# 6   Engineering tighter binding

Summary. Tighter binding of proteins like antibodies to their target can increase the therapeutic effectiveness of biopharmaceuticals and lower the detection limits of diagnostic tests. Engineering a tighter-binding protein involves increasing the complementarity between the protein binding site and target. Both the shapes of the interacting surfaces and non-covalent interactions between them should match.

Key learning goals

- The relative stabilities of the bound state, where the protein and ligand interact with each other, and unbound state, where the protein and ligand interact with water, determine the strength of binding. Non-covalent interactions between the binding site and ligand favor the bound state, while solvation of the binding site and ligand favors the unbound state.
- Equilibrium dissociation constants, K_d, which have units of molarity, reveal the binding strength. K_d corresponds to the target concentration where half of the protein is bound to a ligand. Lower dissociation constants indicate tighter binding. Equilibrium dialysis is one technique to measure equilibrium dissociation constants.
- Increasing the affinity of a protein for a ligand requires increasing the shape and interaction complementarity between the two partners. Close contact between the surfaces creates non-covalent interactions including van der Waals interactions, hydrophobic interactions, electrostatic interactions, and hydrogen bonds. The high cost of desolvating polar groups often makes burying them at the binding-site-ligand interface unfavorable even when they make favorable interactions like hydrogen bonds.
- Binding selectivity is the ability of a protein, often an antibody or receptor, to distinguish between two ligands. Binding selectivity depends on the ratio of the equilibrium association constants for the two binding possibilities.

6.1 Introduction

Tighter binding of a protein to its ligand is a common goal of protein engineering.[1] A ligand is any species (usually a small organic molecule, an ion, a protein or an oligonucleotide) that binds selectively, stoichiometrically and reversibly to a protein. Increased binding of a diagnostic antibody to its antigen can lower detection limits (e.g., a pregnancy test could detect pregnancy earlier), while increased binding of a therapeutic antibody can decrease drug dosage or increase drug efficacy. Traditional methods to improve the binding of antibodies rely on the natural processes that generate antibodies. Mice are immunized with the antigen followed by screening for antibodies with higher affinity. These in vivo methods, called affinity maturation, reach a limit where affinity stops increasing due to biological limits of the innate immune system. A common protein engineering goal is improving the affinity of an antibody from good binding to great binding. Figure 6.1 defines the different parts of an antibody.
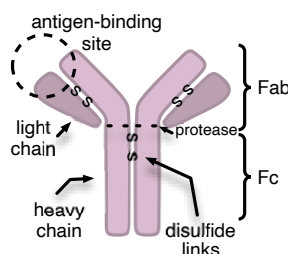
Figure 6.1. Antibodies consist of four protein chains: two identical heavy chains (~ 440 aa each) and two identical light chains (~ 220 aa each). Disulfide bonds connect the two heavy chains to each other and the light chains to their neighboring heavy chain. Antibodies contain two antigen-binding sites, each composed of the tips of the light and heavy chains. One of the two antigen binding sites is circled. Treatment with proteases cleaves both heavy chains as marked into three fragments: two Fab fragments (Fragment antigen-binding) and one Fc portion (Fragment crystallizable). The two protein chains that make up the Fab and Fc fragment remain linked to each other via disulfide cross links.

The strength of binding a ligand to a protein depends on the equilibrium between the bound and unbound states, Figure 6.2. In the bound state the protein and ligand interact with each other, while in the unbound state they interact with solvent. Stronger non-covalent interactions with each other as compared to solvent favors binding.

$$\text{site·ligand} \; \underset{k_{on}}{\overset{k_{off}}{\rightleftharpoons}} \; \text{site} \; + \; \text{ligand}$$
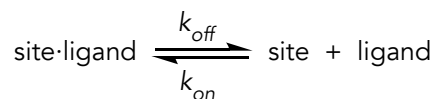
Figure 6.2. The bound state (ligand bound to the binding site) equilibrates with the unbound state (solvated ligand and site).

Two features influence the interactions between protein and ligand. First, the size and

shape of the binding site in the protein must match the size and shape of the ligand. This match creates strong van der Waals interactions between them. Second, the chemical properties of regions within the binding site match those of ligand. Hydrophobic regions match and electrostatic interactions in polar regions are favorable. X-ray structures of antibody-antigen complexes reveal that 15-20 amino acids from the complementarity determining region of the antibody contact the antigen. Most of the contacts are hydrophobic and both the antibody and antigen adjust their conformations slightly upon complex formation to maximize the strength of the interactions. In the perfect case, no adjustments would be necessary since the most stable conformations of both protein and antibody match each another exactly.

The two comparison states that determine ligand-protein binding strength are the unbound state where water solvates the independently diffusing ligand and protein and the bound state where the ligand and protein interact with each other, Figure 6.3. The solvation of polar and charged atoms at the surface stabilize the unbound state. In addition, the ability of the ligand and protein to diffuse independently in solution favors the unbound state. In contrast, the ordering of water at exposed hydrophobic regions of the protein and ligand destabilizes the unbound state as do unfavorable electrostatic interactions in high net charge regions.
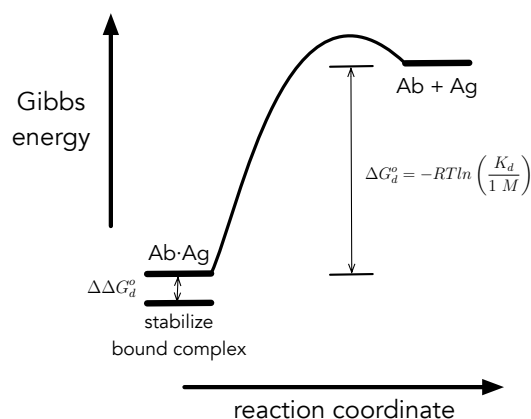
Figure 6.3. Gibbs energy changes associated with binding an antigen (Ag) to an antibody (Ab). The complex, Ab·Ag, is lower in energy than the free species, Ab + Ag, indicating that binding is favorable. Stabilizing the Ab·Ag complex increases the binding strength by an amount $\Delta\Delta G_d^o$.

In contrast, the bound state is stabilized by van der Waals attraction between touching atoms of the ligand and protein due to shape complementarity, the release of ordered water from hydrophobic regions of ligand and protein, electrostatic interactions between oppositely-charged atoms, and hydrogen bond formation between protein and ligand. The bound state is destabilized by desolvation of buried polar and charged atoms, electrostatic interactions between like-charged atoms, and the loss of independent diffusion for ligand.

144

Increasing binding requires increasing the energy difference between the bound and unbound states by stabilizing the bound state, destabilizing the unbound state or both. For example, one could destabilize the unbound state by adding exposed hydrophobic surface or high net charge regions within ligand-binding region. Ligand binding can alleviate this destabilization by binding a complementary hydrophobic surface or opposite charges. This alleviation of the destabilization when the ligand binds is critical to the success of this approach. If the destabilization persists in the bound state, then both states will have been destabilized and the binding strength will remain unchanged. Thus, the destabilizing changes must influence the ligand-binding region so that binding can alleviate the destabilization. One can stabilize the bound state relative to the unbound state by increasing the shape and charge complementarity between the protein and ligand stabilizes the bound state, while having little effect on the unbound state.

In all cases one must consider the effect of change to both the bound and unbound states. For example, engineering a hydrogen bond between the ligand and protein comes at the cost of reduced solvation of the polar atoms involved in the hydrogen bond. If the hydrogen bond is solvent accessible, then it remains solvated in the bound state and is often stabilizing. In contrast, buried hydrogen bonds are usually destabilizing because the cost of desolvating the polar atoms exceeds the gain from the hydrogen bond.

6.2 Measuring binding

Defining binding strength. The strength of binding is measured by the equilibrium constant for the reverse reaction, the ligand dissociation equilibrium constant, K_d, Figure 6.2 and eq. 6.1. For tighter binding, the dissociation equilibrium constant should be more unfavorable, so lower values of K_d indicate tighter binding. K_d has the units of molarity.

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[site] \cdot [ligand]}{[site \cdot ligand]} \tag{6.1}$$

The value of K_d corresponds to the free ligand concentration when half of the sites contain a bound ligand and the other half are empty, eq 6.2. Starting from eq. 6.1, when concentration of empty sites, [site], equals the concentration of filled sites, [site·ligand], these two terms cancel in eq. 6.1 yielding eq. 6.2.

$$K_d = [ligand] \text{ when } [site] = [site \cdot ligand] \tag{6.2}$$

Antibody-antigen complexes are common cases of protein-ligand binding. For example, the mouse monoclonal antibody called 26-10 binds digoxin with high affinity. Digoxin, a hydrophobic steroid with an attached carbohydrate, is used to treat congestive heart failure. Anti-digoxin antibodies have been used to measure serum levels of digoxin and also to neutralize digoxin in cases of an overdose. Since antibody 26-10 contains two independently-acting binding sites for digoxin, the concentration of sites is twice the concentration of antibody. Antibody 26-10 binds digoxin very tightly ($K_d = 10^{-10} M$)

and is half-saturated with digoxin when the concentration of free digoxin is $10^{-10} M$ (only 100 picomolar).

To convert the equilibrium constant in eq. 6.1 into a Gibbs energy, one introduces the notion of a reference state. The purpose of the reference state is to cancel out the molarity unit on the equilibrium constant since one can only apply the logarithm function to dimensionless quantities. The standard reference state is 1 M and the value of K_d is divided by this reference yielding a dimensionless quantity, eq. 6.3.

$$\Delta G_d^o = -RT \ln(K_d/1M) \qquad (6.3)$$

The values of K_d for stable protein-ligand binding interactions are always less than 1 M, so the ratio $K_d/1M$ is always less than one. The logarithm of a value less than one is negative, so the Gibbs energy for dissociation, ΔG_d^o, eq. 6.3, is positive indicating that dissociation is unfavorable. The 'o' superscript on the ΔG_d^o indicates that the K_d is compared to a reference state of 1 M. The free energy of dissociation of the digoxin-antibody 26-10 complex, ΔG_d^o, is +14.2 kcal/mol indicating that dissociation is highly unfavorable.

The Gibbs energy diagram shows that dissociation of an antibody-antigen complex is unfavorable, Figure 6.3 above. To improve binding one seeks to make dissociation even more unfavorable. This approach is analogous to stabilizing a protein by making the unfavorable unfolding reaction even more unfavorable in the previous chapter.

Measuring K_d. One method to measure the strength of binding between a protein like an antibody, and a ligand (or antigen) is equilibrium dialysis. The antigen equilibrates between two compartments separated by a dialysis membrane, Figure 6.4. The small pore diameter of the dialysis membrane allows the antigen to pass through, but not the antibody. At equilibrium, both compartments contain equal amounts of free antigen, but the antibody-containing compartment contains additional antigen as an antibody-antigen complex.

Measuring the equilibrium dissociation constant involves a series of equilibration experiments with different antigen concentrations. At lower concentrations of Ag, less Ab·Ag complex will form. The molar amount of antigen is always in excess of antibody. Since the equilibrium dissociation constants are very low, these experiments may require detecting very low antigen concentrations. For example, measuring the affinity of antibody 26-10 for digoxin requires measuring digoxin concentrations in the range of 100 picomolar.

Each equilibration experiment yields a ratio, Y, of sites complexed with ligand compared to the total number of sites. The concentration of sites complexed with ligand is the difference in the ligand concentrations between the antibody-containing and the antibody-free compartments, eq 6.4. The total concentration of sites is known from the amount of antibody added to the experiment. For the example in Figure 6.4 above, four of the binding sites are occupied by ligand and the total number of sites is eight, so Y = 4/8 = 0.5.
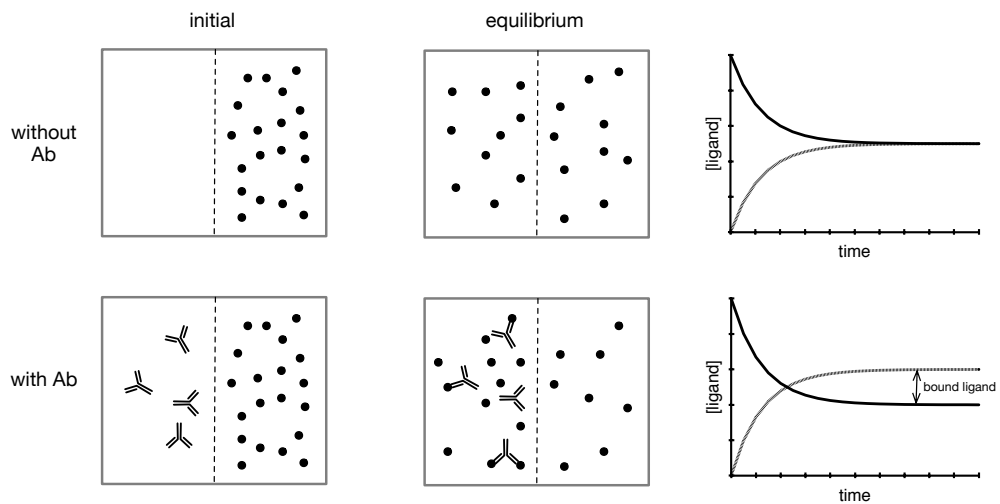
Figure 6.4. Measuring ligand (antigen) binding to an antibody by equilibrium dialysis. A dialysis membrane (dotted line) separates two compartments. Top row shows a control experiment without antibody. The ligand (twenty black circles) that is initially added to the right compartment distributes equally between both compartments (ten in each) because it passes freely through the membrane. Bottom row shows an experiment with antibody (Y-shapes) in left compartment. The dialysis membrane retains the antibody in its compartment. At equilibrium, both compartments contain the same concentration of free ligand (eight in each), but the antibody-containing compartment additionally contains antibody-bound ligand (four bound to the Y-shapes). The difference in the number of ligand molecules in the two compartments is the number of ligand molecules bound to antibody.

$$Y = \frac{\text{sites complexed with ligand}}{\text{total sites}} = \frac{[site \cdot ligand]}{[site] + [site \cdot ligand]} \tag{6.4}$$

Plotting the measured fraction saturation, Y at different concentrations of ligand as a function of free ligand on the x-axis yields a saturation binding curve, Figure 6.5. The sites become saturated with ligand as the concentration of free ligand increases.
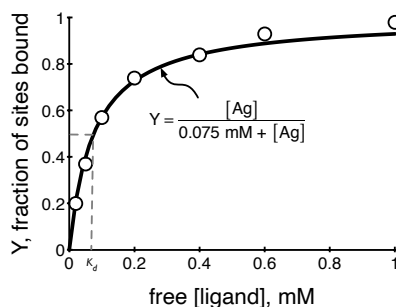
Figure 6.5. Fraction of antibody, Ab, bound to antigen, Ag, increases with increasing concentration of Ag until all of the antibody is saturated. Fitting equation 6.5 to this data yields the equilibrium dissociation constant, $K_d = 0.075$ mM. Inspection of the data table reveals that a bound antigen fraction of 0.5 occurs between 0.05 mM and 0.10 mM of free antigen, so the value of K_d lies within this range.

Equation 6.5 defines the relationship between Y, the free ligand concentration and the equilibrium dissociation constant, K_d. At half-saturation, Y = 0.5, the free ligand concentration equals the equilibrium dissociation constant, 0.075 mM for this example.

$$Y = \frac{[ligand]}{K_d + [ligand]} \tag{6.5}$$

This equation is derived by rearranging eq 6.1 to define the concentration of free sites, eq 6.6, and substituting it into eq 6.4 above.

$$[site] = \frac{K_d \cdot [site \cdot ligand]}{[ligand]} \tag{6.6}$$

Finding K_d from the data in Figure 6.5 involves adjusting the value of K_d so that the eq 6.5 best fits the experimental data.*

Another way to measure the binding of a ligand to a protein is fluorescence polarization.[2] This method requires a fluorescent ligand, but does not require the two-compartment apparatus needed for equilibrium dialysis. The solution of

---

*Non-linear least squares fitting of eq. 6.5 to the data starts with an initial guess for the dissociation constrant followed by iteration to find the best value. The supporting information includes a Python function for this fit.

ligand plus protein is irradiated with plane-polarized light to excite the fluorescent ligand. The emitted light (fluorescence) from the ligand is measured both parallel and perpendicular to the plane of the irradiated light. If the ligand does not turn during the few nanoseconds between the absorbtion and emission of light, then most of the light will be emitted parallel to the plane of the irradiated light. However, it the ligand turns during the few nanoseconds between the absorption and emission of light, then less of the light will be emitted parallel to the plane of the irradiating light. A ligand free in solution tumbles rapidly, but a ligand bound to a macromolecule tumbles more slowly. Fluoresence polarization detects whether a ligand is bound or not by measuring how much of the emitted light is parallel to the plane of irradiated light.

One can extend this method to also measure the binding of non-fluorescent ligands by adding the fluorescent ligand and a competing non-fluorescent ligand simultaneously. The non-fluorescent ligand competes with the fluorescent ligand for the protein binding site. Measuring the decrease in binding of the fluorescent ligand a varying concentrations of the non-fluorescent ligand reveals the relative strength of the two ligands. Since the binding strength of the fluorescent ligand is known, the binding strength of the non-fluorescent ligand can be calculated.

Practical details on binding measurements are available.[3] Other ways to measure binding include surface plasmon resonance and isothermal titration calorimetry.

Increasing binding strength requires increasing the difference in energy between the bound and unbound states. Engineering the antibody for tighter binding lowers the Gibbs energy of the Ag·Ab complex, raises the Gibbs energy of free Ag and Ab or both. The value of ΔG_d^o measures improvement of binding, Figure 6.3. This value is the difference in the Gibbs energy change for the two dissociation reactions, eq 6.7.

$$
\begin{aligned}
\Delta \Delta G_d^o &= \Delta G_{d-new}^o - \Delta G_{d-original}^o \\
&= -RT \ln \left(\frac{K_{d-new}}{K_{d-original}} \right) = -RT \ln(improvement\,factor)
\end{aligned}
\tag{6.7}
$$

Tighter binding corresponds to a larger positive Gibbs energy of dissociation for the improved variant and thus a positive value of ΔG_d^o. The value of K_{d-new} is a smaller (dissociation less favorable) than $K_{d-original}$, so $\left(\frac{K_{d-new}}{K_{d-original}} \right)$ is a value less than 1. For example, equilibrium dissociation constant may decrease from $10^{-6} M$ in the original to $10^{-8} M$ in the engineered case. The improvement factor is 10^{-2} and $\ln \left(\frac{K_{d-new}}{K_{d-original}} \right)$ is -4.61. Then $-RT \ln \left(\frac{K_{d-new}}{K_{d-original}} \right)$ is $+2.73$ kcal/mol at 298 °K.

6.3 Engineering tighter binding

Engineering tighter binding involves modification of the antibody binding site to increase the complementarity between the antibody and antigen, Table 6.1. The first requirement of a binding site is that its size and shape must match the size and shape of

Table 6.1. Strategies to increase non-covalent interactions between protein and ligand.

Strategy	Reasoning
size and shape matching	- creates favorable van der Waals interactions - minimizes bumping - releases bound water - allows direct interactions
preorganization	- binding conformation favored even before binding - no rearrangement needed for binding
hydrophobic matching	- increases hydrophobic effect
optimize H bonds	- compensates for loss of H bonds to solvent - avoids buried H bonds due to poor solvation
optimize electrostatics	- minimizes interactions with like charges - maximizes interactions with unlike charges

the target ligand. This matching excludes solvent water by placing ligand atoms near binding-site atoms to create the opportunity for favorable interactions. The second criteria is that the binding site be preorganized for binding, that is, that the lowest energy conformation be the one that binds the ligand. The remaining criteria refer to optimizing different types of non-covalent interactions: hydrophobic interactions, hydrogen bonds and electrostatic interactions.

Size and shape matching. Matching the size and shape of the protein binding site to the target ligand maximizes their contact surface in order to maximize favorable non-covalent interactions, Figure 6.6. Size matching ensures that all of the ligand interacts with the binding site. If the cavity is too large or too small, then part of the ligand remains exposed to solvent and the interaction between ligand and binding site will be less than the maximum possible. Shape matching refers to the surface contours of the ligand and binding site. Surface complementarity means the surfaces touch to create favorable van der Waals interactions. The close contact also makes additional non-covalent interactions between the ligand and binding site possible.

The intricate shapes of molecules prevents most of the atoms in two contacting molecules to directly contact one another. The sphere representation of a molecule shows its van der Waals surface. The radius of each sphere is the van der Waals radius for the atom and the overlapping spheres show the van der Waals surface of the molecule, Fig 6.7. This surface is intricate and includes narrow crevices too small to fit a water molecule. A water molecule cannot touch the regions within these crevices and a larger molecule will be excluded from even more regions. The *solvent contact surface* is a hypothetical smoother surface that eliminates crevices that are inaccessible to water. The solvent contact surface is the surface mapped by the edge of sphere with a radius of 1.4 Å as it rolls on the van der Waals surface of a molecule. The radius of

a) size   b) shape
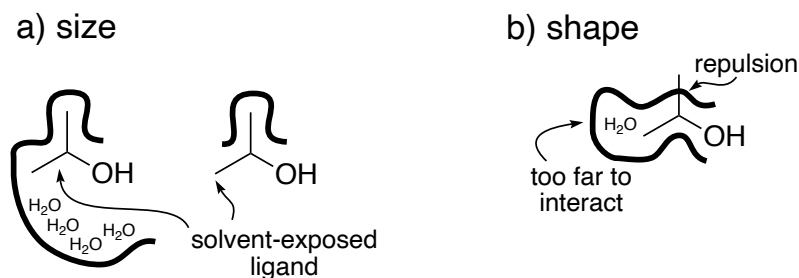
Figure 6.6. To maximize interactions with the target ligand (isopropyl side chain) the binding site must be the same size as the ligand and the contours of the binding site must match the shape of the ligand. a) A binding site that is larger or smaller than the ligand leaves part of the ligand exposed to solvent, which misses opportunities for interactions between ligand and binding site. b) Shape mismatches between the target ligand and contours of the binding site create repulsive interactions in regions where the binding site is too small and an inability to make favorable interactions in regions where the binding site is to large to contact the ligand.

1.4 Å corresponds to the van der Waals radius of a water molecule. This surface shows smooth areas where the water molecule rides over the van der Waals surface as well as regions of protruding atoms where the water can directly touch the molecule.[†]

The majority of the surface atoms, even for highly complementary molecules, will not contact one another. The intricacy of molecular structures make complete direct contact geometrically impossible. Good complementary shapes will exclude solvent water from the interface and make many direct contacts.

An example of high surface complementarity is the interaction between antibody fragment 26-10 complexed with digoxin,[4] Figure 6.8. The structure of 26-10 shows a deep pocket that surrounds the hydrophobic steroid portion. The hydrophilic carbohydrate groups of digoxin remain exposed to solvent and do not contribute to binding. The complementarity between the solvent-accessible surface of the aglycone (compound without the glycoside moiety) of digoxin and 26-10 is imperfect, even in regions that are in contact. The shape complementary is close enough that there are no buried cavities in the complex that are large enough to contain a water molecule. This exclusion of water ensures a strong contribution of the hydrophobic effect to the binding. There are some regions exposed to the bulk water that are large enough to contain water molecules. The conformations of antibody and steroid remain the same in both unbound and bound states indicating that the low energy unbound shapes match each other.

[†]Another surface, the solvent accessible surface, or Connolly surface, lies 1.4 Å outside the solvent contact surface. This surface represents the center, not the edge, of a water molecule on the surface. The solvent accessible surface is more convenient in some cases because positions on this surface will match the x,y,z coordinates of the rolling water molecule.

[‡]Charles Cullen, a nurse and serial killer convicted in 2003, used digitalis (and other drugs) to kill possibly forty hospital patients. This story and that of Amy Loughren, the ICU nurse who helped convict Cullen, is told in the book *The Good Nurse* (2013), and a 2022 movie by the same title currently on Netflix.
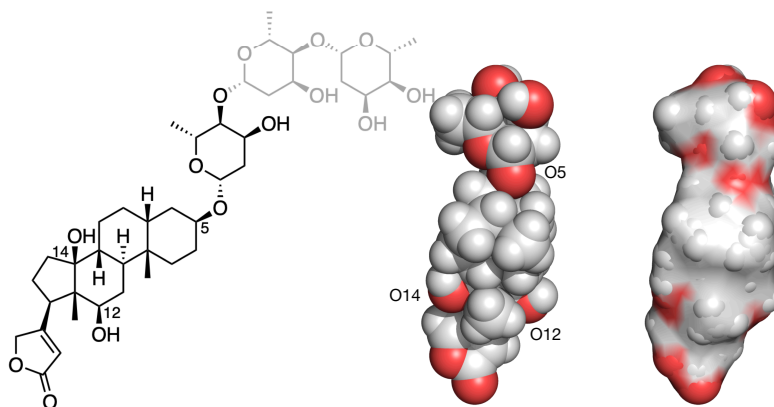
Figure 6.7. Structure of digoxin (digitalis), a cardiac glycoside.[‡] Left: The chemical structure of digoxin consists of a steroid (the aglycone) linked to a trisaccharide. The two sugar moieties in gray are not included in the two structures to the right. Middle: The overlapping van der Waals spheres of each atom create the van der Waals surface of digoxin. It contains numerous crevices that are too small to fit a water molecule. A water molecule would be slightly larger than one of the oxygen atoms shown as red spheres. Right: The solvent contact surface of digoxin is a hypothetical surface that shows the closest approach of a water molecule. Water can directly contact only a small fraction of the van der Waals surface.

Fragment 26-10 binds digoxin tightly, $K_d = 0.1$ nM. The binding is entirely due to the hydrophobic effect and van der Waals interactions. No hydrogen bonds or salt bridges form between antibody fragment 26-10 and digoxin. One consequence of the lack of interactions with the hydroxyl groups at C12 and C14 is a lack of selectivity for analogs without these hydroxyl groups. For example, 26-10 binds digitoxigenin (which lacks the hydroxyl group at C12) with equal affinity to digoxigenin.[§]

Sculpting a complementary surface will likely require multiple mutations and involve readjustment of the main chain and side chains. The smallest change that one can introduce by amino acid substitutions is the addition or removal of a methyl or hydroxyl group. This is a large change - about the size of a water molecule - and likely too large to precisely sculpt a surface. To make smaller changes one needs to make multiple substitutions including substitutions outside the surface being engineered so that a succession of adjusted positions make more subtle changes to the complementary surface. Precisely predicting the adjustments caused by multiple substitutions is more difficult making such sculpting difficult.

Substitutions often introduce readjustments of the main chain and side chains. A larger side chain causes readjustments when it bumps into nearby atoms. A smaller side chain may also cause readjustments because enlarging a hydrophobic pocket is unfavorable. Replacing a large buried hydrophobic side chain with a smaller one destabilizes the protein because the smaller hydrophobic surface 1) reduces the contribution of the hy-

---

[§]The -genin suffix on these two compounds indicates the aglycone only.
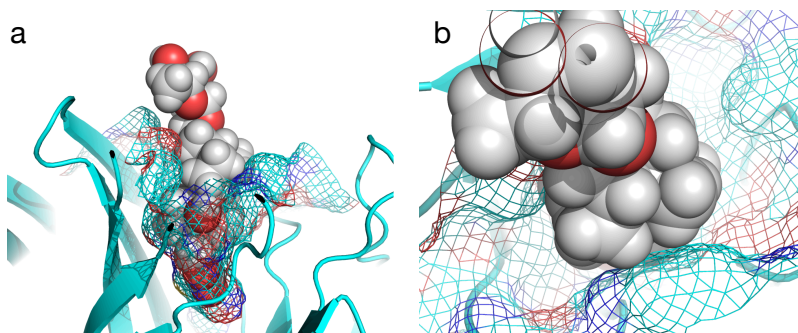
Figure 6.8. The binding pocket of antibody fragment 26-10 (mesh surface, backbone shown as a cartoon) binds the steroid digoxin (spheres) by creating a complementary surface. a) The two terminal saccharides extend into solvent and were not resolved in this structure. b) A top view shows close contact between the 26-10 and the aglycone in several parts, but much of the 26-10 surface does not directly contact the aglyclone. The figure was created using PyMOL from the x-ray crystal structure of the complex (pdb id = 1igj[4]). The surface of the antibody pocket was calculated using the web tool CASTp[5]

drophobic effect to protein folding and 2) reduces the favorable van der Waals contacts between the side chain and the rest of the protein. Structural readjustments often decrease the size of the created space to minimize these changes. The estimated decrease in stability due to removing one buried CH_2 group is 1.1 ± 0.5 kcal/mol.[6] For example, a Leu to Ala substitution removes three carbon atoms from the protein, so this substitution decreases protein stability by an estimated 3.3 ± 1.5 kcal/mol. If the protein structure readjusts, the destabilization will be at the smaller end of the estimate range, while if the structure does not readjust, it will be at the larger end of the estimated range. For example, a Leu to Ala mutation in T4 lysozyme destabilized the protein by 5.0 kcal/mol at position 99, but by only 2.7 kcal/mol at position 46. The x-ray structures revealed that at position 99 the structure did not readjust upon substitution leading to a larger cavity and a larger loss in van der Waals interactions, while at position 46 the structure relaxed leading to a smaller increase in the cavity size and a smaller loss in the van der Waals interactions.[7]

In a few cases, readjustments of the nearby residues reverse the intended change in size. For example, an ancestral hydroxynitrile lyase contains phenylalanines at positions 121 and 178 in the active site, while the corresponding modern enzyme from the rubber tree contains smaller leucines at these positions. One would expect that the substrate-binding site of the rubber tree enzyme would be larger, but the x-ray structures showed the opposite. In the rubber tree enzyme the side chain of Trp128 readjusts into the active site creating a smaller substrate binding site,[8] which is the opposite of what might be expected from comparison of the sizes of phenylalanine and leucine.

Preorganization. Preorganization of a dynamic protein means creating a global energy minimum conformation where the atoms in the apo protein are already oriented to in-

teract with the partner atoms in the ligand. In the ideal case, this conformation has the correct structure with a minimum of flexibility. It must have the correct structure to avoid the energy cost of moving the protein atoms from their most stable postions into the less stable interacting positions. This energy cost lowers the overall favorability of the binding. Flexibility of the apo protein also creates an energy cost. If protein flexibility decreases in the protein-ligand complex, then there is an entropy cost to binding, which also lowers the favorability of binding. Thus, the preorganized apo protein should be no more flexible than the protein-ligand structure.

Constraining a flexible ligand into the three-dimensional shape it adopts when bound to a receptor often strengthens the binding interaction. For example, the phosphotyrosine-containing pseudopeptide **6.1**, Figure 6.9, bound tightly to the SH2 domain of a growth receptor binding protein, $K_D = 2.2\ \mu M$. Introducing a cyclopropyl ring to create compound **6.2** restricted the flexibility of the ligand so that its structure matches the bound conformation. The binding interaction tightened to $K_D = 0.36\ \mu M$.[9] Similar improvements are expected by restricting the flexibility of the binding site.

6.1, $K_D = 2.2\ \mu M$ **6.2**, $K_D = 0.36\ \mu M$
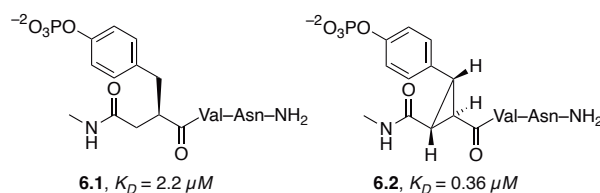
Figure 6.9. Phosphotyrosine containing pseudopeptides bind tightly to the SH2 domain of a growth receptor binding protein. The cyclopropane ring in compound **6.2** restricted its flexibility and increased its affinity as compared to compound **6.1**.

The preorganized protein structure for binding refers to both side chain and backbone orientations. The side chains must point in the correct direction to interact with the ligand, while the backbone must position these side chains at the correct distances. Incorrect backbone positions create binding sites that are too large or too small for the ligand. For example, a molecular dynamics simulation of digoxigenin-binding proteins found that favored conformation of the non-binding proteins was too large to interact with the entire digoxigenin surface due to incorrect backbone positions.[10]

Hydrophobic matching. Besides matching the shape of the ligand and binding site to maximize van der Waals interactions, matching additional non-covalent interactions further strengthens the interaction. For example, matching hydrophobic regions of the ligand and binding site avoids unsatisfied polar interactions. The structure of antibody fragment bound to methamphetamine showed a bound water molecule facing a hydrophobic region of the methamphetamine antigen.[11] Researchers hypothesized that releasing this water molecule would increase the hydrophobic character of the binding pocket and increase binding. Replacement of SerH93, which formed a hydrogen bond to the water molecule, with the slightly larger residue threonine increased the binding affinity 3.1-fold from $K_d = 0.79$ to 0.25 nM. A crystal structure of the improved variant

confirmed that this substitution forced out the water molecule.

Optimize hydrogen bonds. Hydrogen bonds between the binding site and ligand come at the cost losing hydrogen bonds to solvent and desolvation of the hydrogen bonding partners. Buried hydrogen bonds are rarely favorable, but solvent-exposed hydrogen maintain solvation so are more likely to be favorable. Hydrogen bonds often contribute to selectivity of binding.

Electrostatic interactions. Electrostatic interactions can contribute to the tighter binding of the protein and ligand in two ways. One way is to directly stabilize the protein-ligand complex; the other way is to speed up the formation of this complex, Figure 6.10.
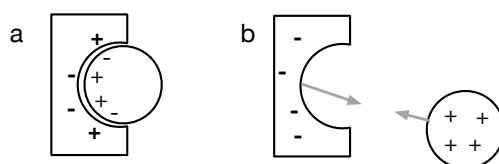
Figure 6.10. Two types of electrostatic interactions strengthen binding between partners. a) Partners with low or zero net charge associate with one another. Electrostatic interactions within the complex strengthen binding. b) Partners have opposite net charges. Favorable, long-range electrostatic interactions steer these oppositely charged molecules toward each other (gray arrows). Uncharged molecules of the same size would likely miss each other because their motions are random.

Direct stabilization of the protein-ligand complex refers to favorable interactions between oppositely charged atoms in the protein-ligand complex after it has formed. For example, improved electrostatic interactions increased the binding of monoclonal antibody hu3F8 (naxitamab) to the surface of cancer cells ~7-fold.[12] The antibody binds to two negatively charged sialic acid groups on glycolipid ganglioside GD2. An Asp32His substitution replaced a negatively charged amino acid residue with a positively charged one just outside the binding site. Computer modeling predicted ~0.5 kcal/mol improvement in electrostatic interaction with the sialic acid groups in good agreement with the measured increase in binding. In another example, computer modeling predicted five substitutions that could strengthen electrostatic interactions between epidermal growth factor receptor and an antibody targeted to it (cetuximab, Erbitux®). Experiments confirmed that three of the five increased binding affinity. Substitutions Ser26Asp and Thr31Glu replaced polar residues with negatively charged residue, while Asn93Ala replaces a polar residue with an uncharged residue. All substitutions were in the antibody light chain. The combination of all three substitutions increased the binding 10-fold by decreasing K_d decreased from 490 pM to 52 pM.[13] The researchers noted that solvent-exposed electrostatic interactions and hydrogen bonds were more stabilizing than buried ones because they remain solvated. The loss of solvation energy made most buried electrostatic interactions and hydrogen bonds unfavorable.

Electrostatic interactions can also strengthen binding by long-range interactions that occur before binding occurs. Electrostatic steering increases binding strength by

increasing the collision frequency between oppositely charged proteins and ligands, Figure 6.10b above. Binding requires the protein and ligand to bump into one another as they randomly move through the solution. Electrostatic forces alter the collision rates between molecules because they act over long distances, that is, beyond van der Waals contact distances (see Chapter 2). This longer range of electrostatic interactions effectively makes charged molecules larger when they interact with other charged molecules. Nearby oppositely charged molecules can move toward each other to make van der Waals contact, while uncharged molecules must meet within their van der Waals radii. This increase in the number of collisions between oppositely charged molecules increases the rate of association of protein and ligand and thereby strengthens the binding interactions. Increasing the electrostatic interaction between separated partners as compared to the complex by 1 kcal/mol increases the association rate by a factor of 2.8.[14] Charges that steer partners toward each other may also contribute to stronger binding after the complex forms. In contrast, partners with low net charge will not be steered even though they form favorable electrostatic interactions after binding.

To engineer such an increase, one increases the charge difference between the protein and ligand. For example, Selzer and coworkers[14] increased the affinity between β-lactamase (negative net charged) and its inhibitor (no net charge) 250-fold by giving the inhibitor a positive net charge. Four substitutions increased the positive charge by six units. A web tool predicts the increase in association rate upon mutagenesis: https://webhome.weizmann.ac.il/home/bcges/PARE.html.

6.4   Computational design of tighter binding

Docking models protein-ligand complexes. Designing improved binding interactions relies on computer modeling of the interactions. An accurate structure of the ligand complexed to the binding protein is the best starting point for computer modeling, but such structures are rarely available. Protein-ligand docking is a molecular modeling technique to create a model of a protein-ligand complex from a structure of the binding protein. SwissDock is a web tool for protein-ligand docking.[15,16] The most common application of docking is drug design, but it is also useful for protein design. Docking predicts where the ligand binds to the protein, the geometry of the bound ligand and the strength of the binding interaction.

As with all protein modeling, the two challenges are sampling and scoring. Sampling finds possible binding locations, orientations and conformations. Scoring ranks these possibilities to identify the best ones. If the sampling does not generate the best-binding structure, then scoring cannot identify it.

SwissDock sampling combines a broad search and a local search of the conformational space. The broad search temporarily reduces the van der Waals and electrostatic repulsions to cross energy barriers. The local search is semi-stochastic, dealing with rotations and translations. SwissDock also avoids revisiting unfavorable conformations to speed up the search.

For scoring SwissDock use the CHARMM force field, which is similar to AMBER and intended for modeling biological molecules. In addition, CHARMM includes accurate
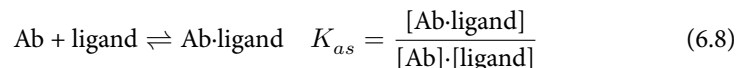
solvation models to help rank the different binding possibilities. The initial search uses a fast calculation that ignores solvation, later evaluation of the bound structures includes solvation to increase accuracy.

Computational design of binding proteins. Predicting substitutions that improve the affinity of a protein for a target ligand has shown some success, but remains challenging. Computational design using Rosetta predicted seventeen proteins that bind to the steroid portion of digoxin called digoxigenin.[17] Two of these proteins showed micromolar affinity for the target, which is relatively weak affinity, but the ability to predict such proteins is an impressive accomplishment. Further optimization did not use computation, but directed evolution methods (see Chapters 6 and 7) and yielded proteins with picomolar affinity for digoxigenin. The designed protein included hydrogen bond interactions with the hydroxyl groups of digoxigenin, so it showed selectivity against analogs lacking these hydroxyl groups. Other computational design were even less successful. Morin et al.[18] designed twelve possible binding sites in *endo*-1,4-β-xylanase for the antibiotic vancomycin, but none of the designs bound vancomycin. X-ray structure analysis of four designs showed the predicted conformation, but molecular dynamics showed that the designs were highly flexible. The authors suggest that the designed sites spent most of their time in conformations not suitable for binding.
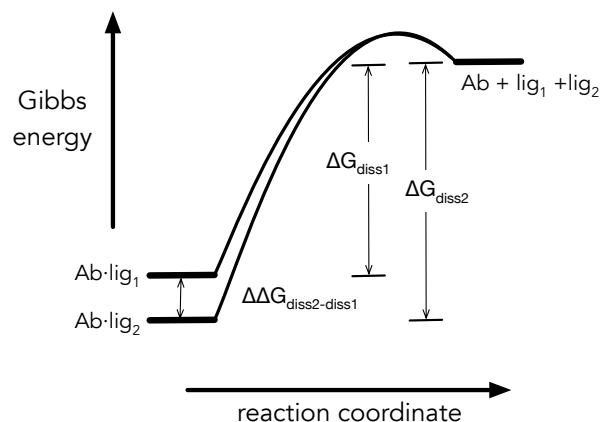
6.5   Selective binding

Define selectivity. Selectivity, S, is the ability to distinguish between alternatives. Binding selectivity is the ability of a protein to bind one ligand more tightly than another one, Figure 6.11. The selectivity value, S, is the ratio of the affinities of the protein for the preferred and non-preferred ligands. This value depends on comparison molecules, so they should always be included with the number. For many proteins, their ability of proteins to favor binding one ligand in the presence of competing ligands is what makes them valuable. For example, the cardiovascular antibody drug abciximab selectively binds to a receptor involved in platelet aggregation to inhibit formation of blood clots.

The affinity of a protein for a ligand is given by the equilibrium association constant, K_{as}, for the binding reaction, eq 6.8. It is the inverse of the dissociation constant defined in Figure 6.2 above.

$$\text{Ab} + \text{ligand} \rightleftharpoons \text{Ab·ligand} \quad K_{as} = \frac{[\text{Ab·ligand}]}{[\text{Ab}]\cdot[\text{ligand}]} \tag{6.8}$$

Binding selectivity, S_{bind}, is the ratio of the affinities for the two protein-ligand complexes being compared, eq. 6.9. By convention, the equilibrium association constant for the tighter-binding ligand, t, is in the numerator and that for the weaker-binding ligand, w, is in the denominator yielding binding selectivity values that are greater than one. The binding selectivity should also mention which ligands are being compared. A selectivity of ten corresponds to ten-fold tighter binding of one ligand as compared to the other ligand. A selectivity of one corresponds to no selectivity; that is, equal affinity for both.

Figure 6.11. Gibbs energy diagram of selective binding of ligand 2 (lig$_2$) over ligand 1 (lig$_1$) to an antibody (Ab)

$$S_{bind} = \frac{K_{as-t}}{K_{as-w}} = \frac{\frac{1}{K_{d-t}}}{\frac{1}{K_{d-w}}} = \frac{K_{d-w}}{K_{d-t}} \tag{6.9}$$

Binding selectivity can also be expressed in terms of the equilibrium constants for dissociation, which are the inverse of the equilibrium constants for association. The dissociation constant for the weaker-binding ligand is in the numerator while that for the tighter-binding ligand is in the denominator. The selectivity of an protein for two competing ligands can be measured by measuring each binding affinity separately, then dividing the two values.

The Gibbs energy change associated with binding selectivity, eq. 6.10, is the difference between the Gibbs energy changes of the two association reactions. The Gibbs energy change is negative since association of the tighter-binding ligand is more favorable.

$$\Delta\Delta G_{S_{bind}} = \Delta G_{as-t} - \Delta G_{as-w} - RT \ln\left(\frac{K_{as-t}}{K_{as-w}}\right) = -RT \ln\left(S_{bind}\right) \tag{6.10}$$

To engineer a protein for increased selectivity between a preferred and a non-preferred ligand, one can either decrease the affinity of the protein for the non-preferred ligand (negative selection) or increase the affinity of the protein for the preferred ligand (positive selection). The best approach will depend on the details of the application.

References

1. Banta, S., Dooley, K., & Shur, O. (2013). Replacing antibodies: Engineering new binding proteins. *Annu. Rev. Biomed. Eng.*, *15*(1), 93–113. https://doi.org/10.1146/annurev-bioeng-071812-152412

2. Rossi, A. M., & Taylor, C. W. (2011). Analysis of protein-ligand interactions by fluorescence polarization. *Nat Protoc*, *6*(3), 365–387. https://doi.org/10.1038/nprot.2011.305

3. Pollard, T. D. (2010). MBOC technical perspective: A guide to simple and informative binding assays. *Mol. Biol. Cell*, *21*(23), 4061–4067. https://doi.org/10.1091/mbc.E10-08-0683

4. Jeffrey, P. D., Strong, R. K., Sieker, L. C., Chang, C. Y. Y., Campbell, R. L., Petsko, G. A., Haber, E., Margolies, M. N., & Sheriff, S. (1993). 26-10 Fab-digoxin complex: Affinity and specificity due to surface complementarity. *Proc. Natl. Acad. Sci. U. S. A.*, *90*(21), 10310–10314. https://doi.org/10.1073/pnas.90.21.10310

5. Tian, W., Chen, C., Lei, X., Zhao, J., & Liang, J. (2018). CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.*, *46*(W1), W363–W367. https://doi.org/10.1093/nar/gky473

6. Pace, C. N., Scholtz, J. M., & Grimsley, G. R. (2014). Forces stabilizing proteins. *FEBS Lett.*, *588*(14), 2177–2184. https://doi.org/10.1016/j.febslet.2014.05.006

7. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science*, *255*(5041), 178–183. https://doi.org/10.1126/science.1553543

8. Jones, B. J., Evans, R. L., Mylrea, N. J., Chaudhury, D., Luo, C., Guan, B., Pierce, C. T., Gordon, W. R., Wilmot, C. M., & Kazlauskas, R. J. (2020). Larger active site in an ancestral hydroxynitrile lyase increases catalytically promiscuous esterase activity. *PLoS ONE*, *15*, e0235341. https://doi.org/10.1371/journal.pone.0235341

9. DeLorbe, J. E., Clements, J. H., Teresk, M. G., Benfield, A. P., Plake, H. R., Millspaugh, L. E., & Martin, S. F. (2009). Thermodynamic and structural effects of conformational constraints in protein–ligand interactions. Entropic paradoxy associated with ligand preorganization. *J. Am. Chem. Soc.*, *131*(46), 16758–16770. https://doi.org/10.1021/ja904698q

10. Barros, E. P., Schiffer, J. M., Vorobieva, A., Dou, J., Baker, D., & Amaro, R. E. (2019). Improving the efficiency of ligand-binding protein design with molecular dynamics simulations. *J. Chem. Theory Comput.*, *15*(10), 5703–5715. https://doi.org/10.1021/acs.jctc.9b00483

11. Thakkar, S., Nanaware-Kharade, N., Celikel, R., Peterson, E. C., & Varughese, K. I. (2014). Affinity improvement of a therapeutic antibody to methamphetamine and amphetamine through structure-based antibody engineering. *Sci. Rep.*, *4*, 3673. https://doi.org/10.1038/srep03673

12. Zhao, Q., Ahmed, M., Guo, H., Cheung, I. Y., & Cheung, N.-K. V. (2015). Alteration of electrostatic surface potential enhances affinity and tumor killing properties of anti-ganglioside GD2 monoclonal antibody hu3F8. *J. Biol. Chem.*, *290*(21), 13017–13027. https://doi.org/10.1074/jbc.M115.650903

13. Lippow, S. M., Wittrup, K. D., & Tidor, B. (2007). Computational design of antibody-affinity improvement beyond in vivo maturation. *Nat. Biotechnol.*, *25*(10), 1171–1176. https://doi.org/10.1038/nbt1336

14. Selzer, T., Albeck, S., & Schreiber, G. (2000). Rational design of faster associating and tighter binding protein complexes. *Nat. Struct. Biol.*, *7*(7), 537–541. https://doi.org/10.1038/76744

15. Grosdidier, A., Zoete, V., & Michielin, O. (2007). EADock: Docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins*, *67*(4), 1010–1025. https://doi.org/10.1002/prot.21367

16. Grosdidier, A., Zoete, V., & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.*, *39*, W270–W277. https://doi.org/10.1093/nar/gkr366

17. Tinberg, C. E., Khare, S. D., Dou, J., Doyle, L., Nelson, J. W., Schena, A., Jankowski, W., Kalodimos, C. G., Johnsson, K., Stoddard, B. L., & Baker, D. (2013). Computational design of ligand-binding proteins with high affinity and selectivity. *Nature*, *501*(7466), 212–216. https://doi.org/10.1038/nature12443

18. Morin, A., Kaufmann, K. W., Fortenberry, C., Harp, J. M., Mizoue, L. S., & Meiler, J. (2011). Computational design of an endo-1,4-beta-xylanase ligand binding site. *Protein Eng. Des. Sel.*, *24*(6), 503–516. https://doi.org/10.1093/protein/gzr006

Problems

1. The equilibrium dissociation constant, K_d, of an antibody and its ligand (a peptide) is 5 μM at pH 5.0 and 25 °C.

   a) At what concentration of the ligand is half of the protein bound?

   b) What fraction of the protein is bound at ligand concentration of 1.25 μM?

   c) When the pH was raised to 6.5, the K_d increased to 20 μM. Is the binding tighter or weaker at this pH compared to pH 5.0? Explain why.

   d) What functional groups/residues are most likely responsible for this change in the binding affinity with pH?

Supporting Information

Code Block S4.1. Python script to fit data for an equilibrium dialysis experiment (Fig 6.5)

```python
#! /usr/bin/env python
# coding: utf-8

# also run the command below in terminal
# sudo chmod +x file-name.py

import numpy as np              #import math functions
   to use arrays
from scipy import optimize      #import non-linear fit
   function
import matplotlib.pyplot as plt  #import plotting
   function

# enter your data here
```

160

```python
freeAg = np.array([0.02, 0.05, 0.10, 0.2, 0.4, 0.6, 1])
Y = np.array([0.20, 0.37, 0.57, 0.74, 0.84, 0.93, 0.98])
Kd = 0.1 # initial guess for Kd

# define equation to be fit
def satCurve(freeAg, Kd):
    return freeAg/(freeAg + Kd)

# fit the data to the equation using non-linear least
    squares with
# default fit settings
popt, pcov = optimize.curve_fit(satCurve, freeAg, Y, Kd)
print("Kd =", "{0:.3f}".format(popt[0]))

# plot data and best fit curve
plt.scatter(freeAg, Y)
xfit = np.linspace(0,1)
plt.plot(xfit, satCurve(xfit, popt[0]), 'r-')
plt.xlim(0, 1)
plt.xlabel('free antigen concentration')
plt.ylabel('fraction of antibody containing bound antigen
    ')
plt.show()
```