BMB 170 2017

Problem Set 2: Proteins II (96 Points)

Due 10/24/2017 12:00pm

Please remember to use ray tracing (set ray_shadows, 0) to generate all PyMol figures.
Turn in as a HARD COPY or as a PDF FILE by 12:00PM 10/25/2015. Points will be deducted otherwise.
PS#2 office hour:

Monday 10/23, 4:00-5:00PM, Sherman Fairchild Library (SFL) Group Study Room 231

TA: Inha Cho (Spalding 333, x4205)
Send questions and PDF file to: icho@caltech.edu

1. Protein Folding [28 points]

A. Levinthal's Paradox (i): Explain Levinthal's paradox, and explain how Levinthal's paradox leads to the conclusion that protein folding pathways more efficient than a random search must exist. [5 points]

Assuming a random search of all possible protein conformations for a 100-residue protein with 10 conformations per residue, it takes 10^80 years for this relatively small protein to fold. Since proteins fold in 10^-6 to 10^2 seconds, there must be folding pathways more efficient than a random search.

B. Anfinsen's Dogma: In his experiment, Anfinsen showed that denatured ribonuclease refolded into the active form. Describe at least three ways in which this experimental setup is different from physiological folding processes. [4 points]

1. The experiment was not done with chaperons present.
2. The experiment was not done in a crowded environment typical to the cytosol.
3. Nascent chains start folding when ribosomes are still attached, and the c-terminus of the protein is not always out of the ribosome while translation and folding occur.

C. Folding in vivo – Single molecule methods: Most protein folding studies on single molecule levels are using either atomic force microscope (AFM) or optical tweezers, for instance, as in the paper by Naqvi et al. (doi: 10.1016/j.bpj.2015.05.028). How does the single molecular technique allow the study of an intermediate state that was not observed in ensemble experiments? Can using an optical tweezer introduce artifact/errors? How so? [7 points]

Single molecular technique allows the measurement of a single molecule, being able to capture less abundant or fast-changing status that might be averaged out by ensemble experiments. The low force applied by the optical tweezers also allows a delicate manipulation that leads closely to the intermediate state. However, as external force is applied to direct the molecule to an intermediate state, the observed intermediate state might not be physiologically real. The molecules also need to be engineered in the first place, so that they can be "grabbed" by the tweezers.

D. Folding in vivo – Protein expression (i): Give three reasons why overexpression of a complex eukaryotic protein might lead to formation of inclusion bodies in a prokaryotic expression system. [6 points]

1. Expression of recombinant proteins requires an artificial induction of the gene. Depending on the rate of induction, the transcription and translation of the protein may be too fast for proper folding.
2. Supply of chaperones is limited, and not all overexpressed proteins have access to chaperones.
3. The overexpressed protein might not be soluble in E.coli cytosol for various reasons; hydrophobicity, or high cysteine content (thus forming disulfide bonds immediately).
4. Reducing environment of the E.coli cytoplasm is not compatible with certain disulfide bond formation that may be required for proper folding.

E. Folding in vivo – Protein expression (ii): Suggest three ways in which one can manipulate E.coli for better expression of properly folded proteins. [6 points]

1. Induce at lower temperature to facilitate slow protein transcription, translation, and folding.
2. Co-express effective chaperones.
3. Add soluble tags to help folding and expression.

2. Protein modifications [31 points]

A. Chaperones: Give three reasons why molecular chaperones do not violate Anfinsen's self-assembly principle while facilitating the correct fate of proteins in vivo. [4 points]

1. No inherent information in chaperone about proteins final fold
2. Prevent incorrect interactions within and between non-native proteins
3. Only assist self-assembly
4. Increase the yield but not the rate, typically

B. Psychrophiles: Most of earth's biosphere is not at 37°C; in fact about 80% of the biosphere is permanently below 5°C. Psychrophiles are extremophilic organisms that can survive and often thrive in such cold temperatures. Explain why psychrophiles cannot share the mesophilic tools of protein synthesis, in terms of how temperature affect the activity and stability of protein structures. What can happen if mesophiles were given with a complete psychrophilic protein synthesis pathway instead of their own, in room temperature? [6 points]

Low temperature results in reduced activity of transcriptional and translational enzymes that would work fine in room temperature. Low temperature also reduces speed of protein folding due to slower prolyl isomerization and the folding needs additional assistance. In lower temperature, DNA and RNA secondary structures are more stable and hinder transcription factor accessibility.

While psychrophilic transcriptional and translational enzymes can be more active in cold environment, they generally have adapted to be much more flexible structure to achieve such activity. As a result (although with exceptions) many components of psychrophilic protein synthesis pathway are not too stable in room temperature, and if a mesophile were given with a psychrophilic protein synthesis pathway, it could easily die.

C. Post-translational modifications: Post-translational modifications play crucial roles in regulating gene functions in lots of biological processes. Name one modification and briefly describe how this modification is involved in regulation of protein function. [5 points]

Phosphorylation, acetylation, methylation, lipidation (other modifications maybe involved too)

Examples maybe: histone phosphorylation, methylation and acetylation can regulate gene expression; phosphorylation in general regulates cell signaling.

D. Psychrophilic post-translational modification: It is vital for psychrophiles to maintain high capacity for post-translational modification. How would (1) the optimal temperature and (2) the expression amount of psychrophilic chaperones compare to those of mesophilic chaperones? [4 points]

Compared to mesophiles, psychrophiles have their chaperones adapted to be the most active at lower temperature. Psychrophiles also exhibit a high level of redundancy in their genomes for chaperones, resulting in chaperone overexpression.

E. Studying post-translational modifications: You plan to express a eukaryotic protein in yeast, and you would like to know if the protein of interest requires extensive glycosylation. Assuming that you know the exact sequence (and therefore the mass) of this protein and that the protein is likely to express very well in yeast, what experiments would you conduct to verify if the protein is heavily glycosylated? [5 points]

If it is obvious that glycosylation is the only major modification on this protein, running an SDS-page might be enough. Further detailed experiments can involve mass spectrometry if the sequence of the protein is known. Using antibodies directed against glycosylation, using metabolic analogues to label modified proteins to help further purification, and quantitative proteomic techniques such as SILAC (stable isotope labeling by amino acids in cell culture) are also reasonable options.

F. O-GlcNAcylation – Drug design: Post-translationally modified proteins are novel targets in drug design. O-GlcNAcylation has been shown to play a significant role in cancer development and is an interesting area to focus on in terms of anti-cancer drug development (Fardini et al., doi: 10.3389/fendo.2013.00099). Describe how the level of O-GlcNAcylation and its cycling enzymes such as OGT (O-glycosyl transferase) relates to cancer. How might targeting OGT be an interesting approach to overcome drug resistance in breast cancer? Why is it challenging to use O-GlcNAcylation of proteins as a mean to design anti-cancer drugs? [7 points]

The level of O-GlcNAcylation and its cycling enzymes both increase in cancer cells compared to normal cells. In breast cancer, significantly higher OGT expression is observed compared to healthy tissues. Tamiproxen is a widely used therapy to breast cancer, and O-GlcNAcylation inducing treatment protect cells from tamiproxen-induced cell death whereas OGT inhibition has the opposite effects, suggesting that targeting OGT might be an interesting approach to overcome tamiproxen resistance.

It is challenging to use O-GlcNAcylation as a mean to design anti-cancer drugs, as we lack the tools necessary to specifically study the impact of O-GlcNAcylation on carcinogenesis. Many proteins are modified by O-GlcNAcylation, it is relatively harder to detect O-GlcNAcylated residues, and strategies to alter O-GlcNAc-cycling enzymes and specific targets among their numerous substrates are yet to be developed.

3. Assemblies, viruses, symmetrical assemblies [10 points]

A. Oligomers: Many of the proteins in the Protein Data Bank (PDB) are oligomeric complexes consisting of two or more subunits that associate by rotational or helical symmetries. List 3 reasons why having a protein with oligomeric structure compared to a monomeric asymmetric sequence might be beneficial for a cell. [4 points]

Less error in translation

Easier to fold small polypeptide chains over large chains
Binding cooperatively and signaling
Genetic economy: reduce genome size by using same subunits
Evolution: mutation effects will be amplified
Smaller solvent exposed surfaced area and higher stability against denaturation

B. Virus assembly: Viruses have protein shells surrounding their genome, and the viral shell is often highly symmetrical. Why is it preferable for a virus to have a symmetrical shell? While it should be easy to obtain the shell structure of a virus thanks to its symmetry, it is hard to get crystals of the shell of an enveloped virus. Why is that? [4 points]

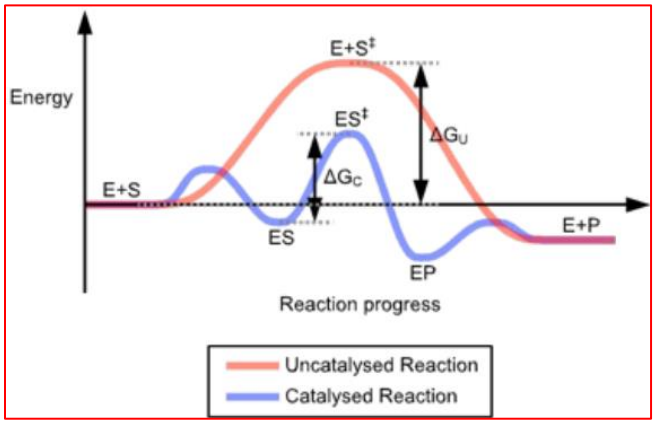
Virus needs a large shell to contain a small genome, so it is preferable for a virus to have a symmetrical shell that it can build the big shell from many copies of just a few different subunits.
Enveloped viruses are surrounded by host membranes on top of its shell. The surface lacks symmetry as a result, and it is hard to get crystals for enveloped viruses.

C. Quasi-equivalence: In 1962, Donald Caspar and Aaron Klug developed theoretical framework accounting for the properties of larger (>60 units) particles with icosahedral symmetry. Explain the principle of quasi-equivalence. [2 points]

Quasi-equivalance: each subunit is facing similar but not identical structural environments. Triangulation number T describes the number of structural units at each facet.

4. Metals and enzymes [27 points]

A. Transition state theory: Enzymes enhance biological reaction rate by 10^5 ~ 10^{20} fold compared to uncatalysed reaction. How does enzyme achieve such high rate enhancement? How does the free energy of reactants and products change for a catalyzed reaction? How about the magnitude of activation energy? Briefly explain the transition state theory. If necessary, include a figure. [6 points]



Enzymes achieve rate enhancement by decreasing the activation energy. The free energy of reactants and products do not extremely alter the free energy of reactants and products (although in some cases enzymes destabilize the reactant to decrease the activation energy barrier). Transition state theory assumes quasi-equilibrium between the activated complexes and the reactant molecules, and activated complexes can also convert into products. Kinetic theory can be used to calculate the rate of conversion

to the product. Details of how the activated complexes and the transition state are formed are not important to studying the reaction rate.

C. Metals in enzymes – crystallographic method: Although a high-resolution structure of nitrogenase Mo-Fe protein was determined in 2002, the central light atom ligand of nitrogenase Fe-Mo cofactor has long been a mystery. In 2011, Spatzal et al. (doi: 10.1126/science.1214025) presented evidence for interstitial carbon in nitrogenase Fe-Mo cofactor. They used atomic-resolution x-ray diffraction data and an electron spin echo envelope modulation (ESEEM) analysis. Describe, briefly, the basics of the techniques employed and explain why these techniques provide a more direct and conclusive evidence regarding the identity of the interstitial atom. [6 points]

X-ray crystallography: obtained a 1.0 A resolution structure

Residual B factor vs. atomic positions – distinct atoms occupy distinct areas; the mystery interstitial atom occupies the same area as Carbon – precise B factor correlation was possible because the structure is very high resolution.

Used probe radius of 1.4 A; varied on fine ED grid and performed statistical analysis.

Resonance spectroscopy of Azotobacter Vinelandii MoFe protein labeled with 13C or 15N was performed.

ESEEM detection of weak hyperfine couplings in paramagnetic moieties

Only 13C labeled sample had an additional spectral pattern.

D. Metals in enzymes – mechanistic insight: In a more recent publication, Spatzal et al. (doi: 10.1126/science.1256679) gain even greater detailed insight into the mechanism of nitrogenase. Describe the tricks that the author used to crystallize the protein in this paper. Why would crystallization with the substrate be so difficult? [4 points]

The authors used an extensively-studied inhibitor, CO, for crystallization. The entire process required continuous presence of CO. MoFe-protein for crystallization was obtained from activity assay mixtures, not from isolated protein. Crystallization occurred rapidly through seeding and in the presence of CO. It is difficult to crystallize with substrate because the isolated protein does not bind substrates; it requires other factors such as the Fe-protein, ATP, and ATPase. The factors, however, lead to turnover, which is too dynamic for the system to be crystallized.

E. Metals in enzymes – low-temperature adaptation: Azotobacter vinelandii, the origin of the nitrogenase studied by Spatzal et al., is a free-living aerobic soil bacteria. Azotobacter species are ubiquitous and some are even found in Arctic and Antarctic soil. The cold climate of Artic and Antarctica poses a huge challenge to the basic Azotobacter nitrogenase with Fe-Mo cofactor. What alternative do Azotobacter species have to the Fe-Mo nitrogenase? Compare the activity of Fe-Mo nitrogenase and Fe-V nitrogenase at room temperature and in a cold environment. [5 points]

Azodobacter species have alternative types of nitrogenases that have Fe-V cofactor or Fe-Fe cofactor. Fe-Mo nitrogenase is more active than Fe-V nitrogenase at room temperature, but as temperature decreases, the activity of Fe-V nitrogenase decreases less than that of Fe-Mo nitrogenase. Eventually, Fe-V nitrogenase is more active than Fe-Mo nitrogenase at very low temperatures.

F. Biological vs. nonbiological catalysts: On one hand, enzymes sometimes provide a much higher reaction rate enhancement than nonbiological catalysts such as palladium. On the other hand, a vast majority of industrial catalysts are metals or nonbiological small molecules. What are advantages of using enzymes instead of a chemical catalyst? Why is a massive scale-up of enzymatic reaction difficult? [6 points]

Enzymes provide milder reaction conditions: physiological pH, temperature close to room temperature. Chemical catalysts often require high pressure and temperature. Enzymatic reactions usually use aqueous buffer and not too acidic or basic conditions, whereas chemical catalysts often require organic solvent and other harsh conditions. In some cases enzymes are much cheaper and accessible than small molecule catalysts.

Enzymes are proteins, hence they're prone to degradation. Sometimes enzymatic reactions could be performed in a living cell so that the host cell can protect and continuously produce enzymes, but an environment favored by this host cell could be favorable to other organisms and the reaction chamber can end up contaminated. Some powerful and stable enzymes often require a very complicated protein expression pathway that cannot be easily achieved in quickly reproducing bacteria, and massive expression of those enzymes can require a lot of labor.