Interactions with Other Molecules

Lecture on March 21, 2016

Interactions with Other Molecules

- The biological functions of proteins almost always depends on their physical interaction with other molecules.
- Proteins interact with a wide variety of molecules: other proteins, nucleic acids, polysaccharides, and lipids.
- They bind with varying specificity and affinities.
- Proteins are generally classified based on their function and consequences of their binding :
- The molecule interacting with the protein is referred to as the **ligand**.
- Protein with bound ligand is known as the **holo** form. (without the ligand, the **apo** form).

Structural proteins

Enzymes

Repressors

Lectin

Membrane transport proteins

Toxins

Hormones

Receptors

Immunoglobulins

Motility

etc....

General Properties of Ligand Binding

- Most cases involve a single defined binding site on each polypeptide chain for a specific ligand.
- While a protein domain is capable of binding more than one ligand (unusual to find more than 2-3), proteins that bind a number of different ligands often bind them on different domains.
- Binding sites are frequently sizable depressions on the protein surface and the interacting surface between the ligand and the protein is maximized.
- Interactions between proteins and ligands demonstrate both steric and physical complementarity.
- Ligands that do not need to dissociate often or at all tend to be bound deep within the protein interior (heme groups and some prosthetic groups).
- The protein-ligand interface is usually as closely packed as the protein interior. Polar groups in the interface are paired in hydrogen bonds and electrostatic charged groups are neutralized.

General Properties of Ligand Binding

- Lock and key analogy.... a bit too rigid, but the general idea fits. A defined protein structure is required for specificity in ligand binding. A very malleable protein would be likely to adjust its shape to accommodate a variety of ligands and specificity would be compromised.
- Structure of a protein domain usually does not change significantly with ligand binding.
- Substantial changes in protein structure upon ligand binding are usually associated with the motion of rigid domains or subunits relative to each other. In some cases the binding site lies at the interface between domains, and the domains shift position to bind the ligand.
- Ligand binding may also induce domain shifts to produce functional alterations at other sites in the protein, as in allosteric proteins.
- From these general principles, it is often possible to guesstimate the structure of a protein-ligand complex, if the structures of the free protein and ligand are known.

Binding Energetics and Dynamics

- A fundamental aspect of the interaction between a protein and its ligand is the affinity of the two for each other.
- Observed ligand affinity varies enormously ranging from very high where dissociation is negligible to very low values where the ligand concentration required for binding is so great that its relevance is questionable.
- If ligand affinity is very great and the protein is likely to be isolated with bound ligand, and the ligand is small, then it is called a prosthetic group.

Binding Affinities

- The affinity between a protein (P) and a ligand (A) is measured by the association constant (K_a) for the binding equilibrium.
- The greater the value of K_a the greater the affinity. K_a has units of conc^{-1}.
- In defining K_a, all species are assumed to be present in sufficiently low concentrations for thermodynamic ideality to apply.
- Often easier to refer to affinities in terms of dissociation constant (K_d). [the reciprocal of K_a and units of concentration]
- Simplest at ligand concentrations much greater than protein binding sites because ligand binding will not significantly impact conc. of free ligand.

$$P + A \underset{}{\overset{K_a}{\rightleftharpoons}} P{\bullet}A$$

$$K_a = \frac{[P{\bullet}A]}{[P][A]}$$

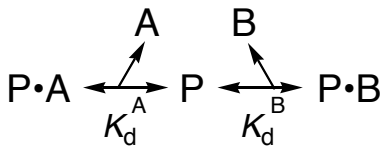
Ratio of bound to free protein is proportional to the free ligand concentration.

$$\frac{[P{\bullet}A]}{[P]} = K_a[A]$$

$$y = \frac{[P{\bullet}A]}{[P] + [P{\bullet}A]} = \frac{K_a[A]}{1 + K_a[A]}$$

y = fraction of protein molecules that are bound to ligand.

Binding Affinities



- Specificity of a protein for one ligand over another depends on their relative affinities.
- If two ligands compete for the same binding site then the ligand with higher affinity will be bound to a correspondingly greater extent when the ligands are present in equal concentrations.
- In such competitive binding, low affinity can be overcome with increased ligand concentration.
- Binding energetics described by the Gibbs free energy of binding (ΔG_{bind}).
- Energetics more explicitly defined as the difference in free energy between the protein without and with bound ligand (ΔG_b). Here, free ligand concentration must be specified.
- Ligand binding may serve to stabilize the protein against denaturation.

$$[P] = \frac{[P\bullet A]\,K_d^A}{[A]} = \frac{[P\bullet B]\,K_d^B}{[B]}$$

$$\frac{[P\bullet A]}{[P\bullet B]} = \frac{K_d^B[A]}{K_d^A[B]}$$

Gibbs free energy of binding

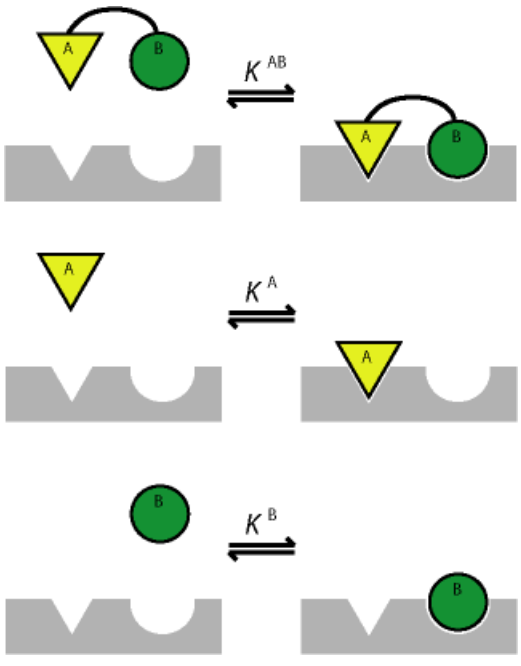
$$\Delta G_{bind} = -RT\ln K_a = RT\ln K_d$$

Difference in free energy between protein without and with ligand.

$$\Delta G_b = -RT\ln(K_a[A]) = RT\ln\left(\frac{[A]}{K_d}\right)$$

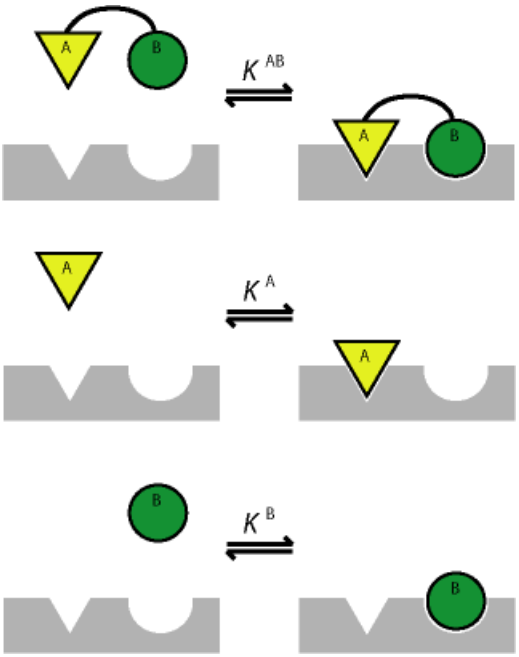
Relative Affinities

- Qualitative insights into ligand affinity and specificity in some cases can be elucidated based on the protein and ligand structures.
- Currently it is impossible to rationalize K_a and K_d values of a protein for a ligand based on their structures and first principles.
- Observed affinities depend on the relative free energies of the complex and of the components.
- Not only a function of the favorable interactions between the protein and the ligand but also conformational changes required for optimal interaction, any differences in their various interactions with solvent, loss of translational and rotational freedom of each component and the displacement of solvent or other molecules that are displaced with binding.



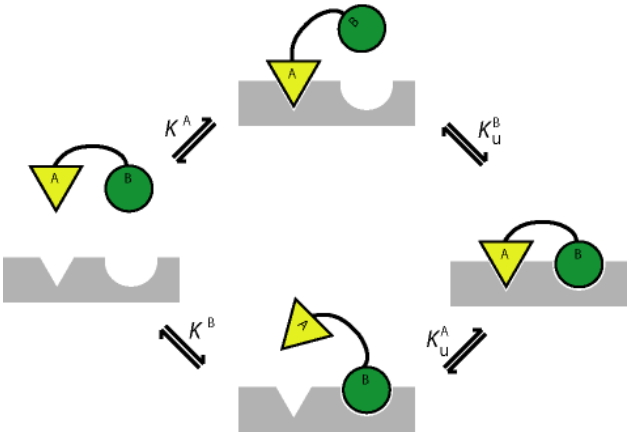
Relative Affinities

- Interactions between a protein and a ligand always involve a substantial number of groups and multiple points of interaction.
- The approach to analyzing the observed affinity has been to dissect it into the contributions of each group by measuring the effect their removal has on affinity.
- It might be thought that the observed binding energy is the sum of the contributions of each group involved, but it is not that easy.
- It is generally the case that the contribution of the various groups to binding is not additive.



Relative Affinities

- The nonadditive nature of specific binding affinities becomes clear if ligand binding is the result of cooperative interactions.
- The first step in the binding of the ligand should be analogous to the binding of each part present alone... bimolecular.
- The second and successive binding events are essentially unimolecular.
- The incremental binding energy contributions calculated in this way give a measure of increased affinity associated with the presence of each group.



K^{A} and K^{B} are bimolecular binding constants

K^{A}_{u} and K^{B}_{u} are unimolecular equilibrium constants

$$K^{AB}_{a} = K^{A}_{a}K^{B}_{u} = K^{B}_{a}K^{A}_{u}$$

$$K^{A}_{u} = \frac{K^{AB}_{a}}{K^{B}_{a}} \qquad K^{B}_{u} = \frac{K^{AB}_{a}}{K^{A}_{a}}$$

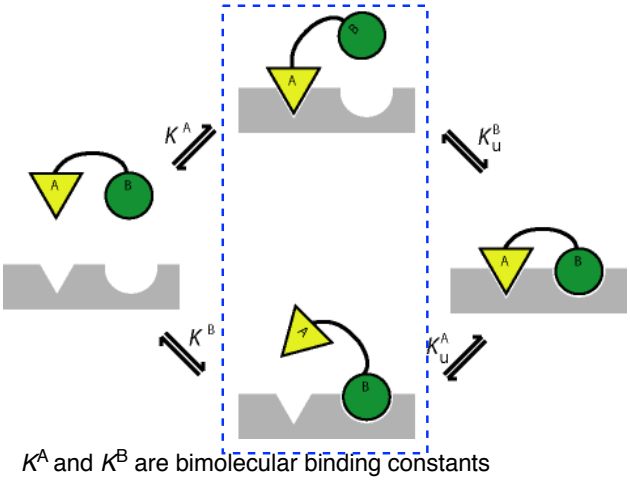
$$\Delta\Delta G^{A}_{i} = -RT\ln K^{A}_{u} = -RT\ln \frac{K^{AB}_{a}}{K^{B}_{a}}$$

$$\Delta\Delta G^{B}_{i} = -RT\ln K^{B}_{u} = -RT\ln \frac{K^{AB}_{a}}{K^{A}_{a}}$$

Free energy contribution of each moiety to binding

Relative Affinities

- The effective concentration of a group in the intermediate complex can conceivably be zero depending on its positioning in the complex.
- Or... the effective concentration can be up to 10^{10}M when the group is optimally positioned in the intermediate complex.
- The large values of effective concentration in intramolecular reactions results from the entropic effect of the covalent linkage.
- Therefore, the greater the rigidity between groups of a ligand, the greater the entropic contribution to the effective concentration is likely to be.



K^A and K^B are bimolecular binding constants

K_u^A and K_u^B are unimolecular equilibrium constants

$$K_a^{AB} = K_a^A K_u^B = K_a^B K_u^A$$

$$K_u^A = \frac{K_a^{AB}}{K_a^B} \qquad K_u^B = \frac{K_a^{AB}}{K_a^A}$$

$$\Delta\Delta G_i^A = -RT\ln K_u^A = -RT\ln \frac{K_a^{AB}}{K_a^B}$$

$$\Delta\Delta G_i^B = -RT\ln K_u^B = -RT\ln \frac{K_a^{AB}}{K_a^A}$$

Free energy contribution of each moiety to binding

Relative Affinities

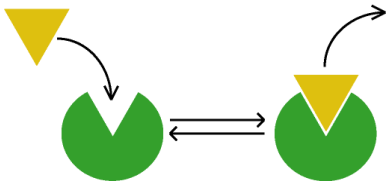
- The effective concentrations of the different groups/parts of ligands are likely to vary substantially in different ligands.
- Therefore it is reasonable to expect the contributions to binding of a hydrogen bond, van der Waals interaction, etc... will not be constant and will be different for each ligand-binding interaction.
- Interpreting binding data in terms of individual interactions is not straightforward.
- Proteins can discriminate very effectively between their proper ligand and one that lacks just one small element.
- There are limits to the specificity of binding, which are set by the energetics of the interaction between the proteins and ligands.
- *These limits can be exceeded if extreme specificity is required.*

Rates of Binding and Dissociation

- Ligand binding is not just a matter of association constant also defined by rates of binding (k_a) and dissociation (k_d). (on/off rates)
- Association constant is the ratio of k_a and k_d.
- These rate constants can vary considerably. Rate of binding influenced by the sizes of the protein and the ligand and any conformational changes associated with binding.
- Many small ligands observed to bind rapidly... with rates approaching diffusion control (k_D). Under normal conditions, k_D expected to be in the range of 10^9 $M^{-1}s^{-1}$.
- If two molecules attract or repel each other, it can affect the binding rate.
- Binding rates lower than k_D suggest that binding requires very specific orientation of the protein and ligand or that changes occur during binding that results in a multistep association reaction.

$$P + A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} P{\bullet}A$$

$$K_a = \frac{k_a}{k_d}$$



$$k_D = 4\pi N_A (D_P + D_A) r_{PA}$$

D_P = protein diffusion coefficient

D_A = ligand diffusion coefficient

N_A = Avagadro's number

r_{PA} = minimum distance at which interaciton occurs

Rates of Binding and Dissociation

- *Binding sites on proteins generally bind ligands in defined spatial orientations.*
- *The binding sites of proteins usually comprise only a small fraction of the total protein surface.*
- This would suggest that the majority of protein-ligand encounters would be unproductive and binding would be a relatively slow process.
- This is not necessarily the case: Electrostatic interactions and asymmetric distributions of charges on the protein surface may serve to draw in, guide and orient ligands for binding.
- Ligand binding likely involves a multistep process involving the diffusion-controlled formation of an unstable *encounter* complex (which is usually not observed). This is followed by rearrangement to form the final complex.
- This would result in the observed binding rate being slower than diffusion-controlled. The encounter complex would disassociate faster than binding is completed ($k_{-1}>k_1[A]$ and $k_{-1}>k_2$). Such a system would allow the two species to orient for binding.

$$P + A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} P{\bullet}A$$

$$K_a = \frac{k_a}{k_d}$$

$$P + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \{P{\bullet\bullet\bullet}A\} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} P{\bullet}A$$

Rates of Binding and Dissociation

- Even in cases where the observed rate of association is apparently diffusion-controlled, additional steps may follow initial association.
- Very little is known about what occurs structurally during the course of ligand binding. In most cases, there is considerable scope for conformational changes in both the protein and the ligand.
- Diffusion-limited encounters occur between all molecules in solution.
- Stable and specific binding is reflected primarily in slow rates of dissociation (k_d).
- Energetically favorable rearrangements associated with initial association and binding decrease the apparent rate of dissociation.

$$P + A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} P{\bullet}A$$

$$K_a = \frac{k_a}{k_d}$$

$$P + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \{P{\bullet\bullet\bullet}A\} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} P{\bullet}A$$

Rates of Binding and Dissociation

- In solution, molecules are free to diffuse in three dimensions, and the rate of diffusion-limited association depends on the size of the target.
- This is not the case when the diffusion is confined to two or one dimensions.
- Mean diffusion times can be significantly shorter in two and one dimensions.
- The diffusion of proteins in cellular membranes represents two-dimensional diffusion.
- One-dimensional diffusion describes the movement of proteins along DNA. Following initial association of the protein with DNA, the protein then moves alone the DNA. Such a strategy would aid these proteins in finding very specific sequences.

$$\tau_3 = \left(\frac{R^2}{3D_3}\right)\left(\frac{R}{a}\right) \qquad \text{in 3 dimensions}$$

$$\tau_2 = \left(\frac{R^2}{2D_2}\right)\ln\frac{R}{a} \qquad \text{in 2 dimensions}$$

$$\tau_1 = \left(\frac{R^2}{3D_1}\right) \qquad \text{in 1 dimensions}$$

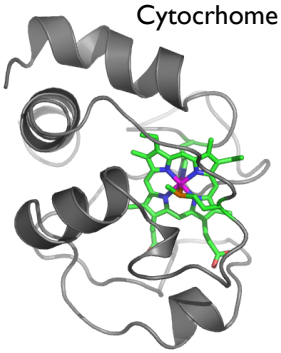
The mean diffusion times τ to reach a small target of radius a in the middle of a space of radius R ($R>>a$) are given in the above equations.

D_i (i=1,2,3) are the diffusion coefficients for the indicated dimensions.

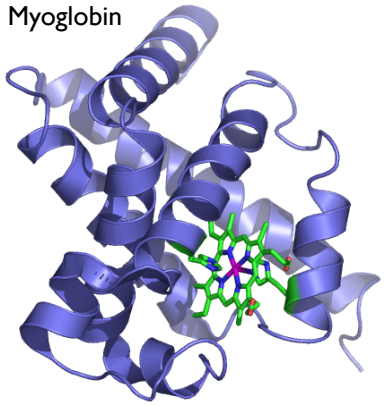
Protein Conformation and Binding

- The three-dimensional structure of a protein provides and ideal scaffold for displaying and positioning amino acid side chains and the resulting unique surfaces.
- Binding sites comprise relatively little of the structure of most proteins.
- These binding sites are often small patches on the protein surface. Small localized alterations can have significant impact on ligand affinity and specificity.
- Homologous proteins can bind different ligands: *Perhaps best illustrated by the immunoglobulins which share the same structural framework, yet bind an immense variety of antigens.*
- Conversely, proteins that share little sequence similarity or even structural similarity may bind the same ligand. *Such is the case in the binding of the heme group by proteins of the globin family, cytochrome c-like proteins and cytochrome b_5, cytochrome b_{562} and cytochrome c_3.*
- Therefore the ability to bind the same ligand cannot be used to imply that two proteins are related.

Cytocrhome C



Myoglobin



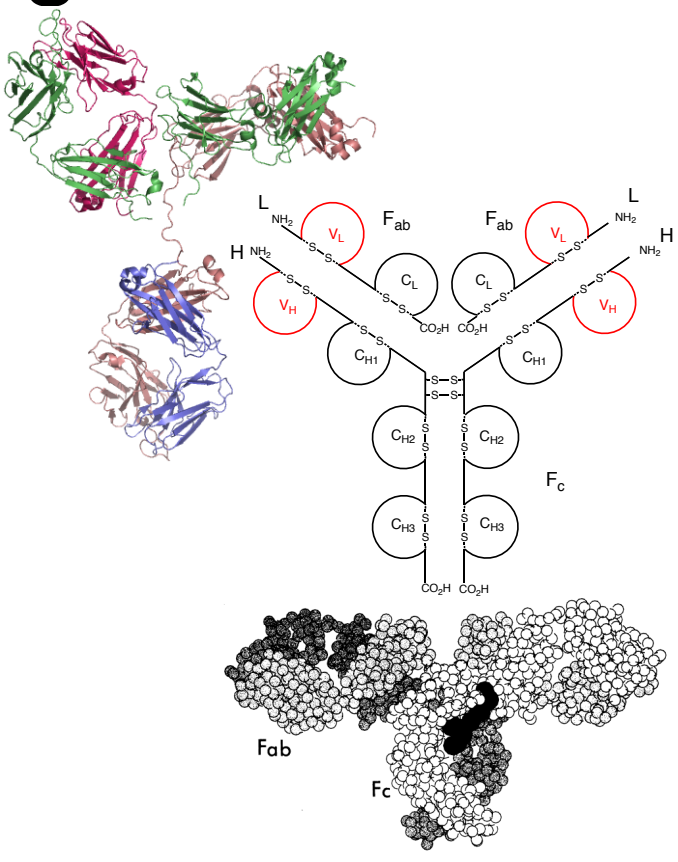
Protein Conformation and Binding

- Members of protein families do tend to bind similar ligands.
- Proteins within each family are known to be evolutionarily related.
- The evolutionary relationship is not always obvious. Proteins that bind similar ligands appear to have on occasion arisen through convergent evolution.
  - The numerous 8-fold αβ barrel proteins represent the most likely case of convergent evolution. They share a common three-dimensional structure and tend to bind negatively charged ligands at similar positions at the same end of the barrel.
  - Nucleotide binding proteins provide another example of convergent evolution. Nucleotide-binding proteins utilize similar structural motifs for ligand binding, yet they have no significant sequence similarity.

Immunoglobulins

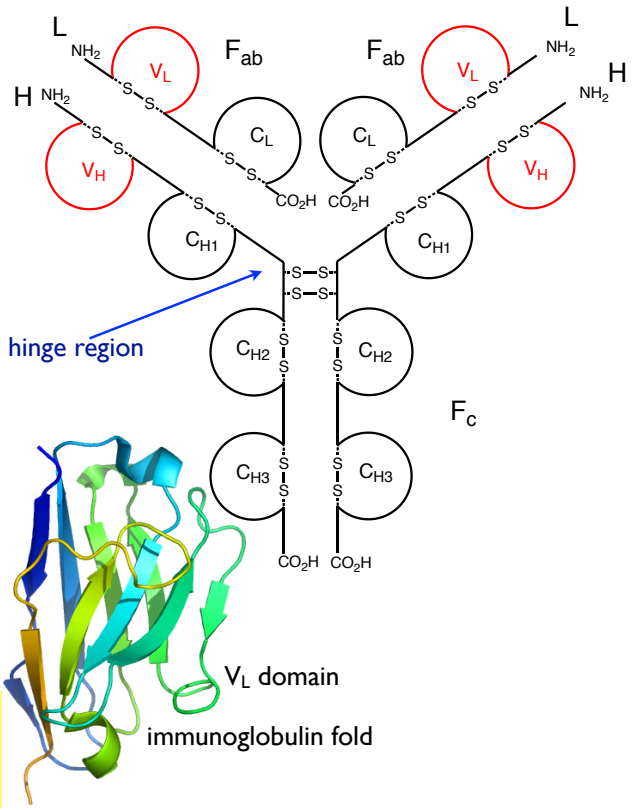
Immunoglobulins

- Individually, antibodies recognize and bind a few antigens with great specificity.
- Collectively, antibodies are capable of recognizing and binding a very diverse spectrum of antigens (virtually any molecule).
- Antibodies share common structural and functional properties. Depending on the class of antibody, ligand binding:
  - leads to elimination of antibody-antigen complex from the bloodstream.
  - leads to complement-induced lysis of cells.
  - triggers histamine release.
  - stimulates secretion of additional antibodies by lymphocytes.
- All immunoglobulins of a given class demonstrate these common qualities linked to different antigen specificities.



Immunoglobulins

- The characteristic structure for an intact immunoglobulin consists of a Y-shaped molecule assembled from four polypeptide chains (two H and two L chains)
  - Each L chain contains two domains (V_L and C_L).
  - Each H chain contains 4 domains (V_H, C_{H1}, C_{H2} and C_{H3}).
  - These domains are approximately 100 residues and share significant sequence homology. Thus, it is not surprising that they share a common folded conformation (immunoglobulin fold).
  - The immunoglobulin domains interact with each other (i.e. the C_{H2} and C_{H3} of one chain interact with their counterparts in the other, similar interactions between C_{H1} and C_L and V_L and V_H).
  - Peptide segments linking the domains demonstrate varying degrees of conformational freedom and are susceptible to enzymatic proteolysis (particularly the hinge region).
  - Cleavage at the hinge region results in release of the two arms (F_{ab} fragments) and base (F_c fragment). [Fv fragments can also be generated containing the V_L and V_H domains].
- This Y-shaped molecule is one of many shapes immunoglobulins can adopt in solution.

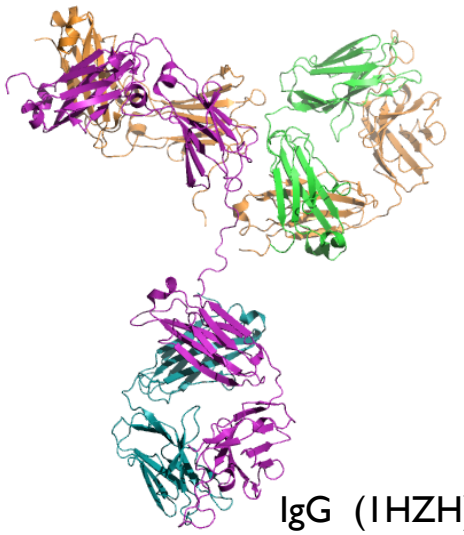


V_L domain

immunoglobulin fold

Antibody Isotypes

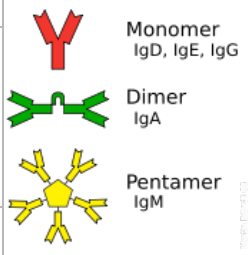
- Five antibody isotypes found in mammals (IgA, IgD, IgE, IgG and IgM).
- Differ in their roles, (i.e. functional location and antigen specificity).
- Reflect different states of B-cell maturity.



IgG (1HZH)

Antibody isotypes of mammals

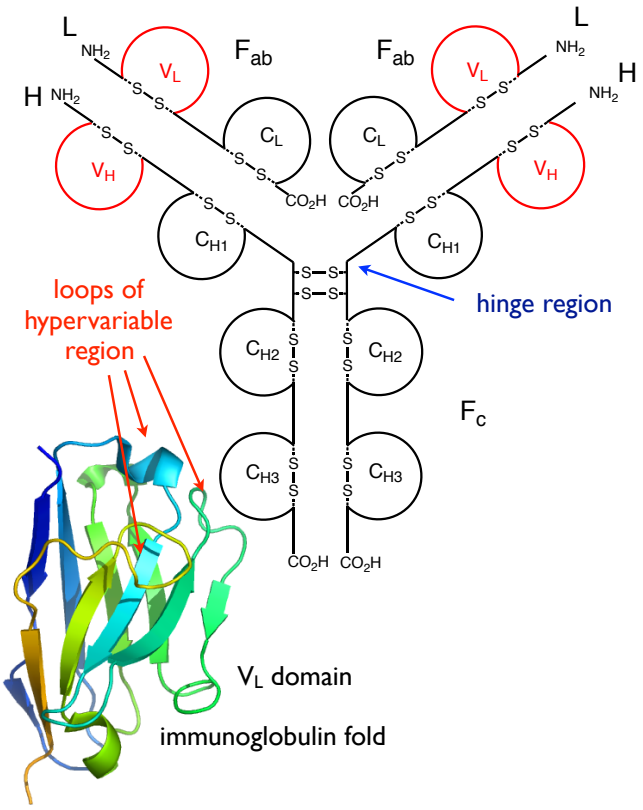
Name	Types	Description	Antibody Complexes
IgA	2	Found in mucosal areas, such as the gut, respiratory tract and urogenital tract, and prevents colonization by pathogens.[8] Also found in saliva, tears, and breast milk.	
IgD	1	Functions mainly as an antigen receptor on B cells.[9] Its function is less defined than other isotypes.	
IgE	1	Binds to allergens and triggers histamine release from mast cells, and is involved in allergy. Also protects against parasitic worms.[7]	
IgG	4	In its four forms, provides the majority of antibody-based immunity against invading pathogens.[7]	
IgM	1	Expressed on the surface of B cells and in a secreted form with very high avidity. Eliminates pathogens in the early stages of B cell mediated immunity before there is sufficient IgG.[7][9]	



from: http://en.wikipedia.org/wiki/Antibody

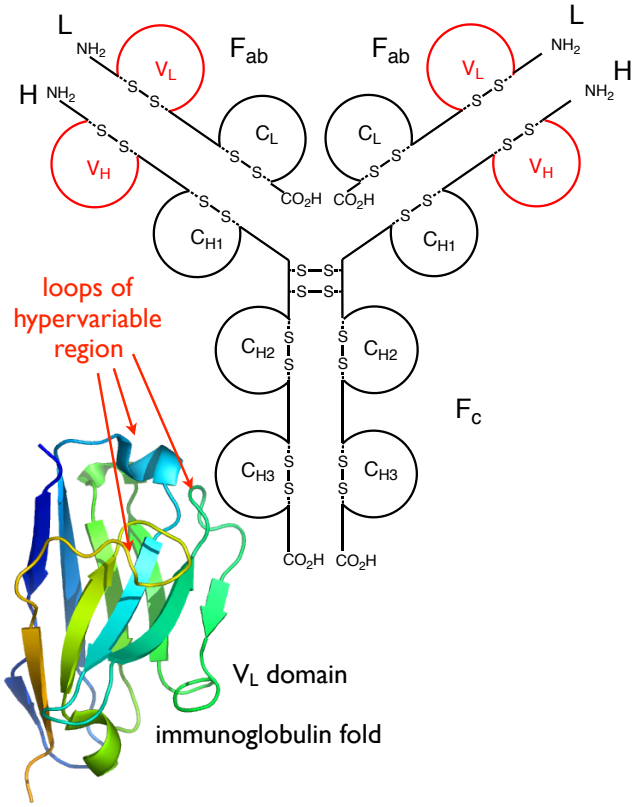
Antigen Recognition

- The antigen binding sites are located at the tips of the F_{ab} arms, situated between the V_L and V_H domains.
- Binding site formed by residues in the loops between β-strands of both of the V_L and V_H domains.
- Known as the complimentarity determining regions (CDR's) or hypervariable regions.
- The positioning and conformations of the loops in the CDR are the result of the scaffold provided by the immunoglobulin fold.
- Antibody diversity achieved through a special process during antibody biosynthesis.
- Variable domains are encoded by separate gene segments. Different antibody molecules are generated by joining together various combinations of these gene elements.
- Each antibody would demonstrate its own unique binding specificities.



Antigen Recognition

- Each antibody producing cell (B cell) produces and displays a single antibody molecule (species).
- When an antigen is encountered that is bound by a displayed antibody (IgM), it serves to stimulate production of soluble antibodies (i.e. IgG) and replication of the cell producing it.
- Individual antibody-producing cells can be isolated and cloned.
- Antibodies are not generally produced that act against molecules normally present in the host.
- Effector functions appear to be triggered primarily by the formation of large antigen-immunoglobulin aggregates.
- C1q is the first protein of the classical complement activation pathway. Appears to bind the C_{H2} domains of immunoglobulins.



Antibody Versatility

- *Antibodies provide powerful tools for molecular biology. If no ligands are known to exist for a given protein (polysaccharide or small molecule), then they can be made to order by generating antibodies against it.*
- All antibodies against proteins are at least somewhat specific to protein conformation.
- Antibodies recognizing the folded protein may be directed against any portion of the exposed surface (unless host has native protein with similar surface).
- Antibodies against proteins in the unfolded state can be generated using proteins that have been irreversibly unfolded.
- Antibodies can be generated that bind transition state analogs. When provided with the starting materials, these antibodies can in turn be used to catalyze the chemical reaction.

Cooperatively and Binding of Dimers to DNA

Sequence-specific Binding to DNA is a Problem

- Specific binding site recognition is the main challenge faced by proteins that specifically bind DNA.
- Following initial association of the protein with DNA, the protein then moves along the DNA. Such a strategy would aid these proteins in finding very specific sequences. (becomes 1 dimensional diffusion.

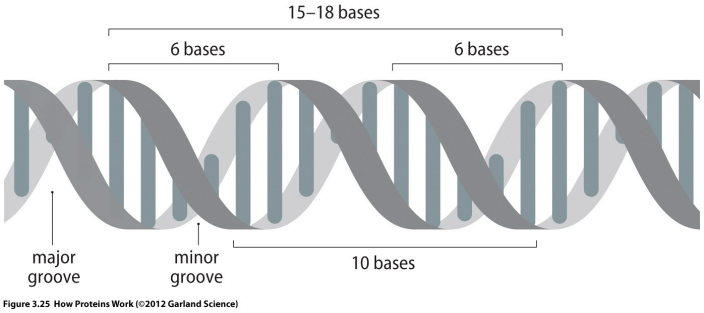


Figure 3.25 How Proteins Work (©2012 Garland Science)

- Protein needs to nonspecifically bind DNA strongly enough for it not to dissociate often, but not so tight so as to prevent it from moving rapidly along DNA. (primarily electrostatic interactions)
- Once the protein finds its specific site, it must bind tight enough to remain affixed at site for a extended period of time and not dissociate or continue moving along DNA (requires a slow off-rate).
- DNA-binding proteins can typically move from one base to the next at a rate of 10^6 bp/sec and are bound to DNA with an affinity of K_d = ~1-2 mM.
- Site specific binding has an effective K_d on the order of ~1 pM.

Sequence-specific Binding to DNA is a Problem

- Only the edges of the nucleotide bases are accessible to solvent and the protein (primarily in major groove).
- Specificity of protein-DNA interaction can be enhanced by increasing the length of the recognition interface.

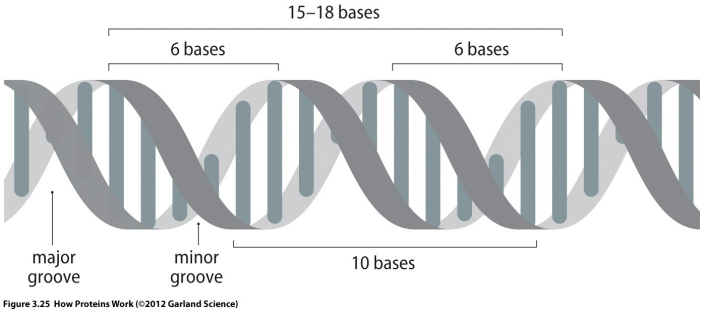


Figure 3.25 How Proteins Work (©2012 Garland Science)

- In order to recognize a single site in a bacterial genome a protein needs to be abel to recognize at least 12 bases.
- Most proteins recognize DNA sequences based on interactions in major groove, which has a periodicity of 10 bp in B-DNA.
- To bind to 12 bp, protein must be able to wrap itself around the DNA for more than one turn.

Assume DNA sequence of *n* bp occurs with probability of 4^{-n}

In a genome of 10^7 bp:

Specific 3-mer: p = 0.0156 or 1.6×10^5 times.

Specific 4-mer: p = 0.0039 or 4×10^4 times.

Specific 6-mer occurs 2500 times.

Specific 10-mer occurs 10 times.

Specific 12-mer occurs <1once.

Sequence-specific Binding to DNA is a Problem

- Most DNA-binding proteins recognize DNA from single direction.
- Avoids the need to wrap around DNA, but creates new problem - how to make a protein that spans a longer DNA sequence.
- In this approach, the two recognized major groove segments (6 bp each = total of 12 bp recognized) are separated by a span of DNA.

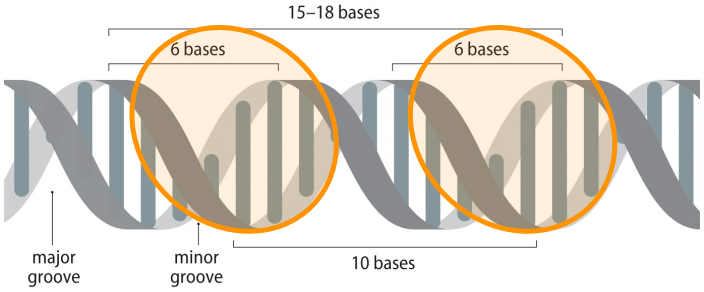


Figure 3.25 How Proteins Work (©2012 Garland Science)

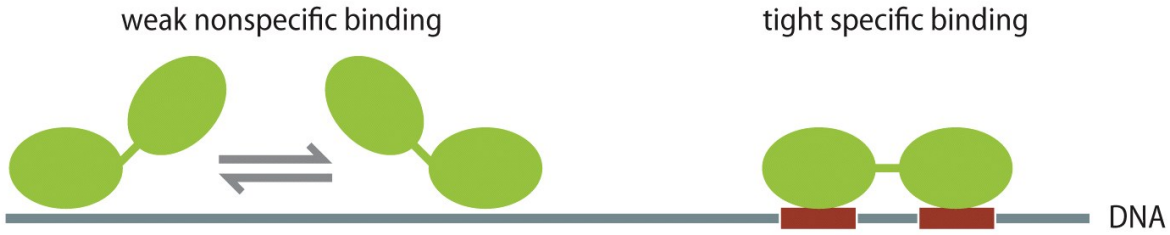


Figure 3.26 How Proteins Work (©2012 Garland Science)

Sequence-specific Binding to DNA is a Problem

- Dimeric protein strategy solves two problems:
  - How to cover an extended sequence with one protein.
  - How to maximize sequence search speed.
- Flexible interface capable of forming through a series of small steps... Zippering onto the DNA.
- Dimeric protein provides for both rapid scanning of DNA and tight specific binding.

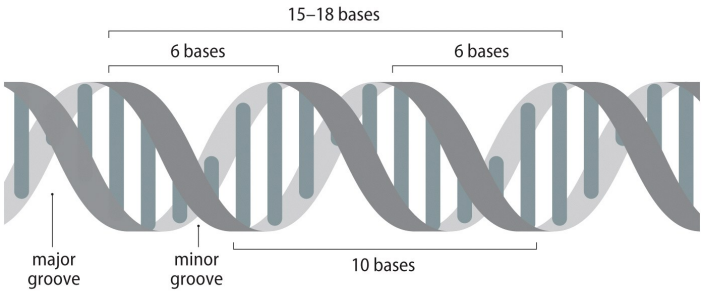


Figure 3.25 How Proteins Work (©2012 Garland Science)

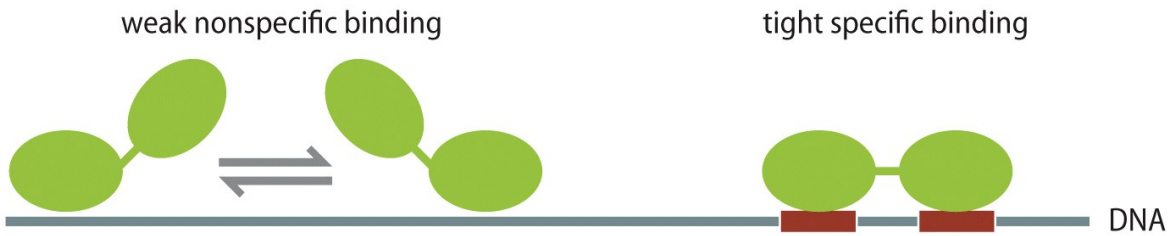


Figure 3.26 How Proteins Work (©2012 Garland Science)

Sequence-specific Binding to DNA is a Problem

- Extra subunits can be added in order to further enhance specificity if needed.
- Homeodomains are transcription factors that regulate development, with many having similar sequences.
- Individually, the yeast transcription factors Mcm1 and Mat **a**1 bind DNA very weakly.
- Strength and specificity of binding is achieved by their binding as a heterodimer (with different geometries) with the homeodomain Mat α2.
- Mcm1 and Mat **a**1bind DNA tightly with specific DNA sequences as dimers with Mat α2 - geometries of the heterodynes are very different.
- Homodimers allow for rotational symmetry and binding of palindromic DNA sequences.

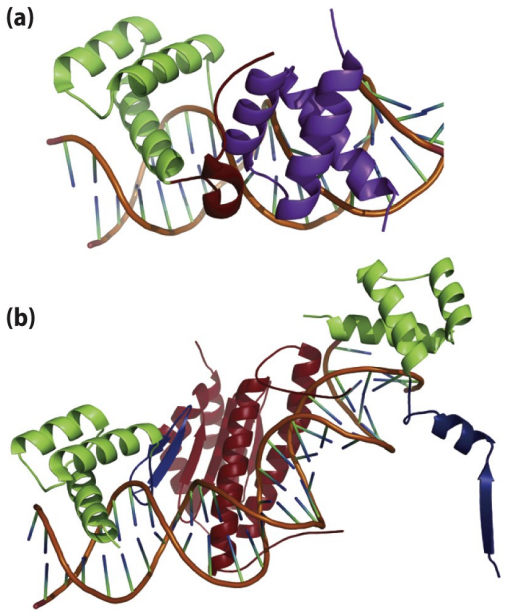
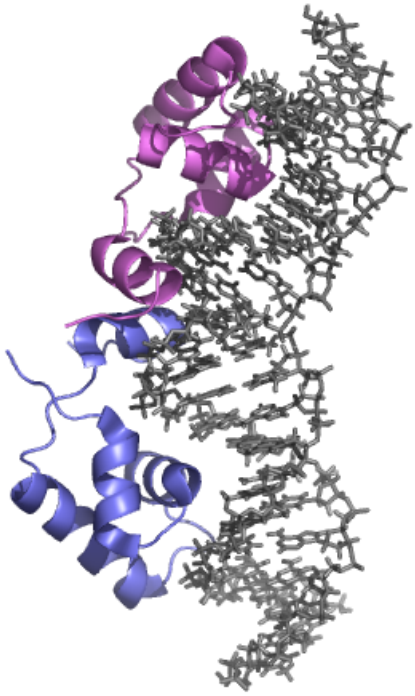


Figure 3.27 How Proteins Work (©2012 Garland Science)

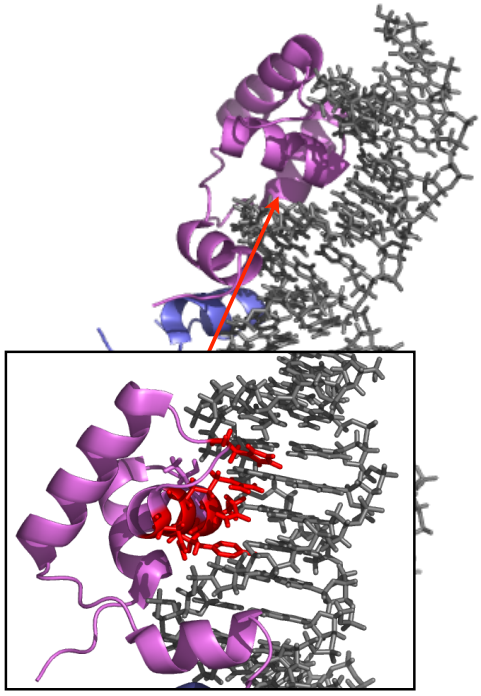
DNA-Binding Proteins

- Proteins that specifically bind DNA are of great biological importance: replication, gene regulation, gene expression.... etc.
- Only the edges of the nucleotide bases are accessible to solvent and the protein (primarily in major groove).
- In order for a protein to discriminate among base-pairs by interacting with their edges in the major groove, it must display interacting groups that project substantially from the protein surface.
- The helix-turn-helix is a common motif used by proteins for interaction with DNA.



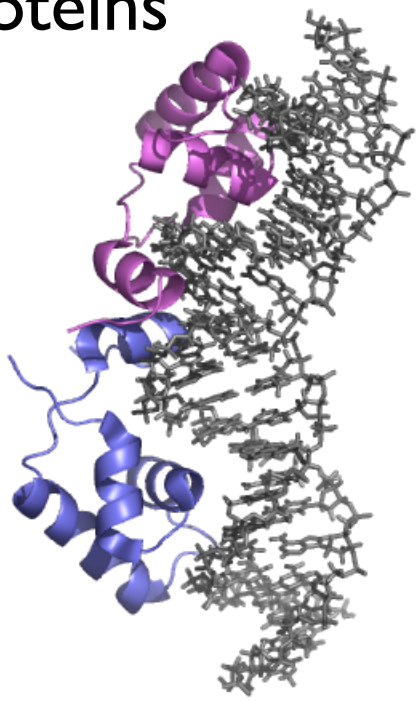
DNA-Binding Proteins

- The helix-turn-helix is a common motif used by proteins for interaction with DNA.
- The specificities of the various helix-turn-helix motifs arises from the different amino acid side chains displayed near the amino terminus of the helix projected into the groove (known as recognition helix).
- These side chain groups participate in hydrogen bonds, electrostatic and van der Waals interactions with the edges of the nucleotides exposed in the major groove.
  - The residues involved with interactions with DNA generally have polar side chains (Arg, Asn, Gln, Asp and Glu).
  - Water molecules are frequently recruited into hydrogen bond networks.
- The other helix crosses the major groove and participates primarily in nonspecific interactions.
- No simple code relating amino acid sequence to the recognized nucleotide sequence.



DNA-Binding Proteins

- Many helix-turn-helix DNA-binding proteins bind as dimers, with both equivalent binding helices making the same interactions with DNA with the same sequence.
- Such proteins tend to recognize palindromic DNA sites. In palindromic sites, the nucleotide sequence of one DNA strand is repeated in a complementary fashion and in reverse order.
- Palindromic sites define areas of localized twofold symmetry. The dimeric protein binds with its twofold axis of symmetry coinciding with that of the DNA binding site.
- Such a strategy should improve binding specificity.



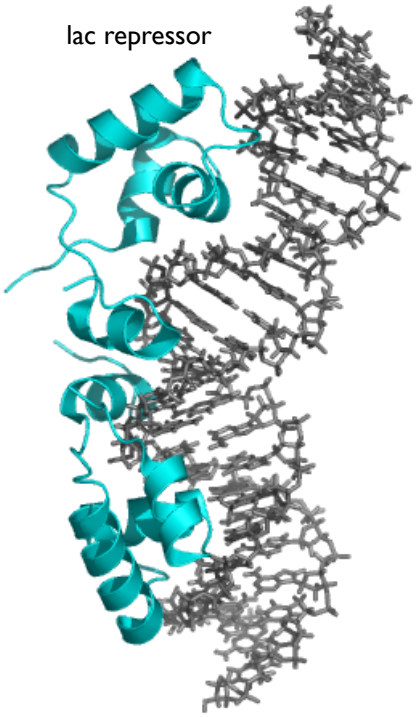
5' TGTGTGGAATTGTX$_9$ACAATTTCACACA
3' ACACACCTTAACA Y$_9$TGTTAAAGTGTGT

⇦ Recognized by *lac* repressor

Interaction between Proteins and DNA

- DNA-binding motifs other than helix-turn-helices and zinc fingers are known.
- While specific protein-base interactions are a major factor, another factor is deformation of the double helix of DNA.
- Binding of DNA by proteins is sometimes associated with distortion of the DNA double helix structure.
- There is evidence that DNA can be structurally pliable, but the degree of plasticity can be sequence dependent.
- Some proteins are thought to favor specific sequences in a distorted conformation that is favorable for that sequence but not for others.

lac repressor



Embo J. 2002, vol 21, 2866

Interaction between Proteins and DNA

- Proteins that bind to DNA irrespective of nucleotide sequence primarily recognize the DNA backbone (phosphate and sugar groups).
- Binding driven by electrostatic interactions and release of counter-ions and solvent molecules.
- Even in the case of sequence-specific proteins, the initial binding of protein to DNA is often nonspecific and is driven by electrostatic interactions and the release of loosely-bound cations. (protein loosely associated with DNA)
- Most DNA binding proteins have their affinity for DNA mediated by another ligand.

Interactions with Small Ligands

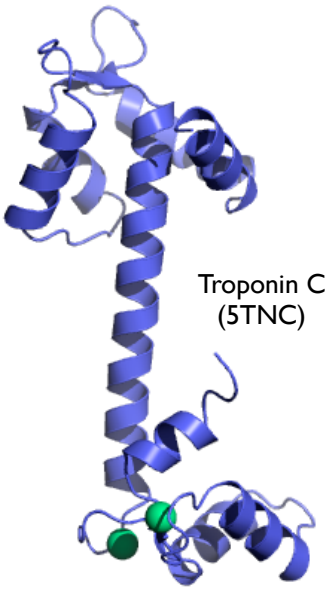
Very Small Ligands

- Proteins are capable of binding very small ligands (i.e. an ion or an electron) with great affinity and specificity.
- Ion binding sites tend to be located in the protein interior, and involve the simultaneous interaction of a number of functional groups.
- The nature of the chemical groups involved tends to contribute to specificity.
  - Ca^{2+} ions tend to be bound to oxygen atoms.
  - Zn^{2+} and Cu ions prefer sulfur atoms and imidazole nitrogen atoms (His).
  - Fe^{2+} and Fe^{3+} bind to sulfur atoms of Cys or disulfide.
  - Mg^{2+} ions are bound along with the phosphate groups of ligands.
- Geometry and positioning of the interacting groups also contributes to specificity and binding.

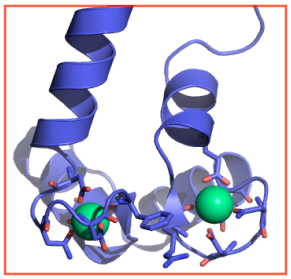
Very Small Ligands

Ca^{2+}-Binding

- Serves in a range of roles and functions in biological systems and binds to a variety of proteins.
- In calmodulin and troponin C, calcium binding site (EF Hand motif) consists of two helices (E and F) that are connected by a loop and lie roughly perpendicular to each other.
- Ca^{2+}-binding proteins often contain 2 or 4 EF hand motifs, with each one binding a Ca^{2+} ion. And two Ca^{2+} coordination sites lying near each other in neighboring EF hand motifs.
- Ca^{2+} coordination by Asp, Asn, Glu, Thr, and Ser side chains and backbone carbonyls (intermediary molecules also used).
- Coordination generally involves positions 1, 3, 5, 7, 9 and 12 in loop.
- Other types of Ca^{2+} binding sites present in other proteins.



Troponin C
(5TNC)



Very Small Ligands

O_2-Binding

- The selective binding of some small ligands require proteins to utilize prosthetic groups to aid in binding (most common example is binding of O_2 and electrons.)
- Oxygen atoms can be bound by ferrous heme groups (i.e. globin proteins), by two Fe^{2+} ions held by His, Glu and Asp and bridged by an oxide anion (hemerythrin), or by two Cu^{2+} ions held in close proximity (hemocyanins).
- In free solution, Fe^{2+}, Cu^{2+} and Fe^{2+} heme groups are oxidized by oxygen. Protein likely serves in part to prevent this oxidation.
- Protein environment determines which ligands bind and the consequences of ligand binding.
- Heme groups are not only used in O_2 binding (as in hemoglobin and myoglobin). Used by other proteins to accept electrons and catalyze redox reactions.

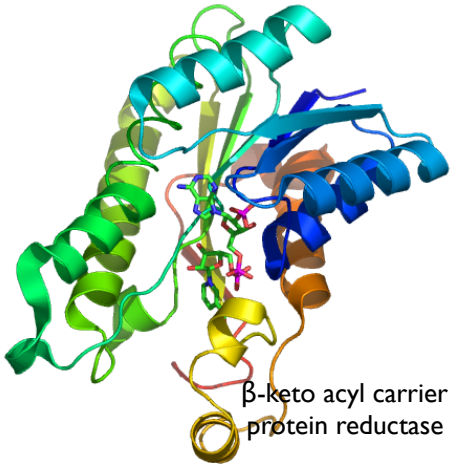
Very Small Ligands

Electron and Redox Proteins

- A variety of proteins bind electrons reversibly. Such proteins play critical roles in photosynthesis and respiration.
- Free electrons are not generally present in solution. They are normally transfered to and from other molecules.
- Most redox proteins utilize prosthetic groups:
  - Flavins or NAD.
  - Metal complexes of Fe^{2+}/Fe^{3+} and Cu^{+}/Cu^{2+}.
  - In some cases Fe ion is in a heme group or part of an iron-sulfur complex.
- The heme group of cytochromes bind and transfer electrons rather than O_2.
- Tightness of binding dependent on nature of the prosthetic group and modulation of its properties by the protein.
- The redox potentials of heme-containing proteins vary enormously (-1100mV to +300mV).

Nucleotide Binding

- Many proteins that bind dinucleotide coenzymes and single nucleotides do so in very similar ways.
- The binding site is composed of loops at the carboxyl ends of two parallel β-strands of a βαβ unit (mononulceotide binding domain or *Rossmann fold*).
- The positive end of the helix dipole appears to interact with the anionic phosphate groups.
- In some cases other βα units are present, and may participate in binding. (helping to orient ligand)
- Evolutionary origin of motif unclear, but sequence similarities appear to be for structural and functional reasons.



β-keto acyl carrier protein reductase

Struct. Fold. Des. **2000**, Vol.8, pp 339-347

Dinucleotide Coenzymes:
Nicotinamide adenine dinucleotide (NAD and NADP).
Flavin adenine dinucleotide (FAD).

Single Nucleotides:
Adenosine mono-, di- or triphosphate (AMP, ADP or ATP).
Flavin mononucleotide (FMN).

Very Small Ligands

Electron and Redox Proteins

- The Cu^{+} ion prefers a tetrahedral geometry, while the Cu^{2+} ion prefers a planar or octahedral ligand geometry.
- The orientations of coordinating side chains in a protein can alter the redox potential of the bound metal ion.
- Binding or release of an electron would be expected to produce a tendency to alter the protein conformation.
- Therefore, the effects of ligand binding on the protein and the effect of the protein on ligand affinity are linked functions.
- The observed perturbations in the intrinsic affinity of the prosthetic groups suggest that the surrounding protein scaffold provides a suitably rigid conformation.
- The actual mechanism of electron transfer remains a subject of interest.

Allostery: Interactions between Binding Sites

- Binding of multiple ligands at multiple sites in a protein can result in very complex behavior.
  - The sites may be identical and have the same intrinsic binding properties.
  - There may be different sites on the protein each with a different affinity.
  - Binding at one site may not affect the binding at any other sites.
  - Binding at one site my increase or decrease the affinity of the other site(s).
    - Ligands may be identical: homotropic interactions.
    - May involve different ligands: heterotropic interactions.
- In most cases where there is interaction between ligands, the ligands involved are not related and their binding sites are not located near each other in the protein structure.
- The term allostery is used to address such interaction between distant binding sites.
- Usually involves multimeric or multidomain proteins.

Interactions between Multiple Binding Sites

When multiple ligands are involved the interaction between them and their impact on binding can be complicated.

Independent Sites

Identical and Independent

- Multiple identical sites usually occur in proteins that are symmetrical oligomers, with each subunit incorporating the same binding site.
- Binding curves for such proteins resemble that of a protein with a single binding site, but with maximum stoichiometry reflecting number of binding sites in oligomer.
- Binding curves often interpreted in terms of association and disassociation constants for each binding event.

Association Constant of First Ligand

$$K_a^1 = \frac{nk_a}{k_d} = nK_a$$

Association Constant of n^{th} Ligand

$$K_a^n = \frac{k_a}{nk_d} = \frac{1}{n}K_a$$

Nonidentical and Independent

- Sites with the highest affinity become occupied at the lowest ligand concentrations followed by sites with lower affinity.
- Unless site-specific binding is measured, total binding curves can be very complex.

Cooperative Binding of Same Ligand at Multiple Sites

- Ligand binding at one site on a protein can increase or decrease the affinities of the remaining site(s) on the protein.
- Interactions between sites are said to be **homotropic** when the binding sites are intrinsically identical.
- **Positive cooperativity** results in a sigmoidal relationship between degree of binding and concentration of free ligand (because affinity for each successive ligand increases).
- Hill coefficient gives a quantitative measure of cooperativity in binding.
- Hill coefficient is given by the maximum slope of the curve of a Hill plot. May have any value depending on cooperativity (not necessarily and integer).
  - Hill coefficient could be as large as the number of binding sites involved, but only if there were complete cooperativity.
  - Partially-occupied protein molecules are always present even if there is positive cooperativity, but at lower concentrations that expected for independent binding.
- **Negative cooperativity** results in Hill plots with slopes of less than unity. Affinity decreases for each successive ligand molecule to be bound.

$$P + nA \rightleftharpoons P{\bullet}A_n$$

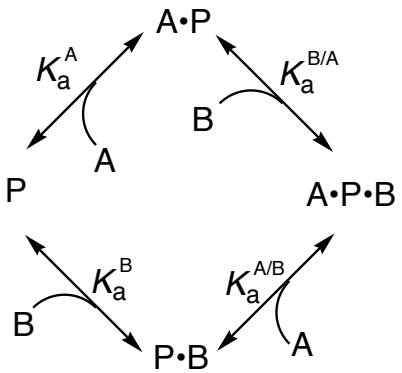
$$K_a = \frac{[P{\bullet}A_n]}{[P][A]^n}$$

$$\frac{y}{1 - y} = \frac{[P{\bullet}A_n]}{[P]} = K_a[A]^n$$

$$\log\left(\frac{y}{1 - y}\right) = \log K_a + n \log A$$

Interaction between Different Ligands

- The binding of one ligand can affect the binding of different ligands at different sites on the same protein (**heterotropic** effect).
- Easier to analyze than the homotropic effects.
- Interactions between the binding sites are linked functions:
- Example:
  - Binding of two ligands (A and B) to different sites on protein P.
  - Free energy change around cycle is zero.
  - The effect binding of A has on binding of B must be exactly the same as the effect of binding of B would have on binding of A.



$$\frac{K_a^A}{K_a^{A/B}} = \frac{K_a^B}{K_a^{B/A}}$$

Allosteric Models

- The interaction between sites on a protein is linked to protein flexibility.
- Two models have been proposed to address allosteric interactions.
- **The sequential model**: The protein is sufficiently flexible that the binding of one ligand at one site can directly alter the conformation at another site.
- **The concerted model**: Ligand binding at one site has no direct affect on other sites, but alters conformational equilibrium between two alternative quaternary structures.