

BIOCHEMISTRY I (CHMI 2227 E)

PROBLEMS and SOLUTIONS

Eric R. Gauthier, Ph.D.

Department of Chemistry and Biochemistry

January 2007

Note:

This problem set has been prepared for students taking the course Biochemistry I (CHMI 2227E), as offered at Laurentian University. It contains several problems taken from textbooks and from the author's imagination.

While the vast majority of the problems found in this book can be relatively easily solved with the help of the class notes, more difficult questions have also been included. Questions marked by a star (*) will require more work from the student. As for the questions labeled with two stars (**), they constitute a good challenge to any student interested in tackling them.

After the « Problems » section, the complete, detailed solution for every question is found. For obvious reasons, we strongly encourage students to look at the solutions only as a last resource.

The list of pK_as and pI for the 20 natural amino acids, as well as the table of the genetic code, can be found after the “Problems” section.

The following texts were consulted while writing this manual:

- 1) Kuchel, P. W. and Ralston, G. B. *Biochemistry*. Schaum Series. McGraw-Hill. 1989.
- 2) Lehninger, A. L., Nelson, D. L., Cox, M. M. *Principles of Biochemistry*. 2nd édition. Worth Publishers. 1993.
- 3) Mathews, C. K. et van Holde, K. E. *Biochemistry*. 2nd édition. Benjamin/Cummings Publishing Company, INC. 1996.
- 4) Rawns, J. D. *Biochemistry*. Editions du renouveau pédagogique. 1990.
- 5) Wood, W. B., Wilson, J. H., Benbow, R. M., Hood, L. E. *Biochemistry. A Problems Approach*. Benjamin/Cummings Publishing Company, INC. 1981.
- 6) Zubay, G. L., Parson, W. W., Vance, D. E. *Principles of Biochemistry*. Wm. C. Brown Publishers. 1995.

More problems and questions can be found in these and other references.

Problems

Chapter 1: Acid-Base Equilibrium and Spectrophotometry

1.1 Acid-Base Equilibrium :

What is the pH of the following solutions?

- 0.35 M hydrochloric acid
- 0.35 M acetic acid ($pK_a = 4.76$)
- 0.035 M acetic acid.

1.2 Acid-Base Equilibrium :

A weak acid, HA, has a total concentration of 0.20M and is ionized (dissociated) to 2%;

- Calculate the K_a for this acid.
- Calculate the pH for this acidic solution.

1.3 Acid-Base Equilibrium :

Calculate the pH of the following mixtures:

- 1M acetic acid and 0.5M sodium acetate
- 0.3M phosphoric acid and 0.8M KH_2PO_4 ($pK_a=2.14$)

1.4 Acid-Base Equilibrium :

You need to prepare a buffer solution at $pH = 7.00$ with KH_2PO_4 and Na_2HPO_4 ($pK_a=7.21$). If you use a 0.1M solution of KH_2PO_4 , what would be the concentration of Na_2HPO_4 needed?

1.5 Acid-Base Equilibrium :

You need to prepare a buffer solution at $pH = 7.00$ with KH_2PO_4 and Na_2HPO_4 . What would be the respective concentration of these substances if you wished to obtain a final phosphate concentration ($[HPO_4^{2-}] + [H_2PO_4^{-1}]$) of 0.3M?

1.6 Spectrophotometry :

What is the concentration of the amino acid tyrosine ($\epsilon=1\ 420\ L\ mol^{-1}\ cm^{-1}$) if you obtain an absorbance of 0.71 with a 1 cm cuvette? With a 0.1 cm cuvette?

1.7 Spectrophotometry :

What would be the absorbance reading of a 37 mM solution of tyrosine?

1.8 Spectrophotometry :

You wish to determine the concentration of haemoglobin in a blood sample by spectrophotometry. You first create a standard curve of the absorbance at 412 nm of several solutions of known haemoglobin concentrations. The data for the standard curve is shown below. What is the concentration (in $\mu g/mL$) of haemoglobin in your sample if the absorbance obtained at 412 nm was 0.303?

Absorbance (412nm)	Concentration of standard solution ($\mu\text{g/ml}$)
0.069	1
0.113	2
0.201	4
0.377	8
0.730	16

Chapter 2: Amino acids

* 2.1. Molecular mass of an amino acid.

1.812 g of a crystallized α -amino acid ($\text{pK}_{\text{a}1}$: 2.4; $\text{pK}_{\text{a}2}$: 9.7) has a pH of 10.4 when dissolved in 100 mL of 0.1M NaOH. Calculate the molecular mass of this amino acid.

2.2. Titration curve

Calculate the pI of histidine and draw its titration curve. Indicate the position of all pKas and the pI as well as the percentages of each ionic form at the start and finish of the titration and at all pKas. The list of pKas for all 20 amino acids can be found at the end of the “Problems” section of this problem set.

2.3. Net charges of amino acids

What is the net charge (+, 0, -) of the amino acids glycine, serine, aspartic acid, glutamine and arginine at:

- a) pH 2.01 b) pH 3.96 c) pH 5.68 d) pH 10.76

2.4. Ionic exchange chromatography

A mixture of lysine, glycine, alanine, isoleucine and glutamic acid are separated by ionic exchange chromatography. What is the order of elution of these amino acids if you use gradient buffer system from pH 10 to pH 2:

- a) with a cation exchange resin?
b) with an anion exchange resin?

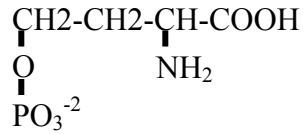
Which column would give the best separation?

2.5. Amino acids

What amino acids can be converted into another amino acid with gentle hydrolysis, resulting in release of ammonia?

2.7. Amino acids

Phosphoserine is found after enzymatic hydrolysis of casein, a milk protein. However, it does not belong to the 20 amino acids coded during protein synthesis. Give a plausible explanation.



Phosphoserine

*2.8. Ionic exchange chromatography

Glycine, alanine, valine and leucine can be successfully separated by ionic exchange chromatography even though their pKas are almost identical. Explain the behaviour of these amino acids.

2.9. Peptides.

A peptide is hydrolyzed and its amino acid content analyzed. Hydrolysis destroys the amino acid tryptophan, therefore the content of tryptophan can be estimated with spectrophotometry. Establish the empirical formula of the peptide with the following information.

<i>Amino acids</i>	<i>mmol</i>
Ala	2.74
Glu	1.41
Leu	0.69
Lys	2.81
Arg	0.72
Trp	0.65

2.10. Peptides.

Draw the structure of the following peptide GWYQR. Indicate the ionic form of the peptide at the following pH:

a) pH 2.0

b) pH 7.0

c) pH 10.5

Chapter 3. General properties and purification of proteins

3.1. Protein Purification

Why do we often use ammonium sulphate precipitation in initial purification steps of proteins?

3.2. Protein Purification

DEAE cellulose columns are rarely used at pH greater than 8.5. Why?

3.3. Protein Purification

6-phosphogluconate dehydrogenase has a pI of 6. Explain why the buffer used for a chromatography on DEAE-cellulose must have a pH greater than 6 but less than 9 in order to ensure the enzyme is efficiently bound to the column.

3.4. Protein Purification.

Would the enzyme, 6-phosphogluconate dehydrogenase bind to a CM-cellulose resin if the same conditions as the previous problem were used? Why?

3.5. Protein Purification.

What pH would the buffer need to be in order to permit the dehydrogenase in the previous problem to bind to the CM-cellulose resin?

3.6. Protein Purification.

We load a DEAE-cellulose column adjusted to a pH of 6.5 with the following mixture of proteins: ovalbumin (pI = 4.6), urease (pI = 5.0), and myoglobin (pI = 7.0). The proteins are eluted first with a buffer of weak ionic strength at a pH of 6.5, and then the same buffer containing increasing amounts of sodium chloride is used to elute the proteins. What order are the proteins eluted?

3.7. Protein Purification.

An enzyme (MW 24 kDa, pI 5.5) is contaminated with two other proteins, one with a similar molecular mass and a pI of 7.0 while the other has a molecular mass of 100 kDa and a pI of 5.4. Suggest a procedure to purify the contaminated enzyme.

3.8. Protein Purification.

A procedure used to purify 6-gluconate dehydrogenase from *E. coli* is presented below.

a) Calculate (1) the specific activity, (2) the percent yield based on the *initial* quantity of the enzyme and (3) the degree of purification for each step (i.e. *fold* increase in purification).

b) Indicate which step purifies the protein the most.

c) Assuming the protein is pure after gel permeation chromatography (on Bio-Gel A), what percent of the initial extract contained 6-gluconate dehydrogenase?

<i>Purification step</i>	<i>Volume (mL)</i>	<i>Total protein (mg)</i>	<i>Enzymatic activity ($\mu\text{g}/\text{min}$)</i>
1- Cellular extract	2 800	70 000	2 700
2- Ammonium sulfate	3 000	25 400	2 300
3- Heat denaturation	3 000	16 500	1 980
4- DEAE chromatography	80.00	390.00	1 680
5- CM-cellulose chromatography	50.00	47.00	1 350
6- Bio-Gel A chromatography	7.00	35.00	1 120

3.9 Protein Purification.

Why is SDS omitted when proteins need to undergo isoelectric focusing?

3.10. Protein Purification.

A series of proteins with known molecular mass and an enzyme of unknown molecular mass are separated by chromatography on a Sephadex G-200 column. The elution volume (V_e) for each protein is indicated in the table below. Estimate the molecular mass of the unknown protein.

<i>Protein</i>	<i>Mr</i>	<i>V_e (mL)</i>
Blue dextran	1 000 kDa	85.00
lysozyme	14 kDa	200.00
Chymotrypsinogen	25 kDa	190.00
ovalbumin	45 kDa	170.00
Serum albumin	65 kDa	150.00
aldolase	150 kDa	125.00
urease	500 kDa	90.00
ferritin	700 kDa	92.00
ovomuroid	28 kDa	160.00
unknown	?	130.00

***3.11. Protein Purification.**

Referring to the previous problem, give a plausible explanation for the bizarre behaviour ferritin's elution from the sephadex column.

3.12. Protein Purification.

A student isolates a protein from anaerobic bacteria and analyses the protein by polyacrylamide gel electrophoresis containing SDS (PAGE-SDS). Following protein staining, a single band appears, which excites the student's supervisor. To be certain, the supervisor suggests that the student run a second electrophoresis under *native* conditions (i.e. non-denaturing, or without SDS). This gel shows two bands after staining. Assuming no errors were committed during these experiments, explain the observations.

3.13. Protein Purification.

A student from CHMI 2227 analyses bovine serum albumin (BSA) with a polyacrylamide gel electrophoresis (PAGE-SDS). During the experiment, the student forgets to add β -mercaptoethanol to the sample. When comparing his sample to those of his classmates he realizes that the molecular mass of his BSA sample determined by PAGE-SDS is 57 kDa, while all the other students (those that added β -mercaptoethanol) found a molecular mass of 68 kDa. Explain this difference.

3.14. Polypeptide sequencing

Consider the following peptide:

A-L-K-M-P-E-Y-I-S-T-D-Q-S-N-W-H-H-R

Indicate the fragments generated after the following digestions :

- a) trypsin b) pepsin c) protease V8 d) cyanogen bromide

3.15 Polypeptide sequencing

Deduce the polypeptide sequence that generated the following results:

a) acid hydrolysis: (Ala₂, Arg, Lys₂, Met, Phe, Ser₂);

b) Carboxypeptidase A digestion: Ala;

c) Trypsin digestion: (Ala, Arg)
 (Lys, Phe, Ser)
 (Lys)
 (Ala, Met, Ser)

d) cyanogen bromide treatment: (Ala, Arg, Lys₂, Met, Phe, Ser)
 (Ala, Ser)

e) thermolysine digestion: (Ala)
 (Ala, Arg, Ser)
 (Lys₂, Met, Phe, Ser)

3.16. Polypeptide sequencing

A polypeptide is reduced by β -mercaptoethanol to yield two peptide fragments with the following sequences :

fragment 1: A-C-F-P-K-R-W-C-R-R-V-C

fragment 2: C-Y-C-F-C

The non-reduced polypeptide is digested with thermolysine and yields the following fragments :

(A,C,C,V)

(R,K,F,P)

(R,R,C,C,W,Y)

(C,C,F)

Indicate the positions of disulfide bridges in the polypeptide.

3.17. Polypeptide sequencing

An analysis of the polypeptide Shawi isolated from the bacteria *Chretientus negativii*, yields the following results :

a) acid hydrolysis: (Ala₄, Val, Lys₂, Arg, Gly, Asp, Met, Pro, Trp)

b) carboxypeptidase digestion: Lys

c) dinitrofluorobenzene treatment: Val

d) cyanogen bromide treatment: generates two polypeptides:

peptide A: (Gly, Arg, Trp, Asp, Lys, Ala); Treatment of this peptide with DNFB and carboxypeptidase yields :

DNFB: Gly

Carboxypeptidase: Lys

peptide B: (Ala₃, Lys, Val, Met, Pro); Treatment of this peptide with DNFB and carboxypeptidase yields:

DNFB: Val

Carboxypeptidase: Met

e) trypsin digestion: yields three peptides

peptide C: (Lys, Trp, Ala); Treatment of this peptide with DNFB and carboxypeptidase yields :

DNFB: Trp

peptide D: (Ala₃, Val, Lys, Pro)

peptide E: (Met, Asp, Gly, Arg); Treatment of this peptide with DNFB and carboxypeptidase yields :

DNFB: Met

Finally, treating peptide D with thermolysine yields the following:

Val
Ala
Ala
(Ala, Lys, Pro)

What is the primary structure of this peptide?

Chapter 4. Three dimensional structures of proteins

4.1. 3-D Structures of proteins

What amino acids among the following would you expect to find a) inside, and b) at the surface of a typical globular protein in an aqueous solution of pH 7?

Glu	Arg	Val
Phe	Ileu	Asn
Lys	Ser	Thr

4.2. 3-D Structures of proteins

According to the structure of urea, deduce how this compound can promote denaturation of proteins.

4.3. 3-D Structures of proteins

Phenylalanine, a hydrophobic amino acid, is frequently found at the surface of natives and functional proteins. Give the most probable role of phenylalanine in this situation.

****4.4. 3-D Structures of proteins***

Aspartic acid, a charged amino acid, is frequently found inside of native and functional proteins. Give the most probable role of phenylalanine in this situation.

4.5. 3-D Structures of proteins

The following table describes the amino acid compositions of three proteins.

	<i>Number of residues per molecule</i>		
<i>Amino acids</i>	<i>protein 1</i>	<i>protein 2</i>	<i>protein 3</i>
<i>Polar residues</i>			
Arg	12.00	4.00	7.00
Asn	9.00	6.00	5.00
Asp	14.00	5.00	9.00

<i>Amino acids</i>	<i>Number of residues per molecule</i>		
	<i>protein 1</i>	<i>protein 2</i>	<i>protein 3</i>
Cys	7.00	2.00	6.00
Gln	8.00	7.00	6.00
Glu	11.00	4.00	6.00
His	4.00	2.00	4.00
Lys	22.00	6.00	15.00
Ser	20.00	5.00	11.00
Thr	15.00	3.00	11.00
Trp	2.00	3.00	3.00
Tyr	7.00	7.00	6.00
<i>Non-polar residues</i>			
Ala	14.00	28.00	25.00
Gly	9.00	9.00	8.00
Ileu	5.00	16.00	9.00
Leu	3.00	19.00	7.00
Met	7.00	11.00	9.00
Phe	9.00	13.00	11.00
Pro	8.00	13.00	10.00
Val	16.00	29.00	21.00

Knowing that protein A has a rod-like form, protein B is a monomeric globular protein, and protein C is a globular protein with four identical sub-units, deduce the corresponding amino acid composition of these proteins.

4.6. 3-D Structures of proteins

Indicate which secondary structure or structures (α -helix, β -pleated, random coil) will the following peptide adopt in an aqueous solution at pH 7

Ileu-Glu-Asn-Glu-Gln-Asn-Met-Ala-His-Phe-Trp-Tyr

4.7. 3-D Structures of proteins

Indicate which secondary structure or structures (α -helix, β -pleated, random coil) will the following peptide adopt in an aqueous solution at pH 7

Gly-Ala-Gly-Ala-Gly-Ser-Gly-Ala-Gly-Ser-Gly-Ala

4.8. 3-D Structures of proteins

Indicate which secondary structure or structures (α -helix, β -pleated, random coil) will the following peptide adopt in an aqueous solution at pH 7

Lys-Gly-Arg-Arg-Lys-Gly-Arg-Gly-Arg-Pro

4.9. 3-D Structures of proteins

Indicate which secondary structure or structures (α -helix, β -pleated, random coil) will the following peptide adopt in an aqueous solution at pH 7

1 10
Gly-Pro-Glu-Ser-Ala-Tyr-Lys-Thr-Leu-Phe-Asp-Val-Pro-Asp-Asp-Glu-Asp-Gly-Gly

Ser-Ala-Gly-Ser-Ser-Gly-Ala

4.10. 3-D Structures of proteins

The following table describes the amino acid composition of three proteins. Determine what structure these proteins will adopt: α -helical, β -pleated or a triple helix of collagen.

<i>protein</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>protein</i>	<i>A</i>	<i>B</i>	<i>C</i>
Ala	29.40	5.00	10.70	Leu	0.50	6.90	2.40
Arg	0.50	7.20	5.00	Lys	0.30	2.30	3.40
Asp	1.30	6.00	4.50	Met	-	0.50	0.80
Cys	-	11.20	-	Phe	0.50	2.50	1.20
Glu	1.00	12.10	7.10	Pro	0.30	7.50	12.20
Gly	44.60	8.10	33.00	Ser	12.20	10.20	4.30
His	0.20	0.70	0.40	Trp	0.20	1.20	-
Hyp	-	-	9.40	Tyr	5.20	4.20	0.40
Ileu	0.70	2.80	0.90	Val	2.20	5.10	2.30

Chapter 5. Enzymology**5.1. Enzymatic kinetics**

With the following enzyme activity results determine:

- V_{max}
- why is the velocity v constant at $[S]$ greater than $2 \times 10^{-3} \text{ M}$?
- what is the free $[E]$ at $[S] = 2 \times 10^{-2} \text{ M}$?

$[S]$ (mol/L)	v ($\mu\text{mol}/\text{min}$)
2×10^{-1}	60.00
2×10^{-2}	60.00
2×10^{-3}	60.00
2×10^{-4}	48.00
$1,5 \times 10^{-4}$	45.00
$1,3 \times 10^{-5}$	12.00

5.2. Enzymatic kinetics

The results for enzyme activity analysis can be found below. **Without using a graph,** determine :

- V_{max};
- K_m;
- initial velocity at $[S] = 1 \times 10^{-1} \text{ M}$;

- d) the amount of product formed during the first 5 minutes at $[S] = 2 \times 10^{-3} \text{ M}$. At a $[S]$ of $2 \times 10^{-6} \text{ M}$?
- e) what is K_m and V_{max} if the free $[E]$ is increased by a factor of 4?

[S] (mol/L)	v ($\mu\text{mol}/\text{min}$)
5×10^{-2}	0.25
5×10^{-3}	0.25
5×10^{-4}	0.25
5×10^{-5}	0.20
5×10^{-6}	0.07
5×10^{-7}	0.01

5.3. Enzymatic kinetics

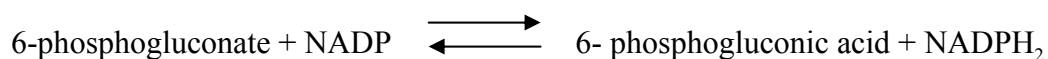
The following table describes the results from an enzymology experiment. Using a Lineweaver-Burke plot determine:

- a) K_m ;
b) V_{max} ;

[S] (mol/L)	v ($\mu\text{mol}/\text{min}$)
1×10^{-3}	65.00
5×10^{-4}	63.00
1×10^{-4}	51.00
5×10^{-5}	42.00
3×10^{-5}	33.00
2×10^{-5}	27.00
1×10^{-5}	17.00
5×10^{-6}	9.50
1×10^{-6}	2.20
5×10^{-7}	1.10

5.4. Enzymatic kinetics

We study the effect of pH on the enzymatic activity of 6-phosphogluconate dehydrogenase. This enzyme catalyzes the reaction:



NADPH_2 absorbs light at 340 nm. The activity of the dehydrogenase is measured

spectrophotometrically by monitoring the absorbance (A) at 340nm, which is proportional to the concentration of NADPH₂.

[S] x 10 ⁴ M	Increase in A at pH 7.6	Increase in A at pH 9.0
0.174	0.074	0.034
0.267	0.085	0.047
0.526	0.098	0.075
1.666	0.114	0.128
4.000	-	0.167

At what pH will the enzyme have more affinity for the substrate?

5.5. Enzymatic kinetics

The following results describe the effect of an inhibitor on enzyme activity of an enzyme. Determine:

- V_{max} in the presence and the absence of an inhibitor
- K_m in the presence and the absence of an inhibitor
- K_i
- type of inhibition

[S] (mol/L)	Without inhibitor v (μmol/min)	With inhibitor [I] = 2,2 x 10 ⁻⁴ M v (μmol/min)
1 x 10 ⁻⁴	28.00	17.00
1,5 x 10 ⁻⁴	36.00	23.00
2x 10 ⁻⁴	43.00	29.00
5x 10 ⁻⁴	65.00	50.00
7,5 x 10 ⁻⁴	74.00	61.00

5.6. Enzymatic kinetics

A biochemist studies the properties of a metabolic enzyme she has just isolated. She obtains kinetic data in the presence and in the absence of two different inhibitors (A and B). The identity of the inhibitors is unknown but we know that one of these is a substrate analog while the other is an alkylating agent.

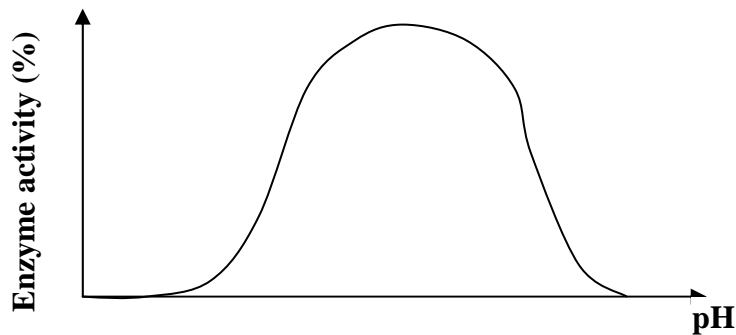
Determine:

- K_m and V_{max} of the enzyme ;
- which inhibitor is the substrate analog? Which is the alkylating agent?
- K_i for both inhibitors;
- what would be the V_o for this enzymatic reaction at [S] = 3 × 10⁻⁴ M and in the presence of the inhibitor [A] = 2 × 10⁻⁵ M?

[S] (mol/L)	Without inhibitor v (μmol/min)	With inhibitor A [I] = 5 × 10 ⁻⁴ M v (μmol/min)	With inhibitor B [I] = 3,2 × 10 ⁻⁶ M v (μmol/min)
5 × 10 ⁻⁴	1.25	0.82	0.48
2,5 × 10 ⁻⁴	0.87	0.49	0.33
1,7 × 10 ⁻⁴	0.67	0.36	0.25
1,2 × 10 ⁻⁴	0.54	0.26	0.20
1 × 10 ⁻⁴	0.45	0.23	0.17

5.7. Enzymatic catalysis

The effect of pH on the activity of an enzyme is demonstrated in the following graph :



How would you explain the effect of pH on enzyme activity?

5.8. Enzyme catalysis

Several enzymes show a dependance on pH similar to the one shown in the previous problem. However, the optimal pH varies a great deal from one enzyme to another. What side chains would you expect to find on active sites of enzymes if the optimal pH is:

- pH 4
- pH 11

5.9. Allosteric enzymes

We study the kinetic properties of two enzymes (A and B). From the results shown below, determine if they constitute an ordinary enzyme or an allosteric enzyme. Explain the shape of the curves representing the velocity, v , in relation to the concentration of substrate, $[S]$.

$[S]$ ($\times 10^3$ M)	v (enzyme A) ($\mu\text{mol}/\text{min}$)	v (enzyme B) ($\mu\text{mol}/\text{min}$)
0.00	0.00	0.00
0.50	8.80	0.30
1.00	14.00	1.00
2.00	19.00	4.70
3.00	21.50	12.40
4.00	22.80	19.00
5.00	22.30	21.80
6.00	23.50	22.80
8.00	23.60	23.30

Chapter 6. Structure and properties of nucleic acids.

6.1. Nucleic acid structure.

Consider the following polynucleotide:

AUUACGUGGUGCACUCGGGAACAUCCCGAGUGCACCACGUAAUGGA

Draw the two most stable **intramolecular** secondary structures this polymer can adopt

*6.2. Nucleic acid structure.

A solution of double stranded DNA is heated and then cooled to room temperature for two minutes. Predict, qualitatively, the variation in absorbance at 260 nm in the following conditions:

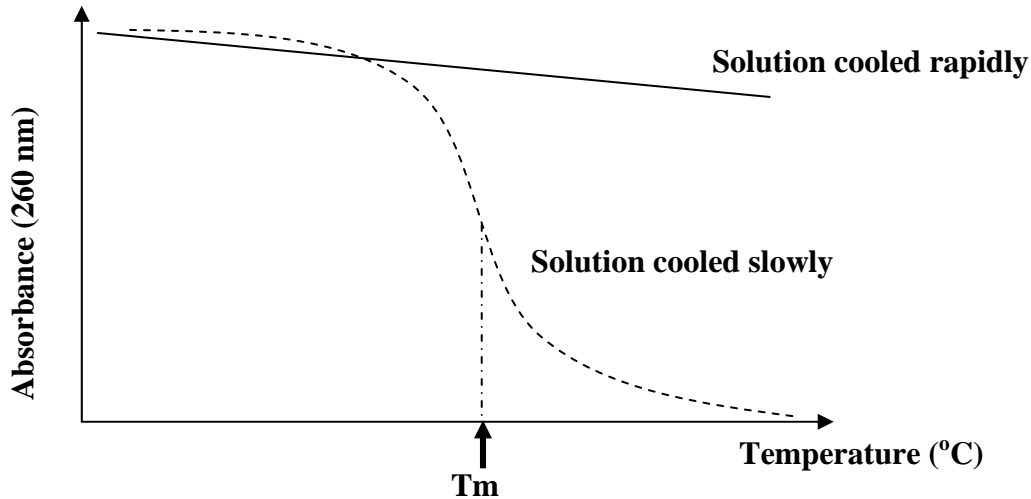
- the solution is heated to a temperature slightly above T_m before being cooled;
- the solution is heated to a temperature way above T_m before being cooled;
- suggest the structure of two polynucleotides (synthetic or natural) which will result in an absorbance profile following a cooling which is the perfect inverse of the pattern obtained in (b).

6.3. Nucleic acid structure.

Explain why, RNA, and not DNA, is hydrolyzed under basic pH conditions.

6.4. Nucleic acid structure.

The following results were obtained during a denaturation/renaturation experiment of a simple nucleic acid (polyA :polyU). How would you interpret these results?



6.5. Nucleic acid structure.

IMP (inosine monophosphate) is present in *chez E. coli* as an intermediate of biosynthesis of purines and it is possible to incorporate IMP to DNA if the ITP (inosine triphosphate) is present in the reaction medium. However, in nature, IMP is never present in DNA. Propose an explanation.

6.6. Nucleic acid structure

What are the products of the digestion of the oligoribonucleotide 5'pACGAUGCUAUC3' by each of the following enzymes:

- pancreatic ribonuclease;
- T2 ribonuclease;
- T1 ribonuclease;

6.7. Nucleic acid structure

Lets proceed to the analysis of an RNA molecule. Its global base composition is 2A, 2C, 1U, 1G.

Its treatment with the serpent venom phosphodiesterase yields pC.

Its hydrolysis by pancreatic ribonuclease yields 1C, a dinucleotide containing A and C, and a trinucleotide containing A, G, and U.

The action of RNase T2 yields pAp, a dinucleotide containing U and C and a trinucleotide containing A, G and C.

What is the primary structure of this RNA?

6.8. Nucleic acid structure

Let's proceed to the analysis of an RNA molecule whose global base composition is 2A, 4C, 2G, 1U.

Pancreatic ribonuclease treatment yields 2Cp, two dinucleotides, one containing G and C and the other containing A and U, and a trinucleotide containing A, C and G.

A mixture of RNase T1 and RNase T2 yields C, Ap, pGp and two trinucleotides, one containing A and C and the second containing CG and U.

The serpent venom phosphodiesterase yields pC.

What is the formula of this RNA?

6.9. Nucleic acid structure

What is the global charge of the trinucleotide ApGpUpC at neutral pH?

6.10. Nucleic acid structure

Why does a circular double stranded DNA renature more rapidly than a linear double stranded DNA?

6.11. Nucleic acid structure

Why does DNA denature in pure water, that is where the ionic strength is close to zero?

6.12. Nucleic acid structure

The size of the E. coli chromosome is 4000 kbp. What length of DNA does it contain?

6.13. Nucleic acid synthesis.

During an experiment similar to that performed by Meselson and Stahl, you grow bacteria for 3 generations (instead of 2 as in the classic experiment) in a mixture containing only ^{14}N . Following DNA isolation and analysis by analytical centrifugation, what proportion of heavy DNA, hybrid DNA and light DNA will you obtain?

6.14. Nucleic acid synthesis.

An isolated strand (+) of DNA (base composition: 10% of A, 20% of G, 30% of C and 40% of T) is replicated by E. coli DNA polymerase into a complementary strand (-). The double-stranded DNA is then used as a model for the E. coli RNA polymerase which transcribes the (-) strand.

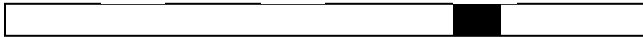
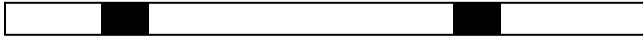
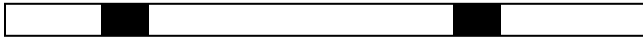
Indicate the base composition of the formed strand (in % of A, C, G of T/U).

***6.15. Nucleic acid synthesis.**

The time required to completely synthesise the E.coli genome is 40 minutes. However, it takes only 20 minutes for these bacteria to produce one generation. Can you explain this paradox?

****6.16. Nucleic acid synthesis.**

You are the first scientist to successfully analyze a micro-organism found on Mars. Because this bacterium contains double-stranded DNA as genetic material, you decide to analyze using Meselson-Stahl techniques. You obtain the following results:

	Generations after ^{14}N transfer
LL HL HH	
	0
	1
	2

a) how would you interpret these results?

b) in order to better understand this phenomenon, you isolate the components implicated in DNA replication in this organism. You identify :

- a RNA polymerase activity;
- a DNA polymerase which functions only on single-stranded;
- a new enzyme which can generate a product sensitive to DNase in the presence of NADH and a product insensitive to DNase and resistant to heat.

According to this information, deduce the mechanism by which this micro-organism replicates its DNA.

6.17. mRNA and transcription

Differently than DNA polymerase, RNA polymerase does not proofread and edit its products.

a) Why does this absence of proofreading/correction in the synthesis of RNA not threaten the cells' viability?

b) How would an enzyme using RNA as a template for DNA synthesis modify the rate of mutations for an organism?

***6.18. mRNA and transcription**

The great majority of mRNAs have a very short half life – in the order of 3 minutes in bacteria. What caused evolution to form mRNA molecules so unstable?

6.19. mRNA and transcription

If RNA polymerase lengthens RNA at a speed of 35 to 70 nucleotides per second and if each molecule of polymerase binds to 70 base pairs of DNA :

- a) What is the maximum speed of transcription per minute where a gene of 6000 base pairs is transcribed into RNA molecules?
- b) What is the maximum number of molecules of polymerase that could be found bound to this gene at any given time?

6.20. Protein coding

Consider the following mRNA:

AGU CUC UGU CUC CAU UUG AAG AAG GGG AAG GGG

- a) indicate the amino acid sequence which would be coded (read from 5' to 3'). The table containing the genetic code can be found in the appendix.
- b) you obtain mutations which consist of additions or deletions of one nucleotide. If we insert G between the third and fourth nucleotide, and we eliminate the 10th nucleotide from the right (it is a G), what would be the peptide sequence?

6.21. Protein coding.

The amino acid sequence from part of lysozyme isolated from a wild type and a mutant bacteriophage T4 is given below:

wild type: -Tyr-Lys-Ser-Pro-Ser-Leu-Asn-Ala-Ala-Lys-

mutant: -Tyr-Lys-Val-His-His-Leu-Met-Ala-Ala-Lys-

- a) can this mutant be the result of a change in a single base pair in the DNA of phage T4? If not how was this mutant produced?
- b) what is the base sequence of the mRNA which codes for the five amino acids in the wild type which are different than those of the mutant type?

6.22. Protein coding.

A strand of DNA has the following sequence:

5' TCGTTTACGATCCCCATTTCGTA C TCGA 3'

- a) what is the sequence of its complementary strand?
- b) what is the base sequence of mRNA transcribed from the first strand?
- c) what is the coded amino acid sequence?

d) what is the coded amino acid sequence if the second T from the 3' end of the DNA is deleted?

6.23. Genetic engineering.

Give the restriction fragments obtained following digestion of the following nucleic acid with the enzyme EcoR I:

5' ATGCTCGATCGATCGAATTCTATAGCCCGGGGCTGGATCCAGGTACCAAGTTAAGCTTG3'
 3' TACGAGCTAGCTAGCTTAAGATATCGGGCCCCGACCTAGGTCCATGGTTCAATTCGAAC5'

6.24. Genetic engineering.

Give the restriction fragments obtained following digestion of the following nucleic acid with the enzyme BamHI:

5' ATGCTCGATCGATCGAATTCTATAGCCCGGGGCTGGATCCAGGTACCAAGTTAAGCTTG3'
 3' TACGAGCTAGCTAGCTTAAGATATCGGGCCCCGACCTAGGTCCATGGTTCAATTCGAAC5'

6.25. Genetic engineering.

Give the restriction fragments obtained following digestion of the following nucleic acid with the enzyme Sma I:

5' ATGCTCGATCGATCGAATTCTATAGCCCGGGGCTGGATCCAGGTACCAAGTTAAGCTTG3'
 3' TACGAGCTAGCTAGCTTAAGATATCGGGCCCCGACCTAGGTCCATGGTTCAATTCGAAC5'

6.26. Genetic engineering.

Give the restriction fragments obtained following digestion of the following nucleic acid with the enzyme KpnI and Hind III:

5' ATGCTCGATCGATCGAATTCTATAGCCCGGGGCTGGATCCAGGTACCAAGTTAAGCTTG3'
 3' TACGAGCTAGCTAGCTTAAGATATCGGGCCCCGACCTAGGTCCATGGTTCAATTCGAAC5'

6.27. Genetic engineering.

You want to map the genome of the λ bacteriophage (a double stranded linear DNA). To accomplish this, you label the genome of phage λ (total length of 48 500 bp) at the 5' end with a radioactive phosphorous (^{32}P). You then digest the marked genome with different restriction enzymes under conditions which will permit partial digestion of the DNA. You analyze the resulting fragments by agarose electrophoresis and then visualize the bands with autoradiography. The results are shown in the table below.

- Calculate the length of each restriction fragment obtained.
- Create the restriction map of the λ phage.

DNA standard		Apa I	Pvu I	BamH I
Length (bp)	Distance migrated (cm)	Distance migrated (cm)	Distance migrated (cm)	Distance migrated (cm)
23 130	3.5	2.76	2.76	2.76
9 416	4.1	4.12	3.02	2.89

6 557	4.5		3.29	3.06
4 361	4.9		3.98	3.24
2 320	5.25			3.43
2 027	6.15			4.65
560	6.7			

pKas and pI Values for Common Amino Acids

	<i>pKa1</i>	<i>pKa2</i>	<i>pKR</i>	<i>pI</i>
G	2,34	9,60		5,97
A	2,34	9,69		6,01
V	2,32	9,62		5,97
L	2,36	9,60		5,98
I	2,36	9,68		6,02
P	1,99	10,6		6,48
F	1,83	9,13		5,48
Y	2,20	9,11	10,07	5,66
W	2,83	9,39		5,89
S	2,21	9,15	13,60	5,68
T	2,63	10,43	13,60	5,87
C	1,71	10,78	8,33	5,07
M	2,28	9,21		5,74
N	2,02	8,80		5,41
Q	2,17	9,13		5,65
D	2,09	9,82	3,86	2,77
E	2,19	9,67	4,25	3,22
K	2,18	8,95	10,79	9,74
R	2,17	9,04	12,48	10,76
H	1,82	9,17	6,00	7,59

The genetic code

<i>Base at 5'</i> ↓	<i>Central bases</i>				<i>Base at 3'</i> ↓
	U	C	A	G	
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr Stop Stop	Cys Cys Stop Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

ANSWERS

Acid-Base Equilibrium and Spectrophotometry

1.1 Acid-base equilibrium :

a) Since HCl is a strong acid, it will completely dissociate when in solution:



Stoichiometry tells us that, since the initial HCl concentration is 0.35M, the final concentration of H^+ in the solution will also be 0.35M. This gives us:

$$\text{pH} = -\log[\text{H}^+] = -\log 0.35 = 0.46$$

b) Acetic acid will also dissociate in solution:



However, since it is a weak acid, it will not completely dissociate, and we have to take into account the association constant (K_a) in our calculations. This constant is described as follows:

$$\text{p}K_a = -\log K_a$$

$$K_a = 1/10^{\text{p}K_a} = 1.74 \times 10^{-5}\text{M}$$

We can now easily determine the H^+ concentration:

$$K_a = \frac{[\text{H}^+][\text{CH}_3\text{COO}^-]}{[\text{CH}_3\text{COOH}]}$$

$$1.74 \times 10^{-5}\text{M} = \frac{[\text{H}^+][\text{CH}_3\text{COO}^-]}{0.35\text{M}}$$

$$1.74 \times 10^{-5}\text{M} \times 0.35\text{M} = [\text{H}^+][\text{CH}_3\text{COO}^-] = [\text{H}^+]^2$$

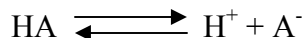
$$[\text{H}^+] = (6.09 \times 10^{-6}\text{M}^2)^{1/2} = 2.47 \times 10^{-3}\text{M}$$

$$\text{Finally: pH} = -\log [\text{H}^+] = -\log 2,47 \times 10^{-3} = 2.61$$

c) Following the same steps as in (b), we get a pH of 3.11.

1.2 Acid-base equilibrium :

a) We have a weak acid. The acid-base equilibrium is:



We can determine K_a as follows:

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

The question stipulates that this acid is only 2% ionised (or 0.02, that's the same thing). This allows us to obtain the respective concentrations of species HA, H^+ and A^- :

$$[\text{H}^+] = [\text{A}^-] = 0.20\text{M} \times 0.02 = 0.004\text{M}$$

$$[\text{HA}] = 0.2\text{M} - [\text{H}^+] = 0.196\text{M}$$

Therefore :

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

$$K_a = \frac{0.004\text{M} \times 0.004\text{M}}{0.196\text{M}}$$

$$K_a = 8.16 \times 10^{-5}\text{M}$$

b) The pH of this solution is: $\text{pH} = -\log [\text{H}^+] = -\log 0.004\text{M} = 2.39$

1.3 Acid-base equilibrium:

a) This mixture is a buffer solution made of acetic acid and its conjugated base, sodium acetate:



This pH of this type of solution can be determined with the Henderson-Hasselbach equation:

$$\text{pH} = \text{p}K_a + \log \frac{[\text{Conjugated base}]}{[\text{Acid}]}$$

$$\text{pH} = 4.76 + \log \frac{0.5\text{M}}{1\text{M}} = 4.46$$

b) We have the following acid-base equilibrium :



Using the same procedure as in (a), we get:

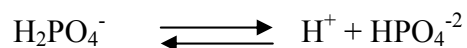
$$\text{pH} = \text{pKa} + \log \frac{[\text{H}_2\text{PO}_4^-]}{[\text{H}_3\text{PO}_4]}$$

$$\text{pH} = 2.14 + \log \frac{0.8\text{M}}{0.3\text{M}}$$

$$\text{pH} = 2.57$$

1.4 Acid-base equilibrium :

We have the following equilibrium:



And the pKa for this equilibrium is 7.21.

The Henderson-Hasselbach equation gives us:

$$\text{pH} = \text{pKa} + \log \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$7.00 = 7.21 + \log \frac{[x]}{0.1\text{M}}$$

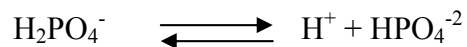
$$-0.21 = \log x - \log 0.1\text{M}$$

$$-0.21 + \log 0.1\text{M} = \log x = -1.21$$

$$x = 10^{\log x} = 0.062\text{M}$$

1.5 Acid-base equilibrium:

We have the following acid-base equilibrium:



According to the question, we have: $[\text{H}_2\text{PO}_4^-] + [\text{HPO}_4^{2-}] = 0.3\text{M}$

Hence: $[\text{H}_2\text{PO}_4^-] = 0.3\text{ M} - [\text{HPO}_4^{2-}]$

From the Henderson-Hasselbach equation, we have:

$$\text{pH} = \text{pK}_a + \log \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$7,00 = 7.21 + \log \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$7,00 = 7.21 + \log \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$-0.21 = \log \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$10^{-0.21} = \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$0.616 = \frac{[\text{HPO}_4^{2-}]}{[\text{H}_2\text{PO}_4^-]}$$

$$0.616 \times [\text{H}_2\text{PO}_4^-] = [\text{HPO}_4^{2-}]$$

Which is identical to :

$$0.616 \times (0.3\text{M} - [\text{HPO}_4^{2-}]) = [\text{HPO}_4^{2-}]$$

$$0.185 - 0.616 \times [\text{HPO}_4^{2-}] = [\text{HPO}_4^{2-}]$$

$$0.185 = 1.616 \times [\text{HPO}_4^{2-}]$$

$$0.114\text{ M} = [\text{HPO}_4^{2-}]$$

And the concentration in H_2PO_4^- will be :

$$[\text{H}_2\text{PO}_4^-] = 0.3\text{M} - [\text{HPO}_4^{2-}] = 0.186\text{ M}$$

1.6 Spectrophotometry

The relationship between the absorbance and the concentration of a solution is given by the Beer-Lambert equation:

$$A = \epsilon cl$$

Where: A = absorbance

ϵ = Molar extinction coefficient (units: litres x mol⁻¹ x cm⁻¹)

c = concentration (units : mol/l = M)

l = light path (thickness of the cuvette; units: cm)

We get the following:

$$0.71 = 1.420 \text{ L mol}^{-1} \text{ cm}^{-1} \times c \times 1 \text{ cm}$$

$$c = \frac{0.71 \text{ mol cm}}{1420 \text{ L x 1 cm}}$$

$$c = 5 \times 10^{-4} \text{ M}$$

If we use a cuvette where c=0.1 cm, we get:

$$0.71 = 1.420 \text{ L mol}^{-1} \text{ cm}^{-1} \times c \times 0.1 \text{ cm}$$

$$c = \frac{0.71 \text{ mol cm}}{1420 \text{ L x 0.1 cm}}$$

$$c = 0.005 \text{ M}$$

1.7 Spectrophotometry

With the Beer-Lambert equation, we have:

$$A = \epsilon cl$$

Therefore:

$$A = 1420 \text{ L mol}^{-1} \text{ cm}^{-1} \times (37 \times 10^{-3} \text{ M}) \times 1 \text{ cm}$$

$$A = 52.54$$

1.8 Spectrophotometry

We first have to graph the standard curve of the absorbance as a function of the concentration of the haemoglobin standards. This graph is shown on the following page.

Since the unknown has an absorbance of 0.303, we can use the standard curve to determine the corresponding haemoglobin concentration, in this case 6.31 $\mu\text{g/mL}$.

A more accurate value can be obtained by using the familiar equation:

$$y = mx + b$$

where y = value on the y axis

x = value on the x axis

m = slope

b = intersect on the y axis

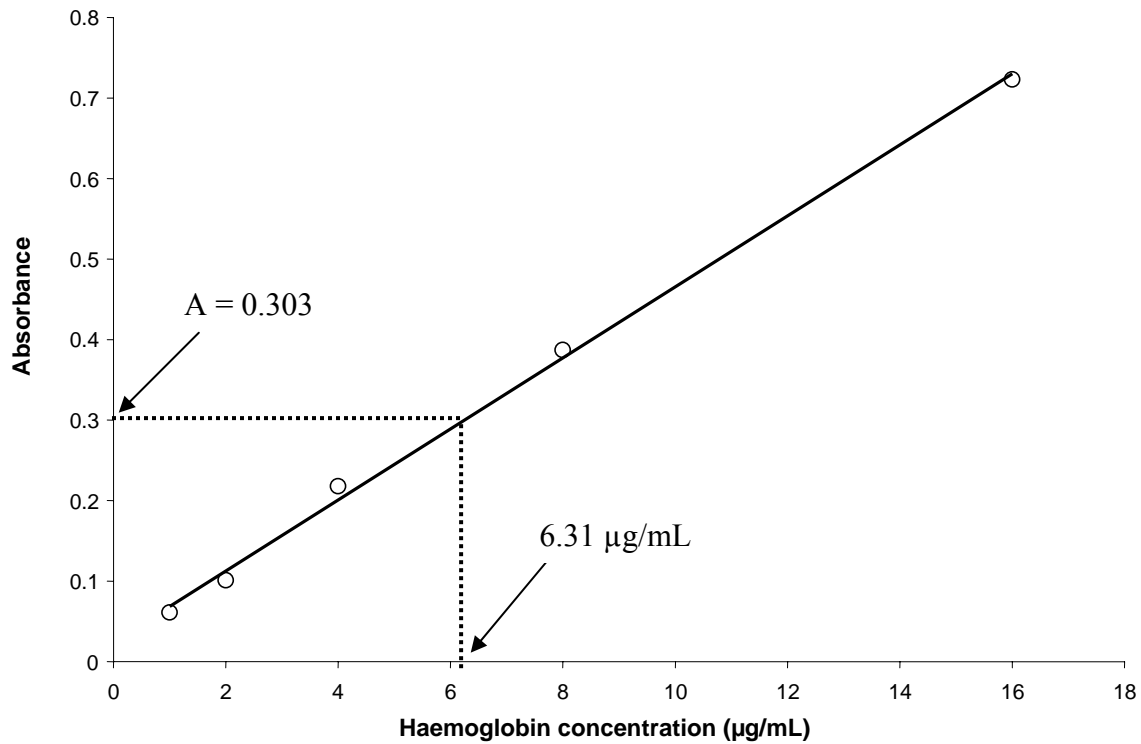
The values for m and b are easily obtained from the graph or by linear regression of the data (the latter being, by far, the best method).

Therefore, we get:

$$y = (0.0441 \text{ mL}\mu\text{g}^{-1})x + 0.0246$$

$$x = (y-b)/m$$

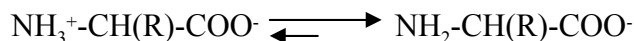
$$x = (0.303-0.0246)/0.0441\text{mL}\mu\text{g}^{-1} = 6.31 \mu\text{g/mL}.$$



Chapter 2: Amino acids

* 2.1. Molecular mass of an amino acid.

By adding NaOH, we shift the acid-base equilibrium of the amino acid towards the base:



With the Henderson-Hasselbach equation, it is possible to obtain the proportion of the amino acid in the base and acid form after the addition of NaOH:

$$\text{pH} = \text{pKa} + \text{Log} [\text{base}] / [\text{acid}]$$

$$10.4 = 9.7 + \text{Log} [\text{base}] / [\text{acid}]$$

$$0.7 = \text{Log} [\text{base}] / [\text{acid}]$$

$$5.011 = [\text{base}] / [\text{acid}]$$

Since 0.01 mol of NaOH (100 mL of a 0.1M solution) was required to obtain a pH value of 10.4, this indicates that 0.01 mol of this amino acid has been converted to the basic form. Therefore:

$$5,011 = [\text{base}] / [\text{acid}]$$

$$5,011 = 0,01 \text{ mol} / [\text{acid}]$$

$$[\text{acid}] = 0,002 \text{ mol}$$

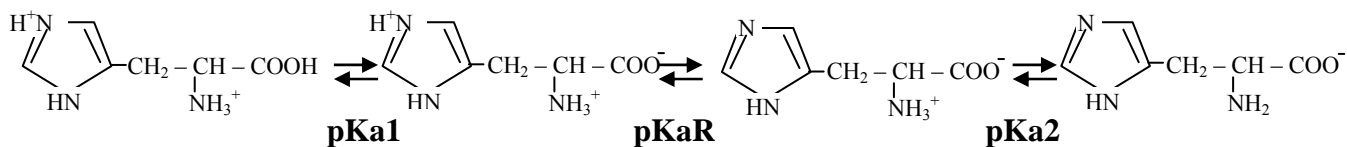
The sum of the amount of amino acid in the base and acid forms will give us the total amount of amino acid in the solution, which is 0,012 mol. All we have to do next is determine the molecular mass:

$$1,812 \text{ g} = 0,012 \text{ mol}$$

Therefore, the molecular mass of the amion acid is: $\text{g} / \text{mol} = 1,812 \text{ g} / 0,012 \text{ mol} = 151 \text{ g} / \text{mol}$

2.2. Titration of amino acids.

The titration of histidine involves the following equilibria:



To draw the titration curve, we need the values for the pKas and the inflexion points. For histidine, the pKas are as follows:

pKa1: 1.82
 pKaR: 6.00
 pKa2: 9.17

By definition, the pKa is the pH where half the the amino acid lost a proton, while the other half is still protonated.

The inflexion points are obtained by the average value of two consecutive pKas:

$$\text{Inflexion poin \# 1: } (1.82 + 6.00) / 2 = 3.91$$

$$\text{Inflexion point \# 2: } (6.00 + 9.17) / 2 = 7.59$$

The pI is defined as the pH where the net charge of the amino acid is 0. It is determined by calculating the average of the pKas preceding and following the inflexion point where the amion acid carries no net charge. For histidine, the pI will be 7,59 (inflexion point #2).

We can now easily draw the titration curve (**see next page**).

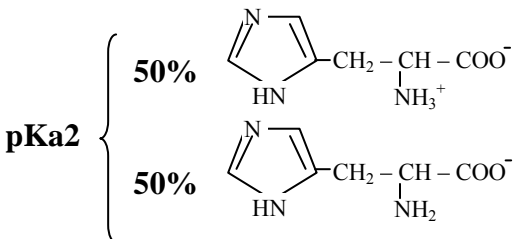
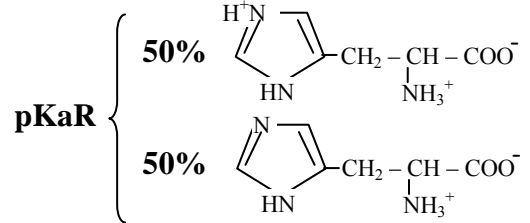
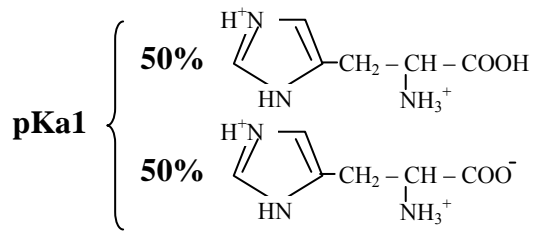
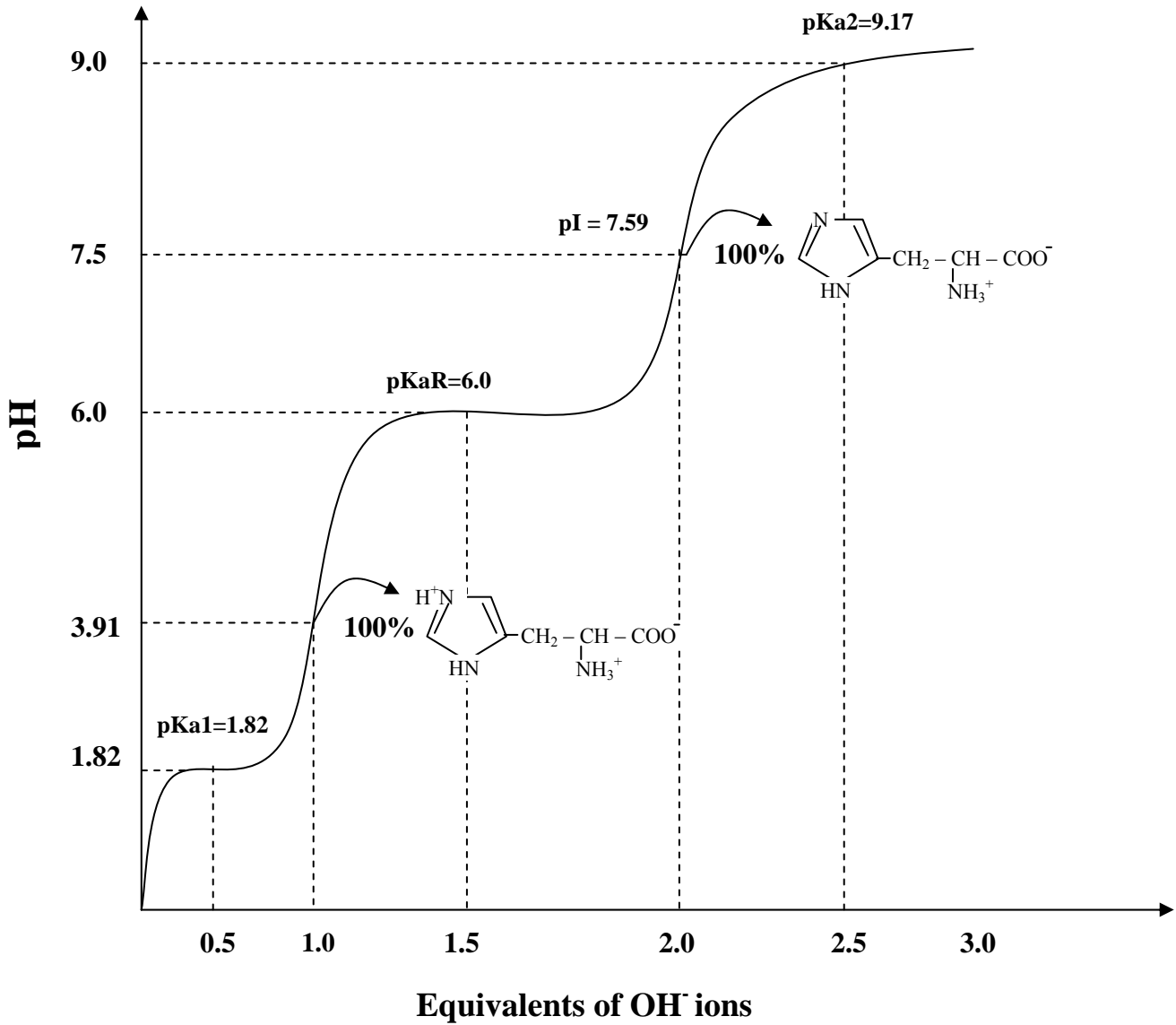
2.3. Net charge of amino acids.

To determine the net charge of amino acids at various pH values, we first need to determine the pI for each amino acid With this information, we can deduce the charge of the amino acids : by definition, the amino acid carries no net charges with the pH equals the pI. When the pH is lower than the pI, the amino acid will carry a net positive charge. When the pH is superior to the pI, then the amino acid carries a net negative charge.

Therefore, we obtain the following:

<i>pH</i>	<i>Glycine</i> (<i>pI: 5.97</i>)	<i>Serine</i> (<i>pI: 5.68</i>)	<i>Aspartic Acid</i> (<i>pI: 2.77</i>)	<i>Glutamine</i> (<i>pI: 5.65</i>)	<i>Arginine</i> (<i>pI: 10.76</i>)
2.01	+	+	+	+	+
3.96	+	+	-	+	+
5.68	+	0	-	0	+
10.76	-	-	-	-	0

Problem 2.2.



2.4. Ion exchange chromatography.

We start the chromatography at pH 10. At this particular pH, all the amino acids in the mixture are negatively charged (the pH greater than the pI of each amino acid).

a) using a *cation* exchange chromatography, none of the amino acids will stick to the resin and they will all be found in the eluent.

b) using an anion exchange chromatography, all the amino acids will bind the resin. As we decrease the pH, the amino acids will progressively elute when the pH of the buffer becomes lower than their pI. We will collect the amino acids in this order:

1- lysine

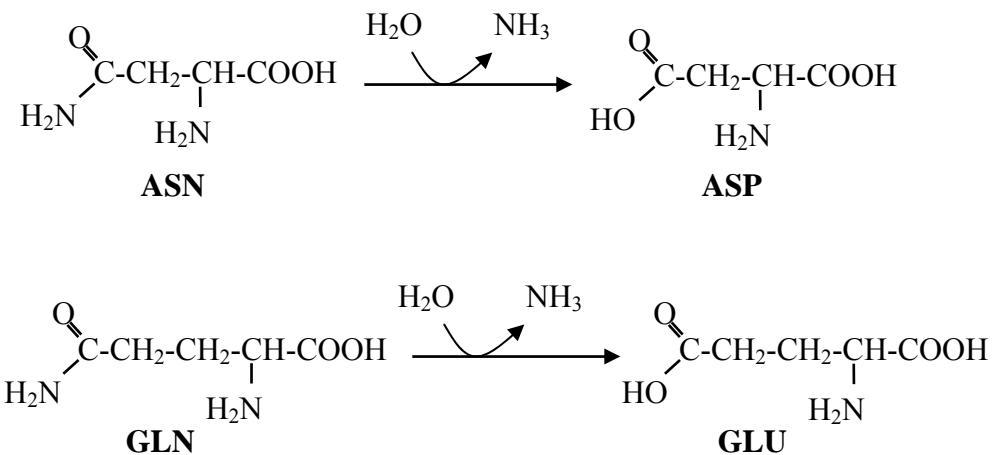
2- isoleucine, alanine, glycine (all three will elute pretty much at the same time since their pI is similar)

3- glutamic acid.

Obviously, the anion exchange resin is the best choice for this experiment.

2.5. Amino acids

Only two amino acids can be converted to other amino acids and at the same time generate ammonia : asparagine and glutamine:



*2.6. Amino acids

This observation can only be explained if we assume that serine is first inserted into casein, and that a phosphate group is then added once protein synthesis is over, producing phosphoserine. The experimental data confirm this hypothesis.

*2.7. Ion exchange chromatography.

Since these four amino acids can be separated by ion exchange chromatography, but that their pI are virtually identical, another physico-chemical property must be responsible for this behaviour. A close look at the structure of glycine, alanine, valine and leucine reveals a progressive increase in the hydrophobic character of their side chain. We can deduce that these amino acids can establish hydrophobic interactions with the ion exchange resin, allowing their separation.

2.8. Peptides.

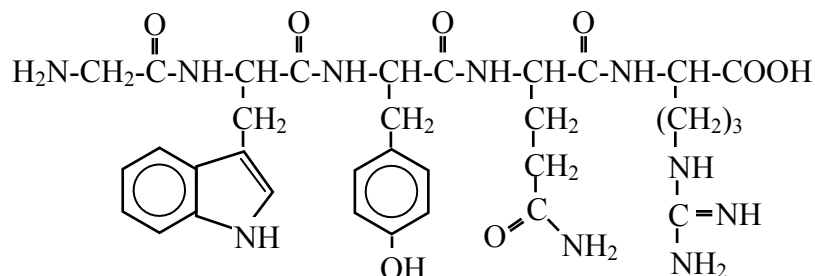
Since the yield in leucine, arginine and tryptophan is similar, we can conclude that they are present in equal proportions in the polypeptide. Twice this amount was obtained as glutamate. Finally, alanine and lysine were recovered in proportions corresponding to four times the amount of arginine/leucine/tryptophan. Therefore, the empirical formula of this peptide would be:



This experiment does not allow us to determine the amino acid sequence of this peptide.

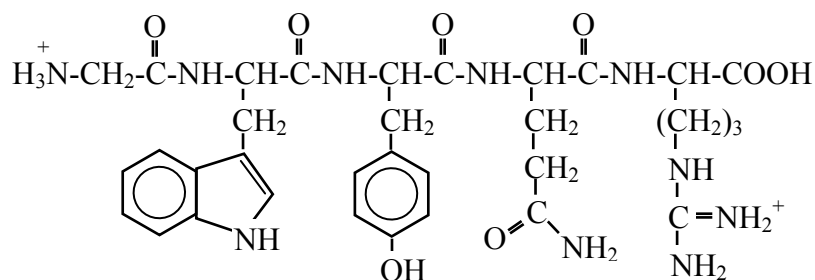
2.9. Peptides.

The structure of the peptide GWYQR (Glycyl-Tryptophanyl-Tyrosyl-Glutaminyl-Arginine) is:

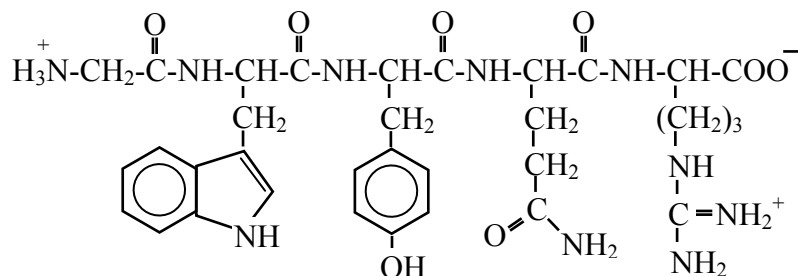


To determine the form of this peptide at each pH, all we have to do is note the pKa values of all the charged groups (including the side chains), and to deduce the ionization status of each of these groups (pH < pKa: protonated; pH > pKa: non-protonated):

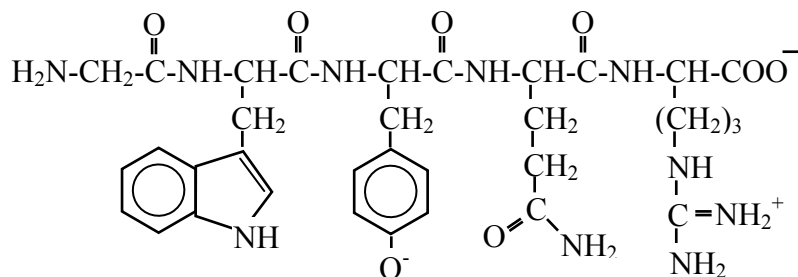
At pH 2: Every group is protonated:



At pH 7: The carboxyl group of arginine is ionized ($\text{pH} > \text{pK}_a$). However, the side chain of arginine and the amino group of glycine remain protonated ($\text{pH} < \text{pK}_a$):



At pH 10.5: Here, the amino group of glycine and the side chain of tyrosine are deprotonated ($\text{pH} > \text{pK}_a$). However, the amino group of the side chain of arginine remains protonated ($\text{pH} < \text{pK}_a$):



Chapter 3. General properties and purification of proteins

3.1. Protein purification.

Ammonium sulphate precipitation enables the concentration of our favorite protein by the non-specific precipitation of a large proportion of the proteins of the extract. We can therefore easily obtain a partial purification of our favorite protein, which can then be purified further using other methods.

3.2. Protein purification.

The diethylamino group ($-\text{CH}_2-\text{CH}_2-\text{NH}^+-\text{CH}_2-\text{CH}_3$) of DEAE-cellulose carries a positive charge which is responsible for the ion-binding properties of this resin. Effectively, negatively charged amino acids/proteins will interact with the diethylamino group (via electrostatic interactions), while positively charged amino acids/proteins will be eluted. Since the diethylamino group has a pK_a close to 8.5, it will be deprotonated at pH values above 8.5 and will use all ability to bind

negatively charged molecules.

3.3. Protein purification.

At a pI above 6, 6-phosphogluconate dehydrogenase has a net negative charge ($\text{pH} > \text{pI}$): it will bind the resin. At a pH value above 9, the diethylamino group of the resin is deprotonated, preventing any separation of the enzyme as a function of its charge.

3.4. Protein purification.

No because CM-cellulose (CM = carboxymethyl = $-\text{CH}_2-\text{COOH}$) is a cation-exchange resin: at a pH above 6, 6-phosphogluconate dehydrogenase is negatively charged (see problem 3.3) and will not bind the resin.

3.5. Protein purification.

In order to separate 6-phosphogluconate dehydrogenase using a CM-cellulose column, we must ensure that the protein has a net positive charge. The pH of the buffer will have to be below the protein's pI, therefore below a value of 6.

3.6. Protein purification.

The biochemist has 2 choices when it comes to eluting proteins bound to ion exchange resins: use a pH gradient, or a salt gradient (usually NaCl).

In the situation where a salt gradient is chosen, one would usually start with a buffer of low NaCl concentration (i.e. low ionic strength), and then progressively introduce a buffer with a greater and greater salt concentration. The Na^+ or Cl^- ions will elute proteins from the column by neutralizing the negative or positive charges on the proteins which interact with the resin. The end result is the elution of proteins as a function of their charge density : to be eluted from a cation exchange resin, those proteins with less positive charges will require less Cl^- ions (thus, a lower NaCl concentration) to neutralize them than those proteins with more positive charges. The same logic can be used to explain the ability of Na^+ ions to elute proteins bound to anion exchange resins (e.g. CM-cellulose).

Regarding question 3.6, we can conclude that:

- myoglobin will elute first, since it will not bind to the resin ($\text{pH} < \text{pI}$: positive net charge);
- the pI of urease (5,4) is closer to the pH of the buffer that is the pI of ovalbumin (4,6): urease will therefore carry a lower number of negative charges than ovalbumin. We can therefore predict that urease will elute at a NaCl concentration which will be lower than the one required to elute ovalbumin.

The order of protein elution will therefore be: myoglobine, urease, and ovalbumin.

3.7. Protein purification.

A fast and simple way to separate these proteins is to first perform an ion exchange chromatography: it will then be possible to get rid to the protein whose pI is different that our favorite enzyme. As a second step, we can perform a molecular sieve (or size exclusion, that's the same thing) chromatography to separate our enzyme (M_r 24 kDa) from the other

contaminating protein (Mr 100 kDa).

Note: Similar results would be obtained if one were to first perform the size exclusion chromatography and then the ion exchange chromatography.

3.8. Protein purification.

a) **Specific activity** is defined as the enzymatic activity per mg protein and is an indication of the relative concentration of the enzyme in the solution (the greater the amount of enzyme in the solution *relative to all other proteins*, the greater the specific activity will be). All we have to do is to divide the enzymatic activity obtained at each step by the corresponding protein concentration. For example, for the heat treatment step, we get:

$$\text{specific activity} = \text{enzymatic activity} / \text{mg protein}$$

$$\text{specific activity} = 1\,980\text{ U}/16\,500\text{ mg} = \mathbf{0,12\text{ U/mg}}$$

The **percent yield** is defined as the amount of enzyme (or protein) recovered at each step with reference to the amount present at the start of the purification procedure. This value is obtained by dividing the enzymatic activity at step X by the initial enzymatic activity. Again, for the heat treatment step, we get

$$\text{Percent yield} = \text{Enzyme activity step X} / \text{Initial enzyme activity}$$

$$\text{Percent yield} = (1\,980\text{ U}/2\,700\text{U}) \times 100 = \mathbf{73,3\%}$$

The **degree of purification** is obtained by dividing the specific activity after step X by the initial specific activity. We therefore obtain the fold increase in enzyme purification after the purification procedure. For the heat treatment step, we get:

$$\text{Degree of purification} = \text{Specific activity step X} / \text{Initial specific activity}$$

$$\text{Degree of purification} = 0,12/0,039 = \mathbf{3,07\text{ fold.}}$$

The results for each step of the purification procedure are shown in the table below:

Purification step	Specific activity (U/mg)	Percent yield	Degree of purification (fold increase)
Cell extract	0,039	100%	---
Ammonium sulphate	0,09	85,2%	2,30
Heat treatment	0,12	73,3%	3,07

<i>Purification step</i>	<i>Specific activity (U/mg)</i>	<i>Percent yield</i>	<i>Degree of purification (fold increase)</i>
DEAE chromato.	4,31	62,2%	110,50
CM-cellulose chromato.	28,72	50%	736,40
Bio-Gel A	32,00	41,5%	820,50

b) To determine which purification step was the most effective, all one has to do is divide the value for the degree of purification of step X by the value for the preceding step. Thus, the DEAE chromatography step was the most efficient, with a 36 fold increase in enzyme purity.

c) Considering that the protein is pure after molecular sieve chromatography, 35 mg of 6-gluconate dehydrogenase were obtained. This corresponds to 41,5 % of the amount of enzyme initially present in the extract (refer to percent yield). Thus, in the initial extract we had:

$$35 \text{ mg} / 41,5\% = 75,9 \text{ mg}$$

And this amount of enzyme was initially present in an extract containing a total of 70 000 mg proteins. Thus, in the initial extract we had:

$$(75,9 \text{ mg} / 70\,000 \text{ mg}) \times 100 = \mathbf{0,108 \% \text{ of 6-gluconate dehydrogenase}}$$

3.9. Protein purification.

The aim of isoelectric focusing is to separate proteins according to their pI, thus according to their charge at different pH values. Adding SDS to the protein sample would give all proteins the same charge density and would prevent their separation by this type of electrophoresis.

3.10. Protein purification.

To determine the molecular mass of our unknown protein, we must first draw the graph of the elution profile of our standards as a function of the log of their molecular mass (see graph on following page).

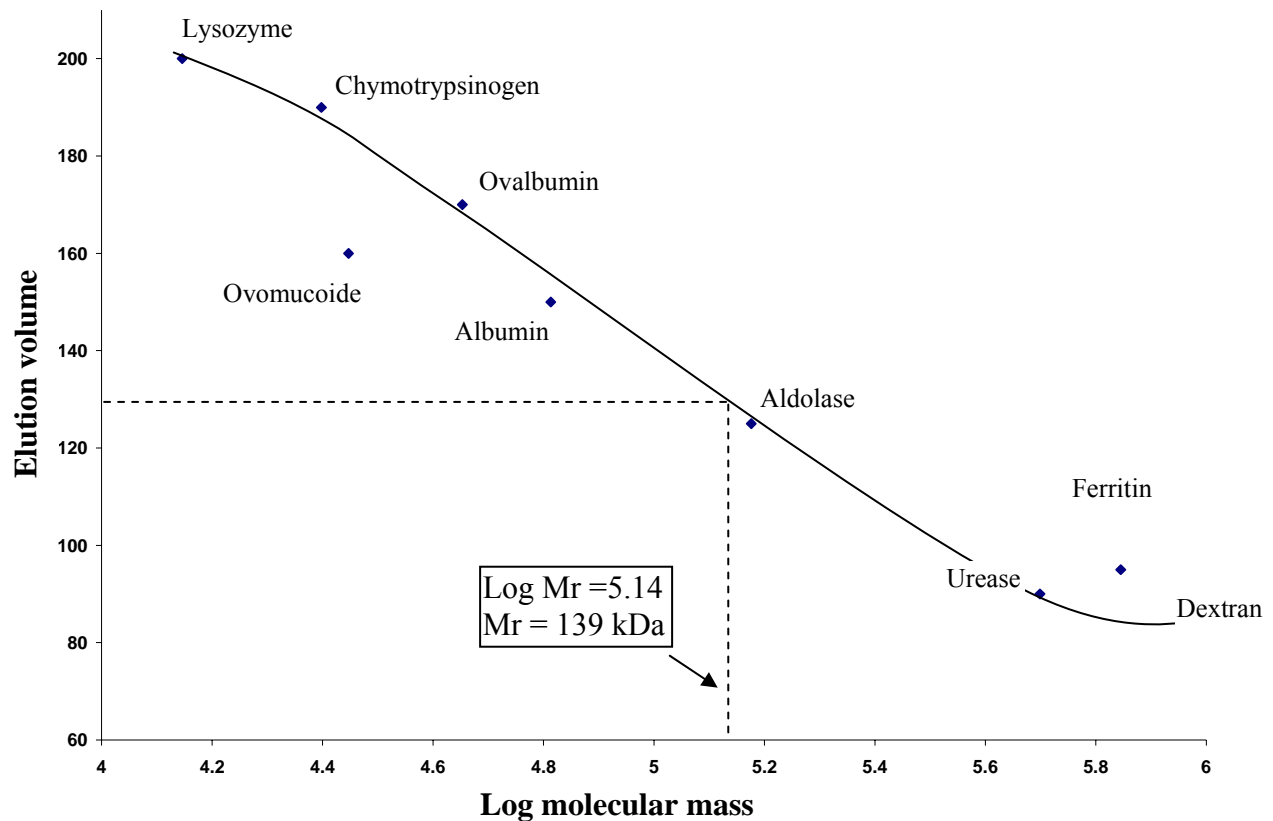
Protein	Log Mr	V_{el} (ml)
dextran blue	6.000	85.00
Lysozyme	4.146	200.00
Chymotrypsinogen	4.398	190.00
Ovalbumin	4.653	170.00
serum albumin	4.813	150.00
aldolase	5.176	125.00

Protein	Log Mr	V _{el} (ml)
urease	5.699	90.00
Ferritin	5.845	92.00
ovomucoide	4.447	160.00

From this graph, we obtain a molecular mass of **139 kDa** for our unknown protein.

Note: ferritin and ovomucoid were excluded of the standard curve because they obviously behaved differently to the other proteins when subjected to molecular sieve chromatography.

Graph, problem 3.10



***3.11. Protein purification.**

Ferritin has an iron core, giving it a greater density than other proteins of similar size and influencing its behaviour when subjected to molecular sieve chromatography.

3.12. Protein purification.

In the presence of SDS, all the proteins have an identical negative charge density: this is what

allows us to use PAGE-SDS to determine the molecular mass of proteins. Therefore, two proteins of different pI but identical molecular mass will co-migrate as a single band on PAGE-SDS. However, in the absence of SDS (thus under *native* or *non-denaturing* conditions), protein migration towards the positive or negative electrodes will be driven by its net charge, in other words by the protein's pI. In this case, two proteins of identical molecular mass but different pI will give two distinct bands upon gel staining.

3.13. Protein purification

In the absence of β -mercaptoethanol, the protein's disulfide bonds remain intact. Thus, the protein will have a more compact shape and will migrate more rapidly during PAGE-SDS than the same protein whose disulfide bonds have been reduced.

3.14. Peptide sequencing.

a) Trypsin hydrolyzes the peptide bond on the carboxyl-side of the basic amino acids lysine and arginine. Therefore, every peptide fragment generated by trypsin will have Arg or Lys at their C-terminus (with the exception, of course, of the fragment corresponding to the C-terminal end of the peptide). Therefore, the peptide given in this example will give us the following fragments:

A-L-K M-P-E-Y-I-S-T-D-Q-S-N-W-H-H-R

b) Pepsin hydrolyzes the peptide bond on the N-terminal side of the aromatic amino acids Phe, Trp, and Tyr. Therefore, the fragments obtained after pepsin digestion will all contain Tyr, Phe or Trp at their N-terminus (with the notable exception of the fragment corresponding to the N-terminus of the initial peptide). Using the peptide shown here, we obtain the following fragments:

A-L-K-M-P-E Y-I-S-T-D-Q-S-N W-H-H-R

c) Protease V8 hydrolyzes the peptide bond on the C-terminal side of the acidic amino acids Asp and Glu. Therefore, every peptide fragment generated by protease V8 will have Asp or Glu at their C-terminus (with the exception, of course, of the fragment corresponding to the C-terminal end of the peptide). Therefore, the peptide given in this example will give us the following fragments:

A-L-K-M-P-E Y-I-S-T-D Q-S-N-W-H-H-R

d) Cyanogen bromide hydrolyzes the peptide bond on the C-terminal side of Met. Therefore, every peptide fragment generated by cyanogen bromide will have Met at their C-terminus (with the exception, of course, of the fragment corresponding to the C-terminal end of the peptide). Therefore, the peptide given in this example will give us the following fragments:

A-L-K-M P-E-Y-I-S-T-D-Q-S-N-W-H-H-R

3.15. Peptide sequencing

Digesting with carboxypeptidase A tells us that the C-terminal residue of the peptide is Ala.

Digesting with trypsin allows us to partially order two of the four fragments (remember: trypsin generates fragments whose C-terminal end is Arg or Lys):

Ala-Arg (Phe, Ser)-Lys

Furthermore, since digesting with trypsin generates a free Lys residue, this indicates that this Lys is on the C-terminal side of either Arg or Lys.

Trypsin digestion also indicates that the tripeptide (Ala, Met, Ser) is the C-terminus of the peptide (it doesn't end with Arg or Lys). Furthermore, CNBr digestion allows us to determine the position of Met in this tripeptide:

Met-(Ala, Ser)

Since Ala is the C-terminal residue of this peptide (see digestion with carboxypeptidase A), the sequence of the last 3 amino acids of the polypeptide will be:

Met-Ser-Ala

Thermolysin cuts the peptide bond on the N-terminal side of hydrophobic amino acids: we can therefore deduce the position of the hydrophobic amino acids of the two fragments:

Ala-(Arg, Ser) and Phe-(Lys, Lys,)-Met-Ser

With this information, and considering the trypsin digestion pattern, we can conclude that the sequence of this polypeptide is:

Ala-Arg-Ser-Phe-Lys-Lys-Met-Ser-Ala

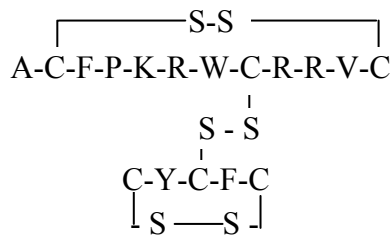
3.16. Peptide sequencing

Thermolysin cuts the peptide bond on the N-terminal side of hydrophobic amino acids. Digesting the two fragments after reduction of the disulphide bonds gives:

fragment 1: A-C F-P-R-K W-C-R-R V-C

fragment 2: C Y-C F-C

Since the disulfide bonds of the peptide were intact when the peptide was digested with thermolysin, some of the peptide fragments shown above will be linked together with disulphide bonds involving Cys residues. From the fragments obtained, we can deduce the position of the disulphide bonds as follows :



Note: Don't forget that both inter-chain and intra-chain S-S- bonds can be present in the molecule.

3.17. Peptide sequencing

Digestion with carboxypeptidase tells us that the C-terminus is Lys.

DNFB treatment indicates that the N-terminus is Val.

Trypsin digestion allows us to partially order the peptides, as follows:

peptide C: Try-Ala-Lys peptide D: Val-(Ala, Ala, Ala, Pro)-Lys (remember: Val = N-terminus)

peptide E: Met-(Asp, Gly)-Arg

We can order the rest of the residues with the results from CNBr treatment:

Met-Gly-Asp-Arg

Finally treating peptide D with thermolysine allows us to order the three Ala and the Pro:

Val-Ala-Ala-Ala-Pro-Lys

Therefore, the sequence of the peptide is:

Val-Ala-Ala-Ala-Lys-Pro-Met-Gly-Asp-Arg-Try-Ala-Lys

Chapter 4. Three dimensional structures of proteins

4.1. 3-D structure of proteins

Generally speaking, hydrophobic amino acids are found buried inside proteins (away from water), while polar and charged amino acids are most often found on the surface of proteins. We will then get the following distribution for the amino acids :

Buried inside: Val, Phe, Ileu

On the surface: Glu, Arg, Asn, Lys, Ser, Thr

****4.2. 3-D structure of proteins***

The structure of urea suggests that this molecule denatures proteins by breaking the hydrogen interactions which stabilize the 3-D structure of the macromolecules (i.e. via the interaction of the amino and ketone groups of urea with the amino and ketone groups of the peptide bonds and side chains).

4.3. 3-D structure of proteins

Even though it is located at the surface of proteins, Phe must avoid contact with water. This can be accomplished if two or more protein subunits interact via hydrophobic regions (which could include Phe), keeping Phe in an hydrophobic environment.

****4.4. 3-D structure of proteins***

The presence of Asp inside proteins is possible if its polar and charged groups are involved in intermolecular interactions. This is possible if Asp is part of a secondary structure like the α -helix.

4.5. 3-D structure of proteins

Protein 1 has a high amount of hydrophilic amino acids, and very little hydrophobic residues (65% hydrophilic/35% hydrophobic). This suggests that a lot of these amino acids will be interacting with the solvent, which is the case for rod-shaped proteins (i.e. protein A).

Protein 3 has the same amount of hydrophobic vs hydrophilic amino acids, while protein 2 has many more hydrophobic than hydrophilic residues (protein 2 = 30%/70%). This suggests that protein 3 would be globular, with several hydrophobic amino acids buried inside and lots of hydrophilic amino acids on the surface. Protein 3 would therefore be protein B.

As for protein 2, its high content in hydrophobic amino acids little content in hydrophilic amino acids suggest that it could be protein C: the association of several subunits would regions of high content of hydrophobic amino acids, protecting them from the aqueous environment.

4.6. 3-D structure of proteins

The primary structure of this peptide doesn't have any features expected from a β -pleated sheet, and no Gly and Pro (which are known to disrupt secondary structures) are present. We can therefore deduce that this peptide would adopt an α -helical structure.

4.7. 3-D structure of proteins

The primary structure of this peptide is typical for those arranged as β -pleated sheets.

4.8. 3-D structure of proteins

The presence of several positively charged residues (Arg and Lys) in addition to Gly indicates that this peptide will most likely be a random coil.

4.9. 3-D structure of proteins

With the criterias use in the 3 preceding problems, we can deduce that:

Amino acids 2-12: α -helix;

Amino acids 13-17: random coil;

Amino acids 18-26: β -pleated sheet

4.10. 3-D structure of proteins

The high amount of Pro and Hypro indicates that protein C will be a collagen-like triple-helix.

Protien A is rich in amino acids with small side-chains (Gly, Ser, Ala): it will adopt a β -pleated sheet type of structure.

Protein B has a lot of amino acids that would be expected to be found in α -helices. However, the integrity of this helix would be severely perturbed by the presence of Gly and Pro and by the presence of several consecutive acidic or basic amino acids.

Chapter 5. Enzymology

5.1. Enzyme kinetics

a) From the available data, we can notice that the reaction rate doesn't increase when the substrate concentration is over 2×10^{-3} M. This is the maximal velocity (V_{max}) of the enzyme, in this case $60 \mu\text{mol}/\text{min}$.

b) v is constant at a $[S]$ above 2×10^{-3} M because the substrates is saturating the enzyme.

c) Since the maximal velocity is achieved at a $[S]$ of 2×10^{-2} M, almost all the enzyme is part of an enzyme/substrate complex. The amount of enzyme free in solution is then negligible.

5.2. Enzyme kinetics

a) $V_{max} = 0,25 \mu\text{mol}/\text{min}$;

b) K_m can be determined using the Michaelis-Menten equation:

$$v = \frac{[S] V_{max}}{[S] + K_m}$$

$$v[S] + vK_m = [S]V_{max}$$

$$vK_m = [S]V_{max} - v[S]$$

$$vK_m = [S] (V_{max} - v)$$

$$K_m = \frac{[S] (V_{max} - v)}{v}$$

Using the data for a [S] of 5×10^{-6} :

$$K_m = \frac{5 \times 10^{-6} \text{ M} \times (0.25 \text{ } \mu\text{mol/min} - 0.071 \text{ } \mu\text{mol/min})}{0.071 \text{ } \mu\text{mol/min}}$$

$$K_m = 1.26 \times 10^{-5} \text{ M}$$

Note: Similar data would be obtained if a different [S] is chosen, as long as $v < V_{\max}$.

c) The initial velocity can be obtained using the Michaelis-Menten equation:

- for [S] = 1×10^{-6} M:

$$v = \frac{[S] V_{\max}}{[S] + K_m}$$

$$v = \frac{1 \times 10^{-6} \text{ M} \times 0.25 \text{ } \mu\text{mol/min}}{1 \times 10^{-6} \text{ M} + 1.26 \times 10^{-5} \text{ M}}$$

$$v = 0.0184 \text{ } \mu\text{mol/min}$$

- for [S] = 1×10^{-1} M: $v = 0.25 \text{ } \mu\text{mol/min}$ (saturating [S]: $v = V_{\max}$).

d) For a [S] of 2×10^{-3} M, the initial velocity will be equal to V_{\max} . We therefore get:

$$v = V_{\max} = 0.25 \text{ } \mu\text{mol/min}$$

After 5 min, we get: $0.25 \text{ } \mu\text{mol/min} \times 5 \text{ min.} = 1.25 \text{ } \mu\text{mol}$ of product.

- For a [S] of 2×10^{-6} M, we must first find the initial velocity using the Michaelis-Menten equation:

$$v = \frac{[S] V_{\max}}{[S] + K_m}$$

$$v = \frac{2 \times 10^{-6} \text{ M} \times 0.25 \text{ } \mu\text{mol/min}}{2 \times 10^{-6} \text{ M} + 1.25 \times 10^{-5} \text{ M}}$$

$$v = 0.035 \text{ } \mu\text{mol/min}$$

After 5 minutes of reaction, we get:

$$v = 0.035 \text{ } \mu\text{mol/min} \times 5 \text{ min.} = 0.175 \text{ } \mu\text{mol}$$
 of product.

e) Since K_m is independent of enzyme concentration, it will not be affected by the increase in $[E]$ and will remain equal to 1.25×10^{-5} M.

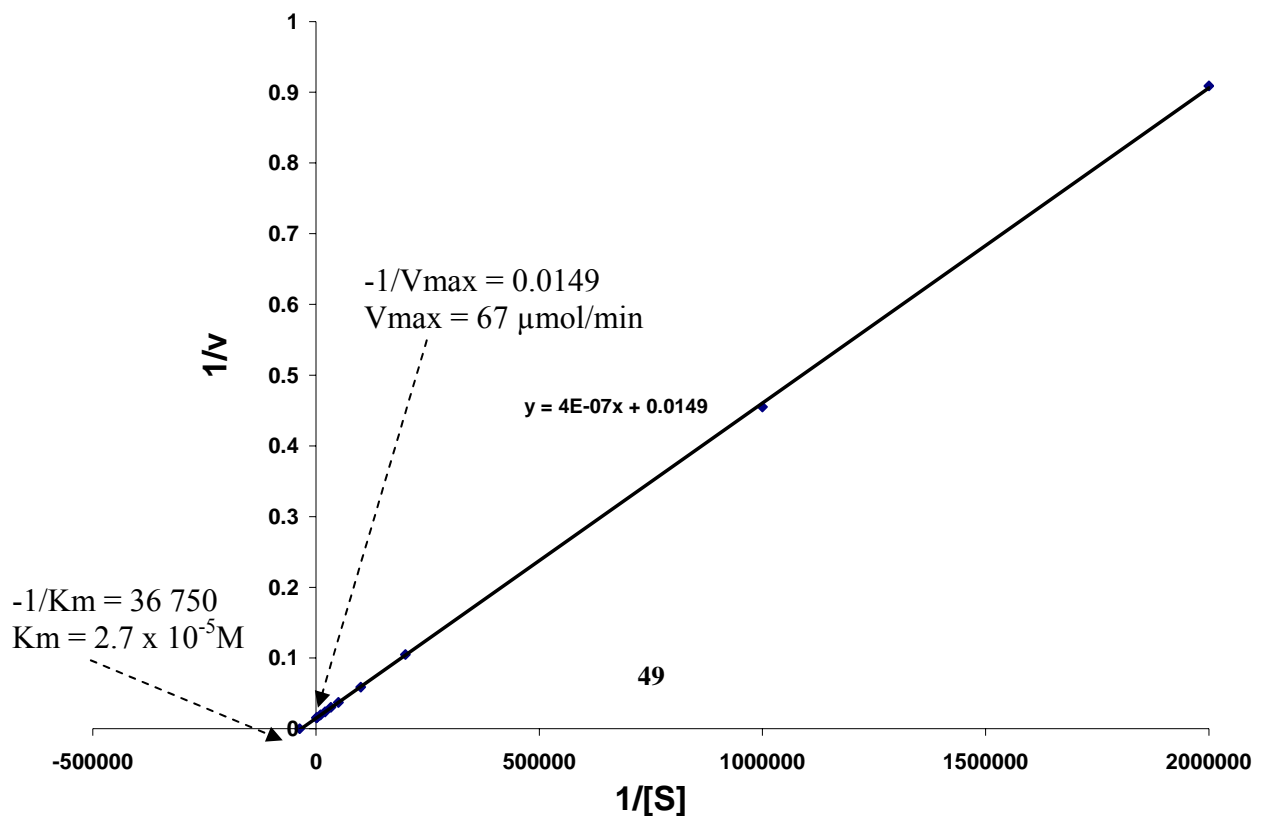
However, V_{max} will be changed because: $V_{max} = k_{cat} \times [E_t]$. Therefore, since k_{cat} is a constant, the 4 fold increase in $[E]$ will also multiply the V_{max} by 4. Therefore, $V_{max} = 1 \mu\text{mol}/\text{min}$.

5.3. Enzyme kinetics

The Lineweaver-Burke plot is graph of the reciprocal of the initial velocity ($1/v$) as a function of the reciprocal of the substrate concentration ($1/[S]$). From the available data, we get:

$1/[S] \text{ M}^{-1}$	$1/v \text{ (mmol}^{-1} \times \text{min.)}$
1000	0.0154
2000	0.0158
10 000	0.0196
20 000	0.0238
33 333	0.0303
50 000	0.0370
100 000	0.0588
200 000	0.1050
1 000 000	0.4550
2 000 000	0.9090

The Lineweaver-Burke plot is shown below. The K_m is easily found as the reciprocal value of the x-intercept. In this case: $K_m = 2.7 \times 10^{-5}$ M. As for V_{max} , we can obtain its value by the reciprocal of the intersection with the y axis. Here, $V_{max} = 67 \mu\text{mol}/\text{min}$.



5.4. Enzyme kinetics

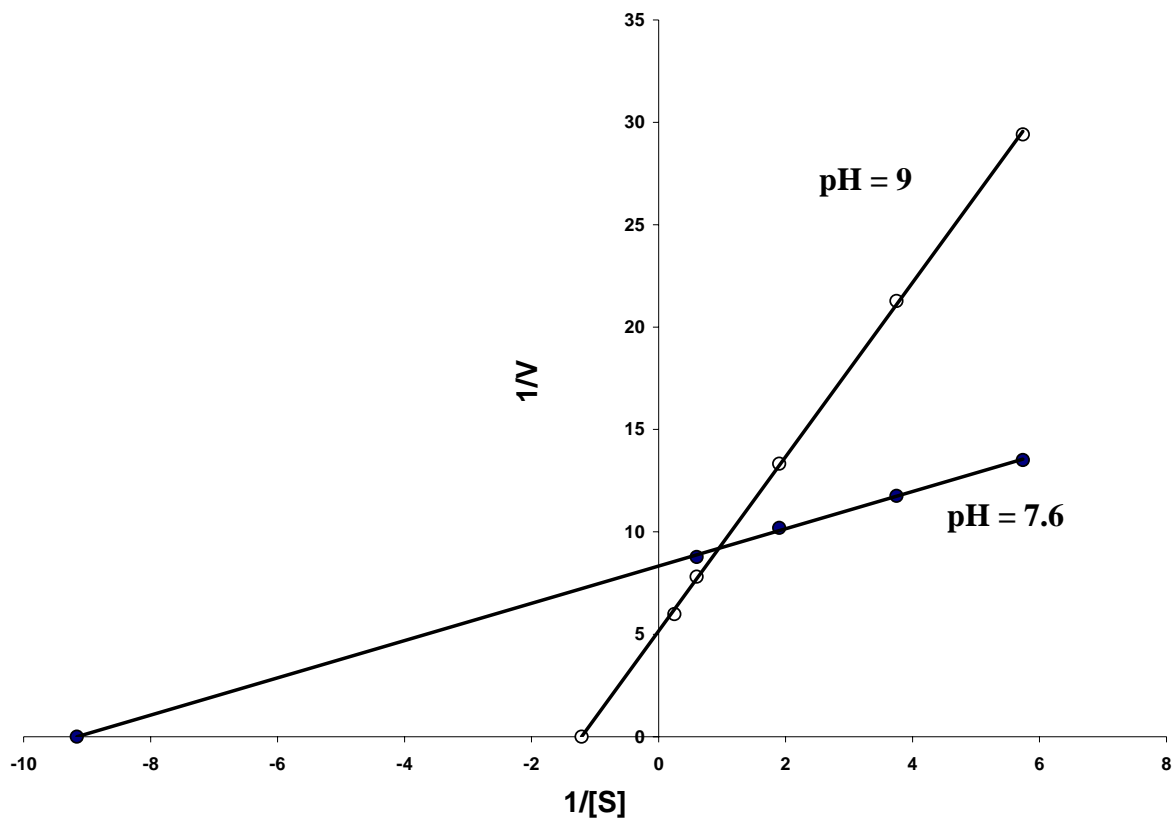
To determine the affinity of the enzyme for its substrate, we must find the value of the Michaelis-Menten constant. This is easily done using the Lineweaver-Burke plot:

$1/[S]$ ($M^{-1} \times 10^{-4}$)	$1/v$ ($\mu\text{mol}^{-1} \times \text{min.}$) (pH 7.6)	$1/v$ ($\mu\text{mol}^{-1} \times \text{min.}$) (pH 9.0)
5.74	13.51	29.41
3.75	11.76	21.28
1.90	10.20	13.33
0.60	8.77	7.81
0.25	-	5.99

The Lineweaver-Burke plot is shown below.

The intersection with the x axis gives us $-1/K_m$. From the graph, we can see that the smaller the value of $-1/K_m$, the greater K_m will be. K_m is a measure of the affinity of the enzyme for its substrate, and is equivalent to the amount of substrate required to reach $1/2 V_{\text{max}}$. Therefore, when K_m is high, the enzyme has a low affinity for its substrate: a lot of substrate is needed to achieve $1/2 V_{\text{max}}$.

From the graph, we can easily determine that the value of K_m is greater at pH 9 than pH 7.6. Therefore, the enzyme will have a greater affinity for its substrate at pH 7.6.



5.5. Enzyme kinetics

To solve this problem, we must first draw a Lineweaver-Burke plot:

$1/[S]$ (M^{-1})	$1/v$ ($mmol^{-1} \times min.$) No inhibitor	$1/v$ ($mmol^{-1} \times min.$) With inhibitor
10 000	0.0357	0.0588
6 666.67	0.0277	0.0435
5 000	0.0233	0.0345
2 000	0.0154	0.0200
1 333.33	0.0135	0.0164

The Lineweaver-Burke is shown below.

a) V_{max} is given by the reciprocal of the intersection on the y axis:

$$1/V_{max} = 0.0101 \mu mol^{-1} \times min.$$

$$V_{max} = 99 \mu mol/min$$

Since the intersection on the y axis is the same in the presence than in the absence of the inhibitor, we can conclude that we are dealing with a competitive inhibition.

b) K_m is given by the negative value of the reciprocal of the intersection on the x axis, in this case:

- In the absence of inhibitor:

$$-1/K_m = 3\,800 M^{-1}$$

$$K_m = 2.63 \times 10^{-4} M$$

- In the presence of the inhibitor:

$$-1/K_{m_{app}} = 2000 M^{-1}$$

$$K_{m_{app}} = 5 \times 10^{-4} M$$

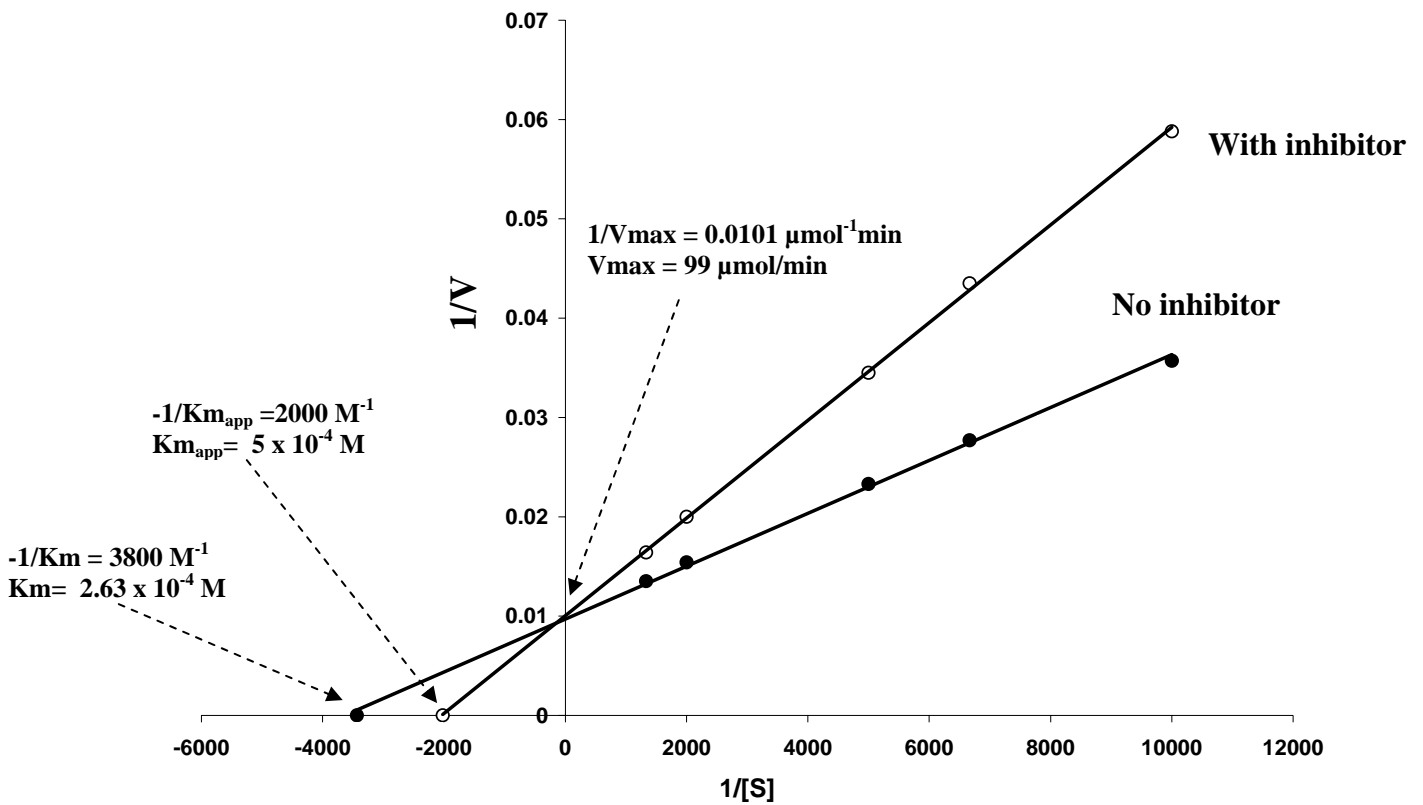
c) K_i can be found with the values determined above and the equation:

$$K_{m_{app}} = K_m + \frac{K_m [I]}{K_i}$$

$$K_i = \frac{K_m [I]}{(K_{m_{app}} - K_m)}$$

$$K_i = \frac{2.63 \times 10^{-4} \text{ M} \times 2.2 \times 10^{-4} \text{ M}}{(5 \times 10^{-4} \text{ M} - 2.63 \times 10^{-4} \text{ M})}$$

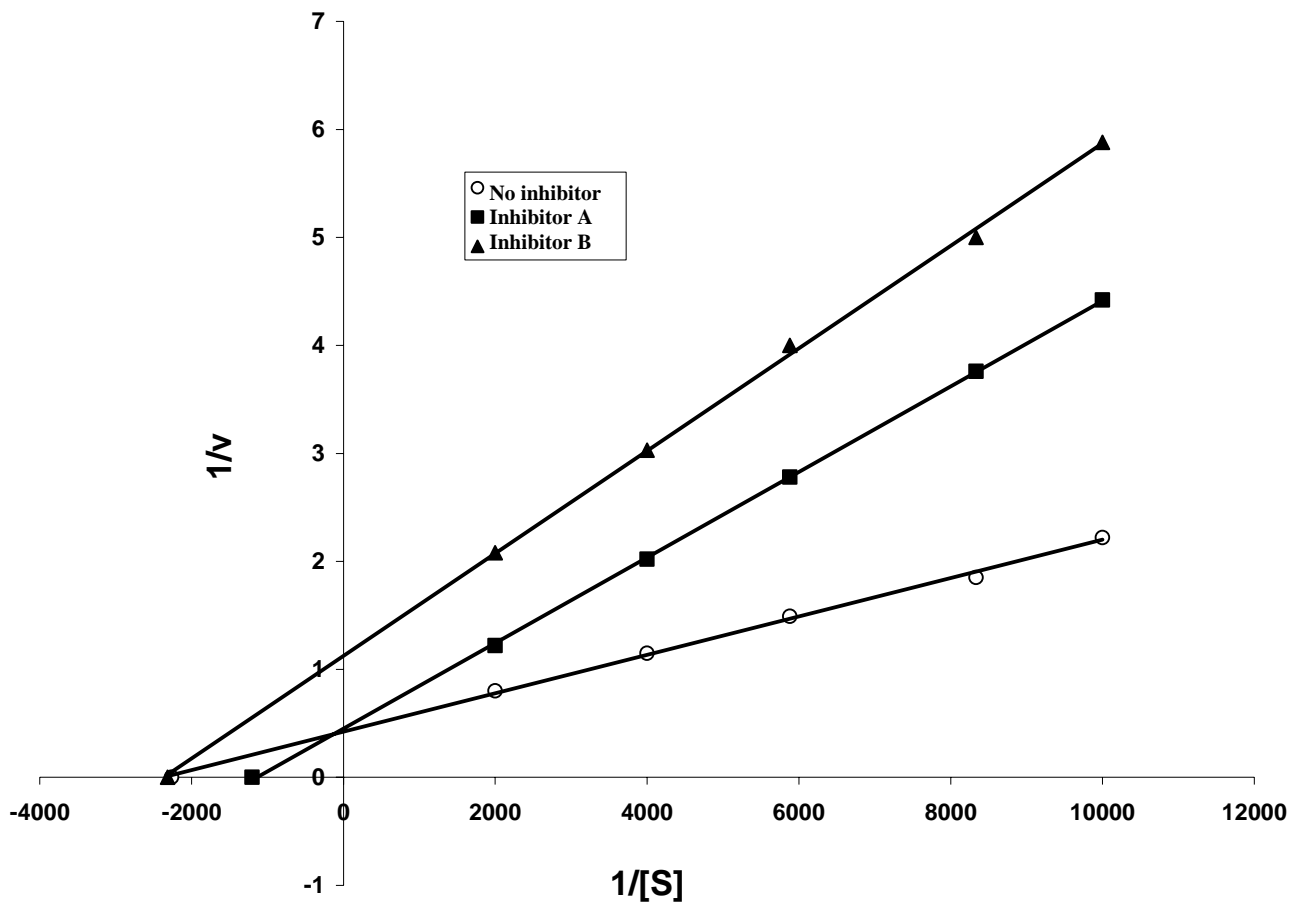
$$K_i = 2.44 \times 10^{-4} \text{ M}$$



5.6. Enzyme kinetics

a) To find K_m and V_{max} , we must first draw a Lineweaver-Burke plot (shown below):

$1/[S]$ (M^{-1})	$1/v$ ($\mu\text{mol}^{-1} \times \text{min.}$)	$1/v$ ($\mu\text{mol}^{-1} \times \text{min.}$) inhibitor A	$1/v$ ($\mu\text{mol}^{-1} \times \text{min.}$) inhibitor B
2000	0.80	1.22	2.08
4000	1.15	2.02	3.03
5882	1.49	2.78	4.00
8333	1.85	3.76	5.00
10 000	2.22	4.42	5.88



The value of V_{max} is given by the reciprocal of the intersection on the y axis:

$$V_{max} = 2.36 \mu\text{mol}/\text{min}.$$

Similarly, K_m is given by the negative value of the reciprocal of the intersection on the x axis:

$$K_m = 4.4 \times 10^{-4} \text{ M}.$$

b) From the Lineweaver-Burke plot, inhibitor A leads to a $V_{max,app}$ of 2.36 mmol/min and a $K_{m,app}$ of 8.3×10^{-4} M. This is expected from a competitive inhibitor (for example, a substrate analog).

As for inhibitor B, $V_{max,app}$ is $0,889 \mu\text{mol}/\text{min}$ $K_{m,app}$ is $4,4 \times 10^{-4}$ M. Since the K_m is identical to the K_m observed in the absence of the inhibitor, we are dealing with a non-competitive inhibition. This type of inhibition is observed when the inhibitor does not interfere with the enzyme-substrate interaction. This is expected from alkylating agents, which would ever so slightly modify the structure of the enzyme, leading to alterations to the active site and a less effective enzyme.

c) The inhibition constant for inhibitor A can be determined as follows:

$$K_{m_{app}} = K_m + \frac{K_m [I]}{K_i}$$

$$K_i = \frac{K_m [I]}{(K_{m_{app}} - K_m)}$$

$$K_i = \frac{4.25 \times 10^{-4} \text{ M} \times 5 \times 10^{-4} \text{ M}}{(1.11 \times 10^{-3} \text{ M} - 4.25 \times 10^{-4} \text{ M})}$$

$$K_i = 5.64 \times 10^{-4} \text{ M}$$

The inhibition constant for inhibitor B can be found as follows:

$$V_{max_{app}} = \frac{V_{max}}{(1 + [I]/K_i)}$$

$$K_i = \frac{V_{max_{app}} [I]}{V_{max} - V_{max_{app}}}$$

$$K_i = \frac{0.889 \text{ } \mu\text{mol/min} (3.2 \times 10^{-6} \text{ M})}{(2.36 \text{ } \mu\text{mol/min} - 0.889 \text{ mmol/min})}$$

$$K_i = 1.93 \times 10^{-6} \text{ M}$$

d) The initial velocity can be determined by a modification of the Michaelis-Menten equation for competitive inhibitors:

$$v = \frac{V_{max} \times [S]}{[S] + K_{m_{app}}}$$

$$\text{and } K_{m_{app}} = K_m + \frac{K_m [I]}{K_i}$$

$$K_{m_{app}} = 4.4 \times 10^{-4} \text{ M} + \frac{4.4 \times 10^{-4} \text{ M} \times 2 \times 10^{-5} \text{ M}}{5.64 \times 10^{-4} \text{ M}}$$

$$K_{m_{app}} = 4.56 \times 10^{-4} \text{ M}$$

Therefore:

$$v = \frac{V_{\max} \times [S]}{[S] + K_{m_{\text{app}}}}$$

$$v = \frac{2.36 \mu\text{mol}/\text{min} \times 3 \times 10^{-4} \text{ M}}{3 \times 10^{-4} \text{ M} + 4.56 \times 10^{-4} \text{ M}}$$

$$v = 0.936 \mu\text{mol}/\text{min}$$

5.7. Enzyme catalysis

The fact that the optimal pH is 8 suggests that the (de)protonation of specific amino acids is important for enzyme activity. For example, at pH below 8, protonation of the side chain of His (pKaR = 6) could alter the active site either directly or indirectly (by inducing a conformational change in the protein). At pH above 8, the deprotonation of other side chains (e.g Tyr or Lys) could have similar effects.

5.8. Enzyme catalysis

a) For an enzyme whose optimal pH is 4, we can propose that the ascending part of the curve would be attributable to the ionization of the lateral group of Asp or the C-terminal carboxyl group of the protein. For the descending part of the curve, it could be caused by the deprotonation of His or the ionization of Glu.

b) For an enzyme whose optimal pH is 11, the ascending part of the curve might be due to the deprotonation of the lateral group of Lys. For the descending portion of the curve, it could be due to the deprotonation of Arg.

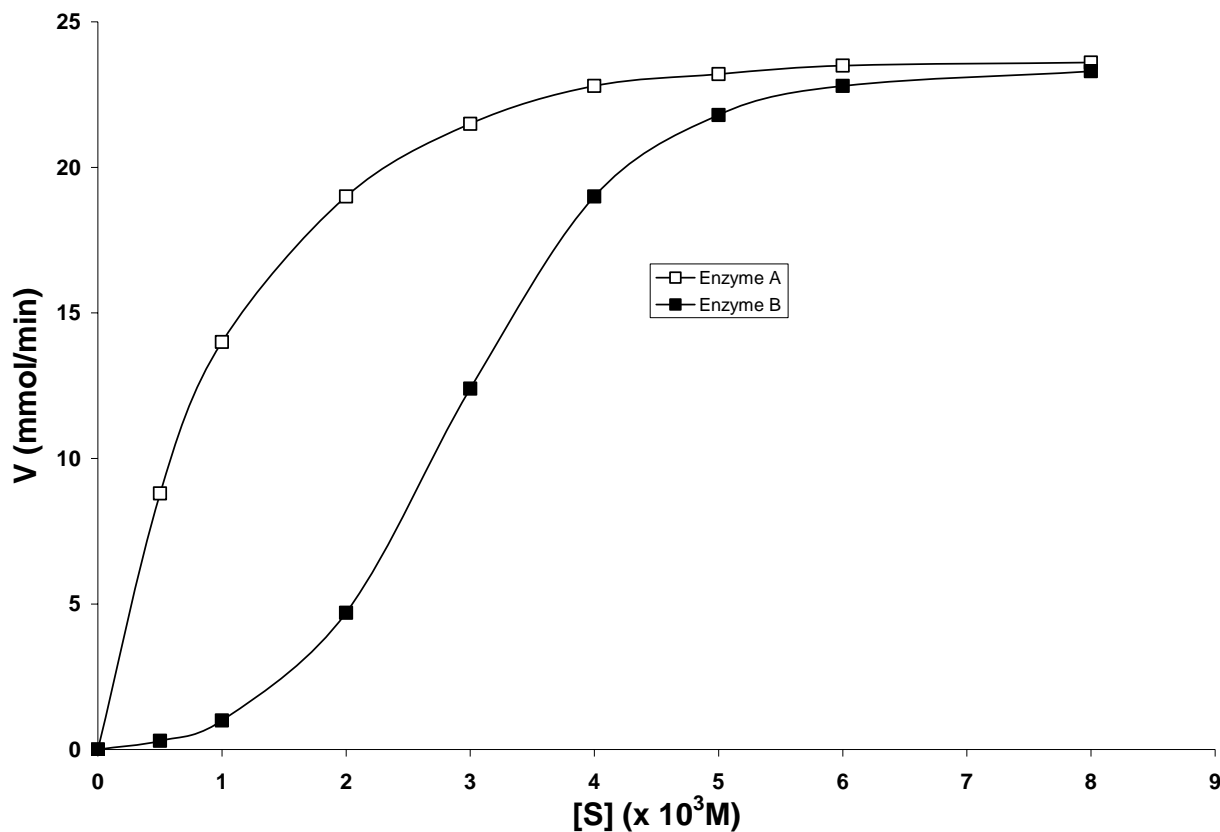
5.9. Enzyme catalysis

To be able to answer this question, we must first draw the graph of the rate of the reaction as a function of substrate concentration (shown below).

From this graph, we can conclude that enzyme A follows the classical pattern of Michaelis-Menten kinetics: the initial velocity of the reaction is only limited by the substrate concentration, and the maximal velocity is reached when the substrate is in excess.

With this graph, we can also conclude that B is an allosteric enzyme. Effectively, at low enzyme concentration, the enzyme (which presumably possesses several binding sites for the substrate) has a conformation that results in low affinity for the substrate, and the rate

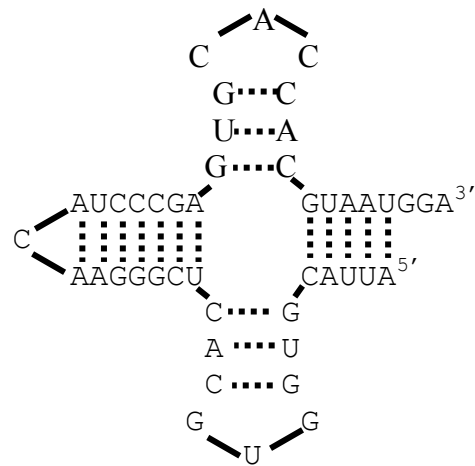
of the reaction is low. However, the binding of one molecule of the substrate causes major conformational changes in the protein that increase the enzyme's affinity for its substrate: the greater the substrate concentration, the more high affinity sites for the substrate will be available, and the greater the rate of the reaction.



Chapter 6. Structure and properties of nucleic acids.

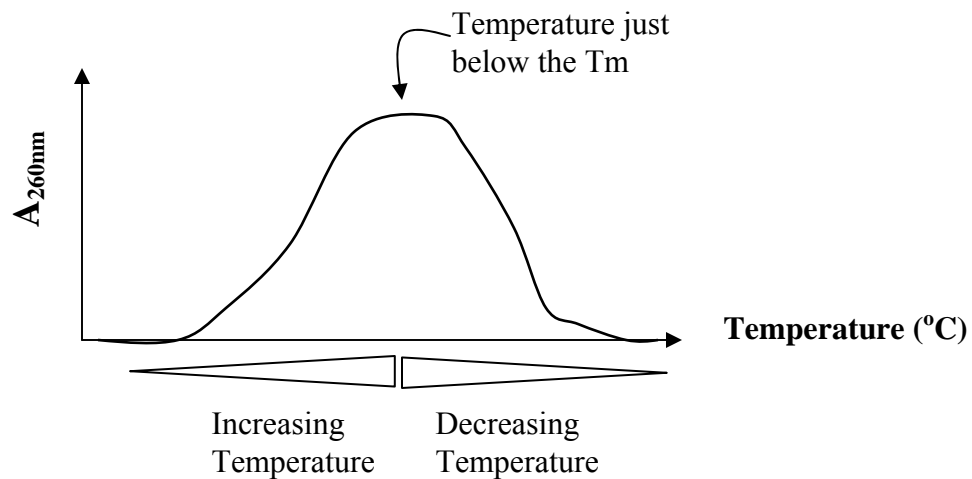
6.1. Structure of nucleic acids.

The polynucleotide shown here is an RNA molecule (notice the presence of uracil). RNA can form double-stranded structures if complementary sequences are present in the molecule. As for DNA, the more stable secondary double-stranded RNA structures will be those with the most base pairs (therefore, the most H-bonds). For the RNA molecule described here, we get these two secondary structures:

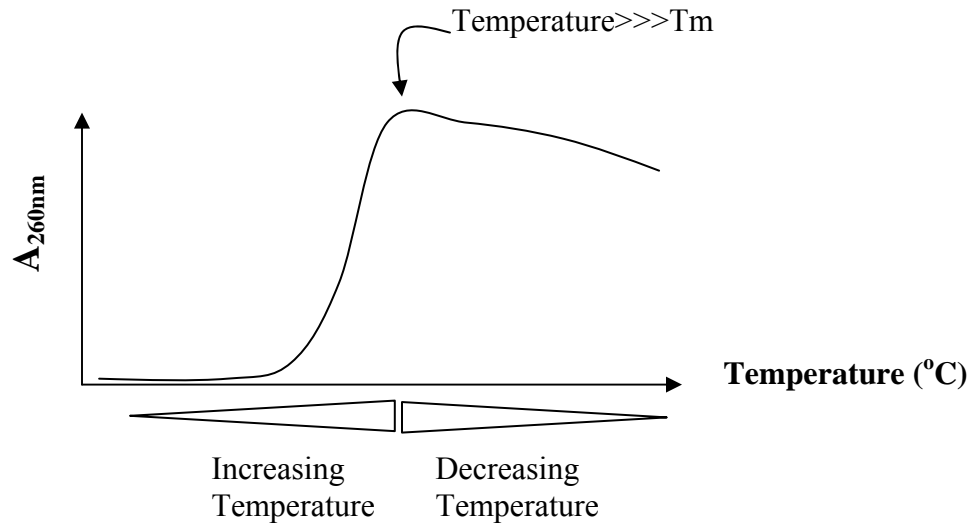


***6.2. Structure of nucleic acids.**

a) Since T_m has not yet been reached, the two polynucleotide chains are not completely separated and are still aligned according to their complementary sequence. We would obtain the following melting curve:



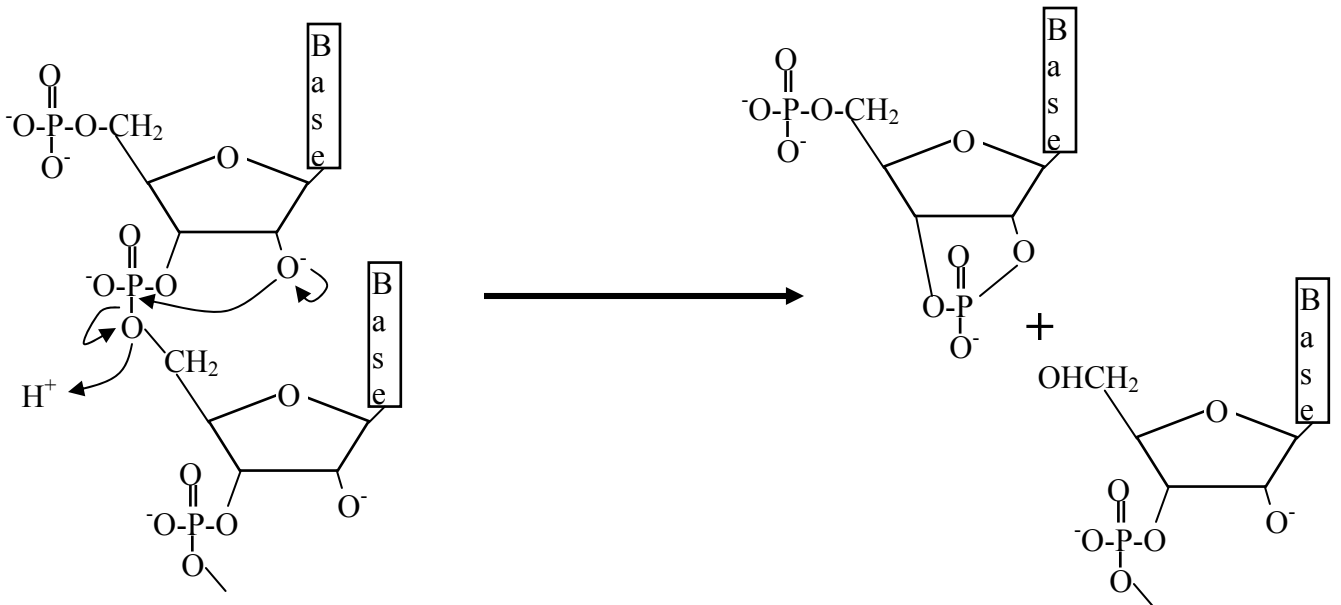
b) Since the temperature is increased way above the T_m , both polynucleotide chains have been separated. Upon cooling the temperature, both chains will have to realign to form complementary base pairs: since this is a random process, it can take a while. Since the time allotted after cooling is short (2 min), the DNA molecule will not have time to completely rehybridize, the absorbance will decrease only very slowly and we will get a curve like this one:



c) A perfectly reversible melting curve can be obtained if the alignment between complementary base pairs can readily be achieved. This is the case when polynucleotide of low complexity are used, like poly(G):poly(C), or poly (AT):poly(AT).

6.3. Structure of nucleic acids.

Nucleic acid hydrolysis involves the break of a phosphodiester bond linking each nucleotide in the 5' → 3' orientation. The only difference between RNA and DNA is the presence of a 2' hydroxyl (2' OH) group on the ribose sugar of RNA. Therefore, in the presence of a base, this group can be easily ionized, generating an O⁻ ion that can easily break the phosphodiester bond by attacking the phosphoryl group:



6.4. Structure of nucleic acids.

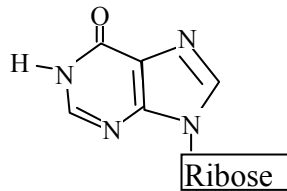
These data can be interpreted in the following fashion:

a) if we *rapidly* cool down the nucleic acid solution, the absorbance at 280nm will not vary by much, indicating that the polynucleotides remained single-stranded (in other words, the DNA molecule stays *denatured*).

b) if we slowly cool-down the nucleic acid solution, we observe a progressive but steady decline in the absorbance at 280 nm, indicating that our polynucleotide has become double-stranded (in other words: the two strands *rehybridized*).

6.5. Structure of nucleic acids.

The structure of IMP is as follows:



IMP can form H bonds with *both* adenosine and cytosine (compare the structure of IMP with AMP and CMP). Therefore, incorporating IMP into DNA would be highly detrimental to the organism by creating an ambiguity in base pairing.

6.6. Structure of nucleic acids.

a) pancreatic ribonuclease cut RNA on the 3' end of *pyrimidine* nucleotides, generating a 3' phosphate end. In the polynucleotide given in this problem, pancreatic ribonuclease will yield the following digestion products:

$5'pACp^{3'}$ $5'GAUp^{3'}$ $5'GCp^{3'}$ $5'Up^{3'}$ $5'AUp^{3'}$ $5'C^{3'}$

b) ribonuclease T2 cuts on the 3' end of adenosine, and generates a 3' phosphate end:

$5'pAp^{3'}$ $5'CGAp^{3'}$ $5'UGCUA^{3'}$ $5'UC^{3'}$

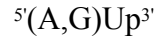
c) ribonuclease T1 cuts on the 3' end of guanosines and generates a 3' phosphate end:

$5'pACGp^{3'}$ $5'AUGp^{3'}$ $5'CUAUC^{3'}$

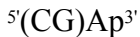
6.7. Structure of nucleic acids.

Snake venom phosphodiesterase only cuts the nucleotide at the 3' end of RNA molecules, liberating a nucleotide with a 5' phosphate. This tells us that Cp is located at the 3' end of our polynucleotide.

Pancreatic ribonuclease cuts on the 3' end of pyrimidines, generating small fragments ending with a pyrimidine with a 3' phosphate end. From the data provided, we can deduce the following nucleotide sequences:



Finally, ribonuclease T2 cuts after adenosine, generating fragments ending with adenosine 3' phosphate. We can use this information to order the sequence of the product ::



In addition, the presence of pAp after digestion with ribonuclease T2 tells us that A is at the 5' end of the polynucleotide.

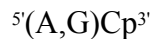
Putting together all these clues, we get the following complete sequence:



6.8. Structure of nucleic acids.

Snake venom phosphodiesterase cuts the nucleotide located at the 3' end of the fragment. The fact that we get pC tells us that this is the nucleotide at the 3' end of the nucleic acid.

Pancreatic ribonuclease cuts RNA molecules on the 3' side of pyrimidines, generating fragments ending with pyrimidine 3' phosphate. With the data at hand, we can deduce part of the nucleotide sequence, as follows:



Also, the fact that we get 2 Cp indicates that these nucleotides immediately follow U or C.

RNAse T1 cuts on the 3' side of guanosine, generating fragments ending with guanosine 3' phosphate. RNAse T2 cuts on the 3' side of adenosine, generating fragments ending with adenosine 3' phosphate. The combined use of these two enzymes will give us a mixture of fragments generated by the action of either enzyme, or both. We can deduce the sequence of the fragments obtained, as follows:

${}^5\text{CCAp}^3$ ${}^5(\text{C,U})\text{Gp}^3$

Note: we can deduce the presence of the two cytidines in the first fragment, since it is mentioned that this fragment is a trinucleotide, and that one C is lacking when we count the nucleotides obtained after digestion with RNAses T1 + T2.

From the data obtained after digestion with pancreatic ribonuclease, we can tell that adenosine precedes uridine. Therefore, we get the following fragment:

${}^5\text{AUCGp}^3$

Also, again from the results obtained after the digestion with RNase T1 + T2, the fact that we obtained pGp tells us that this is the 5' nucleotide. Furthermore, the result obtained with pancreatic ribonuclease tells us that G is followed by C. The fact that C and Ap were obtained also tells us that these two nucleotides immediately follow A or G.

This allows us to order the nucleotides as follows:

${}^5\text{pGC} \dots \text{AUCG} \dots \text{C}^3$

All we have to do now is to place A and C. Since digesting the RNA with RNase T1 + T2 gave us the sequence ${}^5\text{CCAp}^3$, this suggests that C is located between pGC and AUCG. Finally, the last nucleotide, A, would logically be placed before the last C of the nucleic acid.

Taking into account all these pieces of information, we get the following sequence:

${}^5\text{pGCCAUCGAC}^3$

6.9. Structure of nucleic acids.

At neutral pH, the nitrogenated bases are not ionized, and the charges will only come from the phosphate groups. Since, in the example given here, three phosphates for the phosphodiester bonds, we get $3 \times (1-) = 3$ negative charges.

6.10. Structure of nucleic acids.

When a circular DNA molecule is denatured, the result is two tangled-up single-stranded circles. When a linear DNA molecule is denatures, both strands diffuse away from each other. Upon renaturation, it will be much easier for the circular, tangled-up single strands to form complementary base pairs (because they are closer), this compared to a linear molecule of identical sequence (which have to rely on random collisions in order to meet).

6.11. Structure of nucleic acids.

The fusion temperature depends on the ionic strength (i.e. the salt concentration) of the solution. If we decrease the ionic strength, the fusion temperature will also decrease. In

the extreme case where the ionic strength is equal to zero, the the negatively charged phosphate groups will not be neutralized by counter-ions: this will be sufficient to decrease the fusion temperature to below 20°C.

6.12. Structure of nucleic acids.

B-type DNA of 10 base pairs has a length of 3,4 nm (34 Å). The E. coli chromosome, being 4 000 kbp long (i.e. 4 000 000 bp), will therefore have a length of $1,36 \times 10^6$ nm (approximately 1,4 mm).

6.13. Nucleic acid synthesis.

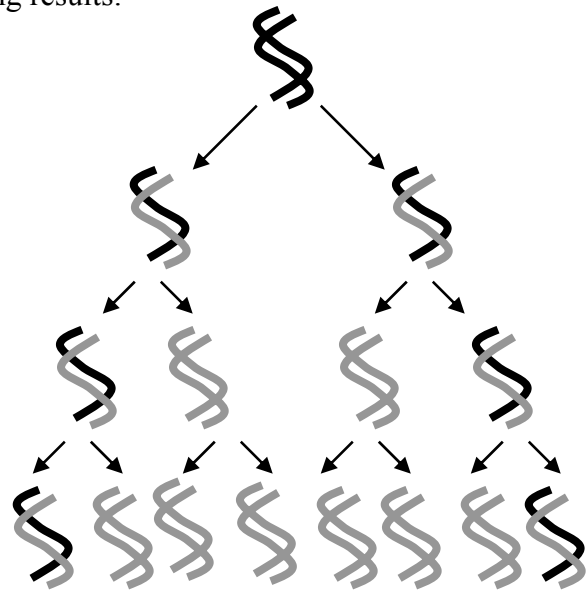
In the Meselson and Stahl experiment, we first incubate bacteria in culture media containing only the ^{15}N isotope of nitrogen. We then transfer those bacteria into a culture media containing only ^{14}N . We obtain the following results:

After 0 generations: 100% heavy DNA:

After 1 generation: 100% hybrid DNA:

After 2 generations: 50% hybrid, 50% light DNA

After 3 generations: 25% hybrid, 75% light DNA.



6.14. Nucleic acid synthesis.

The starting strand (+) is made of: 10% A, 20%G, 30%C and 40%T. After its replication, the complementary (-) strand will have the following composition: 40%A, 30%G, 20%C and 10%T (since A base pair only with T, and C only with G). Finally, after transcription the RNA produced will have a nucleotide composition that will be complementary to the (-) strand, which is 10%A, 20%G, 30%C and 40% U.

6.15. Nucleic acid synthesis.

This observation can only be explained if the E. coli genome is replicated 1.5 times when the bacteria are seperating.

**6.16. Nucleic acid synthesis.

a) since no hybrid molecule is obtained, we must conclude that replication occurs through a *convervative* mechanism, during which the two template (old) strands re-anneal together after each round of replication.

b) The presence of RNA polymerase suggests that the bacterial DNA is first transcribed into an RNA molecule. Afterwards, this RNA molecule (resistant to the action of *DNAse*) is converted into a DNA molecule (*DNAse sensitive*) through the reduction of the 2' carbon via the new enzymatic activity and NADH. Finally, this DNA is used as a template by DNA polymerase to form a double-stranded DNA molecule identical to the initial parental duplex.

6.17. mRNA and transcription

a) The introduction of mutations due to the absence of a proofreading activity is a random process which will affect only a small number of the copies of any given RNA. Furthermore, mRNAs characteristically have a short half-life and will lead to the synthesis of only a handful of protein molecules before being degraded. Therefore, the introduction of mutations in mRNAs will affect only a very small number of the copies of a given protein and will not be detrimental to the cell.

b) The use of an RNA intermediate during DNA synthesis would lead to the rapid accumulation of mutations in the genetic material of the organism, this due to the lack of proofreading activity in RNA polymerases. This explains why some RNA viruses (e.g. HIV) have a very high mutation rate and can, therefore, easily escape anti-viral therapies.

6.18. mRNA and transcription

The short half-life of mRNAs allows 1) the rapid disposal of mutated transcripts, and 2) to rapid regulation gene expression. Effectively, transcription arrest will be rapidly followed by the degradation of the corresponding mRNAs, ensuring the rapid termination of the synthesis of the protein encoded by the gene.

6.19. mRNA and transcription

a) Since the maximal transcription rate is 4,300 nucleotides per minute (70 nucleotides per second x 60 seconds), a 6,000 pb gene will be transcribed in approximately 1.4 minutes.

b) Since RNA polymerase covers 70 pb, 86 RNA polymerase molecules would cover this DNA molecule at any one time.

6.20. Protein coding.

a) From the genetic code, we have the following amino acid sequence:

AGU CUC UGU CUC CAU UUG AAG AAG GGG AAG GGG

Ser - Leu - Cys - Leu - His - Leu - Lys - Lys - Gly - Lys - Gly

b) The mutations will change the amino acid sequence as follows:

\downarrow

 \downarrow

G

AGU **G**CU CUG UCU CCA UUU GAA GAA GGG AAG GGG

Ser - Ala - Leu - Ser - Pro - Phe - Glu - Glu - Gly - Lys - Gly

6.21. Protein coding.

a) Let's start by converting the amino acid sequence of the wild type into its nucleotide sequence:

Tyr	-	Lys	-	Ser	-	Pro	-	Ser	-	Leu	-	Asn	-	Ala	-	Ala	-	Lys
UAU	-	AAA	-	UCX	-	CCX	-	UCX	-	UUG	-	AAC	-	GCX	-	GCX	-	AAG
C		G	-	AGU			-	AGU	-	CUX		U						A
				C				C										

Next, let's convert the amino acid sequence of the mutant into its nucleotide sequence:

-	Val	-	His	-	His	-	Leu	-	Met	-
-	GUX	-	CAU	-	CAU	-	UUA	-	AUG	
			C		C		G			
							-	CUX		

The mutant could not be generated by a single mutation: we need one mutation to change the wt into the mutant, and another one to convert the mutant back to the wt.

If we take a closer look at the wt and mutant sequences, we notice that the mutant sequence can be obtained by the deletion of the 7th base of the wt (i.e. the A of the AGU Ser codon). We then get:

Sauvage :										
				A						
	UCX	-	CCX	-	UCX	-	UUG	-	AAC	
	AGU			-	AGU	-	CUX		U	
	C				C					
Délétion:										
	GU	-	CCX	-	UCX	-	UUX	-	AAU	
	C									

The mutant will give us the following reading frame:

GUC	-	CXU	-	CXU	-	UAA	-	AU
C						G		C

Replacing all the Xs by As, we see that the sequence is identical to the mutant's:

Transcribing this DNA will give us (notice the 5'→3' orientation):

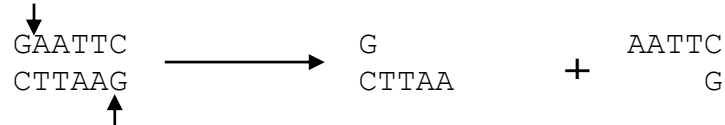
5' UCG AGU CGA AAU GGG GAU CGU AAA CGA 3'

And the corresponding protein sequence will be:

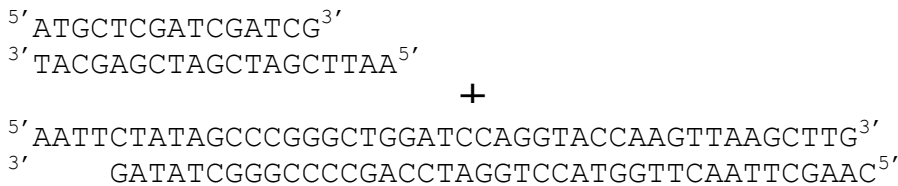
Ser-Ser-Arg-Asn-Gly-Asp-Arg-Lys-Arg

6.23. Genetic engineering.

EcoR I cuts DNA in the following manner:

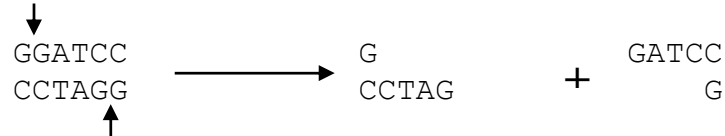


In the fragment given here, there is only one EcoR I recognition site. After digestion with EcoR I, we get the following fragments:

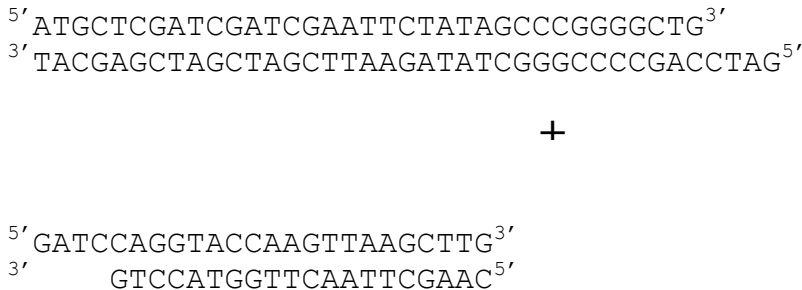


6.24. Genetic engineering.

BamH I only cuts DNA in the following manner:

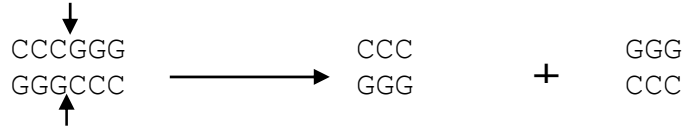


In the fragment given here, there is only one BamH I recognition site. After digestion with BamH I, we get the following fragments:

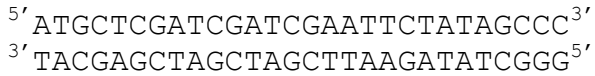


6.25. Genetic engineering.

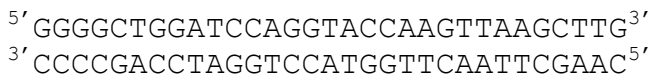
Sma I cuts DNA molecules in the following manner:



In the fragment given here, there is only one Sma I recognition site. After digestion with Sma I, we get the following fragments:

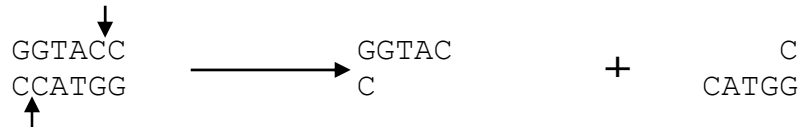


+

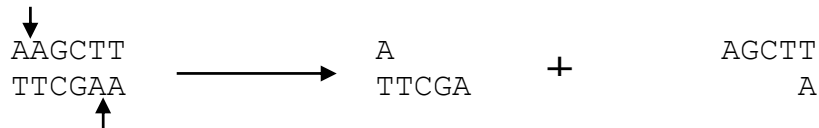


6.26. Genetic engineering.

Kpn I cuts DNA in the following manner:



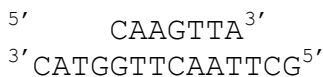
As for Hind III, it will cut DNA as follows:



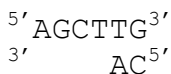
In the fragment given here, there is only one recognition site for each enzyme. After digestion with both enzymes, we get the following fragments:



+

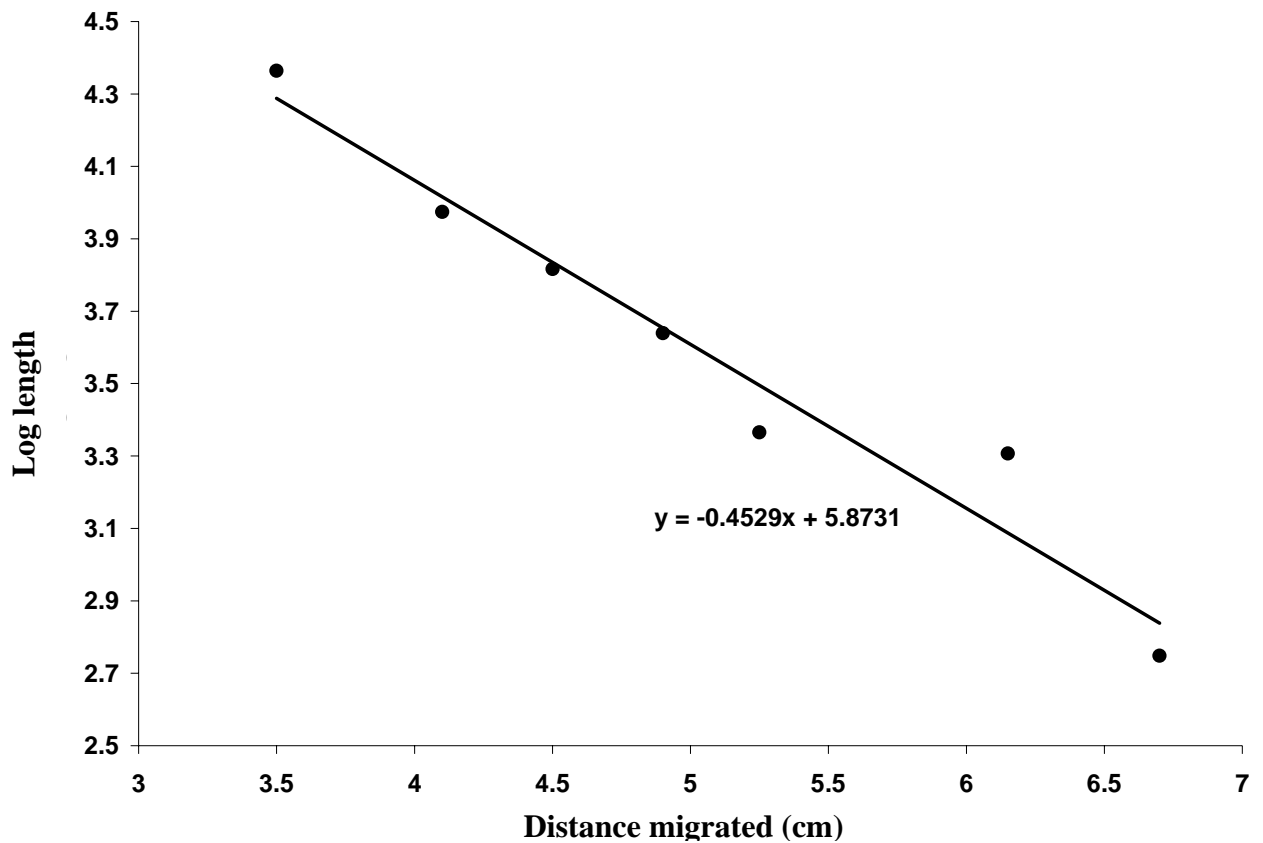


+



6.27. Genetic engineering.

a) To determine the size of the restriction fragments, we must first trace the graph of the log of the length of the standard DNA markers as a function of the distance migrated from the well. This graph is shown below.



With this graph, it is easy to calculate the length of the restriction fragments from the distance they migrated in the gel. The results obtained are shown in the following table.

Fragment	Apa I (bp)	Pvu I (bp)	BamH I (bp)
1	48,500	48,500	48,500
2	10,085	35,790	41,730
3		26,250	34,500
4		11,930	27,920
5			22,345
6			5,505

b) Since only one extremity is radiolabelled, the size of the restriction fragments seen after

autoradiography gives us the distance between the restriction site and the labeled end. Furthermore, since the experiment was done under conditions where only a partial digest was obtained, we get a mixture of fragments of different sizes for those enzymes cutting more than once. The length of each fragment will indicate the position (from the labeled end) of one of the restriction sites. For example, the fact that two fragments of 10,085 bp and 48,500 bp were obtained by digesting the DNA with Apa I tells us that Apa I only cuts once at a distance of 10,085 bp from the labeled end. Applying this reasoning, we obtain the following restriction map:

