AMYLASE ACTIVITY EXERCISE

Functional and Structural Analysis

OBJECTIVES

Upon completion of this exercise, the student should be able to:

1. design experiment(s) with appropriate controls to test their hypothesis.
2. make a standard curve for estimation of maltose concentration.
3. set up enzyme-catalyzed (amylase) hydrolysis of starch into maltose.
4. explain the relationship between amount of maltose produced and the activity of amylase.
5. obtain amino acid sequences of amylase from the NCBI online database.
6. use bioinformatics tools to analyze primary structures, compare **amino acid sequences** and generate a 3D structure of each enzyme.

BACKGROUND

Chemical reactions that take place in a cell do not occur randomly; they are controlled by biological catalysts, called *enzymes.* A catalyst is a substance that speeds up the rate of a chemical reaction without being used up by the reaction. Most *enzymes are proteins* whose primary structure is dictated by genes. Enzymes do not become active until the polypeptide chains are folded into a unique three-dimensional (3D) shape. The 3D structures are held together by various *non-covalent* interactions such as hydrogen bonds, ionic bonds, hydrophobic and Van der Waals interactions as well as disulfide bridges (*covalent*).

Thousands of product molecules may be formed in one second by an individual enzyme. Reactants in an enzymatic reaction are called *substrates* and substrates bind to the enzyme at a specific site, called the *active site*. The active site of the enzyme is like a pocket or cleft in the protein that is shaped in a way that substrates can fit in. Since only properly shaped substrates can fit into the active site, specific enzymes bind *specific* substrates, similar to a key that is shaped to fit a specific lock. Some enzymes may have metal ions (e.g., Ca^{2+}, Cu^{2+}, Fe^{2+}, Mn^{2+}) as part of their active site, which are called *cofactors*, or may use organic molecules as *coenzymes* (derived from vitamins). The substrate(s) is converted into a product at the active site. Substrates bind to enzyme molecules using a combination of weak, noncovalent chemical bonds, forming an *enzyme-substrate (ES)* complex that exists for a fraction of a second. During

this time, subtle changes in the shape of the active site, called ***induced fit***, stress and orient covalent bonds of the substrates in a way that facilitates the formation of a product. (Figure 1). The newly-formed product then leaves the active site, whereas the enzyme remains unchanged and can bind to additional substrate molecules, if they are available.
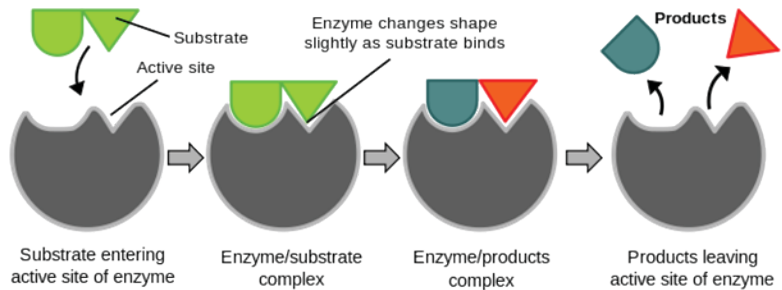

Figure 1: Induced fit mechanism of enzyme action
(Created by TimVickers, vectorized by Fvasconcellos / Public domain)

Enzymes speed up the rate of a reaction by ***lowering*** the initial amount of energy required, called the ***activation energy (E_A)***. Activation energy is required by all chemical reactions to break certain bonds (free up electrons) in the reactant(s) so that new bonds can form, resulting in product formation (Figure 2). Enzymes use energy from the surrounding environment to increase molecular movement of substrates. This reduces the ***activation energy barrier.***
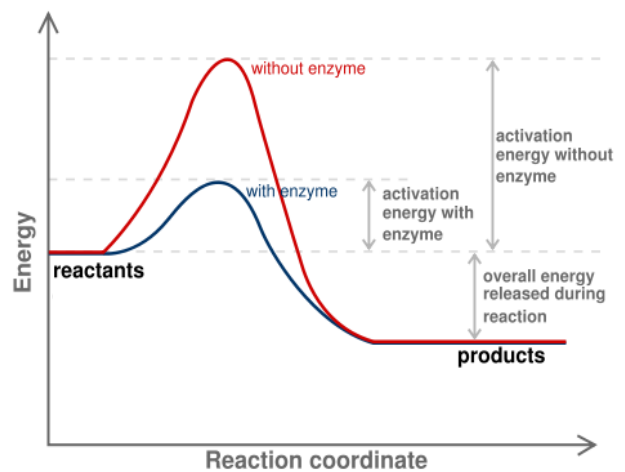

The activity of enzymes, like that of all proteins, is affected by environmental conditions. Factors like ***temperature, pH,*** and ***salt concentration*** interfere with the non-covalent forces that give enzymes their 3D shape and enzymes then start to unfold, or ***denature.*** Alpha helices and beta sheets, for example, are disrupted in denaturing conditions, and the peptide chain takes on a random shape. Denaturation does not break peptide bonds, so the enzyme's primary structure is unaffected. However, because the enzyme has changed shape, the active site is also

Figure 2: Energy diagram of a reaction in the presence and absence of enzyme
Originally uploaded by Jerry Crimson Mann, vectorized by Tutmosis, corrected by Fvasconcellos / CC BY-SA
(http://creativecommons.org/licenses/by-sa/3.0/)

altered and will not complement and bind the substrate as well. The reaction rate therefore decreases accordingly. Enzymes can structurally be similar, and catalyze the same reaction, yet they may have important differences leading to different optimal temperatures, or optimal pH or salt conditions. Organisms regulate the rates of their reactions by regulating the activity of their enzymes.

Factors that affect the rate of enzyme activity:

1. ***Temperature*** affects enzyme activity in two ways. As the temperature rises, molecular motion (kinetic energy) increases and the rate of random collision between enzyme and substrate molecules increases, forming more products. After a certain point, increasing the temperature strains the non-covalent bonds, altering the shape of the active site, and the overall shape of the enzyme. This decreases the rate of product formation. The temperature at which enzyme activity

is the highest is called the optimum temperature. At high temperature, an enzyme will most likely unfold and denature.

2. ***Changes in pH (H⁺ concentration)*** and salt concentration primarily affect the stability of secondary and tertiary structures maintained by hydrogen bonds and disrupt salt bridges held by ionic bonds. As a result, enzymes denature at extreme pH and high salt concentrations. In addition, substrates and/or enzyme active site groups may ionize, which further affects enzyme-substrate binding.

3. ***Substrate and enzyme concentration*** also affect the rate of enzyme reaction. Increasing the concentration of substrate and/or enzyme increases the rate of reaction up to a certain point. As the reaction continues and the substrate molecules are used up, the rate of reaction will decrease regardless of any changes in enzyme concentration. By controlling enzyme and substrate concentration, organisms can regulate their metabolism.

Enzyme activity is measured by monitoring changes in substrate and/or product concentrations:

To learn about the function and structure of enzymes, we will be using the enzyme ***amylase*** as a model in this lab. Alpha amylase, is an enzyme that catalyzes the ***hydrolysis*** (breakdown) of α-1,4 glycosidic linkages of starch (a polymer of glucose), into maltose (a reducing disaccharide made of two glucose molecules). Because starch is one the most abundant carbohydrate polymers on earth, it serves as a major source of energy not only for us but for many other animals, higher plants and microorganisms as well. In order to harvest the energy from starch, the enzyme amylase**,** present in saliva and pancreatic secretions of humans and other mammals, begins the chemical process of breaking (or hydrolyzing) starch down into smaller sugars. Amylase is present in all 3 domains of life - Bacteria, Archaea and Eukarya (plants, animals, and fungi) - with the same catalytic function.

$$\textbf{Starch + water} \xrightarrow[\textit{amylase}]{} \textbf{Maltose}$$
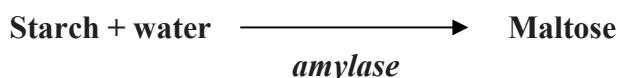
Alpha amylases from three different sources, (i) bacteria (*Geobacillus stearothermophilus*), (ii) fungi (*Aspergillus oryzae*), and (iii) humans (*Homo sapiens*), will be used in this lab. A video on enzyme activity and the amylases we will be studying in this lab can be found [here](#).

The concentration of maltose can be measured using a colorimetric assay; this means combining maltose with a certain reagent that causes a color change. The reagent in this assay is DNS (3,5-dinitrosalicylic acid, also called DNSA) which is yellow in color. First introduced to detect ***reducing substances*** in urine, the DNS assay is commonly used to quantify carbohydrate levels in blood as well as detect alpha amylase activity.

In an alkaline solution, ***reducing sugars*** form aldehyde or ketone groups, which can then reduce different reagents (e.g., Benedicts reagent, dinitrosalicylic acid or DNS). Unlike the ***Benedicts test*** for detecting the presence of reducing sugars, reaction between DNS and a reducing sugar results in a ***soluble, colored product.*** Maltose participates in an oxidation-reduction reaction with DNS due to its carbonyl group (C=O).

DNS is reduced to 3-amino, 5-nitrosalicylic acid (ANSA - will be referred to simply as **reduced DNS**) and maltose is oxidized to maltonic acid (Figure 3).
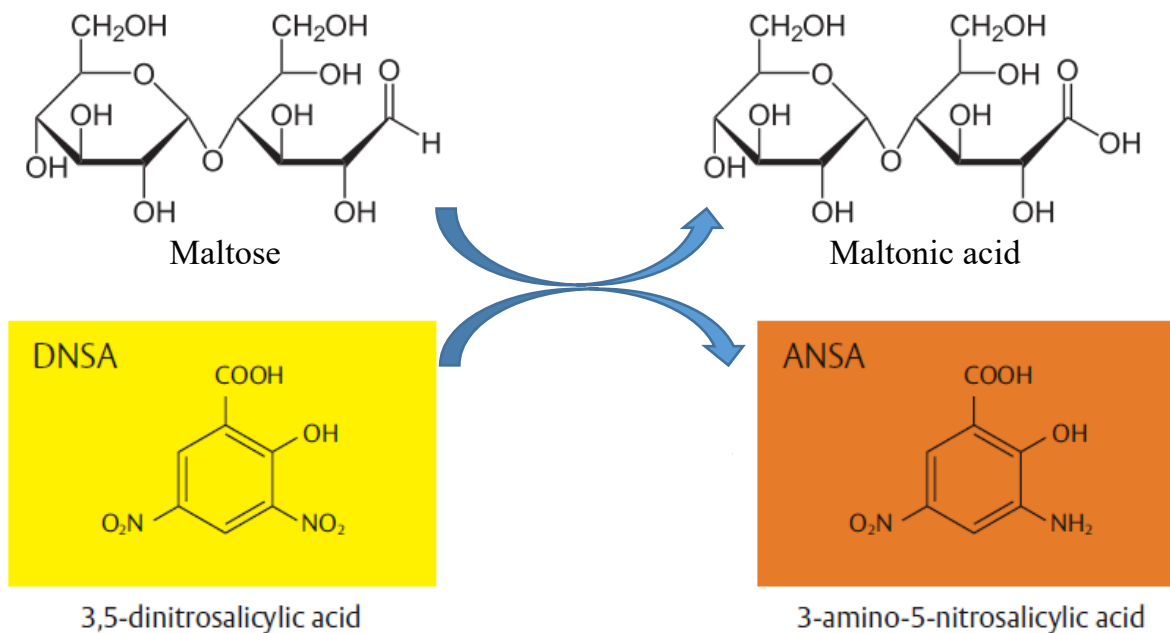

Figure 3: Reduction of DNS by maltose produces reduced DNS and maltonic acid.

Department for Education (April 2014) *GCE AS and A level subject content for biology, chemistry, physics and psychology.* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/446829/A_level_science_subject_content.pdf

This reaction causes a change in color from yellow to orange/red when DNS is reduced (Figure 3). The change in color results in a change in absorption of light, and absorbance is measured in a spectrophotometer at a wavelength of 540 nm. The intensity of color from the reaction, and the absorbance of light, is proportional to the concentration of maltose and is used to estimate the concentration of maltose in any given solution.

The concentration of maltose produced by hydrolysis of starch is directly dependent on amylase activity. Therefore, absorbance data can be used to determine the optimal conditions for amylase activity, i.e., conditions at which this enzyme has the highest activity.

PRE-LAB 1
Submit online or hand in at the beginning of your lab

1. Define Enzymes. What type of macromolecules are enzymes?

2. What are 2 advantages of having enzyme-catalyzed chemical reactions in living cells?

3. What is a substrate? Where on an enzyme does the substrate specifically bind to during a chemical reaction?

4. When an enzymatic reaction is in progress, do you expect to see an *increase*, *decrease* or *no change* in each of following:

    (i) substrate

    (ii) product

    (iii) enzyme

5. How can you measure enzyme activity?

6. What is the relationship between enzyme activity and 3D shape of an enzyme? What type of environment changes affect the function of an enzyme?

7. Define optimal conditions for enzyme activity. How can you determine (i) optimal temperature and (ii) optimal pH of an enzyme?

8. What enzyme will you be studying in lab today? Where can you find this enzyme?

9. Write the reaction catalyzed by amylase. Do you expect this enzyme to hydrolyze cellulose (a polymer of glucose)? Explain.

10. Production of maltose, by _____ (condensation, dehydration, hydrolysis) of starch, in the presence of amylase, can be detected by _____ assay.

PART I: INTRODUCTION TO THE DNS ASSAY TO MEASURE AMYLASE ACTIVITY

In this part of the exercise, you will be introduced to the hydrolysis of starch by amylase, and the DNS assay used to measure amylase activity. You will set up the hydrolysis of 5% starch *in vitro,* at room temperature (25°C), using the *fungal amylase*. Remember this may not be the optimal temperature for this enzyme. *Activity of amylase is directly related to the amount of maltose produced.*

EXPERIMENTAL SETUP:

1. Two tubes are prepared according to the table below. For a video on how to use micropipettes, click here:

Tube	Water (µl)	5% Starch (µl)	Fungal Amylase (µl)	Total Volume (µl)
A	450	450	100	
B	550	450	0	

2. Both tubes are incubated for <u>10 min</u> at 25°C (on bench).

3. 1000 µl (1 ml) **DNS** is added and the tubes are placed in boiling water for <u>5 min</u>.

4. 8 ml of dH$_2$O (deionized water) is added to both samples to dilute them, using a serological pipet. The **absorbance of light at 540 nm** is read for the two solutions using a spectrophotometer. Here is a video on how to use a spectrophotometer like the one we have in our lab:

   Tube B is used as a "blank", to calibrate the machine.

   What treatment is tube B in the experiment? _____

   What treatment is tube A in the experiment? _____

5. The absorbance of <u>tube A</u> you read in the spectrophotometer is **<u>0.358</u>**.

6. After completing Part II, the maltose standard curve, determine the maltose concentration in your tube _____.

PART II: STANDARD CURVE FOR ESTIMATION OF MALTOSE

The objective of this part of the lab activity is to collect data and plot a standard curve. The procedure for this activity is described below.

DNS is added to known amounts of maltose, and the absorbance of the reduced DNS is measured at 540nm, using a spectrophotometer. A reference graph, called a **standard curve**, is made from this data. A standard curve is a graph which shows a relationship between two quantities; in this activity, the relationship between known concentrations of maltose and their absorbance. You may want to review the "line of best fit" section in the graphing skills part of Exercise 1. In part III of this lab exercise, **you will use the standard curve to estimate the amount of maltose produced during hydrolysis of starch, the reaction catalyzed by amylase. The data on the amount of maltose produced will then be used to calculate amylase activity.**

To prepare the standard curve, the concentration of maltose (mg/ml) in each tube has to be determined first. Please watch this video to understand the steps involved in calculations, and to construct the standard curve.

1. **Complete the table below and show a sample of your calculations. You will be using the formula $C_iV_i = C_fV_f$** (where C_i = initial maltose concentration, V_i = initial volume of maltose solution, C_f = final maltose concentration and V_f = final volume of maltose solution).

   The initial maltose concentration (C_i) used is 2.5 mg/ml. After adding 1 ml DNS to 1 ml of maltose, all samples are diluted with 8 ml of water, and their absorbance is determined in a spectrophotometer and indicated in the table. The final volume of maltose (V_f) is, therefore, 10 ml.

   Table 1: Final Concentration of maltose

Volume of initial maltose solution V_i (μl)	Volume of initial maltose solution V_i (ml)	Volume of water added to get a total of 1 ml	Final Maltose Concentration (mg/ml)	Absorbance at 540nm
0				0.000
200				0.228
400				0.487
600				0.751
800				1.031
1000				1.275

2. **Graph 1: graph the Absorbance vs. Concentration of maltose from Table 1. There are two ways to do it, a) manually, on graph paper, OR b) in Microsoft Office Excel, depending on what your instructor has asked you to do.** Remember to use TAILS and insert your graph below. If you are not sure how to do that, **watch the video on maltose standard curve for help.**

   Choose one of these two ways to present your graph:

   a. Construct your graph manually, on graph paper. Draw a line of best fit through your points, making sure that line passes closer to as many data points as possible. Embed a picture below. This graph will be used in Part III, to determine the amount of maltose produced in your experimental tubes.

   OR

   b. Construct your graph on excel, according to the instructions on the video provided on the maltose standard curve. Draw a 'line of best fit' through your data points by clicking on a point and selecting "add trendline". In the menu that opens, select "display equation of line, to get the equation of the line to appear on your graph. Embed your excel graph below. You will use this equation of the line in Part III, to determine maltose produced in your experimental tubes.

PRE-LAB 2
Submit online or hand in at the beginning of the lab

GROUP:

NAMES OF GROUP MEMBERS:

STATE YOUR HYPOTHESIS ON AMYLASE ACTIVITY:

CITED SOURCE (APA):

Part III: Test your hypothesis

Before doing this part, watch this video on testing your hypothesis on the effect of a preassigned variable (temperature, for example) on amylase activity. **You will be assigned an independent variable and amylase enzyme isolated from one of the three organisms described in the background information and the introduction video.**

You will write a question, a hypothesis and a prediction based on the source of amylase (bacterial, fungal or human) and the independent variable you were assigned. You will also define the dependent variable. To write a good hypothesis, research the literature on the organism/enzyme you are studying, and base your hypothesis on the information you find. Make sure to cite your sources (APA style) both in the text and at the end of your report.

Remember:

- **Question:** Asking a question is an important part of the scientific process. The more specific questions you ask, the easier it will be for you to design experiments to test your hypothesis.

- **Hypothesis:** Tentative answer to the question you asked in the previous step. Write a simple but specific statement that is testable and falsifiable. **You should research the literature on the organism/enzyme you are studying, and base your hypothesis on the information you find. Make sure to cite your sources (APA style) both in the text and at the end of your report.**

- **Prediction:** The prediction describes what will happen if a hypothesis is correct. Generate a reasonable prediction by completing this statement: If _____ then _____.

Determine which amylase and which independent variable you are supposed to study, based on your group assignment, according to the table below:

GROUPS	AMYLASE	INDEPENDENT VARIABLE
1	*Geobacillus stearothermophilus (bacterial)*	Temperature
2	*Geobacillus stearothermophilus (bacterial)*	pH
3	*Homo sapiens (human)*	Temperature
4	*Homo sapiens (human)*	pH
5	*Aspergillus oryzae (fungal)*	Temperature
6	*Aspergillus oryzae (fungal)*	pH

3. **Experimental Design: Table 2A** shows how your tubes would be set up if you were studying the effect of **temperature** on amylase activity. **Table 2B** shows how your tubes would be set up if you were studying the effect of **pH** on amylase activity.

    - Tube contents are mixed, and incubated in water baths at appropriate temperatures for <u>10 minutes</u>.

    - 1 ml DNS is added to all tubes and they are boiled for 5 minutes.

    - 8 ml of water is added for diluting, and absorbance readings are collected in a spectrophotometer as indicated in Table 3.

Table 2A: experimental setup for Temperature

Tube	Water (µl)	5% Starch (µl)	Amylase (µl)	Temperature (^{0}C)
1	450	450	100	0
2	450	450	100	25
3	450	450	100	37
4	450	450	100	45
5	450	450	100	65
6	450	450	100	85

Table 2B: Experimental setup for pH

Tube	Buffer (µl)	5% Starch (µl)	Amylase (µl)	pH
1	450	450	100	1
2	450	450	100	3
3	450	450	100	5
4	450	450	100	7
5	450	450	100	9
6	450	450	100	12

The tubes listed in the table are your **experimental** treatment. What will the **negative control** tubes contain?

4. **From the tables in <u>Appendix A</u> (p. 23), <u>choose the one</u> that contains data for your amylase enzyme and your independent variable**, according to the determination you made based on your group. <u>Insert the appropriate table in your report</u>. **Using either the manual graph, or the excel graph of the maltose standard curve you constructed in Part II**:

   a. **Determine the maltose produced in your experiment and fill the data. Show an example calculation below the table.**

   b. **Determine the amylase activity in each tube and fill in that data in. Show an example calculation below the table.**

5. **<u>Graph 2</u>: graph the resulting Amylase Activity data vs. your Independent Variable (on Excel or manually on graph paper) and insert your graph or a picture of it in your report.**

6. **Conclusion:** restate your hypothesis and summarize your results based on your graph of the Amylase Activity vs. your Independent variable. Explain whether your graph supports your hypothesis, by using information from the graph. Comment on other findings from the graph. Comment on what you think happened to the Amylase enzyme and its structure and function where you observe low activity.

Rubric Enzyme Report (40 points)

1. **Introduction: (4 pts)**

    a. What are enzymes and what do they do?

    b. What are the optimal conditions for an enzyme?

    c. What factors affect enzyme activity and how?

    d. How is the structure of an enzyme related to its function?

    e. What is amylase? What reaction does it catalyze?

    f. Research the amylase enzyme you used (organism) and the optimal conditions of temperature, pH, or salt, it requires.

    g. Explain why DNS is used and why it is important for determining enzyme activity. Cite and reference your sources using APA style.

2. **Experimental Design: (5 pts)**

    a. State your question, hypothesis (ensure that it is testable and falsifiable), and prediction.

    b. Identify the independent variable.

    c. Identify the dependent variable.

    d. Identify the experimental group

    e. Identify the control group.

3. **Materials/Methods: (3 pts)**

    List all of the equipment and materials you will need for your experiment. BRIEFLY describe your experimental procedure for setting up the maltose standard curve (Part II) and for testing your hypothesis on amylase (Part III).

4. **Data Analysis/Results:**

    a. List all the <u>formulas</u> you used for your calculations and a <u>sample</u> calculation in each case **(3 pts)**

    b. **Prepare the following tables**:

    *   **Table 1** with maltose standard curve data. **(3 pts)**

    *   **Table 2** with your experimental setup. Explain how your control tubes will be set up. **(2 pts)**

    *   **Table 3** with absorbance measurements of experimental tubes, calculated maltose concentration, calculated amylase activity (mg/ml maltose/min) **(4 pts)**

    c. **Prepare the following graphs:**

    *   **Graph 1** of maltose standard curve: show line of best fit and equation of line if done on excel. If done manually, insert an image of your graph indicating the line of best fit. **(3 pts)**

- **Graph 2** of enzyme activity versus independent variable **(4 pts)**

d. **Summary of results:** summarize your results based on your graphs **(3 pts)**

5. **Conclusion/Discussion: (4 pts)**

Based on your results, is your hypothesis supported? Restate your hypothesis and explain! Discuss any other findings and explain them. Discuss errors, and reasons for data variability.

6. **References: (2 pts)**

List the sources you used at the end of your report, and include in-text citations in your report (APA citation style)
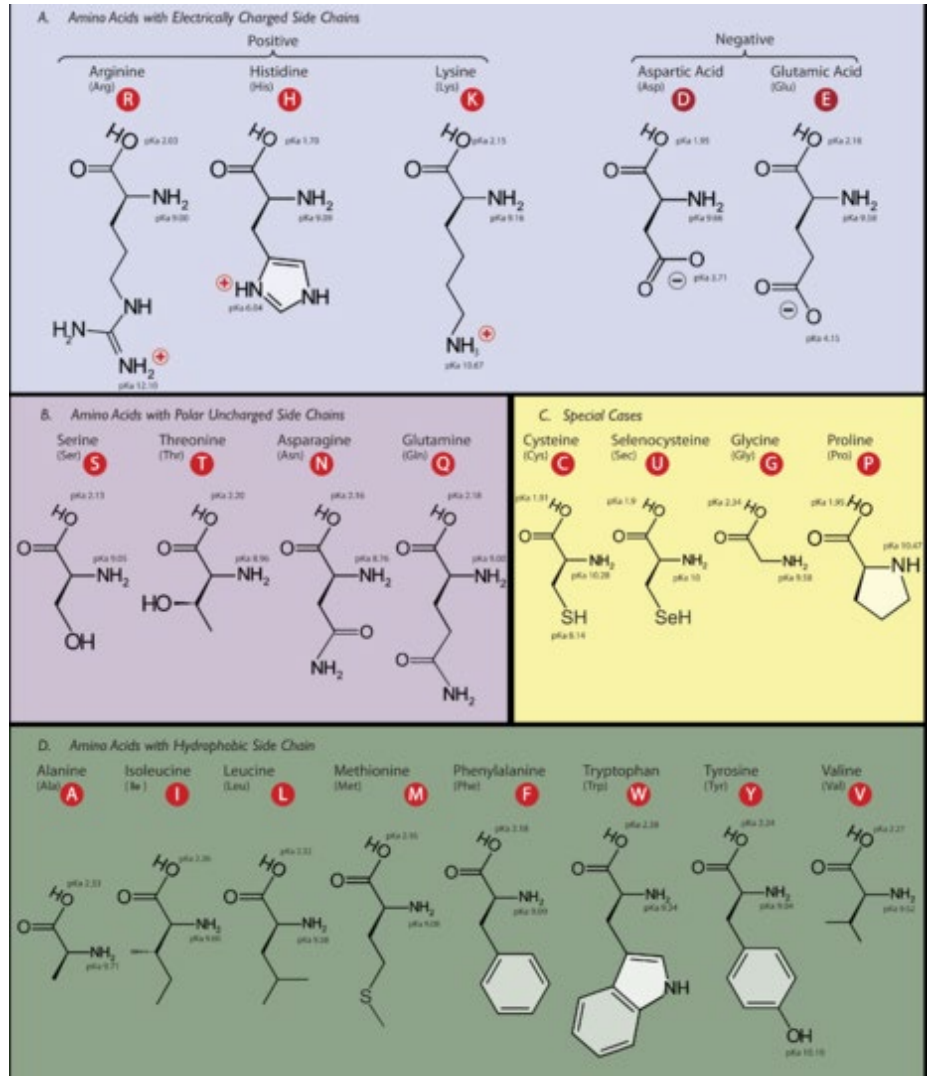
PART IV: BIOINFORMATICS – STRUCTURAL ANALYSIS OF ALPHA AMYLASES

Bioinformatics is the use of computer software and computational tools, for the analysis of protein and nucleic acid sequence information (online databases), through the use of sophisticated, stand-alone and widely available online software packages.

This part of the lab exercise will introduce you to elementary bioinformatics tools for the analysis of three alpha amylase enzymes that were used in parts I-III. You will retrieve the amino acid sequences of all three amylases from an online databank, analyze those sequences and view their three-dimensional structures. Since online databanks use the one-letter amino acid code, please look at the figure, on your right, in order to familiarize yourself with the one-letter system.

Dan Cojocari, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons

Ability to maintain functional 3D structure within a particular range of temperature, pH, salinity, is an intrinsic property of proteins and is determined by the primary structure. Amino acid sequence determines the way a protein folds into unique 3D shape (native conformation). Protein homology means that the proteins are derived from a common ancestor gene and will have the same number and types of secondary structure, oriented in the same way in three-dimensional space. All amylases catalyze hydrolysis of starch, yet their amino acid sequences have significant differences among different organisms.

Increased thermostability can be conferred by a greater number of ionic interactions, disulfide bridges, proline residues, as well as numerous other structural strategies. We will also discuss the factors that give rise to protein thermostability and interpret the data in an evolutionary context.

In many ways, biology has become a "big data" science. Since the human genome project was completed, many more genomes from all domains of life have been sequenced. Furthermore, scientists routinely sequence individual genes as well as their associated proteins. That wealth of genomic data, which includes RNA sequences and non-coding DNA sequences, is stored in a number of databases housed by the National Center for Biotechnology Information or NCBI (http://www.ncbi.nlm.nih.gov/). While the NCBI is known mostly for its searchable databases, another website known as ExPASy provides scientists with computational tools to analyze genomic data. ExPASy bioinformatics portal (http://www.expasy.org/) is maintained by the Swiss Institute of Bioinformatics (SIB). Another popular bioinformatics portal is The European Molecular Biology Open Software Suite or EMBOSS (http://emboss.sourceforge.net/) which is a product of the European Molecular Biology network.

In part III, you tested your hypotheses on the effect of different variables on alpha-amylase activity. Now, let's get some insight on the evolution of thermostability of alpha-amylase. **We will use various biocomputational tools to compare the primary, secondary and tertiary structures of human, fungal and bacterial alpha-amylases.** We will then discuss the factors that may contribute to thermostability and the evolutionary mechanisms that may have led to thermostability of alpha-amylase.

**Retrieval of the amino acid sequences of amylases**: In order to analyze our proteins we must first retrieve their amino acid sequences from the **NCBI** database. As in any archive or library where items are given an accession number for easy retrieval, protein and DNA sequences in the NCBI database are associated with an accession number as a unique identifier (indicated below, in Table 4). The table also includes another unique identifier for the Protein Data Bank or PDB, one of the databases that will be used in this exercise.

Table 4: NCBI ACCESSION NUMBERS AND PDB ID'S OF ALPHA AMYLASES

Source of alpha amylase	NCBI accession number	PDB ID
Homo sapiens (salivary)	GI:157833830	1SMD
Aspergillus oryzae	GI:541881321	3VX0
Geobacillus stearothermophilus	GI:8569361	1HVX

**Analysis of primary structures of amylases**: One of the software tools from EMBOSS, named **PEPSTATS,** computes statistics on the properties of individual proteins based on their primary structures. This tool can provide information about the nature of the amino acids (with polar, nonpolar, or charged side chains, etc).

**Multiple sequence alignment of amylases**: Sequence alignments can be used to detect **homology**. A multiple sequence alignment is a computational method of aligning two or more monomeric sequences of proteins or nucleic acids in order to _**(i) compare identities and similarities and (ii) determine evolutionary relationship.**_ It can also be used when a novel protein is discovered and information is needed about its function. Homologous sequences usually have similar structure and function. _**Significant sequence homology and structural similarity strongly implies common ancestry.**_

Functional similarity 'supports' common ancestry but is not sufficient to demonstrate it. Sequence homology, expressed in percentage, is based on **similarity** or **identity** in the amino acid sequence. Similarity means that the residues in a pair of sequences are chemically similar, and identity means that the residues are exactly the same. Similar amino acids can replace one another over the course of an evolutionary period and still have the same function.

One of the most reliable and versatile bioinformatics tools is called CLUSTAL OMEGA, offered by the European Bioinformatics Institute (EBI). CLUSTAL OMEGA is a general purpose multiple sequence alignment program for DNA/proteins, for alignment of 3 or more sequences. It calculates the best match for the selected sequences, and lines them up so that the identities, similarities and differences can be seen. Evolutionary relationships can also be determined by viewing Cladograms or Phylograms.

Highlighting conserved sequences in the sequence alignment: In sequence alignment, certain amino acid residues are more important to the function than others and thus are highly **conserved** throughout the evolutionary history. The multiple sequence alignment is more useful when it can be viewed with an editor such as BOXSHADE. This is an alignment visualization software that highlights identical amino acids across sequences with black boxes, and similar amino acids with grey boxes.

The ability to compare amino acid sequences between different organisms allows us to see the conserved regions and infer that those sequences may be involved in the function of proteins. The exact roles of highly conserved active site amino acid residues in alpha amylases in the catalytic process is not completely clear. Diversity in amino acid sequences of the 3 amylases reflect flexibility to change amino acids to optimize enzymatic activity under the particular condition that each alpha amylase is required to function.

Three-dimensional structures of amylases: A protein's three-dimensional structure determines its function. The 3D structures of a number of proteins have been determined by X-ray crystallography and/or Nuclear Magnetic Resonance spectroscopy. As stated earlier, the NCBI stores the sequences of millions of proteins and nucleic acids; there is also a repository for the three-dimensional structures of many of those proteins, and it is known as the **Protein Data Bank or PDB**. As in the NCBI database, each 3D structure is assigned an accession number so that it can be easily found in the databank. That number is known as the ***PDB ID;*** the PDB IDs of human, fungal and bacterial alpha amylases can be found in Table 4.

As a polypeptide starts to fold, **secondary structures**, such as **alpha helices**, **beta sheets**, turns, and loops, start to form. Secondary structures form in parts of a polypeptide with the help of hydrogen bonds. **Tertiary structure** is the total 3D conformation (shape) of an entire polypeptide chain including, alpha helices, beta sheets and any other loops, turns or bends.

Organisms found at low temperatures have membrane proteins with a higher percentage of alpha helices (provides flexibility) compared to beta sheets (provides rigidity). On the other hand, **thermostable proteins have a lower percentage of alpha helices compared to beta sheets**, since beta sheets can hold the tertiary structure together more effectively than alpha helices.

A ***fold family*** contains proteins that have the same major secondary structures in the same arrangement with the same topological connections and are clearly related by evolution. Folds are formed because of thermodynamic stability. **Structural domains** are physically independent regions of the tertiary structure, which have a specific function. Proteins may have common domains even if their overall tertiary structures and their overall functions are different.

Procedure for Bioinformatics Analysis of Amylase Sequences

To complete the Bioinformatics assignment below

1. **NCBI:** To obtain amino acid sequences of all three amylases (used in the enzyme lab) in FASTA format from the NCBI database using accession numbers: http://www.ncbi.nlm.nih.gov/
*Select the **Protein** database from the dropdown menu. Type the NCBI accession number (without the GI) into the search box. On the output screen, select **FASTA** for the single-letter version of the amino acid sequence. Copy the amino acid sequence, including the heading but simplify it to ">human amylase", etc. Paste all 3 FASTA sequences into a Word file.*

2. **PEPSTATS:** To compute statistics, e.g., total number of amino acids, number of polar and nonpolar amino acids, percentages polar/nonpolar amino acids of each amylase using the amino acid sequences (from the NCBI database): http://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/
Paste the FASTA format amino acid sequence obtained from NCBI into the query box. Do not include the heading. Complete Table 1. Note: "residues" = amino acids

3. **CLUSTAL OMEGA:** To compare amino acid sequences from all 3 amylases and find out percent identity between the amylases and to understand the differences between homology, identity and similarity: https://www.ebi.ac.uk/Tools/msa/clustalo/
Copy and paste all three FASTA sequences into the query box. Include the heading as well as the ">". Eliminate all extra line spaces between FASTA sequences. Click "Submit". For proper alignment, you may need to widen the margins in Word after you paste the results. To complete Table 2 select the Results Summary tab and the link for "percent identity matrix".

4. **BOXSHADE:** To highlight (for ease of location) specific amino acid(s) that are conserved in all three enzymes using the output from the CLUSTAL OMEGA site: https://embnet.vital-it.ch/software/BOX_form.html
Paste the entire results of the Clustal Omega analysis including the "Clustal" heading into the query box. Change the Output to "RTF-NEW", the Fraction to "1.0", and the Input to "Other". Click on "Run BOXSHADE", and you will be taken to a result page where you will find a link to your alignment editing output. Click on that link and your edited alignment file will be downloaded to your computer as an MS Word-compatible file.

5. **PDB:** To look at computer-generated models of the overall 3D structures of individual amylases and the active site and locate secondary structure motifs, e.g., alpha helices, beta pleated sheets, presence of cofactor, Ca. Use **PDB ID** at this site. http://www.rcsb.org/pdb/home/home.do. *Copy each enzyme's PDB ID from Table 4, and paste it into the PBD website's search field; each protein's 3D structure will appear in a panel on the left. You can copy the image and paste it into your assignment. Click on the tab "Sequence", above the 3D structure, to look at an image indicating alpha helix and beta sheet form in the sequence.*

BIOINFORMATICS ASSIGNMENT

1. Complete Table 1 after using the PEPSTATS tool:

Alpha amylase	Total Number of amino acids	Number of polar amino acids	% of polar amino acids	Number of non-polar amino acids	% of non-polar amino acids
Homo sapiens (salivary)					
Aspergillus oryzae					
Geobacillus stearothermophilus					

2. Complete Table 2 after obtaining the Percent Identity Matrix with CLUSTAL OMEGA:

Source of alpha amylase	*Homo sapiens (salivary)*	*Aspergillus oryzae*	*Geobacillus stearothermophilus*
Homo sapiens (salivary)	■■■		
Aspergillus oryzae		■■■	
Geobacillus stearothermophilus			■■■

3. In <u>human</u> salivary amylase, the amino acids that make up the active site are aspartate197 (D197), glutamate233 (E233) and aspartate300 (D300) (the numbers indicate the order of these amino acids in the human sequence). Using the BOXSHADE alignment, determine if these amino acids <u>are</u> **conserved** in the fungal and bacterial amylases.

4. What secondary structures do you recognize? What is the importance of secondary structures?

5. The percent alpha helix and beta sheet content for all three amylase enzymes is shown in Table 3 below.

	Human	Fungal	Bacterial
% Alpha helices	27	34	23
% Beta sheets	23	21	30

Discuss the differences in alpha helix and beta sheet content between the three amylases. **What do these differences signify?**

6. Obtain the 3D structures of all three amylases, and attach them to your assignment. Draw arrows and label an alpha helix and a beta sheet on each structure.

Acknowledgement

APPENDIX A
Data for use in the study of amylase activity from different sources

Data on the effect of <u>temperature on *Homo sapiens* salivary amylase</u> activity

Temperature (^{0}C)	Absorbance at 540nm	Maltose produced (mg/ml)	Amylase Activity (mg/ml/min)
0	0.107		
25	0.568		
37	0.874		
45	0.342		
65	0.031		
85	0.008		

Data on the effect of <u>pH on *Homo sapiens* salivary amylase</u> activity

pH	Absorbance at 540nm	Maltose produced (mg/ml)	Amylase Activity (mg/ml/min)
1	0.029		
3	0.103		
5	0.267		
7	0.945		
9	0.112		
12	0.015		

Data on the effect of <u>temperature on *Aspergillus oryzae* amylase</u> activity

Temperature (^{0}C)	Absorbance at 540nm	Maltose produced (mg/ml)	Amylase Activity (mg/ml/min)
0	0.129		
25	0.246		
37	0.539		
45	0.839		
65	0.137		
85	0.024		

Data on the effect of <u>pH on *Aspergillus oryzae* amylase</u> activity

pH	Absorbance at 540nm	Maltose produced (mg/ml)	Amylase Activity (mg/ml/min)
1	0.119		
3	0.231		
5	0.945		
7	0.367		
9	0.091		
12	0.016		

Data on the effect of <u>temperature on *Geobacillus stearothermophilus* amylase</u> activity

Temperature (^{0}C)	Absorbance at 540nm	Maltose produced (mg/ml)	Amylase Activity (mg/ml/min)
0	0.116		
25	0.237		
37	0.309		
45	0.391		
65	0.895		
85	1.029		

Data on the effect of <u>pH on *Geobacillus stearothermophilus* amylase</u> activity

pH	Absorbance at 540nm	Maltose produced (mg/ml)	Amylase Activity (mg/ml/min)
1	0.107		
3	0.209		
5	0.637		
7	0.995		
9	0.048		
12	0.037		

APPENDIX B

Links to videos used in this lab

The following videos have been included to help students understand this lab exercise, especially for online delivery.

1. **Introduction to enzymes' role in biochemical reactions:** a review of proteins/enzymes, the amylase enzyme, and the reaction it catalyzes**.**

   https://www.youtube.com/watch?v=V6wHXtO9klA

2. **Maltose standard curve**: constructing and using this graph will allow you to determine amylase activity.

   https://youtu.be/KNMz0pSgYbk

3. **Testing your hypothesis on amylase activity:** review of the scientific method, the experimental design, and data analysis to determine the effect of an environmental factor on amylase activity.
   https://youtu.be/_BacifzIP64

4. **Using micropipettes:**

   https://youtu.be/uEy_NGDfo_8

5. **How to use a spectrophotometer:**

   https://youtu.be/u3C6QXD8gjQ