Lecture 10 - Exercises

Question 1: Biomolecular Interactions

Select a statement that is **TRUE**:

a) Protein-protein interactions are based on:
- Low desolvation energy contribution to binding
- Hydrophobic interactions created by the water molecules trapped in the interface
- Chemically and geometrically complementary molecular surface features
- Identical surface properties in the binding partners

b) Protein-DNA interactions:
- Always involve Watson-Crick base pairing between amino-acids and nucleotides
- Require an optimal distribution of charged and polar amino-acids on the protein side
- Require exclusively helical protein domains because DNA is also a helix
- Are entropically disfavored due to desolvation

c) Electrostatic (ionic) interactions:
- Are created between charged amino acids and phosphate groups in DNA backbone
- Contribute less energy in biomolecular interactions than a single hydrogen bond
- Do not depend on distance between charged groups
- Are formed between Lewis bases in nucleotides and positively charged amino acids

d) When it comes to biomolecule interactions between two binding partners:
- Enthalpy is the primary contributing factor to k_{on} rate while entropy affects k_{off}
- The displacement of surface-trapped water molecules is entropically favorable for binding
- Hydrophobic interactions are the main contributor to enthalpy but do not affect system entropy
- Dissociation constant does not depend on temperature.

e) Determining affinity between two binding partners:
- Also tells you about the type of interactions at their interface
- Allows to infer potential conformational changes that the binding event induces
- Is always based on measurements of thermodynamic properties
- Allows to estimate the fraction of bound and unbound material under different conditions

Question 2: Experimental methods for affinity measurements

Select a statement that is **FALSE**:

a) Nuclear Magnetic Resonance (NMR):
- Can be used to measure entropy and enthalpy contributions to total binding energy
- Allows to map the potential interacting surface between the binding partners
- Is used to measure affinity between binding partners based on changes in chemical shifts
- Is a spectroscopic method for measuring binding affinity

b) Isothermal Titration Calorimetry (ITC):
- Is based on spectroscopic measurements of heat changes upon addition of binding partner
- Allows to determine the thermodynamic parameters of binding
- Can be applied to measure affinity at different temperatures (T)
- Allows to experimentally measure the stoichiometry of binding (n)

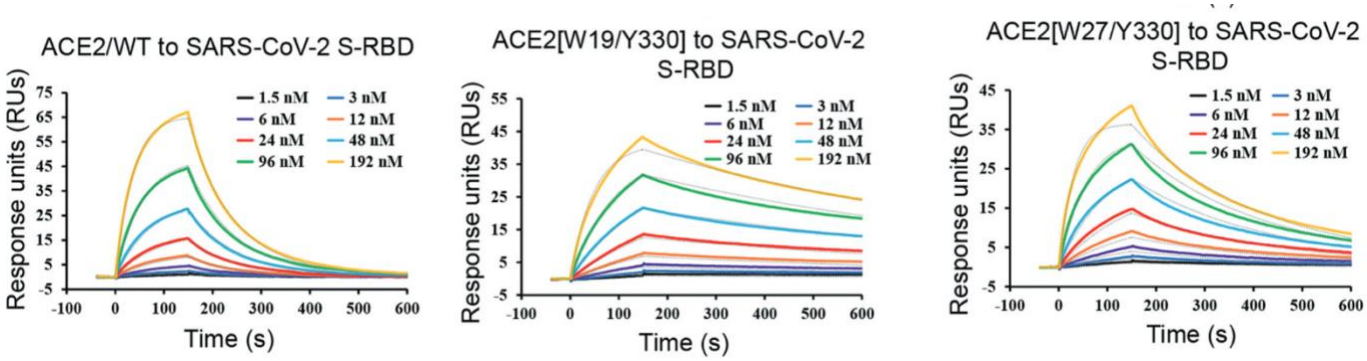
c) Fluorescent polarization (FP):
- Requires fluorescent labeling of one or both binding partners in most cases
- Is based on differing molecular weights of bound versus unbound molecules
- Can be used to measure if two or more ligands compete for the same binding site
- Allows to experimentally measure the stoichiometry of binding (n)

d) Surface plasmon resonance (SPR) is unique compared to other methods because:
- It allows to measure both, the association (K_a) and dissociation constants (K_d)
- It is based on spectroscopic measurement of light reflection at liquid-solid interface
- It can be used to measure kinetic parameters of binding (k_a and k_d rates)
- It requires surface immobilization of molecules for the method to work.

e) When it comes to different experimentally-determined thermodynamic parameters:
- Lower dissociation constants (K_d) indicates stronger binding between two molecules
- Greater reduction in Gibbs free energy indicates stronger affinity between two molecules
- Higher association rate (k_a) automatically indicates lower dissociation rate (k_d).
- For strong binding it is necessary to have high association (k_a) and low dissociation rates (k_d)

Question 3: Application to virus-receptor interactions

During the COVID pandemic scientists spent quite some effort to understand how SARS-CoV-2 virus infects cells. Upon discovery that the ACE2 protein serves as a cellular receptor to which the virus attaches, one research group tested if different mutations in ACE2 would have an effect on binding to the virus. The results for the unmutated ACE2 (WT) as well as the two mutants (W19/Y330 and W27/Y330) are shown below. For each mutant variant multiple concentrations of ACE2 were tested (from 1.5nM to 192nM), while the concentration of SARS-CoV-2 spike was kept constant at 200nM.



a) Can you identify what method was used to test the binding? Can you assign what different portions of the binding response curves refer to?

Answer:

Given that the axes are response units (RU) and time (s), these sensograms came from the surface plasmon resonance (SPR). The portion from 0-150 seconds corresponds to association of ACE2 and SARS-CoV-2, and is used to estimate association rate (k_a). The portion from 150-600s corresponds to dissociation of ACE2 from the SARS-CoV-2 and is used to estimate the dissociation rate (k_d).

b) Why is there an increase in maximum signal with increasing concentration of ACE2 in each of the 3 cases?

Answer:

Given that the amount of SARS-CoV-2 is kept constant and there is an increase in ACE2 concentration, the binding equilibrium is pushed towards generating more ACE2 + SARS-CoV-2 complexes on the surface of the SPR chip which produces higher signal. If the Kd is known, you can use the equation below to calculate the exact amount of each species in equilibrium:

Kd = [Unbound ACE2] * [Unbound SARS-CoV-2] / [ACE2 + SARS-CoV-2 complex]

c) By comparing the binding response curves corresponding to equivalent ACE2 concentration in 3 mutants (use the highest concentrations since they are the easiest), can you rank the mutants based on the association rate (k_a) from highest to lowest? Can you rank them based on dissociation rates (k_d)?

Answer:

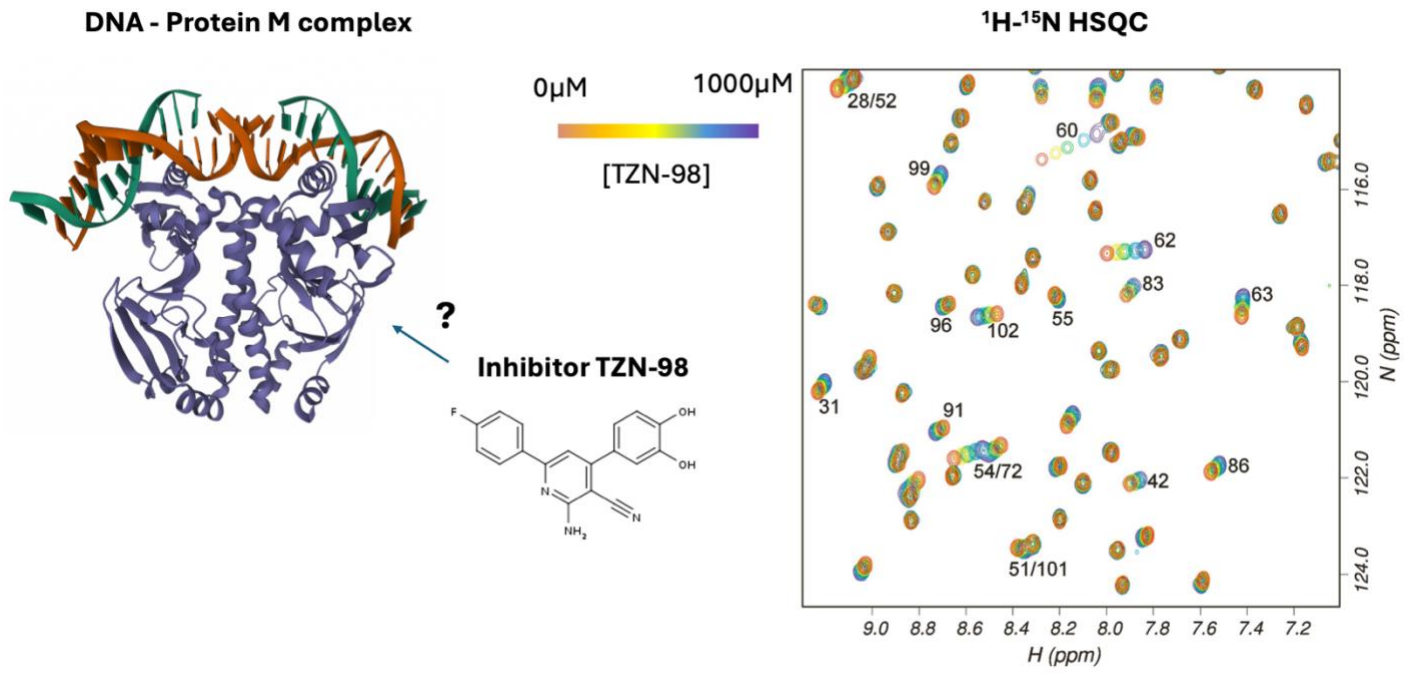
To address this question, you should focus only on the 192nM curve in different graphs and compare the increase or decrease in signal as a function of time. However, you can also use one of the lower concentration curves since it leads to the same result. In reality, the k_a and k_d values are calculated using all the different concentrations.

The association part of the curve is in the time interval 0-150 seconds, and we will use it to compare the k_a of different mutants. If you compare the 3 mutants, they all have slightly different rates at which the signal increases with time. ACE2/WT reaching the highest level (~65 RU at max) compared to ACE2[W19/Y330] (~45 RU at max) and ACE2[W27/Y330] (~35 RU at max). So, the WT will have the highest k_a, followed by [W19/Y330], followed by [W27/Y330].

The dissociation part of the curve is in the time interval 150-600 seconds, and can be used to compare the k_d values of different mutants. If you look at the rates at which the signal reduces with time over the entire period, the highest k_d is with the WT followed by [W27/Y330], and the slowest dissociation is observed for [W19/Y330].

Question 4: Inhibitor characterization

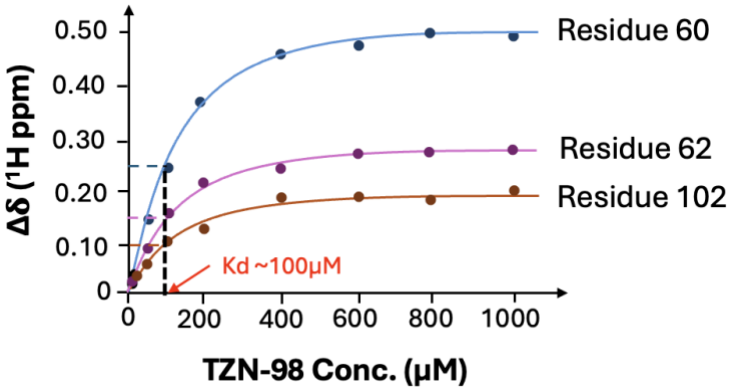
You screened a large library of compounds and discovered a new molecule (**Inhibitor TZN-98**) with potent antibiotic activity against diverse bacterial strains (even the pan-antibiotic resistant ones). Based on computational predictions, you suspect that the molecular target of this molecule is **Protein M**, which is a DNA binding protein essential for bacterial life cycle. Now you wish to confirm that Protein M and TZN-98 interact, and you perform ^{1}H-^{15}N HSQC characterization of Protein M (isotope labelled with ^{15}N) in the presence of increasing concentrations of TZN-98 (ranging from 0-1000μM concentration). The overlay of ^{1}H-^{15}N HSQC spectra obtained at different protein/inhibitor ratios are shown below in different color. The corresponding residue IDs are indicated next to different peaks.



a) Based on the overlay of HSQC spectra under different conditions do you think that there is interaction between TZN-98 and Protein M? How can you tell?
Answer: Some peaks (resonances) shift in a manner that is dependent on the concentration of TZN-98 added to the solution which indicates that a portion of Protein M is interacting with the compound.

b) Why did some peaks move in the spectra in response to TZN-98 addition and the others did not?
Answer: Each peak in ^{1}H-^{15}N HSQC spectra corresponds to a single pair of ^{1}H and ^{15}N atoms bound by a covalent bond, such as the N-H groups found in peptide bonds (= single amino acid). The peaks (resonances) that shifted correspond to amino acids that are directly impacted by the compound binding. Most commonly this is due to direct interaction, but it could also be due to long-range conformational changes (less common).

c) How can you use the data above to calculate the Kd of this interaction? What would you plot?
Answer: For Kd calculation you would take individual amino-acids that moved in the HSQC spectra and plot the change in chemical shift (Y axis) as a function of TZN-98 concentration (X axis). By fitting a single curve for each single peak that moved you can estimate the Kd value as 50% of the max chemical shift change. This fitting can be done using each peak's shifts and the final Kd value calculated as an average of all individual Kd values estimated for each peak.



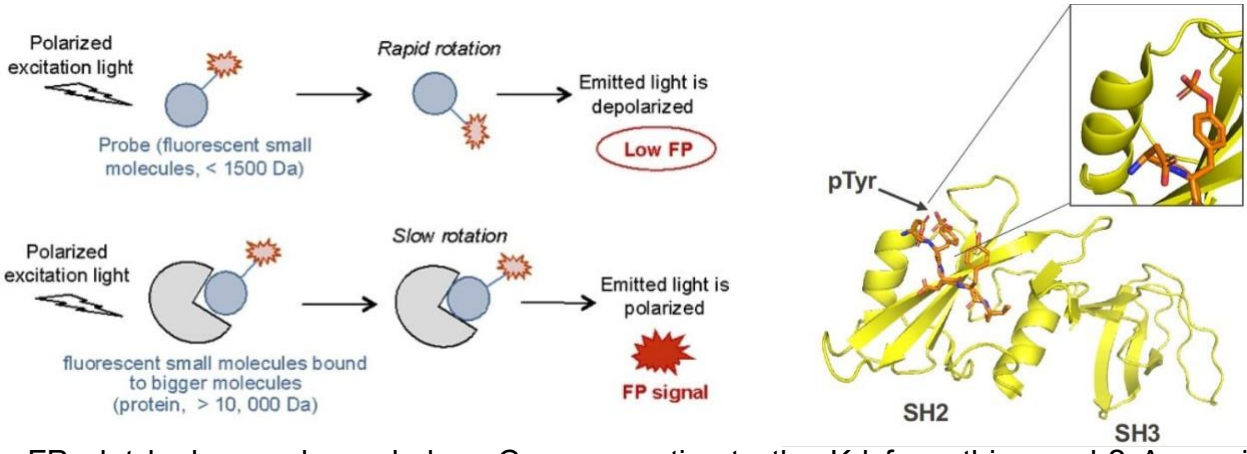
d) Why are some peaks shifting more and others less? Is this an indication of Kd values differing between amino-acids?
Answer: The extent of chemical shift change is dependent on the relative level of local chemical perturbation in the compound binding site. Typically, the closer an amino acid is to the binding site – the more shift it will exhibit as a consequence. So the peaks corresponding to some amino acids will shift more, others less, and most of them will not move at all. This is not an indication of different Kd values existing. In fact if you independently plot the shifts of each peak (amino-acid) as a function of inhibitor concentration, they should all give you the same Kd value. See the plot provided in solution for c) to get a better idea.

e) You suspect that the inhibitor may act by binding at the same site as the DNA molecule and preventing the formation of DNA – Protein M complexes. The structure of this complex is already known. How can you use the NMR data above to see if there is overlap in binding sites?
Answer: We can use the existing structure of DNA-Protein M complex to identify the amino acid residues on the Protein M side that interact with DNA. The same can be done for Protein M and TZN-98 using the HSQC data shown above. The contacts are defined as the amino-acids that exhibited shifts in ^{1}H-^{15}N HSQC spectra upon the addition of TZN-98. In the example above those would be amino acid residues at positions: 28, 31, 42, 51, 52, 54, 55, 60, 62, 63, 68, 72, 83, 86, 96, 99, 101, 102. By comparing the residues DNA and TZN-98 use to intract with Protein M, we can determine if there is overlap which would support or refute the mechanism proposed in the question.

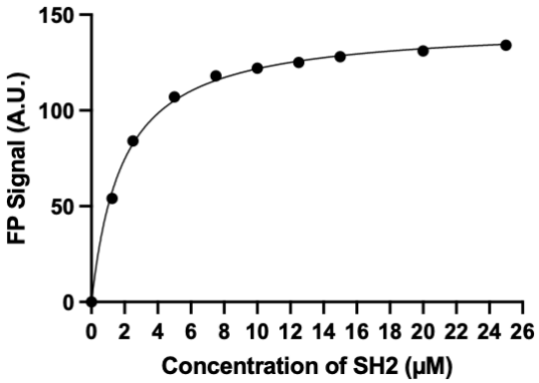
f) Which experiment could you use to directly check if there is competition between DNA and TZN-98 for the same binding site? Describe the details of how the experiment can be performed.
Answer: In principle any experiment described in class can be applied in a way that would allow to test for competition. However, the most straightforward one would be fluorescent polarization (FP). DNA molecule can be chemically modified with a fluorescent dye and used to test the interaction with Protein M in the presence and absence of TZN-98. If TZN-98 prevents the binding of DNA than there will be a reduction in fluorescent polarization signal when the drug is present.
The alternative versions of this experiment (i.e., fluorescently labeling Protein M or TZN-98) are theoretically possible but would not be very optimal. Protein M already has a high molecular weight so the amount of change in fluorescent polarization signal would be low upon binding to either molecule, and potentially not detectable. The labeling of TZN-98 may result in loss of inhibitory activity since this is a very small molecule and every chemical group could be important for binding. Therefore, chemical modification with fluorescent dyes is not recommended here.

Question 5: Fluorescent polarization to study protein-peptide interactions

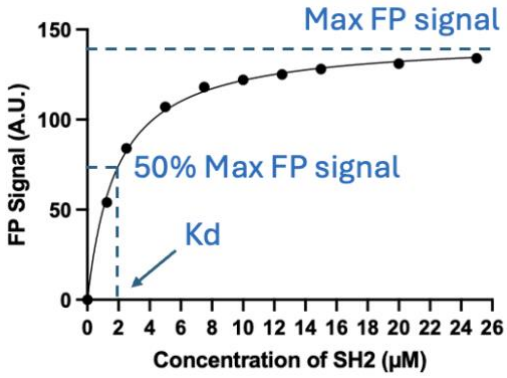
You want to determine the binding affinity of the SH2 domain of the Lck kinase to a peptide that contains a phospho-tyrosine residue. You synthesize the phospho-tyrosine peptide and include a fluorescent dye at the N-terminus. You measure fluorescent polarization (FP) in the absence and in increasing concentrations of SH2 domain.



a) The FP plot looks as shown below. Can you estimate the Kd from this graph? Approximate numbers are acceptable but please explain the process.



Answer: This is usually performed through a program (e.g., Excel). The process involves fitting a curve based on the raw data and calculating the concentration of SH2 that gives 50% maximum FP signal, which would correspond to Kd. In this example the Kd is approximately ~1.8-2.0μM.



b) You have isolated two monoclonal antibodies that bind to SH2 with very high affinity (~1nM). Now you wish to explore if they compete with the peptide so you perform an FP assay, as above. You pre-mix the SH2 domain (at 25μM) with each antibody separately (at 50μM). After the incubation you add the fluorescent peptide and perform the measurement. For reference you also performed the measurement

without antibodies in solution. You obtain the following FP signal values (in arbitrary units):

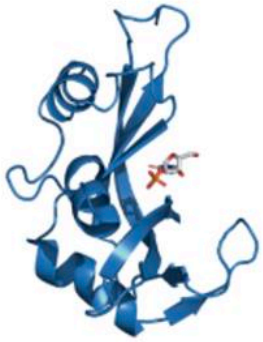
25 µM SH2	**148**
25µM SH2 +50 µM Antibody 1	**52**
25µM SH2 +50 µM Antibody 2	**560**

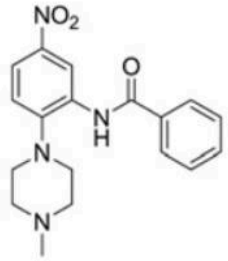
Why is the FP value for SH2 + Antibody 1 lower than for SH2 alone?
Why is the FP value for SH2 + Antibody 2 higher than for SH2 alone?

Answer:
Antibody 1 competes with pY peptide and prevents it from binding to SH2, therefore lower FP.
Antibody 2 does not compete with pY peptide and they can both bind to SH2 simultaneously. Due to the molecular weight contribution of the antibody to the complex, the total amount of FP signal is higher.

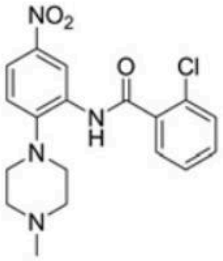
Question 6: Comparing protein inhibitors

WDR5 (WD40 repeat protein 5) plays key roles in development and is abnormally expressed in many cancers. Recent studies discovered small molecules that can disrupt the interaction between WDR5 and peptides from the catalytic domain of MLL (mixed-lineage leukemia protein) and are promising inhibitors for development of drugs for MLL-rearranged leukemias and other cancers. These are three compounds within this class:

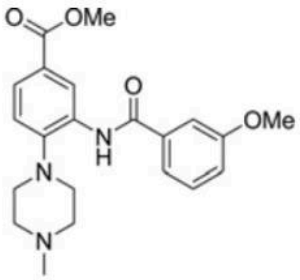




WDR5-0101

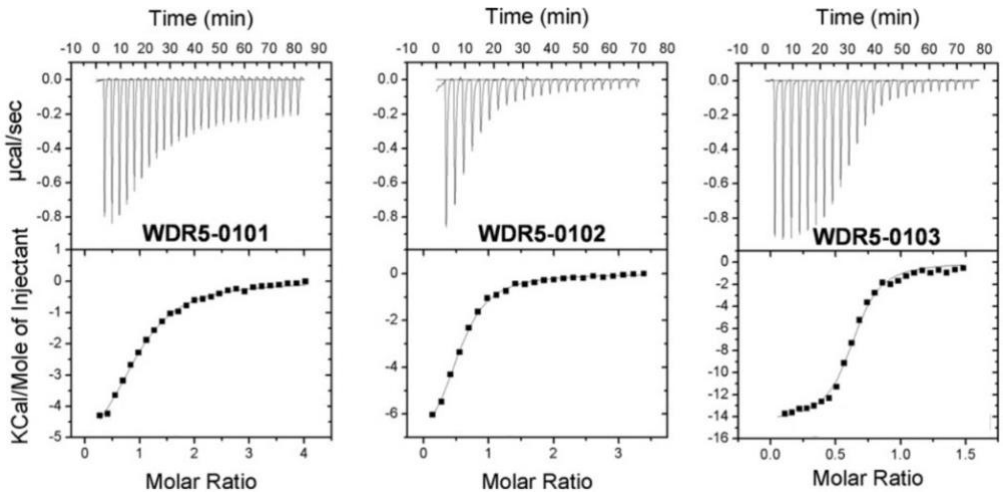


WDR5-0102

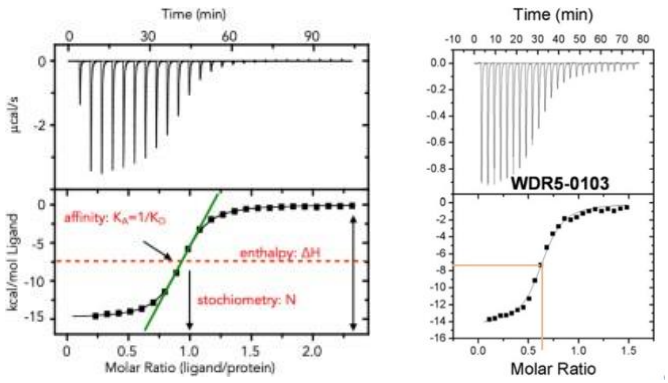


WDR5-0103

a) Scientists performed an assay to determine their binding affinity. See the results below obtained at 300 K. What is the technique used here? Which compound has the highest affinity (greatest Ka – lowest Kd)? Use the graph below to estimate the enthalpy and stoichiometry of binding for this compound? Note that the X and Y axes have different relative values in 3 cases which makes the comparison a bit trickier.



Answer: Use the graphs from class to help with analysis



The technique used is **ITC**. The compound with the highest affinity (Ka) is **WDR5-0103,** based on the sharpest relative increase in signal (note that the energy change on Y axis is higher compared to other 2 compounds).

The enthalpy of binding for this compound can be estimated as the difference between the minimum and the maximum heat energy levels which is approximately **-14 kcal/mol** for WDR5-0103.

The stoichiometry is the molar ratio at which 50% of the maximum binding signal is acquired, which corresponds to **~0.6-0.7 molar ratio (ligand/protein)** in the case of WDR5-0103

b) Assume that the most potent compound from the example above has a Kd of 0.45 μM. What is the free energy of binding (dG) and its entropic contribution (dS) for this compound? Hint: You will need to use some data from answer a).

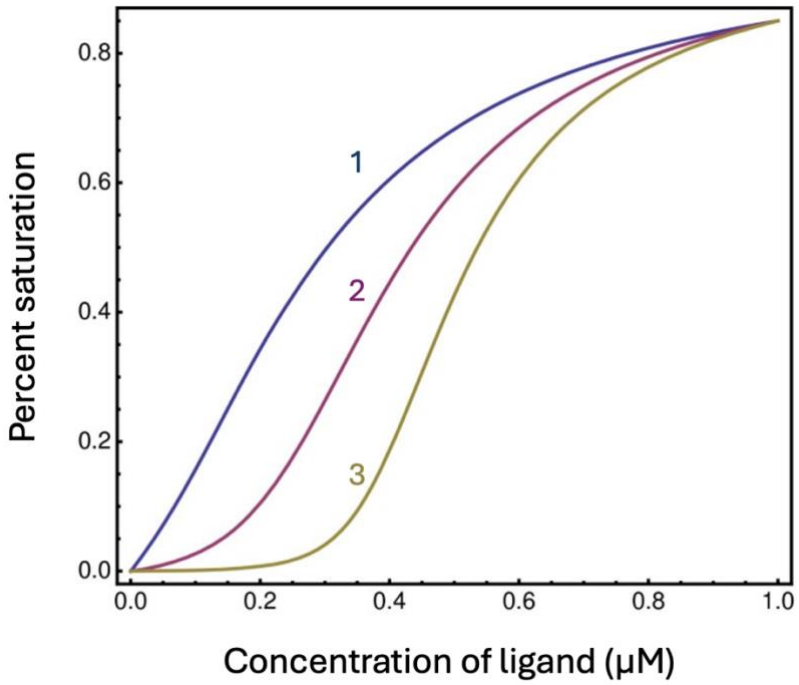
Answer:

Free energy of binding:
$dG = R * T * \ln(Kd)$
$dG = 8.314\ J*K^{-1}mol^{-1} * 300\ K * \ln(0.45*10^{-6})$
$dG =$ **$-36.45\ kJ * mol^{-1}$**

Entropic contribution:
$dG = dH - T*dS$
$dH = -14\ kcal*mol^{-1} = -58.6\ kJ*mol^{-1}$
$dS = (dH - dG)/T$
$dS =$ **$-70\ J*mol^{-1}K$**

Question 7: Cooperative binding

Below are 3 curves showing the response of different mutant variants (1,2,3) of a mutimeric sensor protein to the same ligand.



a) Rank the protein variants based on the degree of cooperativity displayed (from most to least cooperative) ?

Answer : Sample # 3 is the most cooperative as it features the most expressed sigmoidal phenotype, with the sharpest rise in percent saturation in response to ligand concentration. Sample # 2 is the second most cooperative, while Sample # 1 represents the least cooperative variant.

b) If the Hill coefficient for the Sample # 2 (magenta) is 1.5, what is the ratio of K_{d2} and K_{d1} values for the two cooperative binding states in this variant?

Answer : If we use the Hill coefficient equation :

$$n_H = \frac{2}{1+\sqrt{\frac{K_{D2}}{K_{D1}}}}$$

And solve it for K_{d2}/K_{d1}, we get :

$$K_{d2} / K_{d1} = (2/n_H - 1)^2 = 0.11$$

c) Would you expect the Hill coefficient to be higher or lower for Sample # 1 (blue) ? What about Samples # 3 ?

Answer : The Hill coefficient will be higher for Sample # 3 (more cooperative) and lower for Sample # 1 (less cooperative).