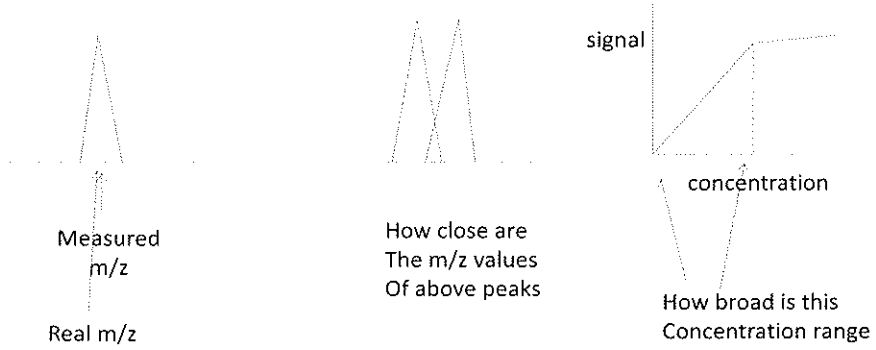
Analytical Cume March 2011 crib

Each question is worth 10 pts.

1. Crude oil and most of its products are complex mixtures of large saturated hydrocarbons. (3 pts) The knowledge on their composition is needed to determine the proper use for each mixture, and/or proper treatment method. (7 pts)
2. MW distribution, elemental composition, structures (3.3 pts each)
3. Difficult to evaporate intact (3.3 pts), difficult to ionize due to lack of acidic and basic functionalities (3.3 pts), when ionization succeeds they often fragment which prevents obtaining MW information, very complex mixtures complicate everything, difficult to determine structures of saturated hydrocarbons (3.3 pts for any one of these answers)
4. FD/FI: gentle evaporation and ionization, can be used to ionize saturated hydrocarbons; LIAD/chemical ionization: same; APCI/charge transfer: same (10 pts for any of these answers)
5. FD/FI followed by CAD; LIAD/chemical ionization followed by CAD; APCI/charge transfer followed by CAD; LIAD/EI; any of these followed by ion-molecule reactions; (10 pts for any two of these answers)
6. Internal standard (5 pts); for unknown mixtures, no good standard may be found; the complexity of the mixtures makes quantitation very difficult; evaporation and ionization efficiencies can be different for different compounds; different ions have different activation energies for dissociation and some may not produce any stable pseudo-molecular ions (5 pts if at least two of these answers are provided)
7. Mass accuracy: how close is your measured m/z value to the true value; mass resolution: how well can you separate ions with close m/z values; linear dynamic range: how broad concentration range can be measured in such a manner that signal correlates linearly with concentration (5 pts)




(5 pts for drawings)

8. High resolution would allow separation of all the components. Without this, for example the measurement of the MW distribution is impossible. High mass accuracy would allow the determination of the elemental compositions of all components. Wide linear dynamic range would allow quantitation of minor as well as major components. (3.3 pts/answer)

9. FT-ICR via ions' cyclotron frequency measurements; TOF via ions' flight time measurements (10 pts for either answer)
10. Space charge (10 pts)

**CUMULATIVE EXAMINATION IN BIOCHEMISTRY** Crib
March, 2011

## Protein folding, unfolding, misfolding, and disorder

It is expected that the answer should contain 1-4 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.

1. A single-residue mutation has been made in the hydrophobic core of a globular protein, Leu → Ala. Predict the most likely effect of this mutation on protein stability. Justify.

First, the unfolded state becomes more favorable. Indeed, there is a large ΔG penalty for solvating Leu, but much smaller penalty for solvating a compact Ala carrying a single methyl group. Second, the folded state becomes less favorable because a cavity is formed (small Ala in place of large Leu leaves a big void). Hence some of the favorable vdW interactions are lost. Let's summarize. Unfolded state becomes more favorable, folded state becomes less favorable – hence, loss of stability as a result of mutation. See Science 255, p. 178.

2. Amyloid plaques are the signature of neuropathological diseases, such as Alzheimer and Parkinson disease. The structure of the amyloid is quite regular, a cross-beta sheet. Yet, there are essentially no high-resolution x-ray crystallographic structures of amyloids. Why?

There is a short-range spatial order in amyloid, but no long-range order which is a necessary condition for high-resolution x-ray studies (i.e. the structure of the cross-beta sheet twists, turns, etc.). In other words, amyloid plaque is not a proper crystal and thus not suitable for crystallography.

3. What is "molten globule"?

A state of the protein where the native-like secondary structure is largely (but not fully) preserved, and the tertiary structure is largely (but not fully) lost. It is distinct from both native state and denatured state.

4. By analyzing the protein primary sequence, one can make a good prediction on whether this particular chain is going to be folded or, alternatively, disordered under the physiological conditions. For instance, a certain sequence with unusually high content of Glu residues is likely to be disordered. Please, explain why.

Electrostatic repulsion between the negatively-charged Glu side chains will tear the structure apart and make the protein unfold. Some people argued that prevalence of Glu means a deficiency of hydrophobic amino acids. This is not necessarily so, but this answer was also counted.

5. Binding of a (disordered) peptide to a (globular) protein typically occurs with association rate $k_{on}$~$10^6$ $M^{-1}s^{-1}$. Sometimes, however, extremely high rates, $k_{on}$~$10^9$ $M^{-1}s^{-1}$, are observed. This exceeds the estimates based on the model of diffusion-controlled binding. What is the source of these very high on-rates?

Electrostatic steering. Electrostatic forces are long-range, they effectively bring the peptide and its target together.

6. Assume that you have successfully prepared and purified a new protein sample (1 ml of protein solution with 100 μM concentration). The first thing that you want to determine is whether the protein is folded or not. How would you answer this question (suggest suitable experimental methods)?

Circular dichroism produces a nice signature of the secondary structure, and lack thereof in the case of the disordered protein. NMR chemical shifts also tell you the story (already $^1$H one-dimensional spectrum is sufficient). If chemical shifts are nicely dispersed, reflecting a variety of environments, this is a folded protein. Otherwise, if chemical shifts are all alike, this represents a random-coil-like protein. Other creative solutions such as fluorescence quenching are also possible.

7. What is prion and how does it cause disease (e.g. mad cow disease)?

Proteinaceous particle containing many copies of the misfolded PrP protein that can trigger avalanche-like misfolding of the normal PrP, eventually resulting in formation of amyloids. The unusual thing about all of this (what differentiates it from more conventional diseases, such as Alzheimer) is that prions are infective.

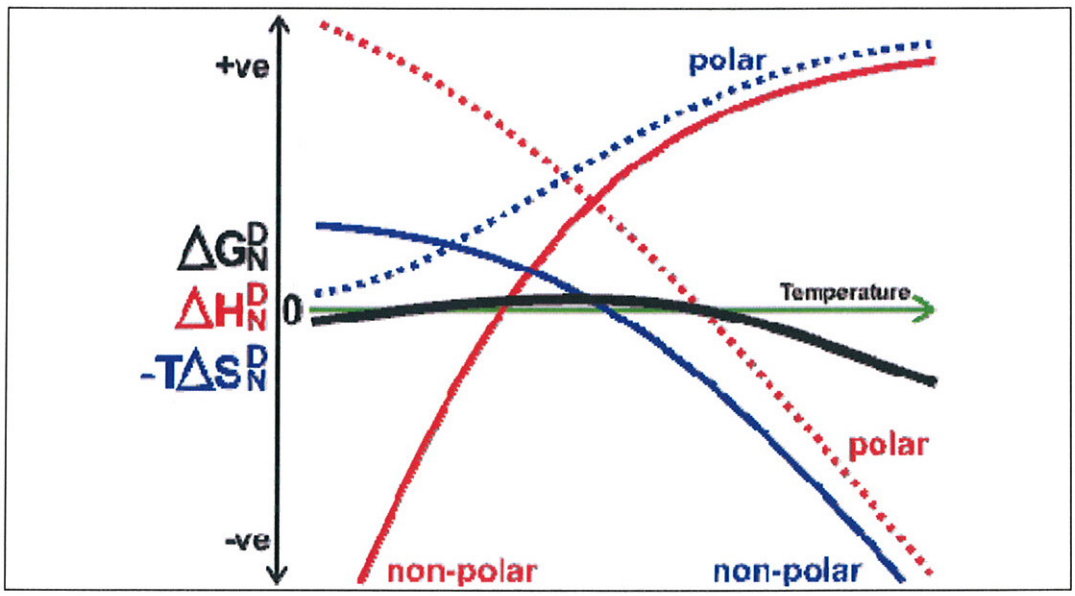
8. Radius of gyration of a protein can be approximately expressed as follows: $R_g \approx (1/\sqrt{6})bN^p$. Here $b$ is the so-called effective monomer length, $N$ is the number of residues in the protein, and $p$ is the exponent (a number). Please, estimate the (typical) value of $p$ for (*i*) folded proteins and (*ii*) denatured proteins. Please, explain and justify your choice of $p$.

For folded proteins it is simple. The volume of the protein is proportional to the number of residues, $V \sim N$. At the same time, the volume is proportional to radius cube, $V \sim R_g^3$. Hence, $R_g \sim N^{1/3}$ or, in other words, $p = 0.33$, which is in a very good

approximation correct.

For disordered proteins it is more complicated. Let us assume that disordered chain follows the random walk pattern, i.e. at each junction the polymer takes a random turn. If that's the case, then end-to-end distance should increase as $\sqrt{N}$. The radius of gyration changes in the same fashion as the end-to-end distance, and hence $R_g \sim \sqrt{N} = N^{1/2}$, and $p = 0.5$. This is a good estimate (would earn you a perfect score). In reality, however, the peptide chain cannot turn randomly at each junction, because of the excluded volume. As a result, in real life the protein is more expanded, $p \approx 0.6$.

9.




The scheme above (© Martin Chaplin) illustrates different contributions into free energy of the transition from the native (N) state to the denatured (D) state, $\Delta G_N^D = \Delta H_N^D - T\Delta S_N^D$ (there was a typo in this formula, but sort of trivial). According to this diagram, how does the folding status of the protein change as a function of temperature?

Protein is (mostly) denatured at low temperature, folded at the intermediate temperature, and denatured again at high temperature. The phenomenon of denaturation at low temperature (cold denaturation) is not frequently observed because the water usually freezes before it happens. There is no need to go beyond the black curve.

10. Using the diagram from Question 9, discuss the entropy term associated with non-polar side chains in the protein interior (continuous blue curve).

At low temperature, $T\Delta S_N^D(non-polar)$ favors folded protein over unfolded. This is well-known, and is actually one of the main forces behind protein folding. Indeed,

protein unfolding make water molecule 'freeze' around the hydrophobic group, which is entropically unfavorable – hence $T\Delta S_N^D(non-polar)$ favors the folded state. As the temperature increases, though, the situation is reversed and $T\Delta S_N^D(non-polar)$ starts favoring denatured conformation. Apparently, the water moves so fast, that it no longer becomes 'frozen' around the hydrophobic groups. On the other hand, the hydrophobic groups themselves become highly mobile in the denatured state, which means favorable entropy. On balance, at high temperature $T\Delta S_N^D(non-polar)$ favors the denatured state.

Note that we discuss here the <u>difference</u> in entropy between the native and denatured states.

11. Many human proteins <u>cannot</u> be successfully expressed in E.coli for the purpose of *in vitro* studies. For instance, this is the case for folate receptor $\beta$, 255 aa protein which is glycosylated, features 8 disulfide bonds and the GPI anchor. Why E.coli is not a suitable organism for expressing this protein*?

E.coli misses much of the necessary chaperone machinery and enzymes for proper post-translational modifications, e.g. disulfide isomerases to properly form the disulfide bonds, etc. The protein will never become properly folded if disulfides do not form in the right order. It will also suffer proteolytic damage without protection from chaperones at the folding stage, etc.

12. Consider a sample of a certain protein (purified and dissolved in the aqueous buffer); the protein is known to be folded at room temperature. As we raise the temperature, the protein becomes unfolded. After that we start lowering the temperature and eventually bring the system back to the room temperature. What will happen to the protein – will it become folded again or not?

It depends – simple small proteins become folded (i.e. unfolding is reversible), more complex proteins, especially those that require chaperones to fold properly, would not fold (they may oligomerize, aggregate, misfold, precipitate, etc).

* Insect cells with baculovirus expression system actually do the trick

# Inorganic Chem Crib March, 2011

A. $N(I): NO$ and $N(0): N_2$

B. $NH_4NO_3 \xrightarrow{\Delta} N_2O + 2H_2O$

at higher temperatures more exothermic formation of $N_2$ can occur explosively

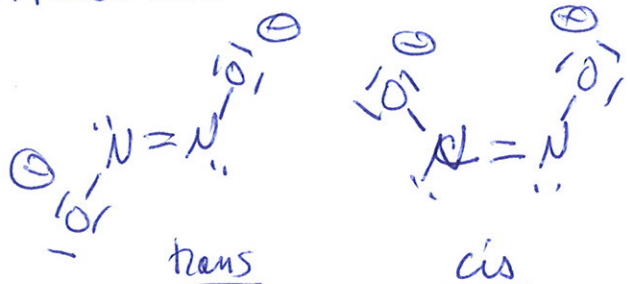
C. The HOMO's have $\pi$-bonding interactions between the 2 nitrogen atoms. (C+d). If the atoms are on the z axis the $2p_x$ and $2p_y$ orbitals are available for $\pi$ bonding and the HOMO + LUMO orbitals are degenerate.

D. See C. Also $|N\equiv N-\bar{\underline{O}}|$ resonance structure shows N-N bonding stronger than N-O bonding.

E. The IR transitions of $N_2O$ occur at lower frequencies and the $\leftarrow N-N-O\rightarrow$ stretch is electric dipole allowed.

F. Anesthetic

G.

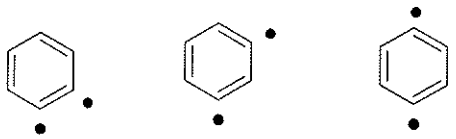


trans cis

**Organic Cumulative Examination** Crib **March 5, 2011** **Kenttämaa**

An article published in *JACS* in late 2010 (volume 132, pp. 17933-17944), titled "Indolyne Experimental and Computational Studies: Synthetic Applications and Origins of Selectivities of Nucleophilic Additions", and co-authored by Ken Houk and Neil Garg with their students at the University of California, Los Angeles, focuses on addition reactions to indolynes and other unsymmetrical *ortho*-benzyne analogs. This research led to the design of a substituted 4,5-indolyne that exhibits enhanced nucleophilic regioselectivity. In order to demonstrate your understanding on the fundamental properties of benzynes and nucleophilic additions, please answer the following questions related to this paper.

1) Draw proper Lewis structures for the *ortho-*, *meta-* and *para*-benzynes.



5 pts

2) These three biradicals have singlet electronic ground states. Explain what this means. Illustrate by an energy diagram showing electrons distributed on the formally singly occupied orbitals. Also show an energy diagram for a biradical with a triplet ground state.

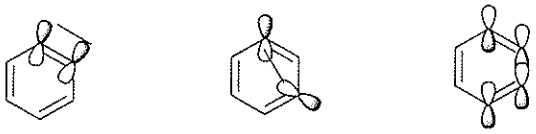
*In the ground state, all electrons are paired, including those on the formally singly occupied orbitals. 4 pts*

S T

3 pts 3 pts

3) Explain why each of the above three biradicals has a singlet electronic ground state. Illustrate your answer by showing how electrons in the formally singly occupied orbitals couple.

*The formally singly-occupied orbitals overlap through space for the ortho- and meta-benzynes, and couple via overlap with C-C $\sigma^*$ orbital for para-benzyne. 5 pts*



5 pts

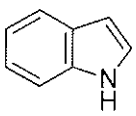
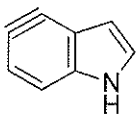
4) Define the singlet-triplet gap.

*The energy difference between the lowest energy singlet and triplet states. 5 pts*

5) Which of the above benzynes has the largest singlet-triplet gap and why?

*ortho-benzyne 3 pts due to direct overlap of the orbitals 2 pts*

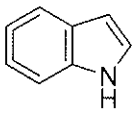
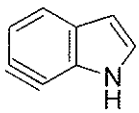
6) Indole and 4,5-indolyne are shown below. Which is which?

5 pts

*Indole* *4,5-indolyne*

7) Draw the structures of 5,6- and 6,7-indolynes.

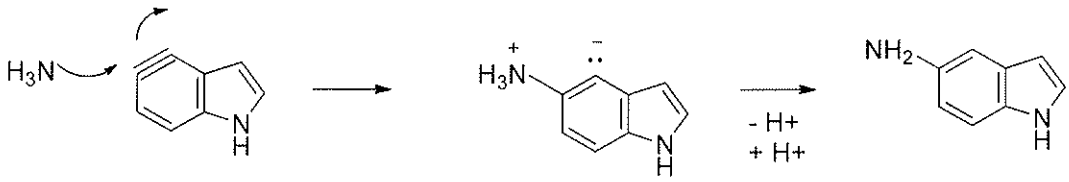



2.5 pts 2.5 pts

8) The indole heterocycle is usually exploited for its nucleophilic character. In contrast, the indolynes can act as electrophiles. Explain why.

*The ortho-benzyne moiety is an electrophile due to sp-carbons* 10 pts

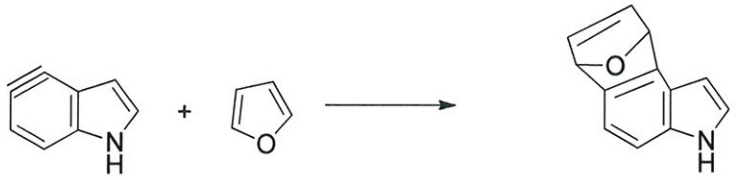
9) The electrophilic behavior of indolynes is of interest because it allows derivatization of the benzenoid ring as opposed to the pyrrole ring of indole. In order to illustrate how, draw a mechanism for the reaction of $NH_3$ with 4,5-indolyne and give the most likely products.




10 pts

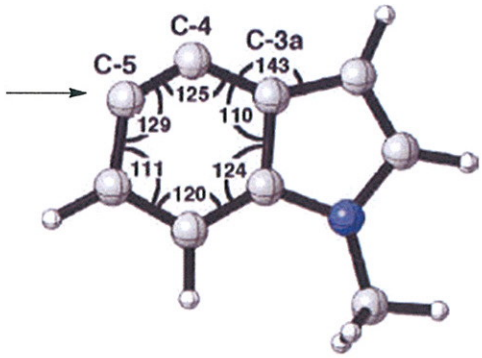
10) Give the most likely product for the Diels-Alder reaction of furan with 4,5-indolyne.

10 pts



11) Nucleophilic addition to 4,5-indolyne was found in this paper to be regioselective. The addition occurs more readily to the carbon with a more distorted angle. Considering the calculated structure of 4,5-indolyne shown below, which is the carbon atom predominantly attacked by nucleophiles?

5 pts




***4,5-indolyne (29)***

12) As opposed to *ortho*-benzyne analogs, what sort of reactions would you expect to dominate for *para*-benzyne analogs, and why?

*Radical reactions due to weak coupling between the biradical electrons. 10 pts.*

13) How about *meta*-benzyne analogs? Why?

*Nonradical reactions due to the strong coupling between the biradical electrons. 10 pts.*

No Physical crib available
March 5, 2011
Written by Professor Wasserman
Ph# 42348