æ

# 13. MULTIMERIZATION

21 July 2022

Although the preceding chapter focused on proteins in a generic way, a substantial fraction of proteins in virtually all organisms operate by binding to other proteins, either transiently or in semi-permanent structures. Here, we consider multimeric assemblages of polypeptide subunits held together at binding interfaces in a noncovalent manner. All cells contain large numbers of symmetric complexes of proteins, called dimers when there are two subunits, trimers when there are three, tetramers when there are four, etc. (Figure 13.1). Indeed, virtually all biological pathways and cell structural features involve one or more multimers. These include a wide array of enzymes, signaling molecules, etc. There are also numerous higher-order complexes built out of lower-order multimers, prime examples including flagella, nuclear pore complexes, and centrosomes, all of which are discussed in separate chapters.

Before exploring the evolution of the basal building blocks of protein complexes, a brief overview of nomenclature will be useful. When the subunits of a multimer are encoded by the same genetic locus, the complex is referred to as a homomer. Complexes consisting of subunits from two or more loci are heteromers, the extreme situation being the case in which each subunit is encoded by a different genetic locus. When considering homomers, it is useful to distinguish between isologous and heterologous forms of interfaces involved in multimeric constructs. In the former case, both participants deploy the same surface patch (although generally rotated) in binding, whereas in the latter case, each binding partner utilizes different interface residues. By definition, all heteromeric interfaces are heterologous.

Odd-mers (e.g., trimers, pentamers, etc.) rely almost exclusively on heterologous interfaces, as these are required to create closed structures (Figure 13.2). Closed structures with just one type of heterologous interface are referred to as cyclical multimers. A special case is the domain-swapping homodimer, which has a geometric configuration that allows two symmetric heterologous interfaces to form a closed loop. In contrast, dihedral multimers have multiple axes of symmetry and require more than one interface type, with each usually being isologous (Figure 13.2).

How and why multimeric structures originate constitute two of the more significant issues in evolutionary cell biology. It is often assumed that the kinds of organized diversity embodied in multimers could only be a product of natural selection, with higher-order multimers sometimes being referred to as the "end points" or "pinnacles" of stepwise evolution. Should this be true, then an obvious question is why only a subset of lineages have been able to achieve such lofty heights. In fact, as will be shown below, the distribution of multimeric states in prokaryotes

is not much different from that in eukaryotes, and there are numerous examples of evolutionary reversions to lower oligomeric states. Moreover, although eukaryotes may have a higher incidence of heteromeric than homomeric structures relative to orthologs in bacteria (Figure 13.3; Reid et al. 2010; Nishi et al. 2013; Marsh and Teichmann 2014), there is no compelling evidence that the former are functionally superior (e.g., Chapters 10, 20, and 24).

Sometimes a multimeric complex achieves a new function by virtue of special features associated with the interface itself. For example, there are numerous cases in which higher-order complexes assemble into containers (some chaperones, and cages for vesicles) or fibrils (actins and tubulins). However, in the case of enzymes, more often than not, all subunits retain their original monomeric functions, in which case for example, a dimeric enzyme would simply have two catalytic sites but no new function.

Because the extent to which multimeric structures are promoted or maintained by selection remains unclear, it is useful to think of them as "biology's snowflakes," to remind us that beauty and diversity can often arise for purely physical reasons. We first consider the incidence of various multimeric types of molecules across the Tree of Life, highlighting the apparent lack of association between molecular and organismal complexity. We then review some of the biophysical considerations relevant to transitions between alternative oligomeric forms, showing that the paths to multimerization can be remarkably simple. Following this entrée into the biophysics of aggregation and the general features of known interfaces, the basic theory for understanding how multimeric architectures evolve will be reviewed.

## The Incidence and Architectural Features of Multimers

There is astounding diversity in the higher-order structure of proteins. Surveys of known quaternary structures indicate that multimers comprise $\sim 60\%$ of all characterized proteins, with homomers being about twice as frequent as heteromers (Levy et al. 2006; Marianayagam et al. 2004; Lynch 2012). Because complex proteins are more difficult to characterize structurally, it is likely that the incidence of multimers is even greater than 60%. The vast majority of multimers are dimers, with a roughly exponential pattern of decline in frequencies with elevated numbers of subunits, and odd-mers tending to be under-represented relative to even-mers (Figure 13.3; Goodsell and Olson 2000; Mei et al. 2004). Strikingly, these distributions are essentially independent of phylogenetic context, with those for prokaryotes being quite similar to those for vertebrates and land plants (Figure 13.3).

Although these distributions are potentially biased by the nature of proteins that have been studied at the structural level, a comparison of orthologous metabolic proteins across taxa supports the conclusion that there is no gradient of molecular complexity (number of subunits) with organismal complexity. Moreover, within and among most major lineages, there is substantial variation in the multimeric states of the same proteins with the same functions (Reid et al. 2010; Lynch 2013). For example, it is not uncommon to see the same enzyme operating as a monomer in some bacteria, a dimer in others, and a tetramer in still others. Hemoglobins provide a textbook view of the extraordinary diversity of complexes that can exist

(from monomers to structures containing more than 100 subunits) (Shionyu et al. 2001). Levy et al. (2006) suggest that at the point of 30-40% amino-acid sequence divergence between orthologs, the probability of two species sharing quaternary structure drops to $\sim 70\%$.

Numerous studies also reveal that orthologous proteins with the same numbers of subunits in different phylogenetic lineages often utilize different binding interfaces. For example, for glycosyltransferases (enzymes involved in the transfer of sugars to acceptor molecules), Hashimoto et al. (2010) found a dramatic diversity of monomeric, homodimeric, and homotetrameric states across the Tree of Life, with many of the independently evolved dimers utilizing different binding interfaces, sometimes even on opposite sides of the monomeric subunit. Examples are also known of multimers in which some species use isologous and others use heterologous interfaces (Vassylyev et al. 2006). For proteins with known multimeric structures in more than one species, $\sim 36\%$ of cases exhibit variants with different interfaces, with $\sim 4\%$ of cases having more than five different binding modes across species (Dayhoff et al. 2010).

Ahnert et al. (2015) produced a "periodic-table" of multimer types based on the total number of repeating units, the number of subunit types (gene products) per repeat, and the number of ways in which these can assemble into multimers. Just a small set of the possible topologies for heteromers, with two to three repeats and two subunit types per repeat, are shown in Figure 13.4. When there are four or more repeats, the final complexes can be cyclical or dihedral in form. For example, a cyclic homomeric tetramer has just one type of face-to-back interface repeated four times in a closed loop, whereas a dihedral homomeric tetramer utilizes two different isologous interfaces twice (Figure 13.2). If there are two subunits per repeat and three repeats, the resultant structure is a hexamer, which can take on two configurations (Figure 13.4). A wide variety of higher-order topologies beyond those outlined in Figures 13.2 and 13.4 are known to exist in biology (Ahnert et al. 2015), but whenever there exists the possibility of either a dihedral or cyclical closed structure (as with tetramers; Figure 13.2), dihedral forms are $\sim 10\times$ more frequent than cyclical forms (Levy et al. 2008). This suggests a greater ease of evolution towards isologous binding interfaces, a point that will be further discussed below.

For heteromeric even-mers beyond dimers (e.g., tetramers and hexamers), there is the possibility of uneven stoichiometry from the contributing proteins (e.g., a tetrameric complex might be AAAB as opposed to AABB). Uneven stoichiometry is thought to be nearly twice as common in bacteria as in eukaryotes, and the incidence of unevenness increases with the number of contributing proteins, e.g., $\sim 10\%$ when just two or three proteins participate, but $\sim 30\%$ when there are five or more contributing proteins (Marsh et al. 2015).

## Propensity to Aggregate

The reliance on multimers imposes some costs and risks upon cells. To achieve multimeric architectures, cellular concentrations of monomeric subunits must be kept at sufficiently high levels to allow reasonable encounter and aggregation rates between partner molecules. Because proteins come at a biosynthetic cost, this may

then entail an excess investment that would otherwise not be required for monomers. If, on the other hand, multimers are less subject to degradation, the overall cost of monomer production could be mitigated. There is, however, an additional fine line between the assembly of functional complexes vs. collateral damage, as adhesive interfaces can promote promiscuous interactions with noncognate molecules and/or lead to the runaway production of harmful self-aggregates, including fibrils.

The proximity of proteins to the edge of misaggregation propensity is demonstrated by experimental work in *E. coli* showing that single amino-acid substitutions to hydrophobic surface residues often shift proteins into supramolecular states (Garcia-Seisdedos et al. 2017). In a well-known case in human biology, a single glutamate-to-valine substitution in the hemoglobin molecule induces the fibril formation that causes sickle-cell anemia. From experimental observations in yeast, Zhang et al. (2008) estimate that promiscuous binding leads to $\sim 25\%$ of proteins being at least transiently bound to inappropriate partners.

The clear implication here is that any involvement of natural selection in the evolution of higher-order protein structures must include not just the enforcement of productive self-binding but the elimination of surface features that promote harmful structures (Zabel et al. 2019). Mechanisms for mitigating the latter problems may include the spatial configuration of adhesive residues on polypeptide chains in ways that influence the relative rates of monomer folding and subunit aggregation. For example, locating interface residues on the C-terminal end of a protein (the last to emerge from the ribosome) reduces the likelihood of premature aggregation of partially folded proteins (Natan et al. 2018; Gartner et al. 2020). The spatially restricted nature of transcription also reduces the likelihood of off-target binding. For example, when mRNAs are cotranslated by groups of ribosomes (so-called polysomes jointly processing the same mRNA), this creates a situation in which monomers are initially colocalized at substantially higher concentrations than expected under a random intracellular distribution, increasing the association rate, and sometimes even leading to co-translational assembly (Wells et al. 2015). In the case of bacterial operons, spatial clustering likely even allows for enhanced assembly of heteromers (Wells et al. 2016).

**Theory of association.** Acknowledging that there are many underlying molecular factors, we now consider in the simplest and most general terms the principles underlying the rate and stability of molecular aggregation, building on the concept of molecular diffusion introduced in Chapter 7. After emerging from a ribosome, amino-acid chains will initially be monomeric in form, and will retain that state until encountering a binding partner. Even then, the maintenance of an appropriate multimeric form requires sufficient binding energy across the interface. In principle, given steady-state conditions (e.g., a constant concentration of protein per unit cell volume), because multimeric subunits are held together in a noncovalent fashion, an equilibrium will be reached within the cytoplasm reflecting equal rates of association into multimers and dissociation of the latter back to monomers. In this sense, few proteins can be viewed as having a single form, and when one is referred to, this should simply be viewed as the dominant phase.

The equilibrium distribution of multimeric forms within a cell depends on at least three factors: 1) the cellular concentration of proteins, which dictates the

encounter rate of monomeric subunits; 2) the rate of particle movement and binding, which dictates the assembly rate; and 3) the stability of the binding interface, which dictates the longevity of a newly formed multimer. The basic principles can be understood by focusing on a system allowing for just monomers and dimers.

Owing to the destabilizing forces of molecular motion, two subunits will have a thermodynamically determined probability of being in complex depending on the rates of association of monomers and dissociation of dimers. Letting [A] and [AA] denote the cellular concentrations of solo A molecules and AA complexes, the fraction of complexes vs. singletons is

$$p_{\mathrm{AA}} = \frac{[\mathrm{AA}]}{[\mathrm{A}] + [\mathrm{AA}]}. \tag{13.1a}$$

By multiplying the numerator and denominator by [A]/[AA], the expected fraction of dimers at equilibrium can then be expressed as

$$p_{\mathrm{AA}} = \frac{[\mathrm{A}]}{K_D + [\mathrm{A}]}, \tag{13.1b}$$

where

$$K_D = \frac{[\mathrm{A}][\mathrm{A}]}{[\mathrm{AA}]} \tag{13.2a}$$

is the equilibrium dissociation constant, which depends on the strength of binding (Foundations 13.1).

Written in this way, $K_D$ is seen to be equivalent to the cellular concentration of monomers at which aggregation leads to equivalent concentrations of dimeric and monomeric complexes, i.e., $p_{\mathrm{AA}} = 0.5$ when $K_D = [\mathrm{A}]$, which requires $[\mathrm{AA}] = [\mathrm{A}]$. Note that $[A]$, $[AA]$, and $K_D$ all have the same units of concentration, usually written below as $\mu$M (where 1 $\mu$M $= 10^{-6}$ M) because cytoplasmic protein concentrations are typically in the $\mu$M range (Chapter 7). In a more mechanistic sense, $K_D$ is equivalent to the ratio of dissociation and association rates, where the latter can be expressed in terms of the random encounter rate (Chapter 7) and the probability of proper binding (Foundations 13.1). Because the rate of encounter is a function of the molecular concentration, proteins with lower binding affinities are expected to require higher cellular concentrations to achieve an effective level of complexation.

To put Equation 13.2a into a more biophysical perspective, the dissociation constant can be written in energetic terms as

$$K_D = e^{-\Delta E/RT}. \tag{13.2b}$$

where $\Delta E$ denotes the excess energy required for dissociation (in kcal/mol), $R$ is the Boltzmann constant $K_B$ (encountered in previous chapters) scaled up to the molar equivalent, and $RT$ denotes the standard background energy associated with Brownian molecular motion (where for most biological temperatures, $RT \simeq 0.6$ kcal/mol) (Foundations 13.1). Thus, as the binding energy across the interface between two subunits increases, $K_D$ asymptotically approaches zero and the probability of complexation approaches 1.0.

It should be emphasized that the quantitative value of the dissociation constant is context dependent. As defined in Equation 13.2b, $K_D$ has a rather precise meaning

from a pure chemistry perspective, and is typically recorded in an aqueous solution containing only A molecules. Written in the form of Equation 13.1b, however, $K_D$ can be thought of in much more general terms, as a simple indicator of the degree of affinity of two A molecules for each other in a system of arbitrary complexity.

Within the cellular environment, which can contain thousands of proteins encoded by other loci, a monomeric subunit is confronted not just with the challenge of adhering to its binding partner, but with the additional problem of avoiding promiscuous engagement with noncognate molecules. Assuming that the molecule of interest is sufficiently adhesive that it is essentially always in complex with either a self or foreign molecule, the methods in Foundations 13.1 can be used to derive a modified expression for the fraction of molecules bound up in appropriate homodimeric complexes. Maintaining the structure of the preceding formula,

$$K_D' = K_D \cdot \phi e^{\Delta E'/(RT)}. \tag{13.3}$$

Here, $\phi$ is the effective concentration of foreign proteins with a capacity for promiscuous binding, and $\Delta E'$ is the strength of binding associated with promiscuous liaisons. In effect, the terms to the right of $K_D$ amount to a weighting factor that increases the level of dissociation as a function of background interference. Because $\phi$ can easily be on the order of 10 or greater, $\Delta E'$ must be much smaller than $RT$ if a high incidence of engagement in nonproductive (and possibly harmful) complexes is to be avoided.

As will be discussed below, a typical value of $\Delta E$ for a binding interface is 10 kcal/mol, which implies $K_D = e^{-10/0.6} \simeq 0.058$ $\mu$M. Under these conditions and assuming no promiscuous binding, Equation 13.1b indicates that when equilibrium monomeric concentrations are 0.1, 1.0, and 10.0 $\mu$M, the expected fractions of dimers are $\simeq 0.63$, 0.94, and 0.99, respectively. Supposing, however, that $\phi = 10$ and that $\Delta E'/(RT)$ is 1 kcal/mol, then $K_D$ is modified to $K_D' \simeq 1.6$ $\mu$M, and for equilibrium monomeric concentrations of 0.1, 1.0, and 10.0 $\mu$M, the fraction of dimers is now reduced to 0.06, 0.38, and 0.86 respectively, owing to the fact that many monomers will be sequestered in inappropriate complexes.

**The physical features of interfaces.** Geometric considerations generally ensure that the interface between two globular molecules constitutes only a moderate fraction of the total surface area. Direct observations from a large number of proteins suggest a range of $\sim 5$ to 30%, with a weak positive linear scaling between interface size and total monomeric surface area (Chothia and Janin 1975; Jones and Thornton 1996; Bordner and Abagyan 2005; Lynch 2013). Whether the latter scaling is a simple consequence of geometry or a result of larger proteins requiring larger interfaces for stabilization remains unclear.

Although little work has been done on the adhesive features of homomeric interfaces, surveys of a diverse set of transient protein-protein interactions (e.g., antigen-antibodies, enzyme-inhibitors, etc., from a variety of taxa; Jones and Thornton 1996; Bogan and Thorn 1998; Horton and Lewis 1992; Kastritis et al. 2011) lead to two conclusions (Figure 13.5). First, the average interfacial binding strength is $\sim 18RT$ with a standard deviation of $5RT$ (i.e., about $18\times$ greater than background thermal energy). We can expect the binding strength of more permanent multimers to be somewhat greater, and Brooijmans et al. (2002) suggest an average of

$22RT$ for homodimer interfaces. Second, there is only a weak relationship between binding strength and total interface size, with the minimum observed interface area being $\sim 1000$ Å$^2$ (equivalent to 10 nm$^2$) and the minimum binding energy being $\sim 6$ kcal/mol, i.e., $10RT$. Drawing from a diverse set of observations, Bahadur et al. (2003) find an average interface area of $\sim 4000$ Å$^2$ for homodimers and $\sim 2000$ Å$^2$ for hetero-complexes.

An approximate mechanistic argument for the lower limit to a functional interface area can be derived as follows (Day et al. 2012). From Figure 13.5, it can be seen that a rough upper limit to the binding strength per Å$^2$ is $\sim 0.012$ kcal/mol. For two proteins to have any affinity at all, the total strength of the interaction must be at least equal to the energetic cost of simply holding the two interfaces in a particular orientation, which is on the order of 8.0 kcal/mol (Janin 1995; Zhang et al. 2008; Day et al. 2012). The minimum interface size necessary to cover this cost is $\sim 8.0/0.012 = 667$ Å$^2$, also in accord with Figure 13.5. Taking an average binding strength of a homodimeric interface to be $22RT$ (from above), Equation 13.2b then implies $K_D \simeq 0.3$ nM, which is of the order of magnitude of protein concentrations often seen in cells (Chapter 7).

Observed interfaces typically consist of no more than a few dozen amino-acid residues (Chothia and Janin 1975). For example, broad surveys over diverse organisms suggest means of $\sim 80$ to 100 for homodimer interfaces and 50 to 60 for heterodimers, with strong asymmetries in the distributions yielding lower modes than means (Bahadur et al. 2003; Zhanhua et al. 2005), such that most interfaces are in the range of 10 to 40 (Bordner and Abagyan 2005). However, these total sizes may exaggerate the number of actual residues involved in adhesion, as the total binding energy is usually concentrated in a small number of hot spots, with $< 10$ residues typically involved (Bogan and Thorn 1998; Hochberg et al. 2020; Pillai et al. 2020).

Why are homomeric structures so common? Many homomers utilize symmetric isologous interfaces rotated 180° with respect to each other, a feature that may reflect a simple geometric "two for the price of one" argument made by Monod et al. (1965). Supposing that residue A binds with a complementary residue B on a parallel flat interface, then rotating around this affinity pair by 180° will yield a complementary B-A match, thereby doubling the strength of binding with just a single mutational change. Extensions of this argument lead to the conclusion that proteins have an innate propensity for self-assembly relative to random surfaces (Ispolatov et al. 2005; Lukatsky et al. 2007; André et al. 2008).

The flip-side of this view is that an isologous interface is also expected to have a 50% reduction in the mutation rate to binding residues (because the mutational target size is reduced by 50%), and that any mutation eliminating a binding residue will have double the effect (because two binding pairs of sites are lost simultaneously). Thus, it remains unclear whether Monod's argument is a sufficient explanation for the excess abundance of isologous interfaces. Instead, it appears that symmetrical isologous interfaces are more common simply because mutations more generally confer stability to such configurations (André et al. 2008; Plaxco and Gross 2009).

Binding can come about in a multiplicity of ways, although utilization of hydrophobic residues and hydrogen bonds are the predominant factors. Interface hydrophobicity is generally intermediate to that of interior cores and exterior surfaces

of proteins, with a typical enrichment of the more hydrophobic and polar uncharged residues, most notably F, C, L, M, I, Y, W, and V (Bordner and Abagyan 2005; Levy 2010). A summary from diverse proteins and organisms indicates that the average numbers of hydrogen bonds are $\sim 11$ and 18 for heterodimer and homodimer interfaces, with $\sim 0.2$ per interface residue (Jones and Thornton 1995; Bahadur et al. 2003; Zhanhua et al. 2005).

Notably, the structural evolution of multimeric proteins involves more than the conjoining interfaces. For example, Marsh and Teichmann (2014) found that the subunits of multimeric assemblies are significantly more flexible than proteins operating exclusively as monomers. The average flexibility of subunits increases with the number of nonhomologous subunits in a heteromeric complex, and as a protein complex acquires new subunits over time, the consecutive additions tend to be increasingly flexible (Marsh and Teichmann 2014). The latter authors' interpretation of these patterns is that flexibility facilitates the conformational changes necessary for successful binding. There are, however, open questions of cause vs. effect here, as selection for monomer rigidity may be relaxed in a protein that is secondarily stabilized by a binding partner.

As in the case of folding stability (Chapter 12), once a highly refined level of binding efficiency has evolved, the fitness advantages of further improvement are expected to become increasingly negligible, rendering a situation in which excess capacity for binding remains unutilized. Several studies have been performed in which alanine residues are individually substituted for the amino-acid constituents of interfaces, and the resultant effects on binding strength typically have distributions in which the highest density is near zero, with $\sim 10\%$ having slightly positive effects on binding (Bogan and Thorn 1998). A number of examples also exist in which multimers have been engineered to have stronger interfaces and increased thermostabilities (Griffin and Gerrard 2012), consistent with the hypothesis of excess capacity.

Given that there are typically unutilized residues at most binding interfaces, these observations suggest that a large number of alternative and nearly energetically equivalent amino-acid compositions exist at individual interfaces. This further implies that interfaces can evolve in multiple ways with minimal effects on molecular functionality, allowing for considerable interface sequence wandering over evolutionary time (Figure 13.6). Consistent with this view, amino-acid sequences at interfaces do not evolve at unusually slow rates – on average not more than 50% more slowly, and sometimes slightly more rapidly than residues on exposed surfaces and in internal cores (Grishin and Phillips 1994; Caffrey et al. 2004; Bordner and Abagyan 2005; Mintseris and Weng 2005). Over time, this combination of multiple degrees of evolutionary freedom and the diminishing fitness advantages of increased binding can lead to effectively neutral interface evolution. The eventual result of such divergence is a situation in which the interfaces on molecules derived from different phylogenetic lineages evolve to become nonfunctional in a cross-species constructs (e.g., as mixtures of the monomeric subunits from two species).

The few attempts to evaluate heterologous compatibilities have led to mixed results, with interpretations made difficult by the fact that most examples focus on constructs where the catalytic function is built from the interface (which will reduce the opportunities for divergence). Cross-species molecular hybrids of thymidylate

synthase from *E. coli* and *Lactobacillus casei*, two very distantly related bacteria, appear to be fully functional (Greene et al. 1993). Functionality has also been demonstrated for hybrid ornithine decarboxylase dimers between the monomeric subunits of *Trypanosoma brucei* and mouse, whereas cross-species dimers between *T. brucei* and *Leishmania donovani* (both trypanosomes) are nonfunctional (Osterman et al. 1994). For triose phosphate isomerase, Sun et al. (1992) find that cross-species hybrid dimers involving mammals, chicken, and yeast have lower levels of catalytic efficiency and enzyme stability than within-species dimers (in this case, the catalytic site is not created by the interface). Careful comparative analyses of this sort involving a gradient of phylogenetic relationships, and extended to situations in which catalysis is independent of interface binding, will be essential to furthering our understanding of how interfaces evolve.

## Evolutionary Considerations

Given the penchant for most biologists to assume that virtually every aspect of biology owes its origin to natural selection, it will come as no surprise that most attempts to explain the existence of multimers start with this implicit assumption (e.g., Monod et al. 1965; Goodsell and Olson 2000; Marianayagam et al. 2004; Mei et al. 2004; Hashimoto et al. 2011; Griffen and Gerrard 2012). The proposed advantages are diverse. First, as noted in Chapter 12, it is generally easier to fold multiple small proteins than a single long one, and cases are known in which multimerization can actually enhance the folding rates and stability of the monomeric subunits (Zheng et al. 2012).

Second, the encounter rate of an enzyme and a small substrate is proportional to the effective radius of the enzyme (Chapter 7), and the elimination of extraneous protein surface may further enhance the frequency of productive encounters between catalytic sites and their substrates. However, these considerations need to be tempered by the fact that multimerization reduces the number of enzymatic complexes subject to diffusion. For example, a dimer with double the volume of a monomer would be expected to have a radius (and hence encounter rate) $\simeq 2^{1/3} \times$ that of the monomer but half the number of particles, which would reduce the encounter rate by a factor of $2^{-2/3}$, unless there were additional favorable factors involved.

Third, multimerization may have secondary advantages. For example, a smaller surface-area to volume ratio might reduce a protein's vulnerability to denaturation or engagement in promiscuous interactions (Bershstein et al. 2012). Higher-order structures might also reduce the sensitivity of catalytic sites to internal motions, thereby increasing substrate specificity. In addition, complexation offers increased opportunities for allosteric regulation of protein activity (with structural changes in one subunit induced by substrate binding being transmitted to another).

Finally, multimerization can sometimes lead to the creation of an entirely new function by inducing secondary structural changes that alter the nature of the catalytic site. For example, members of the archaea have a $CS_2$ anhydrase that converts $CS_2$ into $H_2S$ by a process similar to carbonic anhydrase's conversion of $CO_2$ to $HCO_3$, but whereas the latter is a monomer, the former has a hexadecameric architecture that prevents access of $CO_2$ to the catalytic site (Smeulders et al. 2013).

Weighted against these potential advantages, one must also consider the negative side effects that can result from proteins with a tendency to multimerize. First, unless a newly emerging dimer has a single isologous interface, concatenations into indefinite fibrils can arise. Human disorders involving the production of amyloid fibrils, such as Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis, and sickle-cell anemia, are prominent examples of the negative consequences of overly adhesive proteins (Chiti and Dobson 2009; Jucker and Walker 2013).

Second, as noted above, to achieve a critical concentration of an active multimer, the abundance of monomeric subunits must be raised to a high enough level to ensure an adequate number of encounters for successful complex assembly. Any increase in subunit production will entail an energetic cost, and highly expressed proteins are also vulnerable to promiscuous interactions (Semple et al. 2008; Vavouri et al. 2009). If, however, multimeric proteins are less vulnerable to thermo-instabilities and degradation, multimerization might actually result in a lower energetic demand, as the cost of protein replacement is reduced.

Third, even if a particular multimeric structure is advantageous, it need not follow that a new mutation to such a form can be easily promoted by selection, as the fixation process critically depends on the background conditions in which the mutation first appears. As will be noted below for diploid species, some dimerizing mutations may have deleterious effects when combined with ancestral monomeric proteins in heterozygous individuals (where potentially detrimental heterotypic complexation can occur). Because new mutations in diploid species are always present on a heterozygous background, such a process has the potential to greatly reduce the probability of establishment of a dimerizing allele.

In summary, although the large pool of multimeric structures in today's organisms cannot possibly be strongly maladaptive, this need not imply that they have arisen or are currently maintained by adaptive processes. Indeed, despite the plausibility of many of the above hypotheses, empirical evidence for the adaptive value of alternative multimeric structures is near nonexistent. Moreover, a number of examples can be pointed to in which a more complex structure seemingly operates no more efficiently in its lineage than a simpler structure in others.

For example, the mismatch-repair system, which plays important roles in replication fidelity, DNA repair, and recombination, is comprised of monomeric proteins in bacteria but dimers in eukaryotes (Kunkel and Erie 2005; Iyer et al. 2006), yet the repair efficiency of eukaryotic systems appears to be lower than than that in prokaryotes (Lynch 2011). Freist et al. (1998) point out that the overall fidelity of amino-acid loading of class I tRNA synthetase enzymes (which load specific amino acids onto cognate tRNAs) in translation tends to be much greater than that of the class II enzymes, despite the fact that the former operate as monomers and the latter as dimers. The sliding clamps used in DNA replication are homodimeric in bacteria but homotrimeric in eukaryotes, with both structures having very similar overall architecture (Kelman and O'Donnell 1995), yet replication-fork progression rates are nearly an order of magnitude faster in prokaryotes (Lynch 2007). A more practical example involves the insulin hormone, which normally operates as a homohexamer in humans, but has been engineered to become a monomer by incorporating just one or two amino-acid changes in the interface (Brange et al. 1988); in the treatment of diabetes, the monomers are absorbed much more rapidly than the multimers and

have equivalent efficacy, although they can sometimes lead to the production of amyloid fibers at the site of injection. Finally, although the ribosome has a much more complex protein repertoire in eukaryotes than in prokaryotes (Chapter 6), there is no evidence that translation fidelity is elevated in the former (Chapter 20).

Motivated by these observations, we now consider the more formal evolutionary theory essential to understanding how the evolution of alternative multimeric structures proceeds. Our focus here is primarily on situations in which the gene function is independent of the number of subunits in the complex, i.e., protein function is not a structural outcome of the interface. Quite different outcomes are expected when the interfaces between subunits are essential to function, as this strongly constrains the acceptability amino-acid substitutions (Abrusán and Marsh 2018). Although the basic theoretical issues appear clear, we are still a long way from formally testing the theory, which will require a combination of comparative analysis and experimental work on orthologous proteins across diverse phylogenetic taxa.

**Transitions from monomeric to higher-order states.** In one of the few experimental studies of the causes and consequences of dimerization, Hochberg et al. (2020) compared the features of estrogen and steroid receptors, two hormone-activated transcription factors that diverged following duplication of an ancestral gene at the base of the chordate lineage. The former operate as dimers, and the latter as monomers. By comparing the amino-acid sequences from diverse species, the authors were able to predict, reconstruct, and evaluate the functional properties of ancestral sequences. The key finding was that the alternative functional features of monomeric and dimeric family members are likely not a consequence of multimerization *per se,* but that following its initial establishment, the dimeric form gradually became more entrenched into this structural form as mutations to hydrophobic residues accumulated in the interface.

This proposed scenario is very similar to the kind of evolutionary trajectory postulated under the constructive neutral-evolution model (Chapter 6), which envisions the effectively neutral establishment of higher-order complexes as initially fortuitous interfaces create a permissive environment for the downstream accumulation of mutations that are only deleterious if exposed on the surface of a molecule. What remains to be achieved, however, is the ground-truthing essential to any evolutionary hypothesis – an understanding of the underlying population-genetic processes essential to this kind of transition from monomers to dimers (and vice versa).

Guided by knowledge on the physical features of interfaces noted above, we start with a relatively simple null model in which the architectural features of a homomeric protein can be described as a linear series of alternative structural states, with increasing binding strength of a potentially dimeric interface. Each class has a finite probability (per unit of evolutionary time) of transitioning to an adjacent class by gaining or losing a binding residue (Figure 13.7). The alternative molecular phenotypes range from the extreme case of a pure monomer (state $i = 1$) through a series of states ($i > 1$) in which the dimeric interface becomes increasingly stable owing to the establishment of adhesive amino acids.

Under this model, over evolutionary time, the molecular state will wander over the landscape of alternative phenotypes to a degree that depends on the evolutionary transition rates between states. However, provided the transition possibilities are

bidirectional, given enough time and constant rates, a quasi-steady-state distribution is expected to emerge. At this point, the long-term average evolutionary rates of movement into and out of each class are equal, which further implies that transitions towards a more dimeric state occur at exactly the same rate as transitions towards a more monomeric state. Such a long-term equilibrium condition results from the fact that more abundant states are so because they have lower rates of export to adjacent states, whereas less abundant states have higher rates of export. Given a constant set of conditions, the resultant equilibrium distribution describes both the long-term expected history of a protein in a particular phylogenetic lineage and the diversification across independent lineages.

A useful feature of this model is that the equilibrium probability of being in any particular state is proportional to the product of all of the rate coefficients pointing towards the state from both the upward and downward directions (Foundations 5.3). In addition, these transition coefficients have a relatively simple interpretation, with each being equivalent to the product of the number of mutations arising in the population per generation and the probability of fixation, which in turn depend on the directional powers of mutation and selection, relative to the stochastic force of random genetic drift.

Here, to illustrate the key points in as simple a manner as possible, it will be assumed that mutations destined to fixation arise infrequently enough that populations are almost always in a pure state of one form or another. We will also assume a constant upward mutation rate between classes, i.e., $\mu_{i,i+1} = u$ for all $i$, as would be closely approximated if the number of potential surface residues that can mutate to adhesive states is large. Letting $v$ be the mutational rate of loss of an adhesive residue, the downward mutation rate must increase linearly with the allelic state, where $i = 1$ denotes the monomeric state, i.e., $\mu_{i,i-1} = (i-1)v$, owing to the increase in the number of adhesive residues at the interface subject to loss with increasing $i$. As $u$ is an aggregate mutation rate over a large number of sites, and $v$ is a per-site mutation rate, we expect $u/v$ to be generally $> 1$, and perhaps greatly so.

The logical starting point is a neutral model in which each alternative state ($i$) has equivalent fitness (i.e, all dimers operate with equal efficiencies as each other and as monomers). Three features are immediately apparent under such conditions. First, because each adjacent pair of states is separated by a pair of upward and downward coefficients and all fixation probabilities are equal under neutrality (Chapter 4), the steady-state distribution depends only on the ratios of mutation rates (i.e., independent of their absolute values). Second, because such ratios decrease with increasing numbers of binding sites, the equilibrium frequencies of the higher states must eventually diminish toward zero, ensuring the existence of a quasi-steady-state probability distribution of the array of possible alleles. Third, because the number of mutations arising each generation is a function of the population size, and the probability of fixation equals the inverse of the population size (again assuming neutrality), the transition rates are entirely independent of population size. From Foundations 5.3, the equilibrium probability distribution simplifies to

$$\widetilde{P}_i = \frac{e^{-\alpha}\alpha^{i-1}}{(i-1)!}, \tag{13.4}$$

i.e., a Poisson probability distribution with a single parameter equal to the ratio of mutation rates, $\alpha = u/v$.

The probability of the extreme monomeric state ($i = 1$) is simply $\widetilde{P}_1 = e^{-u/v}$, which with $u/v = 0.01$, $0.1$, $1.0$, and $10.0$, respectively, becomes $\widetilde{P}_1 = 0.99$, $0.90$, $0.37$, and $0.000045$. Thus, even in the absence of adaptive differences among allelic states, the probability of a dimeric structure $(1 - \widetilde{P}_1)$ can be substantial. In addition, because the variance of a Poisson distribution is equal to the mean, for large $u/v$ substantial variation in interface binding strength is also expected among lineages (Figure 13.7), consistent with the observations noted above on variation for interfacial areas and binding strengths. Finally, given the independence of these results on population size, assuming the mutation ratio $u/v$ is reasonably constant across phylogenetic lineages, this neutral model predicts that the probability distributions of monomers and dimers should be independent of phylogenetic context, i.e., approximately the same in bacteria as in multicellular eukaryotes, which is consistent with the observations in Figure 13.3.

What are the consequences of selection towards one extreme or the other? Because the probability of fixation depends on the effective population size ($N_e$) when selection operates, we can anticipate a shift in the form of the distribution from the neutral expectation as the efficiency of selection increases with increasing $N_e$. The key modification that must be made is the weighting of the mutation pressure in each transition coefficient by the probability of fixation of the mutant allele, which is no longer the initial frequency.

There are numerous ways in which fitness might change over the phenotypic array displayed in Figure 13.7. Nonetheless, regardless of the fitness function, for each pair of upward and downward coefficients, the ratio of fixation probabilities in the two directions is equal to $e^{2N_e s}$ assuming haploidy ($e^{4N_e s}$ assuming diploidy), where $s$ is the relative fitness difference between the two adjacent states (Foundations 5.3). If one assumes weak positive selection in the direction of increasing $i$, such that the difference in selective advantage between adjacent states remains constant, Equation 13.4 still holds but with $\alpha = (u/v)e^{2N_e s}$, i.e., the Poisson distribution is maintained, except that the mutation-pressure ratio is multiplied by the selection-pressure ratio. If the prevailing selection pressure is in the direction of the monomeric state, then $\alpha = (u/v)e^{-2N_e s}$. The ratio of the power of selection to the power of drift, $s/(1/N_e) = N_e s$ (for haploidy, and twice that for diploidy) simply enters as an exponential weighting factor, pushing the system further in the favored direction than would be expected on the basis of mutation pressure alone. If the power of drift is substantially greater than that of selection, $1/N_e \gg s$, then $e^{2N_e s} \simeq 1$, and the distribution of state frequencies is still closely approximated by the neutral expectation.

This model can accommodate any alternative pattern of fitness variation among sites. For example, weakly dimerizing states might create structural defects not induced with strong binding interfaces, which could induce a sign change in $s$ with increasing $i$. Moreover, as noted above, one could argue that any incremental selective advantages with increasing numbers of binding residues will progressively decline, owing to the nonlinear relationship between binding energy and the dissociation constant. Any such modifications can be implemented readily through appropriate changes to the selection-pressure weighting terms, but would invariably lead to situations in which the long-term evolutionary distribution is no longer exactly Poisson.

The key point is that, regardless of the selection scheme, the model outlined in Figure 13.7 is sufficiently general to provide mechanistic insight into the driving forces leading to the so-called entrenchment of multimeric states caused by the accumulation of multiple binding residues. Indeed, the theory suggests that entrenchment is a natural expectation in the sense that the cumulative fixed mutations that promote a stronger interface will result in substantial structural defects if suddenly exposed, as for example in an experimental manipulation that exposes an entire hydrophobic patch on the monomer surface (in contrast to more natural single amino-acid substitutions that are more likely to occur in evolution, as envisioned in Figure 13.7).

The basic structure of this model can also be extended to the secondary evolution of homotetramers from homodimers (Lynch 2013). Notably, dimers appear to often constitute the initial step in the evolution of higher-order multimers. For example, Levy et al. (2008) found that whenever a tetramer in some lineage is related to a dimer in another, the dimer interface is usually conserved in the tetramer. This observation motivates the suggestion that tetramers typically evolve via an intermediate dimeric state, with the dominant dimeric interface evolving first (Levy et al. 2008; Dayhoff et al. 2010), followed by the joining of a potentially weaker interface to form a complex of four. Consistent with this view, the order of assembly events of higher-order multimers within cells appears to reflect the postulated order of evolutionary emergence of different interfaces (Marsh et al. 2013; Marsh and Teichmann 2015). Nonetheless, given the steady-state distributional properties noted above, there is need for care in interpretation here, as the sharing of interfaces between orthologous dimers and tetramers could in many cases reflect the reversion of the latter to the former. That is, the theory suggests that the net evolutionary flux rate from dimer to tetramer is equal to that in the opposite direction.

Regardless of the exact details of the model, two central points emerge from the preceding analyses, relevant not just to the issue of quaternary protein structure but to the distributions of all complex traits. First, as already noted for the case of neutrality, substantial phenotypic variation can even arise among lineages experiencing identical intensities of selection, demonstrating how risky it can be to assume adaptive explanations for phenotypic divergence among lineages. For example, if the composite parameter $\alpha = 1$, the probability of being in the pure monomeric state is 0.37, and that of being in the remaining dimeric categories is 0.63. That is, a substantial amount of phenotypic diversity would exist among phylogenetic lineages despite being confronted with identical evolutionary pressures, and any attempt to explain these differences in terms of imagined lineage-specific selection pressures would be quite misplaced.

Second, the most common state is not necessarily the optimum. Even with negative selection against multimers, they will still be common provided the mutational bias towards binding affinity is sufficiently large. The mode of the distribution is entirely determined by the composite parameter $\alpha$, and if $2N_e s \ll 1$, the prevailing molecular phenotype will be essentially defined by the mutation spectrum. Notably, mutation pressure in many phylogenetic lineages is biased toward the production of A/T nucleotides (Hershberg and Petrov 2010; Hildebrand et al. 2010; Lynch 2010; Long et al. 2018). This encourages a bias toward more hydrophobic, and hence more adhesive, surface residues (which, owing to the nature of the genetic code, have more

A/T-rich codons) (Knight et al. 2001; Bastolla et al. 2004; Hochberg et al. 2020).

In summary, the preceding analyses suggest that the substantial phylogenetic variation that exists in the multimeric states of proteins is not necessarily a consequence of idiosyncrasies in modes of selection in different lineages. Rather, it is an expected outcome of the stochastic evolutionary dynamics that arise in finite populations when the combined pressures of mutation and selection are not overwhelmingly large in one direction. If this hypothesis is correct, and one had the ability to sample a single evolutionary lineage over a very long period of time, orthologous proteins in different phylogenetic lineages would be observed to occupy various multimeric states in frequencies reflecting the underlying transition probabilities.

Although we do not have the luxury of making such observations directly, provided enough evolutionary time has elapsed for the Tree of Life to have reached the steady-state distribution, a corollary can be tested – the number of transitions from a monomeric to a dimeric state on the branches of a phylogeny should equal that in the opposite direction, and the same symmetry should hold for dimer-tetramer transitions, etc. Unfortunately, owing to the huge imbalance in the taxa and proteins with existing structural data, such a test cannot yet be made, and if such efforts are to be pursued, the phylogenetic sampling depth will need to be substantially greater than the expected transition times between alternative states. For example, if the likely transition rate between states is on the order of $10^{-8}$ per year, a focus on a lineage that diverged more recently than $10^8$ years ago would be uninformative.

**The domain-swapping model.** The preceding section focused on a gradualistic model for evolution along a gradient of adjacent allelic classes. However, transitions between monomeric and dimeric states can sometimes be precipitated by a major structural mutation that is fundamentally different than the single amino-acid substitutions envisioned above (e.g., an insertion or deletion; Hashimoto and Panchenko 2010; Plach et al. 2017).

One specific and frequently invoked mechanism for the origin of homodimers is encapsulated in the domain-swapping model (Bennett et al. 1994; Kuriyan and Eisenberg 2007), whereby a monomeric protein forming a closed loop with two interfacing domains (within the same polypeptide chain) is physically altered in such a way that intramolecular binding is no longer possible (e.g., by a major deletion in the linker between the two binding domains that prevents them from meeting). In principle, such a modification can be compensated by reciprocal domain swapping between monomeric subunits, the result being a dimer with two, rather than one, heterologous interfaces (Figure 13.8).

An attractive feature of this model is that the well-endowed binding domains already present in the ancestral protein do not have to go through a phase of incremental improvement. In addition, depending on the configuration of mutant monomers, the process envisioned here is not restricted to the origin of dimers, but extends to the establishment of higher-order multimers as well (Ogihara et al. 2001). Finally, as in the model introduced in the previous section, it is plausible that the process is bidirectional, with insertion mutations in the linker sequence sometimes causing reversion of a domain-swapping dimer to the monomeric condition.

There are plenty of seemingly plausible examples of domain-swapping proteins (Liu and Eisenberg 2002). However, the conditions required for such evolution are

particularly sensitive to the population-genetic environment. Here, we consider the simplest case in which an allele for the domain-swapping protein arises by a single deletion mutation that denies self-accessibility within the ancestral monomer. If the dimer is beneficial, such a mutant allele can readily proceed to fixation by positive selection in a haploid species. However, in a diploid species, the mutant allele will initially be present exclusively in heterozygotes, raising potential challenges for establishment.

The key issue is whether heterozygote fitness is compromised by the production of malfunctioning composites of the two alternative monomeric subunits, e.g., chimeras between proteins with and without a deletion. Presumably, the magnitude of any heterozygote disadvantage will depend on the rate of folding of the ancestral monomer and the overall cellular concentration of both allelic products, as slow folding and/or high concentration should magnify the likelihood of chimeric assemblies. Even normally well-behaved domain-swapping dimers can concatenate into harmful fibrils in some cellular environments. The details matter here because reduced fitness in heterozygotes presents a barrier to the spread of a mutant allele unless it can somehow rise to a high enough frequency that the production of beneficial homozygotes becomes likely.

Here, we simply provide a heuristic guide to the most salient issue – the low likelihood of an evolutionary sojourn through a bottleneck in mean population fitness unless the heterozygote disadvantage is overwhelmed by the power of random genetic drift (Lynch 2012). Of special interest is the critical effective population size ($N_e^*$) beyond which the efficiency of selection is so strong that there is effectively no possibility of making a transition to a domain-swapping allele.

With heterozygotes having a fractional fitness reduction of $\delta$, and domain-swapping homozygotes an advantage of $s$, under the assumption of random mating, mean population fitness reaches a minimum when the population frequency of the domain-swapping allele **a** is $\widehat{p} = \delta/(s + 2\delta)$ (Figure 13.8). This is an unstable equilibrium point. When the domain-swapping allele has frequency $< \widehat{p}$, it will be found essentially exclusively in heterozygotes, and will therefore act like a deleterious mutation being removed from the population at rate $\delta$, i.e., there will be net selection against the domain-swapping allele. However, if the domain-swapping allele has frequency $> \widehat{p}$, homozygotes will be sufficiently frequent that there will be net selection in favor of this allele. Thus, because the initial frequency of a novel domain-swapping allele is very low (on the order of the reciprocal of the absolute population size), the key issue is whether a mutant allele can drift against a gradient of negative selection up to frequency $\widehat{p}$, whereupon it becomes subject to net positive selection.

The population-size barrier to the establishment of the domain-swapping protein is

$$N_e^* \simeq \frac{s + 2\delta}{\delta^2}. \tag{13.5}$$

which reduces to $N^* \simeq 2/\delta$ if there is no selective advantage of the domain-swapping homozygote. For example, if the deleterious effect in heterozygotes is just 0.002, unless the effective population size is smaller than 1000, there is essentially no chance of establishment of an otherwise neutral domain-swapping allele. Even if the domain-swapping allele had a 1% advantage ($s = 0.01$), the barrier is still a

very small $N_e^* = 3500$. Thus, a small heterozygote disadvantage is a very strong impediment to the establishment of an allele that is advantageous when fixed.

To sum up, under the domain-swapping model, a transition from a monomeric to a dimeric state is most plausible under two sets of conditions: 1) a haploid population, in which case heterozygote disadvantage is never experienced; or 2) a diploid population in which selection against heterozygotes is inefficient, either because the effective population size is small (which allows selection to be overcome by drift) or because the reduction in heterozygote fitness is negligible. Unfortunately, although a knowledge of the fitness consequences of mixtures of the products of ancestral and derived alleles is essential to resolving how readily domain-swapping can evolve in diploid populations, there appear to be no data on this key issue or even on whether domain-swapping dimers confer greater or lesser fitness than monomers.

Notably, the theory presented above is entirely general in that a simple change in definition of terms is all that is required for considering the reverse transition of homodimer to monomer, a scenario that certainly cannot be ruled out on the basis of existing data. Indeed, for the simplest case in which there is no heterozygote disadvantage, if $u$ is the rate of mutation to dimers and $v$ is the reverse mutation rate, and $s_d$ is the selective advantage of dimers (negative if dimers are disadvantageous), the steady-state probability of being in the dimeric state is

$$\widetilde{P} = \frac{\alpha}{1+\alpha}, \tag{13.6}$$

where $\alpha = (u/v)e^{4N_e s_d}$, and $1 - \widetilde{P}$ is the probability of monomers. Note that this formula follows directly from the theory discussed in the preceding section, being the special case in which there are just two possible states (Foundations 5.3).

**Heteromers from homomers.** Although heteromers can, in principle, arise from promiscuous interactions among nonorthologous proteins, most seem to originate from interactions between paralogs arising from gene duplication. For example, Mcm1 is a transcriptional regulator that operates as a homodimer in many fungal species. Following duplication in *S. cerevisiae,* the paralogous copies acquired complementary mutations that cause heterodimer assembly; the loss of either duplicate is lethal, but ancestral homodimeric constructs are fully functional in *S. cerevisiae* (Baker et al. 2013). Pereira-Leal et al. (2007) found that following whole-genome duplication in yeast, many other homomeric complexes made a transition to heteromeric states. Likewise, experimental work involving historical reconstructions suggests that hemoglobin, deployed as a heterotetramer in a number of metazoan species, evolved from a homodimer, with the transition following gene duplication, and just two subsequent amino-acid substitutions being sufficient to confer a new binding interface (Pillai et al. 2020). A somewhat more complex scenario involves the vacuolar ATPase proton pump, which contains a hetero-hexameric ring composed of two components in most species, but three in fungi, where the additional participant is a paralog of one of the pre-existing components (Finnigan et al. 2012). Many other examples of homomer-to-heteromer transitions in eukaryotes are covered in individual chapters and summarized in Chapter 24.

Transitions to heteromeric structures might emerge in a variety of ways. For example, duplication might occur first in an ancestral gene with no intrinsic tendency

to form dimers, with secondary complementary mutations resulting in complexation of the paralogous products. However, an alternative, and perhaps more likely, scenario involves the situation in which an ancestral gene already engages in homodimerization and therefore has a well-established interface at the outset. The initial steps in developing a heteromeric interface would then require the accumulation of unique interface mutations in both paralogs so as to encourage heterodimerization while discouraging homodimer formation. Such a process might be facilitated if there were no intrinsic benefit to dimerizing, as this would eliminate any negative consequences of relinquishing homodimerization. On the other hand, without some form of reinforcement by selection, the long-term maintenance of the heterodimer would also be evolutionarily unstable owing to the fact that each locus would be subject to loss by degenerative mutations (Chapter 6).

There are several mechanisms by which reinforcement might occur. Suppose, for example, that each monomeric subunit from the ancestral gene had multiple, independently mutable subfunctions. Then, gene duplication followed by complementary degenerative mutations (the process of subfunctionalization) would lead to the joint preservation of both paralogs, with the evolved heterodimer still carrying out the combined subfunctions of the ancestral gene, but with the subfunctions partitioned to each subunit. Alternatively, if for other structural reasons an evolved heterodimer outperformed the homodimeric products of each individual locus, this could lead to positive selection for heterodimeric complexation provided the structural changes necessary for avoiding self-recognition were compatible with those for promoting heterodimerization (Marchant et al. 2019).

Notably, the latter scenario need not always await the arrival of new mutations affecting function (Lynch 2012). Consider, for example, the situation in which an ancestral locus encoding a homodimer harbors two alleles, **A** and **a**, such that the cross-product dimer created within heterozygotes elevates fitness beyond that for either of the two pure types produced in homozygotes. Heterozygous cells would be expected to produce three types of dimeric constructs (e.g., **AA**, **Aa**, and **aa**) in a binomial 1:2:1 ratio. Letting the fitnesses of the three genotypes at the ancestral locus be $1-s_1$, 1, and $1-s_2$, respectively, the two alleles will have been maintained in the ancestral (pre-duplication) population by balancing selection, with frequencies $s_1/(s_1 + s_2)$ and $s_2/(s_1 + s_2)$. There is, however, an intrinsic constraint with such a balanced polymorphism. Because the individuals with highest fitness (heterozygotes) always segregate equal numbers of both alleles into the next generation, there is no possibility for all members of the population to have a pure hetero-allelic state.

Gene duplication removes this barrier by providing the opportunity for each locus to become fixed for a different ancestral allele (Lynch et al. 2001). Once this point has been reached, then every member of the population would have the expression pattern found in the ancestral heterozygote (Figure 13.9) – fitter than the average member of the ancestral population, but in the early stages with every individual still producing the three dimeric types in a 1:2:1 ratio. Following the establishment of this complementing duplication state, subsequent mutational modifications involving the interfaces of one or both loci could then facilitate heterodimerization, eventually to the point at which homodimer assembly no longer occurs. This particular model is, of course, irrelevant for haploid species, which cannot harbor ancestral heterozygosity.

All of this being said, duplication of a homomerizing gene product need not always lead a heteromer. For example, Billerback et al. (2013) created subunit variants of the normally homomeric barrel-shaped, bacterial chaperone, GroEL (Chapter 14), and found that instead of assembling as heteromeric complexes, the resultant assemblies were individualized homomers. Moreover, Hochberg et al. (2018) found that following gene duplication most homomers actually evolve to avoid the construction of heteromeric complexes, apparently becoming preserved by either subfunctionalization or neofunctionalization of different homomeric complexes.

**Summary**

- Across the Tree of Life, at least 60% of proteins assemble into multimeric higher-order structures, with homomers being about twice as frequent as heteromers. Dimers are more common than tetramers, which are more common than hexamers, and so on, with odd-mers being under-represented relative to even-mers.

- These distributions are very similar across all phylogenetic groups of prokaryotes and eukaryotes, indicating a minimal gradient of molecular complexity (number of subunits) with organismal complexity. Moreover, orthologous proteins often have different numbers of subunits in different phylogenetic groups, and even when the level of multimerization is conserved, it is not uncommon for different taxa to utilize different binding interfaces between monomeric subunits.

- A substantial contributor to these patterns is the tendency of proteins to be naturally self-adhesive. This leads to a situation in which monomeric proteins are often just one or two amino-acid substitutions away from switching to a dimeric state (or vice versa) or to the production of harmful open-ended fibrils.

- The binding interfaces of multimers are relatively simple – typically involving fewer than ten key residues, and having binding strengths generally in the range of 15 to 25× the level of background thermal energy.

- Binding interfaces usually have an excess subpopulation of nonadhesive residues. This enables the specific binding sites of any particular lineage to wander in an effectively neutral fashion over evolutionary time, which in turn can lead to incompatibilities between subunits from divergent taxa.

- Although numerous adaptive explanations have been proposed for the widespread use of multimers, other than for the cases in which a new catalytic function is conferred by the interface or in which a functional cage or fiber is produced, there is very little direct evidence that multimeric proteins are selectively advantageous. Moreover, any proposed advantages must be weighed against several costs of

relying on multimers, such as the engagement in promiscuous binding and the necessity of producing elevated numbers of monomers. Resolution of all of these issues will require comparative, experimental work on orthologous proteins with different multimeric structures but conserved functions across the Tree of Life.

- Because of mutation bias towards adhesive amino-acid residues, there is an innate tendency for monomeric proteins to move in the direction of becoming homomeric multimers. This can gradually lead to a situation in which such complexes appear to be entrenched by reinforcing binding sites, even though the process need not have been driven by selection. As a consequence of such directionality and stochastic gains and losses, there can be a broad distribution of molecular phenotypes in different lineages exposed to identical processes of selection, mutation, and drift.

- A celebrated mode of origin of homodimers is domain-swapping, wherein a monomer containing two internal binding domains incurs a deletion in the linker that prevents self-binding and encourages assembly into dimers with two heterologous binding interfaces. The reciprocal route, in which an insertion in a linker in the latter encourages self-binding, is also possible. A major challenge of such transitions in diploid species is the possibility of harmful chimeric complexes between the two allelic products in heterozygous individuals, which imposes a barrier in mean population fitness that can only be overcome by a sufficiently small population size to enable drift across the fitness valley.

- Transitions from homomeric to heteromeric structures are commonly observed, although more so in eukaryotes, and usually following gene duplication with the sister genes then becoming specialized binding partners. Such transitions may initiate when there is a balanced ancestral polymorphism at the locus, with the heterozygote having superior fitness, and fixation of the hetero-complex only becoming possible after duplication enables each locus to adopt a particular ancestral allelic type.

**Foundations 13.1. Association / dissociation equilibria.** To understand a variety of issues with respect to reaction dynamics and equilibria, knowledge of a few basic features of molecular thermodynamics is required. Consider two molecules A and B, with the potential to join together to form a noncovalent complex AB, e.g., a dimer. In a steady-state environment, nearly all such systems will reach an equilibrium containing fixed relative concentrations of A, B, and AB. In this particular chapter, the focus is often on the special case in which A = B, i.e., the two molecules are of the same type, forming a homodimer, but the more general solution is given here. There are two ways to obtain the equilibrium solution.

The first approach takes a macroscopic view of the problem, using only information on the concentrations of the system components and their rates of interchange. Letting $k_{\text{on}}$ be the association rate of A and B to form AB, and $k_{\text{off}}$ be the reciprocal dissociation rate of AB to A and B, at equilibrium the rate of formation of AB must equal its rate of dissociation,

$$k_{\text{on}}[\text{A}][\text{B}] = k_{\text{off}}[\text{AB}], \tag{13.1.1}$$

where the quantities in brackets denote equilibrium concentrations. This general relationship is known as the law of mass action.

The dissociation constant $K_D$ (not to be confused with the dissociation *rate* $k_{\text{off}}$) is the ratio of the reverse and forward rates, which in turn relates to the ratio of reactant molecules under equilibrium conditions. Rearranging Equation 13.1.1,

$$K_D = \frac{k_{\text{off}}}{k_{\text{on}}} = \frac{[\text{A}][\text{B}]}{[\text{AB}]}, \tag{13.1.2}$$

defined at 1 M total concentrations of A and B. Although the underlying on/off rates, $k_{\text{on}}$ and $k_{\text{off}}$, will not be considered further here, it should be emphasized that these are dictated by the structural features of the molecular participants, which determine the rates of encounter and efficiency of binding (critical to $k_{\text{on}}$) and degree of complex stability (critical to $k_{\text{off}}$) (Kastritis and Bonvin 2012).

Letting $p_C$ denote the fraction of A molecules that are in complex, and rearranging and substituting from Equation 13.1.1, at equilibrium,

$$p_{\text{C}} = \frac{[\text{AB}]}{[\text{A}] + [\text{AB}]} = \frac{[\text{B}]}{([\text{A}][\text{B}]/[\text{AB}]) + [\text{B}]} = \frac{[\text{B}]}{K_D + [\text{B}]}. \tag{13.1.3a}$$

This expression shows that $K_D$ is equivalent to the equilibrium concentration of B at which half of the A molecules are in complex with B. If A = B, then

$$p_{\text{C}} = \frac{2[\text{AA}]}{[\text{A}] + 2[\text{AA}]} = \frac{2[\text{A}]}{([\text{A}][\text{A}]/[\text{AA}]) + 2[\text{A}]} = \frac{2[\text{A}]}{K_D + 2[\text{A}]}. \tag{13.1.3b}$$

because there are two A molecules within each complex.

The second approach takes a more detailed, thermodynamical view of the alternative states of the system. The key here is that from the perspective of a single A molecule, there are a number of potential microstates (the set of all possible configurations of the entire population of A and B molecules) involving the states of all B molecules in the system (some of which include a B molecule in complex with A). A classical result from the field of statistical mechanics is that the probability of a particular microstate $i$ of a molecular system is proportional to the function $e^{E_i/(K_B T)}$, where $E_i$ is the energy associated with the state (a more positive number implying a

more energetically favorable state), $K_B$ is the Boltzmann constant, and $T$ is the temperature in degrees Kelvin. At the molecular level, the thermodynamic stabilities of microstates are expressed relative to the background energy related to thermal motion of the solvent molecules ($k_B T$).

Consider the situation in which there are $N$ molecules of B for each molecule of A, with a B molecule having energy $E_{\text{on}}$ when bound to A and $E_{\text{off}}$ when free. Focusing on a specific molecule of A, the total energy of the system will then be $N E_{\text{off}}$ if A is unbound, and $E_{\text{on}} + (N-1)E_{\text{off}}$ if a particular A molecule is bound to a single B. The one significant remaining issue is the number of ways in which this one particular A molecule can be bound with various alternative molecules of B, relative to the number of microstates in which none of the local B molecules is bound. This turns out be simply equal to the concentration of B; see Phillips et al. (2013, p. 237-244) for an explicit derivation. To account for this effect, the probability of an individual AB association must be multiplied by [B], whereas the weighting factor for the situation in which A is unbound is just 1. Thus, an alternative way of expressing the probability that a molecule of A is complexed with B is

$$p_{\text{C}} = \frac{[B]e^{[E_{\text{on}}+(N-1)E_{\text{off}}]/(RT)}}{e^{NE_{\text{off}}/(RT)} + [B]e^{[E_{\text{on}}+(N-1)E_{\text{off}}]/(RT)}} \qquad (13.1.4\text{a})$$

$$= \frac{[B]e^{\Delta E/(RT)}}{1 + [B]e^{\Delta E/(RT)}} \qquad (13.1.4\text{b})$$

where $\Delta E = E_{\text{on}} - E_{\text{off}}$.

Note that there has been a change in notation here. The usual convention is to express molecular concentrations and energies associated with them in terms of molar quantities (mol/liter), and so $K_B$ has been scaled up to its molar equivalent $R$, which is simply $K_B$ times Avogadro's constant ($6.02 \times 10^{23}$ molecules / mol). With $R$ being equal to 1.987 cal $\cdot$ mol$^{-1}$ $\cdot$ K$^{-1}$, at standard temperature 25°C (equivalent to 298 K), $RT \simeq 0.6$ kcal/mol. Throughout, we will adhere to this as an approximate constant, as even a 25°C change in temperature alters $RT$ by $< 10\%$. For Equation 13.1.4b to work properly, $\Delta E$ must also have units of kcal/mol, and [B] must be the molar concentration of B.

Comparing Equations 13.1.3 and 13.1.4b shows that setting

$$K_D = e^{-\Delta E/(RT)} \qquad (13.1.5)$$

provides an alternative definition of the dissociation constant in thermodynamic terms. The binding-energy differential $\Delta E$ is positive for a pair of molecules with an energetically favorable interaction, so that with increasing affinity, $K_D \to 0$ and $p_{\text{C}} \to 1$.

The general approach leading to Equation 13.1.4b can be used to estimate the frequency of alternative states in any localized molecular system at equilibrium. In the example here, there are only two alternative states for any particular molecule, so the solution is relatively simple, but with multiple reactants, the book-keeping for alternative, combinatorial states becomes increasingly complex. The sum of terms in the denominator of Equation 13.1.4a is known as the partition function, as it insures that the probabilities of all possible microstates sum to 1.0. The overall set of probabilities for alternative states is generally referred to as the Boltzmann distribution. In this particular example, there are just two alternatives, A being bound to B with probability $p_C$, and A being unbound with probability $(1 - p_C)$.

Note that Equation 13.1.4b for the probability of one of two particular molecular states in a population of molecules is identical in form to Equation 13.6, which expresses the probability of one particular allelic state in a population (fully derived in Foundations 5.3). There is, thus, a remarkable convergence in the form of these statistical-mechanic and evolutionary-genetic equations, with the prefix terms ([B]

and $u/v$, respectively) being measures of the intrinsic pressure towards the state (owing to molecular concentration and mutation bias, respectively), and the exponential terms denoting the added pressure associated with energetic favorabilities and selective advantages.

**Literature Cited**

Abrusán, G., and J. A. Marsh. 2018. Ligand binding site structure influences the evolution of protein complex function and topology. Cell Rep. 22: 3265-3276.

Ahnert, S. E., J. A. Marsh, H. Hernández, C. V. Robinson, and S. A. Teichmann. 2015. Principles of assembly reveal a periodic table of protein complexes. Science 350: aaa2245.

André, I., C. E. Strauss, D. B. Kaplan, P. Bradley, and D. Baker. 2008. Emergence of symmetry in homooligomeric biological assemblies. Proc. Natl. Acad. Sci. USA 105: 16148-16152.

Bahadur, R. P., P. Chakrabarti, F. Rodier, and J. Janin. 2003. Dissecting subunit interfaces in homodimeric proteins. Proteins 53: 708-719.

Baker, C. R., V. Hanson-Smith, and A. D. Johnson. 2013. Following gene duplication, paralog interference constrains transcriptional circuit evolution. Science 342: 104-108.

Bastolla, U., A. Moya, E. Viguera, and R. C. van Ham. 2004. Genomic determinants of protein folding thermodynamics in prokaryotic organisms. J. Mol. Biol. 343: 1451-1466.

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

Bershtein, S., W. Mu, and E. I. Shakhnovich. 2012. Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations. Proc. Natl. Acad. Sci. USA 109: 4857-4862.

Billerbeck, S., B. Calles, C. L. Müller, V. de Lorenzo, and S. Panke. 2013. Towards functional orthogonalisation of protein complexes: individualisation of GroEL monomers leads to distinct quasihomogeneous single rings. Chembiochem. 14: 2310-2321.

Bogan, A. A., and K. S. Thorn. 1998. Anatomy of hot spots in protein interfaces. J. Mol. Biol. 280: 1-9.

Bordner, A. J., and R. Abagyan. 2005. Statistical analysis and prediction of protein-protein interfaces. Proteins 60: 353-366.

Brange, J., U. Ribel, J. F. Hansen, G. Dodson, M. T. Hansen, S. Havelund, S. G. Melberg, F. Norris, K. Norris, L. Snel, A. R. Sørensen, et al. 1988. Monomeric insulins obtained by protein engineering and their medical implications. Nature 333: 679-682.

Brooijmans, N., K. A. Sharp, and I. D. Kuntz. 2002. Stability of macromolecular complexes. Proteins 48: 645-653.

Caffrey, D. R., S. Somaroo, J. D. Hughes, J. Mintseris, and E. S. Huang. 2004. Are protein-protein interfaces more conserved in sequence than the rest of the protein surface? Protein Sci. 13: 190-202.

Chiti, F., and C. M. Dobson. 2009. Amyloid formation by globular proteins under native conditions. Nat. Chem. Biol. 5: 15-22.

Chothia, C., and J. Janin. 1975. Principles of protein-protein recognition. Nature 256: 705-708.

Day, E. S., S. M. Cote, and A. Whitty. 2012. Binding efficiency of protein-protein complexes. Biochemistry 51: 9124-9136.

Dayhoff, J. E., B. A. Shoemaker, S. H. Bryant, and A. R. Panchenko. 2010. Evolution of protein binding modes in homooligomers. J. Mol. Biol. 395: 860-870.

Finnigan, G. C., V. Hanson-Smith, T. H. Stevens, and J. W. Thornton. 2012. Evolution of increased complexity in a molecular machine. Nature 481: 360-364.

Freist, W., H. Sternbach, I. Pardowitz, and F. Cramer. 1998. Accuracy of protein biosynthesis: quasi-species nature of proteins and possibility of error catastrophes. J. Theor. Biol. 193: 19-38.

Garcia-Seisdedos, H., C. Empereur-Mot, N. Elad, and E. D. Levy. 2017. Proteins evolve on the edge of supramolecular self-assembly. Nature 548: 244-247.

Gartner, F. M., I. R. Graf, P. Wilke, P. M. Geiger, and E. Frey. 2020. Stochastic yield catastrophes and robustness in self-assembly. eLife 9: e51020.

Goodsell, D. S., and A. J. Olson. 2000. Structural symmetry and protein function. Annu. Rev. Biophys. Biomol. Struct. 29: 105-153.

Greene, P. J., F. Maley, J. Pedersen-Lane, and D. V. Santi. 1993. Catalytically active cross-species heterodimers of thymidylate synthase. Biochemistry 32: 10283-10288.

Griffin, M. D., and J. A. Gerrard. 2012. The relationship between oligomeric state and protein function. Adv. Exp. Med. Biol. 747: 74-90.

Grishin, N. V., and M. A. Phillips. 1994. The subunit interfaces of oligomeric enzymes are conserved to a similar extent to the overall protein sequences. Protein Sci. 3: 2455-2458.

Hashimoto, K., T. Madej, S. H. Bryant, and A. R. Panchenko. 2010. Functional states of homooligomers: insights from the evolution of glycosyltransferases. J. Mol. Biol. 399: 196-206.

Hashimoto, K., H. Nishi, S. Bryant, and A. R. Panchenko. 2011. Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization. Phys. Biol. 8: 035007.

Hashimoto, K., and A. R. Panchenko. 2010. Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. Proc. Natl. Acad. Sci. USA 107: 20352-20357.

Hershberg, R., and D. A. Petrov. 2010. Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet. 6: e1001115.

Hildebrand, F., A. Meyer, and A. Eyre-Walker. 2010. Evidence of selection upon genomic GC-content in bacteria. PLoS Genet. 6: e1001107.

Hochberg, G. K. A., Y. Liu, E. G. Marklund, B. P. H. Metzger, A. Laganowsky, and J. W. Thornton. 2020. A hydrophobic ratchet entrenches molecular complexes. Nature 588: 503-508.

Hochberg, G. K. A., D. A. Shepherd, E. G. Marklund, I. Santhanagoplan, M. T. Degiacomi, A. Laganowsky, T. M. Allison, E. Basha, M. T. Marty, M. R. Galpin, et al. 2018. Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions. Science 359: 930-935.

Horton, N., and M. Lewis. 1992. Calculation of the free energy of association for protein complexes. Protein Sci. 1: 169-181.

Ispolatov, I., A. Yuryev, I. Mazo, and S. Maslov. 2005. Binding properties and evolution of homodimers in protein-protein interaction networks. Nucleic Acids Res. 33: 3629-3635.

Iyer, R. R., A. Pluciennik, V. Burdett, and P. L. Modrich. 2006. DNA mismatch repair: functions and mechanisms. Chem. Rev. 106: 302-323.

Janin, J. 1995. Protein-protein recognition. Prog. Biophys. Mol. Biol. 64: 145-166.

Jones, S., and J. M. Thornton. 1996. Principles of protein-protein interactions. Proc. Natl. Acad. Sci. USA 93: 13-20.

Jucker, M., and L. C. Walker. 2013. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. Nature 501: 45-51.

Kastritis, P. L., and A. M. Bonvin. 2012. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. J. R. Soc. Interface 10: 20120835.

Kastritis, P. L., I. H. Moal, H. Hwang, Z. Weng, P. A. Bates, A. M. Bonvin, and J. Janin. 2011. A structure-based benchmark for protein-protein binding affinity. Protein Sci. 20: 482-491.

Kelman, Z., and M. O'Donnell. 1995. Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps. Nucleic Acids Res. 23: 3613-3620.

Knight, R. D., S. J. Freeland, and L. F. Landweber. 2001. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. Genome Biol. 2: RESEARCH0010.

Kunkel, T. A., and D. A. Erie. 2005. DNA mismatch repair. Annu. Rev. Biochem. 74: 681-710.

Kuriyan, J., and D. Eisenberg. 2007. The origin of protein interactions and allostery in colocalization. Nature 450: 983-990.

Levy, E. D. 2010. A simple definition of structural regions in proteins and its use in analyzing interface evolution. J. Mol. Biol. 403: 660-670.

Levy, E. D., E. Boeri Erba, C. V. Robinson, and S. A. Teichmann. 2008. Assembly reflects evolution of protein complexes. Nature 453: 1262-1265.

Levy, E. D., J. B. Pereira-Leal, C. Chothia, and S. A. Teichmann. 2006. 3D complex: a structural classification of protein complexes. PLoS Comput. Biol. 2: e155.

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

Long, H., W. Sung, S. Kucukyildirim, E. Williams, S. F. Miller, W. Guo, C. Patterson, C. Gregory, C. Strauss, C. Stone, et al. 2018. Evolutionary determinants of genome-wide nucleotide composition. Nat. Ecol. Evol. 2: 237-240.

Lukatsky, D. B., B. E. Shakhnovich, J. Mintseris, and E. I. Shakhnovich. 2007. Structural similarity enhances interaction propensity of proteins. J. Mol. Biol. 365: 1596-1606.

Lynch, M. 2007. The Origins of Genome Architecture. Sinauer Assocs., Inc., Sunderland, MA.

Lynch, M. 2010. Rate, molecular spectrum, and consequences of spontaneous mutations in man. Proc. Natl. Acad. Sci. USA 107: 961-968.

Lynch, M. 2011. The lower bound to the evolution of mutation rates. Gen. Biol. Evol. 3: 1107-1118.

Lynch, M. 2012. The evolution of multimeric protein assemblages. Mol. Biol. Evol. 29: 1353-1366.

Lynch, M. 2013. Evolutionary diversification of the multimeric states of proteins. Proc. Natl. Acad. Sci. USA 110: E2821-E2828

Lynch, M., M. O'Hely, B. Walsh, and A. Force. 2001. The probability of fixation of a newly arisen gene duplicate. Genetics 159: 1789-1804.

Marchant, A., et al. 2019. The role of structural pleiotropy and regulatory evolution in the retention

of heteromers of paralogs. eLife 8: e46754.

Marianayagam, N. J., M. Sunde, and J. M. Matthews. 2004. The power of two: protein dimerization in biology. Trends Biochem. Sci. 29: 618-625.

Marsh, J. A., H. Hernández, Z. Hall, S. E. Ahnert, T. Perica, C. V. Robinson, and S. A. Teichmann. 2013. Protein complexes are under evolutionary selection to assemble via ordered pathways. Cell 153: 461-470.

Marsh, J. A., H. A. Rees, S. E. Ahnert, and S. A. Teichmann. 2015. Structural and evolutionary versatility in protein complexes with uneven stoichiometry. Nat. Commun. 6: 6394.

Marsh, J. A., and S. A. Teichmann. 2014. Protein flexibility facilitates quaternary structure assembly and evolution. PLoS Biol. 12: e1001870.

Mei, G., A. Di Venere, N. Rosato, and A. Finazzi-Agrò. 2005. The importance of being dimeric. FEBS J. 272: 16-27.

Mintseris, J., and Z. Weng. 2005. Structure, function, and evolution of transient and obligate protein-protein interactions. Proc. Natl. Acad. Sci. USA 102: 10930-10935.

Monod, J., J. Wyman, and J. P. Changeux. 1965. On the nature of allosteric transitions: a plausible model. J. Mol. Biol. 12: 88-118.

Natan, E., T. Endoh, L. Haim-Vilmovsky, T. Flock, G. Chalancon, J. T. S. Hopper, B. Kintses, P. Horvath, L. Daruka, G. Fekete, et al. 2018. Cotranslational protein assembly imposes evolutionary constraints on homomeric proteins. Nat. Struct. Mol. Biol. 25: 279-288.

Nishi, H., K. Hashimoto, T. Madej, and A. R. Panchenko. 2013. Evolutionary, physicochemical, and functional mechanisms of protein homooligomerization. Prog. Mol. Biol. Transl. Sci. 117: 3-24.

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

Osterman, A., N. V. Grishin, L. N. Kinch, and M. A. Phillips. 1994. Formation of functional cross-species heterodimers of ornithine decarboxylase. Biochemistry 33: 13662-13667.

Pereira-Leal, J. B., E. D. Levy, C. Kamp, and S. A. Teichmann. 2007. Evolution of protein complexes by duplication of homomeric interactions. Genome Biol. 8: R51.

Phillips, R., J. Kondev, J. Theriot, and H. Garcia. 2013. Physical Biology of the Cell. Garland Science, New York, NY.

Pillai, A. S., S. A. Chandler, Y. Liu, A. V. Signore, C. R. Cortez-Romero, J. L. P. Benesch, A. Laganowsky, J. F. Storz, G. K. A. Hochberg, and J. W. Thornton. 2020. Origin of complexity in haemoglobin evolution. Nature 581: 480-485.

Plach, M. G., F. Semmelmann, F. Busch, M. Busch, L. Heizinger, V. H. Wysocki, R. Merkl, and R. Sterner. 2017. Evolutionary diversification of protein-protein interactions by interface add-ons. Proc. Natl. Acad. Sci. USA 114: E8333-E8342.

Plaxco, K. W., amd M. Gross. 2009. Protein complexes: the evolution of symmetry. Curr. Biol. 19: R25-R26.

Reid, A. J., J. A. Ranea, and C. A. Orengo. 2010. Comparative evolutionary analysis of protein

complexes in *E. coli* and yeast. BMC Genomics 11: 79.

Semple, J. I., T. Vavouri, and B. Lehner. 2008. A simple principle concerning the robustness of protein complex activity to changes in gene expression. BMC Syst. Biol. 2: 1.

Shionyu, M., K. Takahashi, and M. Go. 2001. Variable subunit contact and cooperativity of hemoglobins. J. Mol. Evol. 53: 416-429.

Smeulders, M. J., A. Pol, H. Venselaar, T. R. Barends, J. Hermans, M. S. Jetten, and H. J. Op den Camp. 2013. Bacterial $CS_2$ hydrolases from *Acidithiobacillus thiooxidans* strains are homologous to the archaeal catenane $CS_2$ hydrolase. J. Bacteriol. 195: 4046-4056.

Sun, A. Q., K. U. Yüksel, and R. W. Gracy. 1992. Interactions between the catalytic centers and subunit interface of triosephosphate isomerase probed by refolding, active site modification, and subunit exchange. J. Biol. Chem. 267: 20168-20174.

Vassylyev, D. G., H. Mori, M. N. Vassylyeva, T. Tsukazaki, Y. Kimura, T. H. Tahirov, and K. Ito. 2006. Crystal structure of the translocation ATPase SecA from *Thermus thermophilus* reveals a parallel, head-to-head dimer. J. Mol. Biol. 364: 248-258.

Vavouri, T., J. I. Semple, R. Garcia-Verdugo, and B. Lehner. 2009. Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. Cell 138: 198-208.

Wells, J. N., L. T. Bergendahl, and J. A. Marsh. 2015. Co-translational assembly of protein complexes. Biochem. Soc. Trans. 43: 1221-1226.

Wells, J. N., L. T. Bergendahl, and J. A. Marsh. 2016. Operon gene order is optimized for ordered protein complex assembly. Cell Rep. 14: 679-685.

Zabel, W. J., K. P. Hagner, B. J. Livesey, J. A. Marsh, S. Setayeshgar, M. Lynch, and P. G. Higgs. 2019. Evolution of protein interfaces in multimers and fibrils. J. Chem. Physics 150: 225102.

Zhang, J., S. Maslov, and E. I. Shakhnovich. 2008. Constraints imposed by non-functional protein-protein interactions on gene expression and proteome size. Mol. Syst. Biol. 4: 210.

Zheng, W., N. P. Schafer, A. Davtyan, G. A. Papoian, and P. G. Wolynes. 2012. Predictive energy landscapes for protein-protein association. Proc. Natl. Acad. Sci. USA 109: 19244-19249.

Zhanhua, C., J. G. Gan, L. Lei, M. K. Sakharkar, and P. Kangueane. 2005. Protein subunit interfaces: heterodimers versus homodimers. Bioinformation 1: 28-39.

**Figure 13.1.** Some examples of the varied forms of multimeric complexes. In each case, the subunits are presented in different colors, although each subunit might be encoded by the same genetic locus (homomers).
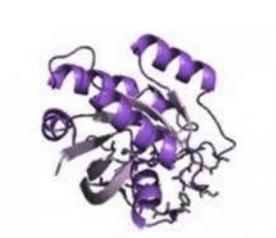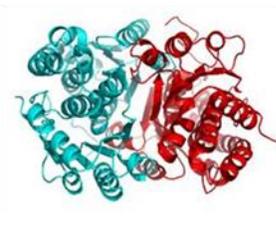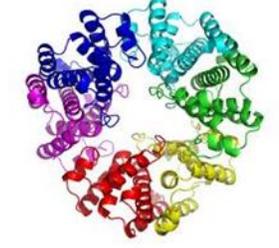


monomer
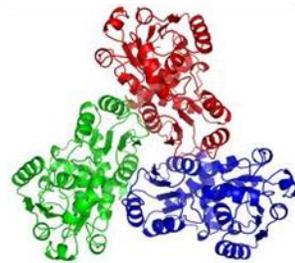


trimer



pentamer



dimer



tetramer



hexamer

**Figure 13.2.** Structures of monomers, dimers, and tetramers for the case of homomeric proteins. Blue-blue and red-red denote isologous interfaces, and blue-red denote heterologous interfaces. For a closed dimer, with one isologous interface, there is only one possible topology. A dimer with a heterologous interface (top) is subject to the development of open fibrils, unless it is geometrically configured to form a closed loop, in which case an odd-mer (e.g., a closed trimer) would have to be produced (as an even-mer would necessitate one interface being isologous). With two interfaces, a tetramer can take on two possible topologies: cyclical when heterologous interfaces attract, and dihedral when isologous interfaces attract. In the latter case, the tetramer can be viewed as a dimer of dimers.
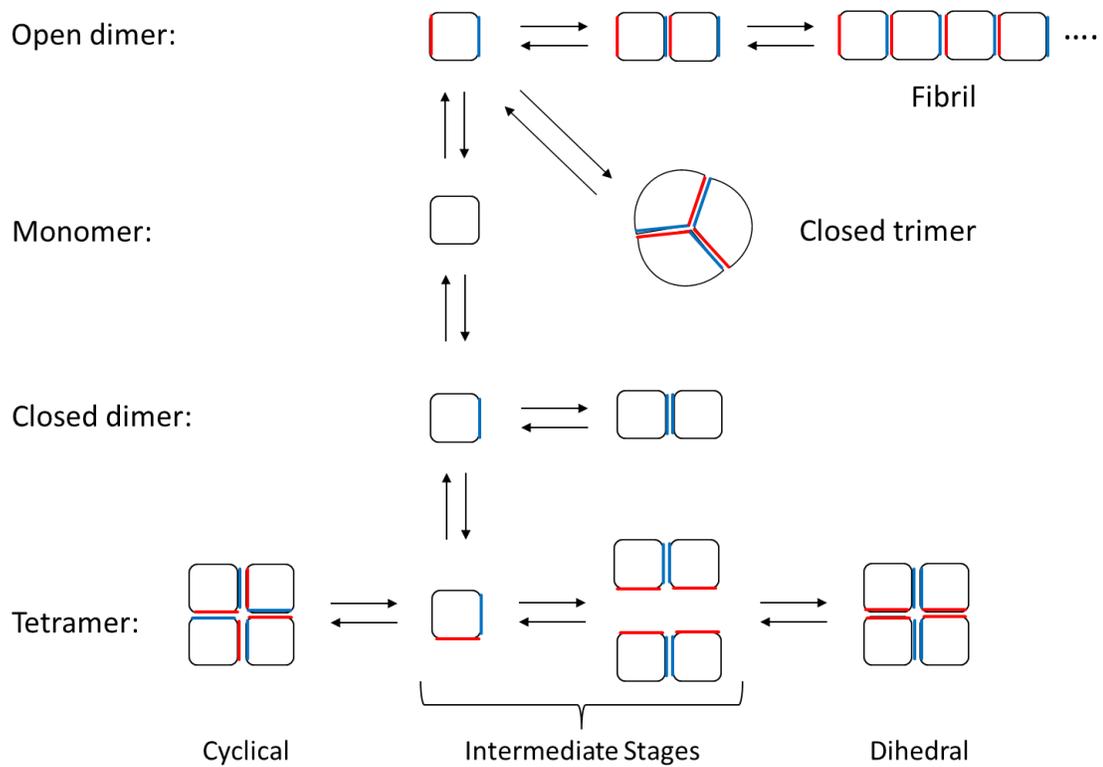
Figure 13.3. **Left**) Relative frequencies of the three main classes of protein structures in three phylogenetically broad groups. **Right**) Frequency distributions of levels of multimeric complexity (including monomers) for homomers for six major taxonomic groups; although the data are more limited, the distributions are similar for heteromers. The data are drawn from a diversity of proteins that have been structurally characterized in a broad range of organisms. From Lynch (2012).
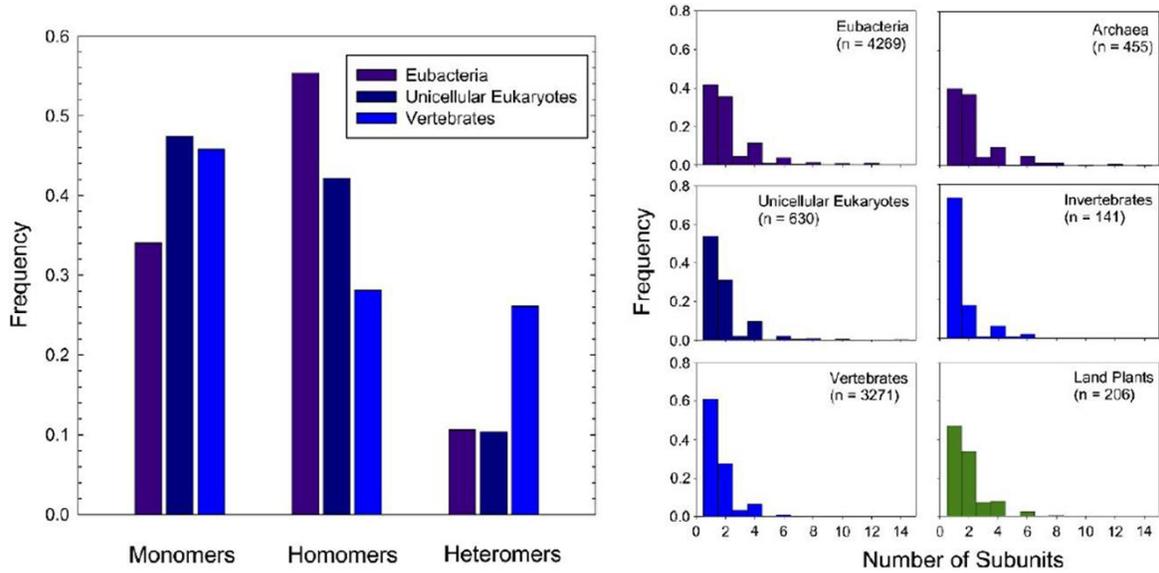
**Figure 13.4.** Structural topologies of heteromers involving two subunits per repeat (one bounded by black solid lines and the other by dashed lines). All interfaces in the structures shown here are heterologous, except the green-green pair for the tetramer on the left.

Heterodimer:

Tetramer
  (dimer of dimers):

Hexamer
  (trimer of dimers):
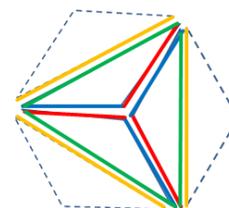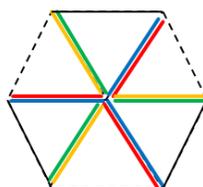
**Figure 13.5.** Binding strengths associated with interfaces of heterodimeric complexes. **Left)** The weak relationship between binding energy of a complex and the area of the interface. Solid circles (Kastritis et al. 2011); open circles (Horton and Lewis 1992). The dashed line denotes an approximate upper limit to the binding energy of $\sim 0.012$ kcal·mol$^{-1}$·Å$^{-2}$, where Å denotes Angstroms (equivalent to 0.1 nm). **Right)** An approximately normal distribution of interface binding strengths, with mean $= 18.3RT$ and SD $= 4.8RT$, where $RT$ is the energy associated with background thermal motion (Foundations 13.1). Data are extracted from a wide variety of sources (summarized in Kastritis et al. 2011).
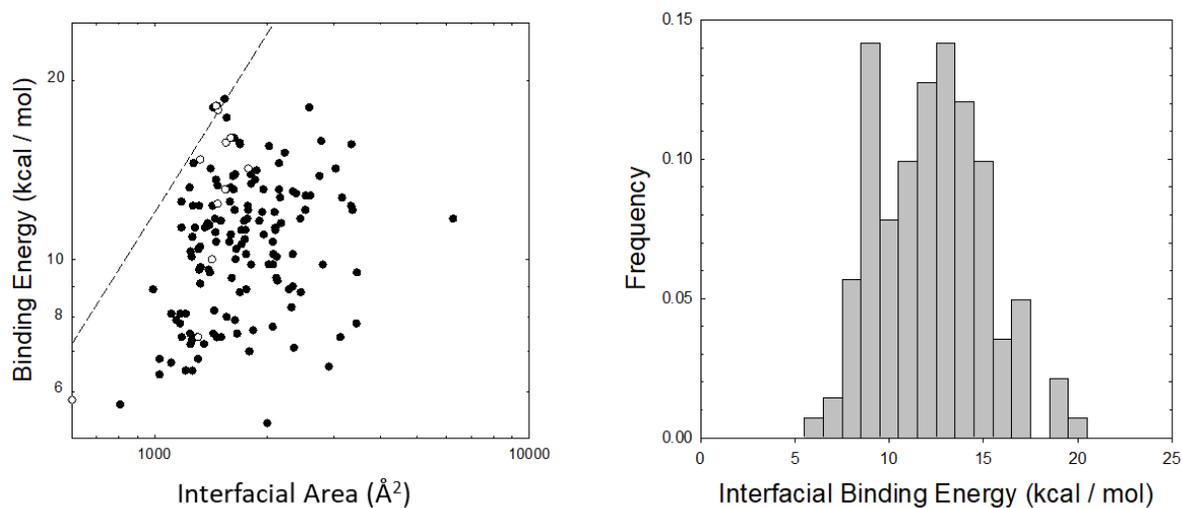
**Figure 13.6.** Effectively neutral evolution of new binding configurations at an interface patch. There are 25 potential binding residues in the patch, but at any point in time here, there are only six adhesive residues (in red). Over time, the specific residues involved in binding are free to wander over the surface, so long as six remain actively engaged. In principle, the binding number could wander above six by mutation and drift, although here it is assumed that six are sufficient to confer maximum fitness.
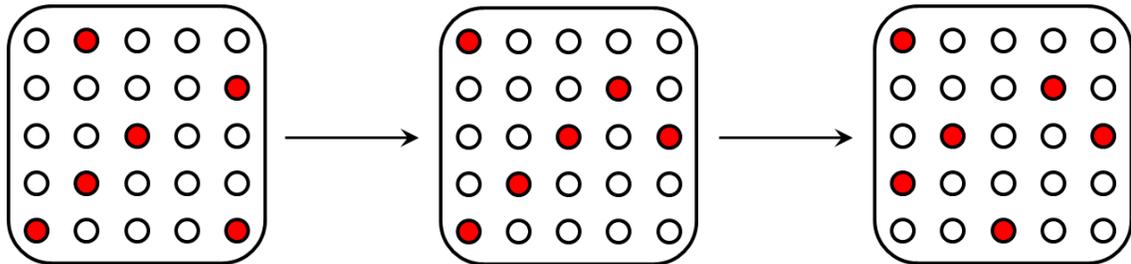
**Figure 13.7.** An idealized model for evolution along a linear array of alternative quaternary states of a protein. Here, the states describe a series of molecular variants for a homodimer with increasing numbers of binding sites at the interface, so that the state on the left is essentially a pure monomer, and moving to the right, the states have progressively stronger binding sites. The coefficients $m_{ij}$ are evolutionary rates of transition from state $i$ to state $j$, the values of which will depend on the power of mutation, random genetic drift, and selection, which together govern the composite parameter $\alpha$ in Equation 13.4. In effect, $\alpha$ is a measure of the net pressure from mutation and selection in the upward direction. The lower panel gives the equilibrium probability distributions of states $(i)$ for different values of $\alpha$ for the model described in the text.
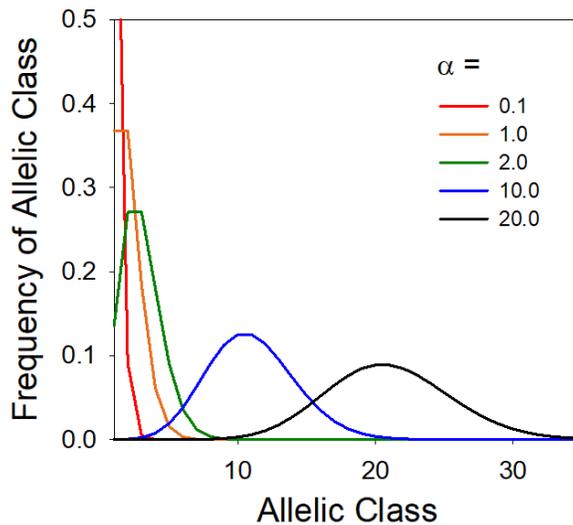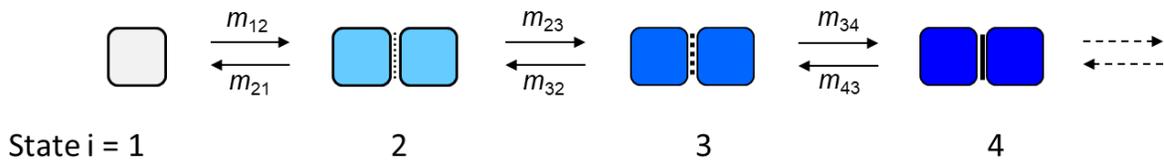
**Figure 13.8.** The domain-swapping model in a diploid population. **Left)** Relative to the ancestral monomeric type **A**, the domain-swapping allele a has a fitness deficit of $\delta$ in the heterozygous state, but advantage $s$ in the homozygous state ($s = 0$ meaning that the alternative fixed states are equivalent in fitness). The configurations following the arrows denote the assembly states of the proteins within diploid cells. **Right)** Such a scenario results in a fitness surface (as a function of the frequency of the domain-swapping allele) with a valley at an intermediate frequency $\hat{p}$ for the domain-swapping allele, owing to the fact that low-frequency alleles are present almost exclusively in deleterious **Aa** heterozygotes. If the frequency of allele a is to the left of the bottom of the valley, the prevailing pressure of selection is to remove the allele from the population, whereas to the right of the valley, selection promotes allele a, potentially driving it to fixation. Two examples are shown, both with $s = 0.01$, and with $\delta = 0.01$ for the dashed curve and $0.10$ for the solid curve.
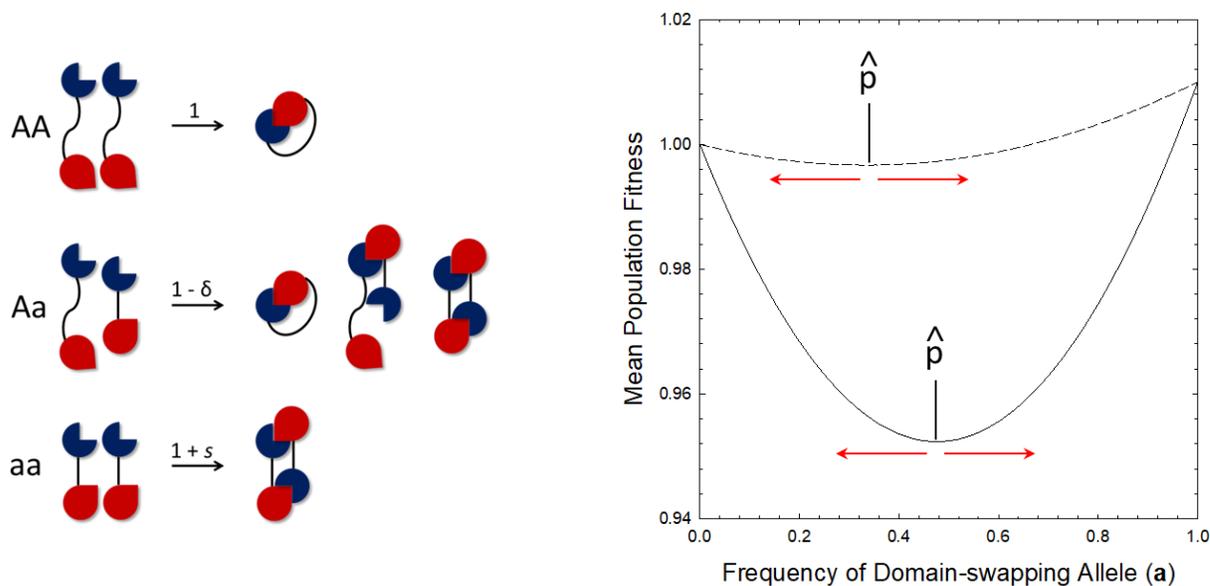
**Figure 13.9.** A potential path to the evolution of heterodimers from a homodimeric state following gene duplication at a heterozygous locus. Solid circles represent individual proteins derived from alternative alleles, which together make dimers consisting of **A** (blue) and/or **a** (red) subunits (four such complexes are shown per cell); ancestral heterozygotes (prior to duplication) produce three types of dimers in a 1:2:1 ratio, assuming random assembly. Following gene duplication, each locus becomes fixed for an alternative allele, and the encoded products at each locus subsequently diverge to the point that self-assembly is avoided, leading to "fixed heterozygosity."