Some equations and formulae that might or might not be useful:

$\Delta G = \Delta H - T\Delta S$	$K_w = [H^+][OH^-] = 10^{-14}\ M^2$
$\Delta G = \Delta G^0 + RT\ lnQ$	$pH = -\log[H^+]$
$\Delta G = \Delta G^0 + RT\ ln(\{[products]\}/\{[reactants]\})$	$K_a = [H^+][A^-]/[HA]$
$K_{eq} = \{[products]\}_{eq}/\{[reactants]\}_{eq}$	$pH = pK_a + \log([A^-]/[HA])$
$\Delta G^0 = -RT\ ln(K_{eq})$	$[A^-]/[HA] = 10^{pH-pKa}$
$K_{eq} = exp(-\Delta G^0/RT)$	$pH = pK_a + \log[x/(c_0 - x)]$
$R = 8.31\ J/K/mol = 0.00831\ kJ/K/mol$	(c_0 is the initial molar concentration of a weak acid, and x denotes molar equivalents of the conjugate base)
$T(K) = t(^oC) + 273.16^o$	

Table 1: Peptide bond cleavage reagents

Treatment	Cleavage points	Cleavage location at
Trypsin	Lys, Arg	carboxyl side of these residues
Chymotrypsin	Phe, Tyr, Trp	carboxyl side of these residues
Asp-N protease	Asp, Glu	amino side of these residues
Pepsin	Phe, Tyr, Trp	amino side of these residues
Cyanogen bromide	Met	carboxyl side of this residue

Problem 1. (12 points)

Consider a hypothetical reaction A + B $\rightleftarrows$ C which takes place at 25ºC. The initial molar concentrations of A, B and C were 60 mM, 40 mM, and 10 mM, respectively. When the molar concentration of C reached 30 mM, the ΔG for the forward reaction was -1.0 kJ/mol. Determine the value of ΔG^o and the equilibrium constant for this reaction.

$\Delta G = \Delta G^o + RT \ln(Q)$ ➔ $\Delta G^o = \Delta G - RT \ln(Q)$, where $Q = [C]/([A]\times[B])$.

Since [C] = 0.03 M at the time point of interest, i.e. increase by 0.02 M (20 mM) compared to the initial conditions, using the matter conservation law we get [A] = 0.06 M - 0.02 M = 0.04 M and [B] = 0.04 M – 0.02 M = 0.02 M at this time point. Substituting these values in the above equation (remember: you need to use the concentrations in M units!) gives:

$Q = 0.03\ M/(0.04\ M \times 0.02\ M) = 37.5\ M^{-1}$

$\Delta G^o = \Delta G - RT\times\ln Q = -1000\ J/mol - RT\times\ln(37.5) = -1000\ J/mol - 8977\ J/mol = -9977\ J/mol \approx -10\ kJ/mol$

Once you know the ΔG^o, calculating the equilibrium constant is straightforward, but don't forget that K_{eq} has units – it is not dimensionless for this reaction:

$K_{eq} = \exp(-\Delta G^o/RT) = 56.15\ M^{-1}$

Problem 2. (25 points) This problem is about glutamate as amino acid. Answer the following questions.

A. For what pH values can glutamate amino acid be used as an effective buffer? (6 points)

Amino acid Glu has 3 titratable groups: the amino group (NH_2, pK_2=9.7), the carboxyl group (COOH, pK_1=2.3) (assuming the "common" pKa values for the terminal groups) and the side-chain carboxyl group (COOH, pK_{Glu}=4.25). Following the convention that the buffering range is ±1 pH unit around the pKa value, these ranges are: 1.3 – 3.3; 3.25 – 5.25 (or combined 1.3 – 5.25 range), and 8.7 – 10.7. (if you used Glu-specific pK_1=2.19 and pK_2=9.67, then the respective ranges are 1.19 – 5.25 and 8.67 – 10.67)

B. What ***fraction*** of glutamate has its side chain in the deprotonated form at pH=5? *Explain your reasoning.*

(9 points)

Using the Henderson-Hasselbalch equation, we get $[E^-]/[E]=10^{pH-pKa}$, where "E" designates Glu with the protonated side-chain carboxylic group and "E^-" designates Glu with deprotonated side chain, and pKa = 4.25. At pH=5, this ratio (let's call it *r*) equals $r = [E^-]/[E] = 10^{0.75} = 5.6234$. The fraction of glutamate molecules with deprotonated side chain is $[E^-]/([E]+[E^-]) = r/(1+r) = 5.6234/(1+5.6234) = 0.849 \approx 0.85$.

C. You have a 250 mL sample of 20 mM solution of glutamate at the pH value that equals the pKa of the glutamate's side chain. You need to bring the pH of the sample to 5 (see previous question). You have at your disposal 1M HCl and 1M NaOH. Which of the two reagents and in what volume you need to add to your sample in order to achieve your goal? *Explain your reasoning and assumptions.* (10 points: 2 for NaOH, 8 for the calculation)

The range of pH values of interest here (from pH 4.25 to 5) is outside the good buffering range for the C-terminal carboxyl group (pH 1.3 – 3.3), thus we can focus only on the side chain carboxyl group of glutamate – this is the *assumption*.

250 mL of 20 mM solution of glutamate contain 0.25 L × 20 mM = 5 mmol of Glu. At the pH = pKa = 4.25, this amount is split equally into 2.5 mmol of base (deprotonated Glu, E^-) and 2.5 mmol of acid (protonated Glu). To increase the pH from 4.25 to 5.0 we need to add a base, hence NaOH.
Based on the answer to the previous question, at pH 5 we will have 0.15 × 5 mmol = 0.75 mmol of protonated Glu (acid, E) and 0.85 × 5 mmol = 4.25 mmol of deprotonated Glu (base, E^-). Thus to go from 2.5 mmol of E^- to 4.25 mmol we need to add 4.25 - 2.5 = 1.75 mmol of NaOH, which corresponds to 1.75 mL of 1M NaOH.

Note that because this added volume is much smaller than 250 mL, this calculation ignores an increase in the total volume (from 200 mL to 201.75 mL).

Problem 3. (20 points) Consider a hexapeptide with the following amino acid sequence: **YKDWTP**

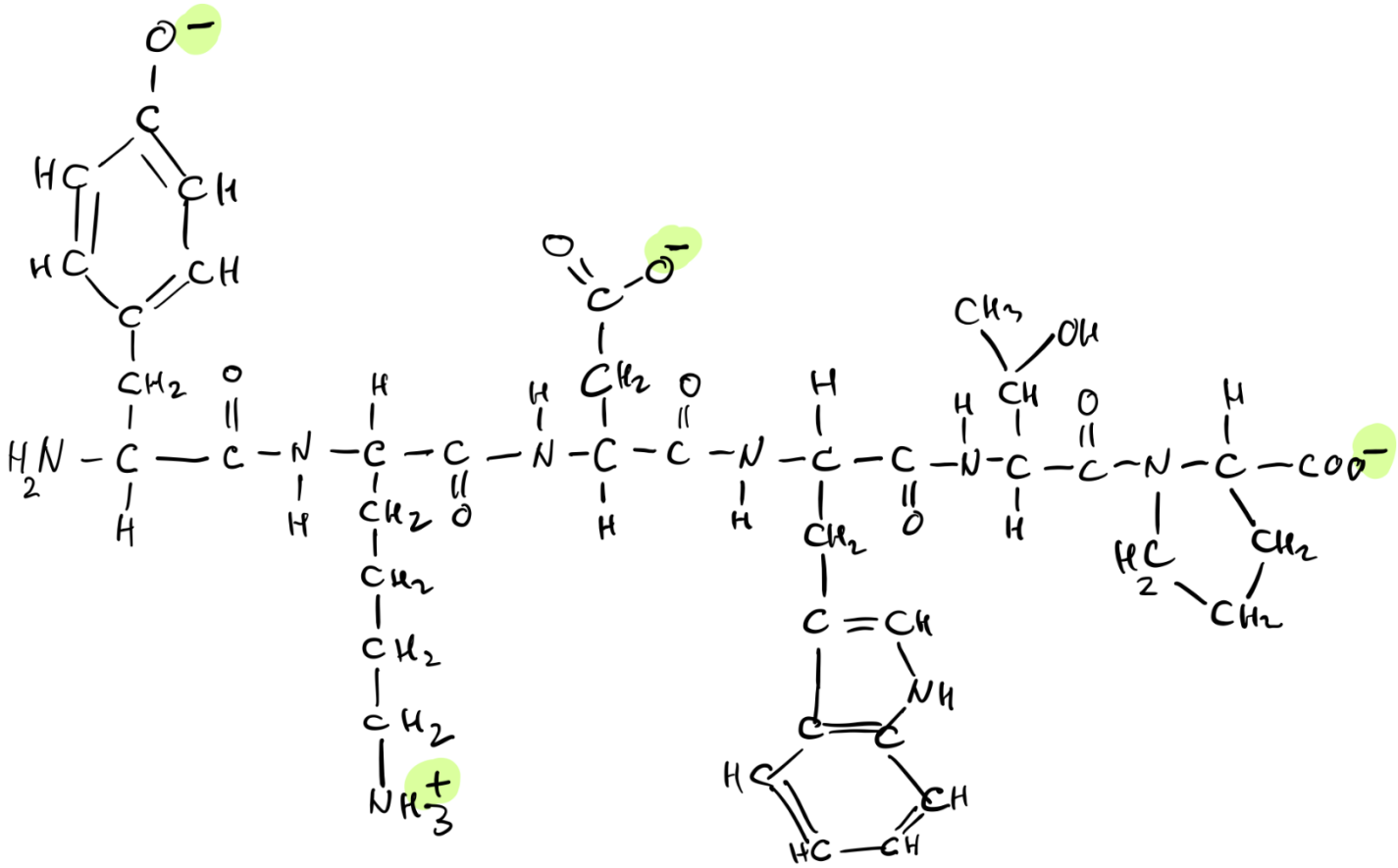
A. Draw the **complete structure** (including hydrogens) of the hexapeptide. Show the main forms of all ionizable groups at pH=10.3.

B. Estimate the isoelectric point of this peptide. *Show your calculation.*

The same questions are repeated below on this and next page to give you more space for your answers

A. Draw the **complete structure** (including hydrogens) of the hexapeptide. Show the main forms of all ionizable groups at pH=10.3.

At pH 10.3, the N-terminal amine will be predominantly deprotonated, i.e. neutral, the Tyr ring will be predominantly deprotonated, i.e. negatively charged, the ε-amino group of Lys will be mainly positively charged, and the C-terminal carboxyl and the carboxyl in the side chain of Asp will be deprotonated, i.e. negatively charged.



B. Estimate the isoelectric point of this peptide. *Explain your reasoning and show your calculation.*

First let's identify all ionizable groups in this peptide (**YKDWTP**): N-terminal amine (pK_2= 9.7), side chains of Tyr (pK_{Tyr} = 10.07), Lys (pK_{Lys} = 10.53), and Asp (pK_{Asp} = 3.65), and the C-terminal carboxyl group (pK_1 = 2.3) (according to the Lehninger textbook values and the "standard" pKas for the N and C termini).

Actually, the Thr side chain hydroxyl group (OH) has the pKa value of ~13 or higher, but because we are focused on lower pH values at which this group is protonated/uncharged, I did not include it into consideration here.

Let's now consider a pH grid and check the predominant charge states of the individual groups and the total charge of the peptide.

pH	NH_2	Tyr	Lys	Asp	COOH	Total charge
1	+	0	+	0	0	+2
3	+	0	+	0	-	+1
5	+	0	+	-	-	0 (zw)
9.8	0	0	+	-	-	-1
10.3	0	-	+	-	-	-2
11	0	-	0	-	-	-3

Thus, **$pI = (pK_{Asp} + pK_2)/2$** = (3.65 + 9.7)/2 = **6.675**

Problem 4. (22 points) Suppose you have a mixture of five proteins listed in the table below.

Protein	pI	Number of amino acids
A	4.0	100
B	4.4	640
C	6.8	77
D	7.8	600
E	10.1	250

1. Predict the order in which these proteins will migrate on the PAGE SDS gel, starting with the fastest. (4 pnts)

Smaller proteins generally migrate faster in SDS-PAGE. So the proteins will migrate in the following order **C, A, E, D, B**

2. Predict the order in which these proteins will elute from a gel-filtration (size-exclusion) column, starting with the fastest. (4 pnts)

Bigger proteins generally migrate faster through the size-exclusion column. So the proteins will elute in the following order: **B, D, E, A, C**

3. You loaded an aliquot (small portion) of this mixture on a cation column (that bears negatively charged groups). The proteins are in sodium acetate buffer at pH 5.4. List the proteins that will bind to the column. (4 pnts)

At this pH proteins C, D, and E will be positively charged, and A and B will be negatively charged. The former three proteins will bind to the cation exchange column: **C, D, E**

4. You also loaded another aliquot of the mixture, this time on an anion column (that bears positively charged groups), and the buffer is phosphate buffer (pH 6.8). List the proteins that will bind to the column. (4 pnts)

At this pH proteins A and B will be positively charged, C will be neutral, and D and E will be negatively charged. Thus, the proteins that will bind to the anion exchange column: **A** and **B**

5. Now it's time to do the separation. Devise a scheme that will allow you to unambiguously separate and purify all these proteins using a combination of the preparative chromatographic methods mentioned in the previous questions. *Explain your reasoning.* Remember, you goal is to separate the proteins unambiguously. (6 pnts)

There are several scenarios. Take advantage of the differences in the size and the pI values. What you want to avoid is trying to separate proteins that are close in size (e.g. 77 a.a. from 100 a.a. or 600 a.a. from 640 a.a. proteins) using size exclusion or those that are close in pI (e.g. 4.0 from 4.4) by ion exchange.

For example, using your answers to questions 3-4, you can
(1) first separate C, D, and E from the rest of the proteins using cation exchange column and a buffer with pH ~ 5.5 (these proteins with bind to cation column);
(2) then take the flow through (which contains A and B) and separate them from each other using size-exclusion chromatography: you should be able to separate 100 a.a. protein from 640 a.a. protein;

(3) then elute proteins C, D, and E from the cation exchange column by adding high-salt buffer and separate them using size-exclusion chromatography: you should be able to separate 77 a.a. protein from 600 a.a. protein from 250 a.a. protein.

Or use an anion exchange column in step 1 above to have A and B bound to the column and C, D, E in the flow through…

Or you can first use size-exclusion chromatography to separate B+D from A+C and from E, and then use the cation exchange column (as in question 3) to separate B from D and A from C.

Or some other scenario….

Note after grading this problem: in question 5 I specifically asked you to design a scheme for *unambiguous* separation and purification of proteins using *preparative* chromatographic methods. For some reason some students ignored this and proposed the use of SDS PAGE to separate and purify the proteins. As discussed in class, the SDS PAGE is an analytical, not preparative, method, and once you get your proteins on a SDS PAGE gel, they are unfolded/SDS treated, so you won't be able to use preparative methods to separate/purify and use them after that. Obviously, points were subtracted.

Problem 5. (12 points) In the course of amino-acid sequence determination of an unknown peptide enzymatic cleavage was combined with subsequent sequencing. You obtained the following results.

1. Trypsin cleaved the peptide into the following three fragments:

EAWK
DNLGQP
QIACNSR

2. Asp-N protease cleaved the peptide into the same three fragments.

A. Based on these data determine the correct amino acid sequence of the peptide.

QIACNSREAWK DNLGQP

B. Now that you *know* the sequence of the peptide, design/propose a reaction using one of the cleavage reagents from Table 1 that will allow unambiguous determination of the peptide's sequence in a single step.

Chymotrypsin: **QIACNSREAW KDNLGQP**

or

Pepsin: **QIACNSREA WKDNLGQP**

Problem 6. (9 points)

You need to determine how many subunits are in a 40-kDa homo-multimeric protein (consisting of several identical subunits). For this you decided to use Fred Sanger's method for identification of the N-terminal amino acid. A 200 mg sample of the protein was treated with an excess of 1-fluoro-2,4-dinitrobenzene under slightly basic conditions until the chemical reaction was complete, and the peptide bonds were then completely hydrolyzed by boiling with concentrated HCl. The resulting hydrolysate contained 8.91 mg of the Leucine derivative of 2,4-dinitrophenyl (DNP-Leu; MW = 297 Da). No 2,4-dinitrophenyl derivatives of α−amino groups of other amino-acids could be found. Based on these data, determine how many subunits are there in the structure of this protein. *Show your calculations and explain how you arrived at your answer.*

Solving this problem requires counting the number of N-termini per one molecule of the protein.
The number of moles of the protein in this reaction is $N_P = 200$ mg/40000 g $= 5 \times 10^{-6}$ moles.
The number of moles of DNP-Leu obtained is $N_{DNP\text{-}Leu} = 8.91$ mg/297 g $= 3 \times 10^{-5}$ moles.
The ratio of the two gives: $N_{DNP\text{-}Leu}/N_P = 6$.
Thus, 6 molecules of DNP-Leu were found per molecule of the protein – this indicates that there are 6 N-termini, i.e. 6 polypeptide chains, hence the protein consists of 6 subunits. It must be a **hexamer**.

Note after grading this problem: Since this problem involves only four numbers, one can manipulate/find a combination of them that gives the correct answer without any reasonable justification or logics. In those cases the "correct" answer was considered a "lucky guess", and only partial credit was given.