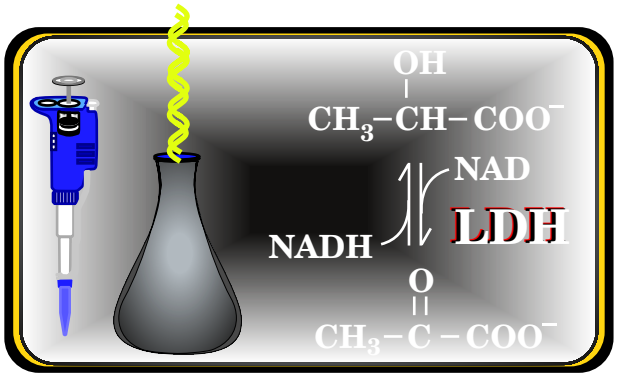
ROSE-HULMAN
INSTITUTE OF TECHNOLOGY
CHEMISTRY



CHEM 433
BIOCHEMISTRY LABORATORY MANUAL

8th Edition
Spring, 2013

Table of Contents

Introduction to the Laboratory....3
Keeping a Laboratory Notebook....5
Laboratory Reports....7
Use of Pipetmen....16
Absorbance Spectroscopy....19
Fluorescence Spectroscopy....21
Dilutions....23

The Art and Science of Protein Purification....25
Purification of Lactate Dehydrogenase from Chicken, Part I....32
Purification of Lactate Dehydrogenase from Chicken, Part II....37
LDH Enzyme Assays....41
Bradford Protein Assay....44
Calculation Hints: Setting Up a Purification Table....46

Characterization of LDH....48
LDH Analysis: SDS PAGE....48
Gel Filtration Chromatography....57
Western Blotting....62
Native Gel Electrophoresis....67
Protein Crystallography....71

Introduction to Enzyme Kinetics....75
[LDH] *versus* Activity and Lactate K_m Determination....86
Pyruvate K_m Determination....89
LDH Inhibition Kinetics....91
LDH Inhibitor Type Determination....97
Chemical Modification of Lactate Dehydrogenase....99

Cloning of Human Lactate Dehydrogenase....101
Polymerase Chain Reaction and Plasmid Preparation....104
Agarose Gel Electrophoresis and Restriction Digests....112
Ligation and Transformation....116
Selection and Screening....120
Screening and Sequencing....124
Activity Measurements....129

Definitions....130

Introduction to the Laboratory

This course is intended to introduce you to some of the most widely used experimental procedures in biochemistry, including protein purification and characterization, enzyme assays and kinetics, instrumental techniques, and, if time allows, DNA isolation and manipulation. You will also gain some familiarity with some of the types of equipment frequently used in biochemistry.

Research is often a collaborative effort in which many people may contribute to different aspects of a given project. Few papers in the biochemical literature are written by single authors; the vast majority of papers have at least two authors, and for many papers more than ten people contributed to the work. In part to provide a more authentic experience of actual lab work, experiments will be done in groups of two or three. You may choose a partner, or you can ask to be assigned to a group.

Prior to each lab period, you will need to spend some time reading the Laboratory Manual. This reading will provide background information and an outline of the procedures to be performed. If you do not do this, you will find yourself wasting large amounts of class time, and annoying your lab partner(s) and your instructor.

The biochemistry laboratory course, like all laboratory courses, is an exploration of procedures. This means that, in order to get full benefit from the course, you will need to read the manual, and you should participate as much as possible in the discussions. You should ask questions in or out of class. You should also try to participate in the actual lab work (and not simply allow your lab partners to do things for you). The more effort you put into the course work, the more you will learn. The class is an opportunity to learn valuable skills; take full advantage of it!

SAFETY: Laboratories contain hazards of various kinds. Everyone is ***required*** to wear **closed-toe shoes, long pants, and safety glasses or goggles with side shields** while performing laboratory work. Students should not work in the laboratory alone. When in a laboratory, you should always behave in a fashion that protects the safety of yourself and those around you.

Some of the chemicals used are toxic, mutagenic, or teratogenic. If you believe that you have a health condition that puts you at exceptional risk, or believe yourself to be pregnant, please see your instructor in private to discuss the issue. If you have questions or concerns about exposure to hazardous chemicals, please consult your instructor.

PHILOSOPHICAL ISSUES: Scientific research involves an exploration of the unknown. In some classes, a question has a single "correct" answer, which is known to the instructor, and is imparted to the students. In research, however, the correct answer is rarely known ahead of time, and must instead be inferred from the experimental results. In some cases, the original conclusion may need to be modified in light of additional experimental data. Researchers must therefore become accustomed to some level of uncertainty about the "correct" answer to any

experimental question, and must always remain open to experimental evidence that contradicts a hypothesis that has arisen from previous experiments. Your task as a scientist will be to consider **your** data, and to attempt to interpret it. In this context, "wrong" answers are answers that are contradicted by your data or that do not arise logically from the data you have collected.

This uncertainty as to the "correct" answer means that you must be careful when reporting what you did and what you observed, especially if you observe something unexpected. Humans are good at fooling themselves; **you need to guard against reporting what you *expect to see* rather than what you actually *did* see**. Scientific fraud, in which people intentionally report false data, is considered very serious because it results in a difficult-to-overcome belief in an answer that is in conflict with the truth. You will occasionally see retractions, in which a scientist publishes a statement that information in a previously published paper is the result of an artifact, and is not a reflection of the "correct" answer. Avoiding the embarrassment of publishing a retraction is one reason for the care that people take in performing experiments and in interpreting the results. More importantly, incorrect research results may cost the people who attempt to apply the results their money, and in some cases, their lives.

Another ethical issue concerns the proper citation of the sources of information you use for any scientific writing. You should ***always*** properly reference the authors of papers or books you consult. This also means that you should cite the inventors of methods that you use for your experiments. If you do not, you are, in effect, claiming credit for work performed by others. Failing to do so is a form of plagiarism, and is a serious ethical offense.

General Information
Keeping a Laboratory Notebook

All students will be **required** to maintain a laboratory notebook. The notebook will be used for the recording of laboratory data and calculations, and will be a critically important source of information while writing your lab reports.

The purpose of a biochemistry laboratory notebook is to allow anyone with some biochemical knowledge to understand **exactly** what you did. You need to record the procedures in sufficient detail so as to be able to repeat them, and you must be able to understand exactly what your results were. You will need good notes to be able to write your lab reports; in addition, as your understanding of biochemistry improves, your notebook may allow you to figure out why some parts of your experiments did not work as expected.

Companies that perform research require their employees to keep proper notebooks. In many of these companies, company policy dictates that any work not recorded in the notebook is treated as though it was never actually performed. As a result, the work must be repeated, which tends to have deleterious effects on the career opportunities of the employees involved. In cases of disputes as to priority, notebook dates are sometimes used to indicate exactly when an experiment was performed. Ownership of patents (and in some cases large amounts of money) can therefore be critically dependent on keeping a proper notebook. Instruction in keeping laboratory notebooks is therefore a major part of most laboratory courses.

In your notebook, you should begin each experiment with a **title**, a **date**, and a statement of the **objective** of the planned work. You should also **record exactly what you did at each step** (being sure to mention anything that you did that differed from the information in the Manual). In addition, you should record any numerical information, such as the weights of reagents used, absorbance readings, enzyme activities, protein concentrations, and buffer concentrations.

Most experiments will extend over several days, and over several pages in your notebook. To allow you to keep track of what you have done, you should include the day's date at the top of each page. Including sub-titles for each page may make it easier to keep track of what you did at each step.

Everything you do should be recorded **directly** into your lab notebook in **pen**. If you make a mistake, draw a line through the incorrect information, and write the correction next to the mistake. (It may turn out that the original information was correct after all, so do not obliterate the original information by erasing it, or by removing the page from your notebook.) Any calculations performed should be written directly into your book. Hard copies of work done on a computer and printouts from laboratory instruments should be taped directly into your lab notebook.

Writing important information on scrap paper, and then recording it in your notebook later is **not** acceptable. **If you are writing something while in the laboratory, you should be writing it directly into your notebook.**

At each step in your experiment (after each assay or measurement), in addition to the results, record your thoughts regarding the experiment and how you think it is going. Record your mistakes, and your attempts to rectify them. Record the calculations involved in any type of data analysis, as well as explanations for both what you did and what you think it means. A research project is a journey into the unknown; your laboratory notebook is usually your only guide through the forests of uncertainty.

It is also a good idea to look over your notebook periodically during the quarter, and to make notes of things that you do not understand, so that you can ask questions before the lab reports are due.

Do not say "well, I will remember what this means"; instead, ***write it down***! Do not say "I will remember what I was thinking while I did this experiment"; instead, ***write it down***! If you use your lab notebook properly, you will find that the writing of your lab reports is much easier; in addition, you will be developing valuable habits for the future.

General Information
Laboratory Reports

The laboratory reports are major written assignments, due at intervals during the quarter. The laboratory reports should be written in the form of a scientific paper. To help you learn to write a scientific paper correctly, the laboratory reports will be due in sections, with each report building on the previous one. **Each report should contain all of the information from the previous report, plus all of the new work**. You should incorporate the instructor's suggestions, using these comments to guide you in the generation of the new sections. Note that the later laboratory reports will be graded more stringently than earlier ones: you are expected to learn from your mistakes!

Laboratory reports should contain the following sections:

Title Page
Introduction
Abstract
Materials and Methods
Results
Discussion
References
Acknowledgments
Appendix

All of the laboratory reports are expected to be well formatted, word-processed documents, written in standard scientific American English. The use of spell-checkers and grammar checkers is strongly recommended.

(Note: the Appendix does not have to be as neatly formatted as the rest of the report, and, if necessary, may be handwritten.)

Content
In scientific research, results are reported to the world in the form of scientific papers published in the peer-reviewed scientific literature. These papers are not only important in disseminating the results of the research, but are critical for essentially all aspects of career advancement for the scientists involved. Learning to write a proper scientific paper is therefore an important part of the education of all scientists.

Scientific papers are expected to be written in a well-defined format. The overall format is generally similar in all journals, although the specific details vary somewhat. In this class, the laboratory reports should be in the form of a paper in the *Journal of Biological Chemistry*. Looking for papers in the *Journal of Biological Chemistry* to use as examples is strongly recommended. (Note that the formatting that you should attempt to emulate applies to **content**; you do **not** need to spend

time generating the specific page layout of a *Journal of Biological Chemistry* paper. The preferred page layout for lab report submission has the body of your paper in double-spaced text.)

Many scientists have their own preferred ways of writing papers. Most scientists, however, use an iterative process of writing, in which they write the paper, and then rewrite it several times before submitting the paper to the journal for review and (hopefully) publication. In addition, most papers are written in an order that deviates from the final format. A common procedure is to write the Methods section first, followed by the Results section. The Methods section is a simple description of procedures and can be written before the experimental results have been analyzed. The Results section contains the observations that constitute the study to be published. Once these sections are written, most people write an incomplete draft of the Discussion section that explains the results in the context of the paper.

After you write a draft version of the Results section, and after you put some thought into interpreting the results, you should write a draft of the Introduction. When writing your Introduction, you should think of the Introduction as an episode of "Jeopardy": the Results are the answers, and now it is necessary to come up with corresponding questions. You do not need to write the "questions" in the form of a question, but you should think about raising questions in the readers' mind that you will then answer in the Results and Discussion sections.

After writing the Introduction, you should then look at how you have written Introduction, and rewrite the Results section to more clearly answer the questions you raised in the Introduction, and then write the Discussion to interpret and clarify the answers. When properly done, each rewrite acts as an impetus for the rewriting of a different section, until all of the sections fit together into a coherent story.

Finally, after all of the other sections have been written, you can write the abstract by extracting the most important information from each section and combining the information into a single paragraph.

You should keep these general concepts for writing a paper in mind while considering the content of each section. The expected content for each section of a scientific paper is discussed below. (Remember that you will probably not write the paper in this order.)

Title Page: This should include the title of your report, the author's name (*i.e. your* name), your lab partner's name(s), and your address (your e-mail address is sufficient).

Abstract: This should be a *brief* version of the **entire** paper. It therefore should include a brief introduction, methods, results, and discussion, expressed in ~200 words. This truncation is normally achieved in part by greatly abbreviating the methods portion, unless the methods involved are novel or are crucial to understanding the findings presented.

Thousands of papers are published every week. Most literature database search engines include the title and abstract, but do not include the remainder of the paper. In writing the abstract, remember that the vast majority of readers probably will **not** read the paper, because they lack the time. Therefore, in order to present your information to the largest possible audience, you need to have an abstract that is clearly written, that is understandable without having to read the paper, and that contains all of the relevant findings from the paper.

The abstract should end with the **overall conclusions** from the paper; once again, this is important because you **want** people to know what you have discovered. Your job/grant funding/promotions/fame and fortune/ability to do more experiments/ability to retire to the exotic locale of your choice may depend on having people understand what you have done. (This applies to the entire paper, but the abstract tends to be at least skimmed by vast numbers of people who will never read the paper.)

Introduction: This section should include background information setting up the scientific problem you are attempting to address and the overall goal of the experiments you performed. What is the hypothesis you are testing? What directly relevant information is necessary to understand this hypothesis and why is it important? What is not known that you hope to address? What are you planning to attempt to accomplish? (Very briefly) How did you accomplish this?

In writing an introduction, you are attempting to orient the readers, so that they will know what to consider as they read the rest of the paper. This means that you should carefully consider whether you are presenting information that is irrelevant or misleading. If you discuss an issue related to your protein in the introduction, the reader will expect you to address that issue in the remainder of your paper. In addition, after having read your introduction, the reader should have an appreciation of the questions you were attempting to address with your experiments and why these questions are important. If someone can read your introduction without wanting to read the rest of your paper to find the answer to the burning questions that you raised, you have not written your introduction properly!

Methods: This section should be a *concise* summary of what you did. It should include enough detail so that any reasonably intelligent biochemist could repeat your work, but not a minute-by-minute recitation of the hours you spent performing the experiment. One common mistake is to include information that belongs in the Results section; the Methods section is for **methods**. For example, a description of a protein assay should describe the procedure used, but generally should **not** include a list of the samples measured in the assay. On the other hand, a common mistake is to fail to include some methods, such as the techniques used to analyze the data obtained during the study.

When most people read a paper, they tend to *skip* the Methods section unless they need to know exactly how an experiment was performed. This means that they will not read the Methods unless they do not believe your description in the Results section, or because they work in the field and want to see if you used a novel technique. Because many people skip the Methods section, the Methods section

should only be a description of the methods used. With the possible exception of one-time events such as plasmid constructions, it is rarely a good idea to include results in the Methods section. If you do include results in the Methods section, these results should be at least summarized in the Results section also.

The Methods section should also contain the source of the important reagents and identifying information for any equipment used. Because research reagents of high quality are available from many vendors, the precise source of most reagents is much less important than it once was. It is common practice, however, to state in the Methods section that, for example, "the ADP-glucose pyrophosphorylase expression vector was a generous gift of Dr. C. Meyer".

Results: This section should be a description of what you did **in words**, illustrated with figures and tables. It is **not** enough merely to have several figures; you need to **explain** what each figure means. Try to avoid merely listing results in the text; instead, explain the findings and briefly fit them into the overall context of the paper.

For each set of experiments, you need to consider the following questions: What are you doing? Why and how are you doing it? What was the rationale for the methods you employed? What is the point of the experiment you are about to describe? What strategy are you using to address the experimental question you are asking?

None of your answers to the above questions should be lengthy, but you do need to consider these questions in writing your report. It may be totally obvious to **you** why you performed your brilliant experiment, but unless you explain the purpose and rationale behind the experiment, your flawless reasoning may not be obvious to your readers.

Remember that you are telling a story to people who have not done the experiments. You cannot assume that the reader will know what you are doing and why. In addition, you are telling a story that people will be predisposed to **disbelieve**. You therefore need to present your information as clearly as possible. If you do so, people will (at worst) understand what they are criticizing, and (at best) see that you have put enough thought and effort into your work as to make it likely that you are trustworthy.

What data do you need to report? Do **not** report data merely because it is available. Instead, report data to make a point. You are trying to tell a factual story. This means that you **cannot** lie to your readers. On the other hand, if you perform an irrelevant experiment, reporting the results may be confusing. For example, if you perform five SDS-PAGE electrophoresis experiments that show essentially the same results, you do not need to include the results of each individual gel.

In reporting the results of an experiment that yielded numerical data, it is poor writing technique to simply list in the text the same values listed in a table or shown in a graph. The raw numbers are meaningless unless put into context. In other words, cite in the text only the *important* numbers, and explain *why* these values are important.

For reporting numbers in the text, convert the numbers to reasonable values. A number such as 0.0014567 mg/μl is **not** reasonable for two reasons: 1) converting the value to 1.4567 mg/ml results in a number that is much easier to read, and 2) the number of significant figures reported seems excessive (unless you **really** believe that your experiment was accurate to five significant figures).

As an example, you will be writing a description of LDH purification and LDH enzyme assays in your Results section. You should consider the following in writing this section.

Purification: Why did you perform the purification? What strategy did you employ for the purification? Why did you use the steps you used and not others? During the purification, what step resulted in the greatest purification? When did you observe the LDH to elute from the column? Was this expected, unexpected, or did you have no basis for making a prediction? Is there a figure you could generate to clarify your results? (Is a figure necessary to clarify your results?) Based on your data, was your purification successful or unsuccessful? Why? Do you have any data other than fold-purification to indicate whether your purification was successful? How did your purification compare to literature values obtained for similar proteins?

Characterization: Scientific research involves *intelligent* observation. In other words, you need to look at your data critically, and to attempt to understand everything it is telling you. Why did you run gel filtration chromatography or SDS PAGE on your protein sample? Simply looking at an SDS PAGE or gel filtration experiment as a method for determining the molecular weight of your protein may result in your missing important information about your protein. If you ran these experiments on a sample that you believe to be highly purified, you should examine the results and compare them to what you would expect to see for a completely homogeneous preparation (in other words, for a preparation containing zero contaminants). Thus, in examining a gel filtration chromatogram, do you see any unexpected peaks? For example, if you expect to have a single, monomeric protein, and you see two peaks on a chromatogram, you need to figure out which peak is your protein and which is a contaminant, how much of the contaminating material is present, where it came from (especially if it was not there previously), and whether it is necessary to perform additional purification steps. It is possible that the second peak is a loosely associated protein that interacts with your protein; in which case, you may be learning useful information about your protein. Do you see evidence on an SDS PAGE on the same sample for similar contaminants? Why or why not?

Enzyme assay**:** what can you learn from each enzyme assay? (If the answer is "nothing", is it worth including these results in the paper?) How do you know that the assay results are valid? What assumptions are you making about the enzyme reaction actually occurring in the reaction tube? Are these assumptions likely to be correct for each assay? Are these assumptions likely to be correct for some assays but not for others? What controls did you run to ensure that the results were at least potentially meaningful?

In some cases, the answers to the above questions do not need to be stated explicitly. However, you **always** need to consider the answers before writing the paper. Knowingly incorporating the results of a flawed experiment in a paper is a good way to lose grant funding or become unemployed, and may result in your finding yourself in court defending yourself in a lawsuit or in a criminal trial. This does not mean that experiments that later turn out to be less informative than you would like are useless, but merely means that you should look carefully at your data, and try to understand the validity of each experiment before mentioning it in a written document.

Discussion: This section should begin with a brief summary of your results, and an explanation of what they mean. What were you hoping to accomplish? What did you discover as a result of your experiments? Which of your results are interesting? What can you say about your hypotheses now that you have additional data? What did you expect to see? Did you see what you expected? Did you find surprising results?

At least in part, the Discussion section should be the section in which you answer the questions you raised in the Introduction. Sometimes the answer is that your original hypothesis turned out to be flawed; in this case, you point out how the data indicate the flaws, and propose a brilliant new hypothesis to account for your observations. Sometimes your original hypothesis is supported by the data, in which case you point out how your original brilliant concept predicted your results.

You should end your discussion section with your conclusions. Did your experiment achieve your goals? How are your results going to change the world?

Figures and figure legends: In writing a paper, figures can be extremely useful. They are rarely, however, self-explanatory. This means that you need to refer to the figure in the text. In addition, you need to include some relevant information in the figure legend, so that people simply glancing through the paper can derive useful information from the figures.

As an example, in a figure of a gel, you should indicate the identity of the samples loaded in the figure legend. If more than one band is present in an important lane, it is often a good idea to highlight the important band in some way. (Note: in doing so, **do not** write on the actual lane, because this may obscure some of your data; instead, place an arrow or other marker beside the gel, or beside the lane.)

Designing figures requires considerable thought. What point are you trying to make with the figure? Is the point necessary? If the point is a necessary one, how can the figure be used to illustrate the point as clearly as possible? Can you design a figure to present more information, or present the information more clearly?

Figure legends can be extremely useful in allowing you to present relevant information that would disrupt the orderly flow of ideas in the text. The figure legends are also necessary in clarifying the information presented in the figure.

References: In any scholarly endeavor, it is customary to give credit to your sources of information. The Reference section allows you to properly credit the originators of the information you are presenting. Where did your introductory information come from? Where did your methods come from? (Note that, unless you invented the method, you should always reference the paper that first described the work.)

Acknowledgments: In scientific papers, it is customary to thank the agency that funded the research. In addition, it is polite to acknowledge gifts of reagents or other supplies. In contrast, if you purchased the reagent, the source of the reagent should be cited in the Methods section, and not in the Acknowledgments.

Appendix: Finally, the report should contain an appendix that contains your raw data and the calculations that you used to reduce your data to understandable form. In a real paper, Appendix sections are only included for the description of novel calculations; in this course, the Appendix is included so that your lab instructor can verify your calculations.

In each section, attempt to organize the information you are presenting logically. Scientific papers are written for intelligent people who have not done the experiments you are describing. If your report is disorganized they may not understand it. If you do not write well, the reader will not believe your conclusions. (In the real world, a poorly written paper will not be published, and you will not get grant funding! In this class, if you instructor does not believe your conclusions, you will not get a good grade.)

The list of questions below is designed to help you write each section of the report correctly. Reading over this list of questions before writing a draft of the report is strongly recommended. Reading these questions after writing your first draft, and using the questions to guide your revisions is also strongly recommended.

Criteria for Judging Lab Reports:

General:

Does it contain the required sections?
Is it clearly written?
Does it use scientific terms properly?
Does it use good grammar?
Are the words spelled correctly?
Are the calculations performed correctly?
Is it unnecessarily long?
Is the title meaningful?
Does the title page contain the author's name and address?
Does the title page contain the name(s) of the author's lab partners?

Abstract:

Does it introduce the overall topic?
Does it explain the hypothesis being tested?
Are the important methods described?
Does it reach logical conclusions supported by the data?
Does it flow well? Is it logically written? Is it concise?

Introduction:

Does it give general background?
Does it point out poorly understood or unknown factors related to the study?
Does it raise questions?
Does it explain the hypothesis being tested?
Does it discuss the significance of the work?
Does it flow well? Is it logically written? Is it concise?

Materials and Methods:

Could the experiments be understood based on the information given?
Does it include the source of the reagents?
Does it include information that belongs in the Results section?
Does it describe all of the methods used?
Is it excessively long?

Results:

Does it explain the rationale and strategy for the experiments performed?
Does it describe, in words, what was done?
Does it answer (or at least address) the questions raised in the Introduction?
Does it flow well? Is it logically written? Is it concise?

Discussion:

Does it summarize the findings obtained in the Results section?

Does it discuss the expected results?
Does it discuss the unexpected results?
Does it answer (or at least address) the questions raised in the Introduction?
Does it reach conclusions?
Does it explain why the conclusions are important?
Does it flow well? Is it logically written? Is it concise?

Figures:

Are the figures well designed?
Do the figures include informative legends?
Do the figures present information useful for understanding the text?

Tables:

Are the tables well designed?
Do the tables present information useful for understanding the text?
Is the information in the tables redundant?

Acknowledgments:

Are the sources of funding given credit?

References:

Is the information obtained from published sources properly referenced?

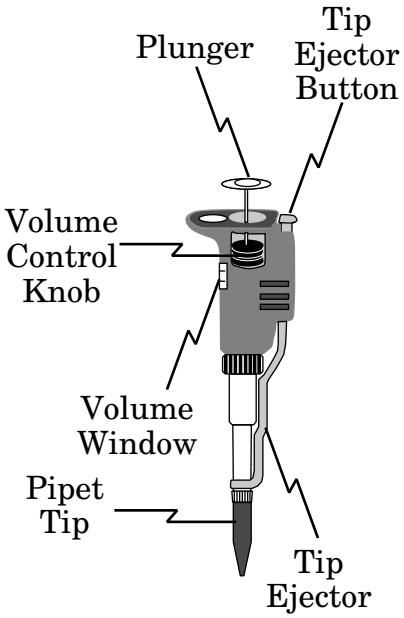
Appendix:

Are the raw data and the calculations included?

Introduction to Techniques
Use of Pipetmen

In biochemistry, the ability to accurately and reproducibly measure and transfer small volumes of liquids is critical for obtaining useful results. For volumes less than 1 ml, the most common method for measuring liquid volumes involves the use of a device known as a pipetman. (Note: "Pipetman" is the brand name of the most commonly used of these types of pipets; however, all of these pipetting devices work on similar principles.)

A drawing of one brand of pipetman is shown at right. (Depending on the equipment in your laboratory, the devices you use may not look exactly like the one shown.) The pipetmen used in this course come in three different types: P1000, P200, and P20.



P1000 are useful for volumes from 200 to 1000 μl. P200 are useful for volumes from 20 to 200 μl. P20 are useful for volumes from 0.5 to 20 μl. Make sure that you are using the correct pipetman for the volume you need. Also, make sure that the pipetman is actually set for the volume you need by looking in the "volume window", and, if necessary, turning the "volume control knob" until the pipetman displays the correct volume (the pipetmen do *not* read your mind; because several people will use the pipets, they may not always be set as you expect them to be). **Do not attempt to set pipetmen for volumes larger than their maximum, or for volumes less than zero**; doing so will damage the pipetman.

All pipetmen use disposable tips (**do not pipet liquids without using the appropriate tip**, because this will contaminate the pipetman and may damage it). When attaching the tip, make certain that the tip is the correct type for the pipetman you are using, and that the tip is properly seated on the end of the pipetman.

Try depressing the plunger. As the plunger depresses, you will feel a sudden increase in resistance. This is the first "stop". If you continue pushing, you will find a point where the plunger no longer moves downward (the second "stop"). When using the pipet, depress the plunger to the first "stop", place the tip into the liquid, and **in a slow, controlled manner,** allow the plunger to move upwards. (Do not simply let the plunger go; doing so will cause the liquid to splatter within the tip, resulting in inaccurate volumes and in contamination of the pipet.)

Now, take the pipetman (carrying the pipetted liquid in the tip) to the container to which you wish to add liquid. Depress the plunger to the first, and then to the second stop. If you watch carefully, you will note that depressing to the second stop expels all of the liquid from the tip. (Actually, this is true for most aqueous solutions. In some cases, however, such as for organic solvents, or for solutions

containing large amounts of protein, it is often difficult to get all of the liquid out of the tip. In these cases, it is best to "wet" the tip, by pipetting the original solution once, expelling it, and then taking up the liquid a second time.)

Although pipetmen are tremendously useful, they have a potential drawback. If used improperly, pipetmen will transfer inaccurate volumes. In addition, pipetmen may lose calibration. If used incautiously, therefore, pipetmen may yield misleading or even totally useless results. Checking the calibration of pipetmen is a simple procedure that can save considerable time, energy, and reagents. The simplest method for checking pipetman calibration is to measure the mass of a pipetted volume of water. (This also checks the technique of the operator.)

Accuracy is a measure of proximity to the true value or the expected value for a measurement. **Precision** is a measure of reproducibility. For example, obtaining weights of 0.5, 1.0, and 1.5 grams for a P1000 set for 1000 μl would be accurate, but not precise; obtaining 0.67, 0.68, and 0.67 g for the same setting would be precise, but not accurate.

Calibration procedure:

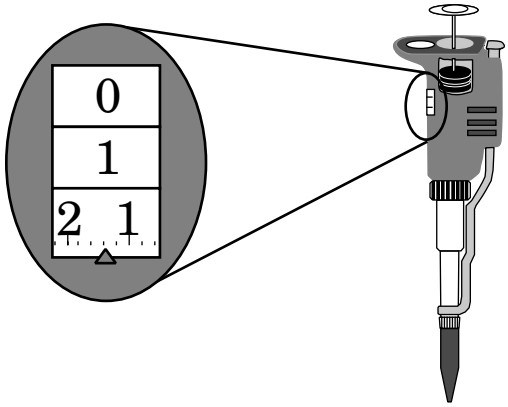
1. Acquire Pipetmen and the correct size tips.
2. Place a weigh boat on the balance and tare the weight to zero.
3. Draw up the designated volume of deionized water into the pipet tip and dispense it onto the weigh boat. Record the weight of the water added.
4. Repeat the procedure twice for each volume (yielding a total of three weights for each volume).

The P20 uses very small volumes, which have very small weights. In order to obtain accurate readings with the relatively low precision balances available, you may need to pipet the volume of water several times (5 or 10 times is recommended) for each volume being tested. For example, pipet 10 μl 10 times in rapid succession (to avoid evaporative losses during the measurement) and record the weight of the 100 μl total volume as one measurement.

Depending on the sample, and on your technique, you should obtain accuracies and precisions of ± 2 μl for P1000, ± 0.5 μl for P200, and ± 0.1 μl for P20.

Study Questions

1. Your lab partner hands you a P200 Pipetman set as shown in the diagram at right. For what volume is it set? Is this the proper Pipetman for this volume? Why (or why not)?



2. You have a 0.5 M stock solution of Tris base. How would you make 100 ml of 0.03 M Tris base?

3. If you perform a 1:4 dilution on 50 mM Tris base, what is the final concentration?

4. Is the extinction coefficient for a molecule the same at all wavelengths?

5. You prepare several dilutions of an unknown compound. You measure the absorbance of each solution at 340 nm using a 1 cm cuvette (your results are listed in the table below). What is the extinction coefficient (in $(M \cdot cm)^{-1}$) of the compound? Are all of the values likely to be equally accurate? Why? (**Hints**: assume that each of the individual values contains some degree of experimental error, and assume that the Ocean Optics Spectrometer used tends to lose accuracy significant at when less than about 3% of the incident beam reaches the detector.)

Concentration (μM)	Absorbance at 340 nm
4	0.023
12	0.077
36	0.225
108	0.670
324	1.681

Introduction to Techniques
Absorbance Spectroscopy

A spectrophotometer is an instrument for measuring the absorbance of a solution. Absorbance is a useful quantity. The Beer-Lambert law states that:

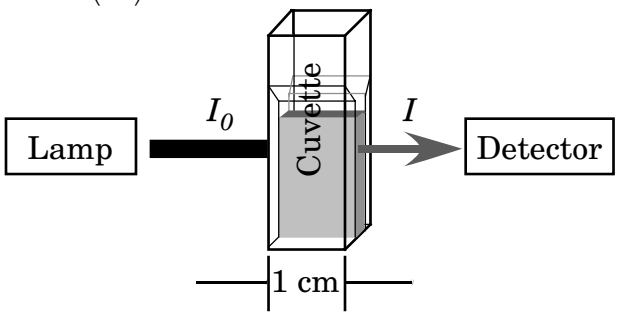
$$A = \varepsilon cl$$

where A is the absorbance of the sample at a particular wavelength, **ε** is the extinction coefficient for the compound at that wavelength in $(M \bullet cm)^{-1}$, ***c*** is the molar concentration of the absorbing species, and ***l*** is the path length of the solution in cm. Thus, if the extinction coefficient of an absorbing species is known, the absorbance of the solution can be used to calculate the concentration of the absorbing species in solution. (This assumes that the species of interest is the only material that absorbs at the wavelength being measured.)

The above is an explanation of *why* we measure absorbance: absorbance allows us to calculate the concentration of compounds in solution. However, it does not explain *what* absorbance is. Another definition of absorbance is:

$$A = \log\left(\frac{I_0}{I}\right)$$

where I_0 is the amount of light entering the sample, and I is the amount of light leaving the sample. Absorbance is therefore a measure of the portion of the light leaving the lamp that actually makes it to the detector. A little thought will reveal that when absorbance = 1, only 10% of the light is reaching the detector; when absorbance = 2, only 1% of the light is reaching the detector.



The typical internal arrangement of a Spectrophotometer

For most instruments, **absorbance values greater than 2 are unreliable, because too little light is reaching the detector to allow accurate measurements.** When measuring absorbance, note the values; if the reading is greater than 2, dilute the sample and repeat the measurement.

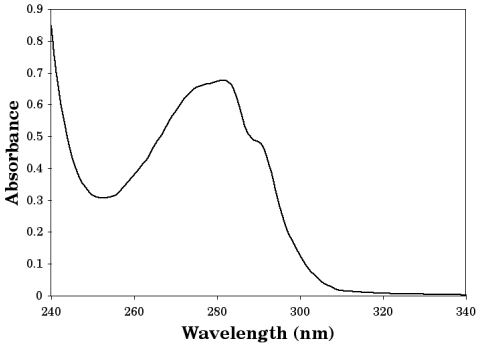
Spectrophotometers measure the decrease in the amount of light reaching the detector. A spectrophotometer will interpret fingerprints on the optical face of the cuvette, air bubbles, and objects swimming in your solution as absorbance; you therefore need to **look carefully at your cuvette** before putting it into the spectrophotometer to make sure that your reading is not subject to these types of artifacts.

Cuvettes are usually rectangular solid objects 1 cm across (as shown in the above figure). In some cases, the liquid reservoir is not square; in those cases, make sure

that the 1 cm dimension is aligned with the light path (note the orientation in the diagram above.)

Some cuvettes are designed for visible light only. When the spectrophotometer is set for ultraviolet wavelengths (wavelengths of 340 nm or less) **make sure that your cuvette does not have a large absorbance when it contains only water**.

The term "spectroscopy" comes from the word "spectrum" which originally referred to the multiple colors of light apparent in an analysis of white light using a prism. "Spectroscopy" therefore implies the use of multiple wavelengths of light. Spectrophotometers have the ability to specifically measure absorbance at specific wavelengths. The most commonly used method to allow this involves a "monochromator", a device (a prism, or more commonly, a diffraction grating) that splits the incident light into its component wavelengths, and allows only light of the desired wavelength to reach the sample. The ability to measure absorbance at different wavelengths is very useful, because the **extinction coefficient of a compound varies with wavelength**. In addition, an absorbance spectrum for a compound can vary dramatically depending both on minor changes to its chemical composition, and on the environment (such as the solvent) around the compound.



The graph at right shows the absorbance spectrum of a protein. The protein has a strong absorbance peak near 280 nm, but exhibits very little absorbance at longer wavelengths. For this protein, the only chromophores (chemical groups within a compound that absorb light) are the aromatic amino acids tryptophan and tyrosine. For many proteins, these two residues are the only significant chromophores; because tryptophan and tyrosine only absorb in the ultraviolet portion of the spectrum, such proteins are colorless molecules. Colored proteins, such as hemoglobin, exhibit their color due to the prosthetic group chromophores (heme, in the case of hemoglobin) that absorb in the visible portion of the spectrum.

The extinction coefficient of a molecule at a given wavelength can be calculated using the Beer-Lambert equation from absorbance measurements for solutions of known concentration.

For proteins lacking prosthetic group chromophores, it is possible to estimate the extinction coefficient of the protein based on its amino acid composition according to the equation:

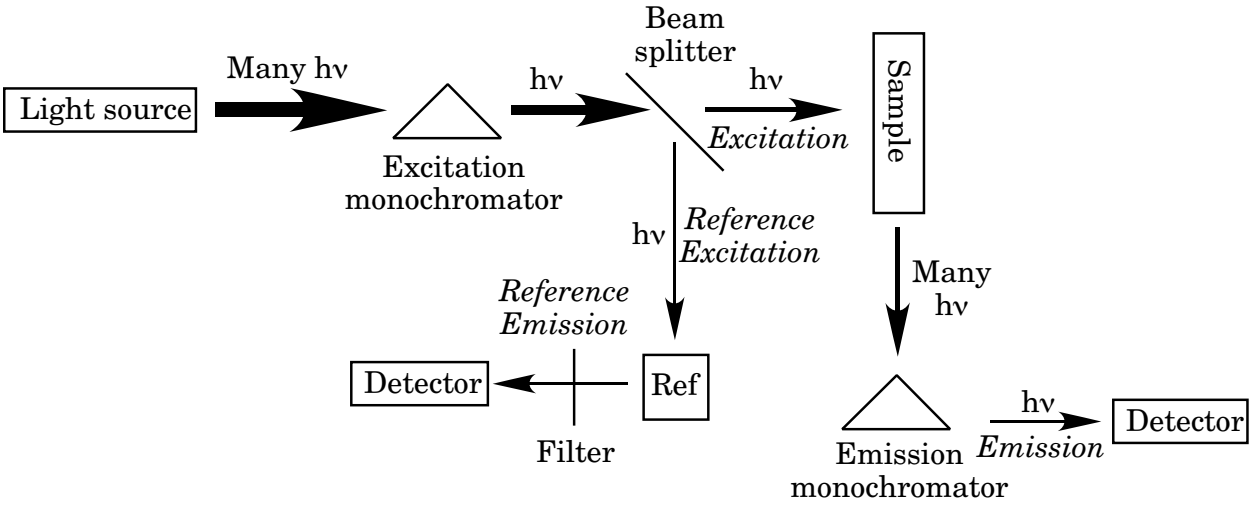
$$\varepsilon_{280} = W \bullet 5615\ (M \bullet cm)^{-1} + Y \bullet 1380\ (M \bullet cm)^{-1}$$

where W is the number of tryptophan residues, and Y is the number of tyrosine residues in the protein.[1]

[1] The values given are the average values from those given in Gill and von Hippel (1989) *Anal. Biochem*. **182**, 319-326, and Mach, Middaugh, and Lewis (1992) *Anal. Biochem*. **200**, 74-80.

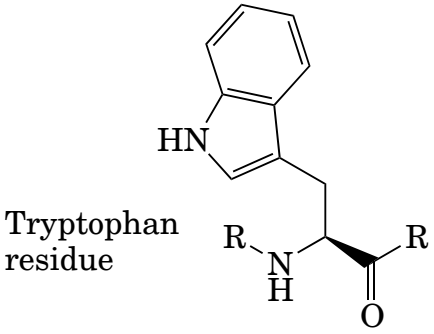
Introduction to Techniques
Fluorescence Spectroscopy

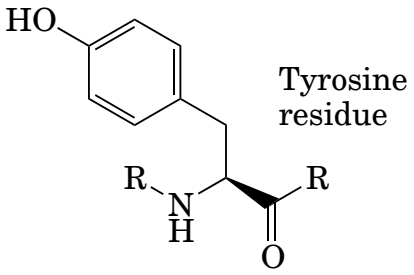
A spectrofluorometer is an instrument for measuring the fluorescence of a solution. In ultraviolet/visible absorbance spectroscopy, the incident light may result in an electronic transition to an excited state of the molecule (the extinction coefficient is a measure of the probability that excitation will occur at that wavelength). Once a molecule is in the excited state, it must lose energy to return to the ground state. One method (among many others) by which a molecule may return to the ground state is by emission of a photon; this emission is known as fluorescence.



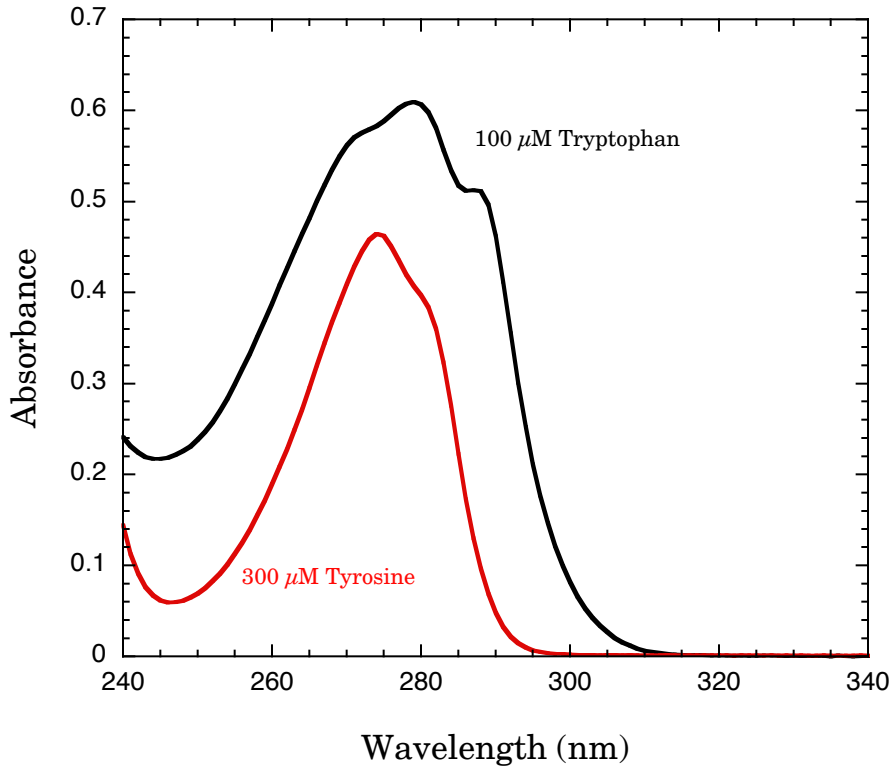
The intensity and wavelength of the emitted light can yield information about the environment of the fluorescence species. For individual molecules in solution, this may be of limited use; however, proteins are large enough that the individual fluorophores may be in different environments. In addition, conformational changes in the protein are often associated with changes in the environment around specific residues within the protein. Finally, binding of a fluorescent molecule to a protein usually results in a change in the fluorescence emission of the molecule, which allows measurements concerning the binding process.

For proteins that lack prosthetic groups, the intrinsic fluorescence is due primarily to tryptophan residues; tyrosine may contribute slightly to the fluorescence, but tyrosine fluorescence is normally observable only in proteins that lack tryptophan.





Tryptophan has an absorbance maximum at about 282 nm; however, excitation at 295 nm is often preferable, because it limits any contributions from tyrosine.



While fluorescence spectroscopy is often more sensitive than absorbance spectroscopy, it is subject to artifacts that are somewhat more difficult to correct for than is absorbance. The lamp output can vary both as a function of wavelength and as a function of time, and both monochromators and detectors generally vary in efficiency as a function of wavelength. While obtaining a baseline absorbance reading corrects for these issues, blank samples generally do not fluoresce, and therefore cannot be used to correct for these problems. While absorbance values should always be directly comparable, fluorescence intensity is relative, and frequently varies considerably between instruments. Performing routine instrument checks is highly recommended. The first is an excitation scan to verify both the excitation monochromator calibration and the intensity of the signal; xenon arc lamps have a maximum output at 467 ± 0.5 nm. Another useful calibration check is to look for the intensity and wavelength of the water Raman peak using a known excitation wavelength. Monitoring changes in the internal reference emission is a necessary to check for variations in lamp output over time. Finally, sample concentrations should be chosen to limit the absorbance at both excitation and emission wavelengths to less than 0.1 (and preferably less than 0.05) to prevent excessive attenuation of light due to absorbance phenomena.

Introduction to Techniques
Dilutions

Many solutions used in biochemistry are prepared by the dilution of a more concentrated stock solution. In preparing to make a dilution (or series of dilutions), you need to consider the goal of the procedure. This means that you need to consider both the desired final concentration and required volume of the diluted material. A simple equation allows the dilution to be calculated readily:

$$C_1V_1 = C_2V_2$$

where $\mathbf{C_1}$ is the concentration of the initial solution; $\mathbf{V_1}$ is the volume of the initial solution available to be used for dilution (this may not be the total volume of the initial solution, and instead may be a small fraction of the initial solution), $\mathbf{C_2}$ is the desired final concentration, and $\mathbf{V_2}$ is the desired final volume.

In most cases, the initial concentration and the final concentration are either known or are chosen in order to work correctly in the experiment being planned. The final volume is usually an amount that is chosen based on the amount required for a given experiment. This means that at least three of the required terms are either known or can be chosen by the experimenter.

Let us consider an example. You are setting up a standard curve. You have a stock solution of 1000 μg/ml BSA, and for one of the points on the curve, you want 200 μl of 20 μg/ml. In this case, C_1 = 1000 μg/ml; C_2 = 20 μg/ml, and V_2 = 200 μl. This leaves V_1 as the unknown value (*i.e*. how much of the stock solution must be diluted to 200 μl final volume to yield the desired concentration). Rearranging the dilution equation gives:

$$V_1 = V_2 \frac{C_2}{C_1} \quad \text{and therefore} \quad 4\ \mu\text{l} = 200\ \mu\text{l}\left(\frac{20\ \mu\text{g/ml}}{1000\ \mu\text{g/ml}}\right)$$

Thus, you need to dilute 4 μl of the stock solution to a final volume of 200 μl (*i.e*. by adding 196 μl).

If, in the example, you wished to make a solution of 1 μg/ml, the same equation would indicate that you need 0.2 μl of the 1000 μg/ml stock solution for 200 μl of the final diluted sample. This is a problem: 0.2 μl is very difficult to measure accurately. You have two choices: change the final volume (*i.e*. if V_2 is larger, then V_1 must also increase), or perform **serial dilutions** (*i.e.* instead of diluting the stock solution by a factor of 1000 in one step, dilute the stock solution, and then make a further dilution of the diluted stock).

In many cases, while the final *concentration* is important, the final *volume* is not (as in the previous paragraph). In these cases, do what was explained in this example: use a **convenient** dilution: a dilution that involves volumes that are easily pipetted. Pipetting 1.3333 μl is usually less accurate than pipetting 4 μl, both because 4 μl is a larger volume, and because it is difficult to set the pipet for 1.3333 μl. In this case, 4 μl is a convenient volume, while 1.3333 μl is not.

In some cases, you may not know the actual starting concentration. If, for example, you need to measure the enzyme activity in a sample, and you find that the activity is too high to measure accurately, you will need to dilute the starting material. Since you don't know the actual starting concentration, all you know is the concentration **ratio** between starting and final solutions. As long as you keep track of the concentration ratio in all of your dilutions, you can easily determine the enzyme activity in the initial solution, even though you cannot measure it directly.

Concentration ratios are frequently of considerable value. For example, you have a stock solution of buffer that contains 450 mM Tris-HCl, 10 mM EDTA, and 500 mM NaCl. You actually wish to use a final concentration of 45 mM Tris-HCl, 1 mM EDTA, and 50 mM NaCl. In each case the concentration of the final buffer is one-tenth that of the original. Simply performing a 1:10 dilution of the stock solution then gives the appropriate final concentration of each component. The stock solution of buffer is typically called a 10x stock, because it is ten-times more concentrated than the final, useful buffer.

Note, in the previous paragraph, the "1:10" dilution. The description uses the chemistry convention for this term, which will be used throughout this course. The 1:10 dilution mentioned is performed by taking one part of the initial solution, and adding nine parts of solvent (usually water). This results in a final concentration that is ten-fold lower than the original.

Experiment 1
The Art and Science of Protein Purification

When attempting to understand how a protein works, it is usually necessary to isolate the protein from other proteins that are present in the tissue. This allows you to study the protein with some assurance that the results obtained reflect the protein of interest and are not due to other molecules that were originally present in the tissue. Protein purification is therefore a commonly used biochemical technique.

Most proteins are fairly large molecules. They are smaller than DNA molecules, but they are tremendously large when compared to the molecules organic chemists are typically concerned with. The three-dimensional structure of most proteins is a consequence of many relatively weak non-covalent interactions. Disrupting this three-dimensional structure, on which the function of the protein depends, is therefore a relatively easy process. Conversely, preventing the loss of the non-covalent structure (and sometimes even the covalent structure) is frequently difficult.

Disrupting cellular structure is required to release the proteins from the cell. However, the process has two side effects that may damage proteins: 1) cell disruption typically involves shearing forces and heat, both of which can damage proteins, and 2) cells and extracellular fluids frequently contain proteases (enzymes that hydrolyze other proteins). In most cells, proteases are carefully controlled; however, disruption of the cell usually also releases the proteases from their control systems, and may allow the cleavage of the protein of interest.

Purification of proteins involves taking advantage of sometimes-subtle differences between the protein of interest and the remaining proteins present in the mixture. Because proteins are all polymers of the same twenty amino acids, the differences in properties tend to be fairly small.

In most cases, current understanding of protein structural properties is insufficient to allow a purification method to be generated theoretically. The “Art” in the title of this section reflects the fact that development of most protein purification procedures is a matter of trial and error. The table below lists some of the general properties of proteins that can be useful for protein purification, and some of the methods that take advantage of these properties. Each of these general methods will be discussed in some detail below. Note that for any given protein, only some of these methods will be useful, and therefore protein purification schemes vary widely.

Property	**Technique**
Solubility	Ammonium sulfate precipitation
Charge	Ion-exchange chromatography
Hydrophobicity	Hydrophobic interaction chromatography
Size	Gel-filtration chromatography
Function	Affinity chromatography
Stability	Heat-treatment, pH treatment

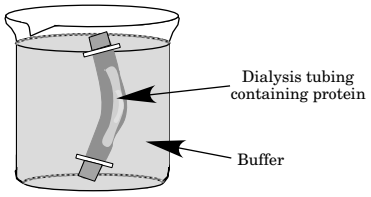
Ammonium sulfate precipitation
In many cases, cell lysates can be loaded directly onto chromatography columns. However, in some cases other molecules present in the lysate interfere with binding of the protein to the resin. In addition, some resins (especially affinity resins and sepharose-based resins) are fairly expensive; loading crude cell lysates on these columns may result in binding of cellular material (*e.g*. lipids and DNA) that are difficult to remove, and which may damage the column. As a result, purification methods often begin with one of several possible simple techniques that remove at least some of these unwanted materials prior to using an expensive column.

One of the most commonly used crude purification techniques involves the use of differential **solubility**. Proteins precipitate with increasing ammonium sulfate concentrations, with most proteins precipitating somewhere between 10% and 60% ammonium sulfate. (The percentages are relative to a saturated solution, which has a concentration of about 4 M; thus most proteins precipitate between 0.4 M and 2.4 M.) This can allow a simple, partial, purification of a protein; if the protein of interest precipitates at 40% ammonium sulfate, many other proteins will remain in solution, as will many other non-protein molecules.

Most proteins are not damaged by ammonium sulfate precipitation, and can be resuspended in a small volume of buffer. Ammonium sulfate precipitation results in a high salt concentration in the protein solution; this may be advantageous or disadvantageous depending on the subsequent step.

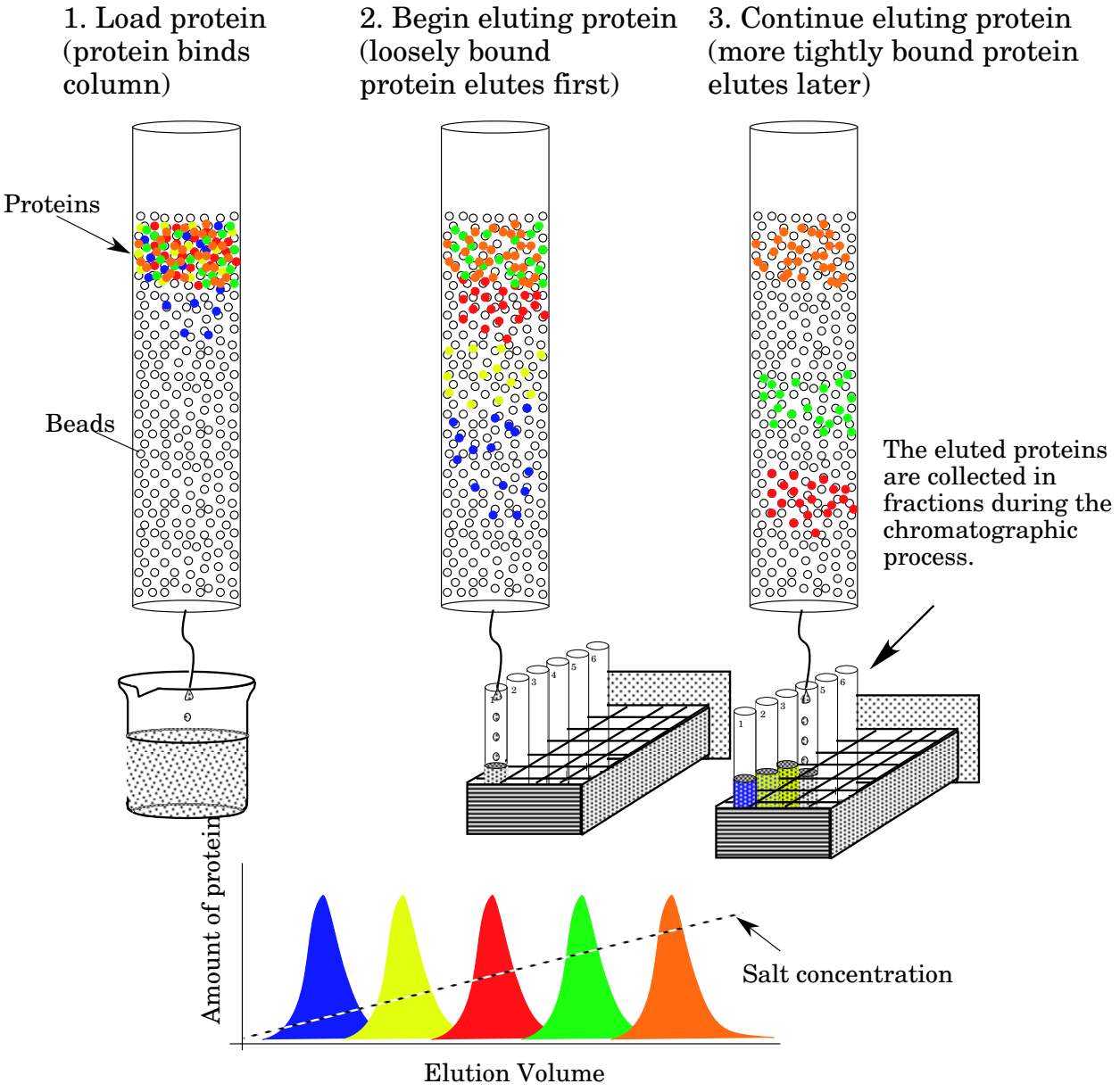
When necessary, two methods are frequently used to remove the salt. One method is gel filtration chromatography (discussed briefly below, and in more detail in Experiment 2B). Another frequently used method is dialysis.

Dialysis
Dialysis involves placing the protein solution in a semi-permeable membrane, and placing the membrane in a large container of buffer. Small molecules (such as salt ions) pass through the dialysis membrane (moving from high concentration to low concentration), while large molecules are unable to cross the membrane. Dialysis membranes come in a variety of pore sizes, and are therefore useful for removing a variety of different sized solutes. In principle, dialysis could allow separation of large proteins from small ones; in practice, however, the pores in the tubing are insufficiently uniform to allow this technique to be used effectively.



Chromatographic methods
Most purification methods involve chromatography. Chromatographic methods involve a column of an insoluble material that can bind molecules based on specific properties common to proteins. The solution containing the mixture of proteins is then allowed to pass through the column; the protein of interest may bind (depending on its properties), while at least some impurities remain in solution and

leave the column. The procedure is completed by eluting (*i.e.* "removing") the proteins that have bound to the column.



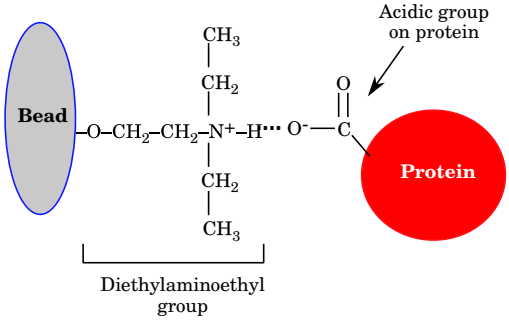
An illustration of a chromatographic run is shown above. The initial sample contains five different proteins (the differently colored filled circles). These proteins are bound to the column fairly tightly. Once elution begins, the proteins begin leaving the column. The graph at the bottom of the diagram shows proteins eluting with increasing salt concentration, in the manner that would occur with an ion exchange column; otherwise, this diagram applies to essentially any type of chromatographic method. Note: most columns do **not** run this neatly, especially in the beginning of a purification procedure.

Ion exchange chromatography: Proteins are charged molecules. Electrostatic

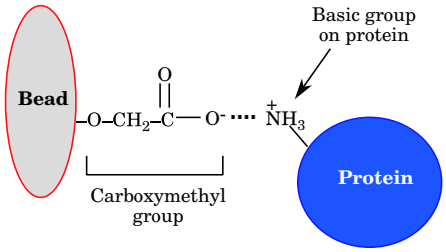
forces will therefore allow proteins to bind to other molecules of opposite charge. Ion exchange columns are produced by covalently attaching charged molecules such as diethyl-aminoethyl (DEAE) groups to insoluble carbohydrate resins. In many cases, small differences in charge can result in significant separations on ion exchange columns. Ion exchange columns are typically loaded at low ionic strength, and the protein removed by raising the ionic strength.

Types of ion exchange chromatography

Anion exchange chromatography

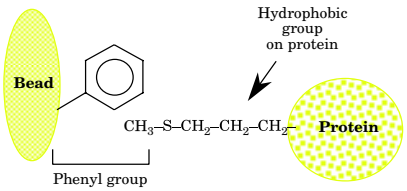


Cation exchange chromatography



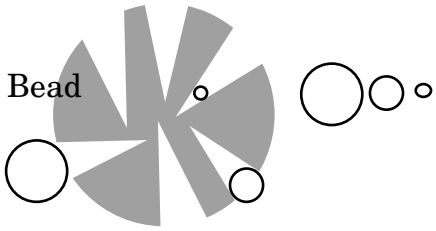
Hydrophobic interaction chromatography: Proteins contain hydrophobic amino acid side-chains, some of which are exposed at the surface of the protein. Proteins will therefore often bind to other hydrophobic molecules. Hydrophobic interaction columns are produced by covalently attaching hydrophobic molecules such as acyl chains or phenyl groups to insoluble carbohydrate resins.

Hydrophobic interaction chromatography



The hydrophobic effect is strongest under high ionic strength conditions; hydrophobic interaction columns are therefore typically loaded at high ionic strength, and the protein removed by lowering the ionic strength (thus, these columns are the opposite of ion exchange columns).

Gel filtration chromatography: In gel filtration chromatography (also known as size exclusion, gel permeation, or molecular sieve chromatography), molecules are separated based on size. Gel filtration columns are made of porous beads packed into a column. Different types of beads have somewhat different physical properties that may make them more appropriate for different proteins. As a solution containing molecules of varying sizes passes through the column, the molecules distribute between the inside and outside of the pores depending on their size. Molecules too big for the pores are totally excluded, and elute from the column first. Smaller molecules fit in the pores, and therefore elute later. The elution volume for a molecule is thus inversely related to the size of the molecule. Because



salts such as ammonium sulfate are much smaller than proteins, gel filtration is frequently used as a method for removing salt from a protein solution.

Affinity Chromatography: Many proteins exhibit specific interactions with other molecules (called ligands); for example, enzymes must have the ability to bind to their substrates, and antibodies exhibit high-affinity interactions with their antigens. In principle, it is possible to covalently attach the ligand to an insoluble resin. A column produced from such a resin is called an affinity column.

Affinity chromatography is somewhat less commonly used than the forms of chromatography discussed above. In many cases, the covalent attachment of the ligand to the column results in steric clashes that prevent the protein from binding. In some cases, although the protein will bind to the affinity resin, the resin is so expensive that other purification methods are used instead. In addition, bacteria find most affinity columns to be admirable growth media, with consequent deleterious effects on the chromatographic usefulness of the columns.

On the other hand, affinity chromatography can be extremely useful for rapidly purifying proteins. While most proteins contain the charged groups that can interact with ion exchange resins, the specificity of protein-ligand interactions means that only a very small fraction of the proteins in a cell will bind to any given affinity resin. Affinity chromatography can therefore be an extremely useful purification technique; in some cases, a single affinity chromatography step may be the only step necessary to completely purify a protein.

Protein Purification Strategies

Developing a scheme for purifying a protein remains an empirical process. However, in purifying a new protein, it is sometimes possible to adapt methods used for purifying similar proteins. In addition, planning the procedure before simply trying different methods can be extremely useful. Examples of this include using an ammonium sulfate precipitation step before a hydrophobic interaction chromatography step, because the high concentration of ammonium sulfate that results from the precipitation will allow the precipitated protein (or the non-precipitated protein remaining in solution) to be loaded directly onto the column. In contrast, an ammonium sulfate precipitated protein must be dialyzed (or otherwise desalted) prior to loading on an ion exchange column. Another frequently used scheme involves an inexpensive technique such as ammonium sulfate precipitation prior to remove bulk contaminants prior to running a higher resolution but more expensive technique such as affinity chromatography.

As with most scientific procedures, the more you know about the protein, and the more you know about protein purification, the more likely it is that you will be able to design a successful purification procedure.

Buffers

Proteins, and especially enzymes, are generally quite sensitive to changes in the concentrations of various solution components. A buffer is a solution that is used to

control the properties of a process occurring in an experimental aqueous medium. The term "buffer" is related to the ability of these solutions to resist changes in the hydrogen ion concentration, but buffers also contain other molecules, and are used to attempt to influence the ionic strength, the activity of proteases, and other parameters of the experiment in addition to the hydrogen ion concentration.

In any biochemistry experiment, the buffer components must be chosen based on their effect on the experiment. Ideal buffer components control pH and ionic strength without interacting in other ways with the system being studied. For example, while phosphate is a common physiological buffer, it may not be appropriate for some biochemical experiments, especially if phosphate is a substrate or product of the reaction being studied. In addition, some proteins interact poorly with some buffer components (a fact usually discovered by trial and error). As an example, Tris is less than ideal because of its high pK_a and the large change in pK_a that it exhibits upon changes in temperature. However, Tris is inexpensive, most proteins are stable in Tris buffers, and Tris rarely reacts with biological compounds; as a result, Tris is commonly used in biochemistry.

You have seen the Henderson-Hasselbalch equation in previous courses. This equation is useful for calculating the **theoretical** pH of a solution. It is also useful for predicting whether a particular compound will be useful as a buffer over a given pH range. However, the Henderson-Hasselbalch equation has its drawbacks. Many buffers used in biochemical experiments deviate significantly from ideal Henderson-Hasselbalch behavior.

$$pH = pK_a + \log \frac{[A^-]}{[HA]}$$ Henderson-Hasselbalch equation

Because of the commonly observed deviations from ideal behavior, buffers are typically prepared by adding the buffer components to a container, adjusting the solution to the desired pH by adding an acid or a base, and then adding sufficient water to reach the intended final volume.

For example, a 1 liter of 50 mM Tris-HCl buffer (pH 7.4) with 200 mM sodium chloride would be prepared by adding 50 millimoles of Tris base and 200 millimoles of sodium chloride to a flask contain about 900 ml of water. HCl would then be added to reduce the pH to 7.4, using a pH meter to monitor the changing pH, followed by addition of enough water to yield a 1 liter final volume. (If the solution contained 1 liter before addition of the HCl, the final volume would be more than 1 liter, and therefore the buffer would be less concentrated than it should be.)

Note that in order to produce most biochemically useful buffers, several components must be added together. This frequently requires careful consideration of the necessary dilutions for each of the components.

When performing experiments with proteins, it is rarely a good idea to dilute the protein with water, unless denaturation of the protein is not a concern. When

denaturation is a concern, such as when performing dilutions for enzyme assays, perform the dilution using a suitable buffer to prevent undesirable alterations in the structure of the protein in solution.

Study Questions #1A

1. You have 47 ml of a protein solution. You want to add 0.39 g of ammonium sulfate per ml. How much ammonium sulfate do you need to add?

2. How would you make up 100 ml of 50 mM Tris-HCl buffer at pH 7.4, using the 0.5 M Tris base stock solution, and any other required materials?

3. You have 5 ml of 3 mg/ml protein solution. (This solution is valuable so you don't want to waste it.) How would you make up at least 100 μl of 0.1 mg/ml using a **convenient** dilution (*i.e*. how much buffer and how much protein solution do you need to add, and is it possible to pipet the amount you need **accurately**)?

4. What do the letters PMSF stand for? What is PMSF used for?

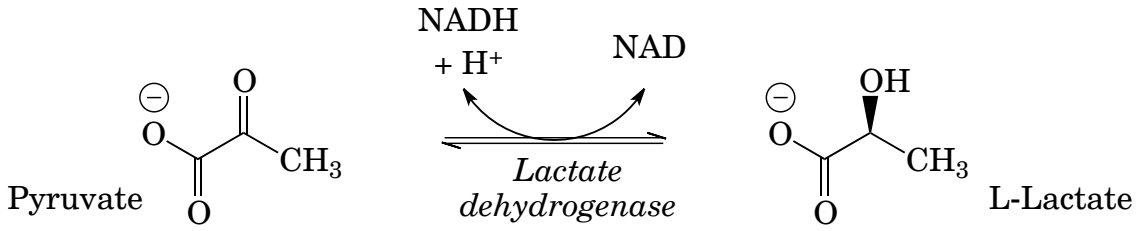
5. Define "supernatant" and "pellet".

6. What is the purpose of adding ammonium sulfate to the protein solution?

7. Why are you saving three 0.5 ml aliquots of your crude homogenate? (Note: you may need to read ahead to find the answer to this question.)

Experiment 1A
Purification of Lactate Dehydrogenase from Chicken, Part I

In this course, you will use lactate dehydrogenase (LDH) as the subject of your studies. LDH (*E.C.* 1.1.1.27) catalyzes the nicotinamide cofactor-dependent interconversion of lactate and pyruvate:



LDH is found in almost all organisms because it plays an important role in carbohydrate metabolism. During conditions in which pyruvate production from glycolysis exceeds the ability of the cell to oxidize the pyruvate, LDH converts the pyruvate to lactate, and thereby regenerates the oxidized NAD required for further glycolysis. LDH also allows the conversion of lactate to pyruvate; both the pyruvate and NADH produced can then be used for other processes.

In most animal tissues, LDH is produced from two genes, designated LDH A and LDH B. The LDH A gene is somewhat more highly expressed in muscle and liver, and its product is referred to as the M isozyme, while the LDH B gene is more highly expressed in heart, and its gene product is referred to as the H isozyme. In most species, the gene products form tetrameric complexes with properties that vary somewhat depending on the relative amounts of the different isozymes present in the tetramer.

LDH activity is readily measurable: the extinction coefficient for NADH at 340 nm (6220 $(M \bullet cm)^{-1}$) is much higher than that of NAD (the ε_{340} for NAD is zero). If the only substrates added to the reaction are NAD and lactate (or NADH and pyruvate), the change in absorbance at 340 nm should be proportional to the change in NADH concentration due to the LDH activity present in the cuvette.

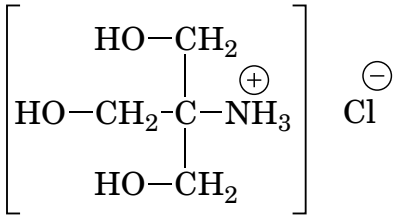
As mentioned in the introduction to this section, in any protein purification protocol it is necessary to take advantage of the way in which the protein of interest (in this case, LDH) differs from the other proteins in the mixture. Most tissues contain thousands of proteins; you need to use the properties of LDH to separate it from all of the other proteins present.

Most tissues contain proteases (enzymes that degrade other proteins). Avoiding proteolytic damage to your protein can be difficult. Three techniques are commonly used to keep proteolysis to a minimum: 1) perform the purification in the presence of protease inhibitors, 2) perform the purification at low temperatures (4°C or on ice), and 3) perform the purification in the minimal amount of time possible. Because you cannot do the last of these (the purification procedure will take more

than one lab period), you should keep your sample on ice or in the refrigerator as much as possible.

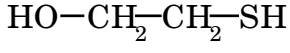
Definitions

Tris-(hydroxymethyl) aminomethane hydrochloride (Tris-HCl) is a commonly used buffer, and is intended to help control the pH of the solution. Tris is inexpensive, water-soluble, and generally inert in most biochemical experiments. These advantages somewhat compensate for its drawbacks. One drawback is the high pK_a value of 8.1 at 25 °C, a value that is well above the typically physiological pH value of 7.4 observed for most biochemical systems. Another drawback is the fact that the pK_a of Tris is far more temperature-dependent than that of most other compounds; the pK_a of Tris changes by –0.031 pH units per °C. (Tris can be used with counterions other than chloride; for most experiments, chloride, in the form of hydrochloric acid, is a convenient ion to add to the Tris base to obtain the desired pH.)



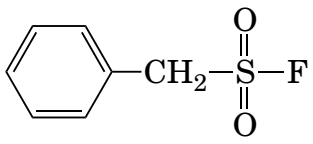
Tris hydrochloride

2-Mercaptoethanol (β-ME) is a reducing agent; it prevents formation of disulfide bonds between free cysteine residues. It also inhibits some proteases.



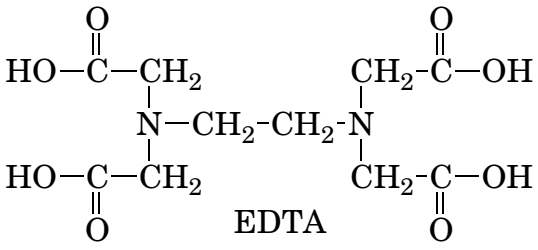
β-Mercaptoethanol

Phenylmethylsulfonyl fluoride (PMSF) is an irreversible inhibitor of serine proteases. PMSF is toxic, because it also inhibits serine esterases such as acetylcholinesterase; avoid getting PMSF on your skin. PMSF is relatively insoluble in water, and is normally dissolved in ethanol or isopropanol before being added to buffers. PMSF is also unstable in aqueous solution, and should be added to buffers immediately before use.



Phenylmethyl sulfonyl fluoride

Ethylenediamine tetraacetic acid (EDTA) is a chelating agent; it is used to remove metal ions from solution. Some proteases are dependent on metal ions (especially calcium ions), so EDTA acts as an inhibitor of some proteases.



EDTA

Reagents

Extraction Buffer:

10 mM Tris-HCl (pH 7.4)
1 mM 2-Mercaptoethanol
1 mM Phenylmethylsulfonyl fluoride (PMSF)
1 mM Ethylenediamine tetraacetic acid (EDTA)

Ammonium sulfate (solid)

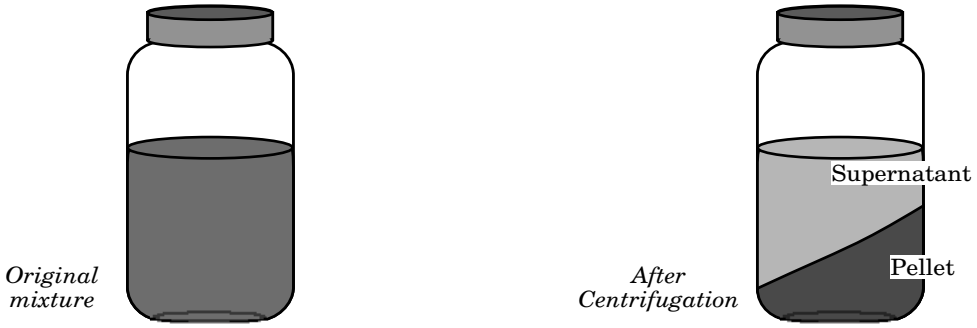
Chicken breast muscle
500 ml Centrifuge Bottles (one per group)
Blender
Microcentrifuge Tubes (1.5 ml)
Pipet Tips
50 ml Falcon Tubes
Pre-soaked Dialysis tubing
Dialysis tubing clamps
Ice buckets
Beakers for ammonium sulfate precipitation
Magnetic stirrers
Stir plates

Procedure

1. **Tissue preparation** – Cut ~50 g of muscle tissue from the tissue source (record the exact weight of tissue used). Cut the tissue into small pieces with scalpel or razor blades. Discard the connective tissue and fat.

2. **Soluble protein extraction** – Place the minced tissue and 75 ml of cold Extraction Buffer in a blender, and put the top on the blender. Disrupt the tissue by homogenizing. Use 4 x 30 second bursts, with at least 10 seconds in between each burst to allow the temperature of the homogenate to decrease.

3. **Centrifugation** – Put the homogenized tissue/buffer mixture into four pre-chilled 50 ml centrifuge tubes (note: the mixture will be the consistency of a thick milk shake, so a spatula will help). Balance the tubes (*i.e.* make sure that each pair of tubes have the same mass). Make sure that the tubes are not too full (you do not want to spill your mixture inside the rotor). **Centrifuges are dangerous and expensive. Consult your instructor before putting your samples into the centrifuge**! Centrifuge your homogenate for 20 minutes at 24,000 *x* g (12,000 RPM in the SLA-3000).



4. **Measure and record the volume of the supernatant**, **and save three 0.5 ml aliquots** (label the aliquots "Crude Homogenate").

5. **Ammonium sulfate precipitation** – *Slowly* (over a period of ~15 minutes) add 0.39 grams of ammonium sulfate per ml of supernatant to your supernatant. It is best to perform this step in the chromatography refrigerator (or on ice) using a magnetic stirrer (obviously, you need to put a stir bar into your sample). Avoid stirring too violently (proteins denature if subjected to shearing stresses; if you see bubbles forming, you are denaturing your proteins). Stir for an additional 15 minutes after you finish adding the ammonium sulfate (this gives the ammonium sulfate a chance to dissolve, and allows the proteins a chance to equilibrate to the presence of the ammonium sulfate).

(Ammonium sulfate precipitates proteins. Different proteins precipitate at different concentrations of ammonium sulfate. You are using "60% ammonium sulfate"; this means that the amount of ammonium sulfate you are adding is 60% of saturation – the maximum amount of ammonium sulfate that will go into solution. Most, although not all, proteins precipitate at 60% ammonium sulfate.)

During the addition of the ammonium sulfate, clean your centrifuge bottle, because you will need it again.

6. **Centrifugation** – Centrifuge the sample (10 minutes at 24,000 x g). Pour the supernatant into a separate container while keeping the pellet in the centrifuge tube. (The LDH should be in the pellet.)

7. **Dialysis** – Gently resuspend the pellet in 5 ml of Extraction buffer. Label a plastic clamp so that you will recognize your sample in the future, and use the clamp to close one end of a piece of dialysis tubing. Use a pipetman to transfer the protein solution into the tubing. Clamp the other end of the tubing, and place the sealed tubing in a large beaker of buffer containing a stir bar. Begin stirring gently, and leave the sample until the next lab period.

Study Questions #1B

1. You put a cuvette into the spectrophotometer and obtain an absorbance reading of 2.4. Is this a believable reading? How much of the light is actually reaching the detector?

Your instructor takes the cuvette, flicks it with a finger, and then replaces it in the spectrophotometer. The reading is now 0.314. What do you think happened?

2. Why does the Lab Manual recommend that you "gently" resuspend the ammonium sulfate pellet?

3. What would happen if you diluted your protein solution to 20 ml and loaded the entire solution on the Desalting column?

4. Why does LDH bind Cibacron blue?

5. Why do you wish to read A_{280} for your Cibacron Blue column fractions?

6. You have stock solutions of 1 M Tris base, 1 M HCl, 0.2 M 2-mercaptoethanol, 0.1 M PMSF, and 0.5 M EDTA. How would you make up 200 ml of extraction buffer?

> **Extraction buffer**:10 mM Tris-HCl (pH 7.4), 1 mM 2-Mercaptoethanol, 1 mM PMSF and 1 mM EDTA.

Experiment 1B
Purification of Lactate Dehydrogenase from Chicken, Part II

Reagents

Tris-PMSF Buffer
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol
1 mM PMSF

NAD Buffer
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol
1 mM lithium lactate
1 mM NAD

NADH buffer
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol
1 mM NADH

Cibacron Blue columns
Test Tubes and racks
Disposable pipets
Pipet tips

Procedure

High concentrations of salts (such as ammonium sulfate) often interfere with subsequent purification steps. Removing salt is therefore frequently necessary; you removed the ammonium sulfate by the slow method of dialysis.[2]

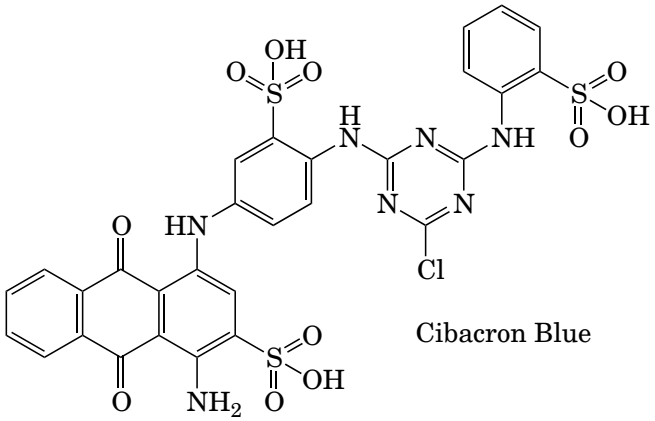
1. **Remove the protein solution from the dialysis tubing** – Unclamp one end, and carefully and gently pour or pipet the protein solution out of the tubing.

2. **Measure and record the volume of the dialyzed solution.**

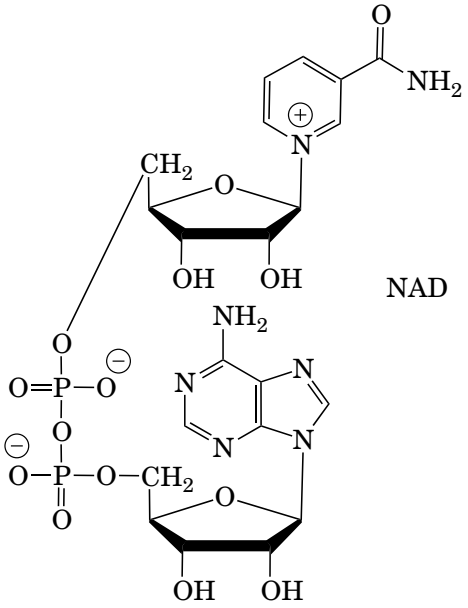
3. **Save three 0.1 ml aliquots** of the dialyzed solution (label them "desalted ammonium sulfate fraction").

[2] An alternative, and generally faster, method for removing salt is to use a small gel filtration column known as a desalting column. Gel filtration columns separate on the basis of size, with large molecules moving through the column more rapidly than small ones. (Note: gel filtration chromatography is discussed in more detail in the introduction to experiment 2B.)

The Cibacron Blue column is an affinity column. Nucleotide-binding proteins such as dehydrogenases frequently bind the Cibacron blue dye.[3] In this procedure the column should allow separation of LDH from the other proteins in the mixture.



Cibacron Blue



NAD

The procedure involves adding the mixture to the column, and then sequentially removing:
1) all proteins that do not bind to the column,
2) all proteins that bind loosely to the column, and finally,
3) LDH from the column.

Proteins, NAD, and NADH absorb at 280 nm. You can determine whether at least one of these materials is present in the eluent from the Cibacron blue column by measuring the absorbance of the fractions. Blank the spectrophotometer against air. Read the absorbance of a cuvette with Tris-PMSF buffer in it (the absorbance should be about 0.4). Read the absorbance of your Cibacron blue column fractions: *if the absorbance is less than ~0.1 above the absorbance of the Tris-PMSF buffer, go on to the next step*. After reading the absorbance, put the sample back in the same tube, cover the tube with parafilm, and store the tube on ice.

Sequentially label about twenty 13 x 100 mm test tubes. As you are running the column, you will be collecting fractions. ***Carefully* note which fractions correspond to each part of the procedure**.

4. **Load the Cibacron Blue Column** – After saving the aliquots, load the remaining desalted protein solution onto the Cibacron Blue column. Begin collecting 5 ml fractions.

[3] See, for example, R. K. Scopes, *Protein Purification: Principles and Practice*, 1982.

5. **Tris-PMSF Wash #1** – When all of the liquid has entered the column (*i.e.* has drained to the frit), fill the column with Tris-PMSF buffer, and continue collecting 5 ml fractions. Measure absorbance at 280 nm until it is less than ~0.1 above the absorbance of the Tris-PMSF buffer. (The absorbance is due primarily to proteins washing off the column.)

6. **NAD Wash** – After removing the remaining Tris-PMSF buffer with a pipet, add 10 ml NAD buffer to the column (note: make sure that you use NAD, and not NADH buffer at this step). Continue collecting 5 ml fractions.

7. **Tris-PMSF Wash #2** – When all of the liquid has entered the column (*i.e.* has drained to the frit), fill the column with Tris-PMSF buffer, and continue collecting 5 ml fractions. Measure absorbance at 280 nm until it is less than ~0.1 above the absorbance of the Tris-PMSF buffer. (The absorbance is primarily due to NAD washing off the column.)

8. **NADH Wash** – After removing the remaining Tris-PMSF buffer with a pipet, add 10 ml NADH buffer to the column. Continue collecting 5 ml fractions.

9. **Tris-PMSF Wash #3** – When all of the liquid has entered the column (*i.e.* has drained to the frit), add Tris-PMSF buffer to the column, and continue collecting 5 ml fractions. Measure absorbance at 280 nm until it is less than ~0.1 above the absorbance of the Tris-PMSF buffer.

10. **Cover** all of the fractions with parafilm (if you have not already done so), and store them in the refrigerator (at 4°C) until you need them again.

Study Questions #1C

1. What reagents (and how much of each) do you need to add to the cuvette when running the LDH enzyme assay?

2. You run an LDH assay on one of your samples, and obtain an activity of 1.5 ΔA/min. You then dilute the sample 1:10, and obtain an activity of 1.2 ΔA/min. Is this the activity that you would expect from a 1:10 dilution of your original solution? Why? What is happening in these assays?

3. Which samples do you need to assay for LDH activity?

4. Why do you need to set the spectrophotometer for 340 nm when running the LDH assay?

5. You obtain a rate of 0.8 ΔA/min from one of your samples (Sample A). You added 10 μl of a 1:10 dilution of Sample A to a total volume of 1.21 ml. What is the corresponding activity (in μmol/min/ml) in the **original sample**? (Hint: you may wish to consult the Calculation Hints page for assistance in calculating this value.)

6. Sample B gave 0.3 ΔA/min when using a 1:50 dilution in the same assay. What is the corresponding activity (in μmol/min/ml) in the **original sample**? Which sample (A or B) has the higher activity?

Experiment 1C
LDH Enzyme Assays

In order to determine whether you have actually purified LDH, you will need to measure the amount of LDH that you started with and the amount you ended up with, and compare these values to the amount of total protein in each of these same samples. You will measure LDH today, and protein concentrations in the next lab period.

How can you specifically measure the amount of LDH?

LDH is an enzyme. As with all enzymes, LDH catalyzes a specific reaction that would occur only *very* slowly without it; if you can measure the rate of that reaction you can obtain a measure of the amount of LDH present in your samples.

As was mentioned earlier in this Manual, LDH converts lactate to pyruvate and NAD to NADH. Although many enzymes convert NAD to NADH, only LDH does so while using lactate as substrate. You can measure NADH production spectrophotometrically; you can therefore supply NAD and lactate, and measure LDH activity by following NADH production.

In any enzyme assay, it is necessary to measure "initial rate". When substrate concentration is much higher than enzyme concentration, product formation is a linear function of time. In the case of LDH, you should see a linear increase in absorbance over time. Since you do not know the concentration of enzyme, you will have to establish initial rate conditions empirically. In other words, you will have to measure the rate; if the rate is too high to be linear (or too low to be accurately measured), you will need to change the amount of enzyme you are using.

The spectrophotometer will calculate rates in units of change in absorbance per unit time (ΔA/minute). You will need to convert these values to μmol/min. Two conversion factors are necessary: ε_{340} for NADH = 6220 $(M \bullet cm)^{-1}$ and M = mol/liter = μmol/μl

The above implies that you need to know some "volume"; which "volume" do you think is the appropriate volume for these calculations?

Reagents

Lactate Stock Solution
120 mM lithium lactate
10 mM Tris-HCl (pH 8.6)

NAD+ Stock Solution
12 mM NAD^+
10 mM Tris-HCl (pH 8.6)

Cuvettes
Ocean Optics

Bicarbonate Stock Solution
18 mM $NaHCO_3$
0.5 M NaCl

Tris buffer for Enzyme Dilution
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol

Procedure

Note: NAD and NADH are expensive, and are consequently in somewhat short supply. Please try to avoid wasting reagents; using excessive amounts may mean that you run out of one of the reagents before the end of the lab period.

You need to run assays to determine the LDH activity on the aliquots you took from the Crude homogenate and the Desalted ammonium sulfate fraction. In addition, you need to find your peak fraction(s) from the Cibacron Blue column. You probably have about 20 fractions; if the chromatography process worked properly, only two or three of the fractions will have significant activity. Because you cannot be certain which fraction(s) will have activity, you need to screen the samples for activity. In principle, you could test all of your fractions. A more efficient approach would be to test every third fraction (*e.g.,* fractions 1, 4, 7, 10, …). It is likely that any peaks that you have will extend over more than one fraction, and that you will find a fraction that has some activity. This fraction might be the peak; to find the actual peak, you will need to check the nearby fractions to find the one with the highest activity.

1. Mix 0.6 ml Lactate Stock Solution, 0.4 ml NAD Stock Solution, and 0.2 ml Bicarbonate Stock Solution in a cuvette.

2. Add 10 μl of the solution to be tested for LDH activity, and mix by inverting the cuvette. This initiates the reaction; you should therefore only add the "enzyme" when you are prepared to actually use the spectrophotometer.

3. Place the cuvette in the spectrophotometer and begin collecting data.

4. Have the spectrophotometer calculate the rate (in ΔA/min), and record the value. (Note: is the rate linear? Are the absorbance values under 2? If the answer to either of these questions is no, you will need to make a dilution of your LDH solution, and repeat the assay. Do not dilute (or discard) all of your Cibacron Blue Peak fraction(s); you will need the concentrated solution(s) for later experiments.

5. Calculate the LDH activity in μmol/min in the cuvette, and μmol/min/ml in the original (*i.e*. undiluted) enzyme solution.

Study Questions #1D

1. Does the Bradford Assay allow you to calculate the LDH concentration? Why?

2. What wavelength do you use for the Bradford assay? Why is this different from the wavelength you used for the LDH assay?

3. You have 0.5 ml of a stock solution of 1 mg/ml BSA. You need at least 200 μl of the solutions listed below. How do you make them up?
25 μg/ml
50 μg/ml
75 μg/ml
100 μg/ml
150 μg/ml
200 μg/ml

4. You set up a Bradford assay. Your protein sample gives an absorbance of about 2.5. What do you need to do?

5. You perform a Bradford assay. You obtain the absorbance values listed below from the BSA samples; your protein sample yields an absorbance of 1.3; what is the protein concentration of your sample?

BSA (μg/ml)	Absorbance
0	0.5
25	0.65
50	0.8
75	0.95
100	1.1
150	1.4
200	1.75

Experiment 1D
Bradford Protein Assay

In biochemistry, it is frequently necessary to know the total protein concentration in your solution. A number of methods have been developed for measuring total protein concentration. The method you will use is called the Bradford Assay, after the name of the scientist who developed the method. The method was first described in a paper written more than 30 years ago[4].

The Bradford reagent binds proteins; when it does so, its extinction coefficient at 595 nm increases. Because the magnitude of the extinction coefficient change is somewhat dependent on the conditions, it is necessary to calibrate the change in absorbance induced by different amounts of protein. This calibration procedure is called "generating a standard curve".

Ideally, the absorbance readings obtained in your standard curve should be linearly proportional to the protein concentration (which means that there are limits as to how much protein you can use; too much protein will give absorbance readings too high to be meaningful). Time constraints mean that you will probably be able to run each sample only once. In an actual experiment, you would run replicates for both the standard curve and the unknown samples, so as to be able to assess the uncertainty in your data.

Supplies

Bradford Reagent
1 mg/ml Bovine serum albumin (BSA)

Cuvettes

Procedure

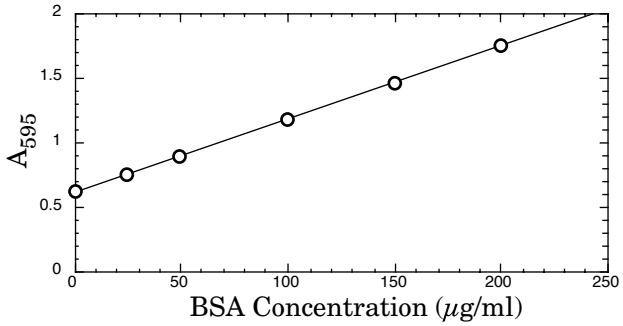
Standard Curve: For this experiment, you will be using a stock solution of 1 mg/ml bovine serum albumin (BSA) as your standard protein. Note that 1 mg/ml = 1000 μg/ml. You will want to make up several dilutions of the BSA stock, and then use a constant volume from each protein dilution in the actual assay tubes. **I suggest making 25 μg/ml, 50 μg/ml, 75 μg/ml, 100 μg/ml, 150 μg/ml, and 200 μg/ml solutions of BSA.** Since you need 100 μl of the protein solution, it is probably best to make convenient dilutions with final volumes of 150 to 400 μl. **Decide how to make up these dilutions *before* you come to class.**

Mix 800 μl of H_2O, 200 μl of the Bradford Coomassie Blue solution and 100 μl of the BSA standard dilution together, allow them to equilibrate for a few minutes, and measure the absorbance at 595 nm. In addition to BSA protein solutions, run one

[4]Bradford, M.M. (1976) *Anal. Biochem*. **72**, 248-254.

sample without protein (*i.e.* which contains the same assay mixture, using buffer instead of protein).

Your plot of the absorbance at 595 nm *versus* protein concentration should look similar to the one shown at right (your data may not fit the line as closely as these do, although they might if you are careful with your dilutions).



You will also need to dilute your LDH purification protein solutions. You will prefer to use small volumes of your precious protein solutions (you worked hard to prepare those protein solutions, and you will need them again for later experiments). You will find that most, if not all of the samples you have will be too concentrated to run undiluted in the assay. If you dilute the initial sample 1:10 (*i.e.* 30 μl sample + 270 μl of H_2O), and then prepare several 1:3 serial dilutions from this sample, eventually you will obtain samples that fall within the standard curve. (Reminder: **serial dilutions** are performed by diluting a sample, then taking an aliquot from the diluted sample and diluting it further, and then taking an aliquot from the second dilution and diluting it further, and . . . .) You will need to make sure that at least one of your diluted protein samples has an absorbance that falls within the standard curve.

Once you have found a useful dilution for each of your important samples, repeat the dilution, and assay the sample again. This is called "performing a replicate"; the replicates will allow you to determine the protein concentration in your samples with more confidence.

You **need** to run assays to determine the protein concentration on the following samples:

Crude homogenate
Desalted ammonium sulfate fraction
Peak fraction(s) from the Cibacron Blue column

Once you have measured the absorbance values for the standard curve and for your unknown protein samples, you will need to use the standard curve to determine the protein concentrations for your unknowns. The best method is to plot the values obtained for the standard curve, determine the slope of the best-fit line, and then use the equation of the line to give you the protein concentrations for your unknown samples. (Don't forget to correct the value obtained in this method for the dilution you performed – your goal is to calculate the protein concentration in the initial sample.)

Calculation Hints: Setting Up a Purification Table

One important feature in scientific papers describing protein purification is a table that describes the success of the purification procedure. This table usually has two goals: 1) to measure the removal of contaminants during the purification, and 2) to assess the efficiency of each step, and the overall efficiency of the entire purification process.

The typical purification table for an enzyme gives the values listed for each step during the purification. In order to set up a purification table, you will need to know the enzyme activity (corrected to μmol/min of product formed per ml of enzyme solution). For your LDH assay, this can be calculated by:

$$Activity\left(\text{in } \frac{\mu mol}{\min} / ml\right) = \left[\frac{Activity\left(\text{in } \frac{\Delta A}{\min}\right)}{6220\left(M^{-1} \bullet cm^{-1}\right)}\right]\left[10^6 \frac{\mu mol}{mol}\right]\left[\text{cuvette volume}(L)\right]\left[\frac{1000 \mu l / ml}{10 \mu l \text{ used}}\right]\left[\text{enzyme dilution}\right]$$

In the equation above, "6220 $M^{-1} \bullet cm^{-1}$" is the extinction coefficient of NADH. The cuvette volume (which must be in liters) is the total volume of solution in the cuvette, and is used to correct from molarity to moles. You will be using 10 μl of the enzyme, but you wish to correct to per ml of enzyme, and therefore must multiply by 100. If you diluted your enzyme preparation, you need to multiply by the dilution factor. You will be performing this calculation repeatedly during the course; you should make certain that you understand the purpose of each term within the equation.

Total Enzyme Activity: the activity per ml of enzyme solution multiplied by the total volume of that fraction. (If you have 50 ml of the original homogenate, the volume is 50 ml, even if you only saved 0.5 ml for the assay.)

Total Protein: the protein concentration per ml of solution multiplied by the total volume of that fraction.

Specific activity: Total Activity/Total Protein **or** Activity/Protein concentration

Fold-purification: (Specific activity at a given step)/(Specific activity of starting sample)

Yield: (Total activity at a given step)/(Total activity of starting sample)*100 (Note: this is intended to track the amount of LDH present in your sample, and **not** the total amount of all of the proteins in the fraction.

Which purification steps do you think should be included in your table? Why do you think that each of the values listed above is important? (What are the values telling you about your purification procedure?)

Your table should include, at minimum, the columns listed below.

Step	Activity (μmol/min/ml)	Protein conc. (mg/ml)	Total volume (ml)	Total protein (mg)	Total Activity (μmol/min)	Specific activity (μmol/min/mg)	Fold Purifi-cation	Yield (%)

Study Questions #2A

1. Why is SDS added to the sample and the gel?

2. What are the two layers of a gel used in SDS-PAGE? What is the function of each layer?

3. What information can you obtain from SDS-PAGE?

4. Why are you running the SDS-PAGE today?

5. In SDS-PAGE, do the proteins migrate toward the anode (positive electrode) or cathode (negative electrode)? Why?

6. Calculate the preparation of all samples to be run on your gel. Prepare a table similar to the one below (**make a copy for yourself**). You **must** complete this table before coming to class.

Sample	**Protein conc. (mg/ml)**	**Are you using a dilution? If so, what dilution factor are you using?**	**Amount (µl) of protein sample (or diluted protein)**	**Amount (µl) of water**	**Amount (µl) of 5x sample buffer**

Characterization of LDH
Experiment 2A
LDH Analysis: SDS PAGE

Once a protein purification procedure has been completed, the resulting protein must be characterized for both purity and physical properties. You will use several techniques to analyze the LDH purified in Experiment 1.

The first technique is SDS polyacrylamide gel electrophoresis. SDS PAGE allows assessment of purity of the preparation, estimation of approximate quantity of the protein, and measurement of the size of the protein.

Electrophoresis is a process in which molecules are exposed to an electric field and separated on the basis of their differential mobilities in that field. The differential mobility is a result of different charge on different molecules, and the result of different resistance to movement through the medium. For molecules with similar shapes, the mobility is proportional to the charge-to-mass ratio of the molecule. For molecules of similar shapes and similar charge-to-mass ratios, the motion through the medium will be proportional to the size of the molecule, because friction increases as a function of size.

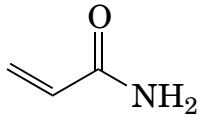
The velocity of a charged molecule is given by:

$$v = \frac{qE}{f}$$

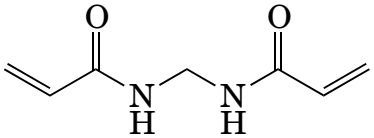
where q is the charge on the molecule, E is the electrical potential gradient, and f is the frictional coefficient of the medium for the molecule. For similarly shaped molecules, f is proportional to size of the molecule. For molecules in which charge increases in proportion to size, larger molecules move more slowly than small ones, because f increases faster than charge.

Gel electrophoresis uses a matrix of large uncharged molecules to provide the required friction. The matrix also serves to inhibit diffusion, and therefore to prevent degradation of the separation that is achieved. The separation of proteins usually involves the use of polyacrylamide as the matrix.

Polyacrylamide is formed by polymerization of acrylamide monomers in the presence of *N,N´*-methylenebis(acrylamide).



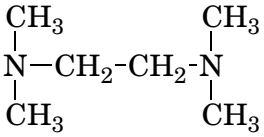
Acrylamide



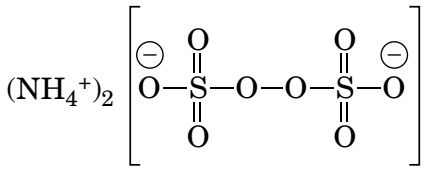
N,N´-Methylenebis(acrylamide)

The bis(acrylamide) contains two double bonds, which allow the compound to act as a cross-linker between polyacrylamide chains. The presence of the cross-linking agent results in formation of a gel matrix rather than a simple linear polymer.

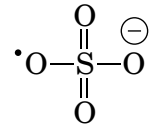
In the synthesis of polyacrylamide, the polymerization occurs via a free radical chain reaction mechanism. The formation of the free radicals is initiated by the unstable compound ammonium persulfate. The sulfate radicals formed then react with *N,N,N´,N´*-tetramethylethylenediamine (TEMED), forming TEMED radicals, which then react with acrylamide molecules to begin the actual polymerization reaction.



TEMED



Ammonium persulfate

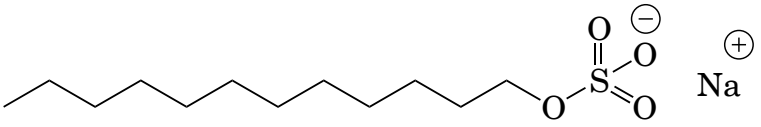


Sulfate radical

Varying the amount of acrylamide monomer and bis-acrylamide cross-linker present controls the formation of the matrix. The use of larger amounts of these components results in a denser matrix. Denser matrices are used for separating smaller proteins; larger proteins may find the pores in a dense matrix too small to enter, and may therefore not enter the gel at all. The table below lists approximate useful ranges for different gel densities.

Percent acrylamide	Useful Molecular weight range
7	30,000 to 200,000
10	20,000 to 150,000
12	10,000 to 100,000
15	5,000 to 70,000

Real proteins have different proportions of charged side-chains. As a result, real proteins do not have constant charge-to-mass ratios. Real proteins also have varying three-dimensional shapes. In order to measure molecular weight, it is necessary to induce the formation of a similar shape and charge-to-mass ratio. Boiling the protein in the presence of the detergent sodium dodecyl sulfate (SDS) and the reducing agent β-mercaptoethanol (which reduces disulfide bonds) results in disruption of the three dimensional structure of the protein. In addition, large amounts of SDS bind to the protein (approximately one molecule of SDS for every two amino acid residues). Given the fact that each SDS molecule has a negative charge at the pH used for electrophoresis, the use of SDS results in a large negative charge that overwhelms any intrinsic charge present in the protein.

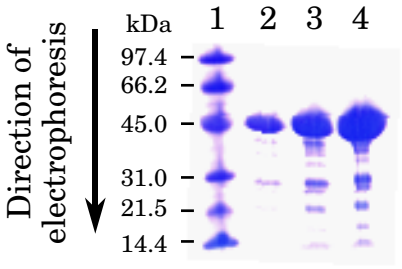


Sodium dodecylsulfate (SDS)

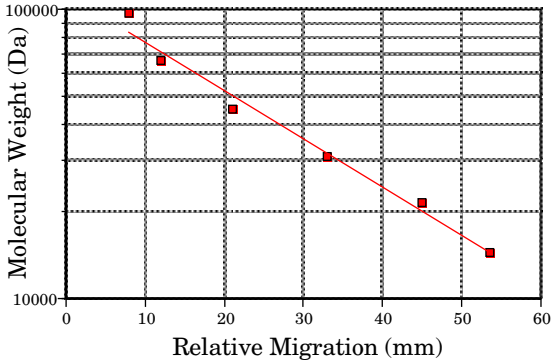
Note that treatment with SDS and β-mercaptoethanol will result in the formation of denatured protein monomers; it is these protein monomers that are separated on the SDS PAGE.

After the electrophoresis has been performed, the protein must be detected. The most commonly used method for detecting protein is Coomassie Blue R-250. Coomassie blue is a dye that binds proteins. Staining is performed by placing the gel in a solution of Coomassie blue in acetic acid and methanol. The function of the acetic acid and methanol is to cross-link the proteins into the gel so that they do not diffuse. Following staining of the proteins, the gel is placed in a solution of acetic acid and methanol, which results in removal of the excess Coomassie blue.

An example of an SDS gel is shown at right. Lane 1 in the example is comprised of proteins of known molecular weight called **molecular weight standards**. Lanes 2, 3, and 4 contain increasing concentrations of an experimental protein sample. Loading increasing amounts of protein makes it possible to see minor impurities, which are difficult to see in lane 2 and fairly obvious in lane 4.



The molecular weight standards can be used to calibrate the migration of proteins of differing sizes on the gel. For any given gel, the migration will be inversely proportional to the log of the molecular weight. Thus, a plot of **log molecular weight *versus* migration distance** for proteins of known size can act as a standard curve to allow measurement of molecular weights of unknown proteins. An example of an SDS PAGE standard curve is shown at right. Note that proteins within an SDS polyacrylamide gel are denatured; the molecular weight determined will be that of the **individual monomers** of multimeric proteins.

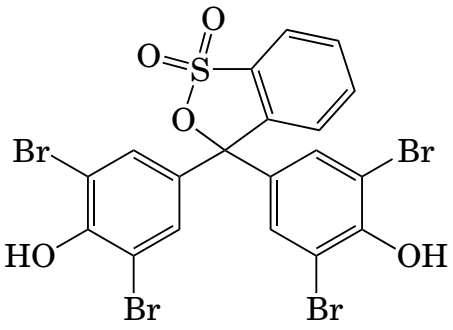


A commonly used set of molecular weight standards for SDS PAGE experiments is shown in the table below.

Standard protein	Molecular weight (kDa)
Rabbit muscle phosphorylase b	97.4
Bovine serum albumin	66.2
Chicken ovalbumin	45.0
Bovine carbonic anhydrase	31.0
Soybean trypsin inhibitor	21.5
Chicken lysozyme	14.4

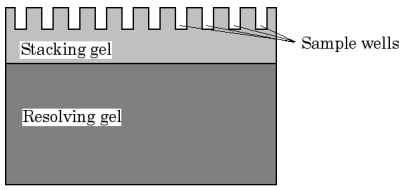
As mentioned above, when running an SDS PAGE, you must denature your proteins and coat them with SDS. This is usually done by boiling the protein sample in SDS PAGE sample buffer. In this experiment, you will be using a 5x Sample

buffer, which contains the relevant reagents at five-times the required final concentration. The sample buffer has three roles. The sample buffer provides the SDS necessary for the uniform charge-to-mass ratio. The sample buffer also provides a tracking dye (usually bromophenol blue, a blue dye that runs at a low apparent molecular weight). The tracking dye makes it easier to see the sample during loading, and lets you know when the lowest molecular weight proteins are getting near the end of the gel. Finally, the sample buffer contains glycerol to make the protein sample denser than the electrophoresis tank buffer, so that the protein sample will sink to the bottom of the well when you load it.



Bromophenol blue
(3´,3´´,5´,5´´-tetrabromo
phenolsulfonephthalein)

The type of SDS polyacrylamide gel you will use consists of two layers: the top layer is the **stacking gel** and the bottom layer is the **resolving gel**, based on a method invented by Laemmli in 1970.[5] The purpose of the two sections differs: the stacking gel concentrates all of the protein in a narrow region, while the resolving gel performs the actual separation of the proteins by molecular weight. The stacking gel also contains the wells into which the samples are loaded.



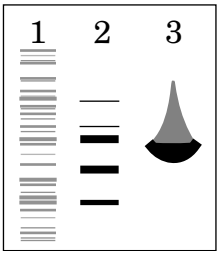
The stacking gel is prepared at a lower pH (6.8) and lower acrylamide percentage (6%). At this low pH the glycine of the electrophoresis tank buffer is in the neutral zwitterionic form and is not an effective carrier of current. The chloride ions also present are highly charged and migrate rapidly toward the anode. The SDS-coated protein molecules and the dye, both of which have charge-to-mass ratios greater than that of glycine but less than that of Cl^-, must migrate behind the Cl^- and ahead of the glycine. This has the effect of concentrating the proteins in a thin band sandwiched between the Cl^- ions and the glycine molecules. In addition, because the acrylamide concentration of the stacking gel is very low most proteins are not retarded and move freely through the gel matrix. When the sample reaches the end of the stacking gel it should appear as a thin blue band.

The resolving gel is prepared at pH 8.8 and has the desired acrylamide concentration for separation of proteins in the appropriate size range (12% for this experiment). When the stacked samples enter the resolving gel the higher pH results in negatively charged glycine molecules that then migrate with the Cl^- ions. The protein samples lag behind and are separated by the sieving effect of the gel. The bromophenol blue tracking dye will migrate faster than the proteins. When the

[5] Laemmli (1970) *Nature* **227,** 680.

tracking dye reaches the end of the gel, the electrophoresis should be terminated so that the proteins do not run off the gel.

Why are you running more than one concentration of protein?
Gels have limits in terms of the amount of protein that can be loaded. Loading too much protein results in "overloading" in which the sample runs unevenly. Loading too little protein makes the protein bands difficult to detect. In choosing the amount of protein to load, you need to consider the number of proteins in your sample; if your sample contains many proteins, you can load more total protein.



In the gel drawn at right, the same mass of total protein was loaded in each lane. In the first lane, there were many proteins, and therefore none resulted in an intense band. You could have loaded more protein for this sample. In the second lane, there are five proteins visible; the amount loaded was appropriate for this sample. The third lane is overloaded; there is too much of the one protein present, and it did not run cleanly. In fact, the lower portion of this band probably ran at a molecular weight smaller than the actual molecular weight of the protein, and it may have distorted the protein migration in adjacent lanes.

The gel run in this experiment may be used to estimate an appropriate amount of protein to load on subsequent gels.

Reagents
Chemicals required:

12% Resolving gel 1.5 ml 1.5 M Tris-HCl, pH 8.8 1.8 ml **40% acrylamide** 0.06 ml 10% SDS 2.94 ml water	**12% Resolving gel** 1.5 ml 1.5 M Tris-HCl, pH 8.8 2.4 ml **30% acrylamide** 0.06 ml 10% SDS 2.04 ml water

Add **last** to initiate reaction:
30 μl 10 % ammonium persulfate and 5 μl TEMED
Pour the gel and layer 1-butanol on top of the polymerizing solution. After polymerization is complete, rinse off the top of the gel to remove the butanol.

Stacking gel 0.875 ml 1.0 M Tris-HCl, pH 6.8 0.437 ml **40% acrylamide** 0.035 ml 10% SDS 2.153 ml water	**Stacking gel** 0.875 ml 1.0 M Tris-HCl, pH 6.8 0.583 ml **30% acrylamide** 0.035 ml 10% SDS 2.007 ml water

Add **last** to initiate reaction:
30 μl 10 % ammonium persulfate and 5 μl TEMED

Pour the gel and insert the comb to create the wells.

5x Sample Buffer

60 mM Tris-HCl, pH 6.8
25% glycerol
2% SDS
14.4 mM β-mercaptoethanol
1% bromophenol blue

Electrophoresis tank buffer

25 mM Tris
192 mM glycine, pH 8.8
0.1% SDS

Coomassie Staining Solution (10% Acetic acid, 25% Methanol, 0.05% Coomassie R-250 or Bio-Safe Coomassie staining solution)

Hardware required:

Electrophoresis apparatus
Gel loading tips
Weigh boats for gel staining
100°C water bath
Microcentrifuge tubes

Procedure

1. Set up the gel apparatus.
   Mark location of wells before adding tank buffer
   Remove bubbles from bottom of gel

2. Prepare your samples.
   You will want to run all protein samples that contain LDH activity. This should include:

   Crude homogenate
   Ammonium Sulfate Pellet/Desalting Column Elution
   Peak Cibacron Blue Elution Fraction(s)

The gel apparatuses you will be using have 10 wells, so you can plan on running 3 lanes for each of your samples. Each sample should contain 20 μl of diluted protein and 5 μl 5x sample buffer. **Loading the same volume of each sample tends to allow the gel to run more evenly**. Prepare **2 μg, 10 μg, and 50 μg aliquots** for the crude homogenate and desalted ammonium sulfate samples; prepare **2 μg and 10 μg aliquots** for the peak fraction(s). (Note: do not prepare the 50 μg sample for your peak fraction(s); this would result in an overloaded lane.) To calculate how much of each sample to use you must refer back to the protein concentration

determined by the Bradford assay in Experiment 1D. If your protein concentration is too low to allow preparation of the more concentrated aliquots, make the most concentrated aliquots you can.

Example: Sample preparation:

Sample #1: The results of the Bradford protein determination indicate a concentration of 9.6 mg/ml. Note that this can also be stated as 9.6 μg/μl.

For the gel, the "2 μg sample" should have a final volume of 25 μl; however, you wish to load 20 μl of the sample, and have this 20 μl contain 2 μg. This means that the concentration of the "2 μg sample should be 2 μg/20 μl, or 0.1 μg/μl. How can you make a 0.1 μg/μl sample from a 9.6 μg/μl stock solution?

A quick calculation indicates that you need to dilute your stock 96-fold. This is a large dilution, and will require pipetting less than 1 μl of your stock solution. Because it is impossible to accurately pipet such small volumes, you should make up a more convenient concentration, such as 1 μg/μl. This is a 9.6-fold dilution of the sample.

9.6-fold dilution preparation: 5 μl protein sample + 43 μl water or buffer (This can be calculated readily. You want a 9.6-fold dilution. Pick a convenient starting volume, such as 5 μl, and multiply by the dilution factor (5 μl x 9.6 = 48 μl) Thus, 48 μl is the desired total final volume. The amount of buffer is then 48 μl – 5 μl = 43 μl. The new concentration is (9.6 μg/μl)/9.6 = 1 μg/μl

You should understand how to do dilutions like this one and how to calculate the protein concentration in the diluted sample!

You can then use the dilution of your stock to prepare the 2 μg sample: 25 μl of 0.1 μg/μl requires 2.5 μg of protein, which is 2.5 μl of a 1 μg/μl solution. To complete the sample, you need to add 17.5 μl of water and 5 μl of 5x sample buffer.

For the "10 μg sample", you could use the dilution prepared above. 25 μl of 0.5 μg/μl requires 12.5 μl protein sample; you then need to add 7.5 μl water and 5 μl of 5x sample buffer. For the "50 μg sample", you can use your original protein stock. 25 μl of 2.5 μg/μl requires 62.5 μg of protein; 62.5 μg)/(9.6 μg/μl) = 6.5 μl of the protein sample.

3. Add 5x sample buffer to each sample. The sample buffer should be added at 1/5 the final volume. Your samples contain 20 μl so you should add **5 μl sample buffer** (5μl is 1/5 of 25 μl). If your samples do not contain 20 μl of diluted protein, you must determine how much sample buffer to add.

Calculate how you are going to prepare all of your samples *before* coming to class!

4. Heat all samples at 100°C for 2 minutes. Spin down the protein solution for 5 seconds, so that all of the liquid is at the bottom of the microfuge tube.

5. Apply samples to wells. Record how the samples were loaded!

6. Attach the electrodes to gel apparatus, and turn on the power supply. Run the gel at constant current (15 mA per gel) until the tracking dye reaches the end of

the gel. **Be careful – the high current during gel electrophoresis is dangerous**!

7. Turn off the power supply and remove the electrodes. Remove the gel sandwich. Carefully separate the plates. Gently remove the gel from the plate and cut one lower corner at lane #1 to distinguish the two sides. Place it in a plastic tray containing stain solution.

Do not handle the gel with your bare hands: **acrylamide is a cumulative neurotoxic agent and is absorbed through the skin!** Residual unpolymerized acrylamide will be present in the gel.

The gel will be stained overnight and transferred to the destaining solution before the next lab period. The results should be visible in the next lab period.

Study Questions #2B

1. What information can be obtained from an analytical gel filtration chromatography experiment? How does this differ from the information obtained from an SDS PAGE experiment? What similar information can be obtained from SDS PAGE and gel filtration?

2. The following protein standards were run on a gel filtration column and their elution times determined. Generate a standard curve from these data.

Protein	**Molecular Wt.**	**Elution time**
Thyroglobulin	670 kDa	8.45 min
Bovine γ-globulin	158 kDa	11.15 min
Chicken ovalbumin	44 kDa	12.7 min
Equine myoglobin	17 kDa	14.25 min
Vitamin B_{12}	1.35 kDa	18.25 min

3. On the same column your sample elutes at 12 minutes. What is the molecular weight of your protein?

4. Why do you want to set the absorbance detector for 280 nm?

5. When pouring a gravity column, you notice that all of the beads have fully settled. What would probably happen if you add more beads? How would this affect any separation performed using the column?

6. What was the apparent molecular weight of LDH from the HPLC gel filtration column? (Please include any calculations and standard curves used to determine this value.)

7. Did you observe any peaks in the gel filtration chromatogram other than LDH? If so, what were their molecular weights? What might this tell you about their identities?

8. What did your gel filtration experiment tell you about possible contaminants in your LDH preparation? Is this information similar to that from your SDS PAGE? Why?

9. What do you need to be careful to avoid when loading a gravity gel filtration column?

Experiment 2B
Gel Filtration Chromatography

Most protein purification procedures, and some protein analytical methods, involve chromatography. Chromatographic methods involve a column of an insoluble material that can bind molecules based on specific properties common to proteins. The solution containing the mixture of proteins is then allowed to pass through the column; the protein of interest may bind (depending on its properties), while at least some impurities remain in solution and leave the column. Eluting (*i.e.* "removing") the proteins that have bound to the column completes the procedure. Separations are achieved as a result of the partitioning of the proteins and other solutes between the solid support material and the mobile phase.

In gel filtration chromatography (also known as size exclusion, gel permeation, or molecular sieve chromatography), molecules are separated based on size. As with SDS PAGE, a gel filtration chromatography column can be calibrated by using molecules of known size; the gel filtration column can then be used to determine the size of an unknown protein. In contrast to SDS PAGE, which measures the molecular weight of denatured protein, gel filtration measures the size of the native protein.

Gel filtration columns are made of porous beads packed into a column. The beads can be polymers of dextran (Sephadex), agarose (Sepharose and Superose), agarose cross-linked to dextran (Superdex), polyacrylamide (Sephacryl), or related compounds. Different types of beads have somewhat different physical properties that may make them more appropriate for different proteins.

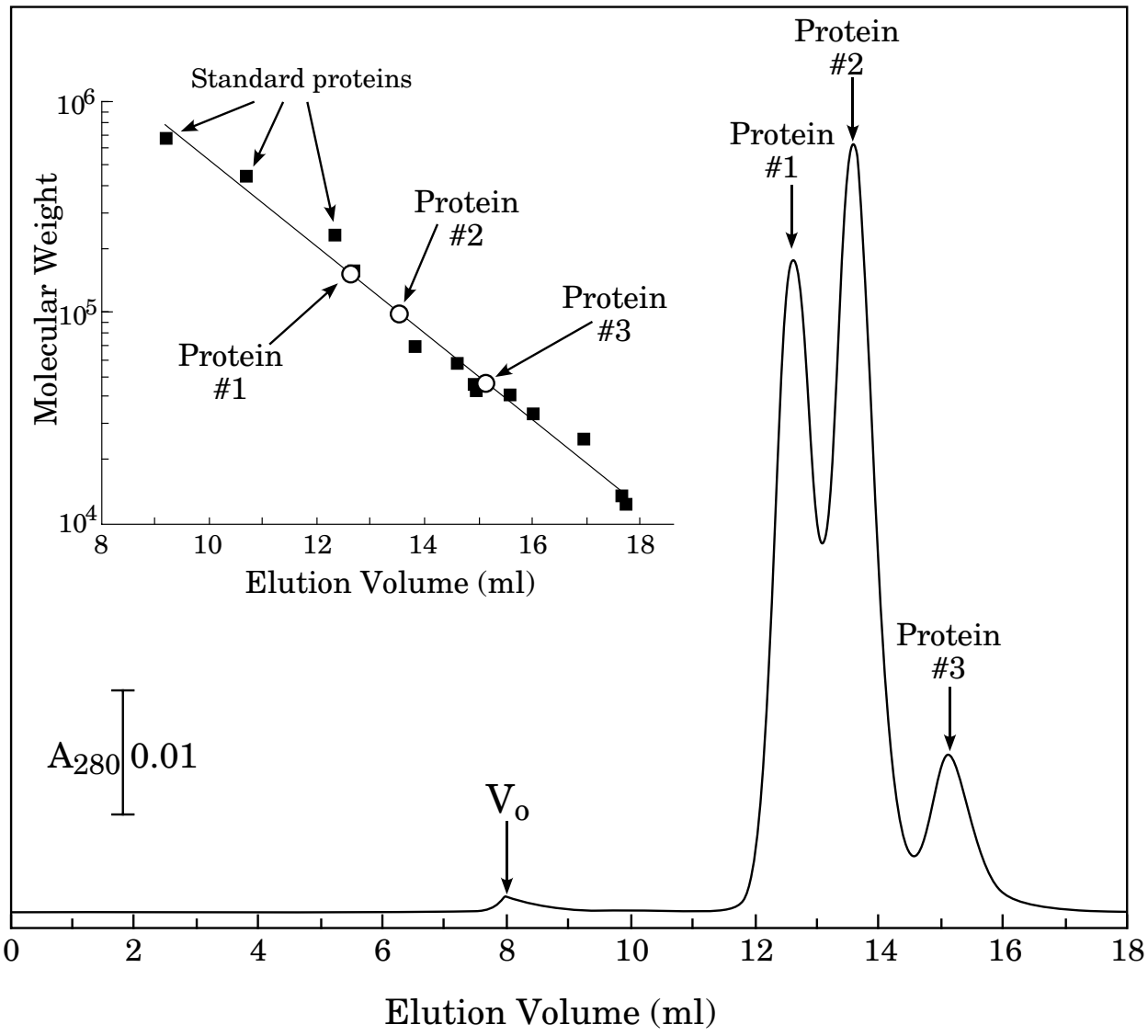
As a solution containing molecules of varying sizes passes through the column, the molecules distribute between the inside and outside of the pores depending on their size. Molecules too big for the pores are totally excluded, and elute from the column first. Smaller molecules fit in the pores, and therefore elute later. The elution volume for a molecule is thus inversely related to the size of the molecule.

Gel filtration beads are characterized by their composition. They are also characterized by their exclusion limit, which is the size of the smallest molecule incapable of entering the beads. Beads with a variety of compositions and pore sizes are available. For example, a Sephadex G-100 bead is made of a dextran polymer and should be able to separate molecules of up to 100 kDa; molecules bigger than 100 kDa will be completely excluded and will run at the void volume.

The "void volume" is the elution volume of a molecule so large that it is totally unable to enter any of the pores. The void volume therefore represents the minimum possible elution volume for any molecule. Another useful parameter for a gel filtration column is "column volume", which is the volume of an empty cylinder the size of the column being used. This represents the maximum elution volume for a very small molecule.

The protein sample should be applied to the column in the smallest reasonable volume to ensure good separation. The molecules do not form tight associations with the column, and if loaded in a large volume, will elute in a large volume with little separation between large and small molecules. The maximum volume that can be applied to a column varies somewhat with bead type (*e.g.*, about 2 to 3% of the bed volume for Sephadex, and about 0.5 % of the bed volume for Superdex).

The figure (below) shows a gel filtration chromatogram separating three proteins. The inset shows a standard curve (a plot of log of molecular weight *versus* elution volume) generated using a number of proteins (black squares); it also shows the migration positions of the three unknown proteins plotted on the standard curve line. From this graph it is possible to estimate the molecular weight of the unknown proteins. In this case, the three proteins shown were known to be the only proteins present in the solution. (Note: V_0 is the measured void volume for this particular column.)



Gel filtration chromatography can be performed using gravity (especially for large columns used for protein purification), or using low pressure or high pressure pumps (especially for analytical purposes). You will use both HPLC and gravity for your gel filtration experiments. HPLC columns have high resolution, and can usually be run fairly rapidly.

Certain problems are inherent in the process of gel filtration. First, even if the sample is applied as a narrow band to the column, the turbulence associated with passing the mobile phase through the column results in a broadening of the bands of the applied molecules as they travel through the column. The broadening of the bands works against the purification of the individual components of the sample, since the broad peaks can partially overlap one another. Thus, it is advisable to apply the sample as a narrow band on the top of the column. This can limit the usefulness of the method, because the sample may in some cases be difficult to concentrate sufficiently to make the method optimal. Another problem with gel filtration is that proteins may interact with the column matrix. This interaction will retard the elution of the protein and the apparent molecular weight will be lower than expected. To prevent interaction of the protein with the matrix the gel filtration should be run under moderate ionic strength to prevent ionic interactions between the proteins and the column matrix.

Separation on gel filtration columns increases in proportion to the square root of the column length. Longer columns therefore result in higher resolution. Unfortunately, both the cost of the column and the time the column takes to run increase in direct proportion to the length of the column.

Separation is inversely proportional to flow rate; thus, improved separation typically requires longer chromatography times. The maximum flow rate for a particular separation on a particular column can be estimated from theoretical considerations, but usually has to be determined empirically (*i.e.* try a flow rate, and see if the results are acceptable).

Materials

Phosphate buffer
    0.05 M sodium phosphate buffer, pH 7

Protein standards

HPLC
Hamilton Syringe
HPLC Gel Filtration Column

Gravity columns
Swollen Sephadex G-100
Column tubing
Column buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.4)

Procedure for the HPLC gel filtration

1. HPLC column will be pre-equilibrated with phosphate buffer before class.

2. Prepare protein standards as per instructor instructions.

3. Inject 50 μl of the protein standards into the injection loop with the manual injector set to "Load". Start the run by turning the injector to "Inject", and simultaneously start data collection by pressing Inject. (The instructor will show you how to use the HPLC). You will use an isocratic elution procedure run at 0.5 ml/min or 1 ml/min with the detector set for 280 nm.

4. After standards have eluted you will follow the same procedure for the LDH sample. The instructor will advise you on preparation of the LDH sample.

Make sure that the gel filtration columns do not run dry at any point. If the HPLC column runs dry, it may be irreparably damaged. If the gravity column runs dry, you will need to re-pour it.

Procedure for pouring the gravity gel filtration column

HPLC columns are packed by machines, and tend to cost several hundred dollars each. Gravity columns, however, are generally prepared by the user, and are usually much less expensive. You will use both types of column during the next two class sessions.

Preparing a gravity column requires preparing the beads. Gel filtration beads usually need to be "swollen". The swelling of the beads is a consequence of hydration of the beads, and usually takes 24 to 72 hours, depending on the bead type. You will use beads that have been swollen prior to the class period.

1. Make sure that the beads and buffer that you will use are at the temperature you will use for pouring the column. If you are pouring the column at room temperature using cold buffer, air bubbles will form as the buffer warms.

2. Gently suspend the beads in buffer using a stir rod. You want an even suspension, but you do not want to damage or crush the beads in the process. (Do not use a magnetic stirrer!) Make sure that a reasonable amount of buffer is present (too little buffer will make the suspension too viscous to pour evenly, while too much will require a long time to pour. A volume of buffer equivalent to about half the bead volume is usually reasonable.

3. Pour the bead suspension into the column. Use a stir rod to help run the suspension down the side of the column. Alternatively, tilt the column to make pouring easier. If air bubbles form, you will need to re-pour the column.

4. Allow the buffer to begin draining from the column. As the buffer drains, the beads will become more concentrated. In addition, the beads will begin to settle toward the bottom of the column. Periodically add more bead suspension to the

column. Do not allow the bead suspension to completely settle before adding more beads; this will create a discontinuity in the column that will adversely affect resolution. In addition, when pouring, try to avoid disturbing the bed of previously settled beads. You want the bed to be as even as possible.

5. After all of the beads have been added to the column, place the top on the column, and begin equilibrating the column with buffer.

Materials

Protein standards
Column buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.4)
Gravity Gel Filtration Column
Fraction collector
Test tubes for fraction collector

Calibrating the gravity gel filtration column is similar in concept, but different in actual procedure from that used in HPLC systems. Ultraviolet spectrophotometer flow detectors, such those commonly used with HPLC systems, are relatively expensive, and have limited usefulness. Many laboratories have general-purpose spectrophotometers but lack chromatography detectors. For these laboratories, someone usually has to measure the absorbance of the column fractions by hand.

1. Remove the top from the column.

2. Allow the buffer above the column bed to drain until no buffer is present above the bed.

3. As soon as the buffer reaches the top of the bed, begin loading the protein sample. **Do not allow the bed to dry out**. Try to avoid disturbing the top of the column bed while loading the sample.

4. Begin collecting fractions. Ideally, begin collecting fractions when ~50% of the sample has entered the bed.

5. Allow the sample to completely enter the bed. As soon as the sample has completely entered the bed, carefully add buffer, so that 2 to 3 cm of buffer is present above the bed. Try to avoid disturbing the top of the column bed while adding the buffer.

6. Replace the top of the column.

7. Make sure that enough buffer is available to complete the run. Run the column no faster than about 50 ml per hour (if possible; the column may not run that fast, depending on how tightly it is packed). Collect 1 ml fractions.

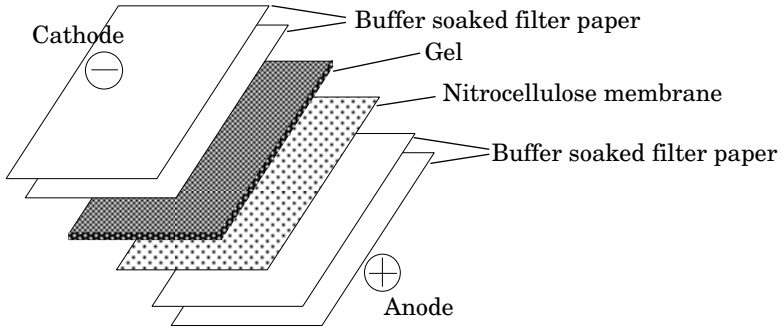
8. Measure the A_{280} of the fractions, and plot A_{280} *versus* elution volume.

Experiment 2C
Western Blotting

Western blotting is a technique that allows the identification of the location of one particular protein from an electrophoretically separated mixture of proteins. It is a technique in which proteins separated on a gel are transferred to a solid support and then probed with reagents that are specific for a particular amino acid sequence. For Western blots, the most commonly used probe is comprised of antibodies raised against the protein of interest; in this case, you will use antibodies raised against LDH. However, the protein-antibody interaction alone will not reveal the location of the protein on the solid support, since no color change occurs. The primary interaction must therefore be coupled with an interaction that produces a detectable product. Detectable products include colored precipitates, emitted light, or radioactive compounds that specifically bind the antibody.

Transfer of the gel

The first step of western blotting involves transferring the protein bands from the gel to a solid support. Various membranes are available, with polyvinyldifluoride (PVDF) and nitrocellulose being the most commonly used. PVDF membranes are favored because of their high mechanical strength, chemical stability, and enhanced protein binding; on the other hand, nitrocellulose is somewhat less expensive. The transfer of protein bands is accomplished by electrophoretic migration of the proteins from the gel to the membrane. The arrangement for a common procedure called "semi-dry blotting" is shown below.



Experimental Procedure:

Chemicals required:

Transfer Buffer
192 mM Glycine
25 mM Tris pH 8.3
1.3 mM SDS
20% methanol

Blocking solution (1% gelatin or 3% BSA in TBST + 0.02% sodium azide)

Hardware required:

Electrophoresis apparatus
100°C Water bath
Gel loading tips
Plastic trays for gel staining
Semi-Dry Transfer apparatus
Filter paper
Transfer membrane

Tris Buffered Saline (TBS)
50 mM Tris-HCl, pH 7.5
150 mM NaCl

TBST: TBS with 0.1% Tween-20

10% SDS polyacrylamide gels (see Experiment 2A)
Electrophoresis tank buffer
5x Sample buffer
Coomassie Staining Solution

Procedure:

1. **Set up the gel apparatus**
Place the gel in the apparatus. After adding the tank buffer, remove the bubbles from the bottom of the gel.

2. **Prepare the samples**
It is usually a good idea to use a previously performed SDS-PAGE gel to determine the best amount of protein to load. You want to choose a protein concentration that yields sharp, easily detectable, individual bands. Prepare samples as you did before, adjusting the amount of protein to be added based on the results from the previous gel.

3. **Load the samples**
If possible, load a sample that has purified LDH to act as a positive control. Try to load the samples in a way that will both allow ready identification of lanes of interest, and that separates samples with large amounts of protein, so that one lane will not contaminate the next.

Remember to record how you loaded the samples!

4. **Run the gel**
Attach the electrodes to the gel apparatus, and turn on the power supply. Run the gel at constant current (15 mA per gel) until the tracking dye reaches the end of the gel. Be careful: the high current levels used during gel electrophoresis can be dangerous!

5. **Disassemble the gel**
Turn off the power supply and remove the electrodes. Remove the gel sandwich and gently separate the glass plates. Remove the stacking gel and discard it.

6. **Transfer the protein to PVDF or nitrocellulose**
To prepare the "transphor sandwich" cut 4 pieces of filter paper and 1 piece of membrane to the same size as the gel. Briefly soak the filter paper and the

membrane in transfer buffer. Place 2 pieces of the soaked filter paper on the anode (positive electrode), followed by the membrane. Then place the gel on top of the membrane and put two more pieces of the soaked filter paper on top. Make sure no air bubbles are trapped underneath; you can roll a glass pipet or test tube over the transphor sandwich to remove any trapped air bubbles.

Transfer will occur for 30 minutes at an appropriate amperage. In most cases, an appropriate amperage can be calculated based on the size of the gel, in order to obtain 0.8 mA/cm^2 of gel. For your gels, this will be approximately 30 mA.

8. **Block the non-specific protein binding sites**
When the transfer is complete, place the membrane in Blocking Solution in a plastic tray on a shaker platform.

Although Western blots can be accomplished in a single day, in most cases, the process is a two-day procedure. The first day is used to run the gel, to transfer the proteins to the membrane, and to block the non-specific protein binding sites using the blocking solution. The second day is used to perform the antibody incubations and to carry out the detection process.

The second part of the Western blot procedure involves incubation of the membrane with the antibody and the detection reagents. An overview of the Western blot procedure is as follows:

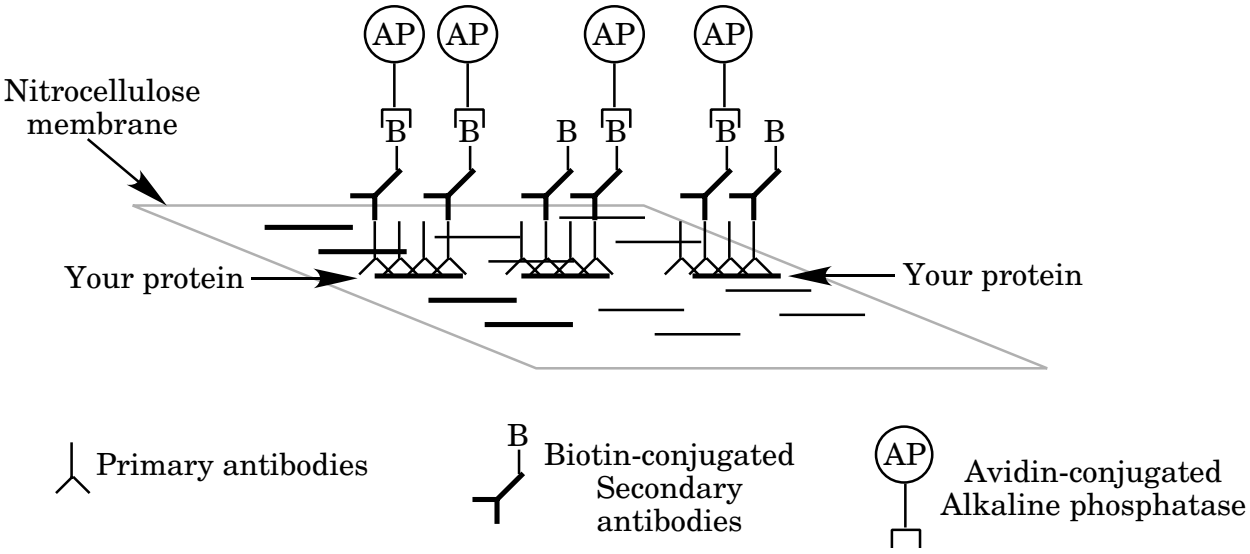
Incubations
1. ***Blocking***. The first step (which should be complete) is to incubate the membrane in a blocking solution (a solution which contains high concentration of protein). The blocking solution is usually prepared with bovine serum albumin, gelatin, or milk proteins[6]. This step is referred to "blocking" because the protein in the blocking solution binds to non-specific protein binding sites on the membrane, and therefore prevents antibodies (which are proteins) from binding to these sites.

2. ***Primary antibody***. After blocking, the membrane is incubated in a solution containing the primary antibody. The antibody typically used is an antiserum from rabbits inoculated with either the specific protein of interest, or a closely related protein. The antibody should bind to any of its antigen molecules present on the membrane. Most antibodies have very high binding affinities for their antigens, and the complex formed should remain intact during the wash steps. Excess antibody is removed by washing in a Tris-buffered saline solution that contains small amounts of the detergent Tween-20 (TBST buffer).

3. ***Secondary antibody.*** To detect the antibody-antigen interaction a secondary antibody is added. An example of a secondary antibody that can be used is a goat-anti-rabbit immunoglobin antibody. This is an antibody raised in goats that

[6] Milk protein is usually the least expensive, since powdered milk purchased at a grocery store can be used. However, the opaque white solution makes some of the later parts of the procedure more difficult.

recognizes rabbit immunoglobin. Incubation of the membrane will create an antigen:primary-antibody:secondary-antibody ternary complex as shown below. Excess secondary antibody is removed by washing in TBST buffer as was done for the primary antibody.



4. ***Color detection.*** One method used for color detection is to take advantage of the strong affinity between the protein avidin and the small molecule biotin. In this approach, biotin is covalently attached to the secondary antibody (creating the "biotin-conjugated goat anti-rabbit immunoglobin antibody" that is frequently used in these types of experiments). After incubation of the membrane with the biotin-conjugated antibody, the membrane is incubated with an enzyme covalently attached to avidin. The membrane containing the multiprotein complex (see the figure above) can then be incubated with substrates for the enzyme reaction that ultimately produces a detectable product. One commonly used enzyme for this process is *alkaline phosphatase*, which catalyzes reactions of the general form:

$$ROPO_3^{2-} + H_2O \rightarrow ROH + HPO^{4-}$$

The substrate 5-bromo-4-chloro-3-indolyl phosphate (**BCIP**) is dephosphorylated by the alkaline phosphatase; the dephosphorylated product can then react with nitro blue tetrazolium (**NBT**) to form a purple/black precipitate. This reaction can detect 25-50 pg of alkaline phosphatase on a Western blot. The sensitivity may be increased by increasing the incubation time with the substrates, which allows more product to form.

Another commonly used detection enzyme is *horseradish peroxidase*, which uses hydrogen peroxide to oxidize a wide variety of substrates. The products of the reaction may form a colored precipitate or may emit light.

Remember to wear gloves when handling the membrane to prevent finger proteins and oils from interfering in the experiment.

After the membrane has been incubated in blocking solution for a minimum of two hours, it is possible to begin the antibody incubations.

Note: The successive washes are time consuming. Work efficiently and be prepared for the next step!

1. Prepare an appropriate dilution of primary antibody in antibody buffer.

2. Transfer the blocked membrane to the primary antibody solution (made up in step 1). Incubate with shaking at room temperature for approximately 30 minutes.

3. Wash the membrane three times (5 minutes each wash) with 30 ml (each wash) of TBST buffer.

4. Prepare the secondary antibody dilution.

5. Transfer the blocked membrane to the secondary antibody solution (made up in step 4). Incubate with shaking at room temperature for approximately 30 minutes.

6. Wash the membrane three times (5 minutes each wash) with 30 ml (each wash) of TBST buffer.

7. Transfer the membrane to 30 ml of Avidin-AP solution and incubate at room temperature with shaking for approximately 45 minutes.

8. Wash the membrane three times (5 minutes each wash) with 30 ml (each wash) of TBST buffer.

9. Transfer the membrane to freshly prepared Development Solution. Let color develop sufficiently (color develops best in relatively low light). When color has developed, stop the reaction by immersing and rinsing the membrane in deionized water for approximately 10 minutes.

10. Photograph the developed membrane to create a permanent record of the experiment.

Experiment 2D
Native Gel Electrophoresis

SDS PAGE allows the separation of proteins based on their size. However, proteins are charged molecules; it is sometimes useful to use the intrinsic protein charge for separation purposes.

The protocol for running native gels varies depending on the proteins to be separated. While in SDS PAGE, all of the proteins are negatively charged due to the SDS, in native gels, both the sign and the magnitude of the charge depend on the specific protein and on the pH used in the experiment. If the incorrect pH is chosen, the protein may not migrate at all, or may migrate in the "wrong" direction, and disappear into the electrophoresis tank buffer. As a result, native gel electrophoresis is more rarely used than SDS PAGE, and its use requires more information about the proteins of interest.

Materials:

Electrophoresis chambers and power supplies

Gel Trays and Combs

Agarose

Reservoir Buffer
- 0.02 M Tris
- 0.02 M Glycine
- 0.002 M EDTA
- pH 8.6

5x Sample Buffer
- 0.2 M Tris
- 0.2 M Glycine
- 0.01 M EDTA
- 25% glycerol
- 0.02% bromophenol blue

Activity stain solutions

Solution A: 0.15 M lithium lactate, 1 mM NAD+, 15 mM NaCl, 8.5 mM $MgCl_2$, 80 mM potassium phosphate, pH 7.4.

Solution B: 5 mM Tetrazolium in 80 mM potassium phosphate pH 7.4

Solution C: 1 mM phenazine methosulfate in 80 mM potassium phosphate, pH 7.4

LDH isotrol standard (LDH 1-5 mixture)

Procedure

Preparation of Native Gels

1. Assemble gel apparatus.

2. Prepare a mixture of 0.8% w/v agarose in reservoir buffer (Tris-glycine pH 8.6). Your instructor will let you know the volume required.

3. Heat the flask in a microwave until the agarose has melted. Swirl to mix. This should take ~1-2 minutes on high. The agarose mixture should be clear with no significant clumps.

4. Allow the solution to cool for ~5 minutes in the cold room before pouring into casting trays. Gels should be poured in the cold room in order to accelerate hardening.

Sample Preparation and Application

1. Samples will be loaded as follows:

    Lane 1: LDH standard (LDH 1-5)
    Lane 2: Crude extract
    Lane 3: Desalt sample
    Lane 4: Purified LDH sample
    Lane 5: LDH standard (LDH 1-5)
    Lane 6: Crude extract
    Lane 7: Desalt sample
    Lane 8: Purified LDH sample

The LDH standard will be provided by the lab instructor. Your LDH samples should be prepared as determined for Experiment 3B except for the use of a different 5X sample buffer that does not contain SDS.

Gently mix the samples and samples buffer. Spin the tubes in the microfuge for a couple of seconds to bring the entire sample down. ***Do not boil samples***!

2. Transfer the casting trays to the electrophoresis unit with sample wells nearest the **black (–) electrode**. Be very careful, agarose is very slippery and has been known to slide off the trays.

3. Slowly fill the chamber with electrophoresis buffer until the buffer is about 1 cm over the gel.

4. Using a Pipetman, apply the samples to the wells. Load the entire sample if it will fit without overflowing, but do not let the solution overflow into the other wells.

5. Put on the lid and connect the leads to the power supply. Electrophorese at 100V until the marker bromophenol blue dye is close to running off of the end. At 125 V the separation would take about an hour and a quarter. However, the gel will get very hot and the bands will not give as clean a separation. At 80 V, the separation will take 2-3 h, but the results will be better.

Staining

1. Carefully cut the gel between lanes 4 and 5. Place one half of the gel in Coomassie Blue Staining Solution. The other half will be used for activity stain.

2. Place the half of the gel for activity stain in a weigh boat and add activity stain.

   The reagents must be added in the following order:

   a. Add 6.5 ml of solution A
   b. Add 2.5 ml of solution B
   c. Add 250 μl of solution C

3. Cover with aluminum foil and place on shaker for 5-15 minutes for bands to develop.

4. Photograph the gel.

Study Questions #2E

1. What is a precipitant? Why are precipitants necessary for crystallography?

2. Does the concentration of the precipitant within a well vary when using the hanging drop method? Why?

3. What is likely to happen if the cover slip is placed on the well without being sealed? How is the seal achieved?

4. In setting up crystal boxes, it is usually necessary to use fairly high concentrations of protein (usually 10 to 20 mg/ml). Why is this necessary?

5. How will you tell if crystals actually form?

Experiment 2E
Protein Crystallography

Proteins, and especially enzymes, have functions because of their precise placement of amino acid side-chains in three-dimensional space. In order to fully understand any protein, it is usually necessary to understand the three-dimensional structure of the protein.

The most frequently used method for determining molecular three-dimensional structures involves the use of X-ray crystallographic analysis. In this technique, a crystal of the molecule is irradiated with short wavelength photons (typical wavelength ~1.5 Å). These high-energy photons are diffracted by the molecular structure, and the diffracted photons can be collected and analyzed.

In order for X-ray diffraction to work, crystals of the molecule of interest are necessary. This is true because individual molecules do not diffract X-rays in sufficient quantities for analysis. In good crystals all of the molecules are arranged in a common pattern, and therefore the diffraction pattern from the crystal is equivalent to that from a single molecule, while providing enough diffracted photons to allow measurement of the diffraction pattern.

For small molecules, crystallization is relatively simple. However, proteins are large molecules, and crystallization can be correspondingly more difficult.

In order for proteins to crystallize, their tendency to remain in solution must be lower than their concentration in the solution. However, if this disparity is too great, the protein will aggregate and precipitate in a non-ordered fashion. It is therefore necessary to find conditions where the protein will leave the solution in an orderly manner.

If it were possible to predict the conditions that would allow the protein to crystallize, X-ray crystallography would be considerably simpler. However, in general, it is impossible to predict conditions that will allow crystallization of any given protein. Instead, the conditions must be determined empirically (*i.e*. by trying many different conditions until a condition under which crystals form is found).

The current method for finding crystallization conditions involves a screening procedure. The protein is tested with many different solutions that have previously been shown to result in crystals for other proteins. These solutions all contain **precipitants**, which are molecules that alter the structure of water in a way that induces proteins to leave the solution.

While the precipitant will induce the protein to leave the solution, the protein must leave the solution slowly. If the protein leaves the solution too rapidly, it will tend to aggregate rather than crystallize. This means that the concentration of precipitant must change slowly from a concentration too low to induce precipitation to a concentration above the precipitation threshold.

The usual method for raising the precipitant concentration slowly involves a technique called **vapor diffusion**. If two separate solutions are confined in a small sealed container, solvent from the more dilute solution will tend to evaporate and condense in the more concentrated solution until both solutions have the same solute concentration. This is especially true if the more concentrated solution has a much larger volume. If the protein is added to the more dilute solution, it will experience a slow increase in solute concentration during this process. If the protein is soluble in the dilute solution, and less soluble in the more concentrated solution, it will tend to leave the solution in a relative slow fashion, and therefore have the opportunity to crystallize.

The method you will use is a commonly used version of the vapor diffusion technique called the **hanging drop method**. In this method, a small volume the protein is mixed with a known ratio of the precipitant solution (usually involving a 1:2 dilution of the precipitant). This small volume of protein/precipitant is then inverted and suspended above a larger, undiluted solution of the precipitant. Assuming that the chamber is sealed, the concentration of precipitant in the protein solution will begin to rise. The time required for equilibration varies depending on precipitant type, and chamber and solution sizes, but typically involves 24-48 hours.

Finding the correct conditions for crystallization frequently involves setting up several thousand of these hanging drops using different types of precipitants, with different types of ions present, and different types of buffers present, while varying the pH and the concentration of all of these components. Precipitants commonly used for the technique include detergents, ammonium sulfate, polyethylene glycols of varying size, and small organic alcohols such as ethanol and isopropanol.

Crystallography projects tend to be relatively lengthy undertakings. The first step is obtaining purified protein (either by purifying it yourself, or by talking someone into purifying it for you). Because large-scale screening of crystallization conditions is usually necessary, it is rarely worth beginning a crystallography project unless at least 100 mg of purified protein are available. The second step is obtaining crystals, the step you are undertaking today. The third step is obtaining diffraction quality crystals. The crystals obtained from screening are usually too small, and are sometimes insufficiently well ordered to allow diffraction analysis. Once crystallization conditions are found, the conditions usually need to be modified to obtain suitable crystals. In general, reasonably large crystals are most useful; for proteins “large” crystals are about 200 μm in their smallest dimension.

The first three steps of a crystallography project involve fairly standard biochemistry. The remaining steps (collecting the diffraction data, analyzing the data, and solving the structure) are more complex, and are most commonly performed by individuals with at least some specialized training in the field of X-ray diffraction analysis.

Reagents and Supplies
24 well plates
Glass cover slips
5 ml syringes
Immersion oil or other viscous oil

Sodium acetate buffer

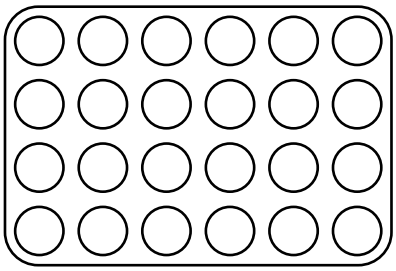
Sodium acetate: 1.361 g of sodium acetate in 50 ml water
Acetic acid: 0.572 ml glacial acetic acid in 50 ml water
Mix the acetic acid and sodium acetate solutions to yield a final pH of 4.7

10% NaCl (weight/volume) in sodium acetate buffer

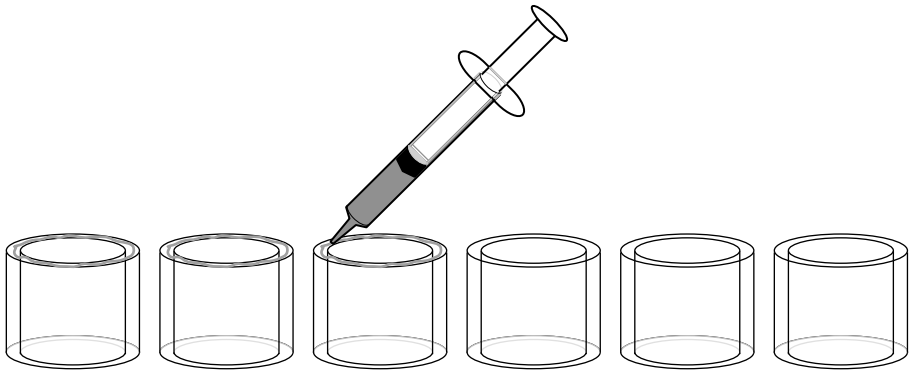
Lysozyme (50 mg/ml in sodium acetate buffer)

Setting up a "crystal box"
Because crystal condition screening requires testing a variety of conditions, the screening is frequently performed in 24-well culture plates. These allow a variety of conditions to be tested simultaneously, and are well suited to the hanging drop method.

The box is arranged in a 6 *x* 4 array. When setting up multiple conditions, this array is frequently used to allow a range of concentrations to be run in series across the box.



Procedure
Use a syringe to put a narrow band of oil around each of the six wells in the first row of the plate.

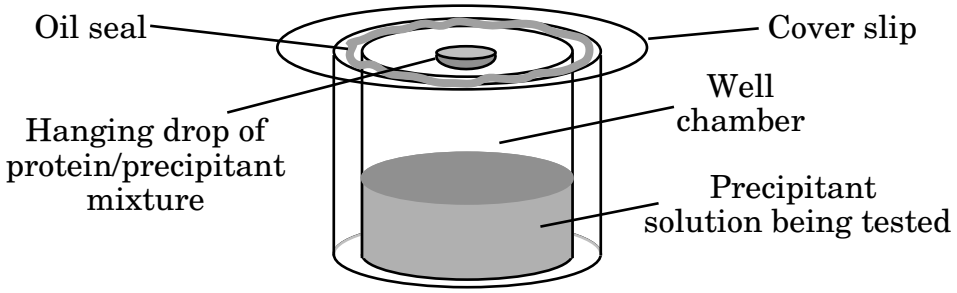


Add 1 ml of the solution to be tested in that well to each of the wells. Use 2%, 3%, 4%, 5%, 6%, and 7% NaCl in sodium acetate buffer as the well solutions.

Lay out six glass cover slips on a small piece of paper. Avoid touching the surface of the cover slips, and try to keep the cover slips as clean as possible. Place 2 μl of protein in the center of the each of the six cover slips.

Add 2 μl of the corresponding well solution to the protein drop.

Pick up the cover slip, and place it, inverted, on the well, making sure that the oil forms a seal around the entire perimeter of the well.



After all of the wells are set up, check the drops under a microscope for the immediate formation of precipitate, and for the presence of dirt, dust, and other objects that may appear as crystalline artifacts. After checking the drops, place the box in a cabinet where it will be protected from traffic.

Periodically (once every day or so) examine the drops under a microscope for the presence of crystals.

Introduction to Enzyme Kinetics

Lactate dehydrogenase is an enzyme. In biochemistry and physiology, enzymes are critically important molecules, and understanding them requires understanding their interactions with their substrates. While it is possible to work with crude preparations, working with purified proteins greatly reduces the probability that impurities in the enzyme mixture will affect the results. Thus far this quarter, you have attempted to purify LDH, and have run several types of experiments to assess the success of that endeavor. Now it is time to examine the enzymatic properties of LDH.

Enzymes are catalysts: they alter the rate of a reaction. When characterizing an enzyme, it is therefore necessary to study the reaction kinetics. A general scheme for a simple enzyme-catalyzed reaction is given in Equation 1:[7]

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \qquad \text{Equation 1}$$

Contemplation of this reaction scheme reveals that product formation is a first order function of the ES complex concentration. This first order rate of product formation, also termed ***velocity***, is usually expressed as:

$$velocity = \frac{d[P]}{dt} = k_2[ES] \qquad \text{Equation 2}$$

Standard chemical kinetic considerations state that concentration of ES complex depends on its rate of formation and its rate of disappearance. Allowing variations in ES complex concentration greatly increases the difficulty of studying the enzymatic reaction. Fortunately, it is usually possible to set up conditions in which ES complex variability is minimized. In most cases, the rate-limiting step is the formation of product. This means that, following initiation of the reaction, the equilibration:

$$E + S \rightleftharpoons ES$$

occurs before significant substrate is converted to product. Thereafter, as long as the substrate concentration does not vary significantly during the period of interest, then the amount of ES complex will be essentially constant. The assumption that

[7]Note that this scheme **is** applicable to reversible enzymes; it merely assumes that the initial concentration of product is zero, and therefore the reverse reaction does not occur. This situation is somewhat unusual in a physiological system, but is readily achieved in experiments performed using purified enzymes. Note also that this scheme assumes a single substrate; while LDH actually has two substrates (lactate and NAD), it is possible to perform kinetic experiments under conditions that allow the substrates to be examined individually. This is discussed more fully below.

the amount of ES complex is invariant is known as the **steady state assumption**, and is expressed mathematically as:[8]

$$\frac{d[\mathrm{ES}]}{dt} = k_1[\mathrm{E}]_f[\mathrm{S}] - k_{-1}[\mathrm{ES}] - k_2[\mathrm{ES}] = 0 \qquad \text{Equation 3}$$

In order for steady state conditions to be possible, the ***enzyme concentration must be a very small fraction of the substrate concentration***. This is necessary so that the conversion of substrate into product does not result in significant changes in substrate concentration. If substrate concentration decreases significantly, the ES complex concentration will also decrease, and Equation 3 will become a poor approximation of the true situation.

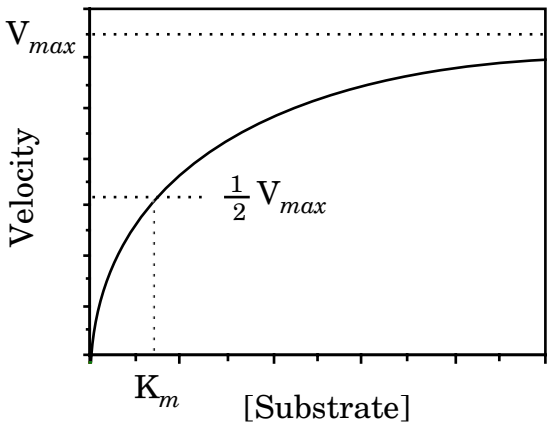
Using the steady state assumption, and some simple kinetics calculations, we can derive an equation for the velocity as a function of substrate concentration:

$$v = \frac{k_2[\mathrm{E}]_t[\mathrm{S}]}{\frac{k_{-1} + k_2}{k_1} + [\mathrm{S}]} \qquad \text{Equation 4}$$

If the amount of enzyme used ($[\mathrm{E}]_t$) is the same in each assay, then $k_2[\mathrm{E}]_t$ will be a constant, usually termed V_{max}. Combining the rate constants in the denominator into a new constant, K_m, results in an equation you have seen before, the Michaelis-Menten equation:

$$v = \frac{\mathrm{V}_{max}[\mathrm{S}]}{\mathrm{K}_m + [\mathrm{S}]} \qquad \text{Equation 5}$$

Contemplation of Equation 4 and Equation 5 reveals that velocity is a linear function of the amount of enzyme present (assuming that steady state conditions hold); for any substrate concentration, increasing the concentration of enzyme will increase the velocity in direct proportion. However, at any enzyme concentration, the velocity depends *non-linearly* on the amount of substrate present. The equation for velocity as a function of substrate concentration is that of a rectangular hyperbola.



The non-linear nature of the Michaelis-

[8]Equation 3 assumes that the reverse reaction can be ignored (a good assumption if product concentration is initially zero). Note the "$[\mathrm{E}]_f$"; this is the concentration of enzyme **not** bound to substrate; in most cases, $[\mathrm{E}]_f$ will be significantly less than the total enzyme concentration (usually abbreviated $[\mathrm{E}]_t$). In contrast, steady state conditions assume that $[\mathrm{S}]_t \approx [\mathrm{S}]_f$ (*i*.e. [ES] is a negligible fraction of $[\mathrm{S}]_t$ because $[\mathrm{S}]_t >> [\mathrm{E}]_t$), and therefore it is not necessary to correct $[\mathrm{S}]_t$ for the amount of S present in the ES complex.

Menten equation has caused some difficulty, because exactly calculating the parameters for non-linear equations from real experimental data by standard analytical methods is impossible. The most common method used for determining the K_m and V_{max} for an enzyme has historically been the use of one of three linear transformations of the Michaelis-Menten equation:

$$\frac{1}{v} = \frac{K_m}{V_{max}[S]} + \frac{1}{V_{max}} \qquad \text{Equation 6}$$

$$v = -K_m \frac{v}{[S]} + V_{max} \qquad \text{Equation 7}$$

$$\frac{[S]}{v} = \frac{K_m}{V_{max}} + \frac{[S]}{V_{max}} \qquad \text{Equation 8}$$

The linear transformations of the Michaelis-Menten equation allow the use of least-squares linear regression to determine the K_m and V_{max} for an enzyme from experimental velocity *versus* [S] data. Linear forms of the Michaelis-Menten equation are useful, because many calculators and most graph-generating computer programs contain linear regression algorithms (such as the "trend-line function" in Excel). In addition, it is much easier to interpret changes for linear plots than for non-linear ones.

Unfortunately, the use of these linear transformations of the Michaelis-Menten equation implies that the enzyme assay data contain **no experimental errors** (a clearly inaccurate assumption), because including an error term makes the mathematical transformation non-linear as well. In addition, the linear equations frequently result in inaccurate values of the parameters K_m and V_{max}, because they tend to be heavily affected by the error in some data points. This is particularly true for Equation 6 (the double-reciprocal transformation often called the Lineweaver-Burk equation), in which **the data point that affects the parameters K_m and V_{max} most strongly is the one the most likely to be inaccurate** (the point corresponding to the lowest substrate concentration).

One partial solution to the difficulties introduced by using linear transformations is to weight the data according to the likely errors. While several algorithms have been constructed to allow this, there is little agreement on the best weighting function to use with real data. Another partial solution to the problem of using linear transformations is to use more than one of them. If your data fit a rectangular hyperbola closely, all of the linear transformations will yield similar values for the kinetic parameters. On the other hand, if your data deviate significantly from hyperbolic behavior, linear regression fits to the different equations will result in markedly different values for K_m and V_{max}.

A much better method for determining the K_m and V_{max} values involves the use of non-linear regression. Until the advent of computers, performing non-linear regression was very difficult, because it is an iterative procedure. While most fairly sophisticated plotting programs include non-linear regression routines that allow data to be directly fit to the hyperbolic Michaelis-Menten equation, most students,

and many scientists, are unfamiliar with the technique. A procedure for performing non-linear regression on enzyme data is given below.

Side Note: Linear Regressions

In dealing with experimentally obtained numerical data, it is frequently necessary to propose a mathematical model that describes the system. In order to be useful, this model should have some basis in reality: the parameters in the model should describe properties of the system being studied. The above discussion summarizes a mathematical model (the Michaelis-Menten equation) that is commonly used to describe simple enzyme-catalyzed reactions. The mathematical model that describes enzyme-catalyzed product formation is a rectangular hyperbola, which is a non-linear equation.

Because linear equations are much easier to solve, and somewhat easier to understand, the hyperbolic function is usually transformed into one of several possible linear equations.

A linear equation is an equation of the form:

$$y = mx + b \qquad \text{Equation 9}$$

A linear equation states that, if x varies, y varies in direct proportion; m is the constant of proportionality (known as the slope of the line), and b is the value of y when $x = 0$ (known as the y-intercept).

When working with experimental data, all of the measured data points will not fit an equation of the line exactly. It is therefore necessary to estimate the slope and y-intercept that best fit the experimental data. For linear equations these parameters can be calculated analytically from the experimental data. The slope and y-intercept can be calculated using Equation 10 and Equation 11, respectively. Equation 10 and 11 are *least-squares linear regression equations:* the slope and y-intercept are chosen such that the values give the smallest possible average errors for each data point.

$$\mathrm{m} = \frac{\dfrac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} - \sum_{i=1}^{n} x_i y_i}{\dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} - \sum_{i=1}^{n} x_i^2} \qquad \text{Equation 10}$$

$$\mathrm{b} = \bar{y} - \mathrm{m}\bar{x} \qquad \text{Equation 11}^{9}$$

[9]In Equation 11, $\bar{x}$ is the mean of the x values, and $\bar{y}$ is the mean of the y values.

"Least-squares" refers to the common technique of squaring the error (*i.e.* the difference between the average value and the actual value) for each measurement. While the error can be positive or negative, the square of the error is always a positive number. The parameter that yields the smallest overall error (the "least" error) is the one chosen. The

$$\text{sum of the squared errors} = \sum_{i=1}^{n} \left(y_i - y_t \right)^2$$

where y_i is each observed data point, and y_t is the theoretical value for that data point based on the fit to the data. Although this procedure can be performed by hand, Equation 10 is moderately complex, and therefore it is usually desirable to use a computer program (or calculator) that deals with this equation automatically.

A useful measure of how well the data points fit the line is the correlation coefficient, usually abbreviated R. This is calculated by:

$$R = m \left(\frac{\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2}}{\sqrt{\frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2}} \right) \qquad \text{Equation 12}$$

A second, related, way of assessing the fit is R^2 (the square of the correlation coefficient); R^2 is the fraction of the variation in y that can be accounted for by the variation in x. If $R^2 = 1$, the data points all fit the line perfectly; if R^2 is less than 1, some of the variation in y is due to other factors (either to errors, or to non-linear behavior of the system). The smaller the R^2, the poorer the fit of the data to the line.

The major advantage of linearized forms of the Michaelis-Menten equation is that the linear regression can be solved exactly (non-linear regressions cannot be solved exactly). The linear forms of the Michaelis-Menten equation have some disadvantages. One problem (a problem common to all linear regressions) is that the slope depends heavily on the points at the end of the lines, especially if those points are well separated from points near the center of the line, while points near the center of the line have little effect on the slope. The other major problem, mentioned above, is that the linear forms of the Michaelis-Menten equation markedly distort the errors in the data points.

Side Note: Non-Linear Regressions
Many real world systems are non-linear. While it is frequently possible to linearize the mathematical model (as with the linear forms of the Michaelis-Menten equation), incautious use of these linear equations can result in incorrect, and sometimes nonsensical, sets of parameters. An alternative approach involves directly fitting the data to the non-linear equation, which has the advantage of allowing a fit of data to any mathematical model.

The technique of least-squares non-linear regression differs somewhat from methods used for analysis of linear equations. Because least-squares non-linear regression equations cannot be solved analytically, it is necessary to make an initial estimate of the equation parameters, and then perform iterative calculations intended to improve the fit between the theoretical and observed functions of the independent variables. Most computer programs (including Excel, see below) that allow this procedure have an algorithm for finding a minimum sum-of-the-squares.

Least-squares non-linear regression algorithms have some potential drawbacks. One potential problem is that, if the initial estimate of the parameters is too far from the correct value, the algorithm may fail to converge on an answer. A second potential problem is that the mathematical function may exhibit local minima of sum-of-the-squares, and therefore converge on incorrect values. Finally, for real data, it may not be possible to distinguish between deviations due to experimental error and deviations due to an incomplete mathematical model.

Catalytic Rate Constants

The Michaelis-Menten equation (Equation 5) contains two parameters: K_m and V_{max}. While K_m is a constant for each enzyme, V_{max} is not really a constant, because it depends on the enzyme concentration. Enzyme concentrations can be chosen arbitrarily; different experiments may intentionally or inadvertently use different enzyme concentrations. Factoring out the enzyme concentration from the V_{max} leaves a rate constant that is, like the K_m, a constant physical property of the enzyme.[10] Comparing Equations 4 and 5 suggests that this rate constant is k_2; this is correct only for enzymes described by the simple reaction scheme shown in Equation 1. A more general term for this is the catalytic rate constant, k_{cat} (also known as the *turnover number*); k_{cat} is the rate constant that characterizes the slowest step of the enzymatic process under investigation. The k_{cat} is the number of substrate molecules converted to product by a single enzyme molecule per unit time.

For any enzyme, k_{cat} is a far more useful value than V_{max}, because k_{cat} is independent of enzyme concentration. Calculation of k_{cat} is straightforward:

$$k_{cat} = \frac{V_{max}}{[E]_t} \qquad \text{Equation 13}$$

Obviously, the use of Equation 13 means that you need to have determined the concentration of the enzyme in some way. Obviously also, the enzyme concentration in Equation 13 is the concentration in the reaction, not the concentration in your stock solution.

You have measured the amount of protein present in your purified LDH fraction. If you assume that this protein concentration (in mg/ml) represents the LDH

[10]Actually, both K_m and k_{cat} are only constant under a given set of conditions. For Michaelis-Menten enzymes, these values are independent of enzyme concentration, but are altered by changes in temperature, pH, and other factors. Small changes in these parameters probably accounts for some of the observed experimental errors.

concentration (in mg/ml), you can use this protein concentration and the molecular weight of LDH to determine the concentration.

Multisubstrate Enzymes

The enzyme kinetics concepts briefly outlined above apply primarily to relatively simple enzymes. Enzymes with more than one substrate, or more than one active site, may not exhibit Michaelis-Menten kinetics. Lactate dehydrogenase has two substrates (the organic acid and the nicotinamide coenzyme in the appropriate redox state); in addition, the tetrameric enzyme may have a total of four active sites for each substrate. Depending on the experimental conditions used, LDH may not be a true Michaelis-Menten enzyme.

Dealing with multiple substrates is *relatively* straightforward: if the concentration of one substrate is varied while the other one is held constant at a high value, the kinetics will exhibit pseudo-first order behavior. Thus, if NAD concentration is high, and lactate concentration is varied, the K_m and V_{max} values for lactate can be calculated from the observed velocity data. Note that the V_{max} value obtained will be somewhat lower than the true value: the true V_{max} value is an asymptote in the relevant equation, and its determination requires extrapolation to infinite concentrations of *both* substrates.

Enzymes with more than one active site may exhibit cooperativity, in which binding of one substrate alters the binding affinity for other substrate molecules and/or the catalytic rate constant for the reaction. Experimentally, differentiating between deviations from hyperbolic behavior due to cooperativity or due to poor experimental technique is often a non-trivial task. For the purposes of this course, you can assume that LDH does not exhibit cooperativity (this turns out to be a reasonable assumption under most conditions).

Procedural Notes: Temperature

Enzyme-catalyzed reactions, like all chemical processes, have temperature-dependent rates. This means that, in order to obtain meaningful results, you need to perform all of the measurements at the same temperature. You will be performing your experiments at room temperature. This means that you should keep all of your reagents (except the LDH) at room temperature. If you keep the LDH assay solutions on ice, you will find it very difficult to control the temperature during the measurements. Note that the volume of LDH solution you will add to each cuvette is small enough to have a negligible effect on the temperature in the cuvette; because of this, and because, like most enzymes, LDH is more stable at low temperatures, it is a good idea to keep your LDH solution cold until you need to use it.

Procedural Notes: Initial Rates

In designing an experiment to determine the K_m and V_{max} values for an enzyme, it is necessary to know something about the enzyme. Obviously, you need to know what substrate to use. In addition, you need to use substrate and enzyme concentrations that result in steady state conditions.

As noted above, the rate of product formation in an enzymatic reaction is dependent on the concentration of the ES complex (see Equation 2). Under steady state conditions, the concentration of ES complex is essentially constant, and therefore velocity should also be constant. Thus, when you measure product concentration, you should observe a linear increase with time. This linear phase (the linear [P] *vs*. time plot) is the velocity that you need to measure. If the enzyme concentration is too high, you will not observe a linear phase in the reaction, and will not be able to measure a meaningful velocity. Alternatively, if you allow the reaction to proceed for too long a period of time, you will reach conditions in which the decrease in substrate concentration is significant, and steady state conditions will no longer exist. It is therefore critical that you examine your data to be sure that you have a linear rate during each measurement.

Variation in ES complex concentration can also be a consequence of varying enzyme concentration. Most experiments assume that enzyme concentration is a constant. Protein solutions are somewhat difficult to pipet accurately, you will need to exercise care in pipetting your enzyme solutions to prevent variations in enzyme concentration that will greatly complicate your analyses.

Procedural Notes: Performing Analyses on Real Data

Using the linear forms of the Michaelis-Menten equation requires figuring out which of the terms should be plotted as x and y, and figuring out how to determine the K_m and V_{max} values from the regression analysis. Table 1 shows the relevant information for each of the three most common linear plots used for analyzing enzyme kinetic data. The names given reflect the authors of the early papers that used each plot. Depending on the source, you may find other names associated with some of these plots (especially the plot for Equation 6).

Table 1

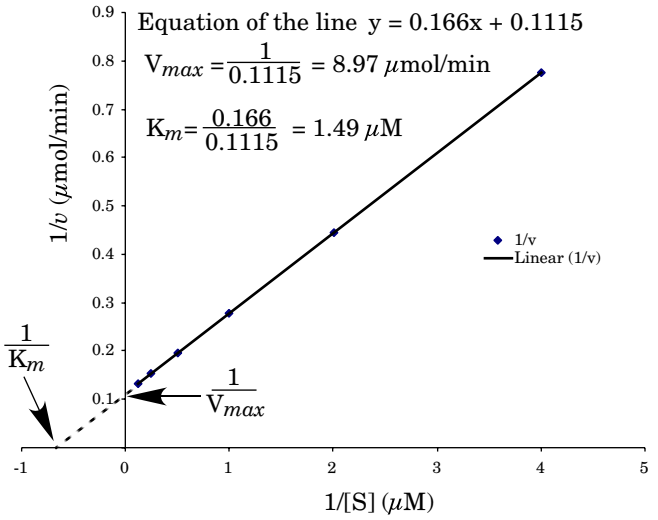
Plot	x	y	Slope	y-intercept
Equation 6 (Lineweaver-Burk)	$\frac{1}{[S]}$	$\frac{1}{v}$	$\frac{K_m}{V_{max}}$	$\frac{1}{V_{max}}$
Equation 7 (Eadie-Hofstee)	$\frac{v}{[S]}$	v	$-K_m$	V_{max}
Equation 8 (Hanes-Woolf)	$[S]$	$\frac{[S]}{v}$	$\frac{1}{V_{max}}$	$\frac{K_m}{V_{max}}$

Setting up one of these plots requires obtaining a series of velocities that correspond to substrate concentrations. These data are then transformed to the value to be plotted according to the relevant equation, and plotted on a graph. The transformed data are then subjected to linear regression analysis (using a function such as the

trend-line algorithm of Excel), and the slope and *y*-intercept values then used to determine the K_m and V_{max} values.

The example below shows this procedure for a set of data using the Lineweaver-Burk plot.

[S] (μM)	***v* (μmol/min)**	**1/[S]**	**1/v**
0.25	1.29	4	0.7752
0.5	2.25	2	0.4444
1	3.6	1	0.2778
2	5.14	0.5	0.1946
4	6.55	0.25	0.1527
8	7.58	0.125	0.1319

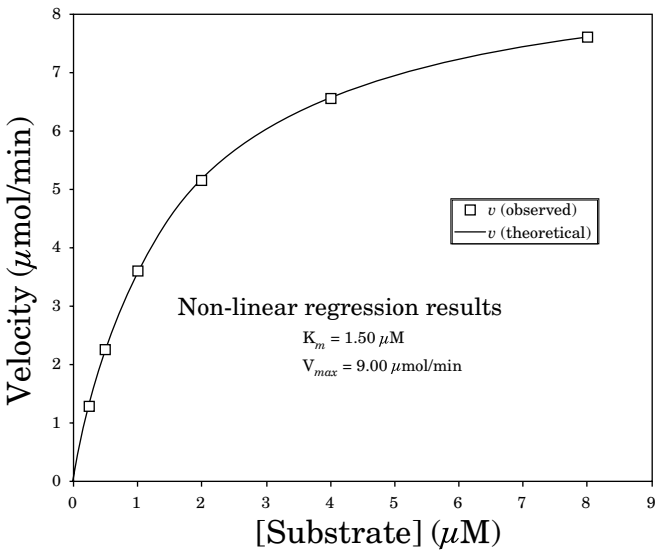


The equation of the line (y = 0.166x + 0.1115) includes both the slope and y-intercept information. In the case of the Lineweaver-Burk plot, V_{max} = 1/(*y*-intercept), and K_m = (slope)/(*y*-intercept). These relationships are shown in Table 1.

Alternatively, the substrate and velocity data can be fit directly to the Michaelis-Menten equation using non-linear regression. The procedure for doing so using Excel is as follows:

Using Non-linear regression in Excel

1. Make sure that the Solver option is present in the Excel Tools menu. (If not, select **Add-ins** from the **Tools** menu, and add the Solver.)
2. Enter the data to be analyzed. (Note that the cell references below are specific for the example spreadsheet from the course website; you may need to use different ones.)
3. Create a column of squared errors (for the Michaelis-Menten equation, the formula would be: (C9-(D$4*B9)/(D$3+B9)))^2, where C9 is the velocity data point, B9 is the substrate concentration data point, and D$4 and D$3 are the V_{max} and K_m values, respectively. (Note that the $ in from of the row number allows you to fill down without changing the row reference.)
4. Create a cell that has the sum of the squared errors column (*e.g.*, D5).
5. Enter initial values for V_{max} and K_m cells. The values should be reasonably close to the actual values, whenever possible.
6. Select **Solver** from the **Tools** menu. In the dialog box, Set Target Cell D5, Equal To: **Min** By Changing Cells D3 and D4.
7. Click **Solve**.



Note that the values for K_m and V_{max} shown on both graphs include units. These units are critically important for comparison purposes. You should **always** include the units for these parameters any time that you report results from a kinetic analysis of an enzyme.

The data given in the example fit a rectangular hyperbola fairly closely, and as a result can be readily analyzed using either the Lineweaver-Burk plot or non-linear regression. However, most real experiments do not yield data this well behaved. You ***therefore need to look at the data and at the resulting graphs***, and not merely accept (and report) the values calculated by Excel. ***There is no substitute for having an intelligent human look at the results*** of a kinetic (or any other kind of) experiment.

References

All biochemistry textbooks have sections on enzyme kinetics. In addition, a number of books have been written on enzyme kinetics. The classic textbook is:

Segel, I.H. *Enzyme Kinetics* John Wiley & Sons, New York, 1975.

Other important literature includes:

Fischer, E. (1894) *Berichte* **27**, 2985.

Henri, V. (1902) *Acad. Sci., Paris* **135**, 916.

Michaelis, L. and Menten, M. (1913) *Biochem. Z.* **49**, 333.

Haldane, J.B.S. *Enzymes* Longmans, Green & Comp. 1930, reprinted in 1965 by MIT Press.

Lineweaver, H. and Burk, D. (1934) *J. Am. Chem. Soc.* **56**, 658.

Eisenthal, R. and Cornish-Bowden, A. (1974) *Biochem. J.* **139**, 715.

Study Questions #3A

1a. You are working with an enzyme, and obtain the data listed in the Table (below). What are the K_m and V_{max} values based on this data? (Hint: using any of the linear transformation plots will work for this problem.)

[Substrate] (mM)	Velocity (μmol/min)
0.25	**1.29**
0.5	2.25
1	3.6
2	5.14
4	6.55
8	**7.58**

1b. Divide the **last** velocity value in the Table by two, while using the other values from the Table. Analyze the resulting data using each of the three linear transformation equations. How does the artificial experimental error affect the apparent K_m and V_{max} values?

1c. Divide the **first** velocity value in the Table by two, while using the other values from the Table. Analyze the resulting data using each of the three linear transformation equations. How does the artificial experimental error affect the apparent K_m and V_{max} values?

1d. Which linear transformation is most reliable for each of the different types of experimental errors? Why?

1e. Which error is more likely to occur: an error at low substrate or at high substrate concentrations?

2. Under the assumptions for steady state conditions, is the substrate concentration allowed to change? Does this make sense for a real enzyme reaction?

3a. You set up a reaction where the concentration of substrate is equal to the K_m. Can this be a steady state condition? Why?

3b. You set up a reaction where the concentration of substrate is equal to the enzyme concentration. Can this be a steady state condition? Why?

4. When running the LDH assay using the lactate stock solution, what is the lactate concentration in the cuvette? Why do you need to know this value? (Hint: this is an important question!)

5. What is a good dilution of your peak fraction to test? Why do you think so?

Experiment #3A
[LDH] *versus* Activity and Lactate K_m Determination

In the next few classes, you will perform several different types of kinetic analyses, including determining K_m and V_{max} values using both lactate and pyruvate as substrates, studying the effects of an inhibitor, and studying the effects of chemical modification of the enzyme.

In order to determine the K_m and V_{max} values of your LDH preparations, you will first need to find a dilution of your LDH preparation that will allow steady-state conditions. You previously ran LDH assays on several dilutions of your purified LDH peak fraction in an attempt to find a useful dilution. Today, you will need to repeat that experiment. **Use your notes as a guide for choosing reasonable dilutions, but you will still need to run the experiment**: it is possible that the activity of your preparations may have changed over the last few weeks. You need to aim for a dilution which yields an activity of about 0.3 to 0.5 ΔA/min using undiluted lactate stock solution (much higher than this will probably result in non-steady-state conditions; much lower will make it difficult to obtain reliable velocity measurements, especially at lower substrate concentrations).

Once you have chosen an enzyme dilution to use, you will need to make an initial estimate of the K_m and V_{max} values. This type of experiment requires varying the substrate concentration, while keeping the concentrations of all of the other assay components constant. Thus, you should make several dilutions of the lactate stock solution, and add these in place of the lactate stock normally used in the LDH assay. You need to run several different lactate concentrations (at least six) that result in a range of velocity values. Although it is impossible to actually assay the enzyme at the V_{max}, you can assume that if you run an assay at a lactate concentration that yields a velocity at least a factor of 2 lower than the activity obtained at the highest lactate concentration, you have probably bracketed the K_m value.

Reagents for LDH Assay:

Lactate Stock Solution
120 mM lithium lactate
10 mM Tris-HCl (pH 8.6)

NAD^+ Stock Solution
12 mM NAD^+
10 mM Tris-HCl (pH 8.6)

Bicarbonate Stock Solution
18 mM $NaHCO_3$
0.5 M NaCl

Tris buffer for Enzyme Dilution
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol

Determining the correct LDH dilution:

Make a series of LDH dilutions in Tris buffer.

Mix 0.6 ml Lactate Stock Solution, 0.4 ml NAD Stock Solution, and 0.2 ml Bicarbonate Stock Solution in a cuvette. Add 10 μl of the enzyme dilution to be tested, and collect data. Attempt to find a convenient enzyme dilution that results in a rate of 0.3 to 0.5 ΔA/min.

Once you have found a satisfactory LDH dilution, you will need to make a series of lactate dilutions, and use these diluted solutions instead of the lactate stock in the assay. Run at least 6 different dilutions of lactate; make sure the dilutions of lactate give you a range of velocities.

Make sure that you make up enough diluted LDH for all of the assays you will need to run. Making a fresh dilution of the enzyme in the middle of the assay often changes the results somewhat, and makes data analysis more difficult.

Repeat the LDH lactate K_m and V_{max} determination. This time, use your data from the last experiment to chose lactate concentrations that will bracket the K_m; try to use at least three substrate concentrations above the K_m and at least three below it.

Check the results of your assay by estimating the K_m and V_{max} for your enzyme. You should have a good rectangular hyperbola; if your data do not appear hyperbolic, repeat the experiment.

Study Questions #3B

1. When you ran the LDH assay using lactate, you could use a high concentration of NAD. In running the LDH assay using pyruvate, you must use a lower concentration of NADH. Why? (Hint: see next two questions.)

2. When you ran the LDH assay using lactate and NAD, what happened to the absorbance of the solution after you added the enzyme?

3. When you run the LDH assay using pyruvate and NADH, what do you expect to happen to the absorbance of the solution after you add the enzyme?

4. Do you expect pyruvate to have the same K_m value as lactate? Why? (Hint: how do enzymes really work?)

5. Do you expect pyruvate to have the same V_{max} value as lactate? Why? (Hint: how do enzymes really work?)

6. Based on your data, what is the k_{cat} for the conversion of lactate to pyruvate by your LDH preparation? In order to calculate k_{cat}, you need to know the amount of enzyme you used, and the concentration of your LDH stock solution. Calculate the concentration of LDH based on the Bradford protein assay you ran on your peak fraction, assuming that all of the protein in that sample is LDH. (Note: this is a moderately complicated calculation; doing it now will greatly help your analysis for your lab report.)

Experiment #3B
Pyruvate K_m Determination

Now that you have completed the lactate K_m and V_{max} determination, you will need to perform a similar experiment running the reaction in the reverse direction. In this case, you will be using pyruvate and NADH as substrates, instead of lactate and NAD. The enzyme dilution you used for the lactate experiment may also work for pyruvate, but it is a very good idea to check.

For the pyruvate assay, use 0.6 ml Pyruvate Stock Solution, 0.4 ml NADH Diluted Stock Solution, and 0.2 ml Bicarbonate Stock Solution in a cuvette.

Solutions Required

NADH Stock solution
0.5 mM NADH
10 mM Tris-HCl (pH 8.6)

Pyruvate Stock Solution
60 mM Pyruvate
10 mM Tris-HCl (pH 8.6)

Bicarbonate Stock Solution
18 mM $NaHCO_3$
0.5 M NaCl

Tris buffer for Enzyme Dilution
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol

Study Questions #3C

1. Collection of % Inhibition *versus* [Inhibitor] data is an iterative process: you usually do not know before you actually run the experiment which concentrations will be useful to run. Which inhibitor concentrations do you think you should run first?

2. Estimate the I_{50} from the data below. How did you come to this conclusion?

[I] (mM)	Activity (ΔA/minute)
0	0.53
3.75	0.38
7.5	0.29
15	0.19
30	0.11
60	0.05
120	0.02

3. Use the I_{50} (from question 2) to determine the K_i, assuming that the inhibitor is competitive, and that the data were collected using a substrate concentration of 15 mM and a K_m of 15 mM.

4. What is the advantage of using a substrate concentration near the K_m in I_{50} determinations?

5. Calculate the k_{cat} for your enzyme for both pyruvate and lactate.

Experiment #3C
LDH Inhibition Kinetics

For an enzyme inhibitor, two properties are of interest. The first is the *type* of inhibitor (competitive or mixed), and the second is the K_i, the *dissociation constant* of the inhibitor for the enzyme.

In many explanations of enzyme kinetics, inhibition type is considered to have three categories: competitive, uncompetitive, and non-competitive. Kinetically, these are all borderline cases. Using only a kinetic definition, "competitive inhibitor" means that the inhibitor only binds to the free enzyme, and prevents binding of substrate, "uncompetitive inhibitor" means that the inhibitor only binds to the ES complex (and inactivates it), and "non-competitive inhibitor" means that the inhibitor binds to (and inactivates) the enzyme with the same affinity regardless of the presence of substrate. The kinetic definition makes no claims about the actual location of the inhibitor binding site.

Although some examples exist of true uncompetitive and non-competitive inhibitors, in most cases, the kinetics are not quite that simple. The reason for this stems from the mechanism by which inhibitors bind to enzymes. While true competitive inhibition is common – especially when the inhibitor binds only to the free enzyme because it binds to the same site as the substrate[11] – true uncompetitive and non-competitive inhibition are relatively rare. In these types of inhibition, the inhibitor usually binds to a part of the enzyme spatially distinct from the active site. While it is possible that inhibitor binding will only occur if the active site is occupied by substrate (uncompetitive inhibition), in most cases, the inhibitor will have some affinity for the unoccupied enzyme as well.

Non-competitive inhibition is also somewhat rare: it requires that the inhibitor affinity be unchanged regardless of whether substrate is bound or not. A true non-competitive inhibitor usually functions by altering the enzyme conformation; in most cases, this conformational change is a two-way process, with the substrate also altering the binding site for the inhibitor. Thus the affinity for the inhibitor usually changes when substrate is bound.

Kinetic experiments can be a method for examining the mechanism and site of inhibitor binding. In many (although not all) cases, either the inhibitor binds to the active site and is a competitive inhibitor, or it binds elsewhere and is a mixed inhibitor[12].
Mixed inhibition has one additional complicating factor: a number of mixed inhibitors do not completely abolish activity. Since the inhibitor binds at a location

[11]Enzyme kinetics experiments usually cannot tell the difference between an inhibitor that actually binds to the same site as the substrate and one which does not; if binding of substrate and inhibitor is mutually exclusive (regardless of the actual location of the inhibitor binding site), the inhibition will appear competitive. Although there are exceptions, most competitive inhibitors do bind to the same site as the substrate.

[12]The term "mixed inhibitor" also avoids the problem associated with the rather poorly chosen names "uncompetitive" and "non-competitive" inhibition; it is difficult to determine from these names what type of inhibition is being described.

distinct from the active site, the substrate can still bind at the normal active site and be converted to product (albeit at a significantly reduced rate). This tends to complicate analyses of mixed inhibition. (Many biochemistry textbooks ignore this possible condition, but many real inhibitors do exhibit this phenomenon.)

In competitive inhibition, the binding of inhibitor and substrate are mutually exclusive. This results in a lowering of the apparent affinity of the enzyme for the substrate; the presence of the inhibitor raises the substrate concentration required to yield any given ES complex concentration. Therefore, the apparent K_m for the substrate increases. However, the maximal amount of ES complex is the same (the maximal amount of ES complex depends only on the amount of enzyme present, a parameter which is not changed by the presence of a competitive inhibitor), and therefore the V_{max} remains constant.

In mixed inhibition the binding of inhibitor and substrate are not mutually exclusive, and the effect of the inhibitor is to lower V_{max} and (in most cases) to alter the apparent K_m.

In order to determine the type of inhibitor, the K_m and V_{max} must be determined in the absence of inhibitor, and in the presence of more than one concentration of inhibitor. These data can then be analyzed for changes in the apparent K_m or V_{max} (*i.e.* a change in the parameters as a function of inhibitor concentration). Note that you are looking for changes in parameters that are already subject to error (multiple K_m determinations rarely yield exactly the same value); you will find determination of inhibitor type and K_i to be difficult unless you obtain good data.

Equation 14 is a general equation for the effect of an inhibitor:

$$v = \frac{V_{max}[S]}{K_m\left(1+\frac{[I]}{K_{i_a}}\right)+[S]\left(1+\frac{[I]}{K_{i_b}}\right)} \qquad \text{Equation 14}$$

Equation 14 has two K_i terms: K_{i_a} and K_{i_b}. K_{i_a} is the inhibitor dissociation constant for the free enzyme; K_{i_b} is the inhibitor dissociation constant for the ES complex. Equation 14 is directly applicable to mixed inhibition. Non-competitive inhibition is the special case in which K_{i_a} and K_{i_b} are equal.

For competitive inhibition $K_{i_b} = \infty$ (*i.e.* the inhibitor has zero affinity for the ES complex), and therefore the equation reduces to:

$$v = \frac{V_{max}[S]}{K_m\left(1+\frac{[I]}{K_{i_a}}\right)+[S]} \qquad \text{Equation 15}$$

Thus, for competitive inhibitors, the apparent K_m should vary from the K_m determined in the absence of inhibitor as a function of inhibitor concentration and K_i:

$$K_{m_{app}} = K_m\left(1 + \frac{[I]}{K_i}\right) \qquad \text{Equation 16}$$

As mentioned above, mixed inhibitors do not always abolish activity. In order to examine this possibility, and in order to choose reasonable inhibitor concentrations for performing K_m and V_{max} determinations, experiments with varying inhibitor and constant substrate concentrations are frequently performed. These experiments are often called I_{50} determinations, since the I_{50}, the concentration of inhibitor required for a 50% decrease in activity, is the major parameter.

In competitive inhibition, and some types of mixed inhibition, the plot of % Inhibition[13] *versus* Inhibitor concentration is approximately hyperbolic, with an asymptote at 100% inhibition. If the plot has an apparent asymptote at significantly less than 100% inhibition, the inhibitor is unlikely to be a competitive inhibitor.

If the inhibitor type is known, the I_{50} can be used to estimate the K_i. (Note: calculation of K_i from apparent K_m and V_{max} values, discussed below, is usually more accurate.) For **competitive** inhibitors, the K_i is given by:

$$K_i = \frac{I_{50}}{\frac{[S]}{K_m} + 1} \qquad \text{Equation 17}$$

In the limiting case where the inhibitor binds to both the free enzyme and ES complex with the same affinity (*i.e.* **non-competitive** inhibition):

$$K_i = I_{50} \qquad \text{Equation 18}$$

In the limiting case where the inhibitor only binds to the ES complex, and the ESI complex is completely inert (*i.e.* **uncompetitive** inhibition):

$$K_i = \frac{I_{50}}{\frac{K_m}{[S]} + 1} \qquad \text{Equation 19}$$

If your data suggest that the ESI complex does, in fact, retain some catalytic activity (for example, if you find that, as inhibitor concentration increases, you are

[13] Percent inhibition $= \left(1 - \frac{\text{Activity with inhibitor}}{\text{Activity without inhibitor}}\right) \bullet 100$

not approaching 100% inhibition), you will probably find it somewhat difficult to determine the K_i, using any of these equations.

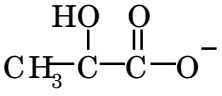
In most cases, the type of inhibitor is determined by measuring velocity for several different substrate concentrations in the presence of several different inhibitor concentrations. These data allow the comparison of the apparent K_m and V_{max} in the absence of inhibitor to the values obtained in the presence of several inhibitor concentrations. The Table below both indicates how the apparent K_m and V_{max} should vary, and gives equations that can be used to calculate the K_i value based in the data obtained from these experiments. In Experiment 3D (below) you will collect such data, which should both allow you to assess the inhibitor type and allow you to calculate the K_i for the inhibitor.

Type of Inhibition	$K_{m_{app}}$	$V_{max_{app}}$
None	K_m	V_{max}
Competitive (Inhibitor only binds free enzyme)	$K_m\left(1+\frac{[I]}{K_i}\right)$	V_{max}
Mixed (Inhibitor binds both free enzyme and ES complex)	$K_m\frac{\left(1+\frac{[I]}{K_{i_a}}\right)}{\left(1+\frac{[I]}{K_{i_b}}\right)}$	$\frac{V_{max}}{\left(1+\frac{[I]}{K_{i_b}}\right)}$
Non-competitive (Inhibitor binds free enzyme and ES complex with identical affinity)	K_m	$\frac{V_{max}}{\left(1+\frac{[I]}{K_i}\right)}$
Uncompetitive (Inhibitor binds only ES complex)	$\frac{K_m}{\left(1+\frac{[I]}{K_{i_b}}\right)}$	$\frac{V_{max}}{\left(1+\frac{[I]}{K_{i_b}}\right)}$

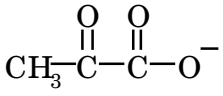
Note: in the Table, "K_m" and "V_{max}" are the values obtained in the absence of inhibitor, while "$K_{m_{app}}$" and "$V_{max_{app}}$" are the apparent K_m and V_{max} values obtained in the presence of the inhibitor.

Experimental procedures:

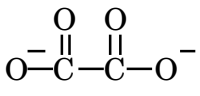
Different compounds may inhibit enzymes. Test each of pyruvate, oxalate, and malonate for their ability to inhibit LDH using lactate (with NAD) as the substrate. If time allows, try testing the compounds for their ability to inhibit LDH using pyruvate (with NADH) as the substrate. Do you observe the same concentration range for the effects?



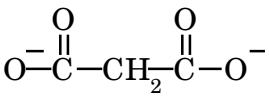
Lactate



Pyruvate



Oxalate



Malonate

Perform an I_{50} determination for each of the potentially inhibitory molecules. When performing the I_{50} determination, it is usually best to use a constant substrate concentration near the K_m, and to vary the concentration of inhibitor over a fairly wide range.

In performing the enzyme assays, add an additional 10 μl of either buffer or inhibitor to each cuvette (this maintains the same volume in all of the cuvettes).

Solutions Required

Lactate Stock Solution
120 mM lithium lactate
10 mM Tris-HCl (pH 8.6)

NAD$^+$ Stock Solution
12 mM NAD$^+$
10 mM Tris-HCl (pH 8.6)

Bicarbonate Stock Solution
18 mM $NaHCO_3$
0.5 M NaCl

Tris buffer for Enzyme Dilution
10 mM Tris-HCl (pH 8.6)
0.5 mM 2-Mercaptoethanol

Putative Inhibitor Stock Solution
300 mM oxalate
10 mM Tris-HCl (pH 8.6)

Putative Inhibitor Stock Solution
300 mM malonate
10 mM Tris-HCl (pH 8.6)

Pyruvate Stock Solution
60 mM Pyruvate
10 mM Tris-HCl (pH 8.6)

NADH Stock solution
0.5 mM NADH
10 mM Tris-HCl (pH 8.6)

Study Questions #3D

1. What were the I_{50} values from your data?

2. Is the value you obtained using pyruvate similar to the value you obtained using lactate?

3. Is there anything obviously different about the inhibition with pyruvate as substrate?

4. Based on your data, what are potentially useful concentrations of the inhibitor to use in K_m determination experiments?

5. The data in the table below were obtained by performing K_m determination experiments in the absence of inhibitor and in the presence of two different inhibitors.

5a. What type of inhibitor is Inhibitor 1? Why?

5b. What type of inhibitor is Inhibitor 2? Why?

5c. What are the K_i values for Inhibitor 1 and Inhibitor 2?

	Velocity (μmol/min)		
Substrate (mM)	**No Inhibitor**	**Inhibitor 1 (1 mM)**	**Inhibitor 2 (1 mM)**
0.25	0.06	0.02	0.02
0.5	0.10	0.05	0.04
1	0.17	0.08	0.07
2	0.25	0.14	0.10
4	0.33	0.22	0.13
8	0.40	0.31	0.16

Experiment #3D
LDH Inhibitor Type Determination

As was mentioned in the introduction to Enzyme Inhibition, the best method for determining the inhibitor type is to perform K_m determinations at a variety of inhibitor concentrations.

Experimental procedures:

Make a series of lactate dilutions, as you did for determining the K_m for lactate. Make up enough of each dilution so that you can run several measurements. Collect the data for a velocity *versus* substrate concentration plot for each of at least three separate inhibitor concentrations (note: one of these "inhibitor concentrations" should be "zero"). The inhibitor concentrations you should use should be low enough to allow you to readily detect activity but high enough so that you will clearly observe inhibition. (Reasonable concentrations are those that should yield about 30 to 40% and about 50 to 60% inhibition.)

Study Questions #3E

1. Why do you want to take a time point at 0.25 minutes?

2. Why are you adding lactate to the reaction mixture for the second chemical modification experiment?

3. What is the function of the "stop solution"? How does it work?

4. What is the composition of the reaction mixture? How much of each reagent will you need to add? (In other words, do the dilution calculations you need, so that you will be ready to perform the reaction before coming to class.)

5. Why might arginine modification alter the activity of LDH?

6. Plot the results of the chemical modification experiment shown in the table below. What is the qualitative effect of the lactate? What might account for this?

Time (min)	Activity (ΔA/min)	
	Enzyme alone	Enzyme + Lactate
0	0.50	0.50
4	0.27	0.43
8	0.15	0.37
12	0.08	0.32
16	0.05	0.27
20	0.02	0.24

7. Most chemical modification reactions follow pseudo-first-order kinetics, in which the substrate concentration remaining at any given time is given by the equation:

$$[S] = [S]_0 e^{-kt}$$

where $[S]_0$ is the initial substrate concentration, k is the rate constant, and t is the elapsed time. If, in a chemical modification experiment, you assume that the change in activity is due to a single modification that inactivates the enzyme, the enzyme activity measured corresponds to the amount of unmodified enzyme remaining. The equation above then allows calculation of the rate constant for the reaction. Calculate the rate constants for the data from question 6. Which gives a higher rate constant for the inactivation reaction (the enzyme alone, or the enzyme + lactate)?

Experiment #3E
Chemical Modification of Lactate Dehydrogenase

The vast majority of enzymes are proteins and are therefore composed of amino acid residues. For many enzymes, chemical modification of specific types of amino acid residues markedly alters the activity of the enzyme. This often lends insight into which amino acids are responsible for the catalytic mechanism of the enzyme.

Chemical modification is a relatively blunt instrument, because chemical modifying reagents will react with several residues. For example, lysine-specific modifying reagents will react with several lysines in the protein, and will often also react with the N-terminal amino group. If a lysine-specific modifying reagent alters the activity of the enzyme, it is not always apparent *which* lysine was responsible; it is also possible that one modified residue is merely *close* to an important residue, and that none of the lysines are critical, or that the process required to modify the lysines merely denatures the protein. Molecular biological mutagenesis techniques, which allow directed changes of selected residues are more specific; they are also far more labor-intensive.

Chemical modification has some advantages over other techniques, however. Because it is a chemical reaction, it allows kinetic analysis of the process. In many cases, it is possible to observe alterations in the rate of modification in the presence of substrate. Careful analysis of these data can often point to specific residues as being important in the activity of the protein.

In this experiment you will use butanedione to alter the activity of LDH. Butanedione reacts with arginine side-chains.

Experimental procedure:

Running a chemical modification time-course:

This is a kinetics experiment. This means that you need to perform the steps with attention to the clock. You will find it much easier if you have all of the reagents and tubes prepared before starting the reaction.

1. You will need to make a dilution of your LDH preparation. You want the LDH concentration to be 10-fold higher than the dilution you normally use for enzyme assays.

2. Make up 200 μl of the reaction mixture containing:
    6.25 mM butanedione
    31.25 mM HEPES buffer
    31.25 mM Potassium Borate
    (Use Tris buffer as necessary to obtain the correct concentrations).

3. Set up 7 microfuge tubes containing 30 μl of Stop solution. You will use these tubes for each of the time points you will be obtaining.

4. Initiate the reaction by adding 50 μl of LDH dilution, and incubate the reaction mixture at room temperature. (Note that this will dilute the other reagents in the reaction mixture to final concentrations that are appropriate for the reaction.)

5. Take carefully timed aliquots at 0.25, 4, 8, 12, 16, 20, and 24 minutes. For each time point, remove a 30 μl aliquot of the reaction mixture and add it to a microfuge tube containing 30 μl of Stop solution. Mix the aliquot with the Stop solution by gently pipeting the sample up and down 3 or 4 times.

6. Check each of your time point samples for LDH activity using the standard LDH assay with lactate as the substrate.

Repeat the above experiment using the same volumes of reagents, except for the inclusion of a final concentration of **75 mM lactate in the reaction mixture**.

Does butanedione inactivate LDH in a time-dependent manner? Does lactate alter the inactivation kinetics?

Reagents:
50 mM butanedione in ethanol

200 mM HEPES buffer, pH 7.5

200 mM potassium borate

Ethanol

Stop solution
20 mM HEPES, pH 7.5
20 mM potassium borate
20 mM arginine

Lactate Stock Solution
120 mM lithium lactate
10 mM Tris-HCl (pH 8.8)

NAD^+ Stock Solution
12 mM NAD^+
10 mM Tris-HCl (pH 8.8)

Bicarbonate Stock Solution
18 mM $NaHCO_3$
0.5 M NaCl

Experiment 4
Cloning of Human Lactate Dehydrogenase

Purifying a protein from the tissue of an organism has certain advantages. You know that the protein was synthesized in its natural environment, and you can assume that it was folded correctly, and was subject to any normal post-translational modifications. Purifying proteins from the tissue of an organism has been the traditional method for most of the history of biochemistry.

However, using tissue as a source of protein also has significant drawbacks. The proteins suitable for study are limited to those expressed by readily available sources, and to those expressed in significant quantities. Studying human proteins tends to be somewhat difficult due to limitations in availability of starting material. In addition, the experimental modifications that can be introduced into the protein are limited both in type and in specificity.

As a result, expression of genes in heterologous organisms (especially bacteria) has become a frequently used technique. Bacterial protein expression generally allows proteins to be expressed in very large quantities, allows the researcher to choose the form of the protein to be expressed, and allows the researcher to introduce mutations in the protein sequence to examine the roles of individual amino acids in the function of the protein.

In the previous experiments you purified and characterized LDH from chicken muscle. Although muscle primarily expresses the LDH A gene, the LDH in muscle is a mixture of LDH isoforms. In addition, you used chicken LDH, not because you particularly wished to study the chicken enzyme, but instead because chicken muscle is far more readily available as a source of protein than human muscle.

In this experiment you are going to attempt to clone and express human LDH A in the bacterium *Escherichia coli*. Assuming that the procedure is successful, you will obtain LDH protein which is exclusively composed of the LDH A isozyme.

E. coli is a bacterium found in the intestinal tracts of many species of mammals, including humans. It has been used for a wide variety of molecular biological experiments, and a large number of specialized laboratory *E. coli* strains have been produced. Techniques for manipulating DNA in *E. coli* are well established.

Comments on Molecular Biological Techniques

Molecular biology in its modern meaning (*i.e.* referring to genetic manipulation and analysis techniques) is a fairly new science; nearly all of the techniques used were invented after 1970. This means that techniques are still being invented, and many of the procedures have changed (often dramatically) during the last few years.

Molecular biological techniques differ slightly from biochemical techniques. In molecular biology, exact quantitation of the reagents is of variable importance. In

some cases (as with some of the buffers) concentrations and pH are critically important; in others, such as ligation reactions (see below), the DNA concentration can vary within a factor of ten (or more) and still allow the procedure to be successful.

On the other hand, some aspects of molecular biology require considerable care. ***Humans (and many other organisms) secrete enzymes that degrade DNA and RNA;*** it is therefore necessary to be completely aware of what you are doing. Allowing the reagents to contact your skin will cause problems – not necessarily to you, but definitely to your samples. DNase (the non-specific enzyme that degrades DNA) is rapidly denatured by heating at 68°C; for this reason, it is a good idea to heat treat your DNA samples if there is any chance of DNase contamination. Your plasticware (the pipet tips and microfuge tubes) has been autoclaved (heated to 121°C at elevated pressure) to denature any DNase associated with them; this means that you should not handle the plastic ware unnecessarily (*e.g.*, pour out a few microfuge tubes; do **not** reach into the container with your fingers).

Molecular biological procedures are very commonly used. Some techniques are so commonly used that all of the required reagents are available in kit form. These kits have made many aspects of molecular biology much easier.

The general outline of the procedure you will use is shown below. The procedure is designed to require seven laboratory periods; however, depending on events, the portions of the procedure may be condensed or expanded.

Outline of procedure to be used:

Day	**Probable experiments**
1	Use PCR to amplify LDH coding sequence.
	Purify plasmid to be used to express the LDH.
2	Run agarose gel on PCR reaction and plasmid.
	Purify DNA from PCR reaction.
	Perform restriction digests on plasmid and PCR DNA.
3	Ligate the plasmid and PCR product together.
	Prepare competent cells.
	Transform cells with ligation mixture.
4	Screen colonies for positive clones by PCR.
5	Screen colonies for protein expression.
6	Quantitate expression of LDH.

Study Questions #4A

1. What is an “expression” plasmid? What properties must it contain to be functional?

2. Why is PCR amplification of the coding sequence necessary for cloning of LDH from the cDNA library?

3. How will restriction sites be introduced into the PCR product?

4. During the plasmid preparation procedure, why is it necessary to be gentle with the sample after lysing the cells?

5. What is DNase, and why should you worry about it?

Experiment #4A
Polymerase Chain Reaction and Plasmid Preparation

The first day will involve the use of two separate procedures: polymerase chain reaction and plasmid preparation.

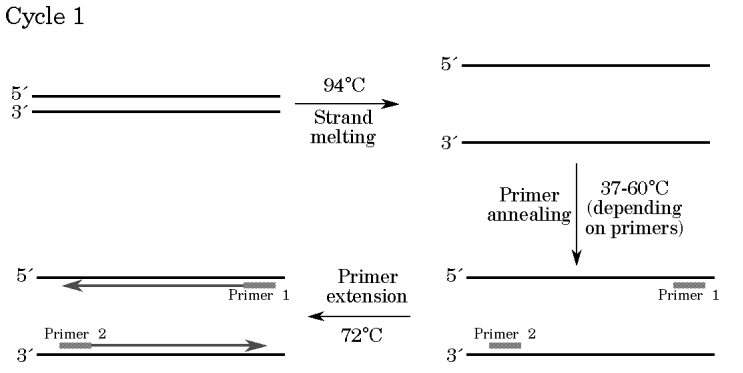
I. Polymerase chain reaction (PCR)

PCR allows the generation of large amounts of a single DNA sequence from a mixture of sequences; the fragment generated can be designed to contain specific starting and ending positions based on the needs of the experiment.

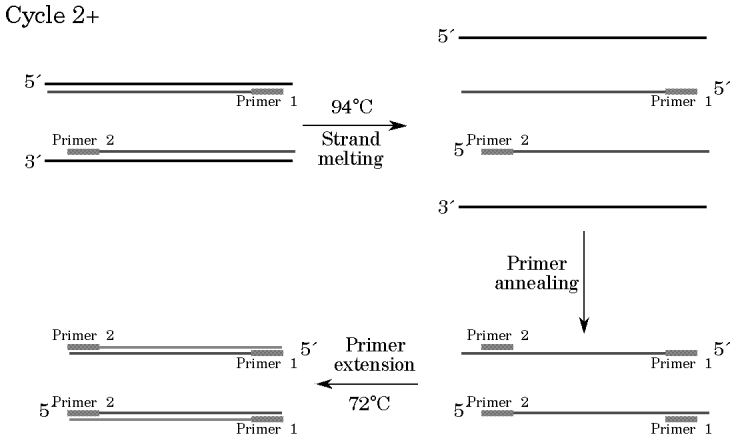
PCR uses a DNA polymerase (an enzyme that synthesizes DNA). DNA polymerases require a "signal" to begin synthesis. The signals you will use are two short fragments of DNA, synthetic oligonucleotides (known as a primer) that correspond to (and thus base-pair to) the ends of the sequence of interest.

If the Cycle 1 reaction (the top of the figure at right) were the entire process, PCR would not be very useful. However, if you examine the process, you will notice that at the end of the first cycle, you have twice as much template DNA as you did at the beginning. Therefore, if you repeat the cycle a number of times, the amount of product DNA will increase geometrically.

In theory, the amount of product will double each cycle; in practice, PCR is not quite that efficient, although it can produce tremendous quantities of DNA. It is literally possible to begin with a single molecule of DNA and generate enough DNA for any molecular biological technique. In addition, the DNA synthesized in the PCR reaction will have specific starting and ending points: the primer sequences define the ends of the fragments.



The two strands of the DNA template are separated by heating (usually to 94°C). The temperature is then decreased to allow the primers to bind to the template DNA. Once the primers have bound, the polymerase is allowed to synthesize new DNA strands (the polymerase most commonly used has a temperature optimum of 72°C.)



Good primers contain approximately 50% G+C and 50% A+T. This reduces problems in inducing template strand separation caused by the high affinity of G for C, and reduces non-specific priming common with high AT content.

In addition, good primers have few regions of complementarity either internal to primer (*i.e.* secondary structure) or between two primers, especially at 3´ end of primer, and avoid repeated sequences (*e.g.*, AAAA or GTGTGT). Primers that are "poor" by these criteria can cause artifacts that prevent amplification of the desired sequence. In some cases, these potential problems are difficult to avoid due to constraints imposed by the sequence of interest; in these cases, the use of longer primers (*i.e.* ≥ 24 bases) may solve specificity problems.

The primer should be long enough to have a reasonable melting temperature (*i.e.* a melting temperature of 55-70°C). Melting temperature depends on a number of variables. In most cases, an approximation [∑(4°C for each G or C) + (2°C for each A or T)] will yield a value close enough to design the PCR experiment.

PCR also allows the generation of mutations at the ends of the fragment, because the primer does not need to be an exact match to the template DNA (it needs to be a sufficiently good match to allow primer binding, but it does not need to be a perfect match). In most cases, the introduction of mismatches at the 5´ end of the primer has little effect. However, mismatches at the 3´ end of the primer may prevent synthesis of the new strand.

The primers you will use for cloning LDH were designed based on the human LDH A sequence downloaded from Genbank, and are:

5´ primer: TCCA<u>cc</u>**ATG**<u>G</u>CAACTCTAAAGGATCAG
3´ primer: CAGA<u>AagCT</u>**TTA**AAATTGCAGCTCC

The lowercase bases in the primers are mismatches (*i.e.* bases that are not present in the actual LDH A sequence). The purpose of the mismatches is to generate specific restriction sites (the underlined sequences) to simplify some of the later steps in the cloning procedure. To avoid the possibility that mutations would destabilize the expressed protein, the mutations were designed in non-coding portions of the sequence. (The **ATG** in bold in the 5´ primer is the LDH A start codon, and the **TTA** in the 3´ primer is complementary to the stop codon.) The 5´ primer is identical (except for the noted mismatches) to the DNA coding strand; the 3´ primer is complementary to the coding strand (*i.e.* is identical to the non-coding strand). The primers were chosen such that the DNA fragment amplified by PCR will contain the entire coding region, with specific restriction sites at each end. The 5´ primer contains an *Nco* I site and the 3´ primer contains a *Hin*dIII site.

In order to amplify the LDH coding sequence, you need a source of the actual coding sequence. Because humans, like most multicellular organisms, have genes with introns, it is rarely possible to use genomic DNA as a starting material for generating the full-length LDH gene. Instead, you will use a human cDNA library, which contains copies of most of the messenger RNAs found in one human tissue.

PCR requires the use of a DNA polymerase to make the copies of the cDNA sequence used as a template. As noted in the figure above, the PCR method involves heating the sample to ~94°C to separate the chains of the double-stranded DNA. In addition, while the oligonucleotide binding and polymerization reactions can occur at a range of temperatures, oligonucleotide binding is much more specific (*i.e.* it is more likely that the oligonucleotide will bind the correct sequence) at higher temperatures. Thermophilic bacteria such as *Thermus aquaticus* prefer to live at temperatures above 70°C, and therefore their proteins (including their DNA polymerases) are stable at elevated temperatures.

Although the *Taq* polymerase is highly thermostable, it does begin to denature at temperatures above 90°C. Its half-life decreases rapidly with increasing temperature. For this reason, melting times of greater than 1 minute for the chain reaction cycles should be avoided. If the thermal cycler melting temperature drifts above 95°C, the enzyme may be inactivated prior to completion of the program.

The *Taq* polymerase has a primer extension rate of 60-100 bases/second under optimum conditions; thus it may be advantageous to use short (1-10 second) extension times, particularly for short products (*i.e.* below 500 bp) to decrease formation of non-specific products. However, for longer fragments (greater than 1000 bases), optimum primer extension rates are rarely achieved, and longer extension times should be used.

Although the *Taq* polymerase is a popular enzyme due to its ability to catalyze primer extension under a wide variety of conditions, it has some drawbacks. Its worst drawback is its relative lack of fidelity; it has a significant error rate, and lacks proofreading functions. As a result, many experiments employ thermostable DNA polymerases that have lower probabilities of misincorporation. Several other thermophile derived polymerases with higher fidelity are in fairly common use, including the *Pyrococcus furious* (*Pfu*) DNA polymerase and the "Vent" polymerase.

PCR conditions, and especially annealing temperature, must be chosen empirically to optimize PCR product formation. Some primer-template combinations work under a wide variety of conditions; others result in product only in a narrow range of conditions. For some reactions the annealing temperature is a critical parameter, with no specific product formed except in a narrow optimum range. It is usually best to begin testing conditions with an annealing temperature of about 55°C, because the use of annealing temperatures above 50°C often prevents certain types of mismatch artifacts. Varying the salt concentration (especially the magnesium concentration may also result in improved yields. However, in most cases, if poor PCR yields are obtained, the problem is either that the template has become degraded, or the primers were poorly chosen.

The temperature profile shown below is a typical one that works for many different primers and templates:

PCR procedure:

Mix reagents in a PCR tube:
16 μl of mixture of 1.25 mM dNTP
10 μl 10x PCR buffer
5 μl 20 μM 5´ LDH primer
5 μl 20 μM 3´ LDH primer
1 μl of DNA template mixture
62.5 μl deionized water
0.5 μl *Taq* polymerase

PCR Temperature Profile

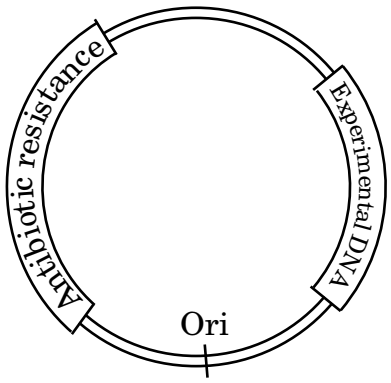
Temp.	Time	Function
94°C	0.5 minutes	Denature DNA
55°C	0.5 minutes	Primer Annealing
72°C	0.75 minutes	Primer Extension

Place the PCR tube in the PCR machine and run the PCR program as shown.

II. Plasmid preparation

PCR is one method for generating large amounts of DNA. A second method is to allow bacteria to replicate the DNA and then purify the replicated DNA from the bacteria. In most cases the DNA of interest for this method is plasmid DNA. A plasmid is a double stranded DNA molecule that will replicate in an organism. A typical plasmid used for molecular biology contains at least four features.

1) The plasmid must be **circular**, because bacteria generally will not replicate linear DNA. 2) The plasmid must contain a sequence that functions as an **origin of replication** (ori). 3) The plasmid must contain a **selection mechanism** that will force the bacteria to retain the DNA; the most common type of selection mechanism used in bacteria is a gene for resistance to an antibiotic such as ampicillin. 4) The plasmid must contain a region for the insertion of the **experimental DNA**. A generic plasmid exhibiting these features is shown at right.



An **expression plasmid** is a specific type of plasmid used to allow expression of heterologous DNA. An expression plasmid must therefore have, in addition to the features listed above, a strong promoter element that will drive transcription of the foreign DNA[14] in the host organism, and an effective ribosome binding site that will allow efficient translation of the transcribed RNA.

Because plasmids are much smaller than chromosomal DNA (for *E. coli*, a typical plasmid contains 5 to 10 kilobase pairs (kb), while *E. coli* chromosomal DNA contains about 4,800 kb), separating the two types of DNA molecule is relatively straightforward. In addition, most plasmids used in molecular biology are "high copy number plasmids"; in other words, each bacterium contains many copies (usually >100) of the plasmid. Therefore, although each plasmid molecule is much smaller than the chromosome, plasmid DNA often comprises ~10% or more of the total DNA in the bacterium.

Plasmid preparation requires several steps:

1) **Growth of bacteria containing the plasmid of interest**. This involves starting a liquid culture using a single clone (each colony on a plate represents a single clone, so cultures are started by picking one colony and adding it to the culture medium). The bacteria are then grown until they reach stationary phase; stationary phase occurs when the bacteria have either used up all of the available nutrients, or when the bacterial waste products have reached levels that preclude further growth. The commonly used laboratory bacterium *E. coli*, reaches stationary phase after 12-18 hours of growth (typically overnight growth is assumed to result in stationary phase). (Note that some other bacterial types, and all yeast and mammalian cells in culture grow more slowly than does *E. coli*.)

[14]Note that "foreign DNA" can be from any organism, including the organism used as a host. Thus, *E. coli* can be used to express *E. coli* proteins from expression plasmids. In most cases, however, the DNA being expressed is from a different organism, and is being expressed in *E. coli* for convenience.

2) **Separation of the bacteria from the culture medium**. The next step has two purposes, concentrating the cells, and transferring them into a buffer that will facilitate plasmid purification. The usual method is to centrifuge the culture and to discard the spent culture medium. The cells can then be frozen for storage (freezing the cells also increases the vulnerability of the cells to the lysis conditions), or resuspended immediately in a buffer for plasmid preparation.

3) **Lysis of the bacteria** (*i.e.* disruption of the bacteria to release plasmid DNA). Bacterial cells are much tougher than human cells, and lysing bacteria requires some effort. The technique usually used to extract plasmid DNA from bacteria involves lysing the cells with a mixture of SDS and NaOH, which disrupts the cell membranes and cell wall. **Once the SDS/NaOH mixture has been added to the cells, it is imperative that the cells be treated gently**; DNA molecules are long and fragile, and vigorous treatment will readily damage the DNA molecules released from the cells. In addition, when performed properly, the alkaline lysis technique does not extract the chromosomal DNA (unless the chromosomal DNA is fragmented by violent shearing forces). The intention in a plasmid prep is to purify the plasmid DNA while obtaining as little chromosomal DNA as possible.

4) **Neutralization of the NaOH**. Once the cells have been lysed, it is a good idea to lower the pH to near neutral. Neutralization is typically performed by adding an acetate buffer. This has two effects: it prevents degradation of the DNA, and it precipitates some proteins and most of the lipids.

5) **Purification of the plasmid DNA**. A variety of different techniques are used for separating plasmid DNA from the other molecules present in the cell. Most of the commonly used techniques involve binding the plasmid DNA to an insoluble material, washing off unbound material (protein, lipids, and small molecules), and then eluting the DNA.

Procedure for Plasmid Miniprep:
The procedure given below should be used as a general guide for the alkaline lysis technique; use the manufacturer's instructions for the exact procedure.

Grow a 3 ml overnight culture of the cells containing the plasmid.
Centrifuge 1.5 ml of the culture in a microfuge tube to pellet the cells; discard the supernatant. Depending on the size of the cell pellet, you may wish to repeat the procedure by adding an additional 1.5 ml from the same culture to the pelleted cells, and recentrifuging.
Resuspend the cells in the Resuspension Buffer solution (the amount to be used varies from 200 to 350 μl depending on the kit). Make sure that the cells are evenly suspended. (The solution should be cloudy with no obvious clumps).
Lyse the cells using a mixture of SDS and NaOH (solution 2). After adding the SDS and NaOH, be gentle with the solution to avoid disrupting the chromosomal DNA. The solution should become fairly clear.
Neutralize the pH using the acetate solution. You should observe the formation of a precipitate. Centrifuge the sample to pellet the precipitate.

The remainder of the technique depends on the exact plasmid prep kit being used; consult the manufacturer's instructions for the recommended procedure.

When the procedure is complete, it is usually a good idea to heat the DNA sample to 68°C for 10 minutes to inactivate any contaminating DNase. After this step, store the plasmid DNA at –20°C.

Required materials:
Human placental cDNA library
LDH 5´ and 3´ primers
PCR thermal cycler
PCR reagents
Autoclaved water
E. coli overnight culture containing pTrc99A plasmid
Plasmid miniprep kit
Autoclaved pipet tips and autoclaved microfuge tubes
Autoclaved PCR tubes
Water bath set for 68°C
Microfuge

Study Questions #4B

1. What is a restriction enzyme?

2. How do you know that the LDH PCR product will be cleaved in only one place by *Nco* I and *Hin*dIII? What would happen if the 3´ primer had been designed to incorporate a *Bam*HI site instead of a *Hin*dIII site? (Hint: you might want to check for the presence of these restriction sites using a computer program.)

3. What is the purpose of the DNA cleanup procedure?

4. What will you use to be able to see the DNA on the agarose gel?

5. Which has a larger molecular weight, the LDH protein or the PCR product that contains the LDH coding sequence? Why is this true? (Hint: this is true for **two** reasons.)

6. What is the expected size of the LDH cDNA PCR product (in base pairs)?

7. Why is the restriction enzyme *Nco* I especially useful for expression plasmids?

Experiment #4B
Agarose Gel Electrophoresis and Restriction Digests

DNA is a charged molecule. Since all DNA sequences have the same phosphate backbone, fragments of all sizes have the same charge-to-mass ratio. As with proteins, it is often useful to run DNA on gels. DNA gels are slightly different from protein gels. Because DNA has an intrinsically constant charge-to-mass ratio, addition of SDS is unnecessary. In addition, DNA fragments are generally much larger than protein molecules; in most cases, DNA is run on agarose gels instead of polyacrylamide gels, because agarose forms a lower density matrix more suited to running larger molecules. Various concentrations of agarose can be used to separate different sized DNA: 0.75% for DNA > 3 kb, 2% or 3% for 50 – 400 bp, and 1% for DNA between these ranges.

Agarose gels allow estimation of DNA fragment size. Running a gel on the PCR product is one way to verify that a fragment of the correct size was produced. It is also useful to run the plasmid DNA on a gel in order to assess both the concentration of the DNA and the quality of the preparation (*i.e.* to look for contaminating nucleic acids or excessive fragmentation of the plasmid DNA).

Agarose gels are usually run in either TBE (Tris-borate-EDTA) or TAE (Tris-acetate-EDTA) buffer. Because the agarose comprises a very small percentage of the gel, while the remainder is the buffer, agarose gels can be run as "submarine gels", in which the gel is merely submerged in the buffer, with the electrical current running through both the buffer and the gel.

Agarose Gel Procedure

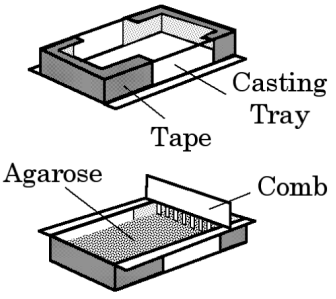
Carefully seal the ends of casting tray. (The diagram at right shows the ends sealed with tape; the apparatus you will probably be using is slightly different, and has its own built-in sealing mechanism.)

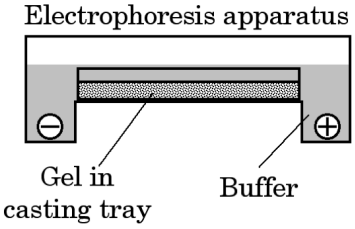
Place the comb in position on the casting tray.

Mix the agarose and TBE at a ratio of 1 g agarose per 100 ml TBE, and heat the solution in microwave oven until the agarose melts and is evenly distributed in the solution.

Add 1 μl of 10 mg/ml ethidium bromide for every 25 ml of melted agarose.

Allow the agarose to cool slightly (to avoid warping the plastic) and pour the agarose onto the casting tray.





After the gel cools (≥20 minutes), remove the sealing tape and the comb. Place the casting tray with the gel into the electrophoresis apparatus, and add 1x TBE to just above the level of the gel.

Load DNA samples containing 10% Loading Buffer. Run the gel at 100 volts (constant voltage).

In the presence of DNA, when excited by ultraviolet light, ethidium bromide fluoresces bright orange. You can therefore place the gel on a UV transilluminator to see the DNA. If you wish to make a permanent record of the gel, take a photograph through a Wratten 22A orange filter that filters out the other wavelengths emitted by the transilluminator.

In order to actually use the DNA generated from the PCR reaction, the DNA fragment needs to be separated from various contaminants. The method you will use involves a modified form of the plasmid preparation procedure.

DNA Clean-up
Perform the procedure according to the manufacturer's instructions. The approximate method is:
Add the PCR reaction mixture to 1 ml DNA clean-up resin and mix. Pipet the mixture into a syringe column, and apply a vacuum to remove the liquid. Add 2 ml of 80% isopropanol, and apply the vacuum again. Centrifuge the column to remove any residual liquid. Place the column in a new microfuge tube, add 50 μl of water at 68°C, and centrifuge again. The microfuge tube will contain the DNA dissolved in the water you added to elute the column.

Restriction Enzyme Cleavage Reactions
Next, both the plasmid and the purified PCR product need to be digested with *Nco* I and *Hin*dIII.

Restriction enzymes are crucial reagents in molecular biology. Restriction enzymes cleave DNA only at specific sequences. If you can control the sequence of the DNA (as you did with the oligonucleotides used for the LDH PCR reaction), you can control the digestion of your DNA. Molecular biology would be extremely difficult without restriction enzymes. Fortunately, different bacteria have evolved a large variety of restriction enzymes, and enzymes with specificity for large numbers of sequences are now commercially available.

You will use two restriction enzymes for the cloning process. One of these, *Hin*dIII, is one of the first restriction enzymes to be discovered; it is widely used because of its reliability. The other, *Nco* I, is also a reliable enzyme; it is most useful, however, because of its sequence specificity. It cleaves the sequence CCATGG. A number of expression plasmids are engineered so that the ATG within the *Nco* I site acts as the start codon for protein synthesis.

Restriction enzymes must cleave both strands of the double stranded DNA. They can do this in a number of ways. Some restriction enzymes cleave both strands at the same location, resulting in a "blunt" end. Other restriction enzymes cleave at different locations on the different strands, leaving short stretches of single stranded DNA. Both *Nco* I and *Hin*dIII leave four bases of single stranded DNA at the 5´-end of their cleaved products. Restriction enzyme products will be discussed in more detail in the next section.

One potential problem with restriction enzymes is that they are most active in specific buffers; the best buffer for one enzyme may not be the best one for another enzyme. In some cases, it is necessary to run one reaction, perform DNA clean-up procedure, and then run the other reaction. Although *Nco* I and *Hin*dIII do prefer different buffers, you can use a compromise buffer to allow both enzymes to cleave the DNA in a single reaction.

Required materials:
Plasmid DNA purified in the previous laboratory session.
PCR product from the previous laboratory session.
Water bath set for 68°C
Water bath set for 37°C
DNA-clean up kit
Microfuge
Autoclaved pipet tips and autoclaved microfuge tubes

*Hin*dIII
Nco I
Restriction enzyme buffer (NEB#2 or equivalent)

Power supply
Agarose gel apparatus
Agarose
DNA molecular weight marker
Agarose gel Loading Buffer:
- 42 mg Bromophenol Blue
- 6.7 g sucrose
- to 10 ml in H_2O

1x TBE:
- 0.09 M Tris-HCl, pH 8.0
- 0.09 M boric acid
- 2 mM EDTA

4x TBE: (4 L)
- 64 ml 0.5 M EDTA
- 174.4 g Tris base
- 89.0 g Boric acid
- QS to 4 L with ddH_2O

Study Questions #4C

1. What is a “compatible end”? Why are compatible ends important in molecular biology?

2. What is ligation?

3. Should you use more plasmid DNA or insert DNA? Why?

4. Why would taking up any random piece of DNA be potentially deleterious to an organism?

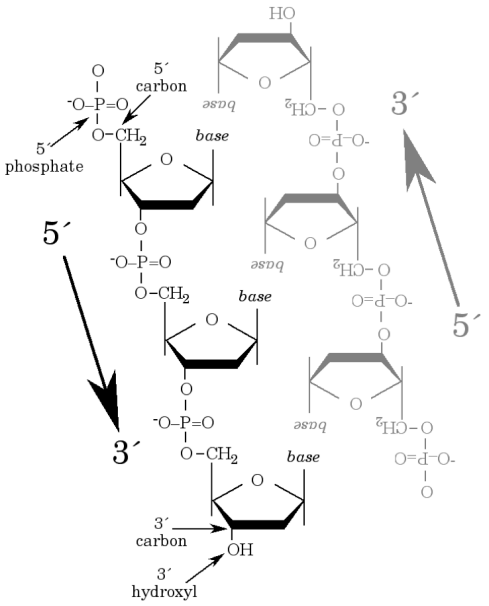
5. How long do you heat-shock your cells during the transformation procedure?

Experiment #4C
Ligation and Transformation

The restriction enzyme digestion of the plasmid and PCR product left 5´ single stranded DNA at the end of each DNA fragment. The next process involves connecting the two pieces of DNA together to generate circular plasmid DNA that contains the LDH cDNA.

The process of connecting two pieces of DNA together is called **ligation**, and is catalyzed by an enzyme called **ligase**. The ligase used most often in molecular biology is derived from the T4 bacteriophage, and uses ATP to supply the energy necessary for the reaction. In addition to ATP, *T4 ligase requires DNA with a 5´-phosphate group and free 3´-hydroxyl groups*.

The drawing at right shows a (very short) region of double stranded DNA. Both strands of the DNA molecule contain 5´-phosphates and free 3´-hydroxyl groups; this DNA molecule is therefore capable of being ligated. This DNA fragment is blunt ended (*i.e.* all of the bases are paired with bases from the opposite strand); note that some restriction enzymes leave blunt ends, while others leave "overhangs", which are short stretches of single stranded DNA at either the 5´ or 3´ end.



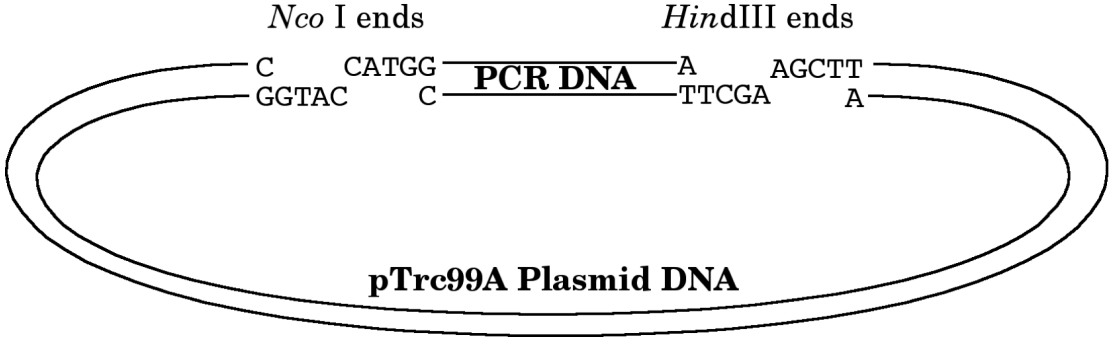
Synthetic oligonucleotides contain free 5´-hydroxyl groups, and therefore must be subjected to phosphorylation prior to ligation. In contrast, most (although not all) restriction enzymes leave 5´-phosphate groups; most restriction fragments can be ligated immediately after digestion.

Ligation also requires compatible ends to the DNA. The drawings below show examples of different types of compatible and incompatible ends (N's imply that any arbitrary sequence could be present).

Compatible Blunt ends		Compatible 4-base, 5´-sticky ends		Compatible 2-base, 3´-sticky ends	
5´–NNNNN	NNNNN	5´–N	CATGN	5´–NNNTG	NNN
3´–NNNNN	NNNNN	3´–NGTAC	N	3´–NNN	ACNNN

Incompatible (blunt/non-blunt ends)		Incompatible 4-base, 5´-overhangs		Incompatible 4-base, 3´-overhangs	
5´–NNNNN	N	5´–N	AGCTN	5´–NCTAG	N
3´–NNNNN	GTACN	3´–NGTAC	N	3´–N	AATTN

When you performed the restriction digests, you used *Nco* I and *Hin*dIII; these enzymes both leave 4-base 5´-overhangs. The end generated by digestion with one of these enzymes is compatible with other ends generated by the **same** enzyme, but not by the ends generated by the other enzyme. You are thus taking advantage of the specificity of the restriction enzymes and of DNA ligase to create a plasmid with the LDH PCR product inserted in the correct orientation.



In setting up a ligation reaction, it is usually desirable to use a molar excess of insert DNA relative to plasmid DNA. The insert DNA will have no effect on the cells; having a molar excess of insert DNA makes it more likely that the ligase will find both plasmid DNA and insert DNA molecule to connect together, and reduces the chance that the incompatible ends of the plasmid DNA will ligate to form circular plasmid DNA. (Ligation of incompatible ends, although very rare, does occur; circularized plasmid DNA will transform cells, and in this experiment, would result in colonies that lack insert.)

Replication of a plasmid requires the insertion of the plasmid DNA into the bacteria used for the replication process. In general, bacteria are reluctant to take up DNA from their environment (at least, they typically will not do so without degrading it first). In order to improve the probability that the cells you will use will actually internalize the plasmid DNA, you must first make the cells "competent" to absorb the plasmid DNA. One procedure that results in efficient competent cells is given below.

Competent cells are significantly more fragile than normal bacteria. Vortexing the cells, heating the cells above 42°C or to 42°C for prolonged periods, or exposure of the cells to any of a number of other even mildly abusive treatments may kill them.

Required Materials:

Ligation buffer
T4 DNA Ligase
Water bath set for 68°C
Water bath set for 42°C
Incubator set for 37°C
LB media (10 to 20 ml)

E. coli competent cells

LB-Ampicillin plates
Sterile microfuge tubes
Cell spreaders
Agarose gel apparatus and solutions

Microfuge
Autoclaved pipet tips and autoclaved microfuge tubes

Procedures:

Ligation procedure:
Mix cleaved plasmid DNA with a ~3- to 10-fold molar excess of cleaved insert DNA. Add ligase buffer (which includes a buffer and the ATP required to support the reaction) and ligase. Incubate at room temperature (actually, T4 ligase prefers 15°C, but for most reactions 20-25°C also allows efficient ligation). After 2 to 24 hours, inactivate the ligase by heating at 68°C for 10 minutes.

Competent Cell Preparation:

TSS: 85% LB medium 10% PEG 8000 5% DMSO 50 mM $MgCl_2$ pH 6.5 *Sterile filter* DMSO may degrade into transformation inhibitors; for best results, fresh DMSO should be stored in frozen aliquots.	Grow an overnight culture of the desired *E. coli* strain under the appropriate conditions (LB + antibiotic, if used). Add 1 ml aliquot of cells to 50 ml LB + 20 mM $MgSO_4$; grow at ~18-20°C (note: growth at higher temperatures results in somewhat lower efficiencies) to $OD_{600} \approx 0.9$ (*i.e.* mid-log phase). Cool the cells on ice for 20 minutes. Spin down the cells (4000 rpm for 5 minutes) in sterile tubes. Resuspend the cells in 1/10 volume TSS (5 ml for 50 ml culture). Cells may be used immediately or frozen for storage. Quick freeze the cells in 310 μl aliquots in a dry ice:ethanol bath and store at -70°C.

Transformation:
The procedure for allowing competent cells to take up DNA is called transformation. Transformation requires mixing a small volume of DNA (usually 1-3 μl) with 100 μl of competent cells. The mixture is incubated on ice for 30 minutes to allow the cells to take up the DNA. The cells are then heat-shocked at 37-42°C for 1 minute, and cold-shocked on ice for 2 minutes. The heat shock/cold shock reverses the effect of the competent cell process, by, in effect, healing the damage inflicted by the competent cell solution. Adding 0.5 ml of culture medium (usually LB) and incubating the cells for 10 - 30 minutes at 37°C assists in this healing process. This cell/medium mixture is then spread onto bacterial culture plates containing agar, LB, and a selection mechanism.

Study Questions #4D

1. What is the difference between screening and selection? Which requires more effort?

2. How can you tell whether a colony has taken up the plasmid DNA?

3. How can you use PCR to screen a colony for the presence of human LDH DNA? What result would you expect if the colony contained the LDH DNA? What result would you expect if the colony did not contain the LDH DNA?

4. Will the PCR screening tell you whether the plasmid will produce LDH protein? Why?

5. The PCR reaction master mix uses the same reagent concentrations as were used in the PCR reaction run in Experiment 5A. Check the calculations given for the master mix in Experiment 5D.

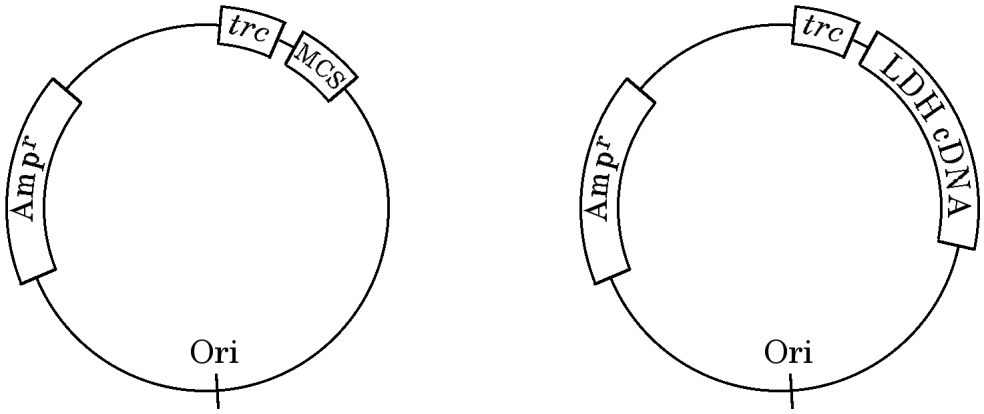
6. How much of each reagent in the master mix would you need in order to screen five colonies while also running a positive control?

Experiment #4D
Selection and Screening

Transformation is usually an extremely inefficient process; in most cases, only a very small fraction of cells actually take up DNA (usually less than 1 in 10^5 cells). If it were necessary to sort out the few cells that took up DNA from the vast majority that did not, cloning experiments would be very difficult. However, selection mechanisms (*i.e.* conditions where cells grow if they express a gene, but die if they do not) make this fairly straightforward.

The selection mechanism most commonly used is antibiotic resistance. If the *E. coli* strain used is not resistant to the antibiotic, but the plasmid DNA contains a gene coding for resistance, then only the cells that have taken up the plasmid will be able to grow. The plasmid you are using contains a gene for resistance to ampicillin; only cells that have taken up the plasmid will be able to grow and form colonies. (Note that, since taking up plasmid DNA is a very rare event, each group of cells visible on a plate represents the offspring of a single cell that internalized a single molecule of plasmid DNA.)

In most cloning procedures, however, only a fraction of the colonies contain the correct plasmid. The selection technique means that only cells containing *some* plasmid DNA have grown; it is still possible for cells to have taken up incorrect plasmids (as long as the plasmid contains an ampicillin resistance gene and origin of replication). The figure below shows the starting plasmid (on the left), and the plasmid that you are attempting to construct (on the right). These are some (although not all) of the possible plasmids that may have resulted in colonies. The next step, therefore, is a screening procedure: an attempt to find cells containing the plasmid of interest.



Screening is more labor intensive than selection; selection techniques only allow cells that *might* be of interest to grow, while screening requires you to actually do something to find the correct plasmid. A large number of screening techniques have been developed; this brief manual cannot cover all of them.

All screening techniques depend on finding a difference between plasmids of interest and other likely plasmids that might be found in cells on the same culture plate. The one you will use takes advantage of the specificity of the PCR technique; in order for any plasmid to be of possible interest to you in this experiment, it must contain the LDH cDNA, and you can therefore use PCR to find *E. coli* colonies that contain human LDH cDNA.

The procedure is to pick the colony with a sterile pipet tip, streak the cells on a new plate (so that, if the colony turns out to be correct, you will still have viable cells to replicate), and then soak the pipet tip in 10 μl of water. This water (actually the few cells that drop off into the water) can then act as a template for a PCR reaction similar to the one you ran to amplify the LDH cDNA initially. Colonies that result in PCR products of the correct size (based on running the samples on an agarose gel) can tentatively be assumed to contain the correct plasmid.

In running the PCR reaction this time, you do not need a 100 μl reaction volume for each sample. A volume this large would waste reagents. Instead, a common method is to prepare a master solution of all of the required reagents except template, followed by adding the same volume of master mix to each PCR tube. The last step should be the addition of the template to the PCR tube (adding the template last reduces the chance of cross-contamination of the PCR samples).

A master mix that allows a final reaction volume of 25 μl and a template volume of 5 μl should be prepared as follows:

Per sample to be run:
4 μl 1.25 mM dNTP mixture
2.5 μl 10x PCR buffer
1.25 μl 20 μM 5´ primer
1.25 μl 20 μM 3´ primer
0.125 μl Thermostable DNA polymerase
11 μl autoclaved water

In addition to the colonies you intend to screen, you should set up a positive control – a reaction that you know **should** work. In this case, the positive control is intended to verify that the reaction mixture was prepared properly. What template do you have that you know should work in the PCR reaction?

Required Materials:
PCR Thermal cycler
PCR reagents
Water bath set for 68°C
Agarose gel apparatus and solutions
LB-Ampicillin plates
Sterile toothpicks
Microfuge
Autoclaved pipet tips and autoclaved microfuge tubes

Study Questions #4E

1. Does running an SDS-PAGE give the same information as a DNA sequencing gel? Why?

2. What would you expect the SDS-PAGE to look like if the bacteria express the human protein? What would the gel look like if the cells do not express the protein?

3. Suggest a method you could use to verify the results of the SDS-PAGE.

4. What positive control do you have that you could run in your proposed method and in the SDS PAGE?

5. How does adding a dideoxynucleotide to a polymerization reaction allow the DNA sequence to be determined?

Experiment #4E
Screening and Sequencing

One potential problem with PCR is that the polymerases commonly used lack proofreading functions. As a result, PCR products often contain errors. In addition, some of the other cloning procedures you used are subject to artifacts. Finally, *E. coli* does not always overproduce the products from foreign genes inserted into expression vectors (the protein may be toxic to the cells, or otherwise unstable). For all of these reasons, the next step is to test the putative positive colonies for human LDH overexpression.

The simple method for doing this is to induce protein expression in a small culture, and then run an aliquot of the cells on an SDS polyacrylamide gel. Overexpressed proteins typically comprise a significant percentage of the total cell protein (5-50%), and therefore are readily visible on a Coomassie stained gel.

The plasmid you are attempting to construct should produce the human LDH gene product under control of the *trc* promoter. The *trc* promoter is derived from the *lac* promoter; the *lac* repressor, a normal *E. coli* protein, binds to the *lac* and *trc* (and related) promoters, and prevents gene transcription. The presence of the *lac* repressor means that protein expression is negligible unless expression is induced by addition of lactose (the normal ligand for the *lac* repressor) or IPTG (a non-metabolizable lactose analog) to the culture.

You will use small (3 ml) overnight cultures treated with IPTG to assay the apparent positive colonies for protein expression by SDS PAGE.

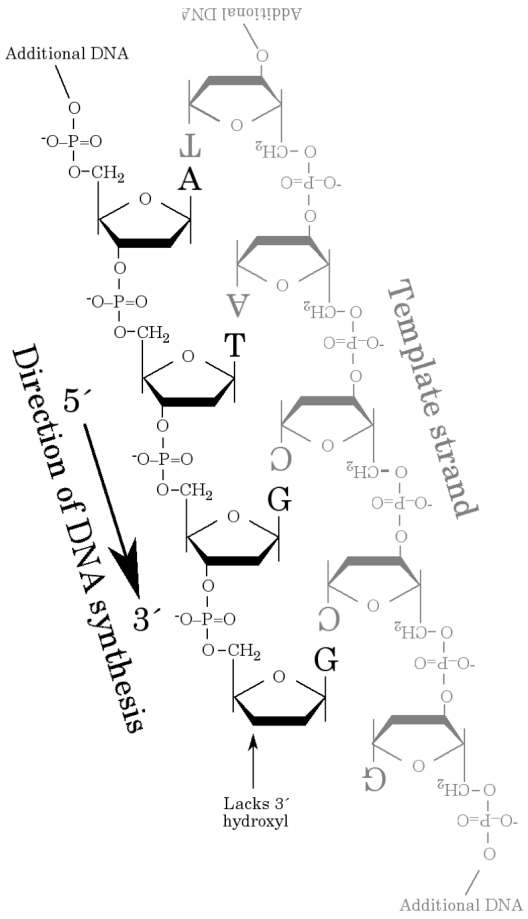
A second method for determining whether the plasmid contains the correct DNA without errors is to determine the sequence of the DNA.

DNA sequencing is a well established procedure; it is far easier to sequence DNA than to sequence either RNA or proteins. DNA sequencing involves a modification of a normal cellular DNA replication process. As with most DNA synthetic reactions, sequencing reactions require an oligonucleotide to use as a primer, a template to use for synthesizing a complementary strand, and dNTPs to use as substrates for synthesis. In addition, the reaction requires a way of identifying each base.

The commonly used methods for DNA sequencing all involve the use of “chain terminators”, which are modified nucleotides that can be incorporated into the new DNA strand, but do not permit continued synthesis of DNA. The chain-terminators lack a 3´-hydroxyl group (in the example on the next page, dideoxyGTP was inserted instead of dGTP, terminating DNA synthesis). If the sequencing reaction is set up such that the identity of the chain terminator nucleotide added is known, it is possible to identify the nucleotide at each position.

One method, being used more and more frequently, is the use of fluorescently labeled chain terminators. If each type (*i.e.* A, C, G, and T) of dideoxynucleotide has

a different fluorescent label, the sequence can be determined very readily using an automated system. Variations on this method are being used in the genome sequencing projects currently in progress. This method is also used by sequencing services (you send your DNA sample to a lab, and they tell you what the sequence is without you having to actually do the work). The fluorescent-tag system, however, requires expensive equipment.

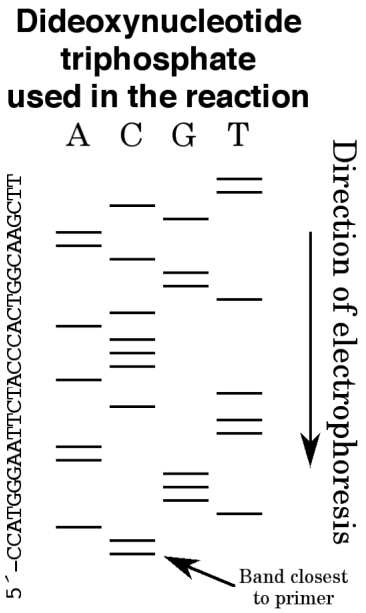


Laboratories that perform small scale sequencing usually use a somewhat older method in which four separate reactions are run, one for each of the possible chain terminator nucleotides; in each case, the newly synthesized DNA is radioactively labeled so that it can be detected (usually by exposing the gel to film). After the reaction is complete, the DNA is run on a polyacrylamide gel; sequencing gels can separate DNA fragments that differ in size by a single nucleotide.

A representation of a sequencing gel autoradiograph (*i.e.* the developed film after incubation of the gel with the film) is shown below. Four reactions were run using the same template DNA and all four dNTPs, but with each reaction containing only one dideoxyNTP. The gel separated the DNA fragments based on their size; as usual, the smallest bands are at the bottom. The sequence can be read directly from the gel; simply start at the bottom, and for each band, note the lane in which the

band occurred (the sequence derived from this "experiment" is shown on the left side of the figure).

The semi-automated sequencing machines use a generally similar method; the difference is that, instead of stopping the gel part-way through the separation (as is shown here), the bands are detected (based on the fluorescent tags on the ddNTPs) as each one runs off the end of the gel.



Depending on how well the reaction runs, on the homogeneity of the gel, on the quality of the template DNA, and on the exact method used, a single sequencing reaction can yield anywhere from 100 to 1000 bases of information. A cDNA of 1000 bp (such as the LDH cDNA you used) would usually require 2 to 4 reactions (using different primer positions) to read the entire sequence. The human genome (3×10^9 bp of unique sequence) would require at least 3×10^6 sequencing reactions; in practice, it is going to require many more sequencing reactions than that, in order to sequence each part of the genome more than once.

The "gel" shown above is an idealized example. Some sequencing gels actually look similar to this; most, however, have artifacts of various types, which make reading the sequence somewhat more challenging. Artifacts can be due to inhomogeneities in the gel (if some lanes run faster than others, it can be difficult to read the sequence). If some copies of the DNA template have holes (or other problems that force the polymerase to stop synthesis), the gel may show bands in all of the lanes (making it difficult to decide which is the correct base). In addition, the sequence itself may cause problems: the GC base-pair binds more tightly than the AT base-pair; sequences with high GC content are especially subject to artifacts where the polymerase has trouble synthesizing the new strand, or where the DNA forms secondary structure while running on the gel which tends to alter its speed of

migration through the gel. While experienced researchers can usually compensate for these problems, determining the sequence of unknown DNA samples can be quite difficult. Verifying a known sequence (as you would be doing for the LDH) is usually much less difficult, but can still be subject to errors.

There is a chance that you may run a sequencing gel in the laboratory. However, due to the cost of the sequencing reagents, it is more likely that the general procedure will be discussed and demonstrated, but that you will not actually run a gel

Required Materials:
IPTG-treated overnight culture of *E. coli* containing apparent positive colonies.
SDS PAGE apparatus
Microfuge
1.5 ml microfuge tubes

Study Questions #4F

1. Is it possible for the LDH cDNA to contain mutations and still be active? Why?

2. Can you use the protein gel to rule out certain types of mutations? Which ones?

3. Why do you need to measure the lysozyme concentration?

4. During the plasmid prep procedure you lysed *E. coli*. Why can't you use the same method to lyse the cells before the LDH assay?

5. Do you expect to observe higher or lower specific activity in your *E. coli* cell lysates than in the chicken crude homogenate? Why?

Experiment #4F
Activity Measurements

The detection of protein expression is a very good sign that you have produced the plasmid you were intending to construct. However, it is still possible that the protein being expressed is inactive, either as a result of mutations, or as a result of an inability of *E. coli* to properly fold the protein.

You can determine whether the protein is folded by running an assay for LDH activity. (You should be quite familiar with the LDH enzyme assay by now.) In order to run an enzyme assay, it will be necessary to lyse the *E. coli* to release the LDH.

In addition, it is often useful to determine the amount of active protein expressed; specific activity is a good measure of the amount of LDH. In order to determine specific activity, you need to measure both the enzyme activity (be sure to correct for background activity using parent *E. coli*, since *E. coli* has its own LDH), and the protein concentration in your preparation (using the same Bradford assay you used to measure protein concentration in your chicken samples a few weeks ago).

Cell lysis procedure:

1. Pellet the cells in a microfuge tube: add 1.5 ml culture, centrifuge for 15 seconds at maximum speed, discard the supernatant, and then repeat the procedure. This will result in the cells from 3 ml of culture in a single tube.

2. Resuspend the cells in 1 ml Lysis buffer.

3. Add 50 μl lysozyme, and mix gently.

4. After 10-20 minutes at room temperature, microfuge for 10 minutes to pellet cell debris.

5. Use the supernatant to test for LDH activity.

6. Measure the protein concentration in the supernatant and in the lysozyme solution; correct the protein concentration in your samples for the lysozyme added.

Required materials:

Lysis Buffer
50 mM Tris-HCl, pH 8.0
10 mM EDTA
2 mM Dithiothreitol

Lysozyme
50 mg/ml lysozyme in Lysis buffer

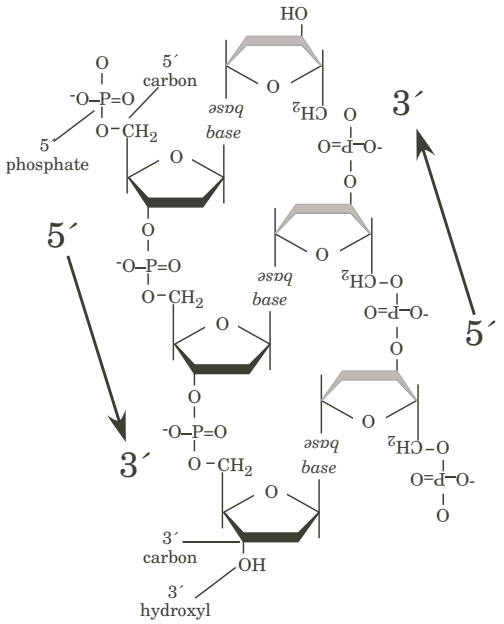
LDH assay reagents
Bradford assay reagents.
Microfuge

1.5 ml microfuge tubes

Definitions

3´ (“three prime”): the 3-carbon of the second ring in a structure. In molecular biology, the ribose ring of the nucleotide is considered to be the second ring. The 3´-end of a DNA strand is the site of new synthesis by DNA polymerases.

5´ (“five prime”): the 5-carbon of the second ring in a structure. In molecular biology, the ribose ring of the nucleotide is considered to be the second ring. Because DNA synthesis and protein translation occur 5´ to 3´, the 5´ end is usually considered to be the starting position; a feature that is 5´ relative to a particular site is considered to be upstream, while a feature that is 3´ is considered to be downstream.



Affinity chromatography: a technique for separating a protein from a mixture on the basis of a property specific to that particular protein. For example, one substrate for lactate dehydrogenase is NADH; a column with NADH or a structurally related molecule may bind lactate dehydrogenase (and probably other dehydrogenases) with high affinity, while not binding the vast majority of other proteins.

Amino acid: strictly, any organic compound containing an acidic function and an amino group; in biochemistry, this term is often used to refer to any of the nineteen amino acid and one imino acid compounds typically used in biological protein synthesis.

Ammonium sulfate: a salt, $(NH_4)_2SO_4$, which has the property of reducing the solubility of proteins in the same solution, usually without causing structural alterations in the protein. Because different proteins exhibit differential solubility in ammonium sulfate, this salt can be used to separate proteins based on gross physical differences. Ammonium sulfate concentration is often given in percent. This “percent” is slightly unusual, in that it refers to “percent of saturation”, rather than grams of solid per 100 ml of solution. The percent ammonium sulfate varies somewhat with temperature.

Anion exchange chromatography: a type of ion exchange chromatography in which the resin is derivatized using positively charged compounds such as DEAE or quaternary ethyl amino groups. The positively charged resin then allows the exchange of anions: it exchanges negatively charged proteins with counterions from the buffer.

Antibiotic: a compound that either inhibits growth of, or is toxic to, bacteria, even when used systemically. Thus, compounds such as penicillin, which can be used

to treat bacterial infections, are considered to be antibiotics, while compounds such as ethanol (which kills bacteria when in direct contact but not if taken systemically) are not considered to be antibiotics.

Antibiotic resistance: the ability to grow in the presence of an antibiotic. In most cases this trait is the result of a gene coding for an enzyme that degrades the antibiotic (for example, the gene for β-lactamase confers resistance to β-lactam antibiotics such as penicillin and ampicillin, because the enzyme inactivates the antibiotic by cleaving part of its structure).

Bacteriophage: a virus that infects bacteria (also called simply "phage"). Bacteriophages, like all viruses, take over cellular machinery as part of their replication process. Engineered bacteriophages are useful for propagating DNA for a number of molecular biological processes.

Base-pair: *n*. a set of nucleotides, one from each strand of a double-stranded nucleic acid, that form a hydrogen-bonded complex with one another. *v*. to form a series of such sets of nucleotides. The average molecular weight of a base pair is about 650 Da (assuming sodium as the counter ion).

β-ME (β-mercaptoethanol): a commonly used reducing agent. Mercaptoethanol and DTT are used to maintain cysteine residues in the free sulfhydryl form.

Cation exchange chromatography: a type of ion exchange chromatography in which the resin is derivatized using negatively charged compounds such as carboxymethyl groups. The negatively charged resin then allows the exchange of anions: it exchanges positively charged proteins with counterions from the buffer.

cDNA: a DNA sequence complementary to another nucleic acid sequence. The term cDNA is usually used to refer to DNA generated by reverse transcribing an mRNA. As such, cDNAs represent actively transcribed genomic DNA but do not contain introns.

cDNA library: a mixture of cDNA fragments comprised of copies of most of the mRNAs expressed within the source tissue. In most cases, the cDNAs that comprise the library are inserted into a modified form of the *E. coli* bacteriophage λ, which allows both the propagation of the DNA and, for some cloning techniques, the isolation of individual cDNAs.

Chromatography: any of a number of techniques in which molecules can be separated as a result of partitioning between the stationary phase and the mobile phase. In protein separations, the mobile phase is an aqueous buffer, while most stationary phases are synthesized by cross-linking carbohydrate polymers, possibly followed by adding functional groups (such as positively charged groups used in anion exchange chromatography).

Chromophore: a chemical functional group within a molecule that absorbs electromagnetic radiation. While the term chromophore applies to groups that aborb radiation of any wavelength, it applies especially to groups that absorb

within the visible portion of the spectrum, because these groups add color ("chromo" is derived from the Greek word for color) to a molecule.

Codon: a sequence of three bases that can be translated into an amino acid. In order to be considered a codon, the DNA (or RNA) sequence must be part of a coding sequence, and must be in the correct reading frame.

Cohesive end: the segment of single-stranded DNA extending 5´ or 3´ from a double stranded DNA fragment resulting from digestion by a restriction enzyme, which is capable of base-pairing to a compatible end of another DNA fragment (or the opposite end of the same fragment). Cohesive ends are typically referred to as "sticky ends" except in formal writing.

Column: a cylindrical apparatus containing chromatography resin that is used for a chromatographic process. Columns typically have an inlet, which allows loading of samples and addition of running buffer, and an outlet, which allows collection of the material that is not bound to the column.

Compatible ends: termini of linear DNA fragments that are capable of being ligated. Compatible ends can be blunt, or can be comprised of sticky ends, where the protruding single stranded DNA sequence of one end can base-pair to the other. (Note: it is possible for the ends of a single DNA fragment to be compatible, in which case, the fragment will tend to circularize if ligated.)

Competent cells: bacteria treated with a solution that greatly increases their likelihood of taking up DNA from their surroundings. Competent cells are significantly more fragile than normal bacteria, and are easily killed by violent treatment.

Complementary: in molecular biology, having a sequence that will base-pair to a sequence of interest. The sequence 5´-GGACTG is complementary to the sequence 5´-CAGTCC.

DEAE: diethyl-aminoethyl, a positively charged functional group frequently attached to resins used for anion exchange chromatography.

Deoxynucleotide: a compound containing a purine or pyrimidine base attached to ribose phosphate, in which the ribose is missing one of the hydroxyl groups normally present. Unless specified, the hydroxyl is missing from the 2´-position. Deoxynucleotides are the monomer units for DNA.

Dideoxynucleotide: a modified deoxynucleotide, in which both the 2´- and 3´-hydroxyl groups are missing. Dideoxynucleotides are used as chain terminators for DNA sequencing. DNA polymerases normally add the next nucleotide to the 3´-hydroxyl of the previous nucleotide; if the previous nucleotide lacks a 3´-hydroxyl, adding another nucleotide is impossible.

DNA (deoxyribonucleic acid): the genetic material of some viruses and all known non-viral organisms. DNA is a deoxyribonucleotide polymer comprised of four types of bases (adenine (A), cytosine (C), guanine (G), and thymine (T)). The

specific base sequence, in the presence of cellular structures, determines the role of the DNA (*i.e*. coding regions, non-coding control regions, regions of other, less well defined functions, or regions with no known function).

DNase: any of a number of enzymes capable of hydrolyzing DNA into small fragments. Unlike restriction enzymes, DNase exhibits limited sequence specificity, and will cleave most DNA strands into smaller fragments. Humans secrete DNase; it is therefore necessary to avoid contact between human skin and any valuable DNA samples. DNase is rapidly inactivated by heating to 68°C.

DTT (dithiothreitol): a commonly used reducing agent. Mercaptoethanol and DTT are used to maintain the cysteine residues in the free sulfhydryl form, although DTT is somewhat more effective and somewhat more stable in aqueous solution.

EDTA (ethylenediamine tetraacetic acid): a chelating agent used in many buffers to sequester metal ions that may affect biochemical systems. EDTA inhibits calcium-dependent proteases by reducing the free calcium concentration.

Elution: the process of allowing protein to dissociate from a column resin. Elution usually involves altering the running buffer to decrease the strength of the interaction between the protein and the resin.

Exon: a DNA sequence that becomes part of the mature mRNA. Exons may include both coding and non-coding sequences.

Expression: in molecular biology, synthesis of RNA or (usually) protein from a DNA coding sequence.

Expression vector: a plasmid designed for expression of foreign proteins in a particular host cell (often *E. coli*). In addition to the normal features of a plasmid, expression vectors contain a strong promoter and ribosome-binding site upstream of unique restriction sites intended to allow the insertion of foreign DNA.

Extinction coefficient (ε): the Beer-Lambert law ($A = \varepsilon cl$) proportionality constant that relates absorbance to concentration for a given molecule at a given wavelength in a cuvette of a given pathlength. The extinction coefficient is dependent on the probability that the molecule will absorb light at the applicable wavelength.

Fold purification: the ratio of the specific activity in a partially purified sample to that of the initial sample.

Frame: short for “Reading frame” (see below).

Gel filtration chromatography: a technique for separating molecules on the basis of size. Gel filtration resins contain small pores; small molecules enter the pores, while larger molecules cannot. Thus, large molecules experience a smaller total volume in the column, and elute first, followed by other molecules in order of

decreasing size. Gel filtration chromatography can be used in purification techniques, or can be used analytically to measure the apparent molecular weight of molecules and folded proteins in solution. (Note: this technique is also called gel permeation or size exclusion chromatography.)

Genetic code: the algorithm that cells use to translate nucleic sequences into protein sequences. Exon sequences are essentially a substitution code; each group of three bases (*i.e.* each codon) defines an amino acid. Since organisms require only 20 amino acids and a stop signal, while the code includes 64 possible codons (4^3), the code contains some redundancy (for example, the genetic code contains 6 codons for the amino acid serine).

Genetic Code

First Position	Second Position T		C		A		G		Third Position
	TTT	Phe	*TCT*	*Ser*	TAT	Tyr	TGT	Cys	**T**
T	*TTC*	*Phe*	*TCC*	*Ser*	*TAC*	*Tyr*	TGC	Cys	**C**
	TTA	Leu	TCA	Ser	TAA	Stop	TGA	Stop	**A**
	TTG	Leu	TCG	Ser	TAG	Stop	TGG	Trp	**G**
	CTT	Leu	**CCT**	**Pro**	CAT	His	CGT	Arg	**T**
C	CTC	Leu	**CCC**	**Pro**	*CAC*	*His*	*CGC*	*Arg*	**C**
	CTA	**Leu**	CCA	Pro	CAA	Gln	**CGA**	**Arg**	**A**
	CTG	*Leu*	CCG	Pro	*CAG*	*Gln*	**CGG**	**Arg**	**G**
	ATT	Ile	*ACT*	*Thr*	AAT	Asn	AGT	Ser	**T**
A	*ATC*	*Ile*	*ACC*	*Thr*	*AAC*	*Asn*	AGC	Ser	**C**
	ATA	**Ile**	**ACA**	**Thr**	*AAA*	*Lys*	**AGA**	**Arg**	**A**
	ATG	Met	**ACG**	**Thr**	AAG	Lys	**AGG**	**Arg**	**G**
	GTT	*Val*	*GCT*	*Ala*	GAT	Asp	*GGT*	*Gly*	**T**
G	GTC	Val	GCC	Ala	GAC	Asp	*GGC*	*Gly*	**C**
	GTA	*Val*	*GCA*	*Ala*	*GAA*	*Glu*	**GGA**	**Gly**	**A**
	GTG	Val	GCG	Ala	GAG	Glu	**GGG**	**Gly**	**G**

Italics indicates preferred tRNA in *E. coli*. **Bold** indicates minor tRNA in *E. coli*.

Adapted from: "Biased Codon Usage: An exploration of its role in optimization of translation." In: *Maximizing Gene Expression*, pp. 225-285, 1986.

Although the control elements (such as promoters) may differ markedly between species, nearly all organisms translate nucleic acid sequences into proteins using the same code, and therefore, foreign DNA expressed in an organism nearly always results in the same protein sequence as is found in the parent organism. However, all organisms do not use all codons with the same frequency. Some prokaryotes use tRNA availability as one method for regulating protein

synthesis rates. Some foreign proteins are poorly expressed in *E. coli* due to large numbers of rare codons (*i.e.* having corresponding tRNAs that are produced in relatively small amounts). A rare codon frequency of 15% or less usually results in high expression, unless the rare codons are in close proximity to one another in the coding sequence.

Genotype: the genetic makeup of an organism. In most cases, the genotype of a given organism is assumed to be identical to that of the wild-type unless explicitly stated to be mutated. A large number of *E. coli* strains with known mutations are available; many of these mutations result in characteristics useful for various molecular biological applications.

Gradient: in column chromatography, a gradual change in the concentration of some component of the running buffer. Gradients can be smooth (typically linear, unless computer controlled pumps are used) or "step". A smooth gradient involves a constantly changing concentration of the component(s) of the running buffer. A step gradient uses a constant concentration of the component, followed one or more times by the use of a new solution containing a different concentration the component.

HPLC (high performance liquid chromatography): a form of chromatographic technique that uses sophisticated high-pressure pumps to move the liquid phase through the column. HPLC pumps are capable of generating pressure of 50 megaPascals or more (over 7000 pounds per square inch). HPLC columns are designed to withstand high pressures, although most HPLC columns will be damaged by the maximum pressure output of the pump.

Hydrophobic interaction chromatography: a technique for separating molecules on the basis of their ability to interact with hydrophobic functional groups covalently attached to a resin.

Incubation: storage under defined conditions, especially at a controlled temperature.

Induction: in molecular biology, to cause initiation of transcription of a gene by some external intervention. For example, IPTG is used to induce expression of genes under control of several *lac*-derived promoters used in expression vectors.

Ion exchange chromatography: a technique for separating molecules on the basis of charge. Ion exchange resins contain charged groups; proteins containing amino acid residues of opposite charge will bind to the charged groups. Raising the ionic strength (usually by raising the salt concentration) of the running buffer causes proteins to elute from the column.

Ionic strength: a measure of the total amount of charged species present in a solution. Mathematically, $\text{ionic strength} = \frac{1}{2}\sum_i C_i Z_i^2$, where C_i is the concentration of the i^{th} species present, and Z_i is the charge on that species.

IPTG (isopropylthio-β-D-galactoside): a non-hydrolyzable carbohydrate derivative. IPTG binds to the *lac* repressor, causing its dissociation from *lac* promoter elements, and consequent activation of transcription from the promoter. Most promoter elements used to drive expression of engineered genes are derived from the *lac* promoter; in most strains of *E. coli*, transcription of the engineered gene only occurs in the presence of lactose or synthetic analogs such as IPTG.

Intron: a DNA sequence that is transcribed and then removed during mRNA maturation. Mammalian genes frequently contain hundreds of kilobases of DNA; after the introns are removed the residual mRNA is usually less than ten kilobases long. Because mammalian DNA contains introns, and because prokaryotes lack the ability to remove introns, mammalian genomic DNA cannot be used directly as a source of genetic material for expression of mammalian proteins in bacteria. Instead, it is necessary to use cDNA (*i.e.* reverse transcribed mature mRNA) as the source of genetic material.

K-12: a wild-type strain of *E. coli*, from which many laboratory *E. coli* strains have been derived.

Kinase: an enzyme that phosphorylates its substrate, generally using ATP as the phosphate donor. Most kinases are specific for certain types of substrates (for example, Protein Kinase C phosphorylates specific proteins on specific serine and threonine residues). Polynucleotide kinase phosphorylates free 5´-hydroxyl groups of DNA, and is therefore frequently used to prepare DNA lacking a 5´-phosphate for ligation reactions.

Lag-phase: the period during which bacteria grow slowly after being taken from an environment in which nutrients are limiting to an environment in which nutrients are plentiful (for normal *E. coli*, this period is usually 1-2 hours).

LB (Luria-Bertani broth): a commonly used “rich” bacterial growth medium. Rich media contain all of the nutrients required for cell growth. (This is in contrast to minimal media, which contain only a few minerals, a carbon source, and a nitrogen source, and therefore require the cells to synthesize all of the required metabolites.)

LDH (lactate dehydrogenase): a ubiquitous nicotinamide coenzyme-dependent oxidoreductase that interconverts pyruvate and lactate. LDH is expressed in relatively large amounts in some tissues, and is an easily purified, stable protein.

Ligase: an enzyme that catalyzes the formation of a covalent bond between a 5´-phosphorylated and a 3´-free-hydroxyl end of a DNA strand. The function of a ligase is thus to connect two DNA fragments. The bacteriophage T4-derived ligase commonly used in molecular biology uses ATP as a co-substrate. Ligase is inhibited by a number of impurities that may be present in DNA samples.

Ligation: in molecular biology, a reaction in which two fragments of (usually) double-stranded DNA are connected together.

Load: in column chromatography, the process of allowing a protein-containing solution to enter a column, usually under conditions in which the protein of interest would be expected to bind to the resin. In gel electrophoresis, the process of placing a sample within a well on the gel prior to the application of a potential gradient to separate the components of the sample.

Log-phase: the process of rapidly growing in an environment rich in nutrients. Bacteria in log-phase divide every 20-40 minutes (in contrast, human cells divide roughly every 24 hours). During log-phase growth, bacteria express a group of genes somewhat different from those expressed during stationary phase.

MCS (multiple cloning site): a region of a plasmid containing a number of unique restriction sites that is intended as the insertion site for foreign DNA. In expression vectors, the MCS is located in close proximity to the promoter and other signal sequences that drive transcription and translation of the inserted gene.

Mismatch: in molecular biology, a base, or series of bases, which do not form base pairs with the corresponding bases on the opposite strand. The presence of mismatches implies that one strand has undergone a mutation. Synthetic oligonucleotides frequently contain a few mismatches to create desired mutations (usually to create restriction sites, or to create modified protein sequences); note that the presence of too many mismatches will prevent the oligonucleotide from binding to its intended complementary strand.

mRNA (messenger RNA): a single-stranded RNA that contains a protein coding sequence and 5´- and 3´-untranslated regions which contain control elements. Most mRNAs in multicellular organisms contain a poly-A tail (astretch of multiple adenosine residues at the 3´-end), which allows the isolation of mRNA; this is not true in bacteria, and therefore bacterial mRNA usually cannot be separated from other bacterial RNAs.

Nucleotide: a compound containing a purine or pyrimidine base attached to ribose phosphate. Nucleotides are the monomer units for RNA; nucleotides are also used for metabolic reactions and for intracellular signaling.

Oligonucleotide: a short, usually single-stranded sequence of (usually) DNA. Most oligonucleotides are synthesized using chemical methods; oligonucleotides are also the product of extensive digestion of nucleic acids with hydrolytic enzymes.

Open reading frame (ORF): a sequence of DNA that begins with ATG and ends with an in-frame stop codon, often used to refer to DNA sequences not certainly identified as being actively expressed (*i.e.* not known to be genes). In most cases, to be identified as an ORF the region must be long enough (>200 bases) to code for a peptide of reasonable size (few coding regions smaller than ~200 bases have been observed).

Ori (origin of replication): a DNA sequence in a plasmid required for replication of DNA in bacteria.

Palindrome: a sequence of characters that reads identically in both forward and reverse direction. In molecular biology, a palindromic sequence is one in which the one strand has the same sequence as the complementary strand. For example, AAGCTT is a palindrome (the complementary strand also reads 5´-AAGCTT). Most (although not all) restriction enzyme recognition sequences are palindromic (AAGCTT is a *Hin*dIII site).

Phenyl sepharose: a hydrophobic interaction chromatography resin, in which phenyl groups are covalently attached to sepharose (a cross-linked carbohydrate derivative).

Plasmid: a circular double-stranded non-chromosomal DNA molecule that bacteria will replicate. Most plasmids contain a gene for antibiotic resistance, an origin of replication, and one or more genes of interest to the researcher.

Plasmid prep: a procedure for purifying plasmid DNA from bacteria. In most cases, the cells are lysed with detergent (usually SDS) and high pH, followed by precipitation of chromosomal DNA; the procedure then uses one of a variety of methods of separating the plasmid DNA from residual contaminants (soluble proteins, carbohydrates, lipids, and other molecules).

PMSF (phenylmethylsulfonyl fluoride): a commonly used protease inhibitor. PMSF is an irreversible inhibitor of serine proteases. It has limited solubility and limited stability in aqueous solutions, and is toxic; alternative inhibitors of serine proteases have been developed that lack these drawbacks, but that are considerably more expensive.

Polymerase chain reaction (PCR): a technique for producing large amounts of a specific DNA fragment from a small amount of mixed DNA sequences. Briefly, a sample of DNA is heated to ~94°C (to separate the strands), the temperature is lowered to 37-60°C to allow binding of specific oligonucleotides to the ends of the sequence of interest, and a polymerase is used to replicate the sequence 3´ to the oligonucleotides; this procedure is then repeated 20-50 times. In most cases, the polymerase used is heat-stable (the polymerase is usually derived from a bacterium that prefers living at 70°-100°C), and therefore the polymerase only needs to be added at the beginning of the process. The “chain reaction” occurs because each cycle results in an increased amount (roughly a doubling) of DNA that can act as a template for further DNA synthesis.

Primer: a short oligonucleotide sequence complementary to a sequence of interest. Most DNA polymerases cannot begin synthesizing nucleic acids without a template and at least a short region of double stranded nucleic acid to act as a starting place. Primers act as the necessary starting place for nucleic acid synthesis for a variety of molecular biological techniques, including PCR and DNA sequencing.

Promoter: a DNA sequence recognized by the transcription machinery (*i.e.* proteins involved in the synthesis of RNA from DNA). Promoters act as signals for initiation of RNA synthesis. Expression vectors typically contain strong

promoters, such as the *trc* promoter, that are used to initiate mRNA synthesis using the inserted foreign gene as a template.

Reading frame: each codon has three bases; if the sequence is read beginning with one base, the translated protein will have a difference sequence from a protein translated beginning with the following base. For example, the sequence ATGTGGTAA codes for Met-Trp-Stop if read from the first base, x-Cys-Gly-x (the "x" refer to the partial codons) if read beginning with the second base, and x-Val-Val-x) if read beginning with the third base. In this example, the TAA stop codon is in-frame with the methionine codon, but not with the cysteine or valine codons in the other reading frames.

Recombinant DNA: genetic material that has been engineered in some fashion. Most commonly, the term recombinant DNA refers to coding sequences taken from one organism and placed in another organism to allow expression of the foreign gene in the new environment.

Replication: the process of synthesizing a new DNA strand using the preexisting strand as a template. In normal cells, the result of replication is a doubling of the total amount of DNA, and only occurs immediately prior to cell division, with each daughter cell receiving one complete set of DNA molecules.

Repressor: a protein that prevents transcription from a promoter element. In most cases, repressors release from DNA in the presence of cellular stimuli. For example the *lac* repressor binds DNA in the absence, but not the presence of lactose; as a result, only when lactose is present in its environment does *E. coli* expend energy synthesizing the enzymes necessary to metabolize lactose. Repressor/promoter pairs are used in molecular biology to allow protein expression only under desired conditions.

Resin: an insoluble material, usually a modified carbohydrate polymer, used to form the matrix of a column. More generally, the term resin applies both to the insoluble polymer, and to the polymer that has been derivatized with functional groups that allow separation of proteins. Thus, DEAE-cellulose is an anion-exchange resin.

Restriction enzyme: an enzyme that cleaves specific sequences of double-stranded DNA. For example, *Nco* I cleaves CCATGG between the two "C"; because it cleaves both strands the same way; digestion with *Nco* I leaves a four base stretch of single-stranded DNA extending from the 5´ end. This four-base overhang is called a "sticky end". (Restriction enzymes are one mechanism that bacteria use to degrade foreign DNA. Each wild-type strain contains a restriction enzyme and a methylase. The methylase tags the host cell DNA with methyl groups on A or C residues of specific sequences; this modification prevents degradation of the host cell's own DNA by the endogenous restriction enzyme. Laboratory strains typically have the restriction enzyme system inactivated to prevent the degradation of introduced plasmid DNA (for example, in K-12-derived *E. coli* strains, *hsdRMS* mutants have both restriction enzyme *Eco*KI and the corresponding methylase genes deleted, while *hsdR17* (r_K^- m_K^+) mutants

have the *Eco*KI gene inactivated, but retain the methylase). Note, however, that some foreign (*e.g.,* commercially available) restriction enzymes will not cleave DNA when the host cell methylation patterns alter the bases in their recognition sequence.)

Reverse transcriptase: a specialized DNA polymerase, usually derived from a retrovirus, capable of using RNA as a template for DNA synthesis. The normal paradigm for information flow within a cell is from DNA to RNA to protein; reverse transcriptases were given their name to reflect the fact that these enzymes alter this standard direction of information flow.

RNA (ribonucleic acid): a polymer of nucleotides normally containing four types of bases (adenine (A), cytosine (C), guanine (G), and uracil (U)), although some forms of RNA include additional types of nucleotide residues. RNA molecules have varying functions, most of these functions being involved with protein synthesis. In some viruses, RNA acts as the sole genetic material.

RNase: any of a number of enzymes capable of hydrolyzing RNA into small fragments. Unlike DNase, most isozymes of RNase are very stable enzymes that are extraordinarily resistant to heat inactivation. Because humans (and most other species) secrete RNase, and because RNase is much more difficult than DNase to inactivate, working with RNA is somewhat more challenging than working with DNA.

Running buffer: the solution (for proteins, almost exclusively an aqueous solution) used for chromatography. Running buffers usually contain a pH-buffering species as well as salts and other molecules designed to either enhance or prevent binding of proteins to column resins. Alternatively, the term running buffer is sometimes used to describe the electrophoresis tank buffer used for running electrophoretic gels.

Scintillation counter: an instrument for measuring radioactivity. In scintillation counting, radioactive decay excites organic molecules (scintillants); the molecules emit the energy in the form of light that is detected by the counter.

Scintillation fluid: a solution that aids in the detection and quantitation of radioactivity. The solution contains scintillants, which are molecules that emit absorbed energy (in this case, from radioactive decay) in the form of light.

Screening: a method for finding desirable cells in a mixture of cells by a process that requires testing by the investigator.

SDS: sodium dodecylsulfate, an ionic detergent used to denature and solubilize proteins in SDS PAGE and other techniques.

SDS PAGE: sodium dodecylsulfate polyacrylamide gel electrophoresis, a technique for separating denatured polypeptides based on their size.

Selection: a method for separating desired from undesirable cells based on ability to survive and/or grow. One common selection technique is to grow cells in the presence of an antibiotic.

Specific activity: the activity of an enzyme corrected for protein concentration (it is usually calculated by dividing the activity by the protein concentration). During purification procedures, determining specific activity allows assessment of the changing amounts of impurities remaining.

Start codon: a sequence that signals initiation of translation. Most genes use the sequence AUG (frequently referred to as ATG, because the AUG is derived from the ATG sequence found in the DNA); a few genes use GUG.

Stationary phase: in column chromatography, the solid resin support material that allows the molecules of interest to separate.

Stationary phase: in molecular biology, the period of little or no growth that occurs when the nutrients in an environment have been consumed, or when waste products have reached toxic levels. Stationary phase involves an adaptive response, in which the cells alter the genes being expressed to allow survival under limiting conditions.

Sticky end: the segment of single-stranded DNA extending 5´ or 3´ from a double stranded DNA fragment following digestion by a restriction enzyme that is capable of base-pairing to a compatible end of another DNA fragment (or the opposite end of the same fragment). Note: "sticky end" is a slang term; the technical term is "cohesive end"; however, few people use the term "cohesive end" except when writing formal papers.

Stop codon: a nucleotide sequence that signals the termination of translation. Three stop codons are commonly used: UAA, UGA, and UAG (frequently referred to as TAA, TGA, and TAG). Many laboratory strains of *E. coli* contain a suppressor tRNA for TAG stop codons; it is therefore preferable to avoid TAG as a stop codon in engineered DNA sequences intended for use in *E. coli*.

Stop solution: a solution containing a reagent that allows termination of an experiment at a defined time.

Suppressor tRNA: a tRNA that binds what is ordinarily a stop codon, but allows protein synthesis to continue by inserting an amino acid instead of terminating translation. In effect, the suppressor tRNA converts the stop codon into a codon for the amino acid. Many laboratory strains of *E. coli* contain TAG suppressor codons (especially supE, which inserts glutamate, or supF, which inserts phenylalanine).

Transcription: the process of synthesizing RNA from a DNA template. (Derived from the standard English term for converting information from one form to another: *e.g.*, verbal English is transcribed into written English; nucleic acid information is converted to a different type of nucleic acid information.)

Transformation: the process of inducing cells to take up DNA from their environment. Transforming bacteria involves using a salt solution to make "competent" cells.

Translation: the process of synthesizing protein from an RNA template. (Derived from the standard English term for converting information from one language to another; information in the form of nucleic acid is converted to a different "language": protein.)

Tris (tris-(hydroxymethyl) aminomethane): a buffer commonly used for biochemical experiments . Tris rarely interferes in biochemical reactions, and is inexpensive. However, Tris has a relatively high pK_a (8.1 at 25°C). In addition, the pK_a value for Tris changes by –0.031 pH units per °C, resulting in a large temperature-dependent pH variation in Tris buffers.

tRNA (transfer RNA): a small RNA molecule that mediates the incorporation of amino acid residues into a growing protein chain. Each tRNA is specific for one type of amino acid, and contains a sequence complementary to the corresponding codon.

Vector: an entity or mechanism for transmitting biological information. In molecular biology, this term is often used to refer to plasmid DNA, especially if the plasmid DNA contains a foreign gene and elements to drive transcription and translation of that gene.