Problem 1

Match the technique or phrase on the left with the correct statement on the right. There is one best answer for each item, and each is used only once.

1. dynamic light scattering
2. 8-anilino naphthalene sulfonic acid
3. differential scanning calorimetry
4. isoelectric focusing
5. reversed-phase high performance liquid chromatography
6. far ultraviolet circular dichroism
7. intrinsic fluorescence
8. near ultraviolet spectroscopic absorption
9. hydrophobic (apolar) interactions
10. static light scattering
11. Raman spectroscopy
12. size-exclusion chromatography (gel filtration)
13. hydrophobic interaction chromatography
14. protein near ultraviolet circular dichroism
15. isothermal titration calorimetry
16. extrinsic fluorescence
17. Amide I band
18. molten-globule state
19. sedimentation velocity analytical ultra centrifugation
20. sedimentation equilibrium analytical ultracentrifugation

a. elution induced by lowering salt concentration
b. can detect deamidation
c. usually due to tryptophan residues in proteins
d. increases upon melting of double-stranded nucleic acids
e. can determine stoichiometry and binding constant as well as enthalpy and entropy changes of a ligand/receptor interaction
f. can measure the absolute molecular weights of both large and small molecules
g. commonly employs the Amide III band to estimate protein secondary structure
h. based on Brownian motion
i. used in FTIR spectroscopy to measure protein secondary structure
j. used to detect apolar sites in macromolecules
k. measures heat capacity changes
l. can be monitored at fluorescence excitation wavelengths
m. based on the rate of movement of a particle in a centrifugal field
n. technique most frequently used to measure protein aggregation
o. employed to separate peptides during peptide mapping
p. intact secondary structure, dramatically reduced tertiary structure
q. characterized by large heat capacity changes
r. signals from aromatic side chains in optically asymmetric environments
s. uses binding of dyes to macromolecular systems
t. used to detect optical activity of protein helices and sheets

Answers, Problem 1
1. h, 2. j, 3. k, 4. b, 5. o, 6. t, 7. c, 8. d, 9. q, 10. l, 11. g, 12. n, 13. a, 14. r, 15. e, 16. s, 17. i, 18. p, 19. m, 20. f

Problem 2
Select a recombinant protein that is a currently marketed drug. A later edition of the Physicians Desk Reference (PDR) should be helpful in this regard. (a) Describe the mechanism of action of the protein and the rationale behind its use. (b) Describe the composition of the formulation of this protein based on the information provided in the PDF. Provide hypotheses about why which each component might be present in the formulation.

Answer, Problem 2
There is no specific or correct answer to this problem. It should, however, be responded to in terms of the information available in the PDR, at least one external source of information and the text material concerning formulation and the role of buffers, pH, stabilizers, etc.

Problem 3
Disulfide bonds have been reported to both increase and decrease the stability of proteins. Explain how such different effects might be induced by this common form of inter- and intra-chain crosslink.

Answer, Problem 3
It was initially thought that stabilization by disulfides was produced by lowering the entropy of the unfolded state. Recent explanations include stabilization of local or packing interactions (i.e. an ethalpic form of stabilization). Destabilization is thought to result by inhibiting conformational changes which would lower the free energy of the folded state. Effects on the dynamic properties (internal motions) of a protein by disulfide bonds could either increase or decrease stability depending on the relative effects on the folded and unfolded states. In other words, if the difference in energy between the two states is increased, the stability is increased and visa versa.

Problem 4
What are molten-globule (MG) states and why are they important? Explain how near and far ultraviolet circular dichroism (CD), infrared and Raman spectroscopy, intrinsic fluorescence, ANS (8-amilino naphthalene sulfonic acid) fluorescence and dynamic light scattering (DLS) can be used to establish the existence of such states.

Answer, Problem 4
MG states are conformationally heterogeneous states of proteins, which contain near native secondary structure, substantially disrupted tertiary structure and slightly expanded (e.g. 10%) sizes. They are important because of their tendency to aggregate and their possible role in protein folding and unfolding. Their maintenance of secondary structure can be detected by far UV CD, FTIR and Raman spectroscopies. Their disrupted tertiary structure can be seen through the use of intrinsic fluorescence, near UV CD and absorption spectroscopy as well as some Raman signals. Their expanded size as well as their propensity to aggregate can be detected by DLS. Such states also tend to bind dyes such as ANS.

Problem 5
Why does drying of macromolecules such as proteins and DNA often greatly improve their stability?

Answer, Problem 5
Removal of water can often stabilize macromolecules because water can be both a key reactant in chemical degradation reactions as well as be an essential component of their structure (i.e. "water of hydration"). A second reason is that removal of water can decrease the internal motions of macromolecules, which slows a wide variety of degradative processes.

Problem 6
What is a virus-like particle (VLP)? Why have VLPs proven to be such effective vaccines compared to monomeric recombinant proteins?

Answer, Problem 6
Virus-like particles are protein complexes formed by one or more surface proteins of viruses. They lack nucleic acids but retain the overall architecture of viral particles. They are thought to be highly immunogenic because they retain the multivalent nature of the viral surface (as opposed to their nomomeric recombinant protein counterparts). They also tend to be highly stable which facilitates their development as vaccines. Their effectiveness is perhaps not surprising when one remembers that the immune system has evolved to deal with similar structures, especially through the "toll" receptors.

Problem 7
Explain how DNA vaccines work.

Answer, Problem 7
DNA vaccines, which in their simplest form consist of plasmids coding antigenic proteins of interest, are thought to enter cells where they produce the encoded protein through a natural transcription/translation process. The plasmids are taken up by some type of endocytic process and escape from the resultant endosomes by an unknown mechanism, which may involve leakage or more extensive disruption of the endosomes. The dendritic cells and macrophages may be especially important in this regard, but other cells (such as muscle cells) may also be able to support their translation. Through mechanisms that are still unclear, the antigenic proteins may be secreted or peptides derived from them presented in the context of receptors on the surface of cells where normal immune responses ensue.

Problem 8

In nonviral gene delivery, a cationic polymer is often added to the DNA plasmid. What is the purpose of such positively charged polymers?

Answer, Problem 8
The surface of cells is negatively charged due to the presence of sulfated polysaccharides, sialic acid residues and negatively charged lipids. Thus, DNA would be expected to be

electrostatically repelled from the cell surface. When cationic polymers and lipids are complexed to DNA, the negative charge on the DNA can be neutralized with the complex even becoming positively charged. This allows such complexes to bind to the cell surface, triggering endocytosis of the complex. The DNA is also collapsed in size by charge neutralized which may facilitate its uptake by cells.

Problem 9
Perhaps surprisingly, even a single amino acid change in a large protein can have a dramatic effect on its solubility. Why?

Answer, Problem 9
Solubility is due to a balance between the interaction of a molecule with the solvent and the forces responsible for the formation of a solid phase. Although a single amino acid change might not be expected to produce a major change in the hydration of a large protein, such a change might well either disrupt one of a small number of the interprotein interactions responsible for the solid phase or produce one or a few novel interactions of this type. This could lead to large increases or decreases in a protein's solubility. In fact, such small effects on the solid phase appear to be responsible for a variety of human diseases such as sickle cell anemia in which a single amino acid change in human hemoglobin results in the formation of insoluble protein polymers that directly produce a serious pathological condition.

Problem 10
One of the major pharmaceutical problems seen with biopolymers is their tendency to form aggregates (particles). One form of this problem occurs when these particles are very small (subvisible). Describe three (3) methods that can be used to detect such particles and discuss the advantages and disadvantage to the use of each. You might wish to consult Carpenter, et.al. [J. Pharm. Sci, 98, 1201-5 (2009) and In Press, (2010)]

Answer, Problem 10
Techniques that can be used include the following:

(1) Size exclusion chromatography (gel filtration) (SEC). This is the most common method employed to detect subvisible particles. Its advantages include its quantitative nature (i.e. the amount of aggregate can be measured simultaneously with unaggregated species) and its usual ready availability and widespread acceptance. Disadvantages include the dilution which occurs during the chromatographic process. This can cause dissociation of aggregates since their formation is usually concentration dependent. Another consequence of dilution is that one is not examining the formulation under its exact storage conditions. A second problem is that proteins can sometimes interact with the chromatography matrix. Thus, the quantitative nature as well as the ability of the method to estimate molecular weight can be compromised. This defect can often be eliminated or minimized by the incorporation of an agent in the SEC buffer such as guanidine HCL, urea or high salt concentrations, which inhibits a protein's interaction with the matrix.

(2) Static Light Scattering. Light scattering is very sensitive to the size of the scattering particles. Thus, it is an excellent method for the detection of the presence of aggregates but

it is not generally quantitative since the signal measured is due to scattering from all of the particles present. An advantage, however, is that measurements can be made with a fluorometer or in the form of turbidity with a simple UV/visible absorption spectrometer. Furthermore, light scattering detectors can be used with chromatography system (especially SEC). This permits the analysis of separated species to be obtained and their sizes and molecular weight determined.

(3) Dynamic Light Scattering (DLS). Like static light scattering, this technique is very sensitive to the presence of aggregates. While it is not yet considered sufficiently quantitative to determine the amount of aggregates present, deconvolution of the autocorrelation functions used to estimate size does to a limited extent permit the analysis of the diameters of heterogeneous mixtures of particles that differ in diameter by at least a factor of two. Instruments for DLS measurements have become increasingly common and are now available in a high throughput format.

(4) Analytical Ultracentrifugation (AUC). Sedimentation velocity AUC can provide a very quantitative description of protein aggregates in terms of both their size and amount. It is however, only functional over a limited concentration range and has a very low throughput. It has, therefore, primarily been used as a research tool in pharmaceutical analysis. The recent availability of fluorescence detection (in addition to the more common absorption technique) somewhat increases its utility. Equilibrium AUC has the advantage that it can provide a quantitative description of association and dissociation phenomenon but is not generally useful for the type of more extensive aggregation seen in pharmaceutical protein formulations.

(5) FTIR spectroscopy can detect aggregates due to the presence of unique intermolecular beta-structure peaks near 1615 and 1692 cm^{-1} but has yet to be developed as a quantitative technique.

A number of other techniques such as field flow fractionation are under development as tools to detect aggregation but have not yet been sufficient employed to evaluate their general utility.

Problem 11

Describe the procedure known as “peptide mapping” and how it can be used to characterize the chemical degradation of proteins. If after treating with a protease such as trypsin, the coverage of the amino acid sequence is less than 100%, what might you do to improve the coverage?

Answer, Problem 11

In peptide mapping, a protein is cleaved by a protease into peptides of defined size and sequence. The peptides are separated by reversed-phase HPLC and their molecular weights are determined to a high degree of accuracy with mass spectrometry. If a chemical change occurs in a specific residue, the molecular weight of the peptide containing that amino acid (or acids) as well as its chromatographic behavior will be altered and consequently identified. The actual sequence of the peptide can also be determined by modern methods of mass spectrometry. If coverage of the amino acid sequence is incomplete, it may be possible to improve the resolution by alteration of

the chromatographic conditions. In addition, a second protease with different specificity (i.e. it cleaves the protein at different positions that the protease initially used) can be used to create a different series of peptides. A combination of the two "peptide maps" will usually dramatically increase coverage of the protein's sequence.

Problem 12
It is often said that biopharmaceuticals such as recombinant proteins are defined by the process used to produce them. Explain this statement. How does this idea impact the controversy over the production of "follow-on" or "biosimilar" protein pharmaceuticals?

Answer, Problem 12
When recombinant proteins are produced in different types of cells or purified in different ways, the final product may differ even if the same DNA sequence is present. This can be due to different forms of chemical modification (e.g. glycosylation) produced by different cell types as well as different trace impurities remaining after isolation of the product. These differences may be quite subtle but still influence the efficacy, toxicity and pharmacokinetics of proteins produced by different types of cells and isolation procedures. As companies produce follow-on biologicals, they hope that a thorough physical chemical and biological characterization of their biosimilar products will reduce or eliminate the need for extensive and expensive clinical studies. The problem with this approach is the lack of complete identity to the original ("innovator") products for the reasons stated above.

Problem 13
Why might proteins gradually change their structure (conformation) when they bind to surfaces such as containers or aluminum salt adjuvants?

Answer, Problem 13
When proteins bind to surfaces, they initially do so through the formation of a small number of contacts with the surface. With time, however, the protein gradually increases its number of contacts with the surface, lowering its free energy. This, in turn, often leads to structural changes or a decrease in stability. A variety of different types of interactions (e.g. electrostatic, apolar, van der Waals, etc.) may be responsible for such phenomena.

Problem 14
A protein displays the typical characteristics of the circular dichroism (CD) of an alpha-helical rich protein with a negative ellipticity double minimum at 208 and 222 nm. When the protein is heated or its concentration is increased (with the spectrum normalized for concentration), the spectrum is reduced in intensity and shifts to higher wavelength. What might be happening?

Answer, Problem 14
When the far UV CD spectrum of a protein changes, this is usually taken as evidence of a change in its secondary structure content (i.e., a conformational change). Although this might be expected to happen with temperature, it is usually independent of protein concentration. One exception is when the association state of the protein is changing with temperature and/or protein concentration. A more likely explanation, however, is that the protein is undergoing significant

aggregation. This often produces the spectral artifact known as “absorption flattening” in which particles “shadow” one another producing a loss in intensity and red shift in the spectrum.

Problem 15
Somewhat surprisingly, when some single-tryptophan containing proteins are excited at 280 nm, they fail to show typical tryptophan containing fluorescence (i.e., a broad emission peak between 320 and 355 nm). Instead, they display what appears to be tyrosine fluorescence near 303 nm. When the excitation wavelength is raised to 300 nm, however, or the protein is unfolded by high concentrations of a reagent such as urea or guanidine hydrochloride, typical tryptophan fluorescence is now seen. What might be occurring?

Answer, Problem 15
It has been found in a number of cases that the fluorescence of tryptophan in single Trp containing proteins is quenched by a neighboring residue (most often one containing a positive charge such as His, Arg or Lys). This permits its weaker fluorescence from Tyr residues, which is normally hidden by the strong fluorescence of Trp, to be seen. When excitation is moved to 300 nm, the loss of Tyr fluorescence now permits the weak Trp fluorescence to be seen. When the protein is unfolded, the indole side chain of Trp is moved away from the quenching side chain allowing stronger Trp fluorescence to be produced.

Problem 16
The initial stability studies of biopharmaceuticals are often conducted at temperatures far above intended storage conditions (e.g. 20-60°C vs. 4°C). Such “accelerated stability studies” are sometimes criticized for their potential inapplicability to the expected real world refrigerator storage. Defend the use of such higher temperature studies.

Answer, Problem 16
Since the desired storage stability of biopharmaceuticals is usually 2-3 years at 4°C, the development of stable formulations would take an inordinate amount of time if one were forced to use low temperature long-term studies for their creation. By using elevated temperature, it is hoped that the rates of both physical and chemical degradation processes can be increased to the point that they can be quickly identified. Using this information, one can then find conditions and agents (excipients) that slow these processes sufficiently that they will produce stable formulations over much longer periods at a lower temperature. Although this approach does not always work since events not detected under accelerated conditions can potentially occur over very long times, it has generally proven to be a reliable approach.

Problem 17
Upon isoelectric focusing, monoclonal antibodies (immunoglobulins) typically display a series of closely spaced bands. How might you determine the origin of this heterogeneity? What is the origin?

Answer, Problem 17
One needs to isolate each peak and determine its covalent structure (amino acid sequence). Isolation is probably best done by ion-exchange chromatography since the differences among the different species is presumably one of change. Preparative isoelectric focusing could also be

used. Peptide mapping could then be used to determine the differences among the isolated proteins. Since only one or a few peptides is expected to be altered, the sequencing necessary should be quite limited. The most common source of such chemical changes is deamidation although differential sialyation is also sometimes observed. The presence of sialyation can also be established by the use of the enzyme neuraminidase while changes in deamidated species seen by induction of further deamidation through the use of high temperature and increased pH.

Problem 18

A recent new approach to the development of protein-based pharmaceuticals and vaccines is based on the creation of "fusion" proteins in which two different proteins or protein domains (fragments) are joined together at the DNA level to create new proteins. Why would one want to create such fusion proteins? Give some actual examples of such proteins which you either find in the scientific literature or are your own idea.

Answer, Problem 18

The idea is to create multi-functionality within a single macromolecule. This is best illustrated with a couple of examples. Recently a number of proteins have been fused to human serum albumin (HSA). Because the serum half-life of HSA is so long, this will extend the life of any fused partner and potentially extend its period of action. A variety of different proteins and peptides have been used to extend protein half lives. Can you name some others? Another example is the fusion of a portion or entire antibody to a second protein. The binding specificity of the antibody combining site can be used to either target the fusion partner to a particular site in the body defined by the Igs specificity or to add a second function. New applications of this type are arising with some frequency.

Problem 19

A protein containing the 3 common ultraviolet absorbing chromophores is exposed to elevated temperature and its absorbance is monitored at 400 nm. No change in the signal is seen at 400 nm for 5 minutes, but the apparent absorbance begins to rapidly increase after this. (a) What are the three chromophores? (b) What is actually being measured at 400 nm? (c) Explain the presence of the lag period before the signal begins to increase.

Answer, Problem 19

(a) Tryptophan, tyrosine and phenylalanine.

(b) Because the 3 chromophores do <u>not</u> absorb above 310 nm, the signal must be coming from light scattering. It is therefore better referred to as optical density or turbidity.

(c) The lag period is due to the need for a nucleation event to facilitate the temperature-induced aggregation of the protein that is responsible for the observed light scattering.

Problem 20

When subjected to differential scanning calorimetry, an RNA molecule is found to produce 4 distinct transitions which we will designate A, B, C and D with Tm (melting temperature) values of 45, 53, 72 and 82°C, respectively. Theoretical calculations reveal 4 hydrogen bonded based-paired regions with the following compositions: 6 A/U, 2 G/C; 7G/C, 1 A/U; 3 A/U, 3 G/C and 2 GC/2 AU. Assume that the differences in stability (Tms) are due entirely to the difference in

hydrogen binding in the different regions. Assign the 4 different compositions to the regions associated with each Tm.

Answer, Problem 20
The critical fact is that A/U pairs have 2 hydrogen bonds while G/C have 3. We can then make the following simple calculation:

	# hydrogen bonds
6 A/U, 2 G/C	12 + 6 = 18
7 G/C, 1 A/U	21 + 2 = 23
3 A/U, 3 G/C	6 + 9 = 15
2 G/C, 2 A/U	6 + 4 = 10

Thus, based on the hypothesis that the number of H-bonds is roughly proportional to the Tms,

		Tm	region
2 G/C	2 A/U	45°C	A
3 A/U	3 G/C	53°C	B
6 A/U	2 G/C	72°C	D
7 G/C	1 A/U	82°C	C

This is not a bad approximation but differences in the interactions between the bases and sequence individualities require the use of more complex algorithms, which work quite well to estimate the Tms of oligonucleotides.

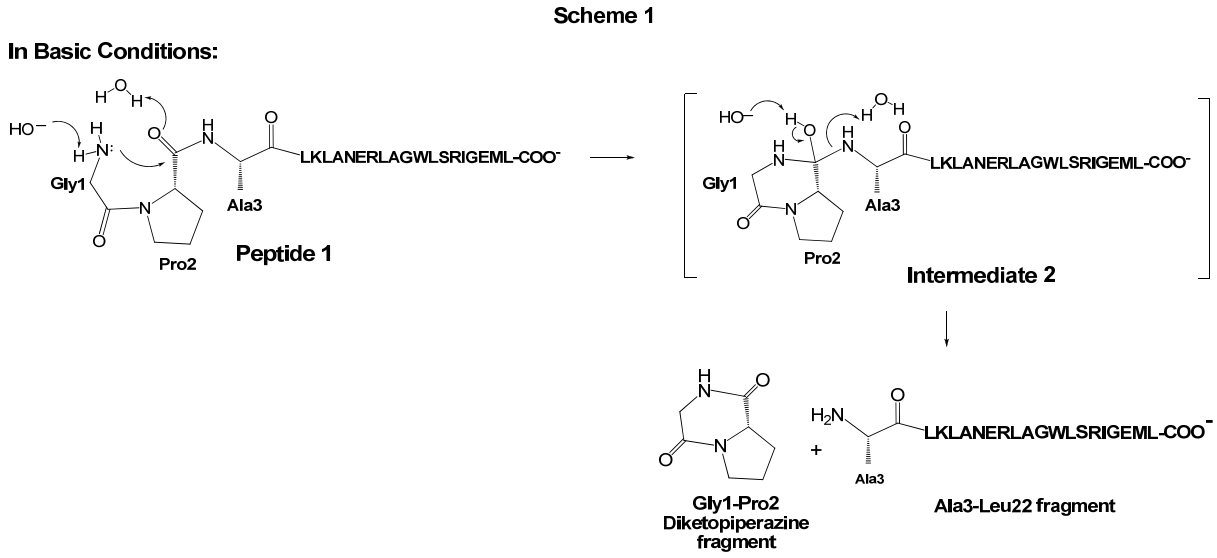
Problem 21-24
You are working to develop **Peptide 1** as a potential drug; the peptide primary sequence is shown below using the one letter code. Chemical stability of this peptide was evaluated using accelerated conditions (i.e., 40–70 °C) at various pH values (i.e., 2–11). The results of the chemical stability study suggest that several residues in **Peptide 1** are prone to degradation reactions (e.g., hydrolysis or oxidation).

Primary sequence: GPALKLANERLAGWLSRIGEML (Peptide 1)
Residue number 1 10 20

Problem 21
Under basic conditions, this molecule degrades to release a peptide fragment Ala3-Leu22. Propose a possible mechanism for the way that this fragmentation reaction can occur.

Answer, Problem 21
Under basic conditions, **Peptide 1** releases the A3-L22 fragment because there is peptide bond cleavage between the Pro2 and Ala3 residues (Scheme 1). The peptide bond between Pro2 and Ala3 is prone to cleavage because the proline residue has a unique five-membered ring structure that causes the amino group of Gly1 to be in close proximity to the carbonyl carbon of Pro2 (see Scheme 1). Therefore, the nitrogen atom of Gly1 (as a nucleophile) can readily attack the

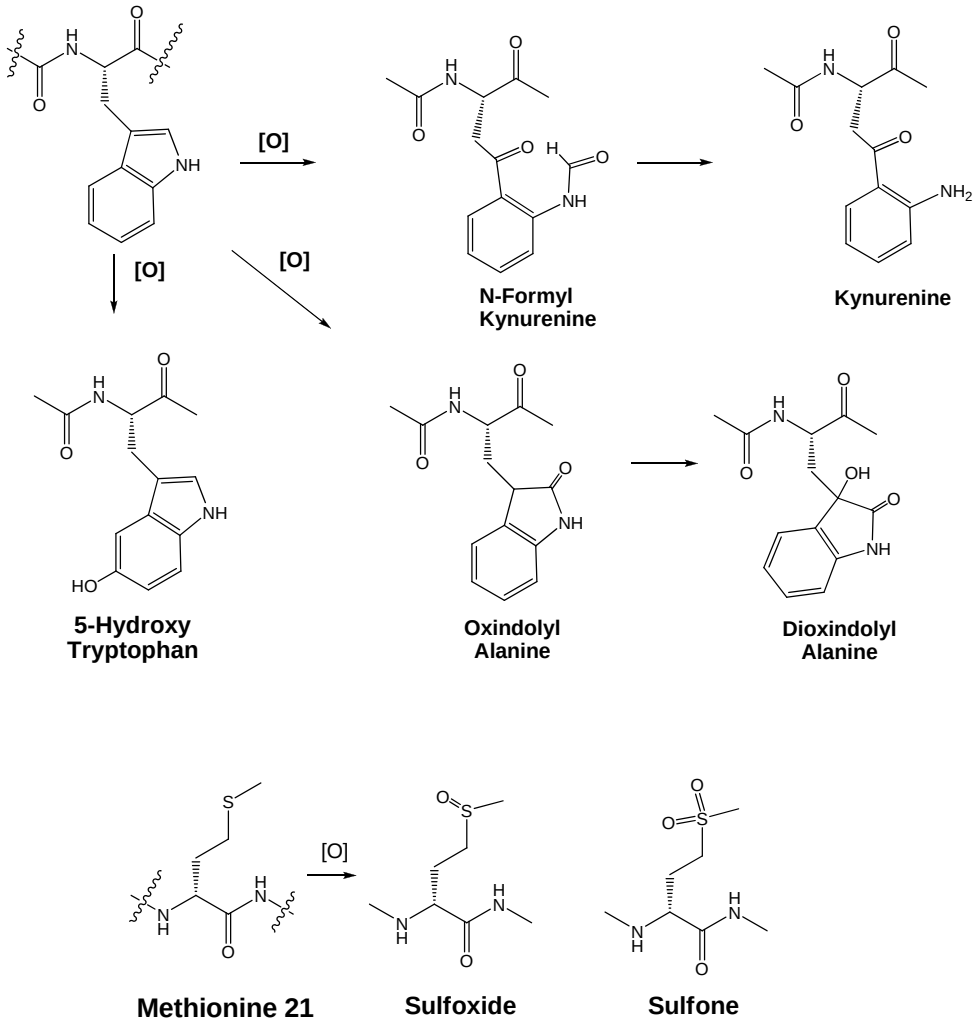
carbonyl carbon of Pro2 to produce **Intermediate 2**. A structural rearrangement of **Intermediate 2** generates two products (Gly1-Pro2 diketopiperazine and A3-L22 fragments) as a result of the peptide bond cleavage between Pro2 and Ala3.



Problem 22
Identify which residue(s) will undergo oxidation reaction. Draw the oxidation product(s).

Answer, Problem 22
The Trp14 and Met22 residues are susceptible to oxidation reaction in this peptide. This is due to the presence of reactive oxygen species in solution and the fact that the oxidation reaction is often catalyzed by the presence of metal ions (i.e., iron or copper). The oxidation reaction increases the molecular weight of the peptide, and this increase can be detected by mass spectrometry. The Trp14 residue can be oxidized to form various oxidation products (Scheme 2) including, *N*-formylkynurenine (NFK), kynurenine (Kyn), oxindolylalanine (Oia), dioxindolylalanine (DiOia), and 5-hydroxytryptophan (5-OH-Trp). The side chain of the Met21 residue can be oxidized to sulfoxide and sulfone (Scheme 2).

Scheme 2


[O]
[O]
[O]
N-Formyl Kynurenine
Kynurenine
5-Hydroxy Tryptophan
Oxindolyl Alanine
Dioxindolyl Alanine

[O]
Methionine 21
Sulfoxide
Sulfone

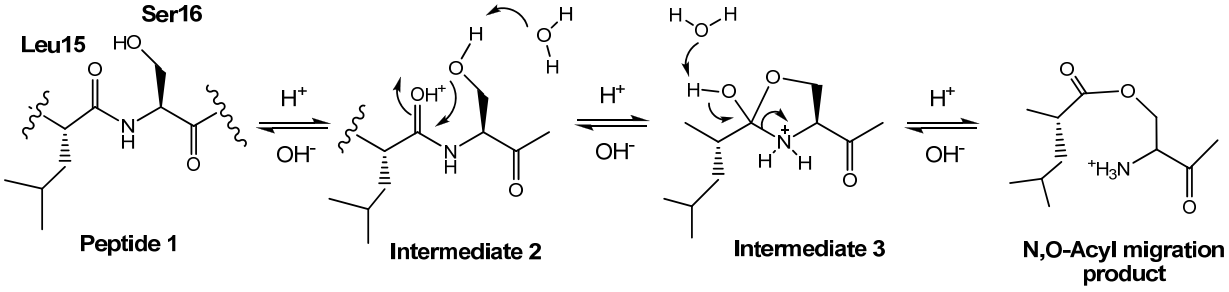
Problem 23

Under acidic conditions, there is a residue that can undergo N,O-acyl migration. Identify this residue and draw the mechanism of N,O-acyl migration. What would happen if the product of N,O-acyl migration were treated with a basic solution?

Answer, Problem 23

The Ser16 residue is susceptible to N,O-acyl migration in acidic conditions to produce a degradation product with the same molecular weight as the parent **Peptide 1**. The degradation product has an ester bond connecting the Leu15 and Ser16 residues that has different physicochemical properties than those of **Peptide 1**. For example, the degradation product has an additional positive charge at the amino group of Ser16 compared to the parent **Peptide 1**. Normally, this degradation product has different polarity than the parent **Peptide 1,** and its retention time in an analytical C18 HPLC column is different than that of **Peptide 1**. The N,O-acyl migration can proceed *via* two different possible mechanisms. One of these mechanisms is presented in Scheme 3, the other mechanism is presented elsewhere in this chapter. In the first mechanism, the carbonyl oxygen of Leu15 of **Peptide 1** is protonated to give **Intermediate 2**, which makes the carbonyl carbon of Leu15 more electrophilic. The attack on the carbonyl carbon of Leu15 by the oxygen of the Ser16 side-chain produces **Intermediate 3** with a five-membered ring structure. Rearrangement of the five-membered ring in **Intermediate 3** yields the N,O-acyl migration product with an ester group in the backbone between Leu15 and Ser16 residues instead of a peptide bond. Treatment of the N,O-acyl migration product in basic condition converts it back to **Peptide 1**.

Scheme 3



Problem 24

During stability studies at pH 6.0, this peptide degrades to release ammonia and give three different products. **Product 1** has a molecular weight 17 atomic mass units (amu) lower than that of **Peptide 1**. **Products 2** and **3** are one atomic unit different than the parent compound. Stability studies of **Product 2** under acidic conditions will produce fragmentation of this peptide. Identify each of these products and the mechanism of peptide degradation that leads to these products.

Answer, Problem 24

At pH 6.0, the peptide releases ammonia, indicating that the peptide undergoes the deamidation reaction. In this case, the Asn8 residue in **Peptide 1** is the unstable residue, which is the site for the deamidation reaction. **Product 1** is the cyclic imide intermediate (Asu8) derived from the deamidation reaction of the Asn8 residue. Hydrolysis of the cyclic imide (**Product 1**) produces degradation **Products 2** and **3**, which are the Asp and iso-Asp derivatives of **Peptide 1**, respectively. **Product 2**, which contains aspartic acid residue, can degrade to give fragmentation products. The first set of the fragmentation products contains Gly1-Asp8 and Glu9-Leu22 fragments; these products are the result of peptide bond hydrolysis between Asp8 and Glu9. The next set of fragmentation products are Gly1-Ala7 and Asp8-Leu22 peptides; these peptides are the result of the peptide bond hydrolysis between Ala7 and Asp8.

Scheme 4