-socket joint’ between V_H and C_{H1} that is believed to play a pivotal role in allowing the variable domains to adjust their position relative to the constant domains. The V_H domains within the dimer are related by a noncrystallographic two-fold symmetry axis of $\sim 180^\circ$, such that the two Fabs are arranged side-by-side with their respective combining sites separated by approximately 35 Å, facing in the same direction. From the analysis of the Fab 2G12 structure three factors emerged that may be promoting domain swapping in this case. First, the closed interface at V_H/V_L contacts may be weakened since a highly conserved interaction involving adjacent glutamines in the V_H and V_L domains is missing. These glutamines usually hydrogen bond to each other at the base of the combining site, however, in 2G12, position H39, at this site, is a rarely observed arginine that is too far away for interaction. Second, the elbow region connecting the V_H and C_{H1} domains comprises an unusual sequence that forms the hinge loop. This causes the V_H domain to pivot around residue Pro113 allowing for stabilization of a hydrophobic contact between Pro and Val that aids in domain swapping. Third, a favorable open interface between V_H/V_H domains is created by an extensive hydrogen bonding and salt bridge network with a total of 10 hydrogen bonds and 136 van der Waals interactions. In addition, π -stacking interactions occur between several aromatic residues creating a substantial buried interface [34].

Biochemical evidence indicated that the unusual neutralizing properties of the 2G12 antibody were intimately connected to the domain exchange of the V_H domains of the two Fab regions. In the structure an extensive multivalent binding surface composed of two conventional combining sites is formed by the homodimeric V_H/V_H that is able to recognize high-mannose sugars. The 2G12 V_H/V_H interface is composed of many conserved germline-encoded residues, with three uncommon mutations that appear to promote stabilization of this interaction. In principle, any IgG molecule could bivalently recognize two oligomannose chains at their tips,

but this would require a near parallel orientation of the two Fab arms which is energetically disfavored. In contrast, the 2G12 domain-exchanged structure is well-adapted for recognition of two oligomannose chains at a spacing of about 35 Å with no entropic penalty for bivalent attachment to the Fab arms, compared to a conventional antibody. In addition, the oligomeric structure of 2G12 explains the apparent high affinity for the carbohydrate antigen by providing a virtually continuous surface for multivalent recognition with interaction sites that match the geometrical spacing of the carbohydrate cluster on gp120.

Llama VHH

Some immunoglobulin isotypes of old (camels, dromedaries) or new (llamas, vicuna) world camelids lack light chains. In addition, their heavy chain is devoid of the CH1 domain. Therefore, the antigen binding fragment of the heavy-chain antibodies comprises a single domain, referred to as VHH, replacing the standard four-domain Fab fragment. This VHH domain is created by a DNA recombination event between dedicated VHH germline gene segments and D and J mini-genes. The crystal structure of a llama VHH domain (VHH-R9) exhibits a domain swapped structure (Figure 2G) [42]. It possesses the shortest CDR3 among camelid VHHs and in its first β strand the first seven residues are missing. In the structure, CDR3 and the last L-strand are swapped between symmetry-related molecules. In addition to dimer formation of the domain swapping kind, a crystal-wide β -sheet is formed. This striking feature consists of trigonal structures that are arranged around the cubic threefold axis involving N- and C-terminal segments of the chain. In this structure, six VHH molecules are linked by a dense and large network of inter-strand hydrogen bonds. Such extensive polymeric contacts would have been impossible in the presence of an intact N-terminus.

The llama VHH sequence displays mutations of four amino acids common to all camelid VHHs and also possesses two cysteines that form a conserved disulphide bridge that is present in all VH and VHH domains. Its CDR1 is nine residues long, a length typical of most VHHs, and its CDR2 contains seven residues, as seen in many camelid VHHs. In contrast, its CDR3 is only six residues long, shorter than the average size of CDR3s of camels or llamas. While the overall structure is similar in architecture to a typical VH domain, the last β -strand of each symmetry-related protein is provided by its closest neighbor. The two symmetry-related molecules are connected by an anti-parallel β -structure and resemble a dumb-bell. Considered as a whole, the incomplete VHH monomer and the last β -strand from the other monomer reconstitute a classical VHH structure, closely resembling other VHH structures.

Since no open interface is observed in the domain swapped VHH dimer and because the VHH fold is conserved, CDR3 strain release and the interactions along the extended hinge could account for most of the favorable dimerization energy. In addition, the association observed in the crystal, via numerous inter-strand hydrogen bonds within and between dimers, may stabilize dimer and higher order oligomers.

Viral capsid CTD

All retroviruses encode the gag polypeptide that directs the formation and release of immature viral particles. During virus budding, the gag precursor is cleaved into three products: matrix (MA), capsid (CA), and nucleocapsid (NC). Concomitant with gag processing, viral particle morphology changes dramatically from its immature to the mature form. Mature infectious particles of different retroviruses are morphologically distinct, characterized by the shape of their CA core structure. The mature virion is conical for HIV-1 and other lentiviruses while spherical or polyhedral for HTLV-1 and MLV.

No high resolution structural information is currently available for the arrangement of CA units within the shell of immature or mature retroviral particles. Important clues with respect to the

architecture of the viral core have come from EM studies of *in vitro* assembled virus-like structures and from models of retroviral assembly, derived from high-resolution studies of isolated CA fragments. Whether protein contacts observed in the crystals of isolated domains authentically represent the arrangement of the protein subunits in the viral shell is, however, still unclear.

In general, the overall protein structure is highly conserved among retroviral CA proteins, comprising two independently folded domains, NTD and CTD, joined by a flexible linker. Both domains of CA are predominantly α helical with the NTD comprising seven α helices and an amino-terminal β hairpin while the CTD is composed of four short α helices and a single-turn 3_{10} helix. Capsid assembly is thought to involve hexamerization of the N-terminal domain and dimerization of the C-terminal domain. Mutations in the individual domains interfere with particle assembly as well as infectivity. The nature of the CTD dimerization interface has been a subject of considerable interest and debate. In early crystallographic studies, face-to-face dimerization was observed for the CTD domain of HIV-1 and from this structure a functional dimer interface for capsid association was inferred. In the crystal dimer, the protein packs through helices 2 and 2' and mutations of this interface completely abolish CA dimerization in solution. Nevertheless, it is still not clear how this interaction can account for the critical role played by the CA-CTD in the assembly of viral particles. 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.

At present, several X-ray crystal structures of the HIV-1 capsid CTD are available and four possible arrangements of the CTD dimer have been observed [43–45]. Three of these dimers exhibit almost parallel packing of helices-2 from each monomer, albeit with slightly different crossing angles. Only the dimer arrangement of 1a43 [45] could be fitted reasonably well into a cryo EM density map of assembled capsid and the best fit was obtained using the CTD as a monomeric unit [46]. These observations indicate that conformational changes in the CTD occur during assembly. In addition, a head-to-tail dimer CA structure has been solved by crystallography [47]; this structure includes a Fab fragment that was present for crystallization.

The X-ray structure of a domain-swapped CTD dimer has recently been determined for HIV-1 CA (Figure 2A). 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 [48]. The domain swapped SCAN structure suggested that an equivalent dimerization mode may represent an alternative conformation of CA-CTD, relevant for viral assembly. Subsequent direct evidence via crystallography supported this notion [49]. A deletion mutant, $\Delta 177$ CA-CTD, exhibits a quaternary structure essentially identical to the SCAN dimer architecture. The swapped unit comprises the N terminal strand, a turn, and helix 1. The hinge region for the domain swapping is located between helices 1 and 2 and the deleted Ala177 resides in this hinge. In the dimer, a striking structural change occurs in helix 2, compared to the previously determined X-ray structures. In the original X-ray structures helix 2 is kinked and resides at the crystallographic dimer interface. In contrast, in the domain swapped dimer structure, helix 2 is a regular, straight helix. The authors suggest that straightening of helix 2 may provide the driving force for domain swapping.

In order to probe the effects of the $\Delta 177$ mutation on assembly, the deletion was introduced into the full-length CA protein and was found to assemble into core-like structures *in vitro*. However, at present, there is little experimental evidence to argue in favor of the domain-swapped model over the conventional side-by-side model of CA-CTD dimerization in the mature virus. Indeed, it may well be that formation of the swapped dimer is predominantly caused by shortening of the hinge and that 'native' capsid assembly involves head-to-head

dimer formation. Further investigation of capsid-mediated dimerization is required to test whether the domain-swapped conformation of the CA-CTD is present and/or plays a functional role in immature virions and their maturation.

Prion Protein

Mammalian prion proteins comprise approximately 210 residues with one conserved disulphide bond. The structure of the human prion protein (PrPC) was solved as a monomer in solution by NMR and also as a domain swapped dimer by crystallography (Figure 2B) [50,51]. The monomer structure consists of three α helices and a small two-stranded β sheet. Helices 2 and 3 are connected by the disulphide bond. In the crystal structure of the domain-swapped dimer the C-terminal helix is swapped. In monomeric PrPC the disulphide bond between helices 2 and 3 is intra-molecular, whereas in the domain swapped dimer it is inter-molecular; for this reason, conversion from monomer to dimer must involve reduction and re-oxidation of the disulfide bond. Interestingly, it is possible to seed the conversion of 'normal' recombinant protein into amyloid fibrils *in vitro* employing a specialized redox process [52]. This conversion process is thought to involve runaway domain swapping, with the hinge loop and adjacent regions of successive subunits forming a continuous β sheet in the centre of the fibril. In this model, the domain swapped PrPC globular domains decorate the exterior of the β sheet and may provide the characteristic templating features of prions. In the growing fibril, two free helices that bear one cysteine each, are always available. They can catalyze sulfhydryl-disulfide exchange at the surface of the fibril, thereby opening a new PrPC monomer and rendering it ready for addition to the growing fibril. Whether the prion protein ever encounters a reducing environment *in vivo* is unclear; however, there is evidence that the intramolecular disulphide bond is required for infectivity.

Antithrombin

The most recently reported domain swapped structure was described by Yamasaki *et al.* [53] for a serpin. Serpins are protease inhibitors that are at the heart of certain protein misfolding disorders - termed serpinopathies - that can cause emphysema, early-onset dementia, and liver cirrhosis. A wealth of structural work over the last 25 years has shown that inhibitory serpins can undergo dramatic conformational changes. The reactive-centre loop region of these proteins can insert itself, either after cleavage in this loop or spontaneously, as an additional β -strand (β_4) into the central β -sheet. The domain swapped dimer structure of a mutant antithrombin shows a surprising twist in this story –here, one monomer donates two long anti-parallel β -strands to another monomer, inserting them into the middle of the principal β -sheet. These two strands comprise the reactive-centre loop, which, in the polymer and the cleaved serpin, forms the β_4 -strand, as well as its neighboring β_5 -strand. Yamasaki and coauthors propose that a polymer can be created by this very long β -hairpin insertion across several molecules, resulting in a highly stable multimeric arrangement. This model of oligomerization explains the rapid propagation of serpin polymers and their extreme stability and provides new insights into their irreversible β -sheet expansion.

Conclusions

More and more domain swapped protein structures are becoming elucidated, and, for several cases, growing evidence supports that the dimer or multimer is an active, biologically important structure. Irrespective of whether domain swapping is a specific mechanism for regulation *in vivo*, it is becoming clear that domain swapping is a means by which stable multimers can be generated under evolutionary pressure.

To date, however, despite numerous efforts, no unifying molecular mechanism of domain swapping has emerged: each protein seemingly behaves in a distinctive and individual fashion,

and a general explanation for how proteins domain swap still remains elusive. What seems clear is the existence of distinct intermediates, in which some hydrophobic part of the monomeric protein becomes exposed and, thereby, is available for interaction with a 'like' molecule. In addition, the fact that high protein concentration (always present during crystallization) promotes domain swapping may indicate that a switch in solute/solvent interaction takes place: the protein itself can act as the solvent for another protein. In this manner, exposed hydrophobic regions no longer undergo unfavorable interactions with the aqueous solvent, but favorable ones with the protein solvent. Thus, domain swapped oligomers in solution involve 'melting' of the protein, with melting being distinct from thermal unfolding. Melting occurs at a much lower temperature, or can be induced by other means, and is equivalent to 'defrosting' the surrounding chains in a polymeric liquid, as described by de Gennes. In the non-ideal solution, an oligomeric structure in a 'frozen' or 'glass-like' metastable state can be trapped, and, given the extremely long relaxation times of glassy materials, such oligomers can persist or convert into the more stable amyloid forms. Analogous behavior may occur *in vivo* under conditions where "monomer" promoting factors are missing or where high local protein concentrations are promoted through compartmentalization or the action of protein-protein-interaction modules.

Acknowledgements

I thank all former and present members of my laboratory for their contributions to the GB1 and CV-N stories. Lin Liu is gratefully acknowledged for the preparation of Figure 2 and Teresa Brosenitsch for critical reading of the manuscript. Work in my laboratory is supported by the National Institutes of Health (GM082251 and GM080642).

References and Recommended reading

1. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000;28:235–242. [PubMed: 10592235]
2. Marianayagam NJ, Sunde M, Matthews JM. The power of two: protein dimerization in biology. *Trends Biochem Sci* 2004;29:618–625. [PubMed: 15501681]
3. Huang DB, Ainsworth CF, Stevens FJ, Schiffer M. Three quaternary structures for a single protein. *Proc Natl Acad Sci U S A* 1996;93:7017–7021. [PubMed: 8692936]
4. Jaffe EK. Morpheins--a new structural paradigm for allosteric regulation. *Trends Biochem Sci* 2005;30:490–497. [PubMed: 16023348] Here, the term 'morpheein' is introduced for proteins that use conformational equilibria between different tertiary structures to form distinct oligomers. Alternate morpheein forms are uniquely assembled native states that contain distinct subunit conformations, and transitions between these are used to regulate function.
5. Lawrence SH, Ramirez UD, Tang L, Fazliyez F, Kundrat L, Markham GD, Jaffe EK. Shape shifting leads to small-molecule allosteric drug discovery. *Chem Biol* 2008;15:586–596. [PubMed: 18559269]
6. Bennett MJ, Choe S, Eisenberg D. Domain swapping: entangling alliances between proteins. *Proc Natl Acad Sci U S A* 1994;91:3127–3131. [PubMed: 8159715]
7. Liu Y, Eisenberg D. 3D domain swapping: as domains continue to swap. *Protein Sci* 2002;11:1285–1299. [PubMed: 12021428] This review by Eisenberg, who introduced the term '3D domain swapping', summarizes structures of proteins involved in diverse biological functions and discusses the unique features associated with domain swapping.
8. Newcomer ME. Protein folding and three-dimensional domain swapping: a strained relationship? *Curr Opin Struct Biol* 2002;12:48–53. [PubMed: 11839489]
9. Bennett MJ, Sawaya MR, Eisenberg D. Deposition diseases and 3D domain swapping. *Structure* 2006;14:811–824. [PubMed: 16698543]
10. Bergdoll M, Eltis LD, Cameron AD, Dumas P, Bolin JT. All in the family: structural and evolutionary relationships among three modular proteins with diverse functions and variable assembly. *Protein Sci* 1998;7:1661–1670. [PubMed: 10082363]

11. Rousseau F, Schymkowitz JW, Wilkinson HR, Itzhaki LS. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc Natl Acad Sci U S A* 2001;98:5596–5601. [PubMed: 11344301]
12. Barrientos LG, Louis JM, Botos I, Mori T, Han Z, O'Keefe BR, Boyd MR, Wlodawer A, Gronenborn AM. 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. [PubMed: 12015150]
13. Chen YW, Stott K, Perutz MF. Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. *Proc Natl Acad Sci U S A* 1999;96:1257–1261. [PubMed: 9990011]
14. Ogiwara NL, Ghirlanda G, Bryson JW, Gingery M, DeGrado WF, Eisenberg D. Design of three-dimensional domain-swapped dimers and fibrous oligomers. *Proc Natl Acad Sci U S A* 2001;98:1404–1409. [PubMed: 11171963]
15. Wolynes P, Luthey-Schulten Z, Onuchic J. Fast-folding experiments and the topography of protein folding energy landscapes. *Chem Biol* 1996;3:425–432. [PubMed: 8807873]
16. Yang S, Cho SS, Levy Y, Cheung MS, Levine H, Wolynes PG, Onuchic JN. Domain swapping is a consequence of minimal frustration. *Proc Natl Acad Sci U S A* 2004;101:13786–13791. [PubMed: 15361578] In this paper, the authors test whether energy landscape principles associated with the folding of proteins into monomeric conformations can also describe how proteins oligomerize into domain-swapped conformations.
17. Cho SS, Levy Y, Onuchic JN, Wolynes PG. Overcoming residual frustration in domain-swapping: the roles of disulfide bonds in dimerization and aggregation. *Phys Biol* 2005;2:S44–55. [PubMed: 16204848]
18. Koharudin LM, Viscomi AR, Jee JG, Ottonello S, Gronenborn AM. The evolutionarily conserved family of cyanovirin-N homologs: structures and carbohydrate specificity. *Structure* 2008;16:570–584. [PubMed: 18400178]
19. Esposito L, Daggett V. Insight into ribonuclease A domain swapping by molecular dynamics unfolding simulations. *Biochemistry* 2005;44:3358–3368. [PubMed: 15736946]
20. Kundu S, Jernigan RL. Molecular mechanism of domain swapping in proteins: an analysis of slower motions. *Biophys J* 2004;86:3846–3854. [PubMed: 15189881]
21. Malevanets A, Sirota FL, Wodak SJ. Mechanism and energy landscape of domain swapping in the B1 domain of protein G. *J Mol Biol* 2008;382:223–235. [PubMed: 18588900] This paper presents an intriguing, novel mechanism for domain swapping. According to this mechanism, swapping in GB1 starts from the C-terminus of the polypeptide chain and progresses by exchanging an increasing portion of the chain until a stable conformational state is reached. In contrast to most other proposed mechanisms, this exchange process does not involve unfolding.
22. Crestfield AM, Stein WH, Moore S. On the aggregation of bovine pancreatic ribonuclease. *Arch Biochem Biophys* 1962;(Suppl 1):217–222. [PubMed: 14023810]
23. Liu Y, Hart PJ, Schlunegger MP, Eisenberg D. The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-Å resolution. *Proc Natl Acad Sci U S A* 1998;95:3437–3442. [PubMed: 9520384]
24. Liu Y, Gotte G, Libonati M, Eisenberg D. A domain-swapped RNase A dimer with implications for amyloid formation. *Nat Struct Biol* 2001;8:211–214. [PubMed: 11224563]
25. Liu Y, Gotte G, Libonati M, Eisenberg D. Structures of the two 3D domain-swapped RNase A trimers. *Protein Sci* 2002;11:371–380. [PubMed: 11790847]
26. Gotte G, Libonati M. Oligomerization of ribonuclease A: two novel three-dimensional domain-swapped tetramers. *J Biol Chem* 2004;279:36670–36679. [PubMed: 15218036]
27. Nenci A, Gotte G, Bertoldi M, Libonati M. Structural properties of trimers and tetramers of ribonuclease A. *Protein Sci* 2001;10:2017–2027. [PubMed: 11567093]
28. Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 1991;253:657–661. [PubMed: 1871600]
29. Byeon IJ, Louis JM, Gronenborn AM. A protein contortionist: core mutations of GB1 that induce dimerization and domain swapping. *J Mol Biol* 2003;333:141–152. [PubMed: 14516749]
30. Byeon IJ, Louis JM, Gronenborn AM. A captured folding intermediate involved in dimerization and domain-swapping of GB1. *J Mol Biol* 2004;340:615–625. [PubMed: 15210358]

31. Jee J, Byeon JJ, Louis JM, Gronenborn AM. The point mutation A34F causes dimerization of GB1. *Proteins* 2008;71:1420–1431. [PubMed: 18076051]
32. Louis JM, Byeon JJ, Baxa U, Gronenborn AM. 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. [PubMed: 15826664]
33. Frank MK, Dyda F, Dobrodumov A, Gronenborn AM. Core mutations switch monomeric protein GB1 into an intertwined tetramer. *Nat Struct Biol* 2002;9:877–885. [PubMed: 12379842] This paper presents the unusual tetrameric GB1 structure, solved by NMR and X-ray. The oligomeric state and quaternary structure of the protein are drastically different from the wild type GB1, and strand exchange, involving all four units, is observed. The amino acid sequence of the GB1 mutant comprises five conservative changes in hydrophobic core residues.
34. O'Neill JW, Kim DE, Johnsen K, Baker D, Zhang KY. Single-site mutations induce 3D domain swapping in the B1 domain of protein L from *Peptostreptococcus magnus*. *Structure* 2001;9:1017–1027. [PubMed: 11709166]
35. Bewley CA, Gustafson KR, Boyd MR, Covell DG, Bax A, Clore GM, Gronenborn AM. Solution structure of cyanovirin-N, a potent HIV-inactivating protein. *Nat Struct Biol* 1998;5:571–578. [PubMed: 9665171]
36. Yang F, Bewley CA, Louis JM, Gustafson KR, Boyd MR, Gronenborn AM, Clore GM, Wlodawer A. Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. *J Mol Biol* 1999;288:403–412. [PubMed: 10329150]
37. Botos I, O'Keefe BR, Shenoy SR, Cartner LK, Ratner DM, Seeberger PH, Boyd MR, 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. [PubMed: 12110688]
38. Kelley BS, Chang LC, Bewley CA. Engineering an obligate domain-swapped dimer of cyanovirin-N with enhanced anti-HIV activity. *J Am Chem Soc* 2002;124:3210–3211. [PubMed: 11916396]
39. Matei E, Furey W, Gronenborn AM. Solution and crystal structures of a sugar binding site mutant of cyanovirin-N: no evidence of domain swapping. *Structure* 2008;16:1183–1194. [PubMed: 18682220]
40. Ziolkowska NE, O'Keefe BR, Mori T, Zhu C, Giomarelli B, Vojdani F, Palmer KE, McMahon JB, Wlodawer A. Domain-swapped structure of the potent antiviral protein griffithsin and its mode of carbohydrate binding. *Structure* 2006;14:1127–1135. [PubMed: 16843894]
41. Calarese DA, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, Zhu P, Wormald MR, Stanfield RL, Roux KH, et al. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* 2003;300:2065–2071. [PubMed: 12829775] This is the first structural report concerning the human antibody 2G12 that neutralizes a broad range of HIV-1 isolates. 2G12 has an unusual architecture in which the arms of the IgG swap variable heavy domains, creating a domain-swapped dimer of Fabs. A series of follow-up studies revealed that 2G12 is capable of binding both the D1 and D3 arms of the Man₉GlcNAc₂ moiety.
42. 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. [PubMed: 15094039]
43. Gamble TR, Yoo S, Vajdos FF, von Schwedler UK, Worthylake DK, Wang H, McCutcheon JP, Sundquist WI, Hill CP. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* 1997;278:849–853. [PubMed: 9346481]
44. Ternois F, Sticht J, Duquerroy S, Krausslich HG, Rey FA. The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat Struct Mol Biol* 2005;12:678–682. [PubMed: 16041386]
45. Worthylake DK, Wang H, Yoo S, Sundquist WI, Hill CP. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr D Biol Crystallogr* 1999;55:85–92. [PubMed: 10089398]
46. Ganser-Pornillos BK, Cheng A, Yeager M. Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell* 2007;131:70–79. [PubMed: 17923088]

47. Berthet-Colominas C, Monaco S, Novelli A, Sibai G, Mallet F, Cusack S. Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *Embo J* 1999;18:1124–1136. [PubMed: 10064580]
48. Ivanov D, Stone JR, Maki JL, 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. [PubMed: 15629724]
49. Ivanov D, Tsodikov OV, Kasanov J, Ellenberger T, Wagner G, Collins T. Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc Natl Acad Sci U S A* 2007;104:4353–4358. [PubMed: 17360528] Here, HIV-1 capsid dimerization, via domain swapping, is proposed as a critical event in immature viral particle assembly. The authors suggest that the interlocked arrangement of viral capsid subunits, conferred by domain swapping, has several advantages over more conventional side-by-side interactions between structural modules.
50. Zahn R, Liu A, Luhrs T, Riek R, von Schroetter C, Lopez Garcia F, Billeter M, Calzolari L, Wider G, Wuthrich K. NMR solution structure of the human prion protein. *Proc Natl Acad Sci U S A* 2000;97:145–150. [PubMed: 10618385]
51. Knaus KJ, Morillas M, Swietnicki W, Malone M, Surewicz WK, Yee VC. Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nat Struct Biol* 2001;8:770–774. [PubMed: 11524679]
52. Lee S, Eisenberg D. Seeded conversion of recombinant prion protein to a disulfide-bonded oligomer by a reduction-oxidation process. *Nat Struct Biol* 2003;10:725–730. [PubMed: 12897768] This elegant experimental study shows that recombinant hamster prion protein can be converted by a redox process, *in vitro*, into PrP^{RDX}, which can seed the conversion of other PrP^C molecules to the PrP^{RDX} form. The converted protein resembles PrP^{Sc}, and X-ray fiber diffraction suggests an amyloid-like structure for the oligomerized prion protein.
53. Yamasaki M, Li W, Johnson DJD, Huntington JA. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 2008;455:1255–1258. [PubMed: 18923394]

Figure 1.
Schematic representation of features inherent to domain swapped structures.

Leger lithographs –
acrobats engage in a variety
of domain swaps

Figure 2.

Selected examples of domain swapped protein structures. A–E: All alpha proteins. (A) HIV-1 capsid; (B) Human Prion Protein; (C) T-SNARE; (D) Designed helical bundle; (E) Bcl-xl. F–J: All beta proteins. (F) 2G12 antibody; (G) Lama VHH Domain; (H) β -Crystallin; (I) Cyanovirin-N; (J) Griffithsin (GRFT).

K–T: Mixed alpha/beta proteins. (K) RNase A C-terminal dimer; (L) RNase A N-terminal dimer; (M) RNase A C-terminal trimer; (N) FOXP2; (O) Guanine deaminase; (P) B1 domain of protein L; (Q) B1 domain of protein G, dimer; (R) B1 domain of protein G, tetramer; (S) Cystatin C; (T) amyloid-like Cystatin C.

The PDB IDs for the structures are (A) 2ONT; (B) 1I4M; (C) 2C5J; (D) 1G6U; (E) 2B48; (F) 1OM3; (G) 1SJV; (H) 1BLB; (I) 3EZM; (J) 2GTY; (K) 1F0V; (L) 1A2W; (M) 1JS0; (N) 2A07; (O) 1WKQ; (P) 1K50; (Q) 1Q10; (R) 1MPE; (S) 1G96; (T) 1TIJ.