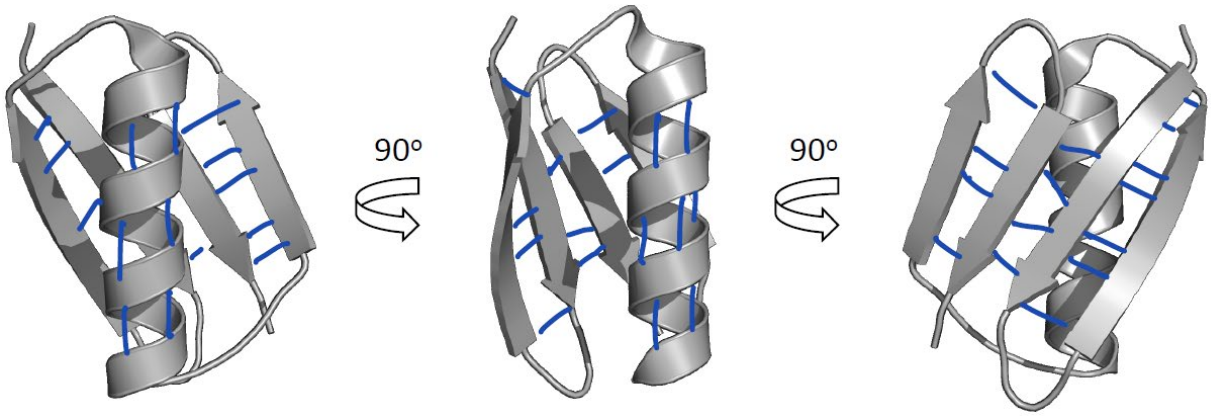
My answers to the exam problems are on the next pages.

I strongly recommend that you try doing these problems yourself BEFORE looking at my solutions.

Problem 1.

Here is a cartoon representation of the tertiary structure of a protein GB3. The pictures represent consecutive 90º rotations of the protein about the vertical axis.



A. List all elements of secondary structure (and how many) that are present in this protein. Be as specific as possible: for a beta-sheet, what type of sheet and how many strands, for a helix, what type of helix and how many turns.

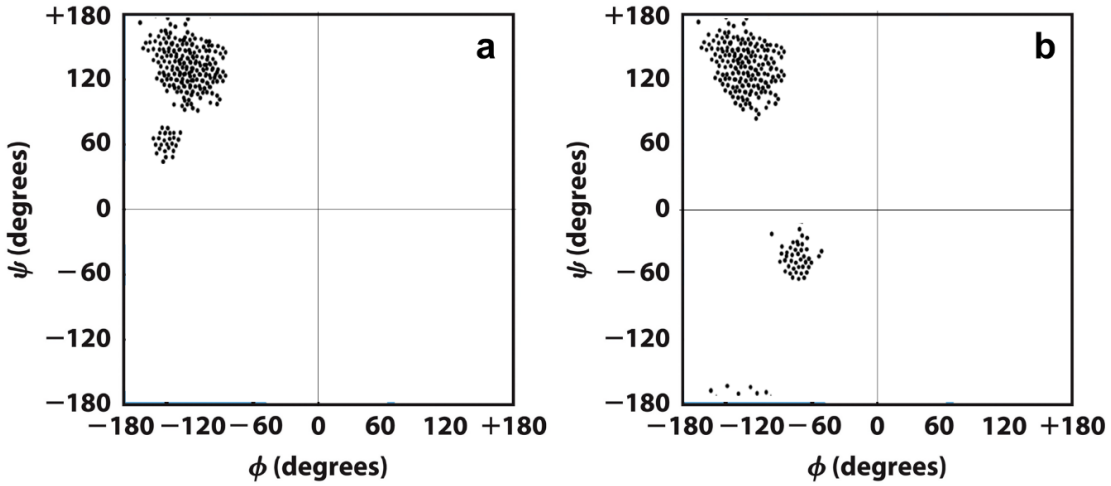
One α-helix (right-handed, approximately 4 turns), one β-sheet (consisting of 4 β-strands: 2 pairs of antiparallel and 1 pair of parallel strands), and possibly 2 β-turns connecting the antiparallel β-strands.

B. On the structural cartoons shown above, sketch backbone-to-backbone hydrogen bonds that you expect in this protein (just draw lines to schematically show hydrogen bonding patterns, no need to show individual atoms connected by these bonds). Explain briefly why you drew the lines the way you did.

The H-bonds in the α-helix connect residues in adjacent turns (carbonyl O of residue *i* and NH of residue *i*+4) and are almost parallel to the helix axis. The H-bonds in the β-sheet connect the neighboring β-strands and are perpendicular to the strand's direction. There are no H-bonds between the backbone atoms in the α-helix and β-sheet.

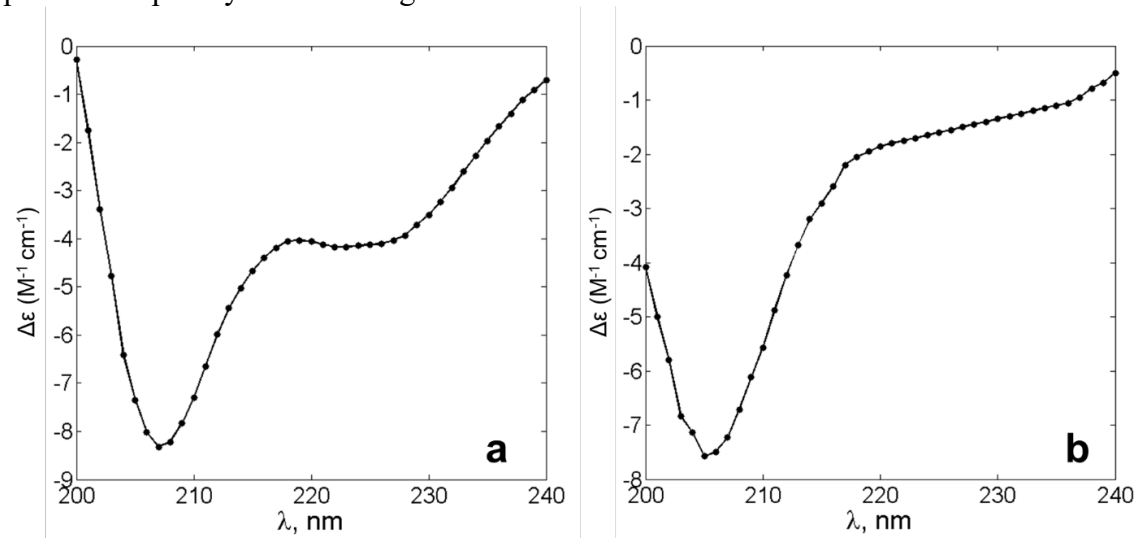
(A more accurate drawing should reflect the fact that the H-bonds between antiparallel β-strands are spaced in pairs, because both the NH and the CO of the same residue of one strand form H-bonds with the CO an NH groups of its bonding partner residue from the opposite strand, and that the H-bonds between parallel strands are a bit tilted from being perpendicular to the direction of strands.)

C. Which of the two Ramachandran maps shown below cannot correspond to this protein? (The dots represent pairs of ϕ, ψ angles for each residue.) Explain your reasoning.



The map in (**a**) **cannot** correspond to this protein because this map contains no dihedral ϕ, ψ angles in the signature α-helix region (ϕ ≈ -60°, ψ≈-50°)

D. Which of the two circular dichroism (CD) spectra shown below likely corresponds to this protein? Explain your reasoning.



Of the two spectra, only the one in (**a**) has two dips (strong negative signals) at λ =208 nm and at λ =222 nm, characteristic for the α-helix. Thus this CD spectrum corresponds to a protein with noticeable helical component. The CD spectrum in (**b**) likely corresponds to a protein with no or very little helical content.

Problem 2.

You discovered an unknown protein and – naturally – you want to characterize its structure. The protein has molecular weight of 45 kDa, and, based on your biochemical data forms a homodimer. The analysis of its primary structure suggests that the protein might contain several independently folded domains connected by unstructured regions. Consider all methods for structure characterization that we covered in class. Answer the following questions.

A. What method allows fast determination of the secondary structure content of a protein even though it provides no details regarding the actual structure?

Circular dichroism

B. What method can provide information about the overall shape of your protein although without revealing atomic-resolution structure?

Small-angle scattering of X-rays (SAXS) or neutrons (SANS).
In principle, cryo-EM also provides a shape-related information (as does X-ray crystallography), however – unlike small-angle scattering – cryo-EM is actually capable of providing structural information, at or approaching atomic-level resolution.

C. What features of your protein could potentially create problems for determining its structure using X-ray crystallography?

Unstructured regions (loops) connecting independently-folded domains are likely flexible/disordered, and therefore might present a problem for getting the protein to crystallize or might be disordered in crystals, resulting in the absence of electron density in these regions of the protein.

D. What features of your protein could potentially create problems for determining its structure using solution NMR spectroscopy?

The observation that your protein forms a homodimer suggest that it will behave/tumble in solution as a ~90 kDa protein, which is beyond the typical range of protein sizes amenable to NMR and could significantly broaden NMR signals due to slow tumbling.

E. What features of your protein could potentially create problems for determining its structure using cryoEM?

The unstructured/flexible regions will present a problem for class averaging needed to improve the contrast and resolution of cryo-EM images. Also the relatively small size of the protein could present a problem for cryo-EM (currently).

Problem 3.

When studying ligand binding to a protein X, you found that at ligand concentration of 35 μM 35% of the protein is in the bound state. Answer the following questions. Show your calculations and briefly explain your reasoning and assumptions.

A. What fraction of the protein can you expect to be bound at the ligand concentration of 70 μM?

Assumption: single ligand binding site on the protein (i.e. 1:1 binding stoichiometry)

Also, the *assumption* that the protein concentration is much lower than the ligand concentration, thus we can approximate [L] = [L_{total}]

$\theta = \frac{[L]}{[L]+K_d}$ ➔ $K_d = \frac{(1-\theta)}{\theta}[L]$. This gives $K_d = \frac{(1-0.35)}{0.35} \times 35\mu M$ = 65 μM. Substituting this value for K_d in the equation for the bound fraction of the protein we get θ = 51.85% at 70 μM.

B. You performed measurements at the ligand concentration of 195 μM and found that 95% of the protein was in the bound state. Compare your result with what you would expect at this ligand concentration. If your experimental result agrees with your expectation, what does this tell you about the binding? If there is a disagreement, what could be a possible reason for it?

The expected bound fraction at this ligand concentration is $\theta = \frac{195\mu M}{195\mu M + 65\mu M}$ = 75% < 95%.

Alternatively, we can estimate the K_d value at this concentration given the bound fraction of the protein: $K_d = \frac{(1-0.95)}{0.95} \times 195\mu M$ = 10.26 μM.

This indicates that at this ligand concentration the binding is tighter than expected. A possible reason for the increase in the strength of binding as the ligand concentration increases is positive cooperativity.

Problem 4.

A hypothetical allosteric protein has several identical subunits, each of which can bind ligand **L**. Each subunit can be in two states, T and R. The T state binds ligand **L** weaker than the R state. In the absence of ligand the T state is more stable. The protein fits the concerted (MWC) model for cooperativity, where all subunits undergo the transition from one state/conformation to another simultaneously. Ligand binding to the T state makes a transition to the R state more likely. Each subunit of this multimeric protein has a second binding site, to which other substances, **X** and **Y,** can bind. **X** binds only to subunit in the T state and stabilizes this state, while **Y** binds only to the R state and stabilizes this state.

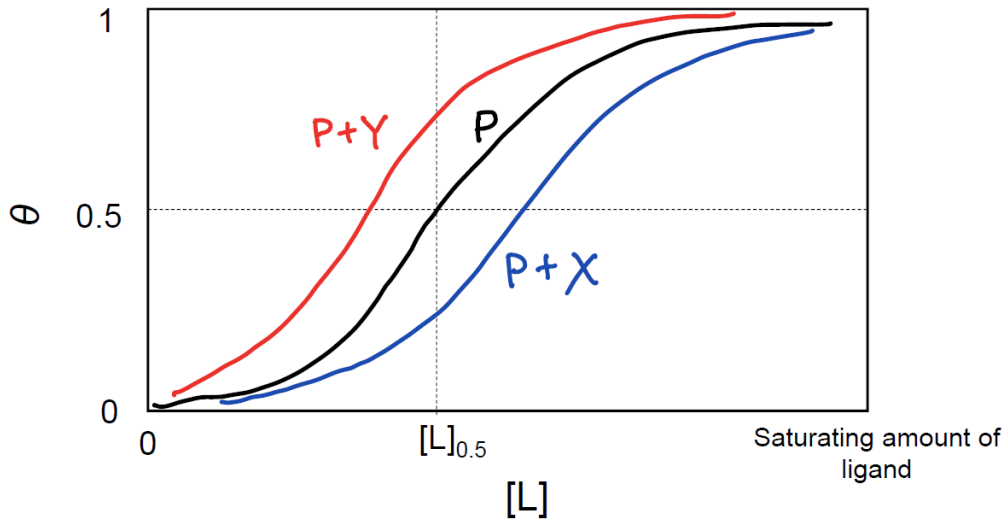
A. Which of the three ligand molecules that can bind to this protein (**L**, **X**, or **Y**) is a *negative heterotropic modulator* for this protein? Explain your reasoning.

Ligand X, because it stabilizes the lower affinity T state, hence binding of X will reduce the affinity for L.

B. Which of the three ligand molecules is a *positive heterotropic modulator* for this protein? Explain your reasoning.

Y, because it stabilizes the higher affinity R state, hence binding of Y will increase the protein's affinity for L. Binding of L will also increase the affinity for L as a result of positive cooperativity, but L is a homotropic not heterotropic modulator.

C. Using the axes below (ligand-bound fraction of protein (θ) versus molar ligand concentration), first draw the binding curve for ligand binding to your protein. Then sketch on this plot how the binding curve in the presence of a certain amount of **X** or **Y** (separately) should look like. Explain why you drew the curves the way you did.



See the drawing. The black curve represents cooperative binding of L to protein alone, when neither X nor Y is present. Answers to A and B provide explanations why I drew the binding curves in the presence of X (blue) or Y (red) like this.

Problem 5.

The enzyme ribonucleotide reductase binds several different nucleoside-triphosphate ligands, with a different affinity for each one. Listed below are the K_d values for three select ligands and their concentrations in cells at 25 °C.

Ligand	K_d (μM)	[L] in cells (μM)
ATP	100	1000
CTP	1	1
GTP	10	1

A. Determine the standard free energy of binding (ΔG'°) for each of the nucleotide ligands. Show your work.

$\Delta G'^{0} = RT \ln(K_d)$ = -22.8 kJ/mol for ATP,
-34.2 kJ/mol for CTP, and
-28.5 kJ/mol for GTP.

B. Assume that the concentration of the enzyme is much lower than the concentration of any of the ligands in the table above. Which enzyme-ligand complex is present at the highest concentration in the cell? Explain how you arrived at your answer.

All three ligands will compete for binding to the same protein in the cell. The relative concentration of each protein-ligand complex will be proportional to the fraction of the protein in each ligand-bound state, [PL]. (I will use the "protein-ligand" terminology here because the fact that ribonucleotide reductase is an enzyme is actually not important for the binding question). The latter, in turn, depends on how the protein is partitioned among its complexes with each ligand, which in turn, will depend on how the ligand concentration in the cell compares to the K_d for this ligand.

Qualitatively, if we use the equation $\theta = [L]/([L]+K_d)$ for the bound fraction of the protein with a given ligand *assuming that this is the only ligand present in the cell,* we can expect that the ATP-bound protein will be present at the highest concentration because the cellular ATP concentration is well above the K_d (hence the enzyme is almost saturated) while the GTP concentration is well below the K_d (hence only a small fraction of the enzyme is bound to GTP), and that of CTP equals K_d (half of the protein will be bound when no other ligand is present).

Quantitatively, we cannot use the above equation when more than one ligand is present because the protein is partitioned among complexes with the ligands: $[P_{total}] = [P] + [P\text{-ATP}] + [P\text{-CTP}] + [P\text{-GTP}]$. However, we can always use the equations that define the K_d through the ratio of the concentrations relevant to protein-ligand binding/dissociation equilibrium: $K_d=[P][L]/[PL]$, from which we get $[PL]= [P][L]/K_d$. This equation applies to each protein-ligand binding regardless of how the protein is partitioned among different complexes because [P] is just the free-protein

concentration. Using this equation for each complex, we get the ratio of the concentrations of protein-ligand complexes in question:

[P-ATP] : [P-CTP] : [P-GTP]= ([P][ATP]/K_{dATP}) : ([P][CTP]/K_{dCTP}) : ([P][GTP]/K_{dGTP}) =

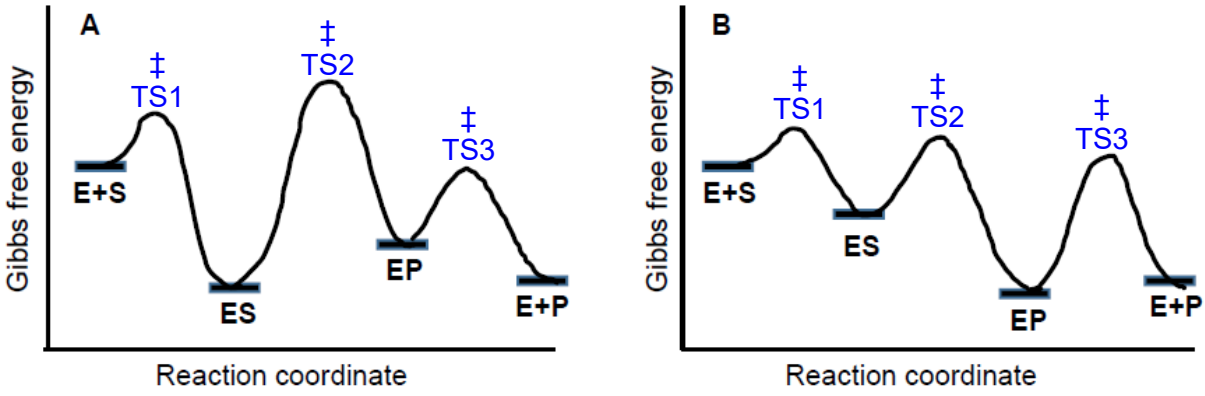
([ATP]/K_{dATP}) : ([CTP]/K_{dCTP}) : ([GTP]/K_{dGTP}) = (1000/100) : (1/1) : (1/10) = 10:1:0.1.

(*Assumption*: I assumed that I can use total the ligand concentration as the estimate for the free ligand concentration because the concentration of the protein is much lower than of any of the ligands)

This result clearly shows that the ATP-bound protein is present in a 10 times higher concentration than the CTP-bound protein and 100 times the GTP-bound protein.

Problem 6.

A reaction S ↔ P is catalyzed by two enzymes, A or B. The reaction coordinate diagrams for these enzymes are shown below. Answer the following questions.



A. Which of the two enzymes binds tighter to the substrate? Explain your reasoning

A, because the ΔG between E+S and ES is larger.

B. Which step in each reaction is the rate-limiting step for the forward reaction? Explain your reasoning

For A: ES → EP because the $\Delta G^{\pm}$ for ES → TS2 is greater than for any other step

For B: EP → E+P because the $\Delta G^{\pm}$ for EP → TS3 is greater than for any other step

C. Which step in each reaction is the rate-limiting step for the reverse reaction? Explain your reasoning

For A: ES → E + S, because the $\Delta G^{\pm}$ for ES → TS1 is greater than for any other step (the EP→ TS2 $\Delta G^{\pm}$ for EP → ES is somewhat lower than this one)

For B: EP → ES because the $\Delta G^{\pm}$ for EP →TS2 is greater than for any other step

D. Which of the two enzymes has higher turnover number? Explain your reasoning

Enzyme B, because the rate constant for the transition depends on the activation barrier as $k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\pm}}{RT}\right)$, and the $\Delta G^{\pm}$ for the rate limiting step (EP→P) of enzyme B-catalyzed reaction is smaller than the $\Delta G^{\pm}$ value for the rate limiting step (ES→EP) of enzyme A-catalyzed reaction.

E. Which of the two enzymes is a better catalyst? Explain your reasoning

Because both reactions use the same S and P pair, the rate of the uncatalyzed S ←→ P reaction is the same. Therefore the enzyme that provides a higher catalyzed reaction rate is the better catalyst because it will have the higher k_{cat}/k_{uncat} ratio. In this case it's B.

Some equations and formulae that might or might not be useful:

$\Delta G = \Delta H - T\Delta S$

$\Delta G = \Delta G'^{0} + RT\ ln\{[\text{products}]/[\text{reactants}]\}$

$K_{eq} = [\text{products}] / [\text{reactants}]$

$\Delta G'^{0} = - RT\ ln\{K_{eq}\}$

$K_{eq} = exp(-\Delta G'^{0}/RT)$

$R = 8.31$ J/K/mol

$k_B = 1.38 \times 10^{-23}$ J K^{-1}; $h = 6.62 \times 10^{-34}$ J s;

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\pm}}{RT}\right)$$

$$V_o = V_{\max} \frac{[S]}{[S] + K_m};\ \frac{1}{V_o} = \frac{K_m}{V_{\max}} \frac{1}{[S]} + \frac{1}{V_{\max}}$$

$K_d = 1/K_a = exp(\Delta G'^{0}/RT)$

$K_d = [\text{P}][\text{L}]/[\text{PL}]$

$$\theta = \frac{[L]}{[L] + K_d}$$

$$\theta = \frac{[L]^{n_H}}{[L]^{n_H} + K_d};\ K_d = ([L]_{0.5})^{n_H}$$

$$\log\left(\frac{\theta}{1-\theta}\right) = n_H \left(\log[L]\right) - \log\left(K_d\right)$$

$$\frac{k_{catalyzed}}{k_{uncatalyzed}} = \exp\left(\frac{\Delta G^{\pm}_{uncat} - \Delta G^{\pm}_{cat}}{RT}\right)$$

$V_{\max} = k_{\text{cat}}\ [\text{E}_{\text{total}}]$; $K_m = (k_{-1} + k_2)/k_1$
