Protein Science (1995), *4*:2455–2468. Cambridge University Press. Printed in the USA.

REVIEW

3D Domain swapping: A mechanism for oligomer assembly

MELANIE J. BENNETT,[1,2] MICHAEL P. SCHLUNEGGER,[2] AND DAVID EISENBERG[2]

[1] Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059
[2] Molecular Biology Institute, Department of Chemistry and Biochemistry, and UCLA-DOE Laboratory of Structural Biology and Molecular Medicine, University of California–Los Angeles, Los Angeles, California 90095-1570

(RECEIVED September 12, 1995; ACCEPTED October 6, 1995)

Abstract

3D domain swapping is a mechanism for forming oligomeric proteins from their monomers. In 3D domain swapping, one domain of a monomeric protein is replaced by the same domain from an identical protein chain. The result is an intertwined dimer or higher oligomer, with one domain of each subunit replaced by the identical domain from another subunit. The swapped "domain" can be as large as an entire tertiary globular domain, or as small as an α-helix or a strand of a β-sheet. Examples of 3D domain swapping are reviewed that suggest domain swapping can serve as a mechanism for functional interconversion between monomers and oligomers, and that domain swapping may serve as a mechanism for evolution of some oligomeric proteins. Domain-swapped proteins present examples of a single protein chain folding into two distinct structures.

Keywords: aggregation; complementation; oligomer evolution; protein dimerization

Since Svedberg's discovery of functional molecules composed of two or more identical protein chains, much effort has been expended in studying their metabolic regulation (Monod et al., 1965; Koshland et al., 1966) and their assembly and disassembly (Kikuchi & King, 1975; Caspar, 1980; Jaenicke, 1995).

Despite this progress, understanding the assembly of oligomeric proteins from monomers remains a challenge. A common observation is that disassembly of an oligomeric protein into its monomeric subunits is accompanied by irreversible unfolding and aggregation. This observation is often interpreted in terms of exposing apolar patches on the monomer surface that are covered in the oligomer, thereby providing binding energy from a hydrophobic interaction. Thus, the question remains of how the oligomer could have been assembled in the first place. We propose an answer to this question for some oligomers based on a mode of association that we have noticed in several proteins of known structure. We term this mode of association *3D domain swapping*, because oligomers are formed from stable monomers by exchanging domains.

A problem related to the formation of oligomeric proteins in a cell is the problem of how oligomeric proteins evolved from monomeric precursor proteins. For an oligomer to evolve, random mutations must change the surface of the monomer so that sufficient free energy is released upon oligomerization to overcome the accompanying entropy loss of immobilizing the monomers. As we discuss in this review, single amino acid replacements must be fortuitous to provide an adequate free energy of interaction. But an evolutionary pathway involving domain swapping can easily lead to stable oligomers.

In this review, we present several examples of domain-swapped proteins that demonstrate 3D domain swapping is a mechanism used by nature to switch monomers to oligomers. We then discuss the case of domain-swapped diphtheria toxin in greater detail, to provide insight into the mechanism of oligomer formation (including destabilization of the monomer by a pH switch and formation of domain-swapped dimers at high protein concentration). We then suggest a possible pathway for oligomer evolution via 3D domain swapping and provide examples. For simplicity, we focus on dimers; however, higher oligomers also may form by domain swapping. We also note that unregulated domain swapping can lead to protein precipitates and even to the abnormal aggregates that characterize amyloidosis.

Reprint requests to: David Eisenberg, Molecular Biology Institute, Department of Chemistry and Biochemistry, and UCLA-DOE Laboratory of Structural Biology and Molecular Medicine, University of California–Los Angeles, Los Angeles, California 90095-1570; e-mail: david@pauling.mbi.ucla.edu.

Abbreviations: BS-RNase, bovine seminal ribonuclease; DT, diphtheria toxin; GM-CSF, granulocyte-macrophage colony-stimulating factor; GST, glutathione *S*-transferase; IF, interferon; IL, interleukin; RNase, ribonuclease.

Definitions

The following definitions are helpful in discussing 3D domain swapping.

Swapped domain

A swapped domain in a protein oligomer is a globular domain (or sometimes one or a few elements of secondary structure) that is intertwined with an identical protein chain, with the swapped domain having an environment essentially identical to that of the same domain in a protein monomer.

Hinge loop

A segment of polypeptide chain that links the swapped domain to the rest of its subunit is a hinge loop. Hinge loops have different conformations in monomers and domain-swapped oligomers.

C-interface

A C-interface occurs between domains in a monomeric subunit (closed monomer) (Fig. 1A, left).

3D domain-swapped dimer

A dimer with a two C-interfaces between two different subunits (open monomers) is a 3D domain-swapped dimer (Fig. 1A, right).

O interface

An O-interface occurs between open monomers in a 3D domain-swapped dimer, NOT present in the closed monomer (Fig. 1B, left).

Examples of 3D domain swapping in proteins

We first described 3D domain swapping as it occurs in metastable dimers of diptheria toxin (DT) formed by freezing in phosphate buffer (Bennett et al., 1994b) (Kinemage 2). However, the structures of other proteins that undergo domain swapping have also been reported (Table 1; Fig. 2). These are well-characterized examples of domain swapping; the structures of several entries in Table 1 have been determined in both monomeric and dimeric states by X-ray crystallography (DT, CD2 [Kinemage 3], and staphylococcal nuclease). Of the remaining entries, the high sequence similarity of the RNases (>80%) makes it likely that monomeric BS-RNase and dimeric RNase A are similar to the X-ray structures of monomeric RNase A and dimeric BS-RNase, respectively. The CksHs2 monomer was observed in solution and its inferred structure is the same as the crystallographic structure of CksHs1 (81% identical) (Arvai et al., 1995). The α-spectrin dimer structure was determined crystallographically and the monomer was observed in solution (Yan et al., 1993). Single-chain Fv molecules form monomers, dimers, and higher oligomers in solution (Raag & Whitlow, 1995). The hinge loop is disordered in the structure of the anti-sialidase single chain Fv molecule, but close interactions between molecules in the crystal suggest it is a domain-swapped dimer (Kortt et al., 1994).

Domain-swapped proteins have a C-interface (see Definitions), generally with many specific interactions. Each pair of proteins in Table 1 has an identical C-interface in the monomeric and dimeric states (Fig. 2), formed between domains linked by a hinge loop. The length of the swapped domains varies greatly from a 150-residue globular domain in DT, to four β-strands in CD2, to a single α-helix of only 15 residues in BS-RNase. However, the buried areas and polar interactions in the C-interfaces are more uniform and extensive, from 850 Å^2 to 3,800 Å^2 and involving 8–19 hydrogen bonds and/or salt bridges.

In addition to illustrating that domain swapping occurs in a variety of unrelated proteins, Table 1 includes domain-swapped dimers with functions altered from their monomers, consistent with the hypothesis that domain swapping can be a means of regulating activity. We discuss two of these proteins in the following.

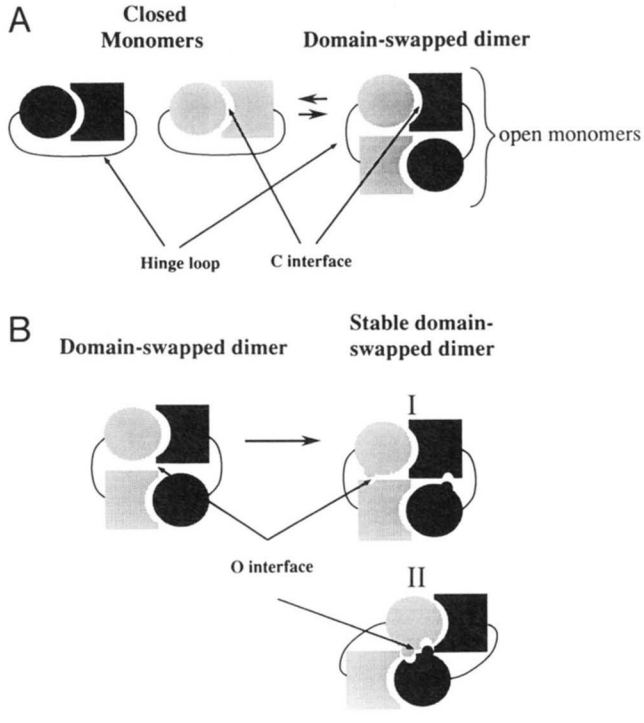

Fig. 1. 3D domain swapping definitions. **A:** Closed monomers are comprised of tertiary or secondary structural domains (represented by a circle and square) linked by polypeptide linkers (hinge loops). The interface between domains in the closed monomer is referred to as the C- (closed) interface. Closed monomers may be opened by mildly denaturing conditions or by mutations that destabilize the closed monomer (see text). Open monomers may dimerize by domain swapping. The domain-swapped dimer has two C-interfaces identical to those in the closed monomer, however, each is formed between a domain from one subunit (black) and a domain from the other subunit (gray). The only residues whose conformations significantly differ between the closed and open monomers are in the hinge loop. Domain-swapped dimers that are only metastable (e.g., DT, CD2, RNase A) may convert to monomers, as indicated by the backward arrow. **B:** Over time, amino acid substitutions may stabilize an interface that does not exist in the closed monomers. This interface formed between open monomers is referred to as the O- (open) interface. The O-interface can involve domains within a single subunit (I) and/or between subunits (II).

RNase A and BS-RNase

RNase (Kinemage 3) is a 124-residue protein isolated from the pancreas of animals, and from bovine seminal plasma. BS-RNase

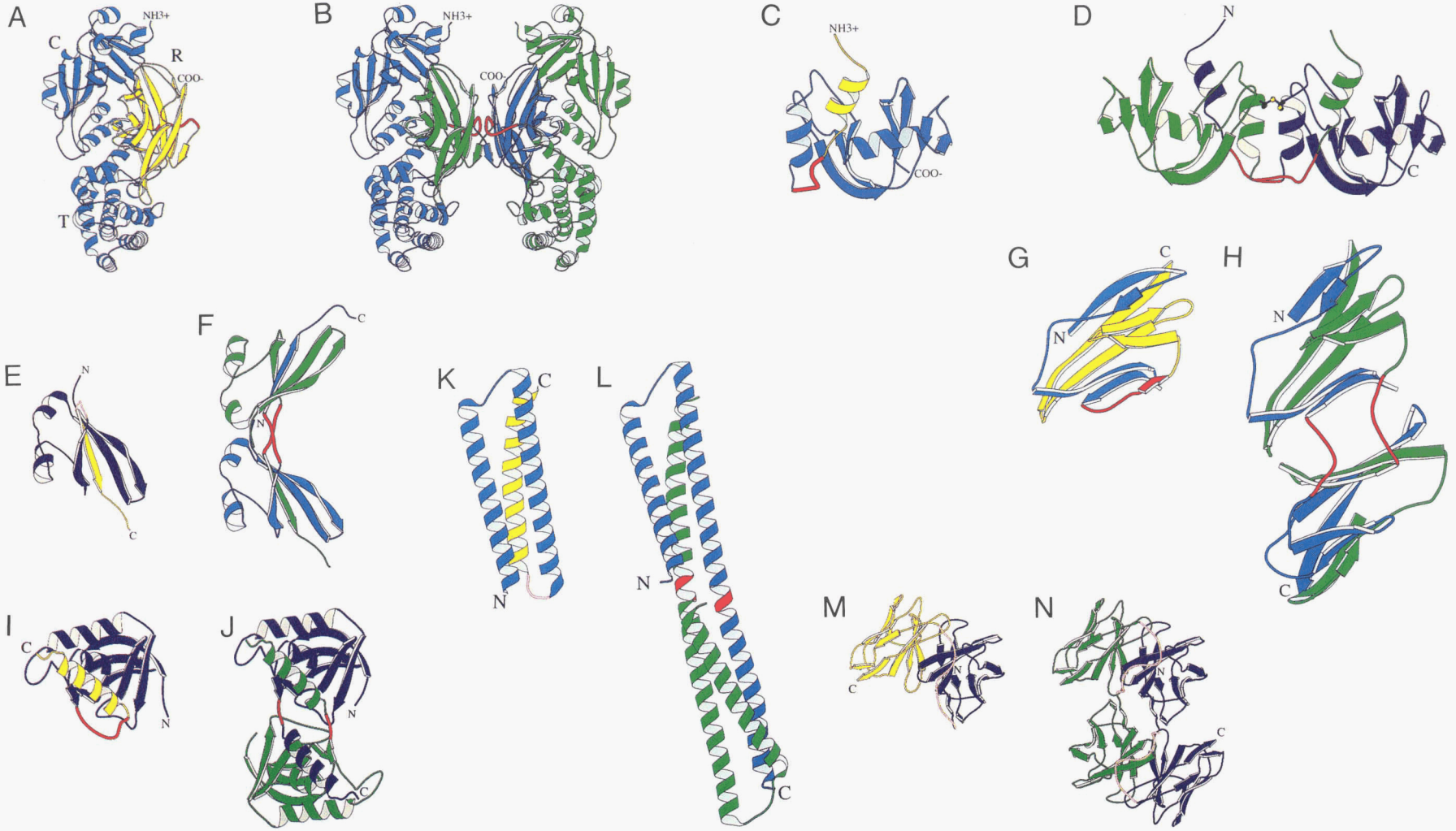

Fig. 2. Monomeric and domain-swapped dimeric proteins in Table 1. The domain that can be swapped is highlighted (yellow) in the monomer structures. The two subunits in the dimer structures are blue and green. N and C termini are labeled for clarity. The only residues whose conformations significantly differ between monomers and dimers are in the hinge loops (red) (Table 1). Where they occur, interchain disulfides in the dimers are shown in ball-and-stick representation. The broken-line hinge loops in CksHs2 (E) and α-spectrin (K) indicate that the monomer structure is only a model based upon the crystallographically determined dimer, so the exact position of the hinge loop is not known. The broken-line hinge loops in the single-chain Fv monomer and dimer (M and N) indicate that this portion of the polypeptide chain is disordered in the crystal structure, so its exact position is not known. **A:** DT monomer at 2.3 Å resolution (Bennett & Eisenberg, 1994). **B:** DT dimer at 2.0 Å resolution (Bennett et al., 1994a). **C:** Bovine pancreatic RNase A at 2.0 Å resolution (Wlodawer et al., 1982). **D:** BS-RNase at 1.9 Å resolution (Mazzarella et al., 1993). **E:** Recombinant human CksHs2 monomer (J.A. Tainer, pers. comm.). **F:** Recombinant human CksHs2 dimer at 2.1 Å resolution (Parge et al., 1993). **G:** N-terminal portion of the soluble domain of rat CD2 at 2.8 Å resolution (Jones et al., 1992). **H:** Recombinant rat CD2 N-terminal domain dimer at 2.0 Å resolution (Murray et al., 1995). **I:** Staphylococcal nuclease monomer at 1.7 Å resolution (Loll & Lattman, 1989). **J:** Staphylococcal nuclease dimer (spontaneous deletion mutant) at 2.0 Å resolution (Green et al., 1995). **K:** α-Spectrin monomer model with a hypothetical hinge loop (modeled by the present authors). **L:** Recombinant *Drosophila* α-spectrin (single repeat) dimer at 1.8 Å resolution (Yan et al., 1993). **M:** Anti-sialidase single-chain Fv molecule at 3.0 Å resolution (Kortt et al., 1994) with a hypothetical hinge loop modeled (by the present authors) to form a closed monomer. **N:** Anti-sialidase single-chain Fv molecule at 3.0 Å resolution (Kortt et al., 1994) with a hypothetical hinge loop modeled (by the present authors) to form an open monomer in a domain-swapped dimer. Drawings were made with MOLSCRIPT (Kraulis, 1991).

Table 1. *Domain-swapped proteins*[a]

Protein (Reference)	PDB code	Residues per subunit[b]	Residues in swapped domain[c]	Residues in hinge loop[d]	Area buried in C-interface[e] (Å^2)	Conserved H bonds and salt bridges in C-interface[f]	RMSD monomer vs. dimer[g] (Å)	Function	Stability[h]	Domain exchanged
DT monomer (Bennett & Eisenberg, 1994)	1MDT	535						ADP-ribosylating toxin	Stable	
DT dimer (Bennett et al., 1994a)	1DDT	535	148	379–387	3,800	14	0.7 (485)	Unknown[i]	Metastable	Globular
BS-RNase monomer (Piccoli et al., 1992)	N.A.	124						Ribonuclease	Unstable[j]	
BS-RNase dimer (Mazzarella et al., 1993)	1BSR	124	14	15–22	1,690[k]	11[k]	0.4 (103)[k]	Ribonuclease	Stable	α-Helix
RNase A monomer (Wlodawer et al., 1982)	5RSA	124						Ribonuclease	Stable	
RNase A dimer (Crestfield et al., 1962)	N.A.	124	14	15–22	1,690[k]	11[k]	0.4 (103)[k]	Ribonuclease	Metastable	α-Helix
CksHs2 monomer	N.A.	79						Cell cycle regulation	Stable[l]	
CksHs2 dimer (Parge et al., 1993)	1CKS	79	14	61–65	2,130	N.A.	N.A.	Cell cycle regulation	Stable[l]	β-Strand
Soluble domain CD2 monomer (Jones et al., 1992)	1HNG	177						Binding domain of lymphocyte adhesion protein	Stable	
N-terminal domain CD2 dimer (Murray et al., 1995)	1CDC	99	43	44–50	2,970	19	0.6 (89)	Unknown[i]	Metastable	4 β-Strands
Staphylococcal nuclease monomer (Loll & Lattmann, 1989)	1SNC	149						Nuclease	Stable	
Deletion mutant staphylococcal nuclease dimer (Green et al., 1995)	[m]	143	29	112–120	1,380	8	0.6 (121)	Unknown[i]	Stable	α-Helix

α-spectrin monomer	N.A.	107						Single repeat of cytoskeletal protein		
α-spectrin dimer (Yan et al., 1993)	2SPC	107	31	72–75	2,170	N.A.	N.A.	Unknown[i]	?	α-Helix
Anti-sialidase single chain Fv monomer	N.A.	252						Single chain variable domain Ig	Stable	
Anti-sialidase single chain Fv dimer (Kortt et al., 1994)	[m]	252	109	123–137	850	N.A.	N.A.	Single chain variable domain Ig	Metastable	Globular

[a] These pairs of structures satisfy the definition of domain-swapped proteins. In column 4, the size of the swapped domain is given. Column 8 shows that the structures of the closed monomers superimpose well with the structures of the open monomers. Columns 6 and 7 show that the swapped domains form large interfaces with the complementary domains on other subunits, and make numerous specific interactions with them. Underlined structures were determined by X-ray crystallography.

[b] Number of residues in the protein; some residues are not included in the PDB coordinates because of disorder.

[c] The smallest of the domains defined between the N or C terminus, and the N- or C-terminal edge of the hinge loop, respectively.

[d] Hinge loops were initially assigned as segments that are not superimposable between monomers and dimers: DT, 379–387; RNase, 16–22; CD2, 45–49; staphylococcal nuclease, 113–119. Hinge loops were extended at both ends to include residues with Φ,Ψ differences >20° until 2 consecutive residues were below the cutoff. Φ,Ψ differences were calculated as $(\Delta\Phi^2 + \Delta\Psi^2)^{1/2}$. For structures with more than one molecule in the asymmetric unit, the average Φ,Ψ difference was used. Exceptions for which the hinge loops do not meet these criteria include: CksHs2 and α-spectrin because the monomer structure is only a model, the hinge loop is estimated; single-chain Fv structure because the hinge loop is disordered.

[e] Surface area per subunit buried in the interface calculated by subtracting the solvent-accessible surface area of the closed monomer from that of the open monomer (one subunit of the dimer structure). Solvent-accessible surface areas were determined using the program HYDROPHOBIC_ENERGY (Eisenberg et al., 1989). A closed CksHs2 monomer model was provided by J.A. Tainer. A closed α-spectrin monomer model was generated from the open monomer by combining residues 1–71 of molecule 1 and residues 76–106 of molecule 2. A closed Fv monomer model was generated from the open monomer by combining residues VL1–VL109 of molecule 1 and VH1–VH122 of molecule 2.

[f] Contacts between domains were calculated using the program CONTACT (CCP4, 1994), with a 4.0-Å cutoff for salt bridges and a 3.5-Å cutoff for hydrogen bonds. Angular cutoffs were 120° at the hydrogen atom and 90° at the oxygen atom.

[g] Structural alignments and best-fit superpositions were carried out using the program ALIGN (Satow et al., 1986). Pseudo-closed monomers were generated from the open monomers in the dimeric structures by combining coordinates from two subunits: DT, residues 1–385 of molecule 1 and residues 386–535 of molecule 2; BS-RNase, residues 1–22 of molecule 1 and residues 23–124 of molecule 2; CD2, residues 48–99 of molecule 1 and residues 4–44 of molecule 2; single-chain Fv, residues VL1–VL109 of molecule 1 and residues VH1–VH122 of molecule 2; staphylococcal nuclease, residues 1–113 of molecule 1 and residues 120–141 of molecule 2. The Cα RMS difference between monomeric and dimeric DT is based upon the superposition of residues 1–39, 41–187, 201–349, 354–378, 388–405, 412–435, 440–497, 506–517, and 519–531. The Cα RMS difference between BS-RNase and RNase A is based upon the superposition of residues 1–15, 23–37, 39–65, 71–87, 95–113, and 115–124 of both. The Cα RMS difference between monomeric and dimeric CD2 is based upon the superposition of residues 4–44 and 50–97. The Cα RMS difference between monomeric and dimeric staphylococcal nuclease is based upon the superposition of residues 7–42, 44–45, 51–79, 81–112, and 120–141 of both. Numbers in parentheses indicate the number of matched Cα atoms used in the calculation of RMS differences.

[h] A monomeric or dimeric state of a protein is classified as stable (predominant state at equilibrium) or metastable (long-lived and separable by standard purification techniques, but slowly converts to the stable state).

[i] This oligomeric state is not found in nature and has no known function.

[j] Monomeric BS-RNase can be stabilized by the presence of dithiothreitol, which prevents formation of intersubunit disulfides.

[k] Interactions are compared between dimeric BS-RNase and monomeric Rnase A.

[l] Purified CksHs2 exists as monomers, dimers, and hexamers in varying proportions depending upon pH and ion concentration (Parge et al., 1993). In vivo, binding of cyclin-dependent kinases and metal ions may differentially stabilize monomers, dimers, and hexamers.

[m] Coordinates supplied by original authors prior to PDB deposition.

is a dimer with >80% sequence identity (Suzuki et al., 1987) and extensive structural homology (Mazzarella et al., 1993) to monomeric pancreatic RNase A (Richards & Wyckoff, 1973). BS-RNase can be isolated as a monomer, which is assumed to have a structure virtually identical to that of RNase A on the basis of its CD spectrum, activity and reactivity with anti-RNase A serum (Blackburn & Moore, 1982). As shown in Figure 2C, RNase A contains two β-sheets and three α-helices. The crystal structure of dimeric BS-RNase (Mazzarella et al., 1993), shown in Figure 2D, reveals that each half of the dimer has a fold virtually identical to that of RNase A. However, the fold is composed of the N-terminal segment (residues 1–14) of one subunit and the "body" (residues 23–124) of the other subunit in the dimer. Therefore, BS-RNase can be viewed as having undergone 3D domain swapping, in which the N-terminal segment is the swapped domain.

Domain swapping endows the BS-RNase dimer with novel properties. Domain swapping yields two composite active sites in which histidine residues 12 and 119 from different subunits contribute to a single active site. This provides the potential for communication between the active sites, and hence the possibility of allosteric regulation. In fact, although dimeric BS-RNase has lower activity, it displays mixed cooperativity, in contrast to monomeric BS-RNase and RNase A (Piccoli et al., 1988). In addition, domain-swapped dimeric BS-RNase displays selective toxicity for tumor cells, whereas monomeric BS-RNase and RNase A do not (Cafaro et al., 1995). Thus, domain-swapped BS-RNase differs both in structure and function from monomeric BS-RNase, even though both are formed from the same protein chain.

CksHs2

CksHs2 is a cell cycle regulatory protein of 79 residues that binds to cyclin-dependent kinases and is essential for their functioning. The crystal structure of dimeric CksHs2, shown in Figure 2F, consists of a four-stranded antiparallel β-sheet with two short helices (Parge et al., 1993). The CksHs2 monomer is believed to have the same fold as one half of the dimer, as shown in Figure 2E, which is the same as the X-ray crystallographic structure of the related (81% identical) monomeric protein CksHs1 (Arvai et al., 1995). In the dimer, the fourth β-strand is extended and exchanged with the identical strand from the other subunit. Therefore, the CksHs2 dimer forms by 3D domain swapping, in which the carboxy-terminal segment (residues 66–79) is the swapped domain.

The domain-swapped dimer is stabilized by metal ion binding as follows. Glu 63 in the hinge loop of each subunit contributes to a metal binding site in the domain-swapped dimer, but not in the CksHs2 monomer (Parge et al., 1993). In the absence of metals, Glu 63 from both subunits would electrostatically repel one another if the dimer formed. Thus, at neutral pH, metal ion binding stabilizes the dimer by compensating this repulsion, whereas at low pH (as in the crystal), the dimer is stable even in the absence of metal ions.

Ultimately, domain swapping in CksHs2 allows ion-regulated assembly of a hexamer that is believed to be the functional unit of this protein (Parge et al., 1993). In the crystals of CksHs2, noncrystallographic three-fold symmetry yields a hexamer from three crystallographically related dimers. Not only does this hexamer expose conserved residues that may be involved in binding cyclin-dependent kinases, but it can also be docked to six copies of a kinase structure homologous to the cyclin-dependent kinases, suggesting that this complex may form in vivo. Based on the crystal structure, it was suggested that ion-regulated dimerization produces structures whose interaction promotes hexamer assembly, consistent with solution studies showing that metal ion binding facilitates hexamerization, whereas the metal ion chelator EDTA prevents it (Parge et al., 1993).

Mechanisms of 3D domain swapping

The pathways leading to domain-swapped dimers are unknown for proteins that undergo dimerization during folding (e.g., BS-RNase and spectrin). However, a simple mechanism can be described for some stable monomers that interconvert with metastable domain-swapped dimers. In this mechanism, the folded monomer is temporarily subjected to conditions that favor open monomers. Then, when closed monomer-stabilizing conditions are restored, the open monomers may swap domains with one or more subunits, particularly if the protein concentration is high. Domain swapping was implicated more than 20 years ago in the formation of dimers and higher aggregates of RNase A (Crestfield et al., 1962), tryptophan synthetase α subunit (Jackson & Yanofsky, 1969), and tryptophanase (London et al., 1974) upon denaturation and subsequent renaturation from urea or guanidine. Unlike these early domain-swapped proteins, the three-dimensional structure of DT has been determined in both the monomeric and dimeric states. DT dimerization is well understood and serves an example of this mechanism, discussed below.

DT dimerization by pH pulse

Diptheria toxin (DT) is a 535-residue protein toxin secreted from a bacterium that causes the disease diphtheria (Collier, 1975). Figure 2A shows the structure of monomeric DT that has three domains labeled C (catalytic), T (transmembrane), and R (receptor-binding). DT is active as a monomer and does not spontaneously convert to dimer even at high concentrations (30 mg/mL). However, DT can be dimerized by freezing in mixed phosphate buffers, which causes a drop in pH from neutral to about 3.6 (van den Berg & Rose, 1959). The decrease in pH converts monomeric DT into an open monomer, in which a globular domain (the R domain) is dissociated from the body of the subunit (Fig. 3, middle). The open monomers form domain-swapped dimers and higher oligomers at the high concentration of the mixture as the pH returns to normal during thawing (Bennett et al., 1994b).

The formation of a domain-swapped dimer such as DT does not release much free energy because the same noncovalent bonds in the C-interface (see Definitions) of the closed monomer (Fig. 3, left) are present in the C-interfaces of the domain-swapped dimer (Fig. 3, right). Also, the loss of rotational and translational entropy of the two monomers in forming the dimer will increase the free energy. However, domain swapping can trap a metastable DT dimer, as shown in Figure 3. The Gibbs free energy of dimerization, $\Delta G^{\circ}_{\text{dimerization}}$, is positive, as shown by the observation that 5–10% of dimeric DT dissociates to monomer over several weeks at 4 °C (Carroll et al., 1986). Monomeric DT converts to dimer only in the conditions of very high concentration and low pH generated by freezing in phosphate buffer, as discussed above. At neutral pH, monomeric DT

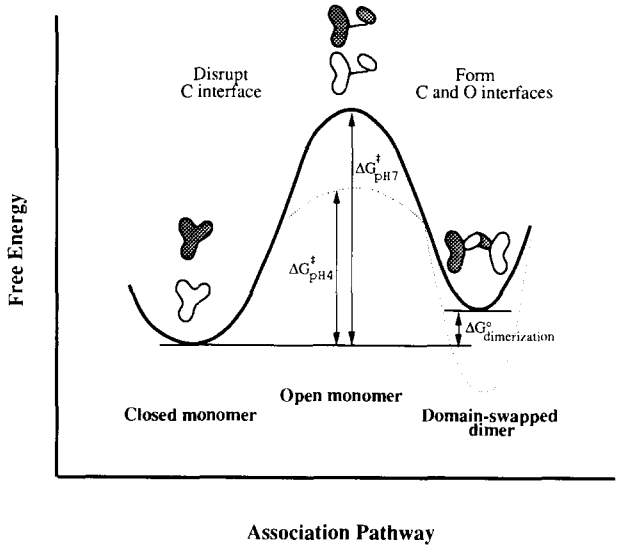

Fig. 3. Free energy relationships in the formation of domain-swapped dimeric DT (right-hand minimum) from monomeric DT (left-hand minimum) as estimated from experiments and calculations. The Gibbs free energy of formation, $\Delta G^{\circ}_{dimerization}$, is positive as shown by the observation that 5–10% of dimer dissociates to monomer over several weeks (Carroll et al., 1986). We calculate a value of 9 ± 2 kcal/mol of monomer for $\Delta G^{\circ}_{dimerization}$ as a sum of two terms: $-T\Delta S^{\circ} + \Delta G^{\circ}_{solvation}$, in which ΔS° is the entropy change of rotation and translation and $\Delta G^{\circ}_{solvation}$ is the free-energy change of solvation of the surfaces in the dimer. Following Erickson (1989), we estimate $T\Delta S^{\circ} = -10 \pm 1$ kcal/mol of monomer at 277 K, and from atomic solvation energies (Eisenberg & McLachlan, 1986), we estimate $\Delta G^{\circ}_{solvation} = -1 \pm 1$ kcal/mol of monomer. The free-energy barrier between monomer and dimer can be estimated from experimental rates of dimer dissociation. We estimate a value for $\Delta G^{\ddagger}$, the activation energy of converting dimer to open monomer, of 27 ± 1 kcal/mol of dimer from the measured rate of dimer dissociation to monomers (Carroll et al., 1986) and the Eyring rate equation, assuming first-order kinetics and that the transmission coefficient is unity. Reprinted in modified form from Bennett et al. (1994b).

does not dimerize and dimeric DT dissociates only slowly, indicating that the activation barrier, $\Delta G\ddagger$, is large. Thus, although the domain-swapped DT dimer is thermodynamically unstable, the large activation barrier endows it with a long lifetime. The dotted-line right-hand well in Figure 3, with a negative $\Delta G^{\circ}_{dimerization}$, represents the case of stable domain-swapped dimers, such as IL-5 and β B2-crystallin. A metastable domain-swapped dimer could evolve into a stable dimer by the accumulation of stabilizing amino acid replacements at the O interface as shown in Figure 1B. Alternatively, a stable domain-swapped dimer could be formed by a mutation in the hinge loop that destabilized the monomer (discussed below).

Mechanisms for domain swapping in vivo

Although the observation of domain swapping under transient denaturing conditions in vitro can be considered an experimental artifact, it is consistent with the possibility that stable monomeric proteins might undergo 3D domain swapping in vivo. This could occur if association of folded domains is slow relative to domain folding or if proteins encountered transient denaturing conditions, for example in an acidic compartment. As an example of dimerization during folding, the N-terminal domain of the cell adhesion molecule CD2 undergoes domain swapping when it is expressed as a fusion protein with GST (Murray et al., 1995). The tethering of two CD2 molecules to a GST dimer during folding permits domain swapping, perhaps mimicking interactions between folding intermediates at high concentrations in the cell. That is, 3D domain swapping might occur in the folding pathways of proteins if intermediates accumulate that have folded but unassociated domains.

Finally, we note that domain swapping may play a role in amyloidosis, a set of diseases characterized by the deposition of protein aggregates. Often, the proteins involved in amyloid aggregates are stable in both monomeric and oligomeric forms, as are domain-swapped proteins. Domain swapping was proposed to be involved in the formation of amyloid-like fibrils in vitro from a Bence-Jones protein (Klafki et al., 1993). The mechanism for forming amyloid fibrils is not yet well understood, but partial denaturation was shown to be sufficient for amyloid formation of transthyretin in vitro (Colon & Kelly, 1992).

Evolution of oligomers

We have documented in Table 1 that domain swapping occurs in a several different proteins, leading in some cases to oligomers with different activities from the monomers. Thus far, we have limited our discussion of domain swapping to dimer/monomer pairs having the same amino acid sequence (Table 1). Now, we will describe two models for oligomer evolution and extend the discussion to include dimer/monomer pairs having different sequences and a common ancestor, to support the idea that domain swapping may play a role in oligomer evolution (Table 2; Fig. 4).

Two models for oligomer evolution

What are the characteristics of a stable oligomer interface and how might such an interface evolve? For simplicity, we focus on dimers in the following discussion. Dimer interfaces are extensive (700–5,000 Å^2/subunit), close-packed, and generally include several hydrogen bonds (Janin et al., 1988). A primitive dimer would have required a stable interface that yielded a free energy of interaction at least equal to the free energy associated with the loss of rotational and translational entropy upon dimerization, which has been estimated as approximately 15 kcal/mol (Erickson, 1989; Finkelstein & Janin, 1989). In the following, we outline two possible models for evolving such a primitive dimer interface, which we call the "sticky billiard ball model" and the "3D domain swapping model."

The simplest view of dimer formation can be termed the sticky billiard ball model: a primitive dimer might arise from the genesis of an interface on the monomer surface by the replacement of residues via random mutations. These amino acid replacements must have the correct chemical structure to bind residues in the other subunit and must be fully accessible at the subunit surface to bind productively upon collision with another subunit. However, the occurrence of several amino acid substitutions on the monomer surface that form a dimer interface is very unlikely in a single generation. Thus, the sticky billiard ball model can account easily for only those dimerizations in which a single mutation produces a stable dimer. However, a single mutation is unlikely to provide the large surface area and multiple specific interactions that characterize stable interfaces. In spite of this seeming contradiction, some intracellular dimers

Table 2. *Monomers and dimers related by divergent evolution*[a]

Protein (Reference)	PDB code	Residues in protein[b]	Residues in swapped domain[c]	Residues in hinge loop[d]	Sequence identity between monomer and dimer	RMSD monomer vs. dimer[e]	Function	Domain exchanged
RNase A monomer (Wlodawer et al., 1982)	5RSA	124		15–22			Ribonuclease	
BS-RNase dimer (Mazzarella et al., 1993)	1BSR	124	14	15–22	83%	0.4 (103)	Ribonuclease	α-Helix
γB crystallin monomer (Summers et al., 1984)	4GCR	174		79–87			Eye lens protein	
βB2 crystallin dimer (Lapatto et al., 1991)	2BB2	204	97	79–87	36%	1.2 (167)	Eye lens protein	Globular
GM-CSF monomer (Diederichs et al., 1991)	1GMF	127		87–99			Granulocyte-macrophage growth factor	
IL-5 dimer (Milburn et al., 1993)	[f]	113	26	82–89	20%[g]	1.9 (73)	B & T cell growth factor	α-Helix and β-strand
IFN-βmonomer (Senda et al., 1992)	1RMI	160		97–114			Fibroblast interferon	
IL-10dimer (Zdanov et al., 1995)	1ILK	160	46	108–118	20%[h]	2.2 (94)	Cytokine inhibitory synthesis factor	2 α-Helices

[a] Each pair of structures resembles a pair of domain-swapped structures (as in Table 1), but the sequences of these pairs are not identical. These pairs support the hypothesis that domain swapping has occurred during the evolution of oligomers. All structures were determined by X-ray crystallography.

[b] Number of residues in the protein; some residues are not included in the PDB coordinates because of disorder.

[c] The smallest of the domains defined between the N or C terminus and the N- or C-terminal edge of the hinge loop, respectively.

[d] Hinge loops were initially assigned as segments that are not superimposable between monomers and dimers: RNase, 16–22; crystallins, 87; IL-5, 84–86; and GM-CSF, 89–96. Hinge loops were extended at both ends to include residues with Φ,Ψ differences of $>30°$ until two consecutive residues were below the cutoff. Φ,Ψ differences were calculated as $(\Delta\Phi^2 + \Delta\Psi^2)^{1/2}$. For structures with more than one molecule in the asymmetric unit, the average Φ,Ψ difference was used. The exception, for which the hinge loop does not meet these criteria, is the IL-10/IFN-β pair; only Cα atoms are available for IFN-β, so the hinge loops can only be tentatively assigned.

[e] Structural alignments and best-fit superpositions were carried out using the program ALIGN (Satow et al., 1986) for the RNases and crystallins, and the program O (Jones et al., 1991). Pseudo-closed monomers were generated from the open monomers in the dimeric structures by combining coordinates from two subunits: BS-RNase, residues 1–22 of molecule 1 and residues 23–124 of molecule 2; β B2-crystallin, residues −2–87 of molecule 1 and residues 88–175 of molecule 2; IL-5, residues 1–85 of molecule 1 and residues 86–112 of molecule 2; IL-10, residues 10–116 of molecule 1 and residues 117–160 of molecule 2. The Cα RMS difference between BS-RNase and RNase A is based on the superposition of residues 1–15, 23–37, 39–65, 71–87, 95–113, and 115–124 of both. The Cα RMS difference between βB2-crystallin and γB-crystallin is based upon the superposition of residues 1–28, 29–68, 69, 72–73, 74–86, 88–106, 107–116, 118, 118A, and 120–172 of βB2-crystallin and residues 1–28, 29–68, 70, 71–72, 74–86, 88–106, 107–115, 116, 118, and 120–172 of γB-crystallin. The Cα RMS difference between IL-5 and GM-CSF is based upon the superposition of residues 6–21, 29–34, 42–55, 68–83, and 87–107 of IL-5 and residues 14–29, 37–42, 52–65, 73–88, and 97–117 of GM-CSF. The Cα RMS difference between IL-10 and IFN-β is based upon the superposition of residues 20–33, 44–47, 61–74, 76–80, 87–107, 118–128, 130–133, and 138–158 of IL-10 and residues 11–24, 33–36, 51–64, 66–70, 76–96, 114–124, 126–129, and 140–160 of IFN-β. Numbers in parentheses indicate the number of matched Cα atoms used the calculation of RMS differences.

[f] Coordinates supplied by original authors prior to PDB deposition.

[g] GM-CSF and IL-5 are both members of the short-chain subfamily of helical cytokines. Although the sequence identity is low, similar positions of cysteines in the 5′ ends of exons 3 and 4 and/or intron-exon organization and/or homology upstream of the TATA box suggests the IL-5 gene is related to the genes for the short-chain cytokines GM-CSF, IL-4, IL-2, and IFN-γ, and less closely to the gene for IL-3 (Tanabe et al., 1987). IL-2 and IFN-γ are not clustered on the same chromosome as the GM-CSF, IL-5, IL-4, and IL-3 genes.

[h] IL-10 and IFN-β are members of the long-chain subfamily of helical cytokines. Although the sequence identity of the proteins is low, the structural, functional, and genetic similarities of the long- and short-chain helical cytokines suggests that they are related by divergent evolution (Sprang & Bazan, 1993).

could form in this way because high intracellular protein concentrations can decrease the free energy of oligomerization (Berg, 1990; Zimmerman, 1993). An example of this might be the polymerization of sickle cell hemoglobin at high concentrations in the erythrocyte.

3D domain swapping (Bennett et al., 1994b) provides an alternative pathway from protein monomers to dimers, in which dimers are formed by exchanging domains. A central feature of the 3D domain swapping model is that it exploits a pre-evolved, tightly binding interface at every stage of oligomer evolution, rather than requiring the gradual accumulation of several random mutations or yielding dimers that are only stable at high protein concentration. Several of the domain-swapped dimers in Table 1 are examples of primitive stable or metastable dimers, which interact through extensive C-interfaces of 850–3,800 Å^2. These dimers could be formed by changes in the solution envi-

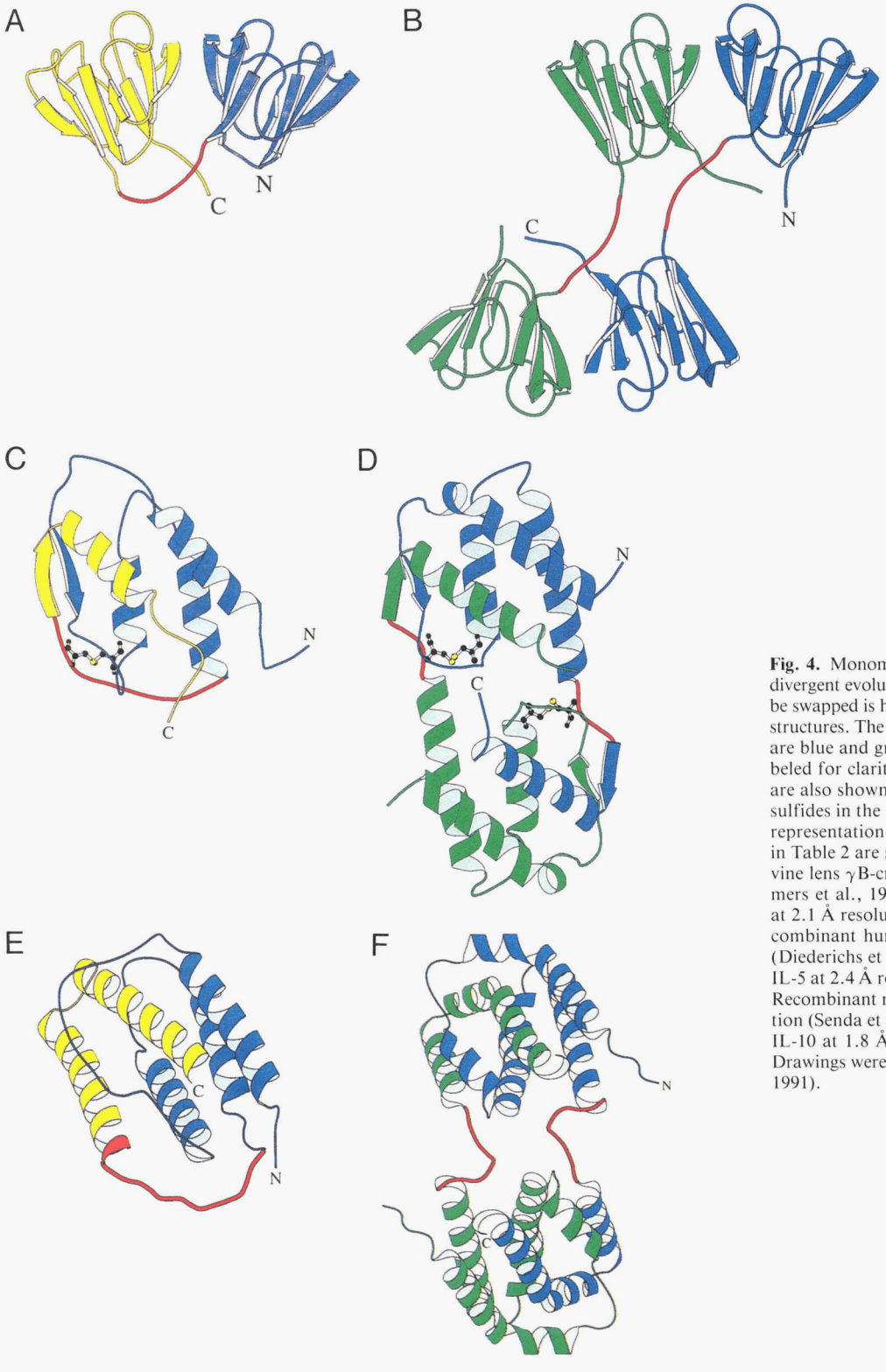

Fig. 4. Monomeric and dimeric proteins related by divergent evolution (Table 2). The domain that can be swapped is highlighted (yellow) in the monomer structures. The two subunits in the dimer structures are blue and green. The N- and C-termini are labeled for clarity. The hinge loops (red) (Table 2) are also shown. Where they occur, interchain disulfides in the dimers are shown in ball-and-stick representation. The RNase structures referred to in Table 2 are shown in Figure 2C and D. **A:** Bovine lens γB-crystallin at 1.5 Å resolution (Summers et al., 1984). **B:** Bovine lens βB2-crystallin at 2.1 Å resolution (Lapatto et al., 1991). **C:** Recombinant human GM-CSF at 2.4 Å resolution (Diederichs et al., 1991) **D:** Recombinant human IL-5 at 2.4 Å resolution (Milburn et al., 1993). **E:** Recombinant murine interferon β at 2.2 Å resolution (Senda et al., 1992). **F:** Recombinant human IL-10 at 1.8 Å resolution (Zdanov et al., 1995). Drawings were made with MOLSCRIPT (Kraulis, 1991).

ronments of the monomers (described above for DT), or by genetic changes that destabilized the monomers (discussed below). If these primitive dimers possessed new activities that conferred an advantage to the organism, stabilizing mutations in the O interface would be favored in natural selection (Fig. 1B).

3D domain swapping in oligomer evolution

Table 2 lists several dimer/monomer pairs in which the dimer is intertwined and resembles the structure of a different monomeric protein (Fig. 4). RNases and crystallins have highly iden-

tical sequences, suggesting divergent evolution from a common ancestor. Helical cytokines lack significant sequence identity, but are genetically, functionally and structurally similar. In the following sections, we describe two of the proteins in Table 2 in more detail.

GM-CSF versus IL-5

Structural, functional, and genetic similarities between the helical hematopoietic cytokines strongly suggest that they are related by divergent evolution (Sprang & Bazan, 1993). GM-CSF belongs to the short-chain helical cytokine subfamily and, like the other members of this family, the protein folds into a bundle of four helices (αA, B, C, and D) with two antiparallel β-strands (β 1 and 2) (Fig. 4C). IL-5 is also a member of the short-chain subfamily. Each half of the IL-5 dimer has the same fold as GM-CSF; however, it is composed of helices A, B, and C and β1 of one subunit, and helix D and β 2 of the other subunit (Fig. 4D). Thus, β2 and αD comprise a domain that was swapped in evolution from a primitive monomer that resembled the structure of GM-CSF and other monomeric helical cytokines.

There is ample genetic evidence that IL-5 and other hematopoietic cytokines are evolutionarily related. The human and murine genes for IL-5, IL-4, IL-3, and GM-CSF are clustered on the same chromosome (Takahashi et al., 1989). The IL-5 and GM-CSF genes also have significant sequence homology in the 5′-flanking regions immediately upstream of the TATA box. In addition, IL-5, IL-4, and GM-CSF have similar exon–intron organization and location of cysteine residues within exons (Tanabe et al., 1987). The similar structures and clustering of the genes on the same chromosome are suggestive of evolution by duplication of a common ancestral gene.

Similarities also exist between the amino acid sequences of IL-5 and GM-CSF. The position of the intrachain disulfide (Cys 54–Cys 88) in GM-CSF is similar to that of the interchain disulfide (Cys 44–Cys 86) in IL-5 (Fig. 4C,D). IL-5 and GM-CSF also share 28% amino acid sequence identity in αD (Rozwarski et al., 1994). The conserved residues in αD are apolar and buried in the cores of the GM-CSF and IL-5 structures. Thus, the tightly binding C-interface between αD and the other helices in GM-CSF is conserved in the dimer interface in domain-swapped dimeric IL-5.

A structure-based superposition of IL-5 and GM-CSF yields an overall RMS difference of 1.9 Å based upon 73 paired Cα atoms (Table 2). The superposition of GM-CSF and IL-5 is shown in Figure 5, highlighting their striking structural similarity. The aligned structures of GM-CSF and IL-5 offer insight into how the IL-5 dimer might have been formed by a loop deletion that forced it to adopt an open monomer conformation: GM-CSF has a 13-residue hinge loop (residues 87–99), whereas IL-5 has an 8-residue loop (residues 82–89) that would prevent formation of a closed monomer unless the secondary structure was disrupted (Fig. 5).

γB-crystallin versus βB2-crystallin

Crystallins are responsible for the high refractive index of the eye lens. β-Crystallins form dimers and higher oligomers, whereas γ-crystallins are always monomeric. The structures of γ B-crystallin (Wistow et al., 1983; Summers et al., 1984) and βB2-crystallin (Lapatto et al., 1991) are shown in Figure 4A and B, respectively and in Kinemage 4. γB-crystallin has two domains (N- and C-terminal), each of which is composed of two Greek key motifs. βB2-crystallin has two similar domains; how-

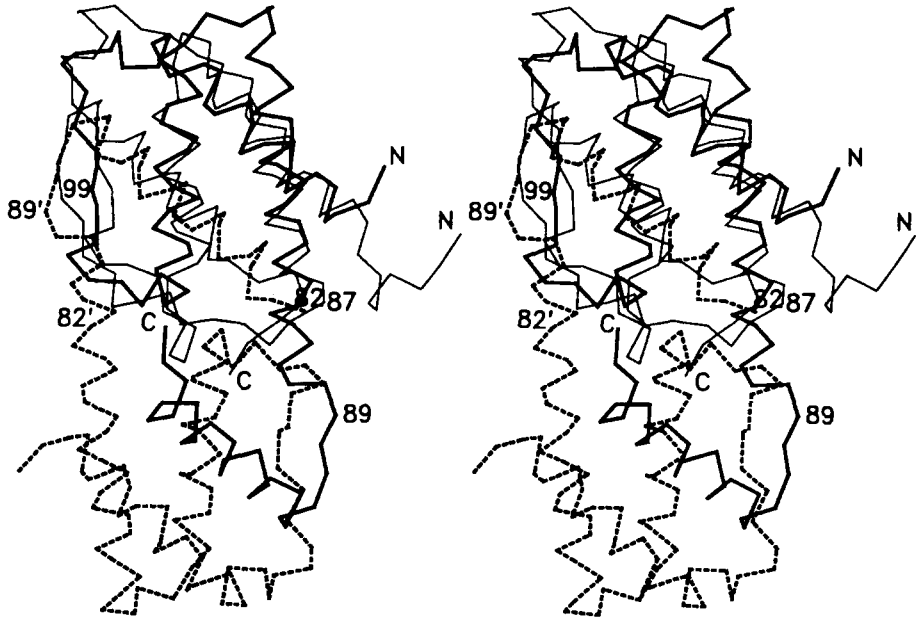

Fig. 5. Similarity of three-dimensional structures of IL-5 and GM-CSF. Stereo figure of the Cα backbones of IL-5 and GM-CSF after superposition of Cα atoms using the program O (Jones et al., 1991). Domain-swapped dimeric IL-5 is shown in bold solid and broken lines and monomeric GM-CSF is shown in thin lines. The N- and C-termini of GM-CSF and one subunit of the IL-5 dimer are labeled, as are the hinge loop boundaries (residues 87–99 in GM-CSF and residues 82–89 in IL-5). Residues in the second subunit of IL-5 are labeled with primed numbers.

ever, approximately half of the βB2-crystallin structure is swapped. Superposition of the entire γB-crystallin molecule with residues −2–87 (N-terminal domain) of one subunit and residues 88–172 (C-terminal domain) of the other subunit in the βB2-crystallin dimer yields an RMS difference of 1.2 Å, based on 167 pairs of Cα atoms. Thus, β B2-crystallin can be regarded as a domain-swapped dimer of a monomer resembling γB-crystallin.

Sequence similarities between β- and γ-crystallins suggest they are evolutionarily related. Alignment of the sequences of βB2- and γB-crystallins reveals their sequence identity is 36%, partly due to conservation of residues necessary for adoption of the Greek key fold (Summers et al., 1984). The gene organization of β- and γ-crystallins suggests they originated from a common ancestor and separated at the stage of a single domain protein, evolving the second domain independently by gene duplication and fusion (Lubsen et al., 1988). Therefore, the separation of the crystallins preceded the evolution of the C-interface between N- and C-terminal domains. The γ- and β-crystallins may have independently evolved similar monomeric forms, with later dimerization of β-crystallins by domain swapping. The idea of 3D domain swapping in the crystallins was essentially suggested previously by Bax et al. (1990): "the β-, γ-crystallin superfamily demonstrates how modification of an existing interface rather than evolution of a new one can give rise economically to novel assemblies during evolution."

Hinge loop mutations can lead to 3D domain swapping

An important distinction between the sticky billiard ball and 3D domain swapping models for oligomer evolution is that stable dimers may be formed by 3D domain swapping simply by destabilizing the monomer, rather than requiring the accumulation of mutations in a dimer interface. Because the C-interfaces are identical in monomers and domain-swapped dimers, destabilizing the C-interface in the monomer would also destabilize dimers. Therefore, the C-interface is an unlikely target for mutations that lead to domain swapping. However, the hinge loop between domains must play an important role in determining whether domain swapping occurs. In the following, we describe two possible ways that changes in hinge loops can promote domain swapping.

First, if a hinge loop connecting two domains is shortened by a deletion, then the closed monomer structure may no longer be sterically possible and the open monomer may be unstable because of the exposure of residues normally buried in the C-interface (Fig. 6A). Domain-swapped dimers would then be favored. An example of a loop deletion mutation that causes domain swapping is found in staphylococcal nuclease. A deletion of six residues in the loop preceding the C-terminal helix prevents the formation of a closed monomer, leading to the formation of a stable dimer by domain swapping (Green et al., 1995) (Fig. 2J). The same phenomenon is observed in the single-chain Fv molecules (composed of variable light (V_L) and heavy (V_H) domains), which form increasing amounts of dimer when the hinge loop between the V_L and V_H domains is shortened: the proportion of dimer is 53%, 34%, and 10% with 12-, 14-, and 18-residue hinge loops, respectively (Raag & Whitlow, 1995).

A second way that changes in a hinge loop can promote domain swapping is the following. The monomer might be destabilized by random mutations, such as substitution of residues that favor open hinge loop conformations because of having

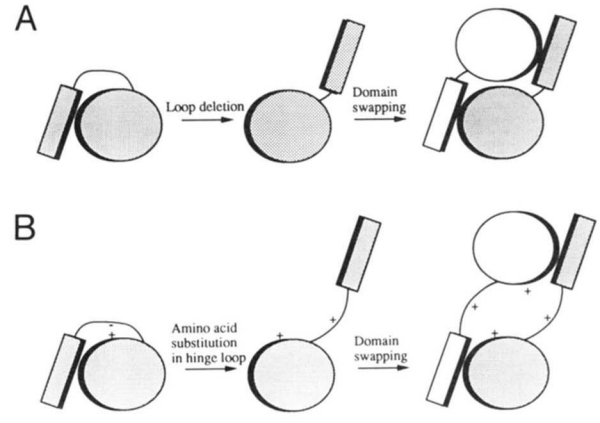

Fig. 6. Hinge loop mutations that lead to 3D domain swapping. The protein is comprised of two domains (circle and rectangle), which could be one or more α-helices and/or β-strands (as in the RNases and CD2) or full globular domains (as in DT and β B2-crystallin). The C-interface in the closed monomer (left structure, each panel) is indicated by black shading. In the open monomer (middle structure, each panel) this interface is exposed to solvent, which is unfavorable. In the domain-swapped dimer (right structure, each panel), the C-interface is buried by interacting with another subunit. **A:** Hinge loop deletion. **B:** Amino acid substitution.

fewer or greater conformational options (e.g., proline or glycine) or that have steric or electrostatic clashes in the closed monomer. As an example of this, βB2-crystallin differs from γ-crystallin in having an inserted aspartic acid (Asp 106A) and an acidic residue at position 87 in the hinge loop between N- and C-terminal domains. It has been suggested that, if the hinge loop in βB2-crystallin formed a closed monomer of the same structure as γ B-crystallin, these two acidic residues would electrostatically repel one another (Lapatto et al., 1991). Therefore, one or both of these amino acid replacements may have been genetic changes that destabilized the closed monomer and led to domain swapping. Replacement of the hinge loop in dimeric βB2-crystallin with the loop from γ-crystallin converts βB2-crystallin to monomers, underscoring the importance of the amino acid sequence in the hinge region in determining oligomerization (Trinkl et al., 1994)

Summary of concepts

The defining characteristic of a 3D domain-swapped oligomer is the presence of an intersubunit interface that is identical to an intrasubunit interface between domains in the monomer. Several protein structures are now known in both monomeric and oligomeric states that display this 3D domain swapping. In some of these structures (for example, DT and single-chain Fv antibodies), the swapped domain is a full globular domain; in other cases (for example RNase, CksHs2, and CD2), the swapped domain is only an α-helix, or a strand of β-sheet, or a few strands of β-sheet. In all cases, the oligomer is formed by the swapping of a domain that occupies the same environment in the oligomer as it does in the monomer. Thus, the same forces that stabilize the association of domains in the monomer must stabilize the oligomer.

The energetics of oligomer formation from monomers by 3D domain swapping are not yet well understood. But it is clear that the monomer and the oligomer generally have comparable en-

ergies. The reason is that the same interfaces are formed by the swapped domain in its monomeric and oligomeric molecules. The relatively small differences in energies of the monomer and oligomer have three origins: (1) differences in the conformations and environments of the hinge loop that leads to the swapped domain; (2) formation of an additional interface in the oligomer (the O interface) that is not present in the monomer; and (3) the loss of translational and rotational entropy in forming the oligomer from monomers. Terms 1 and 2 can favor either monomer or oligomer; term 3 must always favor monomer. Thus, when domain swapping occurs, either the monomer or the oligomer can be the more stable molecular form. In the case of DT, the monomer is more stable (Carroll et al., 1986); in the case of BS-RNase, the dimer is more stable (Piccoli et al., 1992).

There is likely to be a high energy barrier between monomer and oligomer, regardless of whether the monomer or the dimer is the more stable form. The reason is that the domain that swaps is held by noncovalent bonds in both the monomer and the oligomer. These bonds must be broken during the transition between monomer and oligomer; this breakage accounts for the energy barrier. Because of this barrier, the less stable molecular form is likely to be metastable; that is, it is likely to have a prolonged existence in its metastable form. In other words, two (or more) states of the same protein chain can be effectively stable over significant periods of time.

The discovery of domain-swapped proteins forces a minor modification of Anfinsen's "thermodynamic hypothesis" (Anfinsen, 1973). Anfinsen's hypothesis is that a protein folds to its thermodynamically stable structure. With domain swapping, there are two (or more) structures that can be effectively stable over significant periods of time. Of course, the domain-swapped structure resembles the monomeric structure: most of the interfaces in the oligomer are the same as those in the monomer. But there is the important difference that the oligomer is a multiple subunit molecule, capable of cooperativity and other properties that differ from the monomer. The point is that one sequence can form more than one structure, with all structures being close in energy.

Implications of domain-swapped proteins for function

Domain swapping offers the opportunity to cells of regulation of assembly and metabolism by switching proteins between oligomeric structures. For proteins that exist in various oligomeric states within a cell, such as CksHs2, the mechanisms for switching between monomer and dimer are not yet known, but the well-studied examples of RNase A and DT demonstrate that switching is possible by chemical mechanisms (e.g., changes in pH and concentration) similar to those that could operate within organisms. In the case of DT, a drop in pH and the transient high concentration that accompany freezing and thawing cause domain swapping. Changes in pH or ionic concentration may be common events in vivo that destabilize monomers, leading to higher oligomers (see Fig. 3).

The switching between oligomeric states can be transitory, as for the cell cycle protein CksHs2, or can be permanent, as in the evolution of oligomers from monomers by domain swapping. As argued above, oligomer evolution by domain swapping (Table 2, Fig. 4) rather than by the "sticky billiard ball model" avoids the necessity of evolving a set of stabilizing mutations at the intersubunit interface. Instead, a mutation that destabilizes the monomer can lead to oligomers that interact through a C-interface that has "pre-evolved" as an extensive, tight interface in the monomer.

Domain swapping also provides a model for the evolution of oligomeric enzymes that have their active sites at subunit junctions. These include RuBisCO (Schreuder et al., 1993), glutamine synthetase (Almassy et al., 1986), catalase (Fita & Rossmann, 1985), and others. A possible pathway for the evolution of such enzymes is shown in Figure 7. As suggested by Gilbert (1978), the fusion of exons encoding single-domain proteins (perhaps each binding a different substrate) could generate primitive closed monomers with new functions (Fig. 7, I and II). Figure 7, III represents a primitive closed monomer that has an active site in the C-interface, as is commonly found in multidomain proteins. 3D domain swapping could then yield a dimeric form of the protein, which has active sites between subunits (Fig. 7, IV), and also has the possibility of forming an O interface. Figure 7, V represents a present-day dimer, in which amino acid replacements in the O interface further stabilize the dimer. Thus, Figure 7 represents a plausible evolutionary pathway to an enzyme with its active site formed by two protein chains.

Domain swapping also suggests a possible mechanism for protein aggregation and protein fiber formation, as in amyloidosis. Amyloid deposits are insoluble fibers formed from aggregated protein that disrupt cells. The oligomers discussed above are of

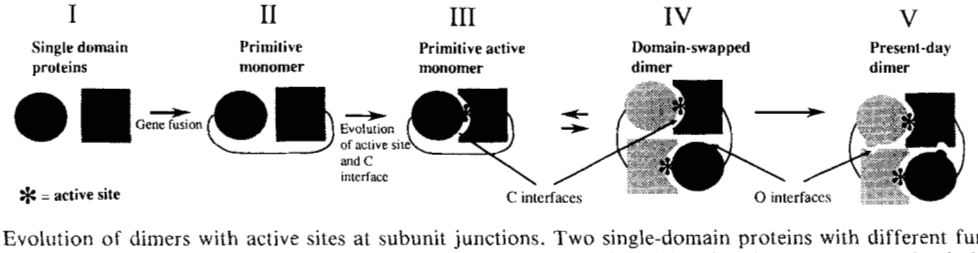

Fig. 7. Evolution of dimers with active sites at subunit junctions. Two single-domain proteins with different functions are represented by a circle and square (I). Fusion of genes for the two proteins yields closed monomers covalently bound by a polypeptide linker (II). Evolution of a C-interface and active site may be favored by covalent linkage of the domains, which have less rotational and translational freedom than the single-domain protein precursors. Monomers having active sites (indicated by an asterisk) at the C-interface are shown in III. If 3D domain swapping occurs, a dimer will be formed that has two C-interfaces with two active sites formed between two subunits (black and gray) (IV). This domain-swapped dimer promotes communication between active sites and hence might favor the acquisition of allosteric regulation, as is found in RNase (Piccoli & D'Alessio, 1984). If this domain-swapped dimer confers some advantage to the organism, then amino acid replacements that stabilize the O interface will be selected in natural selection, leading to a more stable dimer with active sites at subunit interfaces (V).

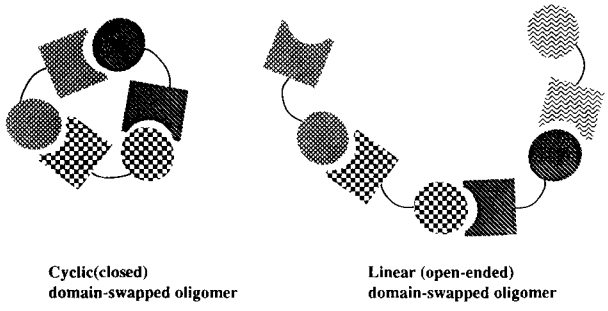

Fig. 8. Formation of open-ended oligomers by 3D domain swapping. Left: Closed domain-swapped oligomer (in this case, a trimer) as in Tables 1 and 2. Right: High molecular weight, open-ended oligomers may also be formed by association of molecules through a C-interface. As described in the text, some protein "precipitates" may be of this form.

the closed sort, such as oligomeric enzymes. But open-ended oligomers can also be formed by 3D domain swapping, as shown in Figure 8, in which each swapped domain leads to a new molecule. Then a large aggregate is formed, which would be expected to precipitate in the test tube or cell. DT (Bennett et al., 1994b), single-chain Fv molecules (Raag & Whitlow, 1995), RNase A (Crestfield et al., 1962), tryptophan synthetase α subunit (Jackson & Yanofsky, 1969), and several other proteins form oligomers higher than dimers, and there appears to be no reason that very large open-ended oligomers of proteins could not be formed by domain swapping. Even though such large domain-swapped aggregates would be difficult to characterize in terms of molecular structure, it may be useful to keep in mind that the protein molecules of precipitates and amyloid deposits may be held together by domain swapping.

One other type of protein association can be explained by domain swapping. This is the phenomenon of protein complementation (Zabin & Villarejo, 1975), in which two different protein molecules, each lacking a different portion, can cooperate to form a functioning molecule. A simple example is from the work of Crestfield et al. (1963), in which two samples of RNase A were treated to inactivate His 12 in one sample, and His 119 in the other sample. Both histidines are required for catalysis and, upon mixing the samples, no activity was found. But after lyophilization from acetic acid and redissolving, partially active RNase, now a dimer, was recovered. Crestfield et al. reasoned that domain swapping must have occurred (see Fig. 2D), leading to some active sites containing unmodified His 12 and unmodified His 119. Other examples of protein complementation have emerged from work on tryptophan synthetase α subunit (Jackson & Yanofsky, 1969), β-galactosidase (Zabin & Villarejo, 1975), and lactose permease (Bibi & Kaback, 1992).

In conclusion, 3D domain swapping is a newly recognized, but not uncommon, mode of protein association that can account simply for a variety of observations on the function and evolution of oligomeric proteins that are otherwise difficult to explain. In particular, domain swapping permits switching among oligomeric forms of a protein having comparable energies.

Acknowledgments

We thank the NIH for grant GM 31299; the Swiss National Science Foundation, Ciba-Geigy Jubilaeums-Stiftung (Switzerland), and EMBO for a fellowship to M.P.S.; Drs. T.O. Yeates, M.S. Weiss, L.J. Beamer, J.U. Bowie, and R. Jaenicke for discussions; and Drs. S.R. Jordan, R.L. Brady, P.M. Colman, E.E. Lattman, and J.A. Tainer for providing protein coordinates prior to deposition.

References

Almassy RJ, Janson CA, Hamlin R, Xuong NH, Eisenberg D. 1986. Novel subunit-subunit interactions in the structure of glutamine synthetase. *Nature 323*:304–309.

Anfinsen CB. 1973. Principles that govern the folding of protein chains. *Science 181*:223–230.

Arvai AS, Bourne Y, Hickey MJ, Tainer JA. 1995. Crystal structure of the human cell cycle protein CksHs1: Single domain fold with similarity to kinase N-lobe domain. *J Mol Biol 249*:835–842.

Bax B, Lapatto R, Nalini V, Driessen H, Lindley PF, Mahadevan D, Blundell TL, Slingsby C. 1990. X-ray analysis of βB2-crystallin and evolution of oligomeric lens proteins. *Nature 347*:776–780.

Bennett MJ, Choe S, Eisenberg D. 1994a. The refined structure of dimeric diphtheria toxin at 2.0 Å resolution. *Protein Sci 3*:1444–1463.

Bennett MJ, Choe S, Eisenberg D. 1994b. Domain swapping: Entangling alliances between proteins. *Proc Natl Acad Sci USA 91*:3127–3131.

Bennett MJ, Eisenberg D. 1994. The refined structure of monomeric diphtheria toxin at 2.3 Å resolution. *Protein Sci 3*:1464–1475.

Berg OG. 1990. The influence of macromolecular crowding on the thermodynamic activity: Solubility and dimerization constants for spherical and dumbbell-shaped molecules in a hard-sphere mixture. *Biopolymers 30*: 1027–1037.

Bibi E, Kaback HR. 1992. Functional complementation of internal deletion mutants in the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci USA 89*:1524–1528.

Blackburn P, Moore S. 1982. Pancreatic ribonucleases. In: Boyer PD, ed. *The enzymes*. New York: Academic Press, Inc. pp 317–433.

Cafaro V, De Lorenzo C, Piccoli R, Bracale A, Mastronicola MR, Di Donato A, D'Alessio G. 1995. The antitumor action of seminal ribonuclease and its quaternary conformations. *FEBS Lett 359*:31–34.

Carroll SF, Barbieri JT, Collier RJ. 1986. Dimeric form of diphtheria toxin: Purification and characterization. *Biochemistry 25*:2425–2430.

Caspar DLD. 1980. Movement and self-control in protein assemblies. *Biophys J 32*:103–138.

Collaborative Computing Project Number 4. 1994. *The CCP4 suite: Programs for protein crystallography*. SERC Daresbury Laboratory, Warrington WA4 4AD, England. *Acta Crystallogr D 50*:760–763.

Collier RJ. 1975. Diphtheria toxin: Mode of action and structure. *Bacteriol Rev 39*:54–85.

Colon W, Kelly JW. 1992. Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry 31*:8654–8660.

Crestfield AM, Stein WH, Moore S. 1962. On the aggregation of bovine pancreatic ribonuclease. *Arch Biochem Biophys Suppl 1*:217–222.

Crestfield AM, Stein WH, Moore S. 1963. Properties and conformation of the histidine residues at the active site of ribonuclease. *J Biol Chem 238*:2421–2428.

Diederichs K, Boone T, Karplus PA. 1991. Novel fold and putative receptor binding site of granulocyte-macrophage colony-stimulating factor. *Science 254*:1779–1782.

Eisenberg D, McLachlan AD. 1986. Solvation energy in protein folding and binding. *Nature 319*:199–203.

Eisenberg D, Wesson M, Yamashita M. 1989. Interpretation of protein folding and binding with atomic solvation parameters. *Chemica Scripta 29A*:217–221.

Erickson HP. 1989. Co-operativity in protein–protein association: The structure and stability of the actin filament. *J Mol Biol 206*:465–474.

Finkelstein AV, Janin J. 1989. The price of lost freedom: Entropy of bimolecular complex formation. *Protein Eng 3*:1–4.

Fita I, Rossmann MG. 1985. The active center of catalase. *J Mol Biol 185*: 21–37.

Gilbert W. 1978. Why genes in pieces. *Nature 271*:501.

Green SM, Gittis AG, Meeker AK, Lattman EE. 1995. One step evolution of a dimer from a monomeric protein. *Nature Struct Biol 2*:746–751.

Jackson DA, Yanofsky C. 1969. The formation and properties of dimers of the tryptophan synthetase α subunit of *Escherichia coli*. *J Biol Chem 244*:4526–4538.

Jaenicke R. 1995. Folding and association versus misfolding and aggregation of proteins. *Phil Trans R Soc Lond B 348*:97–105.

Janin J, Miller S, Chothia C. 1988. Surface, subunit interfaces and interior of oligomeric proteins. *J Mol Biol 204*:155–164.

Jones EY, Davis SJ, Williams AF, Harlos K, Stuart DI. 1992. Crystal struc-

ture at 2.8 Å resolution of a soluble form of the cell adhesion molecule CD-2. *Nature 360*:232–239.
Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A 47*:110–119.
Kikuchi Y, King J. 1975. Genetic control of bacteriophage T4 baseplate morphogenesis. *J Mol Biol 99*:645–647.
Klafki HW, Pick AI, Pardowitz I, Cole T, Awni LA, Barnikol HU, Mayer F, Kratzin HD, Hilschmann N. 1993. Reduction of disulfide bonds in an amyloidogenic Bence Jones protein leads to formation of "amyloid-like" fibrils in vitro. *Biol Chem Hoppe-Seyler 374*:1117–1122.
Kortt AA, Malby RL, Caldwell JB, Gruen LC, Ivancic N, Lawrence MC, Howlett GJ, Webster RG, Hudson PJ, Colman PM. 1994. Recombinant anti-sialidase single-chain variable fragment antibody. *Eur J Biochem 221*:151–157.
Koshland DE Jr, Nemethy G, Filmer D. 1966. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry 5*:365–385.
Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.
Lapatto R, Nalini V, Bax B, Driessen H, Lindley PF, Blundell TL, Slingsby C. 1991. High resolution structure of an oligomeric eye lens β-crystallin. *J Mol Biol 222*:1067–1083.
Loll PJ, Lattman EE. 1989. The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+} and the inhibitor pdTp refined at 1.65 Å. *Proteins Struct Funct Genet 5*:183–201.
London J, Skrzynia C, Goldberg ME. 1974. Renaturation of *Escherichia coli* tryptophanase after exposure to 8 M urea. *Eur J Biochem 47*:409–415.
Lubsen NH, Aarts HJM, Schoenmakers JGG. 1988. The evolution of lenticular proteins: The β- and γ-crystallin super gene family. *Prog Biophys Mol Biol 51*:47–76.
Mazzarella L, Capasso S, Demasi D, Di Lorenzo G, Mattia CA, Zagari A. 1993. Bovine seminal ribonuclease: Structure at 1.9 Å resolution. *Acta Crystallogr D 49*:389–402.
Milburn MV, Hassell AM, Lambert MH, Jordan SR, Proudfoot AEI, Graber P, Wells TNC. 1993. A novel dimer configuration revealed by the crystal structure at 2.4 Å resolution of human interleukin-5. *Nature 363*: 172–176.
Monod J, Wyman J, Changeux JP. 1965. On the nature of allosteric transitions: A plausible model. *J Mol Biol 12*:88–118.
Murray AJ, Lewis SJ, Barclay AN, Brady RL. 1995. One sequence, two folds: A metastable structure of CD-2. *Proc Natl Acad Sci USA 92*:7337–7341.
Parge HE, Arvai AS, Murtari DJ, Reed SI, Tainer JA. 1993. Human CksHs2 atomic structure: A role for its hexameric assembly in cell cycle control. *Science 262*:387–395.
Piccoli R, D'Alessio G. 1984. Relationships between nonhyperbolic kinetics and dimeric structure in ribonucleases. *J Biol Chem 259*:693–695.
Piccoli R, Di Donato A, D'Alessio G. 1988. Co-operativity in seminal ribonuclease function. *Biochem J 253*:329–336.
Piccoli R, Tamburrini M, Piccialli G, Di Donato A, Parente A, D'Alessio G. 1992. The dual-mode quaternary structure of seminal RNase. *Proc Natl Acad Sci USA 89*:1870–1874.
Raag R, Whitlow M. 1995. Single chain Fvs. *FASEB J 9*:73–80.
Richards FM, Wyckoff HW. 1973. *Ribonuclease S*. Oxford: Clarendon Press.
Rozwarski DA, Gronenborn AM, Clore GM, Bazan JF, Bohm A, Wlodawer A, Hatada M, Karplus PA. 1994. Structural comparisons among the short-chain helical cytokines. *Structure 2*:159–173.
Satow Y, Cohen GH, Padlan EA, Davies DR. 1986. Phosphocholine binding immunoglobulin Fab McPC603. *J Mol Biol 190*:593–604.
Schreuder HA, Knight S, Curmi PMG, Andersson I, Cascio D, Sweet RM, Branden CI, Eisenberg D. 1993. Crystal structure of activated tobacco rubisco complexed with the reaction-intermediate analogue 2-carboxyarabinitol 1,5-bisphosphate. *Protein Sci 2*:1136–1146.
Senda T, Shimazu T, Matsuda S, Kawano G, Shimizu H, Nakamura KT, Mitsui Y. 1992. Three-dimensional crystal structure of recombinant murine interferon-β. *EMBO J 11*:3193–3201.
Sprang SR, Bazan JF. 1993. Cytokine structural taxonomy and mechanisms of receptor engagement. *Curr Opin Struct Biol 3*:815–826.
Summers L, Wistow G, Narebor M, Moss D, Lindley P, Slingsby C, Blundell T, Bartunik H, Bartels K. 1984. X-ray studies of the lens specific proteins: The crystallins. In: Hearin MTW, ed. *Peptide and protein reviews, Vol 3*. New York: Marcel Dekker. pp 147–168.
Suzuki H, Parente A, Farina B, Greco L, La Montagna R, Leone E. 1987. Complete amino-acid sequence of bovine seminal ribonuclease from seminal plasma. *Biol Chem Hoppe-Seyler 368*:1305–1312.
Takahashi M, Yoshida MC, Satoh H, Hilgers J, Yaoita Y, Honjo T. 1989. Chromosomal mapping of the mouse IL-4 and human IL-5 genes. *Genomics 4*:47–52.
Tanabe T, Konishi M, Mizuta T, Noma T, Honjo T. 1987. Molecular cloning and structure of the human interleukin-5 gene. *J Biol Chem 262*: 16580–16584.
Trinkl S, Glockshuber R, Jaenicke R. 1994. Dimerization of βB2-crystallin: The role of the linker peptide and the N- and C-terminal extensions. *Protein Sci 3*:1392–1400.
van den Berg L, Rose D. 1959. Effect of freezing on the pH and composition of sodium and potassium phosphate solutions: The reciprocal system KH_2PO_4-Na_2HPO_4-H_2O. *Arch Biochem Biophys 81*:319–329.
Wistow G, Turnell B, Summers L, Slingsby C, Moss D, Miller L, Lindley P, Blundell T. 1983. X-ray analysis of the eye lens protein γ-II crystallin at 1.9 Å resolution. *J Mol Biol 170*:175–202.
Wlodawer A, Bott R, Sjolin L. 1982. The refined crystal structure of ribonuclease A at 2.0 Å resolution. *J Biol Chem 257*:1325–1332.
Yan Y, Winograd E, Viel A, Cronin T, Harrison SC, Branton D. 1993. Crystal structure of the repetitive segments of spectrin. *Science 262*:2027–2030.
Zabin I, Villarejo MR. 1975. Protein complementation. *Annu Rev Biochem 44*:295–313.
Zdanov A, Schalk-Hihi C, Gustchina A, Tsang M, Weatherbee J, Wlodawer A. 1995. Crystal structure of interleukin-10 reveals the functional dimer with an unexpected topological similarity to interferon γ. *Structure 3*:591–601.
Zimmerman SB. 1993. Macromolecular crowding effects on macromolecular interactions: Some implications for genome structure and function. *Biochim Biophys Acta 1216*:175–185.