CUMULATIVE EXAMINATION IN ANALYTICAL CHEMISTRY
Apr 26, 2008

It is expected that the answer should contain several succinct straight-to-the-point sentences. If you feel like elaborating, please stay within an 8 sentence limit anyway. All questions carry the same weight. After each individual exam is graded, two questions with the lowest score will be dropped. In other words, you can completely miss two questions and still earn a perfect total score.

1. A (non-covalent) protein complex involving toxin A and antitoxin B is studied as a part of a drug development program. As a first step, the binding interface needs to be 'mapped out' (i.e. the portions of the polypeptide chain that form the binding interface need to be identified). For this purpose one can employ a mass spectrometry experiment which relies on H/D exchange of amide protons. Describe, in general terms, the concept of such an experiment.

Two measurements need to be carried out: the first is performed on A alone, while the second is performed on the A-B complex. In each case, the stop-flow apparatus is used to dissolve the protein material in D$_2$O. The H/D exchange is thereby initiated and allowed to proceed for a period of time. The exchange is then quenched (for example, by rapidly altering pH) and the protein material is analyzed by mass spectrometry (e.g. electrospray or MALDI).

In the case of the A-B sample the uptake of deuterium is lower than in the A sample because certain amide sites are shielded by the ligand B. Using controlled proteolysis (e.g. with pepsin) it is possible to identify the specific fragments showing lower deuterium uptake - these are the fragments located at the A/B interface.

Note that the experiment is not completely fool-proof. Imagine that the binding interface is formed by $\alpha$-helices, where amide protons are 'locked' into hydrogen bonds. For such amides the H/D exchange is greatly slowed down and one may not be able to discern the effect of the ligand B.

2. A (non-covalent) protein complex involving toxin A and antitoxin B is studied as a part of a drug development program. As a first step, the binding interface needs to be 'mapped out'. For this purpose one can employ a popular NMR experiment, which relies on the chemical shift information ('chemical shift mapping'). Describe, in general terms, the concept of such an experiment.

Two spectra need to be recorded: that of A alone and that of the A-B complex. Usually, 2D or 3D spectra are needed in order to resolve many spectral lines. Typically, 2D $^1$H-$^1$N correlation maps (HSQC spectra), which provide the footprint of the amide groups, are used for this purpose.

Comparing the chemical shifts of the individual amide groups (i.e. peak positions in the 2D spectra) we note that the chemical shifts of amides located at the interface
change upon binding of B, whereas the chemical shifts of amides located away from the interface remain essentially unchanged. This is understandable since the binding of B alters the local environment for the interface amide moieties. In this fashion the residues comprising the interface can be identified directly from the change in the chemical shifts.

Note that this experiment is not completely fool-proof. Sometimes binding causes a global conformational change which leads to chemical shift changes throughout the protein.

3. A (non-covalent) protein complex involving toxin A and antitoxin B is studied as a part of a drug development program. To determine the binding affinity, a titration is carried out and the inherent Trp fluorescence is measured through the course of the titration. How can these data be used to determine the binding affinity (i.e. the dissociation constant $K_d$)?

Footnote: for simplicity, assume that A contains a single Trp residue, while B contains none.

The total fluorescence signal is:

$$F_t = \frac{[A]}{[A]+[AB]}F_{free} + \frac{[AB]}{[A]+[AB]}F_{bound}$$

where $F_{free}$ and $F_{bound}$ are (presumed to be different!) signal intensities from free and bound $A$, respectively. The theoretical expression for titration profile can be obtained if one recalls that:

$$K_d = \frac{[A][B]}{[AB]}$$

and, for analytical concentrations,

$$c_A = [A] + [AB]$$

$$c_B = [B] + [AB].$$

The resulting formula for $F_t$ can be used to fit the titration data with three fitting parameters - $F_{free}$, $F_{bound}$, and $K_d$.

4. A (non-covalent) protein complex involving toxin A and antitoxin B is studied as a part of a drug development program. An ultimate characterization of the complex is offered by the X-ray crystallographic structure. As it turns out, all but a few crystallographic structures deposited in the Protein Data Bank do not contain proton coordinates (only heavy atom coordinates). What is the fundamental reason for this?

X-ray radiation is scattered efficiently by electrons (and only to a negligible extent by much heavier nuclei). Hence, protons are very inefficient scatterers compared to carbons, nitrogens, and oxygens (in fact, the scattering by each atom is, to a very good approximation, proportional to its atomic number).
5. A (non-covalent) protein complex involving toxin A and antitoxin B is studied as a part of a drug development program. To confirm the purity of the sample, an SDS-PAGE gel has been used. What bands may one expect to see in this gel (monomers of A and B, complex A·B, dimers, trimers, higher-order oligomers of A, B, or A·B)?

SDS-PAGE gel means denaturing conditions. Most likely, only monomers of A and B will be detected under these conditions. It is possible, however, that denatured proteins will form higher-order species such as oligomers or complexes – those may also be visible in the gel. Most often this happens due to the formation of disulfide bridges between the extended polypeptide chains.

6. Protein A has been extracted from a biological tissue and purified using liquid chromatography (LC). In operating an LC column it is important to maintain the optimal flow rate. If the flow is too slow or too fast, the outcome of the separation procedure can be unsatisfactory. Why it is not a good idea to run the column with an exceedingly low flow rate?

If the solute ‘plug’ is allowed to stay in a column for a very long time, it is going to ‘diffuse out’ leading to a broad chromatogram band (cf. longitudinal diffusion term in van Deemter equation).

7. Protein A has been extracted from a biological tissue and purified using liquid chromatography (LC). In operating an LC column it is important to maintain the optimal flow rate. If the flow is too slow or too fast, the outcome of the separation procedure can be unsatisfactory. Why it is not a good idea to run the column with an exceedingly high flow rate?

If the ‘plug’ moves too fast the performance suffers from statistical fluctuations involving the exchange between mobile and stationary phases. Say, over a short period of time, \( t = 1\ s \), some solute molecules interact with stationary phase once, others two or three times, yet others not at all. As a result the ‘plug’ spreads out. If the flow rate is \( u = 1\ cm/s \) then the spread is \( \sim 1\ cm \); if the flow rate is \( u = 10\ cm/s \) then the spread reaches \( \sim 10\ cm \), etc. Hence, faster rates lead to broader bands (cf. mass transfer resistance term in van Deemter equation).

8. The \( ^1H \) NMR spectrum of a certain protein sample recorded with 16 scans has a signal-to-noise ratio \( S/N=5.0 \). How many scans are needed to produce a spectrum with \( S/N=20.0 \)?

256 scans. \( S/N \) increases as a square root of the number of scans (signal increases as \( n \), noise increases as \( \sqrt{n} \), hence \( S/N \) increases as \( \sqrt{n} \)).

9. Certain small protein contains a sole Glu residue. For this specific residue, the \( pK_a \) of
the side-chain carboxylic acid is $pK_a=4.3$. Suppose that the protein is dissolved in a buffered solution, with pH carefully adjusted to pH=4.3 ($pH=pK_a$). Under these conditions, the Glu side chain appears as a mixture of protonated (glutamic acid, $-COOH$) and deprotonated (glutamate, $-COO^-$) species. What is the proportion of these two species?

$$K_a = \frac{[H^+][A^-]}{[HA]}.$$ 

If $pH=pK_a$, then $[H^+]=K_a$ and, therefore, $[A^-]=[HA]$. Thus, the proportion of protonated and deprotonated species is 50%-50%. (This is essentially the definition of $pK_a$).

10. A series of $i=1, 2, ..., N$ samples with variable concentration of ferrocene were used to measure optical absorption. The resulting experimental data (concentrations $X_i$, absorbances $Y_i$) were fitted, in a least-square sense, with the equation $y = x + \alpha$. The procedure involved a single fitting parameter, $\alpha$ (intercept). Please, derive the expression for $\alpha$ ensuring the best fit quality.

$$\chi^2 = \sum_{i=1}^{N} (Y_i - (X_i + \alpha))^2.$$ Now find a minimum of this expression with respect to $\alpha$. Take the derivative with respect to $\alpha$ and set it to zero, $\sum_{i=1}^{N} (Y_i - (X_i + \alpha)) = 0$. This leads to

$$\alpha = \frac{\sum_{i=1}^{N} (Y_i - X_i)}{N}.$$
Post-translational modifications are often involved in the mechanism of diseases.

1. Why would that be, i.e. how could post-translational modifications play a role in a disease progression in general?

**Answer:** Post-translational modifications (PTMs) are involved in many cellular processes, among these being intra- and extra-cellular signaling, structural organization, cellular recognition, oxidative stress, and a variety of allosteric phenomena. Failure to achieve a particular PTM, an excessive amount of a PTM, or aberrations in the PTM as in glycosylation can be involved in the mechanism of a disease.

2. Describe two diseases (other than the example given below) in which a change in the degree of post-translational modification (PTM) plays a role and explain how the PTM contributes to disease mechanism.

**Answer:** Multiple students referred to oxidative stress (OS) diseases. OS diseases are the result of a dramatic increase of reactive oxygen species (ROS) in a cell. ROS alter proteins at many levels including sulfhydryl group oxidation, nitration of tyrosine, cleavage along the primary structure, amino acid side chain oxidation involving lysine, arginine, proline, threonine, and histidine, fatty acid oxidative cleavage and alkylation of proteins with the resulting HNE, and conformation modifications that lead to protein aggregation. OS triggers many neurological diseases. Other students noted the role of aberrations in glycosylation as playing a role in cancer progression, particularly metastasis.

3. How would you recognize a PTM in a protein?

**Answer:** One way would be by noting differences in the separation properties in isoelectric focusing (IEF) of liquid chromatography (LC). This would work in many cases except when a PTM has no charge difference. In this case neither IEF or LC would work very well. Another way would be by matrix assisted laser desorption ionization (MALDI) mass spectrometry. If the molecular weight of the protein is higher or lower than expected it is likely there has been a PTM. Unfortunately some proteins do not ionize well and it is difficult to get a MALDI. Yet another way is to look at the tryptic digest of a protein. Reversed phase chromatography of the digest along with mass spectral analysis would show peptides bearing a PTM to behave in a different many than those without the PTM. Again the difference in mass would give a clue about the nature of the PTM. Still another way is through affinity selectors that recognize a particular type of PTM. Immobilized metal affinity chromatography recognizes phosphorylation. Antibodies recognized nitration, phosphorylation, HNE, methylation, glycosylation, and many more.

4. PTM driven diseases involve changes in the concentration a PTM in a number of ways. One is with an increase in the relative amount of the PTM
modified form of a protein while the total amount of all forms of the protein remains constant. In another, the ratio of the PTM and total amount of parent protein remain the same while the amount of both proteins increases. Describe a quantification procedure that would differentiate between these two cases.

**Answer:** You would need a method that quantified both forms of the protein, i.e. the form lacking the PTM and the PTM form. One way to do this is with an antibody that targets an epitope present on both forms and another that targets the PTM alone. A second way would be mass spectrometry in which a tryptic peptide common to both forms is quantified in a digest of both proteins and the PTM modified peptide is quantified independently. You could also quantify these peptides with either SDS-PAGE in a tightly cross-linked gel or by reversed phase chromatography or hydrophilic interaction chromatography.

5. Splice variant forms of proteins in which a PTM is part of the variable region often play a role in cancer. This is particularly true of glycoproteins on the outside of tumor cells. Explain how such splice variant glycoproteins might impact metastasis.

**Answer:** Changing the primary structure of a protein through a splice variations could cause multiple effects. One would be a change in protein conformation that could in turn change the biological activity of a protein. Because the splice variant region contained a PTM, the total number of modifications in the protein would be altered. This too could have a large impact on protein activity.
Answers

INORGANIC CHEMISTRY CUM EXAM 4/20/08

1. (a) IrCl₆³⁻ is extremely slow to undergo substitution reactions. Ir(III) has a 5d⁶ electronic configuration and requires a large ligand field activation energy to undergo substitution reactions. (This is similar to the behavior of Co(III)) complexes which also have a 3d⁶ low-spin electronic configuration.

(b) The outer-sphere oxidation of IrCl₆³⁻ by O₃ gives IrCl₄²⁻ and O₂⁻ as immediate products. (The O₃⁻ reacts with H₂O to give HO₃ that in turn forms O₂ + OH⁻. The OH radical is a very strong oxidizing agent.) These reactants can form more IrCl₆³⁻.

(c) The second-order rate constant of 1.7 x 10⁴ M⁻¹s⁻¹ for the O₃ reaction with IrCl₆³⁻ indicates that the reactions are fast (i.e. If IrCl₆³⁻ is millimolar, k₂₁₂ = 1.7 x 10⁴ M⁻¹s⁻¹ × 10⁻³ M = 170 M⁻¹ and t¹/₂ = 0.693/170 = 0.011 s or 41 milliseconds). Therefore, stopped-flow mixing is required with observation times of milliseconds to seconds.

(d) * IrCl₄²⁻ + IrCl₆³⁻ → IrCl₆³⁻ + IrCl₄²⁻  E° = 0.867

Self-exchange rxn.  k₁₁ = 2.3 x 10⁵ M⁻¹s⁻¹

(e) O₃ + e⁻ → O₂⁻  E° = 1.02 V

O₃ + O₂⁻ $\overset{\text{k}_2}{\longrightarrow}$ O₅⁻  (need to calculate)  K₂ = 10²⁻⁵⁹

IrCl₆³⁻ + O₃ $\overset{\text{k}_{12}}{\longrightarrow}$ IrCl₄²⁻ + O₂⁻  1.7 x 10⁴ = (2.3 x 10⁵ k₁₂ T₁₁)¹⁄₄  k₁₂ = 1.7 x 10⁴ M⁻¹s⁻¹

k₁₂ = (k₁₀ k₂₂ k₁₁ T₁₁)¹⁄₄  k₁₀ = 1.7 x 10⁴  k₁₁ = 2.3 x 10⁵ M⁻¹s⁻¹

k₂₂ = 2.39 x 10⁸  2.30 x 10⁵ x 10⁻⁵⁹ = 3.23 M⁻¹s⁻¹
(a) Relative substitution reactivity

\[
\begin{align*}
\text{Cr}^{3+} & > \text{V}^{2+} > \text{V}^{3+} > \text{Cr}^{3+} \\
\text{Cr}^{3+} & > \text{V}^{2+} > \text{V}^{3+} > \text{Cr}^{3+} \\
& \text{most inert}
\end{align*}
\]

Most stable

(b) Factors that affect $K_{\text{M-H}_2\text{O}}$

1. The ligand-field stabilization energy is large for $d^4$ species and $d^3$ ions have much greater metal ion - water dipole attraction than $d^2$ ions.

Thus, $\text{Cr}^{3+}$ is more inert as $\text{V}^{2+}$ is next.

2. $d^4$ metal ions have Jahn-Teller distortions that weaken some $\text{Cr}^{3+}$-$\text{OH}_2^-$ bonds relative to others, increasing some $\text{M-OH}_2$ bond distances.

(c) The factors in (b) greatly affect the $K_{\text{M-H}_2\text{O}}$ values because of the number of $3d$ electrons greatly influence the ligand-field stabilization with $d^3$ causing $\text{Cr}^{3+}$ as $\text{V}^{2+}$ to be most stable relative to metal-\text{OH}_2 bonding.

On the other hand, the Jahn-Teller distortions for $d^4$ systems plus the lower metal ion charge weakens the metal-\text{OH}_2 bond and gives faster substitution reactions.
3. (a) **Boltzmann Distribution Law**

\[
\frac{N_i}{N_{total}} = \frac{e^{-\frac{\varepsilon_i}{k_B T}}}{\sum_i e^{-\frac{\varepsilon_i}{k_B T}}} \\
N_i = \text{no. of molecules with energy } \varepsilon_i \text{ at temp. } T \\
N_{total} = \text{total no. of molecules with energy } \varepsilon
\]

The relative population of energy states depends on \( T \) and the difference in the levels of the energy states.

(b) How are \( K^\pm, \Delta H^\pm, \) and \( \Delta S^\pm \) related to each other?

1. \( \Delta G^\pm = \Delta H^\pm - T \Delta S^\pm = -RT \ln K^\pm \)

   These are transition-state functions or activated state functions.

2. How do these terms differ from \( K, \Delta G^0, \) and \( \Delta S^0 \), which are thermodynamic functions as opposed to transition-state functions? \( K \) is the equilibrium constant for products/reagents, where the species are at energy minima. \( \Delta G^0 \) is the thermodynamic free energy change for a reaction, where \( \Delta G^0 = \Delta H^0 - T \Delta S^0 \). Whereas \( \Delta H^\pm \) and \( \Delta S^\pm \) are for the transition state species at energy maxima as opposed to \( \Delta H^0 \) and \( \Delta S^0 \) for energy minima.

3. \( k = \frac{K^\pm}{h} e^{\frac{\Delta S^\pm}{R}} e^{-\frac{\Delta H^\pm}{RT}} \)

   where \( k = \text{rate constant} \)
   \( K^\pm = \text{Boltzmann constant} \)
   \( h = \text{Planck's constant} \)
   \( R = \text{gas constant} \)
   \( T = \text{absolute temp.} \)

**\( \Delta S^\pm = \text{activation entropy} \)
\( \Delta H^\pm = \text{activation enthalpy} \)**
No Organic crib available
April 26, 2008
Written by Professor Ghosh
\[ H = -\frac{1}{2} \partial_x^2 - \frac{1}{4} \partial_x^2 - \frac{1}{k_1} - \frac{1}{k_2} + \frac{1}{k_2} \]

\[ \eta = 15(1) 15(2) \left[ \alpha(1) \beta(2) - \alpha(2) \beta(1) \right] \frac{1}{\sqrt{2}} \]

\[ E(R \to \infty) = \frac{\hbar^2}{2m} \left( \frac{1}{k_1} - \frac{1}{k_2} \right) \]

\[ E(R = R_0) = \frac{\hbar^2}{2mL^2} \]

\[ E = E(R = R_0) - E(R \to \infty) \]

\[ \text{or} \]

\[ H_2^+ \rightarrow H + H^+ \]

\[ m \gg m_e \quad E(R) \quad \eta(\beta; R) \]

\[ 40 \]

\[ 20 \]

\[ 3 \]

\[ 4 \]

\[ 5 \]