CUMULATIVE EXAMINATION IN BIOCHEMISTRY

1. Amyloid plaques are the signature of neuropathological diseases, such as Alzheimer disease. The plaques are built of peptide (protein) material. What is the arrangement of peptide chains in the plaques (i.e. the principal secondary-structure motif)?

Cross-beta sheet. The chains are stacked on top of each other, bonded with interchain β-sheet bonds.

2. Controlled trypsinolysis is often used as a probe of change in conformational state of proteins. What is the conceptual basis of this approach?

Cleavage sites (in the case of trypsinolysis, centered at Lys and Arg) can be either exposed to proteolytic attack or shielded – depending on the protein conformation. By analyzing the fragmentation pattern, we can therefore obtain the evidence of conformational changes.

3. An aqueous solution of a protein is prepared for biochemistry studies. In order to improve the solubility of the protein it is recommendable to adjust the pH of a solution such that the pH is substantially higher or substantially lower than the protein isoelectric point, pI. What is the reason for this recommendation?

Footnotes: (i) the buffer should be selected accordingly (ii) extreme pH values should be avoided to prevent protein unfolding

When far away from the pI, the protein surface tends to be charged (either positively or negatively). The electrostatic repulsion between the protein molecules would prevent aggregation and thus improve the solubility.

4. The so-called ‘Levinthal paradox’ suggests that it will take peptide chain an infinitely long time to sample its entire conformational space – and therefore an infinitely long time to find a global free energy minimum (i.e. to fold). How to explain this paradox?

The protein does not fold via a random conformational search. Instead, there is a ‘folding funnel’ – i.e. the peptide chain first forms certain elements of secondary structure and certain local contacts, and from there the system quickly moves downhill toward the global energy minimum (toward the bottom of the funnel).

5. The so-called ‘Anfinsen’s dogma’ states that the protein’s 3D structure is uniquely determined by its primary sequence. Is this an absolute principle or are there exceptions to it?

There is nothing fundamental about it. It is possible to have two co-existing folds with the same free energy. Those proteins are called “metamorphic proteins” (see Biochemistry 2011, 50, 7077–7079). Perhaps there are evolutionary reasons that favor the Anfinsen’s principle (predictability of the protein structure), but no fundamental thermodynamic reasons.

6. Burial of hydrophobic surface upon formation of a protein complex is entropically favorable. Conversely, dissociation of a complex incurs an entropic penalty. Explain the origin of this entropic penalty.
When water comes in contact with a hydrophobic protein surface, a number of possible arrangements of water molecules on the surface is reduced (compared to bulk or to hydrophilic surface). Hence, the entropic penalty comes from the ordering of the water molecules.

7. Disulphide bridges are rare in cytosolic proteins, but common in secretory proteins. Why?

The environment in the cytosol inside the cell is reducing (e.g. through the presence of glutathione), whereas outside the cell it is oxidizing (e.g. through the presence of molecular oxygen).

8. For protein-protein association the maximum on-rate $k_{on}$ is described as “diffusion-limited rate” which amounts to $10^9 - 10^{10}$ M$^{-1}$s$^{-1}$. Typically, $k_{on}$ is several orders of magnitude slower than that. What are the factors that slow down the binding (relative to the diffusion limit)?

Diffusion limit assumes that the two proteins are “fully reactive spheres”, i.e. if they come in contact with each other they would necessarily form a complex. In reality, they need to be properly aligned to form a complex, i.e. to have a correct mutual orientation. Also the successful binding may depend on the dynamics of the surface loops and the side-chains—only the right conformation (which may occur infrequently) may be conducive to binding.

9. The protein Databank coordinate set 1MBO shows the structure of myoglobin with an O$_2$ molecule bound to the heme group. Close inspection of the structure reveals that there is no path available to O$_2$ to leave myoglobin. If so, then how can myoglobin unload O$_2$?

There is dynamics in actual protein that creates a “dynamic channel” (i.e. O$_2$ can travel through small cavities that are transiently opening in the structure). This has been extensively investigated by laser spectroscopy (e.g. PNAS 101, 18000).

10. What is the difference between competitive and noncompetitive enzyme inhibitors?

Competitive inhibitor competes with the substrate for the same binding site. Non-competitive inhibitor binds to a different site, but renders the enzyme:substrate (ES) complex unproductive.

11. In order to study folding kinetics, a stopped-flow fluorescence experiment has been performed on a small protein domain. In doing so, 1 part of a denatured protein solution was rapidly mixed with 9 parts of a refolding buffer; then a fluorescence signal (mainly from a single tryptophan present in the protein) was monitored as a function of time. Predict the shape of the experimental curve (fluorescence intensity vs. time).

Monoexponential. Small proteins typically fold according to two-state, ‘all-or-nothing’ model. This gives rise to a monoexponential curve, which simply reflects the build-up of the folded species. See Biochemistry 36, 3373 for the real-life example.

12. A protein sample was dissolved in D$_2$O and left on a bench for 10 mins. Then the sample was rapidly digested and the obtained peptide mixture was used to take a MALDI-TOF mass spectrum. The result: loop regions in the protein showed a higher level of deuteration than α-helices. Why?
In loop regions, backbone amide $^1$H are readily exchanged for $^2$H. Not so in α-helices, where $^1$H atoms are protected by hydrogen bonds and the exchange takes much longer than 10 mins.

13. A protein has two binding sites, BS1 and BS2, located close to each other on the surface. Small molecule M1 binds to BS1 with dissociation constant $K_d^{(1)} = 0.1$ mM. Small molecule M2 binds to BS2 with $K_d^{(2)} = 1.0$ mM.

We now make a construct where M1 and M2 are covalently linked through a tether (assume that the tether is sufficiently long and flexible and does not interact with the protein). This construct, M1-M2, binds to the protein such as shown in the figure (right portion). Estimate the dissociation constant of the construct, $K_d^{(12)}$, and briefly describe your reasoning.

Ideally, we expect $K_d^{(12)} = 0.1$ μM. Here is the reasoning. To break off the M1 fragment we need $\Delta G^{(1)} = -kT \ln(K_d^{(1)} / K_0)$ (in this expression $K_0$ describes the standard reference equilibrium where concentrations of all species are 1 M, i.e. $K_0 = 1$ M). To break off the M2 fragment we need $\Delta G^{(2)} = -kT \ln(K_d^{(2)} / K_0)$. To break off both M1 and M2 we need $\Delta G^{(12)} = \Delta G^{(1)} + \Delta G^{(2)} = -kT \ln(K_d^{(1)} K_d^{(2)} / K_0^2)$. Hence, $K_d^{(12)} = K_d^{(1)} K_d^{(2)} / K_0^2$, which is 0.1 μM. Of course, in real life there could be binding cooperativity, interference from the linker, etc.

14. Please, describe the process that is illustrated in the sketch.

Receptor on the cell surface loads the ligand & dimerizes, becomes internalized, then captured in an endosome. Then there are 2 routes – it can be recycled (returned to the surface) or routed toward late endosome and then lysosome where it is degraded.

15. One attempts building a homology model for a protein with an unknown structure:
   (a) using a template with 51% identity to the target sequence;
   (b) using a template with 29% identity to the target sequence;
(c) using a template with 14% identity to the target sequence. Please analyze the feasibility of the homology modeling for all of these cases and describe the expected quality of the models.

(a) High-quality model comparable to the experimental NMR structure. (b) Modeling is highly problematic (so-called ‘twilight zone’). The model is likely to be crude, as characterized by a large rmsd, and may be incorrect altogether (incorrectly predicted fold). (c) Homology modeling is impossible (‘midnight zone’).
-Answer Key-
Inorganic Chemistry Cumulative Exam
Purdue University
December 8, 2018

Question 1:
Lobsters use hemocyanin, a copper protein, to transport oxygen. People and whales use iron-containing proteins such as hemoglobin and myoglobin.

Question 2:
EDTA chelates the metals in food, thereby preventing bacteria from having access to these metals. If bacteria can’t get metals, they will not thrive on the food.

Question 3:
A) MRI works by examining the variability of nuclear relaxation of water molecules. In the presence of Gd$^{3+}$, some water molecules will bind to this paramagnetic center (as a ligand), thereby relaxing the water protons. Thus the Gd$^{3+}$ enhances the water relaxation signal wherever this ion is located.

B) The Gd$^{3+}$ ion has 7 unpaired electrons, more than any other ion. Thus Gd$^{3+}$ is a more effective agent for influencing the relaxivity of water.

Question 4:
A) Zinc fingers are protein loop structures used to bind DNA. These groups are made from Zn$^{2+}$ bound by two histidines and two cysteines.

B) Zinc avoids problems with redox chemistry. Zinc fingers are inert structures, not meant to have any reactivity. If the metal center was, say iron, there might be oxidation or reduction reactions, leading to unwanted (and possibly toxic) chemistry.

Questions 5:
Burying the heme in the hydrophobic protein interior stabilizes the neutral Fe$^{2+}$ heme complexes. Conversely, exposing the heme to the polar water environment helps to stabilize the higher charged Fe$^{3+}$ complex. Thus reduction potentials decrease with increasing heme exposure.
Org. Lett. 2018, 20, 6970 - 6974
Org. Lett. 2018, 20, 6750 - 6754
Org. Lett. 2018, 20, 6736-6740
6. (cont'd)

\[
\begin{align*}
\text{Org. Lett.} & \quad 2018, 20, 6812-6816 \\
\beta - H \text{ elimination}
\end{align*}
\]

\[
\begin{align*}
(Cp^* \text{Rh}^I & \xrightarrow{\text{Cu(OAc)}_2} Cp^* \text{Rh(OAc)}_2) \\
\end{align*}
\]

7.

\[
\begin{align*}
\text{Org. Lett.} & \quad 2018, 20, 6886-6890
\end{align*}
\]

\[
\begin{align*}
\text{[4+2]}
\end{align*}
\]