

# **20.320 Exam 2**

Thursday November 15<sup>th</sup> 2012

9:35-10:55

***Instructions:***

- 0. Write your name on the front cover of the blue book.**
  
- 1. Answer all questions in the blue books. This exam paper will not be graded.**
  
- 2. All questions can be answered in at most a few sentences. We will deduct points for excessively long replies, even if they contain the right answer.**
  
- 3. State all assumptions for each problem.**
  
- 4. In order to pace yourself please note that the maximum possible score on this exam is 100 – note that there are 4 total questions.**

# Question 1

---

*25 points*

As discussed in class, the amount of receptor on the surface of a cell directly affects the maximal activation of downstream signaling cascades when the cell is pulsed with ligand. To quantify surface receptors, assume that receptor (R) is constitutively synthesized at rate  $V_s$ , that endocytosis of receptor occurs with rate constant  $k_{er}$ , and that activated endocytosis of receptor-ligand complex (C) occurs with rate constant  $k_{ec}$ . C is formed by binding ligand (L) with rate constant  $k_f$ ; ligand disassociates with rate  $k_r$ . Internalized receptors and complexes are degraded with rate constant  $k_{deg}$ . The internalized receptor can be recycled to the cell surface with rate constant  $k_{rec,R}$  while the internalized complex can be recycled to the cell surface with rate constant  $k_{rec,C}$ .

- Create a cell-level diagram of this system including all relevant terms.
- Write a system of ODEs including the rates of change for each species in the above system.
- A western blot experiment is performed to determine the total amount of receptors expressed by the cell at steady state. From this experiment, it is established that these cells have  $2 \times 10^8$  copies of receptor in a population of 10,000 cells. Assuming that  $V_s = 102 \text{ min}^{-1} \text{ cell}^{-1}$ ,  $k_{deg} = 1 \times 10^{-2} \text{ min}^{-1}$ ,  $k_{er} = 2 \times 10^{-2} \text{ min}^{-1}$ , and  $k_{rec,R} = 9 \times 10^{-3} \text{ min}^{-1}$ , calculate the fraction of total receptors that are internalized per cell at steady state conditions, in the absence of ligand. State any additional assumptions.

# Question 2

---

25 points

ETV3, an ETS-family transcription factor, is a transcriptional repressor: it represses transcription of the DNA to which it is bound. One of the genes to which ETV3 binds is the gene for ETV3 itself (with a  $K_d$  value of  $10^{-6}$ M), representing a nice example of transcriptional regulation.

- Plot the population distributions for negative, simple, and positive regulation and explain each curve. Which type of regulation does ETV3 exhibit?
- Provide an equation describing the rate of change of ETV3.
- Given a basal rate of synthesis (e.g. in the absence of ETV3 protein) of  $\beta \frac{\text{receptors}}{\text{cell*second}}$  and a rate of degradation of  $\alpha \frac{\text{receptors}}{\text{cell*second}}$ , how much ETV3 is present at steady state?
- (5 points) Provide a rough plot of the dynamics of this system prior to reaching steady state. Be sure to label your axes.

# Question 3

---

30 points

You are interested in designing two highly specific proteases each of which will target a different sequence. You intend to use these protease-target pairs in a synthetic biology experiment. The goal is to express each protease only under specific conditions, when it will cleave its target without causing any off-pathway effects.

You begin with the TEV protease, which cleaves the sequence ENLYFQG after the glutamine residue. This is an incredibly specific protease, and you will focus on altering that substrate specificity. Your goal is to mutate the TEV protease to develop a second version that is specific for the sequence RNLYFQG. In order to do this, you will start working with a catalytically inactive form of TEV, one which binds the substrate but does not cleave it.

- 1) In the **wild-type** protein, what properties of the binding pocket do you expect around the first amino acid in the target sequence?
- 2) How much of a free-energy difference do you need so that the mutant will bind RNLYFQG 10,000-fold better than ENLYFQG?
- 3) In order to speed up the design process, you use pyRosetta to test out the predicted changes in binding energy for each mutant.
  - a) Draw the free energy cycles for the wild type and mutant TEV protease binding the ENLYFQG and RNLYFQG peptide sequences.
  - b) Write an equation for the  $\Delta\Delta G$  of each protein for the two substrates.
  - c) To save computing time, you would like to eliminate the “alchemical” reactions for the free peptide. Is that legitimate? Justify your answer.
- 4) When your design is complete, you find that the system works very well for isolated target peptides. Both TEV and the mutant have high binding affinities for their own targets and low affinities for the wrong peptides. For example, the mutant peptide (TEV<sub>M</sub>) will bind well to the mutant substrate (Pep<sub>R</sub>) but not to the wild type substrate (Pep<sub>E</sub>). The ratio of  $K_1 / K_2$  for the reactions below is approximately 1,000.

Reaction	Equilibrium constant
$TEV_M + Pep_R \rightleftharpoons PepC_{MR}$	$K_1$
$TEV_M + Pep_E \rightleftharpoons PepC_{ME}$	$K_2$

Both proteins are much less efficient at binding their targets when the targets occur inside the sequence of a globular domain, but not equally so. In fact, the mutant binds to a target protein containing either the ENLYFQG or RNLYFQG sequences with the same affinity. These two versions of the target differ only in the one amino acid substitution (E to R). Specifically, the ratio  $K_3 / K_4$  for the reactions below is approximately 1:

Reaction	Equilibrium constant
$TEV_M + Prot_{R, fold} \rightleftharpoons Prot_{C_{MR}}$	$K_3$
$TEV_M + Prot_{E, fold} \rightleftharpoons Prot_{C_{ME}}$	$K_4$

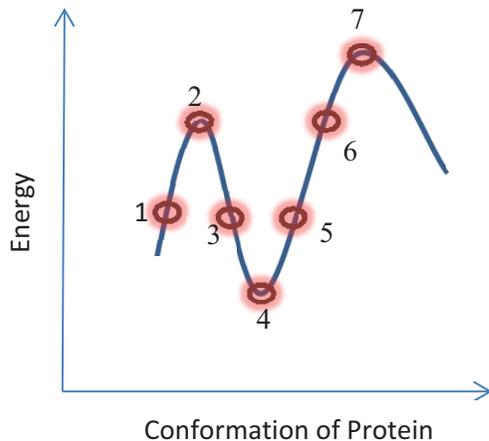
You wonder whether folding might contribute to this strange overall effect, and decide to include it in your model. Let  $K_{unfold}^E$  be the equilibrium constant for the unfolding of the protein containing ENLYFQG, and  $K_{unfold}^R$  the equilibrium constant for the corresponding reaction of the target containing RNLYFQG.

Reaction	Equilibrium constant
$Prot_{E, fold} \rightleftharpoons Prot_{E, unfold}$	$K_{unfold}^E$
$Prot_{R, fold} \rightleftharpoons Prot_{R, unfold}$	$K_{unfold}^R$

- Write an equation for the fraction of each peptide that would be bound by the catalytically inactive form of each enzyme in terms of the equilibrium constants above (these equations should include the effects of both folding and binding; write one equation for each peptide).
- Explain, based on the relative magnitude of  $K_{unfold}^E$  and  $K_{unfold}^R$ , why the enzyme's binding preferences change when the target sequence is part of a globular protein.

# Question 4

20 points



State #	Energy
1	100
2	105
3	100
4	95
5	100
6	105
7	120

Consider a system in which a protein can adopt seven discrete conformations. The relative energy of each conformation is depicted in the schematic and the table above. We're interested in applying the Metropolis-Hastings algorithm discussed in lecture to identify the globally optimal conformation which corresponds to state #4. In setting up the algorithm, we only allow adjacent states to transition to one another. In other words, state 2 can transition to states 1 or 3, while state 1 can only transition to state 2. In the algorithm, when a test state is generated, both of the states that neighbor the current state are equally likely to be generated. If the current state only has one neighbor (such as state 1), the test state corresponds to its only neighbor.

- a) Calculate the natural log of the probability that the current state transitions to the test state for each of the 12 combinations listed below in terms of  $k_B$  and  $T$ .  $P_{AB}$  represents the probability of transitioning from state A to state B when A is the current state and B is the test state.

Transition Probability	Natural Log of Probability
$P_{12}$	
$P_{21}$	
$P_{23}$	
$P_{32}$	
$P_{34}$	
$P_{43}$	
$P_{45}$	
$P_{54}$	

$P_{56}$	
$P_{65}$	
$P_{67}$	
$P_{76}$	

**For the parts that remain, assume that when we initialize our system, we're located in state 7.**

- b) What is the probability that we will **be in state #3 after 4 iterations** of the algorithm? Clearly explain your logic and express your answers in terms of the transition probabilities defined above. Omit transition probabilities of 1 from your final expression and don't neglect the probability associated with generating a test state.
- c) What is the temperature required for the algorithm to be in state #3 after 4 iterations with a probability greater than 0.05? For simplicity's sake, assume  $k_B = 0.04$  and that the units are in accordance with those used for the energy.
- d) What is the temperature required for the algorithm to be in state #3 after 4 iterations with a probability greater than 0.1? For simplicity's sake, assume  $k_B = 0.04$  and that the units are in accordance with those used for the energy.
- e) Rationalize the results obtained in parts c and d by explaining why a higher/lower temperature increases/decreases the probability of being in state #3 after 4 iterations.

MIT OpenCourseWare  
<http://ocw.mit.edu>

20.320 Analysis of Biomolecular and Cellular Systems  
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.