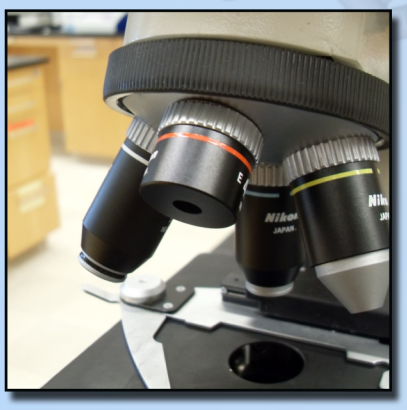
Investigations in Molecular Cell Biology

Author: Clare O'Connor

2011



Investigations in
Molecular Cell Biology

Boston College | Clare O'Connor
2011-2012

Table of Contents

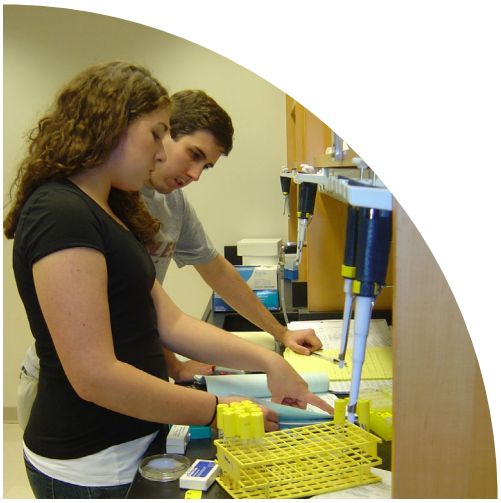
Table of Contents

Part II. Theory and Techniques

1 Introduction

Investigations in Molecular Cell Biology

- Welcome
- Course goals
- Our research project
- Course overview
- Acknowledgments



BI204 is a new kind of lab course. The Biology Department believes that authentic research experiences provide the best training for future scientists. In BI204, we will use some of the strain and clone resources generated by the yeast genome project to investigate the evolution of genes involved in the synthesis of methionine, an essential amino acid for all living organisms.

Chapter 1

Welcome

Welcome to BI204 - Investigations in Molecular Cell Biology. The BC Biology Department believes that research experiences are an important element of undergraduate biology education. Today, we live in a "post-genomic" era in which large-scale genome projects have generated tremendous amounts of sequence data as well as extensive collections of strains and DNA clones. Many of the "genes" annotated by the genome projects were identified by their similarity to known genes in other organisms, but their functions have not been tested directly. These loose ends provide opportunities for undergraduate students to participate in authentic functional genomics research. This course is designed as a research project in which students study the evolution of the genes involved in methionine biosynthesis. We hope that you enjoy the research experience and we look forward to your experimental results!

Course goals

The experiments in this course incorporate the core concepts and competencies of undergraduate biology education that were recently articulated in a national report sponsored by the National Science Foundation and the American Association for the Advancement of Science (Bauerle et al., 2011). Our project incorporates the themes of evolution, cellular structure/function, information transfer/gene expression, energy transformation and systems biology. In this course, you will also learn to design controlled experiments to test your hypotheses. You will also learn to use basic laboratory equipment, to collect and critically analyze experimental data, to access information in biological databases, to read the primary scientific literature, and to communicate scientific results in both oral and written presentations.

Our research project

During the semester, we will be investigating the evolution of the genes involved in the synthesis of methionine, an essential sulfur-containing amino acid in proteins. Most of our experiments will involve the budding yeast, *Saccharomyces cerevisiae*. We will take advantage of a large collection of *S. cerevisiae* strains with defined defects in methionine biosynthesis, which were generated as part of the yeast genome project. We will use these mutant strains as hosts for methionine genes cloned from other organisms. If the foreign gene restores the ability to synthesize methionine to its *S. cerevisiae* host, in a process known as functional complementation, we will know that gene function has been conserved over the evolutionary time frame that separates the two species.

During the 2011-2012 academic year, we will explore the conservation of methionine biosynthetic enzymes between *S. cerevisiae* and the fission yeast, *Schizosaccharomyces pombe*. Both *S. pombe* and *S. cerevisiae* are ascus-forming yeast from the phylum Ascomycota. The two species are thought to be separated by close to a billion years of evolution. During the time since the two species diverged from a common ancestor, the lineage that includes *S. cerevisiae* has undergone whole genome duplications and reductions. By contrast, *S. pombe* is considered to be much closer to ancestral members of the phylum. Our results during this coming year will establish which genes involved in methionine synthesis have been functionally conserved between the species. Results from this year will also guide the future direction of the project into other branches of life. Over time, student results should allow us to construct evolutionary trees for the genes involved in methionine synthesis.

Course overview

The first part of the semester will introduce you to the basic techniques used in molecular cell biology and genetics. You will receive three *S. cerevisiae* mutants, each of which is deficient in of the genes involved in methionine synthesis. You will also receive three yeast expression plasmids carrying the coding sequences for the the *MET* genes that are missing in the three *met* strains. Your task will be to identify the mutant strains and plasmids. As you carry out these experiments, you will learn to culture microorganisms, to analyze DNA sequences and to monitor gene expression. In the second part of the semester, you will design your own experiments, testing whether *S. pombe* genes can substitute for *S. cerevisiae* genes in these same mutant strains.

This manual is divided into two parts. The first section of the manual describes the series of experiments with *S. cerevisiae* genes and plasmids from the first part of the semester. The second part of the manual contains experimental protocols, together with some theoretical background for the procedures. We have chosen to place the techniques in a separate section from the experiments because you will use the techniques several times during the semester. During the second part of the semester, you may refer to the techniques as you design your own experiments.

This is the first edition of this manual, so some errors are inevitable. Your comments are welcome and will help us to improve the manual for future classes.

References

Bauerle, C. et al. 2011. Vision and Change in Undergraduate Biology Education: A Call to Action. (National Science Foundation/American Association for the Advancement of Science, Washington, D.C.)

Acknowledgments

It has been a pleasure to work with many Boston College colleagues on this manual. In particular, Dr. Michael Piatelli and Meghan Rice helped to design the experiments in this manual and offered many valuable comments on its contents. Thanks are also due to the students and teaching assistants who participated in the pilot semesters of BI204 and provided many useful suggestions for improving the course design.

The appearance of this manual owes a great debt to Arthur Hwang of the Boston College Media Technology Services. Arthur designed the layout of this manual and patiently provided advice during the writing process.

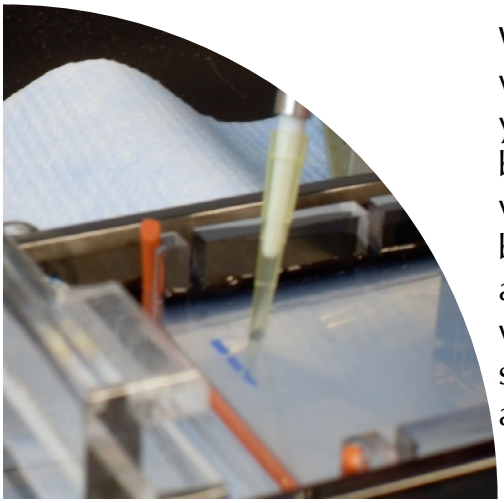
Part I

Project Experiments

2 Mastering the Micropipette

Objectives

- Learn to select and adjust micropipettes.
- Accurately transfer microliter volumes.
- Use the spectrophotometer to measure absorbance.
- Understand experimental errors in measurements.



Welcome to the microworld! In this class, you will be working with microorganisms, including yeast and bacteria, which are much too small to be detected with the naked eye. Indeed, millions would fit into a period on this page. You will also be working with costly reagents, such as plasmids and enzymes. Therefore, in every experiment, you will be required to accurately measure volumes as small as a few microliters (μL). Micropipettes will allow you to do this accurately and precisely.

Chapter 2

Introduction

In many respects, adjustable micropipettes are the most important scientific equipment that you will use this semester. These micropipettes have been calibrated to deliver small volumes of liquid with great accuracy and precision. Since you will be using micropipettes for all of your experiments, the quality of your results will depend on proper operation of the micropipette. Today's laboratory will lead you through some exercises that will show you how to use micropipettes correctly and point out some common pitfalls associated with their use. Your results will also provide information about whether the pipettes are functioning properly.

In these exercises, we'll be using the spectrophotometer to determine if your pipetting is accurate and precise. You will be given some deionized water and two solutions of the blue dye, bromophenol blue (BPB). In today's exercises, you'll combine various volumes of water and the two solutions of BPB, and you'll measure the absorbance of the resulting solutions at 590 nm (A_{590}), which is close to the absorbance maximum of bromophenol blue at neutral pH. Measuring errors will be reflected in the spectrophotometer readings.

Before starting this lab, read the background material on micropipettes in Chapter 13.

Carefully record your results in your lab notebook as you perform these exercises. Since this is the first experiment of the semester, the instructors have generated a template that you can use to record your results. In the future, you'll be expected to prepare your lab notebook in advance of the experiment. The template follows the general recommendations for lab notebooks in the Communication Resources section of this manual, but keep in mind that lab notebooks are a personal record of your work in the lab and their format is more flexible than a lab report. Lab notebooks are handwritten documents subject to modification during the course of an experiment. Lab notebooks are not always neat, but they need to be accurate! ***As a rule of thumb, someone who reads your lab notebook should be able to reproduce your experiment.***

Exercise 1

Getting the feel of micropipettes

Concept: Micropipettes work by an air displacement mechanism

1. Set the P200 to deliver 200 µL. Be careful not to overshoot, which could damage the pipette piston.
2. Grip the pipette by wrapping your fingers around the barrel.
3. Use your thumb to depress the plunger to its first stop.
4. Next press the plunger to the second stop.
   Compare the distance that the plunger moved during the first and second strokes.
5. Set the P200 to deliver 20 µL and depress the plunger to its first stop.
   Compare the distance that the plunger moved when the P200 was set to 200 or 20 µL.

6. Depress the plunger to the second stop.
   How does the distance between the first and second stops compare for 200 and 20 µL?
7. Set the P20 to deliver 20 µL. Depress the plunger to the first stop.
   Compare the distance to the first stop when a P20 and P200 are set to deliver 20 µL.

Concept: The filling and dispensing strokes are different.

1. Place a tip on the shaft of the P200.
2. Set the P200 to deliver 50 µL.
3. Draw up 50 µL of 0.05% BPB solution into the pipet.
4. Dispense the BPB into a microcentrifuge tube down to the first stop, holding the tip against the wall of the tube. Note whether all of the dye has been expelled. Push the plunger down to the second stop to release any remaining BPB.

Exercise 2

How NOT to pipette!

In this exercise, you'll determine the error that results from deliberate incorrect pipetting. (Hopefully, this is the ONLY time that you ever do this!)

1. Use the P1000 to add 990 µL of water to two microcentrifuge tubes. Label the tubes A and B. Dispose of used tips in the containers provided.
2. Use a P20 to correctly transfer 10 µL of 0.05% BPB to tube A. Make a mental note of what fraction of the pipet tip is filled with the dye. Use the vortex mixer to disperse the BPB in the water.
3. Use a P20 to ***INCORRECTLY*** transfer 10 µL of 0.05% BPB to tube B. Do this by depressing the plunger to the second stop before you take up the BPB solution. Make a mental note of how well the dye fills the tip this time.
4. Set the wavelength of the spectrophotometer to 590 (A_{590}). Pipette 1 mL of water into a plastic cuvette and blank the spectrophotometer at this wavelength.
5. Read the A_{590} of the solutions in tubes A and B, in the spectrophotometer.
   How do the two readings compare?
   What kind of error results from drawing solution into the pipette incorrectly?

Exercise 3

Determine the precision of micropipette transfers

The standard method used to determine the accuracy of a micropipette is to weigh a specified volume of distilled water on an analytical balance. Unfortunately, we don't have enough analytical balances available for this, so we will transfer specified volumes of BPB to deioinized water, and we'll then measure the A590 of the solution. (We will assume that there is no error in the measurement of the water!)

Work in groups of three. One person in the group should work with the P-20, another with the P-200 and the third with the P-1000. Each person should prepare three identical samples and then determine the A_{590} of the three samples. From the data, you will be able to determine if the micropipette is measuring volumes correctly.

1. Set up three microcentrifuge tubes and label them appropriately. The final volume (water + BPB) in each tube will be 1.0 mL. Calculate the volume of water that will need to be combined with each of the following to give 1.0 mL, and record your calculations in your lab notebook:
   P-20: 10 µL of 0.05% BPB
   P-200: 100 µL of 0.005% BPB
   P-1000: 300 µL of 0.005% BPB
2. To minimize our plastic waste, strategize how to minimize the number of tips that you use without contaminating the stock solutions. A tip can be used multiple times, but a tip that has been used for BPB cannot be used to subsequently transfer deionized water. Combine the BPB solution and water to give a final volume of 1.0 mL.
3. Measure the A_{590} of the three solutions and record the data in your notebook.
4. Compute the mean and standard deviations for your three measurements, using either a calculator or Excel.
5. Enter these values on the chart that your TA has prepared on the whiteboard. Compare the individual values that you obtained for the three pipettes with the combined class measurements for each tube.

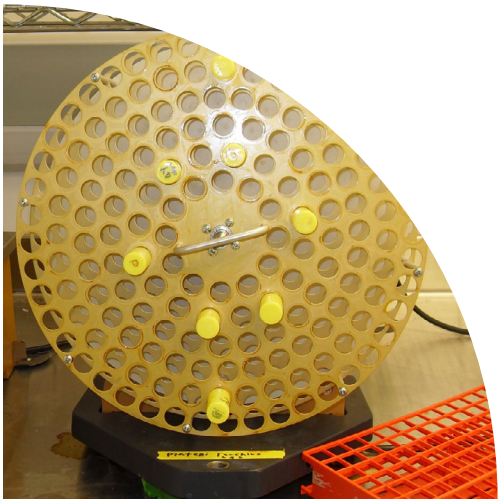
What conclusions can you make about the behavior of your micropipettes?

Notify your TA if any of the micropipettes are not performing properly. Your TA will follow up on your observations and test the micropipettes with the gravimetric test described at the beginning of this exercise.

3 Working with Yeast

Objectives

- Use sterile technique to plate yeast on nutrient media
- Use serial dilutions to determine the number of viable cells in a culture
- Estimate the number of cells in a culture with a spectrophotometer
- Use the light microscope to identify yeast at different phases of the cell cycle

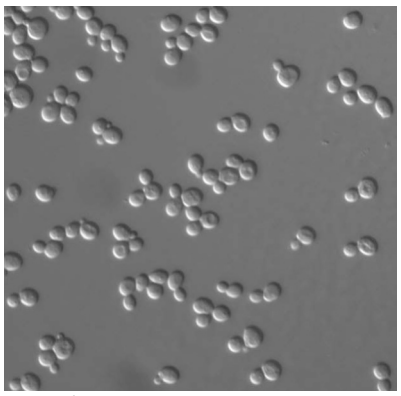


The budding yeast, *Saccharomyces cerevesiae*, is an important model organism that is easily cultured in the laboratory. This lab will introduce you to some of the techniques used for culturing, counting and observing yeast. You will also use the light microscope to compare *S. cerevisiae* with the fission yeast, *Schizosaccharomyces pombe*.

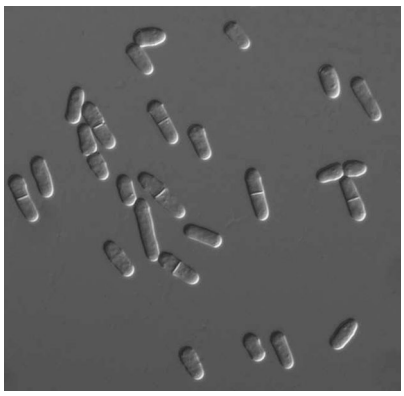
Chapter 3

Introduction

The budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*, derive their common names from their mechanisms of cell division. These differences in growth properties are clearly apparent using a light microscope. The images below were obtained using differential interference contrast (DIC) microscopy, a kind of light microscopy that provides detail about the surface properties of living cells.



Saccharomyces cerevisiae



Schizosaccharomyces pombe

Throughout the semester, you will be working with cultures of *S. cerevisiae*. This lab will introduce you to some of the techniques used for culturing yeast. Variations on these same techniques are used to culture other microorganisms. Success in the lab will depend on your ability to use sterile technique, which is needed to maintain the integrity of the different strains that you will work with throughout the semester. An equally important element in laboratory success is careful bookkeeping! ***Make sure that all plates and culture tubes are clearly labeled AND that the labeling information is included in your notebook.***

This lab will introduce you to some of the techniques used to culture microorganisms. You will also learn how to use the light microscope to look at yeast cultures. Our laboratory microscopes do not have the sophisticated optics used to obtain the images above, but you will be able to distinguish cells at various points in the cell cycle and to distinguish the two yeast species. At the end of this lab, you should be able to perform sterile transfers, to estimate the number of cells in a yeast culture, and to distinguish *S. cerevisiae* and *S. pombe* by their growth properties and appearance in the light microscope.

Before starting this lab, read the information on yeast culture in Chapter 14 and the information on spectrophotometers and light microscopes in Chapter 13.

You will receive two cultures of both *S. cerevisiae* and *S. pombe* before beginning these experiments. One culture will contain cells in vegetative growth and the other, denser culture will contain cells in stationary phase. All of the cultures have been grown in YPD medium. *Can you distinguish the two phases by looking at the culture tubes?*

Students should work in groups of two for these experiments.

Exercise 1

Prepare spot plates of *S. cerevisiae* and *S. pombe* cultures

Concepts: The number of colonies on spot plates can be used to calculate the number of viable cells in a culture. Colony size gives information about growth rates.

1. Prepare a series of five 1:10 dilutions of the stationary phase cultures of *S. cerevisiae* and *S. pombe*, using STERILE water to dilute the cultures. ***Be careful not to confuse sterile water with deionized water!***
2. Spot 5 µL aliquots of each dilution series on a YPD plate.
3. After the cells in the spots have sufficient time to settle and adhere to the agar, invert the plates. You should allow at least 30 minutes.
4. Follow your TA's instructions for incubating the plates. The growth rate of yeast is heavily dependent on temperature. It is important the plates do not over-grow, so you will be able to distinguish individual colonies. The culture schedule that you will use will vary, depending on the day that your class is scheduled.

During your next class period:

5. Record your data, following the instructions in Chapter 14.
6. Use the data to calculate the density of the original cultures.

Questions to consider:

S. cerevisiae cells divide approx. every 90 minutes at 30˚C. Estimate how many cells are present in a colony of *S. cerevisiae* cells.

Are there any differences in the number or sizes of the colonies from the *S. pombe* and *S. cerevisiae* cultures? Hypothesize what the differences may mean.

Compare the number of cells estimated from the spot plates with your estimate from the spectrophotometer in exercise 2. What might account for any differences that you see?

NOTE: Exercises 2 through 4 do not require sterile technique, since these cells will not be used for further cultures. Nonetheless, practice good pipetting techniques.

Exercise 2

Estimate cell density with the spectrophotometer

Concept: Light scattering of a yeast culture can provide a rough estimate of the culture density.

Before starting this exercise, review the information on spectrophotometer use in Chapter 13.

1. You will use two cuvettes for this experiment. One cuvette will serve as the reference cuvette. Fill this cuvette with 1.0 mL deionized water and place a small piece of Parafilm over the top.
2. Use the arrows labeled "nm" to adjust the wavelength of the monochromator to 600 nm.
3. Place the reference cuvette in the cuvette holder. Orient the cuvette so that the light beam passes through the flat face of the cuvette.
4. Close the lid and press the "0 Abs/100%T" button to establish a baseline value for further measurements. Remove the reference cuvette from the holder and place it on the bench next to the spectrophotometer. Do NOT discard it.
5. Place a second cuvette containing water in the sample position of the spectrophotometer. Close the lid and record the OD_{600} value in your lab notebook. Ideally, this value is close to zero. Any variation from zero is due to variations in the cuvettes. You will need to subtract this value from all subsequent measurements using the same sample cuvette.
6. Remove the cuvette from the instrument and replace the water with one of your samples. Close the lid and read the OD_{600}, as in step 5. Record this value in your lab notebook. If the optical density of the sample is greater than 1.0, dilute the sample 1:10 with distilled water and read the optical density again. (The linear relationship between the OD_{600} and cell density is lost when OD_{600} values exceed 1.0) Record the new value in your lab notebook, noting how you diluted your sample. Dispose of all cell material in the white liquid waste container.
7. Repeat step 6 with each of your samples, using the same cuvette.
8. Calculate the cell density of the original culture, assuming that 1.0 OD_{600} corresponds to a density of 1-2 x 10^7 yeast cells/ml. Use only data where the OD_{600} is less than 1.0 for these calculations.

How do the culture densities calculated with the spectrophotometer compare with those calculated from spot plate data? If you see any differences, hypothesize why this is so.

Exercise 3

Using the compound light microscope

Concept: Magnification of a specimen can be adjusted by changing the objective lens

1. Identify the parts of the microscope. Note the positions of the objectives, the coarse and fine focus adjustments, the adjustable light switch and the condenser diaphragm.
2. Locate the four objective lenses on the microscopes. The magnification of each lens (4x, 10x, 40x, and 100x) is stamped on its casing. Rotate the 4x objective into position. Adjust the position of the aperture diaphragm to the corresponding 4x position.
3. Turn on the microscope lamp and adjust the dimmer switch until the light is not too intense when you look through the eyepieces. You may need to adjust the distance between the eyepieces to fit your eyes.
4. Place the transparent ruler on the microscope stage. Use the coarse focus knob to bring the ruler into focus. You may also need to adjust the light. Make additional adjustments with the fine focus knob. The ruler is graduated into 1 mm divisions. (Remember that 1 mm is equal to 1000 μm, the unit of distance usually used by microscopists.)
   What is the diameter of the field of view with the 4x objective?
5. Use the stage manipulators to move the ruler to either the right or the left.
   What direction does the image move?
6. Dial the 10x objective into position and adjust the condenser diaphragm.
   How does the distance between the specimen and the objective change?
   Adjust the focus with the coarse and fine focus adjustment knobs.
   What is the diameter of the field of view with the 10x objective?
   Can you find a mathematical relationship between the magnification and the diameter of the field of view?
7. Swing the 40x objective into position and adjust the condenser diaphragm. Adjust the focus using ONLY the fine focus knob.
   What happens to the working distance and the field of view?

Exercise 4

Using the compound light microscope to observe yeast cultures

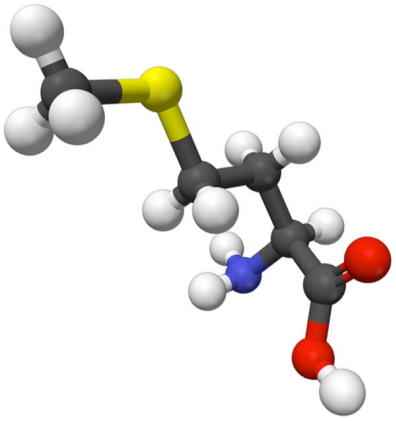
As yeast progress through the cell cycle, they undergo characteristic morphological changes. A log phase culture will have cells in the G1, S, G2 and M phases of the cell cycle. In this experiment, you will observe both log phase and stationary phase cultures of *S. cerevisiae* and *S. pombe*. The two yeasts use very different modes of division. In *S. cerevisiae*, buds begin to form when cells enter S phase. The size of the bud, which will become the daughter cell, continues to grow until the cells divide in M phase. At the time that the cell divides, the daughter cell is still smaller than the mother cell. The daughter cell will need to grow a bit before it enters another round of cell division. By contrast, *S. pombe* divides by medial fission. Cells grow in length until they are 12-15 μm, at which point a septum forms and the cell divides. In this experiment, you will use the compound light microscope to observe samples of the four cultures. Make careful note of the sizes and shapes in each culture.

1. Concentrate the cells in your log phase yeast cultures by spinning them for a count of 10 in a microcentrifuge set at top speed. Use a transfer pipet to remove most of the culture medium, until the medium just covers the cell pellet. Resuspend the cells with the vortex mixer. (It is not necessary to centrifuge the stationary phase cultures, which are more concentrated than the log phase cultures.
2. In this experiment, you will prepare a series of slides, each of which contains two different samples for easy comparison. The slides are large enough to accommodate two samples (and two coverslips). In your notebook, be sure to record which sample is closer to the frosted area on the slide. (If your slide does not have a frosted area, mark one end of the slide with a magic marker.)
   First, let's compare the two yeast species. Prepare a slide with log phase cultures of *S. pombe* and *S. cerevisiae*. Spot 5 μL of each cell suspension and cover each sample with a coverslip.
3. Use the same sequence of microscope adjustments that you used in the ruler exercise to visualize the cultures. Start at low magnification and gradually increase the magnification, making changes in the condenser diaphragm as needed. (Play with the position of the aperture diaphragm a bit to maximize the quality of the image. The cultures have not been stained, so they may show up better with less light than you expect.) In your notebook, draw some examples of the forms that you see in the cultures and the relative proportions of each form. Comment on both the sizes and shapes of the cells.
4. Rinse the slide and coverslips with deionized water and dry them with a paper towel. Spot 5 μL of a stationary phase culture of *S. cerevisiae* on the slide and apply a coverslip. For the second spot, combine 2.5 μL of stationary phase cells with 2.5 μL of log phase cells. *Can you distinguish stationary phase cells from exponentially growing cells?* Comment on any differences in size and morphology.
5. Rinse and dry the slide and coverslips. Prepare a third slide with *S. pombe* cultures similar to the one that you prepared in step 4.

4 Yeast *met* Mutants

Objectives

- Become familiar with the biochemical pathways involved in methionine synthesis
- Design and carry out a strategy to distinguish *met* mutants by their nutritional requirements
- Streak yeast strains on selective media containing different sources of organic sulfur



The growth properties of mutant organisms can often provide information about the gene products involved in biochemical pathways within cells. In this experiment, you will use selective media to identify the mutated genes in yeast strains unable to grow in the absence of methionine.

Chapter 4

Introduction

Mutant organisms provide powerful tools to study biochemical pathways in living cells. This semester, we are working with yeast strains that are unable to synthesize methionine because one of the genes involved in the biosynthetic pathway has been inactivated. Methionine is an essential amino acid for all organisms. Consequently, these *met* mutants must receive methionine or another compound that they can convert to methionine from their diet.

The biochemical pathway for methionine synthesis in yeast has been elucidated from the combined experiments of many different biochemists and geneticists. In each of our met mutant strains, a *MET* gene has been replaced with a bacterial kanamycin resistance gene by homologous recombination. Depending on the position of the missing gene in the methionine biosynthetic pathway, which is shown on the opposite page, these *met* mutants may be able to grow in the absence of methionine if another source of organic sulfur is present. In this lab you will use selective media to identify the affected *MET* gene in three different mutant strains.

Before starting this lab, read the material from Chapter 15 on genetic screens, auxotrophs and genetic nomenclature. Be familiar with the streak plate technique described in Chapter 14.

Exercise 1

Devise a strategy to identify met mutants by their growth properties

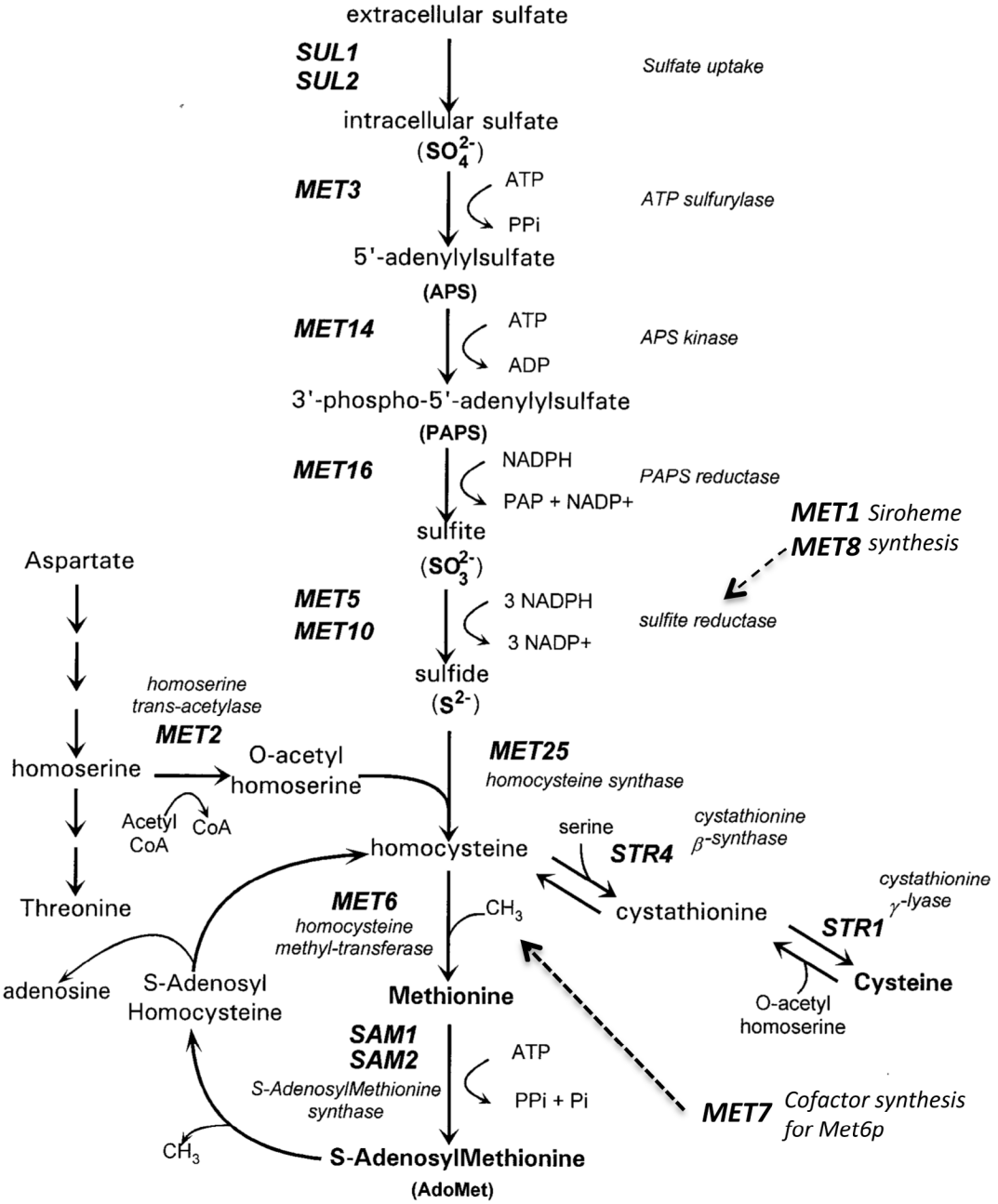
The *met* mutants that we are using for our experiments were derived from the parent strain BY4742, which has the genotype *MATa his3-Δ1 leu2Δ0 lys2Δ0 ura3Δ0*. Each strain contains one additional mutation in a *MET* gene.

You will be receiving six agar plates with different media to use for screening your *met* mutants. The plates vary in their content of methionine (Met), cysteine (Cys) and homocysteine (Hcy), as noted in the table below. With the exception of YPD, the other plates contain synthetic YC media with additions or subtractions as noted below. The complete recipe for YC media is included in Chapter 15.

Media	YPD	YC - Met	YC - Met +Cys	YC - Met +Hcy	YC -Ura	YC Complete
methionine	+	-	-	-	+	+
cysteine	+	-	-	+	-	-
homocysteine	?	-	+	-	-	-

Construct a table in your notebook that predicts the ability of various *met* strains (*met1,2,3,5,6, 7,8,9,10,14,16,25*) to grow on each of the six plates. Place a plus (+) when you predict that the strain will grow on the plate and a minus (-) when you do not expect the strain to grow.

Methionine biosynthesis in yeast.
The proteins catalyzing individual steps in the pathway are listed next to each step. The names of the genes encoding the activities are shown in capital letters,according to *S. cerevisiae* convention. *MET1* and *MET8* are involved in synthesizing siroheme, an essential cofactor for sulfite reductase. The *MET7* gene product is involved in synthesizing a different cofactor for Met6p, homocysteine methyltransferase.



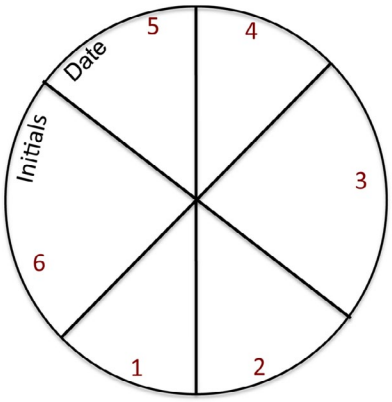
Exercise 2

Streak out strains on selective plates and YPD

Two groups will share each of the selective plates. (You will be able to collect more data this way and we will conserve reagents.) Each group will receive three strains identified by a code. Be very careful recording the code details in your notebook!

1. Label the bottom of each plate with a magic marker before you begin to streak out the strains.

   - Use small letters and keep the labels at the rims of the plate.
   - Label each plate with the date and your initials.
   - Divide each plate into six sectors.
   - Write the number of the strain that will be streaked into each sector.
     Use the SAME pattern for each of the six plates!



2. Streak a different strain into each of the sectors, as described in Chapter 14.
   - Use two toothpicks for each strain. The first toothpick will be used to transfer yeast from another plate.
   - Be careful not to transfer too much! You should barely see the strain on the tip of the toothpick.
   - After the first streak, change to a new toothpick and use this for the second and third streaks.
3. Incubate the plates at 30˚C or room temperature until you can detect individual colonies. Place a piece of Parafilm™ around the rim of each plate and put the plates in the refrigerator or cold room. Do not allow the plates to overgrow!
4. During the next period, scan the plates as directed. Make sure that all of the plates have the same orientation, so that you can easily determine the results by looking at the image.
5. Hypothesize which strains could be streaked into each sector.

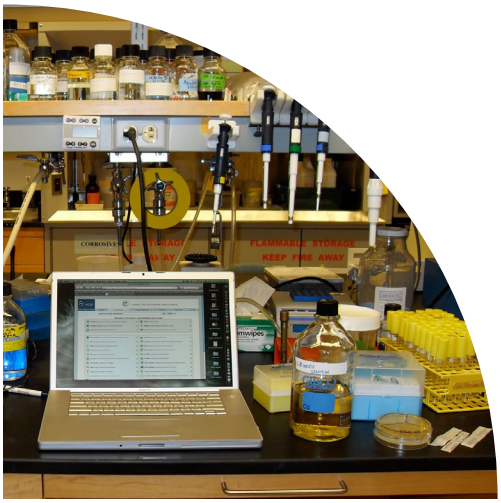
References

Thomas, D. and Surdin-Kerjan, Y. 1997. Metabolism of sulfur amino acids in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **61**: 503-532.

5 Introduction to Databases

Objectives

- Learn how to search for information in the NCBI databases
- Learn how to find protein information in the UniProt Knowledgebase
- Learn how to use the *Saccharomyces* Genome Database to find information about a yeast gene



In today's biology lab, the computer belongs on the benchtop along with other essential equipment. A network of online databases provides researchers with quick access to information on genes, proteins, phenotypes, and publications. In this chapter, you will learn to use several databases as you search for information on a *MET* gene and the protein it encodes.

Chapter 5

The best way to learn about databases is to use them! This lab will bring you through a kind of scavenger hunt, where you will search for information on a *S. cerevisiae MET* gene. The end of this chapter contains a series of questions asking about gene-specific information from the databases. Fill in this information as you are guided through the databases. In this chapter, you will be given a quick overview of several databases designed for different communities. As you progress through this lab, you will realize that there are many different ways to access the same information, and you may develop preferences for certain databases over others. You will also develop familiarity with database accession numbers as you wend your way through the records. You may feel like you're going in circles at points, because you are! The records in databases are extensively hyperlinked to one another and you may find the same record via multiple paths. We will start the lab with a visit to the National Center for Biotechnology Information (NCBI), which houses a collection of databases with different specialties.

Before starting this lab, read Chapter 16 on Databases.

NCBI Databases

1. Point your browser to the NCBI homepage: ***ncbi.nlm.nih.gov***
   Hint: Bookmark this page. You will use it often.
2. This is the homepage of the NCBI, which is actually a large collection of databases. Note the search box at the top with the dropdown box. Clicking on the dropdown box brings up a list of individual databases for more targeted searching. For now, leave the "All databases" setting. Write the name of your *MET* gene in the search box and click "Search."
3. The search brings you to the NCBI Entrez page, which summarizes the number of hits in each of the many NCBI databases. The number is probably quite large! Take a look at the results. In your notebook, record the number of records in the PubMed, Nucleotide, Protein, and Structure categories.
4. Modify the search term by adding "AND Saccharomyces cerevisiae" to the search box. Record the number of records in each of the categories used above. The numbers have probably dropped significantly! *Why do you think that this happened?* This simple comparison may give you some idea of the sheer volume of records in the NCBI databases. You may not receive any hits in the Structure category, since the vast majority of proteins have not been crystallized or studied with NMR.
5. Your next challenge is to narrow down this list to sequences of interest. In this course, we will be focusing on proteins involved in methionine biosynthesis. Proteins sequences are subject to more evolutionary constraints than gene sequences, so there are usually less "Protein" links than "Nucleotide" links. Click on the "Protein" database link.
6. From the search results, find and click on the record with the accession number that begins with NP_. This is the RefSeq protein record. The RefSeq record is a very rich source of information about your gene. Record the NP_ accession number.

RefSeq protein records

In the first line of a RefSeq NP_ record, you can find the accession number, the length of the protein and the date of the the last revision of the record. All RefSeq records are actively curated as new data becomes available and in response to comments from the scientific community. As you look through the record, note the dates that modifications were made and the nature of the modifications. The yeast sequences in RefSeq were generated by the genome sequence project, which set a high standard for accuracy. Nonetheless, some errors were made in the 12 million bases!

RefSeq protein records contain multiple fields, depicted in capital letters. We will point out just a few features of the records. Feel free to explore the hyperlinks under the various fields.

- DBSOURCE: The NM_....record is the nonredundant transcript record from the RefSeq database. Note that NM_ records contain only transcript coding regions. GenBank contains many other full or partial transcript records, including EST sequences, that contain untranslated portions of mRNA sequences. As you might expect, there is greater variability in untranslated regions of transcripts than in coding regions. Record the NM_ number of your gene.
- REFERENCE: This field contains publication information about the Met protein, with active hyperlinks to the publications referenced in PubMed. PubMed assigns accession numbers sequentially to its entry. Note that over 8.8 million records were already entered into PubMed before the yeast genome sequence was reported.
- FEATURES: This section contains quite a bit of information about your protein. You will find the molecular weight of the protein, its enzymatic classification (indicated in the E.C. number), conserved domains (CDD), and its mRNA coding sequence (CDS - this is identical to the NM_ record for the gene). Under the CDS subfield, note links to gene records in two other important databases. The Gene ID link brings you to a summary page that NCBI has prepared for each gene within its genomic context. The SGD accession number brings you to the record for your gene in the *Saccharomyces* organism database, which integrates molecular and genetic data for your gene. We'll visit the SGD later in these exercises.
- ORIGIN: This field contains the amino acid sequence of the protein.

The right side of the NP_ record contains many links to other information on your gene that the NCBI curators have prepared. Feel free to explore here.

UniProtKB

Researchers seeking out protein information frequently turn to the UniProt Protein Knowledgebase, which is jointly maintained by European Bioinformatics Institute, the Swiss Institute of Bioinformatics and the Protein Information Resource. There are two portions to UniProt. The Swiss Protein Database is an older database that predates the whole genome sequencing era. The Swiss Protein Database is a nonredundant database that contains information from both protein and translated nucleotide sequences. The TrEMBL portion of UniProt automatically translates nucleotide sequences in the EMBL nucleotide database. Some TrEMBL records enter Swiss Protein, but only after curators determine that the sequence

information is nonredundant. In this course, we will use the nonredundant records in the Swiss Protein Database.

1. Direct your browser to the UniProt homepage: ***uniprot.org***.
2. Enter the name of your *MET* gene in the search box. How many records does your search retrieve? How many are in Swiss Protein (yellow stars)? How many in TrEMBL (gray stars)?
3. Refine your search by adding "AND Saccharomyces" to the search box. How does this affect your results?
4. Click on the Swiss Protein record for your gene. (This should be the top entry.) Record the accession number, which can be used to query other databases.
5. The Swiss Protein record is a rich source of information on proteins. Record the recommended name for the protein and its E.C. (Enzyme Commission) number. The E.C. numbers classify enzymes in a hierarchical system according to the kind of reaction that they catalyze. Both the enzyme nomenclature and E.C. numbers conform to international standards set up by the International Union of Biochemistry and Molecular Biology.
6. Look through the general annotation for the enzyme. Does the enzyme have a cofactor? Is the enzyme composed of multiple subunits?
7. Look through the sequence annotation. What kinds of domains are present in the protein? Which (if any) amino acids are involved in cofactor binding? Are any amino acids sites of post-translational modification?
8. Go down to the Sequence field of the record. You should see the complete sequence of the protein. Note the dropdown box under Tools. UniProt offers some powerful tools for analyzing protein sequences in the biochemistry lab. You can calculate the protein's MW and isoelectric point (pI), generate hydrophobicity plots, predict the products of proteolytic digests as well as the masses of peptides generated by proteolytic fragmentation. Use the tool to compute the pI and MW for your protein.

The Uniprot and NP_ submissions should have the identical amino acid sequences. Either accession number can be used with BLAST and other bionformatics tools. Now that you've seen both records, which site do you prefer?

Saccharomyces Genome Database

The *Saccharomyces* Genome Database (SGD) is an invaluable tool for yeast researchers. It is extensively curated by yeast biologists, who link information from the primary literature to information in NCBI, Uniprot and other databases. Members of the yeast research community also contribute information about phenotypic and molecular information about yeast strains. Since the completion of the Yeast Genome Project in 1998, all of the genes have been systematically organized by locus names that correspond to the positions of the genes on the yeast chromosomes. The SGD is the perfect place to start learning about a *MET* gene and its function, since much of the hunting has been done for you by these professionals.

Before proceeding, review the tutorial on the SGD at www2.bc.edu/clare-oconnor/tutorials

1. Direct your browser to the SGD home page ***yeastgenome.org***.
2. Type the name of your *MET* gene into the search box. This will bring up the summary page for your gene.
3. Use the information from the tutorial to answer the questions at the end of this chapter.

Database search questions

Answer the following questions from your database searches.

NCBI records

1. Which MET gene are you following?

2. How many records did Entrez retrieve when only the gene name was used in the search?
   PubMed Nucleotide Protein Structure

2. How many records did Entrez retrieve when Saccharomyces was added to the search terms?
   PubMed Nucleotide Protein Structure

3. What is the NP_ accession number for your gene?

4. What is the NM_ accession number for your gene?

UniProtKB records

1. How many records did UniProt retrieve when only the gene name was used in the search?
   Total Swiss Prot records TrEMBL records

2. How many records did UniProt retrieve when Saccharomyces was added to the search terms?
   Total Swiss Prot records TrEMBL records

3. What is the Swiss Protein accession number for your protein?

4. What is the recommended name for your protein? ...its E.C. number?

5. What reaction(s) does the protein catalyze?

6. Does the enzyme contain cofactors? If so, what are they? Where in the sequence do the cofactors bind?

7. Use the tools in the sequence portion of the record to calculate the molecular weight and isoelectric point of the protein.

Saccharomyces Genome Database records

1. What is the chromosomal location of your gene? (Use the systematic name for the answers.)

2. How many introns are in your gene?

3. How was the gene originally identified?

4. List two phenotypes associated with deletion of the gene.

5. Does the gene have any other aliases?

6. List two conditions where expression of your gene is increased at least 2-fold

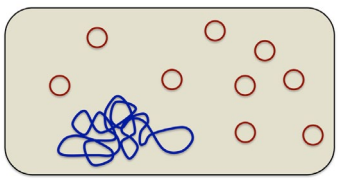
7. List two conditions where expression of your gene is decreased at least 2-fold

8. Is the protein known to interact with other proteins?

6 Plasmid Isolation

Objectives

- Learn how the physical properties of plasmids are used in their purification
- Isolate a plasmid from *E. coli*
- Use ultraviolet spectroscopy to estimate the concentration of plasmid DNA



Plasmids play a central role in molecular biology, and bacteria hosts allow them to be produced in large quantities. In this lab, you will use a miniprep procedure to isolate plasmids that you will use later for yeast transformation.

Chapter 6

Introduction

Plasmids are the workhorses of molecular biology. Plasmids are small, circular DNA molecules that replicate independently of the chromosomes in the microorganisms that harbor them. Plasmids can be engineered to carry large amounts of foreign DNA, they can be transferred between organisms, and they can be easily isolated from microorganisms for manipulation in the lab. Today, they provide an essential tool in molecular biology. In this laboratory, you will isolate plasmid DNA from a bacterial culture, taking advantage of the distinctive physical properties of plasmids.

Before beginning this laboratory, read the background material on plasmids in Chapter 17.

The plasmids that we are using this semester are shuttle vectors that are capable of replicating in both yeast and bacteria. Plasmids are routinely propagated in bacteria, because bacteria tolerte more copies of the plasmid, they replicate more rapidly than yeast and they grow to a higher density in culture. In this lab, you will isolate plasmids from transformed bacteria. In a few weeks, you will use these plasmids to transform yeast. The procedure that we will use to isolate the plasmids exploits the distinctive physical properties of the bacterial chromosomal DNA and the plasmid DNA.

The Zyppy™ miniprep kit uses a proprietary version of a classical alkaline lysis protocol. Plasmids purified with the Zyppy protocol are pure enough to be used for a variety of procedures, such as restriction enzyme digestion and transformation. The exact components of the Zyppy kit are proprietary, but we can make some educated guesses from the literature and the components of competitors' kits.

Plasmid isolation begins with a culture of *E.coli* that have been transformed with plasmids. In the first step of the procedure, the bacteria hosting the plasmid are lysed with sodium hydroxide, which destroys the cell membrane and denatures DNA and proteins. The lysate is then neutralized, which allows plasmid DNA, but not chromosomal DNA or proteins, to reassume its native structure. In the next step, denatured proteins and denatured chromosomal DNA are removed from the lysate by centrifugation. Plasmids are then purified from the clarified lysate by absorption to a silica resin. Once the plasmids are bound to the resin, it is washed with solutions that remove salts and any residual protein. In the last step of the procedure, DNA is eluted from the resin in a small volume.

NOTE: Sterile technique is not required for plasmid isolations.

Plasmid purification

Concentrate the plasmid-bearing bacterial cells

1. Your group will receive a bacterial culture that was grown overnight in Luria Bertani (LB) media containing 100 µg/mL ampicillin. The density of cells in the culture is expected to be 3-4 X 10^9 cells/mL.

   What is the purpose of the ampicillin? How does it work?
2. Transfer 1.5 mL of the 3 mL culture to a microcentrifuge tube. Concentrate the cells by centrifuging them at maximum speed (~14,000 rpm) for 1 min. The cells will form a white pellet at the bottom of the tube.
3. Remove the culture medium above the cell pellet by gently decanting the supernatant. Add the remaining 1.5 mL of bacterial culture to the tube containing the cell pellet.
4. Concentrate the cells once more by centrifugation at maximum speed for 1 min.
5. Use a P200 to aspirate off as much of the culture medium as possible.

Alkaline lysis of bacterial cells harboring the plasmids

6. Re-suspend the pellet in 600 µl of TE buffer (Tris-HCl, EDTA pH=8.0) using the vortex mixer.
7. Add 100 µL of 7X Blue Zyppy Lysis buffer to the tube. Mix the buffer and cells by gently inverting the tube 4-6 times. ***Be gentle!*** Too much mechanical stress will fragment the bacterial chromosomal DNA and contaminate your plasmid preparation. The solution should turn from a cloudy blue to a clear blue.

Separate plasmid DNA from denatured proteins and chromosomal DNA

8. Add 350 µL of cold Yellow Zyppy Neutralization buffer (w/RNAase A) to the tube, and mix the contents thoroughly by inverting several times. The solution will turn yellow when neutralization is complete, and a yellowish precipitate will form. Invert the sample an additional 3-4 times to ensure complete neutralization.
9. Centrifuge the mixture at maximum speed for 3 minutes to remove denatured proteins and chromosomal DNA. Notice that the tube contains a yellow precipitate that has collected on one side of the tube. The pale yellow supernatant contains the plasmid DNA.

Purify plasmid DNA by adsorption to a silica resin.

10. Using a P200, carefully transfer the pale yellow supernatant (~900 µL) onto a Zyppy spin column. Be careful not to transfer any of the yellow precipitate!
11. Place the column with the collection tube attached into a centrifuge and spin at maximum speed for about 15 seconds. It is best to use the "pulse" button on the centrifuge and count to 15 or 20 seconds for this centrifugation step.
12. Remove the column and discard the flow through in the collection tube.
13. Place the column back into the collection tube and add 200 µL of Zyppy Endo-Wash solution. (Endo-Wash contains guanidine hydrochloride and isopropanol, which will remove denatured proteins from the resin.)
14. Centrifuge for 15-20 seconds, and discard the flow through.
15. Place the column back into the collection tube then add 400 µL of Zyppy Column Wash buffer. (This steps removes contaminating salts.) Centrifuge for 30-40 seconds.

Elute the plasmid DNA

16. Transfer the Zyppy column to a clean (and appropriately labeled) 1.5 mL centrifuge tube, leaving the lid of the tube open.
17. Carefully, add 100 µL of TE buffer directly on top of the white column bed. Place the pipette tip as close as you can to the white column bed without poking it. Slowly dispense the TE on top of the resin bed.
18. Allow the buffer to percolate into the column by letting the column stand upright in the microcentrifuge fuge tube for 10 minutes.
19. Centrifuge the column at maximum speed for 30 seconds. Again, it's fine to leave the cap open during this spin.
20. Remove the column, cap the tube and place it on ice. This tube should now contain plasmid DNA.

Measure the DNA concentration

The concentration and purity of DNA samples can be estimated from their absorbance of ultraviolet (UV) light. Nucleic acids absorb light strongly in the UV range, with an absorbance maximum of ~260 nm. An A_{260} of 1 corresponds to ~50 µg/mL DNA. Proteins also absorb UV light due to their aromatic side chains, but much less strongly than nucleic acids. Proteins have a different absorbance maximum of ~280 nm. An A_{280} of 1 corresponds to 1 mg/mL protein. Because of the differences in protein and nucleic acid absorbances, the A_{260}/A_{280} ratio of a sample is useful in determining the purity of a DNA sample. Pure DNA has an A_{260}/A_{280} of ~1.8. Protein or chemical contamination will lower this ratio.

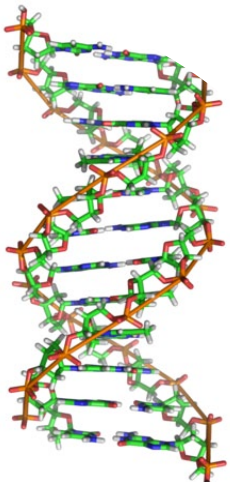
In this laboratory, you will use an instrument called the Nanodrop™ to obtain an ultraviolet absorbance spectrum for your plasmid DNA samples. The Nanodrop software will also calculate the concentration of DNA and the A_{260}/A_{280} ratio. The NanoDrop uses fiber optic technology and the surface tension of a sample to provide accurate absorbance measurements over a large dynamic range of DNA concentrations. From a practical point of view, this means that you need to sacrifice only small amounts of your precious DNA sample. The NanoDrop is amazingly simple to use, and your TA will demonstrate the appropriate use of the instrument.

1. Pipette 1 µL of your plasmid sample on the pedestal on top of the pin (shown at the end of the arrow at right).
2. Close the arm and record the absorbance.
3. When the measurement is complete, open the arm and wipe the pedestal surface with a lint-free wipe in preparation for the next sample.
4. Record the absorbance readings in your notebook.

7 Plasmid Forensics

Objectives

- Choose restriction endonucleases for mapping plasmid DNA
- Set up reactions containing plasmid DNA and restriction endonucleases
- Use PCR to amplify MET genes in plasmids
- Analyze restriction digests and PCR reactions on agarose gels



Microbiologists hunting down the causes of foodborne illnesses and detectives hunting down criminals all use methods developed by molecular biologists to analyze DNA. Restriction endonuclease digests and the polymerase chain reaction (PCR) are routinely used to distinguish DNA samples from one another. During the next few laboratories, you will use these same tools to identify *MET* genes carried by plasmids.

Image from http://en.wikipedia.org/wiki/DNA#mediaviewer/File:DNA_orbit_animated_static_thumb.png

Chapter 7

Introduction

In the last experiment, you isolated plasmids carrying *MET* genes from bacterial cell lysates. Each group was given three plasmids that carry coding sequences for different *MET* genes. These *MET* genes correspond to the same genes mutated in the met strains whose growth you analyzed on selective plates. Your task over the next few lab sessions is to identify the *MET* genes encoded by the plasmids. To do this, you will use some of the standard tools used in DNA forensics laboratories. In the first part, you will design and carry out a strategy to distinguish between the plasmids using restriction endonucleases (REs). The products of the restriction digests, or restriction fragments, generate a restriction map when they are separated by agarose gel electrophoresis.

In the second part, you will use the polymerase chain reaction (PCR) to differentiate between the plasmids. All of the *MET* genes have been cloned into the same cloning vector, the plasmid pBG1805 (Gelperin et al., 2005). In the PCR experiment, you will be able to determine the size of the *MET* genes that have been cloned into your plasmids. From the results of the two experiments, you will hopefully be able to match the plasmids with specific strains for future experiments.

Before starting this lab, read the background material in Chapters 17 and 18.

Exercise 1

Plan the restriction digest

Assign each person in your group a different *MET* gene. You will use the sequence information for your *MET* gene and the sequence information for pBG1805 vector to predict which REs will cut the plasmid and/or your gene coding sequence. You will also be able to predict the sizes of restriction fragments generated with different REs.

1. Find the coding sequence of your *MET* gene from the Yeast Genome Database.
   http://yeastgenome.org
   Use the Retrieve Sequences drop-down box on the right side of your gene's summary page to retrieve the coding sequence for your gene. Record the number of nucleotides in your gene's coding sequence.
2. Retrieve the pBG1805 sequence from either GenBank or the course website. Paste the *MET* coding sequence at the end of the BG1805 vector sequence. After you finish pasting, delete the last three nucleotides of the MET coding sequence, which is the gene's stop codon.
   Why are we deleting the stop codon?
   The BG1805 sequence contains 6573 nucleotides. The first nucleotide of your *MET* gene will be nucleotide 6574.
3. Paste the sequence from step 2 into the search box in the NEBCutter tool:
   http://tools.neb.com/NEBcutter2/
   Check the box to indicate that the plasmid is circular. If you wish to store your query on the NEB site temporarily, give your plasmid a name. ***Click submit.***

The search tool will return results for a bewildering number of REs. The vast majority of the RE sites are not very useful, because the fragments are very large or small, the enzyme is not available in the lab, or the endonuclease is sensitive to DNA methylation.

4. Perform custom digests with REs that look promising and ***are available in the lab.***
   Click the custom digest link. This brings up an interactive chart of REs that cut the plasmid, their recognition sites, the number of recognition sites, and the enzyme activity in four different buffers. The most useful REs for our experiment are those that cut in both the vector and the ORF sequence, generating a small number of fragments. We will be analyzing the restriction fragments on agarose gels, which do well resolving fragments ranging in size from ~200 bp to ~5000 bp. See which of the REs available in the lab are on the list.
5. With your team, prepare a table summarizing the RE predictions for your three plasmids.
   Make a table indicating the sizes of the restriction fragments generated with each RE. It may be helpful to include the length of the *MET* gene coding sequence in the table. One possible organization for the table is shown below. The table uses a fictitious RE named PacMan and notes the lengths of the predicted restriction fragments.

	MET 3	***MET 14***	***MET 16***
size	1533 bp		
Restriction Enzyme			
RE #1 - PacMan	2533, 5573*		
RE #2 - name			
etc.			
as many REs as you like			

*PacMan cuts the vector at 2 positions, generating 2 fragments. Note that the total size of the two fragments is equal to 8106, the sum of the BG1805 sequence and the MET3 coding sequence.

Exercise 2

Restriction digests of plasmid DNA

1. Use the table above to select possible REs for restriction digests.
2. Identify the buffers which give optimum activity with your REs.
   In your lab notebook, note which RE(s) you have decided to use. Check which 10X salt solution that your should be using from the manufacturer's chart of enzyme activity:
   http://www.neb.com/nebecomm/tech_reference/restriction_enzymes/buffer_activity_restriction_enzymes.asp?
   In some cases, it may be necessary to use two REs to generate fragments with acceptable sizes. (Try to avoid using two REs, if possible.) If you are using two enzymes, use the online double digest calculator to find the recommended buffer:
   http://www.neb.com/nebecomm/DoubleDigestCalculator.asp

3. The concentrations of RE and plasmid DNA need to be matched in a restriction digest. Manufacturers assay the activity of each batch of RE and express the activity in units of activity per µL. A unit of activity (U) is assessed in a standardized assay. Restriction digests are usually set up to contain 2-5 U per µg plasmid DNA. This protocol is set up for 5 µL of plasmid miniprep DNA in each reaction, which should correspond to 0.2-0.4 µg DNA. The REs that you will be using have been diluted to a concentration of 0.5 U/µL. Use your Nanodrop data to check the amount of DNA that you would be adding with 5 µL. If the amount of DNA is outside of the desired range, you can adjust the volume of deionized water in the reaction upwards or downwards, as long as the final volume of the reaction is still 10 µL.
4. Set up the reactions to have a final volume of 10 µL (10.1 µL if you're adding BSA). For each reaction, combine in this order:

   ____ µL deionized water (to bring total volume to 10 µL)*
   1.0 µL 10X buffer
   0.1 µL BSA (from a 10 mg/mL 100x stock) - IF recommended
   5.0 µL plasmid - to give a final concentration of 0.2 - 0.4 µg plasmid DNA
   2.0 µL (1.0 U) restriction enzyme

   *If you are doing a double digest with two enzymes, omit the water and add 2.0 µL of the second RE.
5. Add the restriction endonuclease(s) last and mix by gently tapping the tube. Incubate the samples at 37 °C for 2 hr. Store the samples in the freezer for analysis by agarose gel electrophoresis during a future lab.

Exercise 3

Agarose gel electrophoresis of restriction digests

Before starting this lab, read the background material and protocol in Chapter 20 on agarose gel electrophoresis. Additional information is available in the online tutorial on agarose gel electrophoresis.

In this laboratory, you will analyze your plasmid restriction digests by agarose gel electrophoresis. Agarose gel electrophoresis is commonly used to analyze the composition and quality of DNA samples. We will be using 1% agarose gels constituted in TBE, containing SYBR™ Safe as an intercalating agent. These gels provide good resolution of DNA fragments ranging in size from about 500 bp to 7 kb.

Although SYBR Safe has been shown to have very low mutagenicity, wear disposable gloves when working with it. If any spills on your skin, rinse the affected area immediately with copious amounts of water and inform your TA.

1. Plan your gel. Each group of students will prepare one agarose gel. Each student will run one lane with undigested plasmid and a second lane with plasmid that has been digested with restriction endonuclease(s). An additional lane should be reserved for molecular weight standards. Record in your notebook how you plan to load your gels.
2. Prepare, load and run your gels as described in Chapter 20. Be sure to record the details of every step in your lab notebook.
3. Remove your gel from the gel apparatus. Leave the buffer in the apparatus if it will be used again. If you are the last to use the buffer, dispose of it in the appropriate waste container.
4. Photograph your gel on the transilluminator. Send yourself a copy of the gel by email.
5. Dispose of the agarose gels in the appropriate waste container.
6. Calculate the sizes of the DNA fragments in your restriction digests. Sum up the sizes of the fragments to estimate the total length of the plasmid.

   How do the observed sizes agree with your predictions?
   How large is the MET gene encoded by the plasmid?
   Can you identify the MET genes in the three plasmids?

Exercise 4

In this lab, we will use PCR to amplify the *MET* genes that have been cloned into plasmid pBG1805. Since all of the genes have been cloned into the same plasmid, we will be able to use PCR primers based on pBG1805 for all of the reactions.

Before beginning this lab, read the background material and protocol in Chapter 19.

1. The PCR primers have the following sequences:

   Primer 1 - 5' - AAC CAC TTT GTA CAA GAA AGC TGG G - 3'
   Primer 2 - 5' - AAT ATA CCT CTA TAC TTT A AC GTC - 3'

   Use the BLASTN program to find the positions of the primers in the pBG1805 sequence. The PCR procedure will add extra nucleotides to your PCR fragment with the *MET* gene.

   How many nucleotides are added with each of the primers?
   What is the projected size of your PCR fragment?

   Determine how long an extension time at 72˚C will be needed to amplify your genes. As a general rule, allow 1 minute of extension time for each kilobase of length. There will be two thermocyclers set up in the lab. The first thermocycler will use a 1 minutes extension time, and the other thermocycler will use a 4 minute extension time. ***If you have any doubt about the length of your PCR product, use the longer extension time.*** Note: *Taq* polymerase is no a highly processive DNA polymerase. It is less efficient at copying long templates than short templates, so long reactions may be less efficient.
2. Assemble the PCR reactions as described in Chapter 19. Each group should have three PCR reactions and a blank.
3. Run the PCR reactions as described in Chapter 19. When the reaction is finished, store the reactions in the freezer for later analysis on agarose gels.

Exercise 5

Analysis of PCR products by agarose gel electrophoresis

1. In this exercise, you will analyze the results of your PCR experiment on an agarose gel. Each group should have four samples: amplification products from three plasmids as well as a blank. You will also need to include a size standard on your gel. Plan the organization of your gel and record the pattern in your lab notebook.
2. Prepare and run an agarose gel as you did for Exercise 3.
3. Photograph your gel on the transilluminator.
4. Calculate the sizes of your PCR products.

   Are the sizes consistent with your expecctations?
   Can you identify the MET genes in your plasmids?

From the combined exercises in this chapter, you will hopefully have some testable hypotheses about the *met* mutations in your strains and the *MET* genes carried by the plasmids. In the next experiment, you will be able to test your hypotheses by complementation.

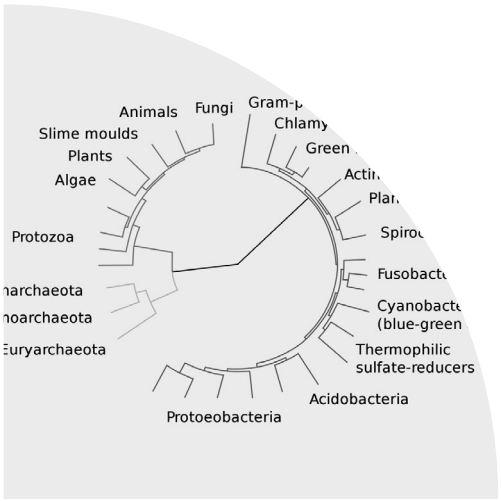
References

Gelperin, D.M. White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., Gerstein, M., Dumont, M.E., Phizicky, E.M., Snyder, M., and Grayhack, E.J. 2005. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop.* **19**: 2816-2826.

8 Protein Conservation

Objectives

- Learn the one letter code for amino acids
- Understand the functionalities of amino acid side chains
- Understand how amino acid side chain chemistry is reflected in the BLOSUM62 matrix
- Use BLASTP to find phylogenetically diverse Met proteins
- Construct a phylogenetic tree of Met proteins within the Ascomycota



The function of a protein is determined by its primary sequence, which folds into a unique, three-dimensional conformation. In this lab, you will learn the 1-letter code for amino acids, an essential part of a modern biologist's vocabulary. You will also use amino acid sequences for *S. cerevisiae* Met proteins to query protein databases, searching for homologous proteins within the Ascomycota. You will then use those sequences to construct a phylogenetic tree for your Met protein.

Chapter 8

Introduction

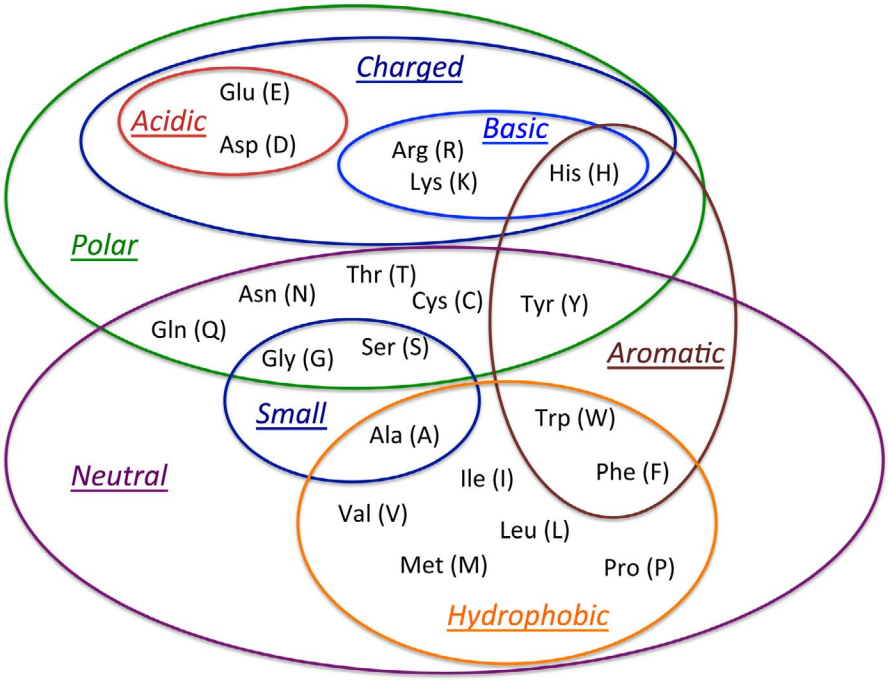
As we look over species that have diverged from one another over evolutionary time, we see that protein sequences have changed. The rate that an individual protein changes is widely variant, reflecting the evolutionary pressures that organisms experience. Our goal this semester is to determine if genes for methionine biosynthesis are functionally conserved between *S. pombe* and *S. cerevisiae*, two species separated by close to a billion years of evolution. Protein function is intimately related to its structure. When the function of an enzyme is conserved between species, one expects to find significant conservation in the three-dimensional structure of the proteins. You'll recall that the final folded form of a protein is determined by its primary sequence, the sequence of amino acids. Strongly conserved proteins generally contain regions with closely related amino acid sequences.

In this lab, we will use information in protein databases to identify potential homologs of the methionine biosynthetic proteins from *S. cerevisiae*. To use the protein databases, it is essential to first learn the one letter code for amino acids. We will also study the scoring matrices used to query protein databases and see how evolution has favored substitutions that involve amino acids with similar chemistries.

Before starting this lab, read the material in Chapter 25 on the BLASTP algorithm.

The side chains of amino acids have distinct chemistries

Each of the 20 amino acids commonly found in proteins has a distinctive R group with its own unique chemistry. R groups differ in their size, polarity, charge and bonding potentials. When thinking about evolutionary changes in proteins, it is often convenient to group the amino acids within overlapping sectors of a Venn diagram, such as shown on the opposite page. In general, changing one amino acid in a given sector into a different one within the same sector can be considered a largely conservative change. Watch out for size, however! R groups vary considerably in the bulkiness of their chains. Replacing a small side chain with a large side chain may seriously perturb the structure of a folded protein.



Functional groupings of amino acids

Exercise 1

Become acquainted with the 1-letter code for amino acids

Met-Glu-Asn-Asp-Glu-Leu-Pro-Ile-Cys-Lys-Glu-Asp-Pro-Glu-Cys-Lys-Glu-Asp

What is the net charge of this peptide? (Assign -1 for each acidic amino acid and +1 for each basic amino acid. Add up the total charges.)

How many hydrophobic amino acids are found in this peptide?

Write the name of a music group that you enjoy. Then transpose the name into an amino acid sequence. Pass the amino acid sequence to a friend and have him/her decode it. (Note: the 1-letter code uses all of the alphabet, except B, J, O, U, X and Z

Exercise 2

Bioinformaticists use the BLOSUM 62 matrix to compare divergent proteins

Another way to look at protein evolution is to compare the frequency of amino acid substitutions that have appeared in naturally occurring proteins. Statisticians first set a threshold for protein identity and then compare the frequencies at which pairs of amino acids have substituted for one another in protein sequences above the threshold. The most common amino acid substitutions are given higher positive values, while very rare substitutions are given the most negative values. The threshold of 62% amino acid identity was used to generate the popular BLOSUM 62 matrix shown below. In general, the most high scoring substitutions are ones that would be considered conservative from the Venn diagram on the previous page. There are some notable discrepancies, however, which you may wish to ponder.

Note how few of the values are positive! Most substitutions are likely to disrupt protein function.

BLOSUM 62 scoring matrix

Ala	**A**	4																			
Arg	**R**	-1	5																		
Asn	**N**	-2	0	6																	
Asp	**D**	-2	-2	1	6																
Cys	**C**	0	-3	-3	-3	9															
Gln	**Q**	-1	1	0	0	-3	5														
Glu	**E**	-1	0	0	2	-4	2	5													
Gly	**G**	0	-2	0	-1	-3	-2	-2	6												
His	**H**	-2	0	1	-1	-3	0	0	-2	8											
Ile	**I**	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
Leu	**L**	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
Lys	**K**	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
Met	**M**	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
Phe	**F**	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
Pro	**P**	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
Ser	**S**	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
Thr	**T**	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
Trp	**W**	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Tyr	**Y**	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
Val	**V**	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4
		A	**R**	**N**	**D**	**C**	**Q**	**E**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**

Henikoff, S. and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**:10915-10919.

NCBI's Amino Acid Explorer is a rich source of information on amino acids
Access the Explorere at: http://www.ncbi.nlm.nih.gov/Class/Structure/aa/aa_explorer.cgi

1. Click on the biochemical properties to obtain a table that summarizes many valuable facts about the amino acids.

   What are the three most common amino acids in proteins?

   What do these amino acids have in common?

   Which two amino acids are least common in proteins? Why do you think this is?

2. Use the common substitutions matrix in the Amino Acid Explorer to see which amino acids most frequently substitute for leucine. Note the BLOSUM62 color scale at the top of the page.

   Which common leucine substitutions have positive BLOSUM62 scores?

   What do these amino acids have in common?

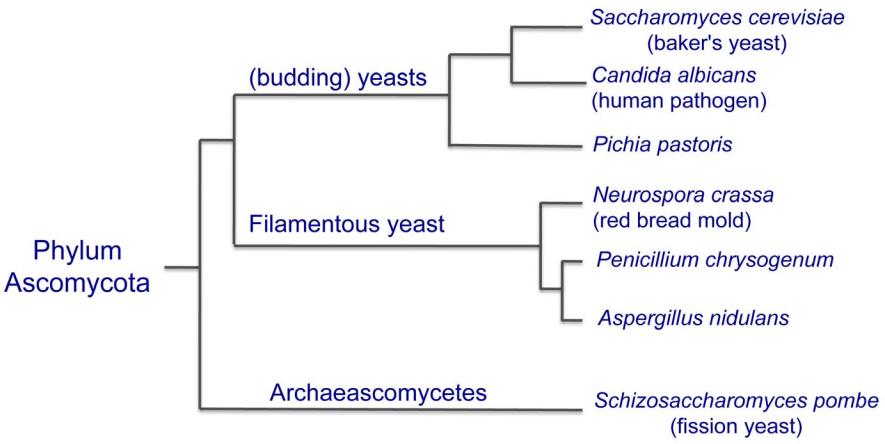
Look at the common subsitutions for cysteine. How do the results differ from the results for leucine? Why do you think this is so?

Exercise 3

BLASTP searches of the NCBI database

In this exercise, you will use the *S. cerevisiae* protein sequence in BLASTP searches to find homologs within the Ascomycota. The figure below shows one possible phylogenetic tree for 7 sac-forming yeast based on DNA sequences. In this exercise, you will collect records for these common yeast species, and we'll use them to build a phylogenetic tree in the next exercise.

- Go the BLASTP home page at NCBI: **http://blast.ncbi.nlm.nih.gov/Blast.cgi**
- Under Basic BLAST, click the protein blast link
- Use the Ref Seq number (NP_xxx) for your *S. cerevisiae* Met protein as the query sequence
- For the database, choose reference protein (refseq_protein)
- In the organism box, enter one of the species in the figure below.
- Click the BLAST button
- Complete the grid on following page with output information from the searches.



Accession number	Species	Bit score (maximum)	Sequence Coverage	E-Value
NP_	*Saccharomyces cerevisiae*			
	Candida albicans			
	Pichia pastoris			
	Neurospora crassa			
	Penicillium chrysogenum			
	Aspergillus nidulans			
	Schizosaccharomyces pombe			

Repeat some of the BLASTP searches using the genus name, rather than the species name, in the organism box. *What differences do you see in the results? Why do you think this is?*

Exercise 4

Construct a phylogenetic tree of your Met protein

1. Open a new tab or window in your browser. You will be working with material at two different sites, so you need two operational browser pages. Direct a browser page to http://www.phylogeny.fr. This site contains a suite of programs that you will use to compare your sequences. Click the "Your workspace" tab and set up an account to store your work.
2. Under the heading Phylogeny analysis, select One Click. Your sequences will be automatically brought through multiple alignment and phylogenetic tree building algorithms. The advanced option on this page would allow you to adjust the parameters associated with each program. *We'll let Phylogeny make the decisions for us!*
3. The first step in the process is to construct a FASTA file containing all the sequences that you would like to compare. You will be pasting files from NCBI records into the Phylogeny data entry box. The title line of a FASTA file begins with a ">" character and end with a hard return. The title lines of NCBI FASTA files contain a GenBank reference number (gi_____), the RefSeq accession number (NP_xxxxxx), protein name and the species name in brackets.

   The first file in our comparisons will be the *S. cerevisiae* MetXp sequence.
   - Direct your browser to the NCBI Protein Database
   - Use the RefSeq accession number to retrieve the protein record.
   - Click the FASTA link at the upper left side of the record.
   - Copy the title line (begins with >) and the ***entire*** amino acid sequence
   - Paste the FASTA record into the Phylogeny data entry box
4. Repeat step 3 with each of the sequences that you would like to compare.
5. When you are finished, give your project a title, enter your email address (the analysis can take a little time) and click the Submit button. Your results will be posted on a web page when they are ready.
6. Work with your results. A link will bring you to output data from the multiple sequence alignment and tree-building programs. Separate tabs will allow you to view the outputs of each program. The last tab brings you to a graphics program that allows you to manipulate the appearance of your tree. Adjust the look of your tree. When you are satisfied with it, copy it to a program that handles images or download the file to your computer.

   How does your tree compare to the lineage shown earlier in this chapter?

Exercise 5

Construct a WebLogo for your sequences

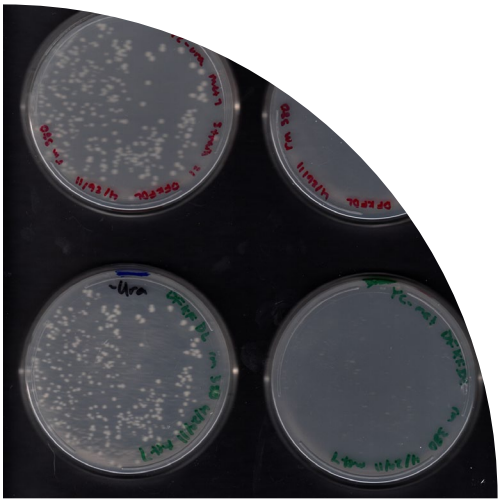
WebLogos provide graphical views of the multiple sequence alignments that are very useful in identifying strongly conseved regions in a protein.

1. Open a new browser window or tab. Direct the browser to http://weblogo.berkeley.edu
2. Read the Introduction and click the Create link.
3. Paste the FASTA file for the multiple sequence alignment into the entry box. ***This is NOT a list of individual FASTA files pasted together!*** Find the multiple sequence alignment FASTA file by returning to your project at Phylogeny.fr.
   - Click on the Alignment tab, which brings up your aligned sequences in a shaded box.
   - At the bottom of the page under output options, click the "Alignment in Fasta format" link.
   - A page with the FASTA file appears. Note the large number of dashes that have been introduced to account for gaps in the alignments. Copy the contents into the WebLogo data entry box.
4. Under Advanced Logo Options, check the Multiline Logo option. For starters, you may want to use the 32 symbols per line - this can be increased later, but letters will be smaller. Click Create Logo.
5. Your logo appears quickly. Use the magnifier to increase the size of the letters. The output should give you a good sense of conserved regions. How does this data compare with the domain information that you collected in Lab 5? You may want to adjust the number of symbols per line or image options. When you are satisfied, paste or save the logo for further use.

9 Plasmid Complementation

Objectives

- Prepare transformation-competent yeast cells
- Transform yeast strains with plasmid vectors carrying *MET* genes
- Select transformed cells on selective media
- Identify mutated *met* genes in yeast strains by complementation



Complementation is a powerful tool for identifying genes. In this lab, you will identify the gene mutated in an unknown *met* strain using plasmid complementation. This experiment builds upon the data that you collected from the selective plating and plasmid mapping experiments of the past few weeks.

Chapter 9

Introduction

In this lab, you will use plasmid complementation to test your hyptheses about the identities of your unknown yeast strains and plasmids. Each group has analyzed the growth properties of three different *met* strains on selective media, which have provided some clues to the positions of the genes in the process of methionine biosynthesis. Each group has also used restriction mapping and PCR to analyze the three overexpression plasmids that contain the coding sequences of the three deleted *MET* genes. From these experiments, you may have already formulated some hypotheses about which strain is missing the gene carried in each plasmid. In this lab, you will transform each of the three plasmids into one of the yeast *met* strains. All of the overexpression plasmids carry a *URA3* gene, so all of the transformed strains will be able to grow on selective plates that lack uracil, but contain methionine. The only transformed strains that will be able to grow on plates lacking methionine are those in which the overexpressed *MET* gene is the same as the gene deleted in the host strain.

Before starting this lab, read Chapter 22 on Yeast Transformation.

In the first part of this lab, you will transform the three yeast strains with your plasmids and isolate transformants on media lacking uracil. The ability of the transformed cells to grow in the absence of uracil is due to complementation. The BG1805-based plasmids all carry a normal copy of the yeast *URA3* gene, including its promoter, so the gene is regulated much like a normal chromosomal gene. Our yeast strains are derived from BY4742, which has the *ura3Δ0* allele. Uracil selection depends on complementation. The Ura3p protein produced from the plasmid *URA3* gene compensates for the *ura3* deletion in the yeast chromosome, allowing transformed cells to grow in the absence of uracil.

In the second part of the lab, you will perform another complementation experiment, in which you test the ability of the transformed strains to grow in the absence of methionine. For complementation to occur, the *MET* gene introduced by the plasmid must be identical to the mutant chromosomal gene. The *MET* genes in the plasmids are under the control of the powerful *GAL1* promoter. The *GAL1* promoter is induced by the presence of galactose in the culture medium. Because the *GAL1* promoter is so powerful, however, some transcription may occur even in the absence of galactose. We will therefore analyze the growth of transformants on spot plates containing either glucose or galactose. Yeast have a strong preference for glucose as a carbon source, so galactose may have some effects on growth that are unrelated to *MET* gene expression. To ensure that any growth differences that we observe are due to the presence of the plasmid, rather than the carbon source, we will also grow the parental strains on both glucose and galactose.

Exercise 1

Transformation

Refer to the tranformation protocol in Chapter 22 for important details of the procedure.

1. Label three microcentrifuge tubes with a code of your own devising so that you are able to identify the yeast strain and the plasmid used for each transformation. ***Be sure to record the code in your lab notebook!*** Label a fourth tube "Blank."
2. Prepare a transformation master mix.
3. Set up a transformation mixture for each of the three strain/plasmid combinations.
4. Set up a blank transformation in which you substitute 5 µL water for plasmid DNA. Choose one of your three strains for the blank transformation. Record the name of the strain that you used for the blank in your notebook.
5. Transform the strains and plate the transformants on YC-Ura plates.
6. Determine the numbers of viable cells in your transformation mixtures using a spot plate on YPD media.
7. Calculate the transformation efficiency of each reaction.

Did you note any differences between the strains in the transformation reactions?

Exercise 2

MET gene complementation

Late in afternoon preceding the laboratory, begin overnight cultures of your transformed yeast strains.

1. Obtain three culture tubes, each containing 2 mL of YC-Ura. Label each of the tubes with the name of the strain that will be grown in it.
2. Inoculate each tube with the appropriate strain. To do this, sterilize an inoculation loop in a Bunsen burner. Before picking up a colony from a plate, touch the loop to an unoccupied area on the YC-Ura plate to cool the loop. Next, inoculate the culture medium with a single transformed colony.
3. Incubate the culture at 30 ˚C on either the rotating wheel in the warm room or on the shaker in your lab room.

NOTE: Remember to turn the wheel back on after removing your samples!

Next laboratory session

1. Prepare a series of five 10-fold dilutions of each overnight cultures with sterile water.
2. Prepare three different spot plates with the three dilution series. Each row on the spot plate will represent a different transformation reaction. Use the same pattern of rows on each of the three plates. ***Record the order that the strains are spotted on the plates in your notebook.*** The cells will tend to settle to the bottom of the tube between platings. Make sure that the cells are uniformly suspended before you make each spot. Label each plate with your initials and the date.

   Plate 1 - YC minus Met containing glucose
   Plate 2 - YC minus Met containing galactose
   Plate 3 - YC minus Ura containing glucose

   Why are we spotting the samples on these three plates? Wouldn't one plate be enough?
3. Prepare two similar spot plates with an untransformed parental strain on YC complete media. Your TA will give you an aliquot of an overnight culture of one of your unknown strains in YC-complete.

   Plate 1 - YC complete containing glucose
   Plate 2 - YC complete containing galactose

   NOTE: Two groups should share the YC-complete spot plates.
4. Incubate the plates at 30˚C until colonies appear.

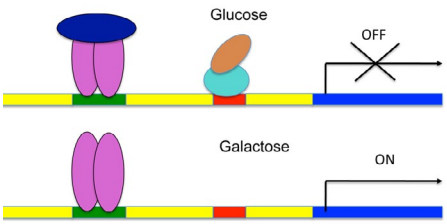
   Are you able to identify your mutant strain from the spot plates?
   How does growth on glucose compare to growth on galactose?
   Do you see any differences in growth rates between parental and transformed strains?
5. Record your data using the scanner. Remove the covers from each plate and place them upside down on the scanner. (Keep the tops of the plates upside down on the bench while you do this.) Make sure that all of the plates are in the same orientation.

Use this complementation data, together with your data from the earlier experiments to identify your strain!

10 Protein Overexpression

Objectives

- Understand regulation of the yeast *GAL* promoter.
- Culture yeast with different carbon sources to induce and repress expression from the *GAL* promoter
- Prepare extracts from yeast grown under repressed and induced conditions.



In this lab, you will use various carbon sources to manipulate the expression of Met fusion proteins in cells that have been transformed by overexpression plasmids. You will then prepare extracts from these cells for protein analysis.

Chapter 10

Introduction

In this lab, you will analyze the expression of *S. cerevisiae* Met fusion proteins in your transformed strains under various conditions. Your goal is to determine if the cloned Met fusion proteins can replace the enzymes that are missing in the *S. cerevisiae met* mutants. Complementation depends on the presence of functional Met proteins. The fusion proteins expressed from the pBG1805 plasmid have three tags added to the C-termini of the naturally occurring proteins. The biochemical activities of these fusion proteins have not been evaluated. This experiment will tell you if the proteins can replace the missing Met proteins, but you will not be able to tell if the fusionproteins are less active than the naturally occurring proteins. *(Think about how one might do this.)* The fusion sequences are very convenient for identifying the fusion proteins on SDS-PAGE gels and western blots, since they encode epitopes that can be recognized by antibodies. It is especially convenient that all of our fusion proteins have the same tags at their 3' ends!

Protein expression from the pBG1805 and pYES2.1 plasmids is controlled by the powerful S. cerevisiae *GAL1* promoter. The *GAL1* promoter is an inducible promoter that is positively regulated by galactose and negatively regulated by glucose (Johnston, 1987). In yeast, as in most eukaryotic cells, glycolysis plays a major role in energy production, and glucose is the preferred carbon source. Genes involved in the metabolism of other carbon sources are usually repressed when glucose is available. Glucose repression is a complex phenomenon that includes the binding of repressor proteins to control regions of these genes. When glucose is not available, however, yeast activate genes that metabolize other energy sources. Galactose increases the transcription of several genes involved in the conversion of galactose into glucose-6-phosphate (G6P), an intermediate in glycolysis. The *GAL1* gene encodes galactokinase, the first enzyme in the pathway that converts galactose into G6P. In the presence of galactose, expression of the *GAL1* gene increases ~1000-fold above the level observed in the presence of glucose.

In this lab, you will grow transformed *met* strains grown with three different carbon sources; glucose, raffinose, and galactose. Raffinose is a trisaccharide composed of galactose, fructose and glucose monomers. Raffinose relieves glucose repression, but does not activate expression from other inducible promoters. For this reason, raffinose is commonly used to "wean" yeast from their dependence on glucose. In this experiment, we will grow the transformed strains in medium containing raffinose over several nights. Cultures will then be given glucose or galactose for a few hours before class to manipulate expression of the Met proteins from the *GAL1* promoter. We will then prepare cell extracts for protein analysis on SDS-PAGE gels (Chapter 23). The procedure for preparing extracts is a simple one in which cells are first weakened by a brief treatment with NaOH, after which the cells are lysed by boiling them in SDS-PAGE sample buffer. SDS, or sodium dodecyl sulfate, is an anionic detergent that denatures all the proteins in a cell and surrounds them with a negative charge.

Read the background information in Chapter 21 before starting this lab.

Prepare cell cultures

First lab session of the week

1. Use an inoculation loop to transform cells from one transformed colony to 1 mL of YC-URA containing raffinose.
2. Repeat step 1 for each of your transformed strains.
3. Place the cultures on the wheel, taking care that tubes are properly balanced. Turn the wheel on.

~Four hours before the next lab session

1. Remove your cultures from the wheel. Divide the culture into three equal portions in new culture tubes.
2. To each culture, add 0.5 volume of YP containing either 6% glucose, 6% galactose, or 6% raffinose. Be sure that the tubes are labeled appropriately labeled. Place the cultures back on the wheel until class.

Don't forget to turn the wheel back on after removing or adding cultures!!

Determine cell densities

Ideally, all of your cell extracts will contain the same amount of protein, which will make it easier to compare differences between the proteins from each of the extracts. In this experiment, cells have been cultured with three different carbon sources. Although the size (and protein content) of a cell is not expected to be affected by the carbon source, cells may have grown at different rates under the three conditions. Therefore, it is imporant to get a rough idea of the cell density in each culture before preparing the cell extracts. This can be done by measuring the light scattering activity of each culture in the spectrophotometer.

NOTE: Yeast cells will settle out of cultures when they are not agitated. Be sure that the cells are uniformly suspended in the cultures before taking samples for spectrophotometer readings.

1. Transfer 100 µL of each culture to 900 µL of deionized water in a test tube.
2. Read the OD_{600} of the diluted cell suspension in the spectrophotometer. Record these values in your lab notebook.
3. Calculate the volume of cell culture that will be required to obtain 2.5 OD_{600} equivalents of cells.

Prepare cell extracts

1. Transfer the volume from each culture that you will need to obtain 2.5 OD_{600}s of cells to a fresh test tube.
2. Follow the protocol in Chapter 21 for preparing cell extracts. Be sure to write down the steps in the procedure into your lab notebook.

References

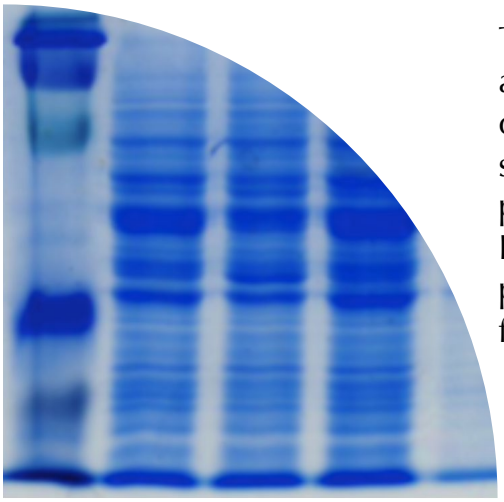
Johnston, M. 1987. A model fungal gene regulatory mechanism: the *GAL* genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.* **51**: 458-476.

11

Protein Electrophoresis

Objectives

- Assemble and cast SDS-PAGE gels
- Separate proteins in yeast extracts using electrophoresis.
- Compare proteins in different extracts on stained SDS-PAGE gels.
- Use densitometry to quantify differences in protein expression



This lab will introduce you to SDS-PAGE, a simple and inexpensive method for resolving proteins in complex mixtures. SDS-PAGE gels provide the starting materials for western blots and for some proteomic techniques. In this lab, you'll use SDS-PAGE to analyze the protein extracts that you prepared from yeast strains overexpressing Met fusion proteins.

Introduction

In today's laboratory, you'll use SDS-PAGE to analyze the proteins in yeast extracts. In the last lab, you prepared extracts from an untransformed "parental" *met* mutant as well as *met* strains that had been transformed with either a pBG1805-*GAL1:METx* overexpression plasmid. The pBG1805 plasmids were constructed to overexpress *S. cerevisiae* Met proteins. In the last lab, you prepared extracts from cells that were grown under uninduced (raffinose), repressed (glucose) or induced conditions (galactose). In today's lab, you will construct SDS-PAGE gels and separate the proteins in the extracts by electrophoresis. The gels will be stained overnight with Brilliant Blue G-250 dye and then analyzed by densitometry of the stained image of the gel. The gels should reveal major differences in MetXp expression between the strains and responses of individual strains to carbon sources.

As you analyze your results, it is important to keep in mind that yeast cells express thousands of proteins, and significant changes in the expression of a particular protein may be masked by the many other proteins detected with Brilliant Blue staining. In the next lab, we'll use antibodies to specifically analyze Metxp expression in the various extracts on western blots. Today's lab will provide important information about cellular protein expression that will be useful for the western blots. In this lab, you'll make two SDS-PAGE gels. If all goes well, you'll be able to save one of the gels for western blots in the next lab.

Before starting this lab, read the background information on SDS-PAGE in Chapter 23 and review the SDS-PAGE tutorial.

Plan the SDS-PAGE Gel

Before running your gel, you'll first need to determine the predicted sizes of your MetXp fusion proteins. The pBG1805 plasmids encode fusion proteins that add significant numbers of amino acids to the C-termini of the Met proteins. The BG1805 plasmid encodes HA and His_6 epitopes as well as an immunoglobin binding domain. Together these sequences add a walloping ~19 kDa to the expected mass of S. *cerevisiae* Met proteins (Gelperin et al., 2005). Consult the plasmid information at the class website to compute the expected size of your fusion proteins. Record the calculations in your lab notebook.

You'll next need to decide if you should use a 12% or 15% gel to separate your proteins.. Figure 1 shows a set of standard proteins that have been separated on 12% and 15% SDS-PAGE gels, together with the MWs of the proteins. Use this information to decide whether a 12% or 15% SDS-PAGE gel will give you the best resolution for proteins with a similar molecular weights to your Met fusion proteins.

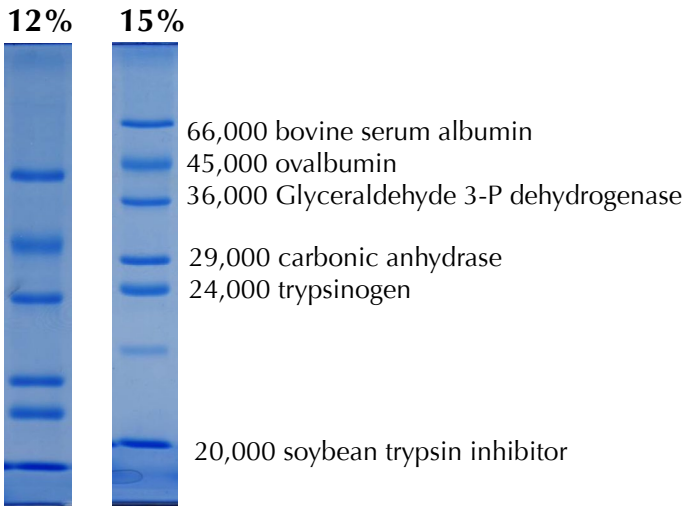


Figure 1.
Separation of MW standards on 12% and 15% SDS-PAGE gels

Cast SDS-PAGE Gels

SAFETY NOTE: Acrylamide and bisacrylamide monomers are neurotoxins. Gloves, goggles and aprons should be used when working with acrylamide.

Follow the protocol in Chapter 23 to cast two SDS-PAGE running (resolving) gels with the acrylamide concentration that you selected in Exercise 1. You will be casting two gels, although you will only be using one for this experiment. The second gel provides insurance against leaks! If the first gel leaks, you will have a second gel to work with. If not, you can save the second gel for the western blot in our next experiment. (You'll need to pour the stacking gel just before you use it.) Wrap any gel that you're saving in plastic wrap, write your initials on a piece of labeling tape, and store the gel upright in the refrigerator for the next lab.

Decant the water overlay from the first running gel and pour a stacking gel. After the stacking gel polymerizes, you're ready to run the gels. Two groups will share each electrode apparatus and electrophoresis tank. Remove the comb from your stacking gel and position the gel on one side of the electrode apparatus. After both groups have positioned their gels in the apparatus, place the gel assemblies in the electrophoresis tank and add SDS-PAGE to the upper and lower reservoirs, as described in Chapter 23.

Run and Stain SDS-PAGE Gels

With your group, decide which samples will be loaded in each lane of the gel. Each gel should have one lane with MW standards. You should have nine different samples: uninduced, repressed and induced cell extracts from each of your three strains. As you plan your gel, keep in mind that it's easiest to compare proteins in extracts that have been run in adjacent gel lanes. You also want to consider the position of the MW standard on the gel. Lanes containing cell extracts are likely to look very similar after the gels are run. Placing the MW standards in an asymmetric position on the gel can be helpful in distinguishing the order of lanes on the gel. ***Be sure to record which samples are in each lane in your lab notebook.*** Run and stain the gels as described in Chapter 23.

Analyze SDS-PAGE Gels

Record an image of your stained gel with either a camera or the scanner. Carefully inspect your gel and look for differences in protein expression between the lanes.

Can you detect the expression of Met fusion proteins in extracts of transformed cells?
How do the molecular weights of fusion proteins compare with the expected values?
Can you detect any differences in induced cell extracts compared to the induced controls?

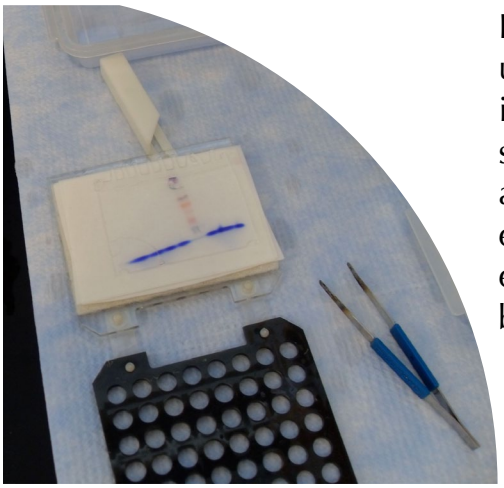
References

Gelperin, D.M. White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., Gerstein, M., Dumont, M.E., Phizicky, E.M., Snyder, M., and Grayhack, E.J. 2005. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop.* **19**: 2816-2826.

12 Immunochemical Detection

Objectives

- Transfer proteins from an SDS-PAGE gel to a PVDF membrane
- Use primary and secondary antibodies to identify epitope-tagged proteins on western blots
- Use western blots to determine the effects of carbon sources on expression from the yeast GAL1 promoter



Immunoblots, also known as western blots, make use of antibodies to identify proteins of interest in complex mixtures of proteins that have been separated on SDS-PAGE gels. In this lab, you will analyze Met protein expression in the yeast cell extracts on western blots, taking advantage of the epitope tags that were introduced into the proteins by molecular cloning.

Chapter 12

Introduction

In the previous lab, you ran an SDS-PAGE gel that enabled you to visualize the proteins in your yeast cell extracts. You may or may not have been able to identify proteins on the stained gel with the molecular weight (MW) expected of your Met fusion protein. In this lab, you will use western blots to identify overexpressed Met fusion proteins in the extracts, using antibodies that recognize hemagglutinin (HA) epitope tags. The experiment will take place over several lab sessions.

In the first lab session, you will run a new SDS-PAGE gel and use an electric current to transfer proteins from the unstained gel to a PVDF (essentially Teflon) membrane. Proteins adhere tightly to this membrane, which can either be used directly for antibody visualization or be dried and stored for later analysis. Before the second lab session, you will block the membrane and apply the primary antibody. In the second lab session, you will visualize the bound antibody on the membrane.

Read the material in Chapter 24 before beginning this experiment.

Plan the western blot

Before beginning this experiment, plan your SDS-PAGE gel. One lane of your gel should contain the Kaleidoscope™ standards. Next, strategize which samples to load on the gel and the order in which the samples should be loaded on the gel. You should have extracts from three different transformed strains grown with three different carbon sources. The three sugars are expected to have different effects on transcription of the plasmid genes.

You have nine gel lanes available for samples. Decide with your group which extracts should be compared to each other.

What results do you expect to see with each extract?
Is it important to analyze all of the extracts?
Which comparisons are the most meaningful?

Your stained SDS-PAGE gel from the last lab may have already yielded some relevant information about the protein composition of your extracts. *Did you see any interesting differences that you would like to pursue?*

Separate proteins by SDS-PAGE

1. During the last lab, you prepared an SDS-PAGE resolving gel. The gel has been stored in the refrigerator wrapped in plastic, with a small amount of water over the gel to prevent the interface from drying out. Unwrap the gel and remove the water overlay with a paper towel or Kimwipe™.
2. Place the gel in a casting frame and clamp the assembly into the casting stand.
3. Pour the stacking gel like you did in the last lab. ***Combine:***
   - 1.40 mL deionized water
   - 0.675 mL 0.5 M Tris-HCl, pH 6.8
   - 0.42 mL 30% acrylamide:bisacrylamide (37.5:1)

   To catalyze polymerization, add and gently mix:
   - 20 µL 10% ammonium persulfate
   - 5 µL TEMED
4. When polymerization is complete, place the gel in the electrophoresis tank, add samples and run the gel as you did in the last lab. Be careful to note which samples are loaded in each lane. Include Kaleiodoscope molecular weight standards in one lane.

NOTE: Do NOT stain the gel when it is finished. Staining fixes proteins in the gel and renders them non-reactive with antibodies!

Detect epitope tags on western blots

1. Follow the protocol in Chapter 24 for transferring proteins from the gel to the PVDF membrane.
2. Wearing gloves, disassemble the transfer apparatus. Try not to let the membrane dry out. If the membrane dries out, it will need to be re-wet in methanol.
3. Transfer the membrane to a tray containing TBS-T. Make a careful note of the Kaleidoscope standards bound to the membrane. (This will be the only lane with proteins visible.) The side of membrane with bound standards should be facing up in the tray. Place a small piece of tape with your initials on the tray.
4. Proceed with the antibody binding and washing steps as described in Chapter 24.
5. Allow the blots to dry on a piece of filter paper.

Chapter 12

Analyze western blots

Record the data using with the scanner.
Use ImageJ to analyze the blot.

Questions to consider:

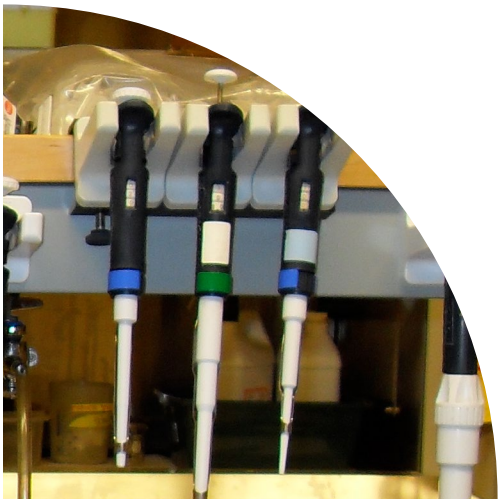
- Which extracts contain epitope-tagged proteins?
- How do the sizes of the tagged proteins compare with the expected values?
- How do carbon sources affect the amount of Met protein expression?
- How do the western blot results compare with those on the SDS-PAGE gels?

Part II
Theory and Techniques

13 Laboratory Equipment

Objectives

- Understand how to use micropipettes accurately.
- Understand how to use a spectrophotometer to measure absorbance
- Understand how to use a compound light microscope
- Understand how to use a centrifuge



The proper use of laboratory instruments is essential for successful experiments. This chapter will introduce you to our laboratory equipment. Instruments covered in this chapter include micropipettes, the Leica DM500 light microscope, the Genesys 20 spectrophotometer and benchtop centrifuges.

Chapter 13

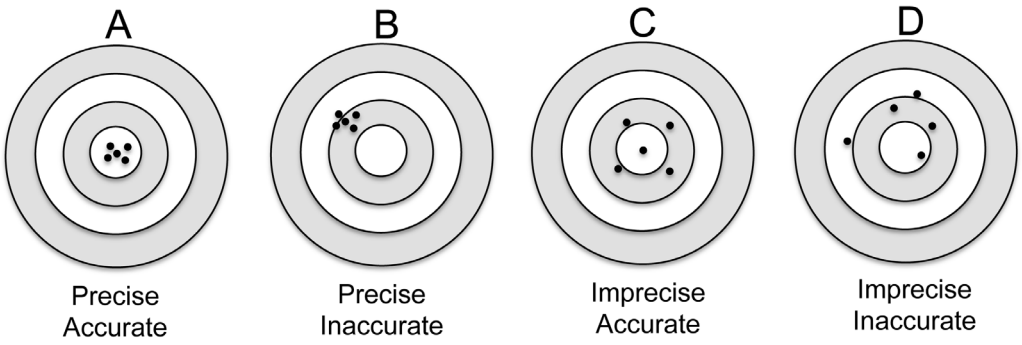
Micropipettes

Arguably, the most important scientific equipment that you will use this semester are adjustable micropipettes. Micropipettes are precision instruments that are designed to accurately transfer volumes in the microliter range. You may use microliters or milliliters as the units of volume in your lab notebooks and lab reports, but be careful to always state the volume unit that you're using. Recall the relationships between volume units:

1 microliter (abbreviated μL) = 10^{-3} milliliter (mL) or 10^{-6} liter (L)
(A useful tip for Mac users: The keyboard shortcut for the Greek letter μ is Alt-m)

Accuracy and precision

Ideally, micropipettes will deliver liquids with accuracy and precision. Accuracy depends on the micropipette delivering the correct volume. Precise results are reproducible. Let's use a target analogy to demonstrate the difference between accurate and precise results. Imagine that four students try to hit the bulls-eye five times. Students A and B are precise, while students A and C are accurate.



The best way to determine the accuracy and precision of micropipettes is to use them to weigh set volumes of distilled water on an analytical balance. The density of water is 1.0 gram per mL at 25˚C. The process is repeated several times during the calibration process, and the data is used to calculate the accuracy and precision of a micropipette.

Accuracy refers to the performance of the micropipette relative to a standard (the intended) value. It is computed from the difference between the actual volume dispensed by the micropipette and the selected (intended) volume. Note that this can be a negative or positive value. When micropipettes are calibrated, the accuracy is normally expressed as a percent of the selected value. In general, micropipettes are designed to operate with accuracies within a few percent (generally <3%) of the intended value. The accuracy of a micropipette decreases somewhat, however, when micropipettes are set to deliver volumes close to the lowest values in their range.

Precision provides information about reproducibility, without any reference to a standard. Precision reflects random errors that can never be entirely eliminated from a procedure. Precision is expressed as the **standard deviation** of the measurements. Like accuracy, precision is often expressed as a percent of the selected volume in calibration documents.

Choosing the micropipette

There are three different sizes of micropipettes in the laboratory, which we will refer to as the P20, P200 and P1000. Our micropipettes have been purchased from several different manufacturers, but the principles of operation are the same. The numbers after the "P" refer to the maximum number of microliters that the micropipette is designed to transfer.

Use the chart below to select the correct micropipette for an operation. Note that there is some overlap in the ranges of the different micropipettes. For example, both the p200 and p20 can be used to transfer 15 µl, but the p20 is more accurate within that range. ***As a rule of thumb, always select the smallest volume pipette that will transfer the volume, since the measurement will be more precise.***

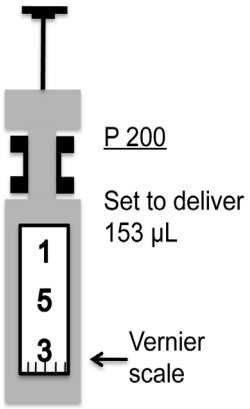
Micropipette	Recommended range (µL)	Smallest increment (µL)
P20	1 - 20	0.02
P200	20 - 200	0.2
P1000	100 - 1000	2.0

Micropipettes use disposable plastic tips. The P1000 tips are larger than those used with P200s and P20s. P1000 tips may be either white or blue, depending on the vendor, while P20 and P200 tips may be either yellow or natural in color.

Specifying the transfer volume

There are three numbers on the volume indicator. With each of the micropipettes, you will specify a volume to three digits by turning the volume adjustment knob. You will also be able to extrapolate between the lowest numbers with the vernier marks on the lower dial. Most of the measurements you will make with the micropipettes will be accurate to four significant figures!

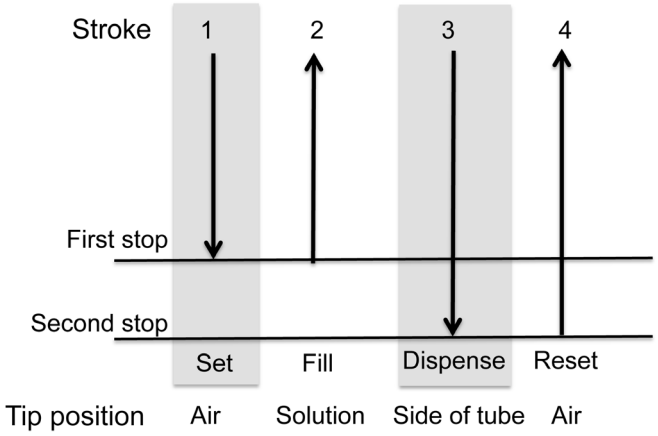
NEVER turn the indicator dial beyond the upper or lower volume limits of the micropipette! This could damage the piston.



Transferring volumes accurately

Micropipettes work by air displacement. The operator depresses a plunger that moves an internal piston to one of two different positions. The first stop is used to fill the micropipette tip, and the second stop is used to dispense the contents of the tip. As the operator depresses the plunger to the first stop, an internal piston displaces a volume of air equal to the volume shown on the volume indicator dial. The second stop is used only to dispense the contents of the tip.

The transfer of a solution with a micropipette can be summarized in four strokes, as shown in the figure and described in more detail below.



Filling the micropipette

1. Remove the lid from the box containing the correct micropipette tip. Try to avoid touching the tip (especially the business end), because the tips are sterile. To attach the tip, insert the shaft of the micropipette into the tip and press down firmly. This should produce an airtight seal between the tip and the shaft of the micropipette. Replace the lid of the tip box to keep the remaining tips sterile.
2. Depress the plunger of the micropipette to the FIRST stop.
3. Immerse the tip a few millimeters below the surface of the solution being drawn up into the pipette. Pipetting is most accurate when the pipette is held vertically. Keep the angle less than 20˚ from vertical for best results.
4. Release the plunger S L O W L Y, allowing the tip to fill smoothly. Pause briefly to ensure that the full volume of sample has entered the tip. Do NOT let the plunger snap up. This is particularly important when transferring larger volumes, because a splash could contaminate the shaft of the micropipette. If you inadvertently contaminate the shaft, clean it immediately with a damp Kimwipe.
5. Dispense the contents of the micropipette by following the steps below. NEVER rest a micropipette with fluid in its tip on the bench!

Dispensing the contents of the micropipette

1. Place the micropipette tip against the side of the receiving test tube. Surface tension will help to dispense the contents of the micropipette. Do NOT attempt to eject the contents of the micropipette into "thin air. "
2. Smoothly depress the plunger to the first stop. Pause, then depress the plunger to the second stop. The contents of the pipette should have been largely released at the first stop. The second stop ensures that you've released the "last drop."
3. Use the tip ejector to discard the tip.

WARNING: The most common-and serious- operator error is depressing the plunger to the second stop before filling the micropipette tip. DON'T DO THIS!!!

Spectrophotometer

Spectrophotometers measure the amount of light absorbed by a sample at a particular wavelength, which depends on the electronic structure of the molecules in the sample. In a spectrophotometer, light first passes from a source to a monochromator, which allows light of only a defined wavelength to pass through. The light then passes through a sample cuvette to a detector. The instrument compares the fraction of light passing through the monochromator (I_0) to the light reaching the detector (I) and computes the transmittance (T) as I/I_0. Absorbance (A) is a logarithmic function of the transmittance and is calculated as:

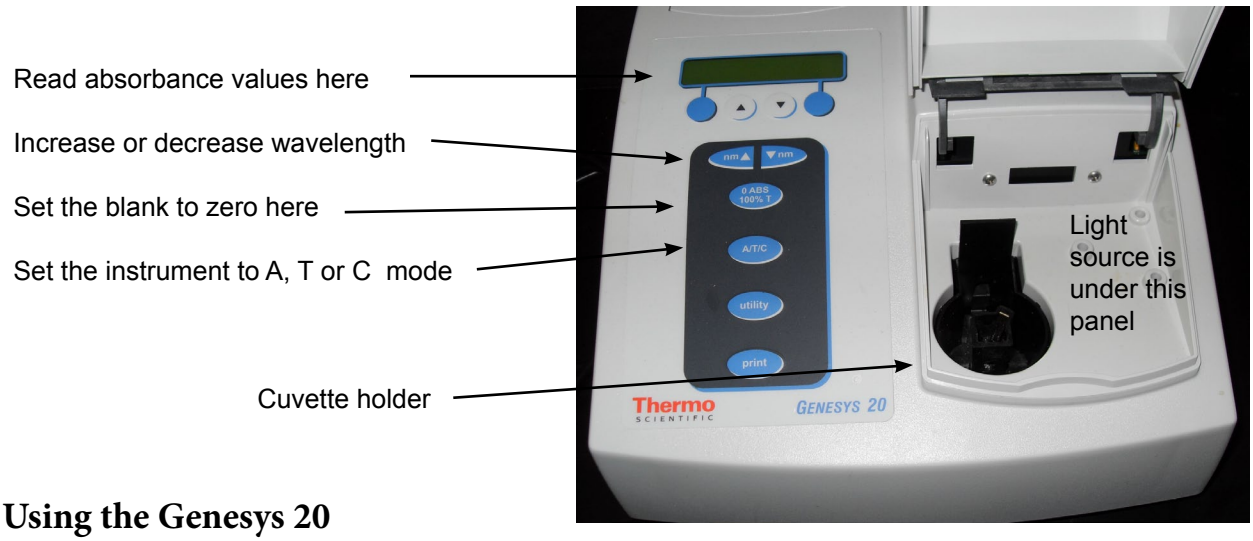
$$A = \log_{10}(1/T) = \log_{10}(I_0/I)$$

Spectrophotometers can express data as either % transmittance or absorbance. Most investigators prefer to collect absorbance values, because the absorbance of a compound is directly proportional to its concentration. Recall the Lambert-Beer Law, traditionally expressed as:

$$A = \varepsilon\, b\, C$$

where ε is the molar extinction of a compound, b is the length of the light path through the sample, and C is the molar concentration of the compound. Cuvettes are formulated to have a 1 cm light path, and the molar extinction coefficient is expressed as L/moles-cm. Consequently, absorbance is a unitless value.

Our labs are equipped with GeneSys 20 spectrophotometers. Programming is very simple and is accomplished with a few buttons, as shown in the figure below. In our labs, we will be using the instruments in absorbance (A) mode, rather than the transmittance (T) or concentration (C) modes.



Using the Genesys 20

1. Set the mode to absorbance.
2. Adjust the wavelength using the arrow keys.
3. Prepare a cuvette containing deionized water to serve as a blank. Be sure the cassette is oriented correctly in the light path. Insert the cuvette into the cuvette holder. Close the lid. Press the 0 Abs/100% T key. This will zero the instrument. Save this blank for others to use.
4. Place a second cuvette with water into the holder and record the absorbance. You will need to substract this value, which corrects for differences between cuvettes, from all subsequent readings. Empty the cuvette and use this same cuvette for all of your other samples.
5. Repeat step 4 with all of your samples, recording the absorbance readings in your notebook.

Chapter 13

Light microscopes

Microscopes are essential for viewing microorganisms. Yeast cells typically have diameters of ~10 μm, while bacteria have diameters of ~1 μm, far too small to be seen without considerable magnification. Light microscopes use a system of lenses to gather and focus light passing through a specimen and to project the image on the viewer's retina. (The lens in the viewer's eye forms part of this path as well.) Our labs are equipped with Leica DM500 brightfield light microscopes that effectively allow one to view microorganisms. The microscope has four, interchangeable objective lenses, with magnifications of 4X, 10X, 40X and 100X. Ocular lenses in the eyepieces magnify specimens an additional 10-fold, producing final magnifications of 40X, 100X, 400X and 1000X. When working with the microscope, ***always begin with the lowest power objective***, which is easiest to focus, and work your way to the higher power objectives. The lenses on the DM500 are parfocal, meaning that specimens remain reasonably well-focused when the lenses are changed. (We will not be using the 100X lens, which requires immersion oil.)

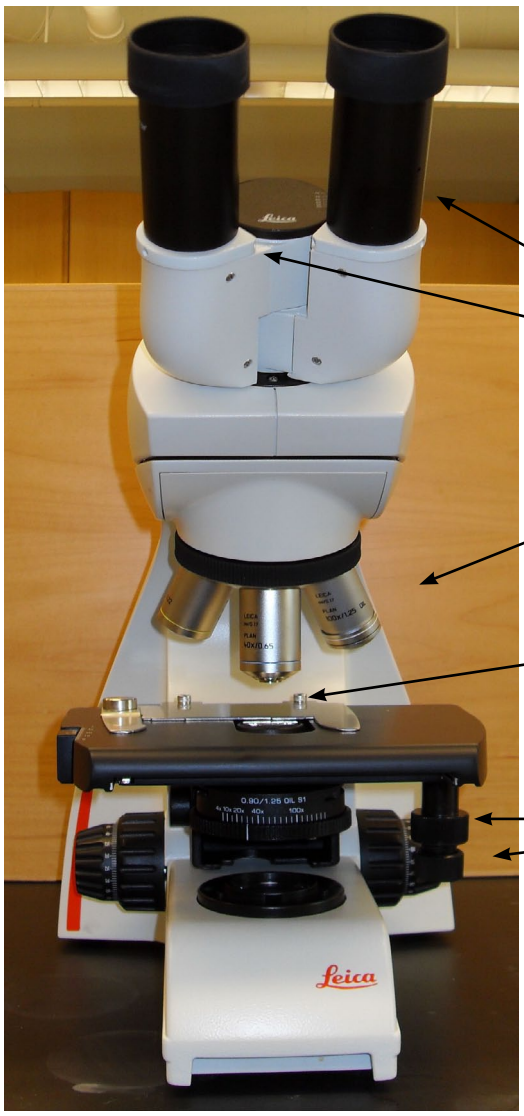
In the DM500 microscopes, light from an LED source at the base of the microscope enters a condenser that focuses the light reaching the specimen on the microscope stage. Users are able to control the amount of light reaching the specimen by adjusting the intensity of light emitted from the source and/or by opening and closing the aperture diaphragm, essentially an iris that controls the diameter of the light beam that enters the condenser. Slides containing specimens are mounted on the stage in a slide holder. The position of the slide with respect to the light path can be finely adjusted by moving the stage with X- and Y-control knobs.

Using the microscope

1. Adjust the positions of the eyepieces to fit the distance between your eyes.
2. Rotate the 4X objective lens into position, and adjust the aperture diaphragm to the recommended opening for the 4X lens. Turn on the light source using the switch on the lower right of the microscope stand. Place a slide in the slide holder and use the XY stage controls to center the coverslip in to the microscope light path.
3. Use the coarse focus knob to bring the specimen into focus. Use the fine focus knob to sharpen the image. Adjust the light intensity to a comfortable level by using the illumination control on the left side of the stand or by turning the aperture diaphragm.
4. Swing the 10X objective into position and focus on the specimen. Because the lenses are parfocal, you may be able to sharpen the image with only the fine focus knob. Adjust the light, as in step 3. Yeast cells should be visible, but small.
5. Swing the 40X objective into position and readjust the lighting and focus. You may find that closing the aperture slightly improves the appearance of unstained yeast preparations.

Precautions to take with lenses

- Be careful while focusing. The objective should not touch the slide at any time! Lenses are fragile and expensive!
- Clean lenses with lens paper only. Kimwipes™ and other paper may scratch a lens.
- The 100 X objective contains an oil immersion lens. A drop of immersion oil must be placed on the coverslip before using the lens. If you use the 100X lens, be sure to wipe the oil off the lens with ***lens paper*** when you are finished. Do NOT use oil with any of the other lenses, which will be damaged by the oil.



Leica DM500
Light microscope

Eyepieces have 10X magnification
Interpupillary distance is adjustable

Parfocal **objective lenses**
4X, 10X, 40X and 100X

Slide holder is mounted on microscope stage

XY Controls for moving stage

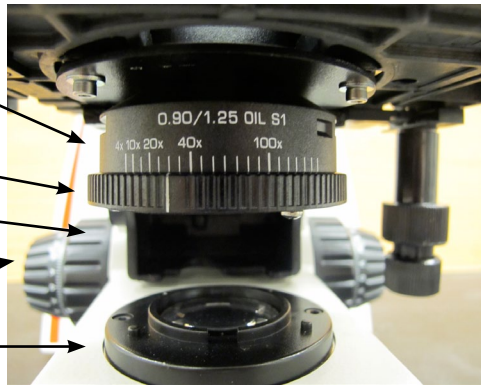
Condenser focuses the light reaching the specimen

Iris diaphragm regulates the amount of light reaching the condenser

Coarse focus

Fine focus

Light source



Chapter 13

Centrifuges

We will be using several different of small-scale preparative centrifuges this semester. Centrifuges are instsruments that use centripetal force to accelerate the sedimentation of components in a mixture. The sedimentation rate of an object is determined by its size and density, as well as the viscosity of the matrix. In this course, we will often use centrifugation to collect cells from cultures. Cells will settle out of cultures that are not agitated over time, but the centrifuge speeds up the process and produces a more compact cell pellet. We'll also use centrifugation to collect insoluble molecules into a pellet known as a precipitate. The clarified liquid layer that remains after the precipitate is removed is known as the supernatant.

In a centrifuge, samples are placed in rotors that contain sample holders designed for particular sizes of test tubes or bottles. The rotors are spun around a central spindle at a high speed, generating a centripetal force in a direction perpendicular to the vertical axis of the centrifuge. The relative centrifugal force (*RCF*) produced by a centrifuge is expressed relative to gravitational force (*g*) and depends on several variables, including the radius of the rotor (*r*) and its angular velocity (ω). The angular velocity of the rotor depends on the number of revolutions per unit time (*N*). The equation below shows the derivation of *RCF*.

$$RCF = \frac{r\omega^2}{g} = \frac{r(2\pi N)^2}{g}$$

As you can see from the equation, increasing either the radius of the rotor or the speed of the centrifuge will increase the centrifugal force exerted on the sample. The product of RCF and time, *g-min*, is often used by scientists when they are adapting a someone else's procedure to their own centrifuges and rotors.

We will be using two kinds of centrifuges in the lab. The microcentrifuges are designed specifically for 1.7-2.0 mL microcentrifuge tubes. Depending on the model, the microcentrifuges can be used at speeds up to 13,000 rpm, generating centrifugal forces as high as 17,000 *x g*. The ThermoScientific tabletop centrifuges are designed to carry larger samples. The TX400 rotors have four buckets, each of which carries an adaptor designed for either nine 15 mL or four 50 mL culture tubes. Samples can be spun at speeds up to 5000 rpm, corresponding to a centripetal forces of ~4600 *x g*.

Because of the large forces generated in centrifuges, it is critical they are properly balanced! Improperly balanced centrifuges are a safety threat!

- When loading a centrifuge, place the tubes in symmetric pattern with respect to the spindle. Most often, tubes are balanced in a pairwise fashion - two tubes are placed in a direct line that passes through the rotor spindle. Certain rotors allow 3-point symmetry as well.
- Tubes that balance each other must contain the same volume of fluid. If you have only one sample to concentrate, you will need to prepare a balance tube that contains a similar volume of water.
- Make sure that the rotor cap (if any) is placed on the rotor and the lid is closed before starting the centrifuge.
- Stop the centrifuge IMMEDIATELY if it begins to wobble or shake more than usual.

14 Yeast Culture Techniques

Objectives

- Be familiar with sterile techniques used in microbiology
- Understand the phases of microbial growth
- Understand how growth media are designed to satisfy nutritional requirements of microorganisms
- Know how to isolate single colonies on streak plates
- Know how to estimate cell densities of microbial cultures

This chapter will introduce you to standard techniques used in microbiology. Very similar techniques are used with yeast and bacteria, although the culture conditions are optimized for different organisms. In this chapter, you will learn basic sterile techniques required for maintaining the integrity of strains in the lab, as well as methods for culturing cells and estimating cell numbers.

Methods for culturing yeast in the laboratory are well-developed. The techniques used for yeast and bacteria are basically similar, except that the media composition and culture temperature are optimized for individual organisms. In general, all culture media contain a carbon source, nitrogen source, salts, vitamins and essential minerals. To work successfully with microorganisms in the lab, scientists need to be familiar with sterile techniques that allow them to maintain pure cultures of a particular genotype. Careful bookkeeping is equally important, since strains can rarely be distinguished by their appearances.

Sterile technique

Sterile technique is ESSENTIAL when working with microorganisms! This semester, we will be working with many different strains of yeast and bacteria. The strains have defined genotypes that have been generated by careful planning and experimentation. It's important to protect the integrity of these strains from contamination with other laboratory strains and from the many undefined microbes in the environment. Large numbers of diverse microorganisms are all around us - in the air, on laboratory surfaces, on your skin and on your clothing. True to their name, microorganisms are too small to be detected by the eye, but they grow rapidly in laboratory culture media. Correct transfer techniques and the use of sterile reagents is usually enough to prevent contamination of valuable laboratory strains.

Some simple precautions will reduce the possibility of contamination:

- Before working with strains, it's a good idea to wipe down a small working area on the lab bench with 70% ethanol.
- Use sterile reagents, micropipette tips, and test tubes. Tips and test tubes should be kept in covered containers until used.
- Minimize contamination from clothing and body surfaces. Pull back and secure long hair. Avoid touching or breathing on sterile surfaces that will contact microorganisms. Avoid talking when you are transferring strains.
- Minimize the time that the caps are removed from vessels containing microorganisms or sterile media. Caps should always be kept right-side up to prevent contamination from airborne microbes.

The culture media and reagents that we will use have been sterilized by either autoclaving or by filtration. An autoclave is a chamber that uses pressurized steam to kill cells on surfaces and in solutions, using temperatures near 121˚C and pressures from 30-40 psi. (For comparison, atmospheric pressure is ~15 psi.) Filtration is used in the place of autoclaving when solutions contain temperature-sensitive compounds. The pores in the filters used to remove microorganisms are typically 0.2 or 0.45 μm, which are sufficiently small to prevent passage of bacteria. It's not difficult to keep stocks of media and reagents sterile as long as you work quickly and follow the directions above.

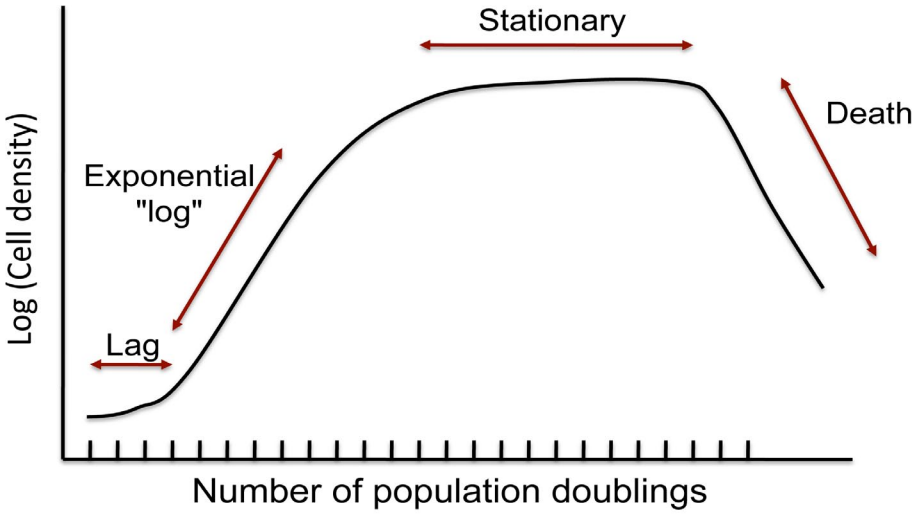
Yeast growth media

For routine culture, scientists employ rich media that supply all the nutrients that the cells need. The individual components of rich media are often undefined. For example, yeast are often grown in a medium known as YPD, which is simple and inexpensive to prepare. The "Y" in YPD refers to yeast extract and contains the water-soluble compounds generated when yeast are forced to self-digest. (Those of you who have visited Australia may have encountered yeast extract in the popular spread, Marmite.) The "P" refers to peptone, a mixture of peptides and amino acids prepared by digesting animal protein with proteases. The "D" refers to dextrose, or glucose, which is the favored carbon source of yeast. YPD is used for both liquid and solid agar-based culture media.

Because YPD is composed largely of crude extracts, its composition may show significant batch-to-batch variation. This variation is rarely a problem, however, because YPD contains more than enough essential nutrients to satisfy the metabolic requirements of cells. Many experiments, however, require media with a more defined composition. Over the years, the yeast community has developed several varieties of synthetic media, in which all the components have been defined. Cells are able to grow well in these media, but their growth rates are generally lower than those in YPD. In our experiments this semester, we will be manipulating the concentrations of several media components, including methionine and uracil, to select for particular genotypes. More information on synthetic media is available in Chapter 15 of this manual.

Yeast growth phases

When yeast are grown in liquid YPD medium, the culture follows a well-established pattern for microbial growth. Cultures are usually started by inoculating media with a small number of cells. A lag phase follows the inoculation, during which cells become acclimated to the new environment and cells begin to condition the media with their own metabolites. ***Lag phase*** is followed by an ***exponential, or log, phase***, when the number of cells increases exponentially.



The exponential growth of yeast can be described by the equation:

$$N = N_0 e^{kt}$$

where N represents the number of cells at any time (t), N_0 represents the number of cells at time zero. Scientists often find it convenient to think of the growth constant k in terms of the doubling time of the culture. In this rendering, k = ln2/T (T = the doubling time of the culture). The growth rate of yeast varies with temperature. Yeast grow well at room temperature, but they grow more rapidly at 30˚C, which we'll use for our experiments. At 30˚C, wild-type yeast have a doubling time of ~90 minutes in YPD. We will also make sure that the cultures are well-aerated by growing them on either rotary shakers or a rotating wheel.

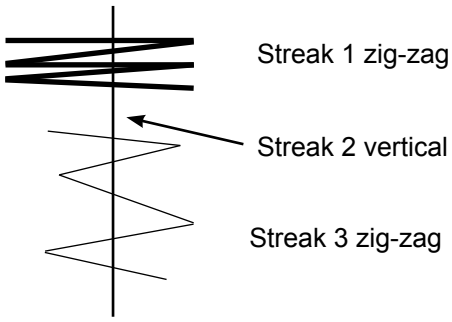
After a few doubling times, cells begin to deplete the nutrients in the culture, their growth rate slows, and the cells enter ***stationary phase.*** Cells in stationary phase have a different transcriptional program that allows them to survive in the less favorable environment, sometimes for considerable lengths of time. In stationary phase, the rate of cell division is similar to the rate of cell death, so the number of cells does not change appreciably. Eventually, cells enter ***death phase*** if conditions do not improve.

Streak plates

Yeast geneticists like to begin their experiments with a single yeast colony, because the cells in the colony are the progeny of a single cell, and they are therefore likely to be genetically similar. A concern in all genetic experiments is unknown mutations that arise spontaneously and may affect the phenotype under study. Spontaneous mutations arise constantly in all cells, with a rate of approximately 10^{-8}/base/generation. For *S. cerevisiae*, with a genome of 12 Mbp, most cells will have accumulated at least one mutation by the time that they've undergone 9-10 divisions. A colony, which has hundreds of millions of cells, is therefore a population of genetically similar, but not necessarily identical, organisms.

Researchers commonly use streak plates to isolate single colonies. A streak plate is actually a serial dilution of an existing culture on solid media. Researchers begin a streak by picking up a small sample of yeast or another microorganism with a sterile loop, wooden applicator stick or toothpick. They then spread the culture by making a series of zig-zag strokes across the surface of the plate. The number of cells on the loop or toothpick decreases as the streak progresses. Consequently, streaks appear thickest at their starting points, and the streak thickness decreases until it is possible to detect well-isolated single colonies near the end of the streak. Because it may be difficult to resolve colonies from a single streak, many labs use a series of streaks on the same plate to separate colonies. The second streak is done with a freshly steerilized loop that picks up cells by crossing over the tracks of the first streak before beginning the new series of zig-zags. In our experiments, we'll use a multi-streak protocol to conserve culture media. We will also streak out multiple strains on the same plate. (See the figure on the next page.) As you use this techniques, pay careful attention to detail to avoid cross-contamination or confusion about the identities of individual strains.

Streak plate with three sectors.
Plate has been divided into three clearly labeled sectors. Three streaks were used to spread the cells in each sector. The third streak in each sector contains well-separated colonies that can be used for genetics experiments.



Preparing a streak plate

1. Divide the plate into sectors by labeling the bottom of the plate with a magic marker. CLEARLY label each sector with a code for the strain that will be streaked in that sector. Keep the labels at the rim of the plate and use small letters. ***Note your initials andthe date as well.***
2. Gather the parent culture(s) that will be propagated, an inoculation loop or sterile toothpick, and an agar plate with the appropriate culture medium. The parent culture can be another agar plate, a liquid culture or a frozen stock of a particular genotype.
3. If you are using an inoculation loop, sterilize the loop by holding it the flame of a Bunsen burner until it glows red. Cool the loop by briefly touching the surface of the agar plate before proceeding. (Note: toothpicks and applicator sticks will have been sterilized in the autoclave, so they should not be placed in the flame.)
4. Dip the tip of the toothpick or inoculation loop into the frozen or liquid culture. The culture should be barely visible on the tip. Avoid removing too large a sample. Parent cultures are very concentrated, and a barely perceptible drop contains millions of cells. It will be more challenging to obtain nice single colonies if the starting volume is too high.
5. Remove the lid of the petri dish holding the medium with one hand and hold it at an angle over the bottom of the plate. With the other hand, make several zigzags across the outside edge of one sector with the toothpick or loop. LIGHTLY touch the agar surface as you move the toothpick. Think of pushing a hockey puck across an ice rink, rather than digging a ditch. Replace the lid and discard the toothpick. If you are using an inoculation loop, sterilize it in the flame, and then touch it to the agar plate in an area without cells to cool it down.
6. With a fresh toothpick or sterile loop, make a single vertical streak from the rim of the plate toward the center, staying within the sector. The streak should cross the zigzags in the first streak.
7. Make a new series of zigzags that cross back and forth over the straight second streak, beginning at the outer edge of the plate and moving toward the center. Be careful to stay within the sector. Invert the plate and incubate it at 30˚C until individual colonies are apparent.

Chapter 14

Determining cell densities

Investigators use a variety of methods to calculate the number of yeast cells in a culture. The quickest way to estimate the number of cells in a sample is done with a spectrophotometer, taking advantage of the ability of cells to scatter incident light. Cell numbers can also be calculated using a hemacytometer, a special counting chamber that is viewed under the microscope. (In this lab, we won't be using hemacytometers, which are expensive and delicate.) None of these methods, however, can distinguish dead cells from living cells. Investigators estimate the number of viable cells in a culture by spread plating. In spread plating, dilutions of cell cultures dispersed across the surface of an nutrient agar plate with a sterile spreader or sterile glass beads. In our experiments, we'll use a variation of spread plating known as spot plating, which allows investigators to calculate the cell densities of multiple cultures on a single agar plate.

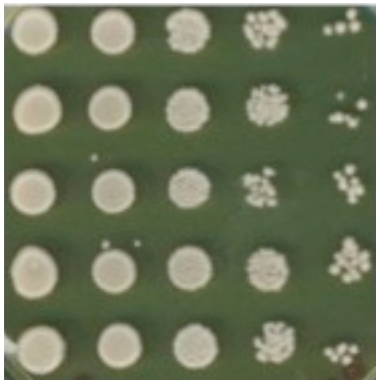
Spectrophotometric method

The spectrophotometric method to determine cell number is based on light scattering. Most visible light cannot penetrate a cell. When the light beam in a spectrophotometer hits a cell, the light is deflected from the light path, so some of the light never reaches the detector. The greater the number of cells in a sample, the more light scattering that occurs. The light scattering ability of a cell depends on its size and geometry, so a calibration curve is necessary to extrapolate optical density measurements to cell number. For example, the same number of yeast cells would scatter light more than the same number of bacterial cells, because the bacterial cells are much smaller.

Light scattering is measured with the spectrophotometer set to report absorbance. Because the principles used to measure light scattering and absorbance are different, the amount of light scattered by a solution is referred to as its "optical density" rather than its "absorbance." The optical density of a sample analyzed at 600 nm is abbreviated OD_{600}, with the subscript indicating the wavelength used for the measurement. See Chapter 13 for details on how to use the spectrophotometer.

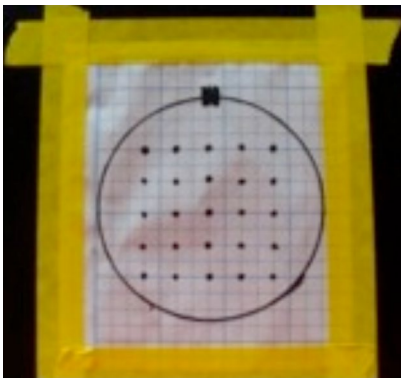
Spot plates

Scientists use spot plates both to calculate the number of cells in cultures and to obtain information about the growth properties of strains on different media. The figure on the following page shows an example of a typical spot plate. Each of the four rows on the plate represents a dilution series from a different yeast culture. The same volume of diluted culture is used for each spot. The dilution series is planned so that the most dilute spots contains a small number of individual colonies that can be distinguished from one another, typically less than ten. Most commonly, investigators make a series of 1:10 dilutions in sterile (NOT deionized) water and then spot a few microliters of each dilution in a row. In this experiment , 5 µL aliquots were spotted from the serial dilutions. Note that it's possible to count individual colonies in the most dilute samples. This in turn enables you to calculate the number of viable cells in the original culture. In the top row, you can distinguish 4 colonies in the sample that has been 100,000-fold. The original culture would have contained 400,000 cells in 5 µL, which corresponds to 80 million cells per mL (8×10^7 cells/mL).



Spot plate.
Each row on the plate represents a series of 1:10 dilutions of a liquid culture of *S. cerevisiae*. Five µL of each dilution was spotted on the plate. The plate was incubated for two days at 30˚C. Individual colonies are apparent at the highest dilution of each extract.

Day1 - Preparing the spot plate



Alignment for preparing spot plates

1. It can take a little practice to prepare good-looking spot plates! To help with the placement of the spots, the agar plate is often positioned over a grid with the target positions noted. Prepare an alignment grid from a piece of graph paper, as shown in the figure on the right. Tape the grid to the lab bench so that it doesn't move during the spotting procedure.
2. Label the petri dish containing nutrient agar with your initials and date with small letters around the rim on the BOTTOM of the dish. Put a hash mark on the bottom edge of the plate as an alignment marker. Place the dish right side up on the alignment grid that you are using for the spot plate, with the hash mark facing away from you.
3. Prepare a series of five 1:10 dilutions from your original culture using sterile distilled water. (Diagrams in your lab notebook are often helpful here.) First, pipette 90 µL sterile water into five microcentrifuge tubes. Next, transfer 10 µL from the yeast culture into the first tube, mix, and then transfer 10 µL from this tube to the second tube in the series, and so on. Use the same P20 pipette tip for the dilution series. Eject the tip into the appropriate waste container.
4. Beginning with the last dilution in the series, spot 5 µL spots in a row. Again, you will be able to use a single pipette tip for a dilution series, since you started with the most dilute sample.
5. Repeat step 3 for each culture that you are analyzing. Be careful to note in your lab notebook which culture has been spotted into each row on the plate!
6. Leave the plate right side up for ~30 minutes, to allow time for the yeast to settle and adsorb to the medium.
7. Invert the plates and incubate them at 30°C. Plates are inverted to prevent water droplets that form on the inner surface of the lid from falling on the colonies. Plates can also be kept at room temperature, but cells will grow more slowly. Do NOT incubate the cells above 30°C, which stresses the yeast.

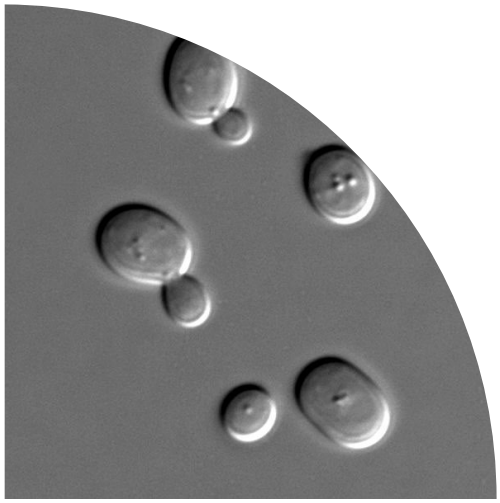
Day 2 and/or 3 - Analyze the spot plates

6. When colonies are large enough to count (typically 2-3 days), remove the plates from the incubator. If you're not ready to analyze the data, seal the plates with parafilm and place them in the refrigerator or cold room.
7. Record your data with the scanner. To do this, remove the top from the plate and invert both the plate and the lid. Place the bottom half of the dish on the scanner and leave the inverted lid on the bench. (The lid is inverted to avoid contamination from spores and microorganisms that may be present in the air.) Place a black piece of cardboard or a folder over the plates before lowering the lid of the scanner.
8. Use spots where you can count individual colonies to calculate the density of cells in the original extract, correcting for the dilutions that you used and the volume of the spot.

15 Genetic Analysis

Objectives

- Know the correct nomenclature for *S. cerevisiae* genes and proteins
- Understand how selective media are used in genetics
- Understand the principles of genetic screens
- Understand how complementation is used to identify genes
- Understand how homologous recombination is used for gene replacement



The budding yeast *Saccharomyces cerevisiae* has been a popular model organism for many years because of the the many tools available for genetic manipulation. This chapter will introduce you to the conventions used to describe *S. cerevisiae* genes and proteins as well as some of the methods used to analyze biochemical pathways in yeast.

Yeast geneticists often talk about the "awesome power of yeast genetics," because of the many techniques available for genetic manipulation as well as the wealth of genetic information available for yeast. The experiments in this project make extensive use of the genetic resources developed for yeast over several decades. The *met* strains that we are using in this project were generated by "reverse genetics," meaning that *MET* gene sequence information was used to generate mutant strains. The strains were generated in a genome-wide homologous recombination project (Winzeler *et al.*, 1999), which is described in more detail below. The original *met* mutants, which were actually used to identify the pathway of methionine synthesis in yeast, were generated by classical or "forward" genetic screens. In the classical genetics approach, investigators begin with a phenotype and work their way back to a genotype in a process that involves complementation. (Interestingly, although we know the genotype of our strains with reasonable certainty, they have not been subjected to the kind of rigorous phenotypic analysis used to classify the original *met* mutants. In fact, there is a good possibility that we're the first to study their phenotypes in any detail! Be open for possible surprises!)

Unlike other chapters in this manual, this chapter does not contain protocols. Instead, the chapter is intended to provide you with basic information about genetic analyses that will help you to understand the mutants that we are working with in the laboratory. You may also want to refer to a recent review by Sherman (2002).

Genetic nomenclature

When working with strains, it's important to use the correct genetic nomenclature. Pay close attention to italics and capital letters. Gene names are placed in italics, while proteins and phenotypes are referred to with normal font. Gene names that begin with capital letters refer to dominant alleles, while gene names beginning with lower case letters refer to recessive alleles. Gene names begin with three letters and are followed by a number. There may be many different gene names that begin with the same three letters, but the number at the end is specific for a particular gene. If some molecular information is available for the particular mutation, the number may be followed by a hyphen and additional information.

For an example, let's look at the nomenclature that would be used for the *MET6* gene from *S. cerevisiae*. *MET6* is one of over twenty different *MET* genes. The *MET* prefix is used because most of the *MET* genes were isolated in genetic screens based on the inability of mutant strains to live in the absence of methionine. The *MET6* gene acquired its name after genetic complementation studies, but before its gene product had been identified as homocysteine methyltransferase, the last step in methionine synthesis. The list on the following page describes the naming conventions for genes, proteins, and strains related to *MET6*. These same rules apply for other genes in *S. cerevisiae* as well.

(One oddity about budding yeast: *S. cerevisiae* gene names are unique in that dominant alleles are described with three capital letters. In almost all other species, dominant alleles would be referred to as *Met6* with only the first letter capitalized.)

MET6	Dominant allele of the *MET6* gene or the chromosomal locus
met6	Recessive allele of the *MET6* gene (allele found in a *met6* mutant)
met6-12	Recessive allele - number after the parentheses refers to specific mutation
met6-Δ1	Recessive allele - met6 allele has a specific deletion (Δ indicates a deletion)
met6::LEU2	Recessive allele -insertion of a dominant *LEU2* gene into the *MET6* locus on the chromosome has inactivated the host *MET6* gene
Met6p	Protein encoded by the *MET6* gene, *i.e.* homocysteine methyltransferase

To write the genotype of a particular strain, begin with the mating type and follow with the mutant alleles in the strain. For example, we are using *met* strains constructed by inserting a bacterial kanamycin resistance gene into yeast strain BY4742, which is auxotrophic for histidine, leucine, lysine and uracil. Thus, the genotype of our *met6* mutant would include the BY4742 mutations and be written: *MATa his3-Δ1 leu2Δ0 lys2Δ0 ura3Δ0 met6::KAN^R*

Auxotrophs and selective media

The *met* mutants are methionine auxotrophs, meaning that they are unable to grow in synthetic media lacking methionine. Today, many common laboratory strains are auxotrophs with mutations in genes involved in the synthesis of essential nutrients. Some common auxotrophic strains are unable to synthesize amino acids, such as leucine, methionine, histidine and lysine. Other auxotrophic strains are unable to synthesize bases required for nucleic acids, such as adenine and uracil. Thus, auxotrophic strains have specific nutritional requirements that must be supplied in their culture media. Auxotrophic strains have proven to be very useful in the laboratory. Molecular biologists often use auxotrophic strains as the host strains for plasmid transformation (Chapter 20). In these experiments, the plasmids carry functional versions of the defective yeast genes, and transformants are identified by their ability to grow on media lacking an essential nutrient.

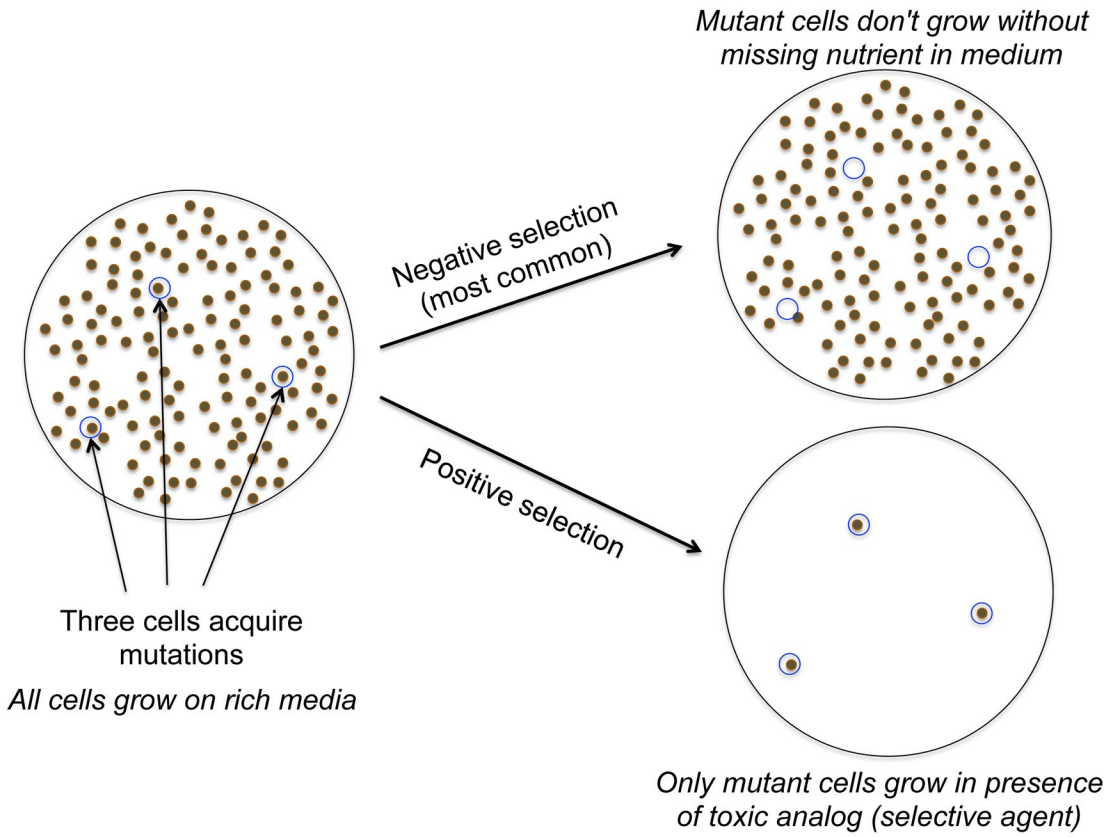
Synthetic media are an essential tool for culturing and studying auxotrophs, because all of the components are defined. Yeast researchers have developed a variety of different formulations for synthetic media. The most minimal synthetic media contain D-glucose as a carbon source, ammonium sulfate as a nitrogen source, and a variety of vitamins and minerals. The growth rate of wild type strains in these minimal media is much slower than that in rich media like YPD, but the strains survive. Strains grow more rapidly in complex synthetic media, such as the YC media that we use (http://labs.fhcrc.org/gottschling/Yeast%20Protocols/yc.html), which supplements minimal media with a variety of additional amino acids and nucleotide bases. Most common auxotrophic strains are able to grow in complete YC medium, whose composition is shown on the top of the following page. When investigators want to use YC to identify mutants, they "drop out" the nutrient of interest. For example, *met* mutants will grow in complete YC, but not in YC - Met. To determine which media will support growth, one looks at the strain's genotype. (Many experiments have foundered on this point - always look at the complete genotype of the strain when predicting growth!)

Composition of Yeast Complete (YC) Medium					
Component	**grams/liter**	**Component**	**mg/liter**	**Component**	**mg/liter**
YNB*	1.7	arginine	100	tyrosine	50
$(NH_4)_2SO_4$	5	aspartic acid	50	lysine	100
D-glucose	20	isoleucine	50	methionine	50
		phenylalanine	50	tryptophan	100
		proline	50	leucine	100
		serine	50	histidine	50
		threonine	100	uracil	10
		valine	50	adenine	10

YNB is a complex mixture of vitamins, minerals and salts. Final concentrations in YC:
Vitamins (µg/liter): biotin (2), calcium pantothenate (400), folic acid (2), inositol (2000), niacin (400), p-aminobenzoic acid (200), pyridoxine hydrochloride (400), riboflavin (200), thiamine hydrochloride (400).
Minerals (µg/liter): boric acid (500), copper sulfate (40), potassium iodide (100), ferric chloride (200), manganese sulfate (400), sodium molybdate (200), zinc sulfate (400).
Salts (mg/liter): potassium phosphate monobasic (1000), magnesium sulfate (500), sodium chloride (100), calcium chloride (100).

Genetic screens

Synthetic media provide an important tool for genetic screens designed to identify mutants in biochemical pathways. Classical genetic screens begin with a mutagenesis, in which the parent strain is treated with a chemical or radiation that induces mutations in DNA. Yeast, like all organisms, have a low spontaneous mutation rate of ~10^{-8}/base/generation. The yeast genome is ~12 Mbp, so most cells undergo cell division without having accumulated a mutation. The spontaneous mutation rate would be much too low for a practical genetic screen, so investigators adjust the mutagen dose to one that kills up to ~50% of the cells. The mutagenized cells are then used for the genetic screen. In the screen, large populations of cells (***many*** plates with ***many*** colonies per plate) are plated on selective media that will identify mutants. The easiest kinds of screens employ positive selection, which allows only mutant cells to grow. More often, investigators are analyzing pathways important for cells to grow - in these cases, mutant cells will fail to grow on selective media. In these cases, which employ negative selection, cells are first plated on permissive media that allows all cells to grow. Replicas of these master plates are made, and copies of the colonies are then plated on both permissive and restrictive (selective) media. Since only wild-type cells grow on the restrictive media, researchers look for colonies on the permissive plate that are missing on the restrictive plate. These negative screens are more difficult to do than screens that use positive selection, but they are critically important for sorting out essential biochemical pathways. The figure at the top of the next page provides a graphical comparison of genetic screens using positive and negative selection strategies.



Selection strategies used to isolate yeast mutants.
After the initial mutagenesis, yeast are plated on a plate containing rich (or complete synthetic) media. In this figure, the mutagenesis has generated three different mutants in the gene of interest. The mutant colonies are surrounded by an empty circle. Replicas of the master plate are copied to selective media. In a negative selection scheme, the selective plate lacks a component that is normally present in complete synthetic media. In a positive selection scheme, the media contains a selective agent, which is toxic to normal cells, but tolerated by mutant cells. The selective agent is often a toxic analog of a normal cellular metabolite.

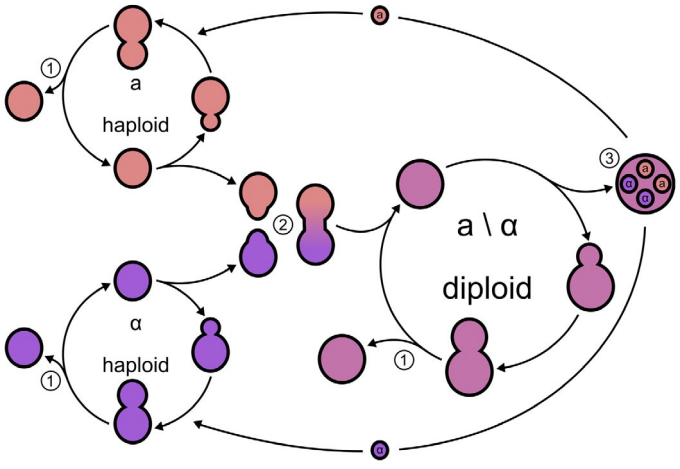
The genes required for methionine synthesis were primarily discovered with negative selection schemes. Wild type yeast are able to synthesize methionine using only sulfate as a sulfur source, but *met* mutants require an external source of methionine (or an alternative source of organic sulfur) to grow. Large numbers of *met* mutants were isolated in a variety of screens carried out in different laboratories. Investigators exchanged strains, did complementation experiments, and performed biochemical analyses that elucidated the metabolic pathways involving methionine.

Chapter 15

Complementation

After a successful genetic screen, investigators often have sizeable collections of mutant strains with the same phenotype, *e.g.* the inability to grow in the absence of methionine, but they have no idea which genes have been mutated. Because mutations occur randomly and the genome is large, it's extremely unlikely that any two mutant strains carry exactly the same mutation, but it is likely that two different strains could carry mutations in the same gene. How can the investigator determine if the mutations have occurred in the same gene? To answer this question, investigators need to put the two mutant alleles into the same cell and see if the cell lives under restrictive conditions (*e.g.* medium lacking methionine). If the cell functions normally, the mutations are in different *MET* genes. If the cell is unable to live, the mutations must be in the same *MET* gene. This process, known as complementation testing, has been used to identify thousands of genes involved in many different biochemical pathways.

Complementation testing is easier in yeast than in many organisms, because investigators are able to take advantage of the fact that yeast have both haploid and diploid phases in their life cycles. During genetic screens, investigators usually work with haploid strains, because all mutations are evident in the phenotype. After the screen is over and mutant strains have been established, these haploid strains can be mated with other mutant strains of the opposite mating type (opposites attract!) to form diploid strains.



Yeast alternate between haploid and diploid phases.
Yeast mating type is determined by the *MAT* locus, which has two major alleles, *MAT*α and *MAT***a.** Yeast of either mating type can be propagated indefinitely as haploid strains in the laboratory (1). Under stress or starvation conditions (2), haploid yeast strains form diploids with yeast strains of the opposite mating type. Depending on the genotype and the culture conditions, the diploid may be able to divide indefinitely in culture. If not, the diploid enters meiosis and produces four spores (3), each of which possesses a novel combination of the parental genes.

For complementation tests, investigators conduct a series of matings (we're skipping some steps here.....). In the first round of matings, investigators introduce the undefined *met* mutations into strains with different mating types. Consider a situation where investigators have constructed strains of opposite mating types with mutations in either the *MET3* or *MET5* genes.

(1) Haploid strains with different mating types are methionine auxotrophs:

MATα met3 MET5 *MAT***a** *MET3 met5*

Investigators then mate the strains, generating diploid strains that can be screened on selective media. If the diploid is unable to survive in the absence of methionine (or if it does not produce some spores capable of surviving without methionine), the *met* mutations are in the same gene. If the diploid (or some of its spores) survives without methionine, then the mutations are in different genes.

(2) Diploid strains are prototrophs (strains don't require methionine in media)

*MATα/MAT***a** *met3/MET3 MET5/met5*

This diploid might be able to propagate indefinitely in selective medium, since it has both *MET3* and *MET5* functions required to synthesize methionine.When nutrients become limiting, however, the diploid might enter meiosis and produce spores with eight different genotypes.

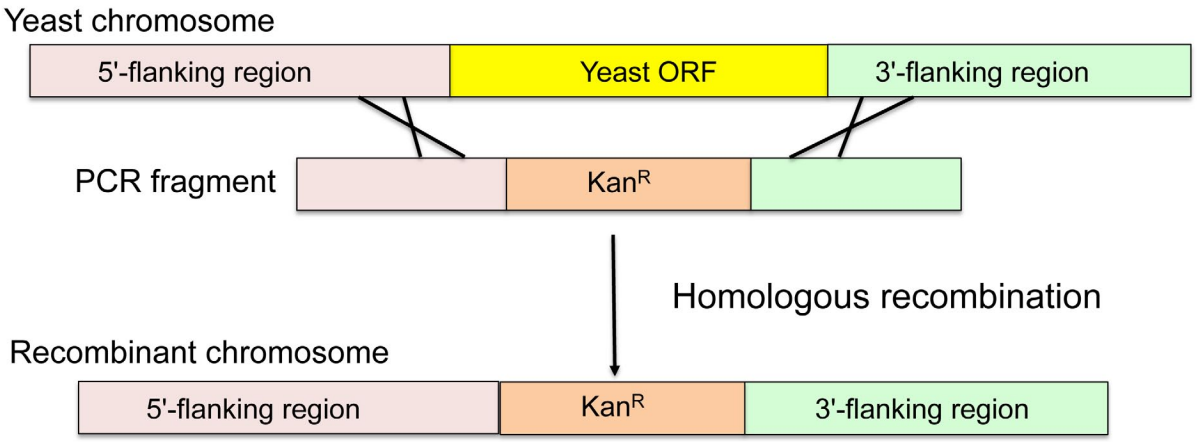
Can you name the eight different genotypes of the haploid spores?
What fraction of the spores would be able to survive without methionine?

Classical complementation experiments attest to the power of using phenotypes to identify genes in a pathway. To date, investigators have identified over 20 genes involved in yeast methionine synthesis. Not surprisingly, many of the MET genes encode proteins that catalyze steps in methionine biosynthesis (see Chapter 4). However, other MET genes encode transcriptional regulators or proteins involved in synthesis of enzyme cofactors, which are equally important for methionine synthesis.

Homologous recombination

The publication of the yeast genome sequence opened new opportunities for yeast geneticists. Knowing the DNA sequence of the yeast genome, geneticists could take advantage of the high frequency with which yeast exchange genes using homologous recombination to generate mutants of their own design. Homologous recombination normally occurs during meiosis and during certain kinds of DNA repair. During homologous recombination, two closely related DNA sequences align with one another, the DNA strands break and strand exchange occurs when the breaks are healed. Investigators who want to use homologous recombination for gene replacement use molecular cloning to construct a replacement cassette in which a marker gene is flanked on either side by sequences that flank the gene to be replaced in the target chromosome.

The strains that we are using were generated by this approach during the *Saccharomyces* Genome Deletion Project (Kenzeler *et al.*, 1999). After the publication of the DNA sequence, yeast researchers began to systematically replace all of the predicted ORFs in the *S. cerevisiae* genome with a kanamycin resistance (*KAN*^{R}) gene. The figure below outlines the process used to generate the deletion strains. For each ORF, researchers used a series of PCR reactions to construct cassettes in which the *KAN*^{R} gene was flanked at either site by short DNA sequences upstream and downstream of the targeted ORF on the *S. cerevisiae* chromosome. The PCR fragments were used to transform the BY4742 strain of *S. cerevisiae*, and strains that had incorporated the *KAN*^{R} gene were selected on plates containing kanamycin.



All of the deletion strains were analyzed by PCR to confirm that the targeted ORF was replaced by the *KAN*^{R} cassette. Thus, we can be reasonably certain about the genotypes of our *met* mutants. The Deletion Project also verified that strains were able to grow and mate, showed the appropriate auxotrophies and had functional mitochondria. (See the Deletion Project website for additional details: http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3.html.) With the large number of strains generated in the project, it was not possible to do a more thorough phenotypic analysis of individual strains. Those analyses will be left for you to do!

References

Sherman, F. 2002. Getting started with yeast. *Methods Enzymol.* **350**: 3-41.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R. *et al.* 1999. Functional characterization of the *Saccharomyces cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**: 901-906.

16 Databases

Objectives

- Understand the importance of databases in modern molecular biology
- Be aware of the kinds of information in different biological databases
- Understand the role of curators in constructing databases and processing the information in databases
- Understand that databases are extensively linked to one another

An explosion of information in modern biology has driven the construction of online databases that provide various kinds of information to researchers. A prudent investigator will conduct a thorough investigation of the relevant databases when planning an experiment. This chapter will introduce you to the various databases available for research and the way that information is processed in the databases.

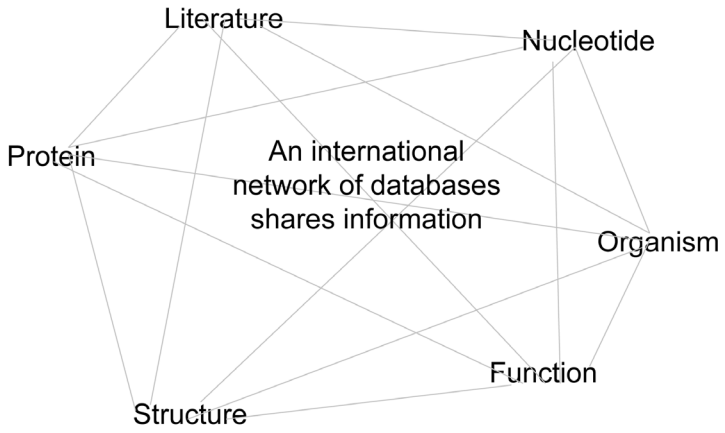
Chapter 16

Databases: an overview

Drowning in data! Technological advances over the past few decades have spurred the development of high throughput experiments in biology. Today, robots and 96-well plates are not uncommon in the research laboratory. With big experiments, comes a flood of data that must be processed and organized in a manner that allows easy retrieval. Bioinformatics is the rapidly growing interdisciplinary science that develops tools for organizing and analyzing a wide variety of different kinds of biological data.

An international network of linked databases

All of this information is organized into databases, which have various missions that vary widely in their size and scope. This chapter will introduce you to some of the large, international databases that are freely accessible to the biological research community. The large, interntional databases receive support from governments, and they work cooperatively with one another. The figure below shows the various kinds of databases that are widely used in molecular biology.



By far, the largest collection of databases is housed at the National Center for Biotechnology Information (NCBI) in the United States. NCBI includes literature, nucleotide, protein and structure databases, as well as more specialized, derivative databases. The European Union supports the EMBL-EBI (European Molecular Biology Laboratory-European Bioinformatics Institute) collection of databases and bioinformatics tools. Both NCBI and EMBL-EBI exchange nucleotide sequence information on a daily base with each other and the smaller DNA Database of Japan (DDNJ). Researchers need to submit their data to only one of the databases. Once the data is entered into a database, it is shared by all of the other databases.

A database user can access information in the network from many different starting points. The user's entry point is usually a matter of personal choice. For example, the NCBI and EMBL-EBI maintain protein and nucleotide databases with their own look and feel. The curators at the two institutions organize and present the data somewhat differently, but the basicinformation is the same. One user may prefer NCBI's Protein database record, while another prefers the EMBL-EBI's Uniprot database record.

Professional curators process the information submitted to databases

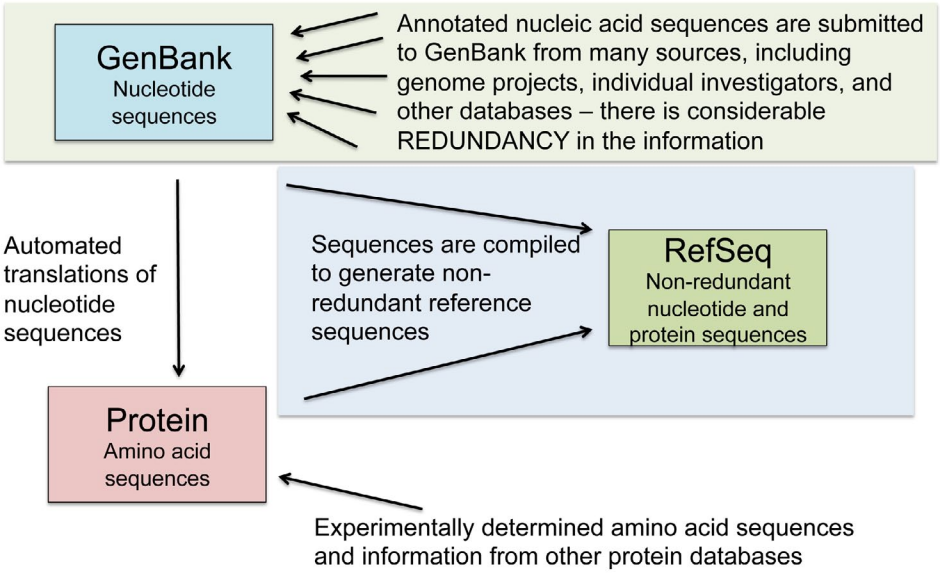
The information in the databases comes from the research community. There are a few databases that accept raw data, such as environmental DNA sequences, but most databases require some annotation to accompany the information being submitted. For example, when an investigator submits a DNA sequence to NCBI's Genbank database, the investigator completes a form that asks for the source of the DNA and some basic features of the sequence. Does the DNA contain gene coding sequences? ...promoters? ...introns? Structure databases require detailed atomic coordinates for the molecular models that investigators submit. Organism databases accept experimental information on mutant strains that meets the community standards for genetics notation. All databases ask for references to research publications that provide experimental evidence for the submission.

A record that meets a database's standards is accepted and assigned a unique accession number that will remain permanently associated with the record in that database. Each database has its own system of accession numbers, so it is often possible to identify the source of a record from its accession number. Once a record is accepted into a database, professional curators take over the record and the curators add considerable value as they link the record to information in other databases. Curators also process data from primary databases into derivative databases with specialized foci. Derivative databases are very useful to the research community because they eliminate the redundancy in primary records. No database record is ever static. The information may change over time as investigators correct errors and new supporting data becomes available.

In this course, we will use many records from the Refseq databases at NCBI. The Refseq records eliminate the redundancy in GenBank records, which is the primary repository for investigator submissions at NCBI. GenBank is a huge databank with records ranging from plasmids and individual gene records to chromosome sequences. Gen Bank was established in 1982, well before the scientific community began whole genome sequencing. Over the years, investigators contributed many separate records for individual genes, including genomic sequences, transcripts, and coding sequences. (As of April, 2011, there were over 135 million records in GenBank.) The RefSeq database was established to bring together records on a single gene in one location.

Many RefSeq records begin with NM_ or NP_, depending if the sequence represents a transcript or protein sequence, respectively. These NP_ and NM_ records are supported by experimental data. The *S. cerevisiae* records that we will be using in this course have NP_ and NM_ records. The *S. cerevisiae* records use results from the genome project records as the standard sequence, but curators have incorporated information from other primary records as well. RefSeq records for many other species begin with XM_, XP, YM_ or YP_ prefixes. These latter records were also derived from genome sequencing projects. Protein translations and other annotation was computationally generated using automated algorithms. Consequently, the annotation is less well-supported by experimental evidence.

The figure on the following page outlines the data flow through the NCBI's GenBank, Protein and RefSeq databases.



Databases for different needs and different communities

The best way to learn about databases is to use them! The list below gives a few examples of popular free sites. The list is not comprehensive, and you will be sure to find others. Explore these sites and develop your own favorite approaches to finding information. Note that all NCBI sites are available through its Entrez portal (***http://www.ncbi.nlm.nih.gov***).

Literature - PubMed (at NCBI) references over 5000 biomedical journals and is referenced by all the major international databases. NCBI does not reference many journals with an environmental focus. Those databases, such as Web of Science and Biosis, are available to you through the BC libraries. *Hint: Log on to the BC library system before working with ANY literature database, as this will make the BC Find-It services available to you.*

Nucleotide - NCBI maintains other specialized nucleotide databases in addition to the more general GenBank and RefSeq nucleotide databases, which are described above. Other nucleotide databases can be found in the EU's EMBL-EBI database collection (***www.ebi.ac.uk/***). EBI has a very nice set of user-friendly set of bioinformatics tools available at its site.

Protein - The UniProtKB Protein knowledgebase (***uniprot.org***) is a favorite nonredundant site for protein information. Each protein's summary page includes structural, functional and sequence information. Tools for analyzing the protein can be accessed directly.

Function - Proteins can be placed into metabolic pathways at the searchable MetaCyc: Encyclopedia of Metabolic Pathways (***metacyc.org***) and KEGG: Kyoto Encylopedia of Genes and Genomes (***www.genome.jp/kegg/***) sites.

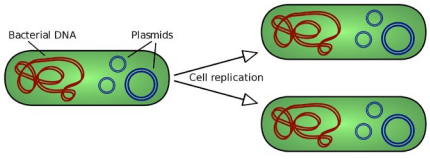
Structure - Atomic coordinates for protein and nucleic acid structures obtained by X-ray crystallography and NMR are deposited in the Worldwide Protein databank, which can be accessed through sites in the U.S. (***rcsb.org***), Europe (***pdbe.org***) or Japan (***pdbj.org***). The sites also contain software for molecular visualization that is easy to use.

Organism - Organism-specific databases provide "one-stop shopping" for information about genes in model organisms. Data from all of the different types of databases in this list are collected and hyperlinked at these sites. Databases are available for *E. coli*, *S. cerevisiae* (yeastgenome.org), *Drosophila*, *C. elegans*, and zebrafish, among others.

17 Plasmids

Objectives

- Understand the structure of plasmids
- Understand the mode of plasmid replication
- Understand the function of selectable markers in plasmids
- Appreciate various features engineered into cloning vectors



Adapted from http://en.wikipedia.org/wiki/Plasmid#mediaviewer/File:Plasmid_(english).svg

Plasmids were the first cloning vectors to be widely used in molecular biology. Plasmids are small, circular pieces of DNA that replicate independently of the host chromosome. The first plasmids used in the lab were derivatives of naturally-occurring plasmids. Since then, investigators have added many bells and whistles to plasmids to suit various purposes in the lab.

Background

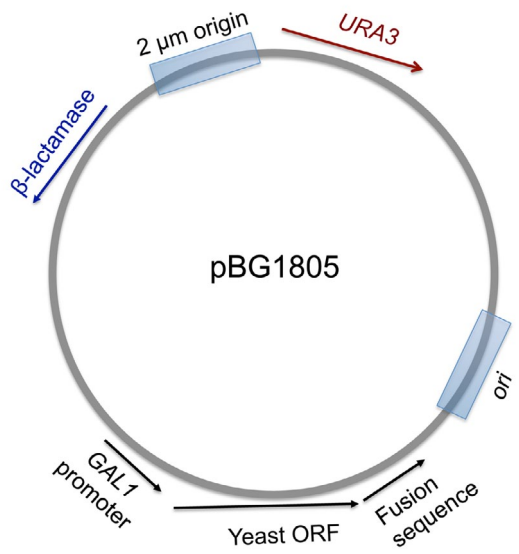
Plasmids replicate independently of the host chromosome

Plasmids are extrachromosomal pieces of DNA found in many microorganisms. In nature, plasmids can be transferred between species by transformation or conjugation, but they generally have a restricted host range. When you think of plasmids, you probably also think of bacteria, but plasmids are not restricted to bacteria. In fact, most *Saccharomyces cerevisiae* strains carry a large plasmid known as the 2 micron or 2 μm plasmid. Multiple copies of the 2 μm plasmid are usually present in the nucleus of a yeast cell, and the plasmid number is stable through many rounds of cell division. Although plasmids replicate independently of the chromosomal DNA, they rely on host enzymes to catalyze their replication. Host DNA polymerases bind to origins of replication within the plasmid replicon. Plasmids that replicate in bacteria have origins of replication that bind bacterial DNA polymerase, while plasmids that replicate in yeast have origins of replication that bind the very different yeast DNA polymerase. Other sequences in the replicons determine the copy number of the plasmid, which has practical ramifications for molecular cloning. Most laboratory plasmids are multi-copy plasmids, which are present anywhere from tens to hundreds of copies per cell.

Plasmids used in molecular cloning are commonly referred to as vectors, because they carry DNA sequences of interest into a cell. In this class, we will use shuttle vectors that contain both bacterial and yeast origins of replication. We will propagate the shuttle vectors in bacteria, because bacteria grow more rapidly than yeast and the yield of plasmid is usually higher from bacteria than from yeast. We will purify the plasmids from bacteria and then use them to transform yeast *met* strains for complementation analyses.

Selectable markers are used to identify transformed cells

For plasmids to be useful in molecular biology, it's essential that investigators be able to distinguish transformed cells that carry the plasmid from nontransformed cells. Molecular biologists have therefore mimicked nature by engineering various selectable markers into laboratory plasmids. Without a selective advantage, plasmids would most likely be lost from host cells, because of the extra toll they place on host metabolism. The markers used to select tansformed bacteria often carry genes that confer antibiotic resistance, while the markers used to select transformed yeast strains often carry genes that complement nutritional defiiencies in auxotrophic strains. The plasmids that we will use in our experiments are based on pBG1805 (Gelperin et al., 2005). The "p" in a its name denotes that it is a plasmid, while the remainder of the plasmid name is a code used by the laboratory that constructed it. Often, the letters in a plasmid's name contain the initials of the researcher who performed the last step in its construction. In our plasmids' case, "BG" refers to Beth Greyhack, one of the paper's authors. The structure of pBG1805 is outlined on the next page. The pBG1805 series of vectors were generated in a genome-wide experiment in which each of the *S. cerevisiae* ORFs was amplified by PCR and then cloned into the BG1805 cloning site. The pBG1805 vector was designed to drive overexpression of the yeast ORFs in host yeast cells. To accomplish this, the vector contains the powerful *S. cerevisiae GAL1* promoter, which is normally regulated by glucose and galactose in cells. The vector also adds amino acid sequences with different functionalities to the C-termini of cloned sequences (See Chapter 22 for additional details.)



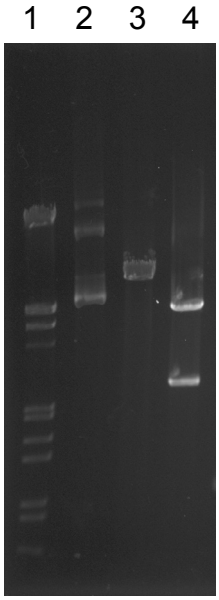
Plasmid BG1805.
The plasmid without a yeast ORF is 6573 bp long. Features of pBG1805 include:

- the *E. coli* gene for the beta-lactamase protein that degrades ampicillin
- the *S. cerevisiae URA3* gene with its natural promoter
- a bacterial origin of replication (ori)
- the yeast 2 µm origin of replication
- the *S. cerevisiae GAL1* promoter
- a cloning site for yeast ORFs
- fusion sequences at the C-terminus of the cloned ORFs

Native plasmids have a supercoiled structure

Plasmids are circles of double-stranded DNA twisted into supercoiled structures. The supercoiling makes it impossible to determine the size of a native plasmid by agarose gel electrophoresis. The figure below shows an agarose gel containing a sample of native plasmid DNA in lane 2. The native plasmid DNA in lane 2 does not migrate as a single band. Instead the plasmid electrophoreses in several discrete bands, reflecting a variety of structures. Investigators must break open the circular structure and convert the plasmid to a linear form for analysis on agarose gels. To do this, investigators use restriction endonucleases that recognize specific nucleotide sequences within the plasmid and catalyse double-strand breaks at or near these recognition sites (see Chapter 18).

Incubations of the plasmid in the figure with restriction endonucleases generate linear DNA fragments whose sizes can be calculated from their migration relative to size standards. The restriction enzyme BglII cleaves the plasmid at a single site, converting it to a linear form with the predicted size. The BsaAI restriction endonuclease recognizes two sites in the plasmid DNA, generating two fragments.



Separation of plasmid DNA by agarose gel electrophoresis

1 - Size standards
2 - Intact plasmid DNA
3 - BglII restriction digest
4 - BsaAI restriction digest

Plasmid isolation

Plasmids are easy to isolate from bacteria because their physical properties are quite distinct from that of the bacterial chromosome and because they are quite resistant to permanent denaturation. Today, most laboratories use commercial kits for plasmid isolations, because the kits are convenient and relatively inexpensive. The kits give good yields of high-quality DNA, while avoiding the need for organic denaturatants. A variety of less expensive, but somewhat more time-consuming, procedures have been described for investigators who want to make their own reagents. These procedures generally give good yields of DNA that is slightly less pure than DNA purified with the kits. Whatever the isolation procedure, the general principles of plasmid isolation are the same. The paragraphs below summarize the general principles used for plasmid isolation.

1. ***Denaturation*** - The first step in plasmid isolations is to lyse that bacterial cells carrying the plasmids. Strong denaturating conditions are required for this step, because of the bacterial cell wall. The most common procedures use a combination of strong base and a detergent, such as SDS, to lyse the cells. An alternative is to boil the cells in the presence of detergent for several minutes. These treatments not only lyse the cell wall, but they denature both DNA and protein.
2. ***Neutralization*** - The extracts are then neutralized (or cooled in the boiling method) to renature the plasmid and precipitate denatured proteins and chromosomal DNA. Plasmids renature when the extract is neutralized because they have supercoiled structures that remain partly associated even when denatured. Samples must be mixed gently at this step to prevent fragmentation of the long, chromosomal DNA into pieces that would co-purify with the plasmids. Precipitated proteins and DNA are removed from the extracts by centrifugation or by phenol extraction.
3. ***Purification of plasmid DNA*** - Most kits for plasmid purification contain anion exchange resins that bind the plasmid DNA under high salt conditions. Contaminants are washed from the column before the plasmid DNA is eluted. When the procedure does not involve resins, the extracts may be extracted several times with phenol and/or chloroform, followed by precipitation of the plasmid DNA with either ethanol or isopropanol. The purity and yield of the extraction procedure can be monitored by ultraviolet spectroscopy.

References

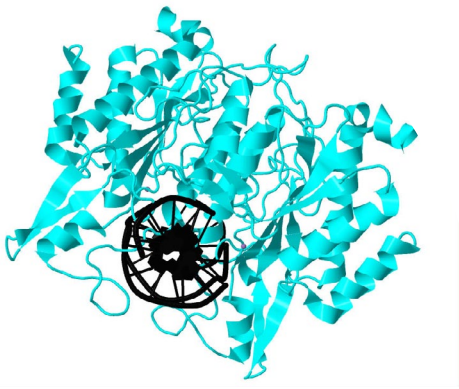
Gelperin, D.M. White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., Gerstein, M., Dumont, M.E., Phizicky, E.M., Snyder, M., and Grayhack, E.J. 2005. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop.* **19**: 2816-2826.

18

Restriction Endonucleases

Objectives

- Understand the biological origin of restriction endonucleases
- Understand the restriction endonucleases and bind and cleave specific sequences in DNA
- Understand the structure of restriction sites in DNA
- Understand some applications of restriction endonucleases in molecular biology



Restriction endonucleases have become the scissors for cutting and pasting DNA in the laboratory. In nature, restriction endonucleases are part of a defense system that protect bacteria from infection. Bacteria have evolved thousands of these enzymes with distinct recognition sites in DNA sequences. The figure shows *EcoR*I straddling its recognition site in DNA (PDB accession 1QPS).

Background

Restriction endonucleases are part of a bacterial defense system

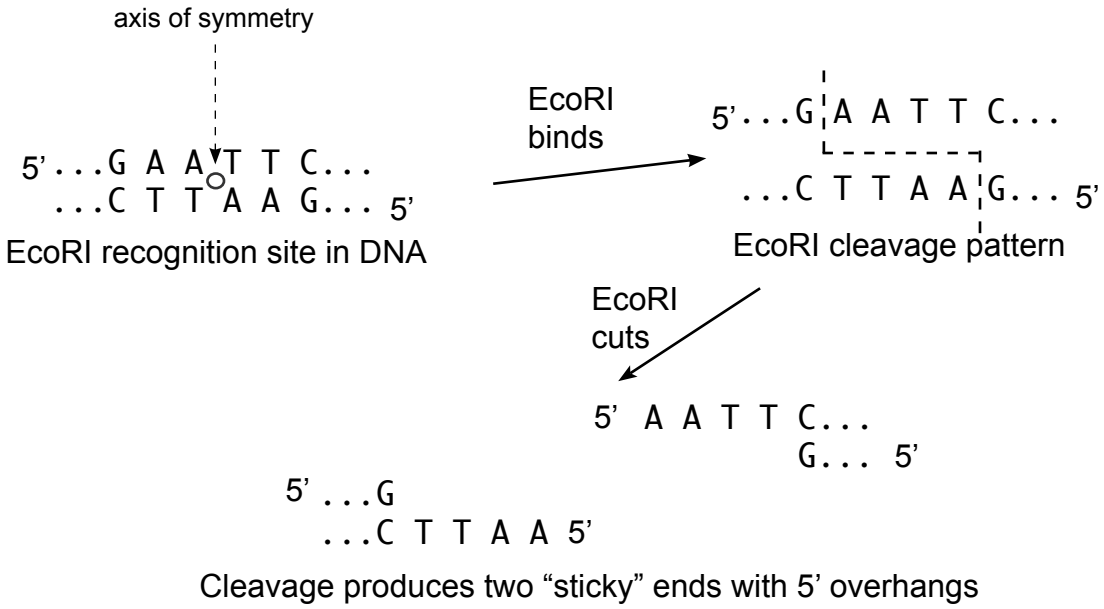
The discovery of restriction enzymes, or restriction endonucleases, was pivotal to the development of molecular cloning. Restriction enzymes are endonucleases from bacteria that specifically recognize short stretches of nucleotides in DNA and catalyze double-strand breaks at or near the recognition site. The sequence of nucleotides recognized by a restriction enzyme is known as a restriction site. To date, thousands of restriction endonucleases with distinct specificities have been described. You might wonder why bacteria harbor these potentially destructive enzymes. Restriction enzymes are part of a bacterial defense system against foreign DNA, such as might be carried by an infectious bacteriophage. The bacteria protect their own DNA from the endonuclease by adding a methyl group to the same restriction sites in its own DNA. The methylated restriction sites are no longer recognized by the endonuclease. The combined activities of the endonuclease and methyltransferase are referred to as a restriction/modification system. In Type I enzymes, the endonuclease and methyltransferase activities are part of the same complex. The restriction endonucleases used in laboratory are Type II enzymes. The Type II endonucleases and methyltransferases do not form a complex and the two enzymes function independently of each other.

To understand how restriction endonucleases work, it may be helpful to look at a single enzyme in greater detail. One of the best-studied restriction enzymes is EcoRI. Although the names of restriction enzymes sound a bit like baby talk, the nomenclature is actually very systematic and is based on the biological source of the enzyme. EcoRI is found naturally in the RY13 strain of *Escherichia coli*, or *E. coli*. Its name begins with the genus and species (Eco for *E. coli*), followed next by a strain identifier (R for RY13), and ends with a Roman numeral that distinguishes the different enzymes in the strain. Strain RY13 of *E. coli* contains multiple endonucleases, but only EcoRI and EcoRV have specificities that are useful in the lab. Purification of restriction enzymes from their natural sources is a laborious procedure that requires large quantities of bacteria, since the restriction endonucleases are not abundant proteins in bacteria. To facilitate the production of these useful enzymes for routine laboratory use, molecular biologists therefore cloned the coding sequences of the endonucleases into plasmids. Today, most commercially available enzymes are recombinant proteins overexpressed in bacteria. (Along the way, scientists also introduced some mutations that generated enzymes with enhanced stability or higher activity.)

Restriction enzymes cleave specific sites in DNA

Restriction enzymes like EcoRI are frequently called 6-cutters, because they bind to a 6-nucleotide sequence. Assuming a random distribution of A, C, G and Ts in DNA, probability predicts that a recognition site for a 6-cutter should occur about once for every 4000 bp (4^6) in DNA. Of course, the distribution of nucleotides in DNA is not random, so the actual sizes of DNA fragments produced by EcoRI range from hundreds to many thousands, but the mean size is close to 4000 bp. A DNA fragment of that length is very useful in the lab, since it is long enough to contain the coding sequence for some proteins, yet it is easily resolved on common agarose gels. The sequence that EcoRI recognizes in double stranded DNA is G A A T T C. The sequence is a palindrome with a two-fold axis of symmetry, because reading from 5' to 3' along either

strand gives the same sequence. The palindromic nature of the restriction site is more obvious in the figure below. The circle in the center of the restriction site denotes the axis of symmetry. EcoRI binds DNA at the restriction site and catalyzes the hydrolysis of the phosphodiester bond between G and A on either strand.



EcoRI catalyzes the cleavage of a palindromic recognition site.
The recognition site for EcoRI has a two-fold axis of symmetry. Cleavage generates two fragments with 5′-sticky ends.

The structure of EcoRI is consistent with an enzyme that catalyzes identical reactions at two different sites. Crystallization data shows that EcoRI is a homodimer, an enzyme with two identical subunits. The figure below (PDB datafile 1QPS) shows two different views of EcoRI bound to DNA. In the left view, we're looking down the axis of the DNA helix. The right view is a side view. Take a close look at both structures, using the DNA helix as the center. In these models, alpha helices and beta sheets are given different colors or shades. The symmetry in the enzyme is quite apparent.

Crystallization data reveals that EcoRI is a symmetric homodimer.

Selecting restriction enzymes for experiments

The first step in an experiment involving restriction enzymes is to analyze the nucleotide sequence of the target DNA molecules with a software program that detects restriction sites. These programs usually generate a graphical output as well tabular data with details about the lengths and positions of the restriction site in the DNA sequence. The list of enzymes that cut a particular sequence is always impressive, but only a few enzymes usually turn out to be practical for the purpose of the experiment. When choosing enzymes, investigators consider the number and sizes of the fragments that will be generated in a digest, the position of the restriction sites relative to a gene of interest, as well as the kind of ends in the restriction fragment.

The kind of end generated by a restriction enzyme is particularly important when investigators are constructing a recombinant DNA molecule. In the figure on the previous page, you saw how EcoRI produced DNA fragments with 5' overhangs. Other enzymes produce fragments with 3' overhangs, while still others produce blunt ends with no overhangs. Ends with overhangs are often referred to as "sticky" ends, since they will form hydrogen bonds with another sticky end with the same sequence. For example, fragments generated by EcoRI will bind another EcoRI fragment, but not a fragment generated by HindIII, which produces a 5'-overhang with a different sequence. By contrast, fragments with blunt ends can anneal with any other blunt end, regardless of the adjoining sequence. DNA ligases supply the "paste" for stitching restriction fragments together, by catalyzing the formation of phosphodiester bonds.

In this course, we will be using plasmids that were constructed by a series of steps, each of which involved restriction digests and ligase reactions. We will not be constructing any recombinant molecules in this course, but it is helpful to understand the process by which they were generated.

Handling restriction enzymes in the laboratory

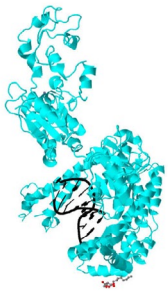
The restriction endonucleases that we are using in the lab are highly purified (and expensive!) proteins that have been purified from recombinant bacteria. Like all enzymes, each restriction endonuclease functions optimally under a defined set of reaction conditions, including variables such as temperature, pH, metal ions and salts. The manufacturer's supply restriction enzymes with buffers that generate reaction conditions close to those that are optimal for the enzyme. Some of the buffers contain bovine serum albumin (BSA), an abundant protein from cow serum. BSA is included largely for its stabilizing properties. In general, enzymes are more stable at higher protein concentrations. BSA does not affect the enzymatic activity of the restriction endonucleases, but protects particularly sensitive enzymes from denaturation. BSA may also saturate non-specific binding sites for REs on the sides of the test tubes used in the digests.

Like all enzymes, restriction endonucleases are subject to spontaneous denaturation over time, so they need to be handled with care. (By comparison, DNA is an exceptionally stable molecule.) The rate of protein denaturation increases as the temperature increases, so enzymes are kept on ice while reaction mixtures are being prepared. Denaturation is also accelerated at air/water interfaces, so it's important to minimize air bubbles and freeze/thaw cycles for solutions containing restriction enzymes. Manufacturers attempt to minimize denaturation by supplying the solutions in a solution of 50% glycerol, which does not freeze when the enzymes are stored at -20˚C.

19 Polymerase Chain Reaction

Objectives

- Understand the role of thermostable DNA polymerases in PCR
- Understand the reactions occurring in each PCR cycle
- Understand the considerations used in designing primers for PCR
- Understand how PCR is adapted for different DNA templates



The discovery of a thermostable DNA polymerase from the bacterium *Thermus aquaticus* paved the way for a revolution in molecular biology. The properties of *Taq* polymerase (left) were well-suited to the polymerase chain reaction (PCR) procedure, which allows researchers to amplify DNA sequences present at vanishingly small concentrations in biological samples.

Chapter 19

Background

The polymerase chain reaction (PCR) revolutionized molecular biology. With PCR, researchers had a tool for amplifying DNA sequences of interest from vanishingly small amounts of material. Indeed, billions of copies can be synthesized from a single DNA molecule in a typical PCR reaction. The development of PCR grew out of research on DNA polymerases and the discovery of thermostable DNA polymerases that can withstand extended heat treatments, which would denature most proteins (Sakai *et al.*, 1988). Today, PCR is a standard technique in most laboratories. PCR is an enormously versatile technique that has been used for a wide variety of applications involving the analysis of DNA molecules as well as the construction of novel recombinant molecules.

Thermostable DNA polymerases

Thermostable DNA polymerases are central to PCR. The first description of PCR used a mesophilic DNA polymerase from *E. coli* (Sakai et al., 1985), but the procedure was much-improved by the introduction of a DNA polymerase that had been isolated from thermophilic *Thermus aquaticus* bacteria living in thermal springs at Yellowstone National Park. The *T. aquaticus* DNA polymerase, or *Taq* polymerase, functions best at temperatures of 75-80˚C and is able to withstand prolonged (but not indefinite) incubation at temperatures above 90˚C (Sambrook and Russell, 2001). Within a few years, *Taq* polymerase had been cloned and overexpressed in *E. coli*, greatly expanding its availability. Today, the selection of polymerases available for PCR has increased dramatically, as new DNA polymerases have been identified in other thermophilic organisms and genetic modifications have been introduced into existing polymerases. Nonetheless, *Taq* polymerase is a robust and reliable enzyme that continues to be widely used for many procedures, including the protocol described later in this chapter.

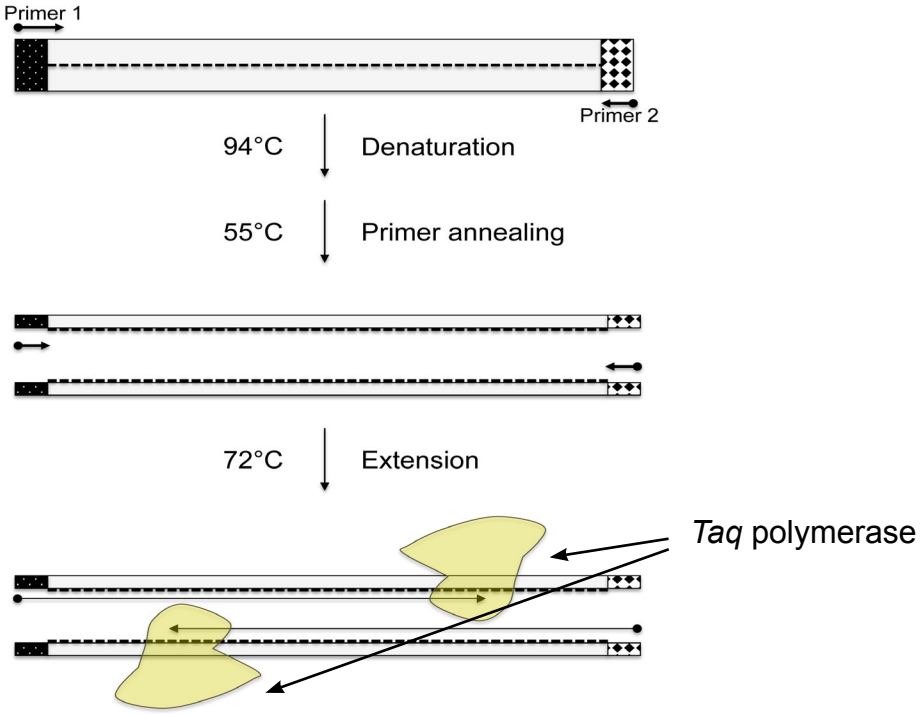
The thermostability of the DNA polymerases is critical to PCR. PCR involves multiple cycles of DNA synthesis that include temperature transitions from 55˚C to 72˚C to 94˚C. During each cycle, the DNA polymerase processes along a single-stranded DNA template, extending the 3'-end of a DNA primer. At the end of each cycle, the double-stranded DNA produced by the polymerase must be melted into single-strands that can serve as the template for DNA synthesis during the following cycle. In the early days of PCR, investigators (or their students!) manually transferred tubes between three water baths programmed for one of the three temperatures used in the procedures. Today, PCR is performed in automatic thermocyclers that rapidly adjust temperatures in a block containing the PCR reactions.

Polymerase chain reaction overview

PCR reactions typically begin with an initial segment in which the reaction mixture is heated for several minutes at a temperature high enough to break the hydrogen bonds that hold the two strands of the DNA helix together. Effective denaturation of DNA is critical, because DNA polymerase requires single-stranded DNA to serve as a template. The first cycle of denaturation is longer than subsequent denaturation steps, because PCR templates are often long, complex molecules held together by many hydrogen bonds. In subsequent cycles, shorter templates will predominate (see below).

Following the initial denaturation, PCR involves a series of 30-35 cycles with three segments, as outlined and discussed in greater detail below.

- A denaturation step - commonly 94˚C
- A primer annealing step - commonly 55˚C
- An extension step - commonly 72˚C



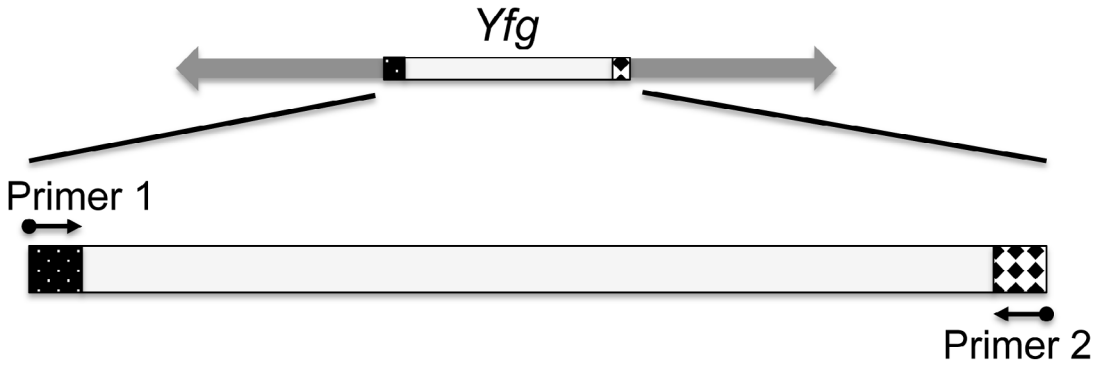
PCR reactions include multiple cycles at three different temperatures
Each cycle of a PCR reaction includes three different temperatures. During the first denaturation step, the hydrogen bonds holding DNA helix together are broken. In the following annealing step, oligonucleotide primers bind to single-stranded template molecules, providing starting points for processive DNA polymerases that will extend the primer sequence. DNA polymerases become more active at the extension temperature, which is close to its temperature coptimum. Investigators frequently adapt the temperatures and times of the steps above for different primers, templates and DNA polymerases.

Primer design

Good primer design is critical to the success of a PCR reaction. When designing PCR primers, researchers need accurate sequence information at both termini of the projected PCR product. PCR works best when the primers are highly specific for a target region in the template DNA. Specificity is particularly important when the template DNA is very complex, such as genomic DNA. The more complex the template, the greater the probability that it will contain another sequence that is very similar to a primer sequence. False priming from mismatched hybrids can lead to PCR products with the wrong sequence. To guard against false priming, investigators often perform a BLAST comparison of the primer sequences against the target DNA. BLAST is an alignment program that detects close matches between DNA sequences (see Chapter 25). If a particular primer will bind to multiple sites in the target DNA, it's best to redesign the

primer or alter the temperature cycling parameters to favor hybrids with greater stability. With low-complexity templates like plasmids, false priming is rarely a problem.

Two primers are required for a PCR reaction. Both primers will be extended at their 3' ends by DNA polymerase, so they must be designed with opposite orientations. During the annealing step, the primers will bind to opposite strands of the template DNA. The figure below shows an experiment designed to amplify the sequence of *Your favorite gene*, *Yfg*. The two primers that will be used to amplify the *Yfg* sequence are located on either side of the gene. If the experiment is intended simply to detect the presence of the gene, investigators will have greater choice in selecting priming sites, and the primers may correspond to sequences outside or within the coding sequence of the *Yfg* gene. If the PCR products will be used to place the *Yfg* gene in a new genetic context, the primers will probably need to precisely correspond to a particular stretch of nucleotides. For example, PCR was used to place the *MET* genes into the BG1805 plasmids that we are using in this project. One primer for each *MET* gene included the initiation codon and the codons for several other N-terminal amino acids, while the second primer included the reverse complement for the last few codons in the *MET* gene. Because the BG1805 plasmids encode fusion proteins, the stop codons of the *MET* genes were not included in this second primer (Gelperin *et al.*, 2005).



PCR primer design.
PCR requires two oligonucleotide primers with opposing orientations. The primers will bind to opposite strands of the target DNA during the annealing step, and they will be extended by DNA polymerase at their 3′ ends during the extension step.

The primers used for PCR are usually synthetic oligonucleotides between 18 and 25 bases long. When designing primers, it is important to estimate the T_m, or melting temperature, for the hybrid formed between the primer and its complementary sequence in the target DNA. (Many online calculators will help you to do this.) In general, the stability of a hybrid increases with the length of the primer and its GC content. The following formula provides a good approximation of the melting temperature of oligonucleotide hybrids for most experiments. More precise calculations depend on the actual nucleotide sequence and take thermodynamic parameters into account. In this formula, *n* refers to the number of nucleotides, and the concentration of monovalent cations is expressed in molar (M) units.

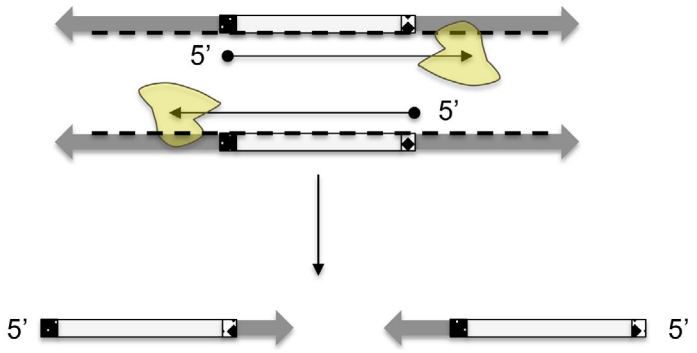
$$T_m = 81.5^{\circ}C + 16.6\,(\log_{10}([K^+] + [Na^+])) + 0.41\,(\%[G + C]) - (675/n)$$

When possible, researchers design primers that are similar in length and have a 40-60% GC composition. The sequences are designed so that the T_ms of the primers are a few degrees higher than the annealing temperature. Adjusting the T_ms of the primers to be close to the annealing temperature favors specific hybrids over less specific hybrids that may contain a few mismatched bases. This consideration is especially important with more complex templates, such as genomic DNA, which may contain other non-target regions with similar sequences to the primers.

PCR products accumulate exponentially

PCR is indeed a chain reaction, since the DNA sequence of interest roughly doubles with each cycle. In ten cycles, a sequence will be amplified ~1000 fold (2^{10}=1024). In twenty cycles, a sequence will be amplified ~million fold. In thirty cycles, a sequence can be theoretically amplified ~billion fold. With this tremendous amount of amplification, it's not hard to see how a single molecule can be detected in a biological (crime scene?) sample by PCR. PCR reactions in the lab typically involve 30-35 cycles of denaturation, annealing and extension. To understand PCR, it's important to focus on the first few cycles. The first two cycles set the stage. Exponential amplification of the desired end-product only begins in the third cycle.

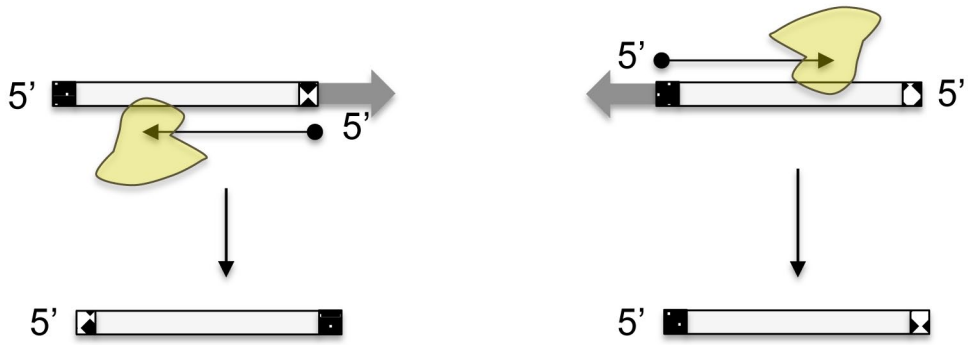
During the first cycle, the thermostable DNA polymerases synthesize DNA, extending the 3' ends of the primers. DNA polymerases are processive enzymes that will continue to synthesize DNA until they literally fall off the template. Consequently, the complementary DNA molecules synthesized in the first cycle have a wide variety of lengths. Each of the products, however, has a defined starting position, since it is "anchored" in the primer sequence. These "anchored" sequences will become templates for DNA synthesis in the next cycle, when fragments of the intended length first appear.



First cycle of PCR
During the first PCR cycle, DNA polymerases synthesize complementary strands of the template DNA, beginning at the primer site. The lengths of the products are quite variable and depend on the processivity of the DNA polymerase.

The starting template for PCR will continue to be copied in each subsequent cycle of PCR, yielding wo "anchored" products with each cycle. Because the lengths of the anchored products are quite variable, they will not be detectable in the final products of the PCR reaction.

DNA strands of the desired length first appear during the second cycle. Replication from the "anchored" fragments begins at one primer site and terminates at the second primer site. The number of these defined length fragments will double in each new cycle and quickly become the predominant product in the reaction.



PCR fragments of the desirect length first appear in the second cycle.
The "anchored" fragments generated during the first cycle of replication begin with either the primer 1 or primer 2 sequence. During the second cycle, replication begins at the other primer site, generating a PCR fragment capped at both ends with primer sequences.

Most PCR protocols involve 30-35 cycles of amplification. In the last few cycles, the desired PCR products are no longer accumulating exponentially. As in any enzymatic reaction, substrates have been depleted and the repeated rounds of incubation at 94˚C have begun to take a toll on *Taq* polymerase, which begins to denature.

Designing the PCR reaction

PCR is an enormously versatile technique. A variety of DNA templates are used for PCR. Depending on the purpose of the experiment, investigators could choose to use genomic DNA, a plasmid or a cDNA (complementary DNA generated by a reverse transcriptase from mRNA). PCR can also be done with much cruder preparations of DNA, such as a bacterial or yeast colony. This kind of "colony PCR" provides a rapid method for screening the results of transformation reactions.

The number of target sequences in the template should be carefully considered when setting up a reaction. PCR has been used to detect a single target molecule, but reactions are usually set up with thousands of target molecules to act as substrate. The presence of many target molecules "kick starts" the reaction, but can work against the investigator. Remember that the initial cycle of PCR generates products of indeterminate length from the original template. The original template continues to be copied in all subsequent cycles. Over time, this background replication will deplete reagents that could be used to synthesize the desired product, and undesirable side-products will accumulate. The presence of too much template will also promote mis-priming. Even though the primers hybridize with lower affinity to mismatched sites, equilibrium considerations will cause more mismatched hybrids to form.

It's easy to calculate the nubmer of target molecules when purified DNA is being used as the template for PCR, since the DNA concentration is already known. The molecular weight of a DNA base pair is 660 g/mole. Use the length of the template DNA (NOT the target DNA, which resides within the template DNA) and Avogadro's number to calculate the number of molecules in a reaction mixture. Much more DNA will be required when using a complex template, such as genomic DNA, than a simple template, such as a plasmid, for PCR. Imagine that you have a 6 kbp plasmid carrying a yeast *MET* gene. The yeast genome, with a size of 12 Mbp, contains a single copy of the same *MET* gene. You would need 2000 times more yeast genomic DNA than plasmid DNA to obtain the same number of target molecules.

It's much more harder to estimate the number of target molecules in colony PCR, an important screening technique in the laboratory. In colony PCR, cells are transferred with a sterile toothpick from a colony to the PCR reaction. Cells should be barely visible on the toothpick. A typical yeast colony probably contains close to a million cells, so the toothpick will probably be carrying many thousands of cells. In colony PCR, the temperature cycling program is adjusted to include a longer initial denaturation period. The high heat breaks open the cells and releases the DNA into the reaction mixture. Many potentially interfering substances are also present in these crude extracts - another reason to avoid transferring too many cells.

The components of a PCR reaction are simple, consisting of the DNA template, primers, dNTPs, a buffer containing $MgCl_2$ (polymerases use dNTPs complexed with Mg^{2+}), and the thermostable polymerase. For our experiments, we'll be using a master mix that contains all of the components except the template DNA and the primers. The use of a master mix ensures that all reactions have identical reagents and it also reduces the number of transfers requiring micropipettes. The smaller number of transfers is particularly advantageous, because it reduces the opportunities for cross-contamination of reagents. PCR is an exquisitively sensitive procedure. Many researchers use special barrier tips for their micropipettes, which contain filters that prevent samples from reaching the barrel of the micropipettes.

Protocol

PCR using a commercial master mix

This protocol describes very general conditions that can be easily adapted for particular templates and target sequences. The protocol uses a commercial master mix, which contains all the components of the PCR reaction, except the primers and DNA template. The final concentration of *Taq* polymerase is 25 Units/mL; the concentration of each dNTP is 200 μM and $MgCl_2$ is 1.5 mM.

NOTE: Gloves should be worn when setting up PCR reactions. PCR is an exquisitely sensitive procedure and care must be taken to prevent contamination from foreign DNA.

1. Label a 0.5 mL microcentrifuge tube for each reaction. Use a labeling code that takes into account the primers and template that will be used in the reaction. Be sure to write the record the meaning of the code in your lab notebook.

2. Set up an additional tube to serve as a blank. The blank will contain all of the reaction components ***except*** the DNA template.
3. For a 25µL reaction, combine:

Component	Volume	Final Concentration
PCR Master Mix, 2X	12.5 µL	1X
Primer 1, 10µM	1 µL	0.4 µM
Primer 2, 10µM	1 µL	0.4 µM
DNA Template*	5 µL	variable*
Nuclease Free Water to	5.5 µL	

* The concentration of DNA depends on the kind of template being used in the reaction. Use concentrations of 1 µg/mL for yeast genomic DNA, 0.1 µg/mL for bacterial genomic DNA, and 5 ng/mL for plasmid DNA. The reactions will contain 5 ng, 0.5 ng and 25 pg of DNA, respectively.

4. Program the thermocycler. Note that the optimum extension time at 72˚C depends on the length of the target sequence. Allow 1 minute per kb of target length.

   Initial denaturation segment: 94°C for 2 minutes
   35 cycles of: 94°C for 30 seconds (denaturation)
   55°C for 30 seconds (primer annealing)
   72°C for 1 minute (extension)
   Final extension segment: 72°C for 5 minutes
5. Turn off the thermocycler when the reaction is over. Store the reaction tubes in the refrigerator or freezer.
6. Analyze 5 µL aliquots of the reaction products on 1% agarose gels.

References

Gelperin, D.M. White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H., Gerstein, M., Dumont, M.E., Phizicky, E.M., Snyder, M., and Grayhack, E.J. 2005. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Develop.* **19**: 2816-2826.

Sakai, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Amheim, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.

Sakai, R.K., Gelfand, D.H., Stofel, S., Scharf, S.J., Higucki, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.

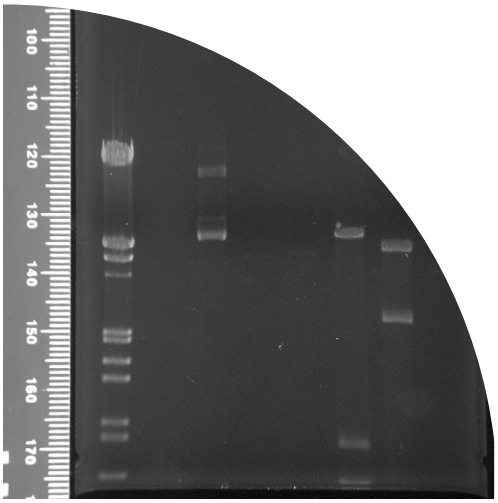
Sambrook, J., and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd Ed. , Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

20

Agarose Gel Electrophoresis

Objectives

- Understand the theory of nucleic acid separation by electrophoresis
- Know how to prepare and run an agarose gel
- Understand how intercalating dyes are used to visualize nucleic acids
- Know how to calculate the sizes of nucleic acids from agarose gels



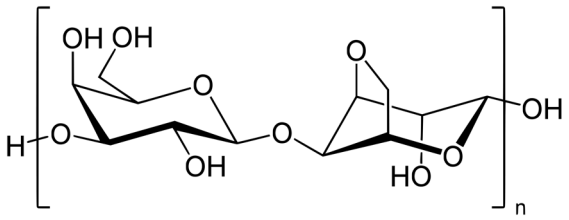
Agarose gels provide a simple and reliable method for separating nucleic acids according to their sizes. The positions of nucleic acids in the gels can be determined by staining the gels with intercalating agents that fluoresce when bound to DNA. This chapter will introduce you to the theory and practice of DNA electrophoresis.

Background

Agarose gels provide a simple method for analyzing preparations of DNA. Although the base compositions of individual DNA molecules vary, the basic chemical structure of DNA is the same for all DNA molecules. DNA molecules share the same charge/mass ratio, which imparts similar electrophoretic properties to DNAs of widely varying lengths.

Agarose gels are porous matrices

Agarose is a polysaccharide purified from red algae, or seaweed. Agarose is more highly purified (and significantly more expensive!) than agar, which is obtained from the same seaweed. Agarose molecules are long, linear polymers of the repeating disaccharide (1 3)-β-D-galactopyranose-(1 4)-3,6-anhydro-α-L-galactopyranose (right). A typical agarose molecule contains over one hundred monomers. Agarose used for electrophoresis has been highly purified. The purification process removes contaminants that would interfere with the enzymes used in molecular cloning, such as restriction endonucleases. The process also generates an agarose preparation with desirable electrophoretic properties and minimal background fluorescence, which is important for visualizing DNA molecules.



Agarose molecules are able to form gels with relatively defined pore sizes because of the chemical properties of agarose molecules. Agarose demonstrates hysteresis - its melting temperature is higher than its gelling temperature. Agarose molecules dissolve at about 90°C, forming random coils in solution. Gels form when the temperature falls to approximately 40°C. As the gel forms, the agarose molecules first assemble into helical fibers, which then further aggregate to form networks of supercoiled helices stabilized by hydrogen bonds. The sizes of the pores, which typically range from 50 to 200 nm, depend on the concentration of agarose. As the agarose concentration increases, the average diameter of the pore decreases.

Several factors affect the migration of DNA through agarose gels

Because of the negative charge of the phosphate residues in the DNA backbone, DNA molecules move toward the positive pole (anode) of the electrophoresis apparatus. The uniformity of DNA structure gives molecules the same charge/mass ratio, which makes it straightforward to calculate the sizes of DNA fragments. (Molecular weights are not very meaningful with DNA fragments, since a single molecule can contain the sequences of several genes.) In general, the migration of DNA molecules is inversely proportional to the logarithm of their lengths, or number of base pairs.

Although the relationship between migration rate and DNA length holds true in most cases, the actual migration rate of DNA molecules in a particular experiment is affected by multiple factors. Some of these factors are intrinsic to the DNA molecules, while other factors relate to the electrophoretic conditions. Because of this variability, DNA standards with known sizes should always be run on the same gels with samples being analyzed. Importantly, the standards need to have a similar structure (e.g. linear or supercoiled) and be subjected to any chemical modifications as the DNA samples being analyzed.

Size and conformation of DNA

The migration rate of a DNA molecule depends on both its size and its conformation. Within a certain size range dictated by the gel conditions, the migration rate of linear DNA molecules is inversely proportional to the $\log_{10}$ (number of base pairs). The migration of more structured DNA molecules, such as circular plasmids, is much less predictable. The migration rates of these more structured forms are influenced by the density of coils, the presence of nicks, and other structural features.

Intercalating agents

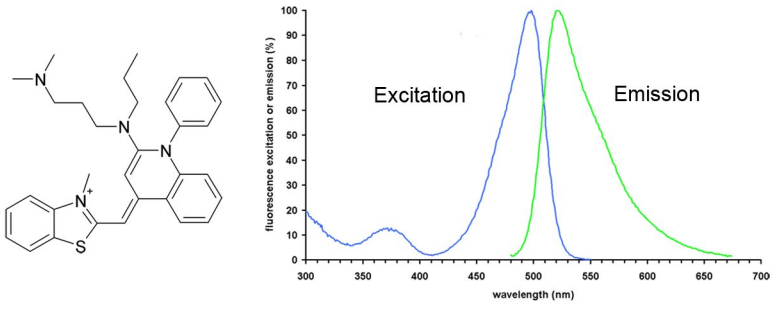
The migration rates of DNA molecules are reduced when they are complexed with intercalating agents such as ethidium bromide and SYBR™ Green. DNA molecules bind large quantities of these intercalating agents, which increases both the length and stiffness of the DNA molecule.

Buffer systems

The migration rates of DNA molecules are also affected by the composition of the gel and running buffer. The migration rate of a DNA molecule decreases as the concentration of agarose in the gel increases. The migration rate of a DNA molecule is somewhat higher in a gel made with TAE (Tris: acetate: EDTA) buffer than with TBE (Tris: borate: EDTA) buffer.

Fluorescent intercalating agents are used to visualize DNA fragments in gels

Nucleic acids are visualized by fluorescent dyes that bind strongly to DNA. The dyes are intercalating agents that insert into the DNA helix and into structured regions of single-stranded nucleic acids. The fluorescence of these dyes increases by an order of magnitude when they bind nucleic acids, so the background fluoresence remains low. The most commonly used dyes are ethidium bromide and variants of SYBR Green. DNA intercalating agents have an absorbance maximum in the long ultraviolet range, so they are viewed with transilluminators that emit light with wavelengths close to the absorbance maximum of the dyes. Bands on the gel have an orange-like color. The figure below shows the structure and spectral data for SYBR Green I, which is closely related to the SYBR Safe dye that we use in the lab. (The structure of SYBR Safe is a trade secret.) ***Be careful to wear protective eyewear if you look at an unshielded transilluminator. UV light is damaging to the eye!***



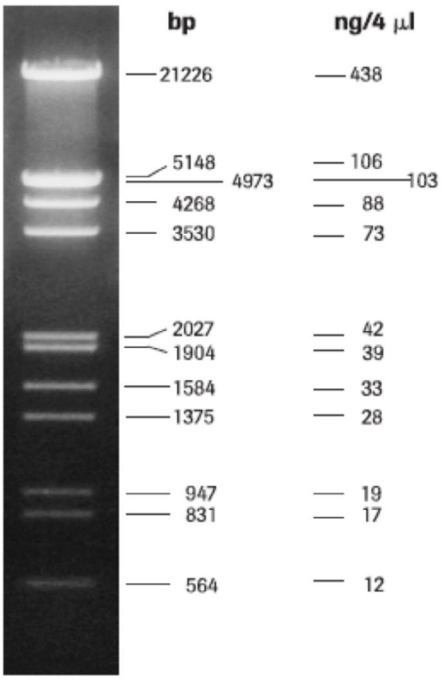
SyBr Green I fluorescent dye
Left: Chemical structure
Right: Absorbance and emission spectra of SyBr Green I. SyBr I has an absorbance maximum of 488 nm and an emission maximum of 522 nm.

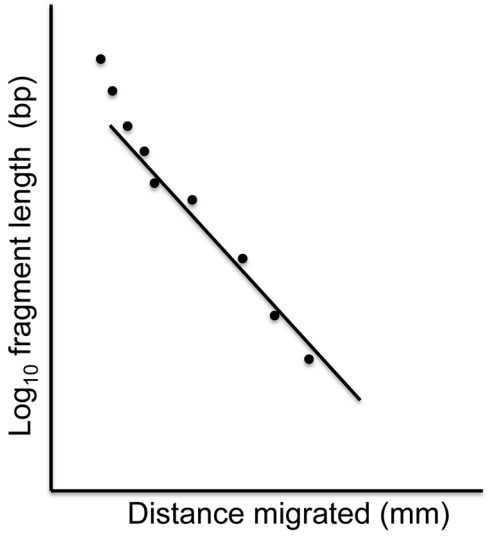
SAFETY NOTE: All DNA intercalating agents are potential mutagens. Be sure to wear gloves when handling solutions with intercalating dyes. We are using SYBR Safe for our experiments, which is reported to be significantly less mutagenic than other intercalating dyes, but you should still use precautions.

The sizes and quantities of DNA fragments can be calculated using standards

Agarose gels provide information about the sizes of DNA fragments, if they are calibrated with appropriate size standards. Agarose gels also provide information about the quantity of DNA in a particular fragment, because the amount of intercalating agent that a DNA molecule binds is proportional to its mass. Consequently, the intensity of a band reflects the amount of DNA in the band.

For our standards, we are using a preparation of bacteriophage λ that has been digested with two restriction endonucleases, *Eco*RI and *Hind*III. The intact λ phage is 48,502 bp in length. The enzymes generate a series of fragments, varying in length from 564 to 21.2 kbp, as shown on the right. Because all of the fragments are present in equimolar amounts, it is also possible to use this standard to visually estimate the quantity of DNA in an unknown sample. In the figure, the standard is diluted so that 4 µl contains 1000 ng DNA. (We'll use half this amount, or 500 ng, in our gels.)





To calculate the sizes of bands in other lanes of the gel, you first construct a standard curve like the one at the left from the migration of the λ restriction fragments. Plot the $\log_{10}$(number of base pairs) for each fragment on the y-axis against the distance that each fragment migrated on the x-axis. You will be able to estimate the sizes of fragments in your sample by interpolating on the standard curve.

Protocol

Plan the gel

As you plan your gel, there are several things to consider. The good news is that there is no single correct way to formulate gels. The choice of a particular system usually depends on convenience and expense, and a variety of different gels will yield satisfactory results. However, careful planning may lead to better resolution of DNA fragments in a particular size range. This section will lead you through the decision-making process.

1. Predict the sizes of the DNA fragments that you anticipate generating in your experiment. Use this information to guide your further decisions.
2. Choose an agarose concentration. The most common agarose gels are 1% w/v (1 g agarose/100 mL) gels, because this pore size allows good resolution of DNA fragments ranging in size from ~500 bp to 7 kb. The agarose concentration can be increased or decreased to allow better resolution of low or high molecular weight species, respectively. If you anticipate many fragments less than 1000 bp long, you may want to increase the agarose concentration *slightly*. (Gels may become rigid and hard to handle at concentrations of 1.5%.) Conversely, if you need to resolve DNA fragments greater than ~5 kb, you may want to use an agarose concentration as low as 0.5%. (Caution - these gels will be difficult to handle and may break easily.)
3. Choose a buffer system. Either of the two most commonly used buffer systems, TAE (Tris:acetate:EDTA) or TBE (Tris:borate:EDTA), will likely give you satisfactory results. Both TAE and TBE can be stored as concentrated stocks and diluted as needed. TAE is easier and less expensive to make, but TBE provides better buffering power.
4. Decide which intercalating dye you will use to visualize the gel. Ethidium bromide is an inexpensive, sensitive and reliable dye that has been used for decades. Newer SYBR Green dyes are more expensive than ethidium bromide, but they have been designed to have improved spectral properties. The SYBR Safe variant is also reported to be less mutagenic and more environmentally friendly than other common intercalating dyes, but it should still treated as a potential mutagen.
5. Decide if you will incorporate the dye into the gel itself or whether you will stain the gel after it is finished running. If you decide to incorporate the dye into the gel, you will be able to analyze the gel as soon as it is finished. On the other hand, dyes have a slight positive charge, so they will electrophorese toward the negative pole while the gel is running, reducing the resolution of small DNA fragments. The dyes will also diffuse into the running buffer, generating more hazardous waste. Post-staining is a more sensitive detection method, but it requires additional time. Bands are usually apparent after ~30 minutes, but sensitivity can be improved by increasing the staining time. Sensitivity can also be improved by destaining the gel in deionized water, which preferentially removes dye that is not bound to DNA. (Note: If no bands are apparent on a gel made with intercalating dye, the gel can still be post-stained to detect very faint bands.)

Practical note: Intercalating dyes are light-sensitive, so they should be stored in the dark. Dye fluorescence will also decay during prologed illumination of the gels with UV light.

Prepare the agarose gel

SAFETY NOTE: Wear safety goggles, lab aprons and disposable gloves when preparing agarose gels. Gloves are important when working with intercalating dyes, which are potential mutagens.

The following protocol refers to a situation where SYBR Safe is incorporated into a 1% agarose gel in 0.5 X TBE buffer. The protocol can be easily adapted for other agarose concentrations, dyes and buffers.

1. Determine the amount of agarose that you will need for your gel, which depends on the size of the gel frame. Most of the gel apparatuses in the lab are designed for 60 - 75 mL gels. Determine the volume that will work best with your apparatus, and calculate the amount of agarose you will need to weigh out to make a 1% (1 g/100 mL) gel.
2. Fill a graduated cylinder with the appropriate volume of 0.5 X TBE containing SYBR Safe. Pour the solution into a flask that accommodates at least twice the volume of the gel solution.
3. Weigh out the appropriate amount of agarose. Sprinkle the agarose onto the surface of the gel solution in the flask.

   Note: the agarose will not dissolve until it is heated.
4. Dissolve the agarose by heating the solution for 15-20 second intervals in a microwave oven. After each interval, remove the flask and gently swirl it around a bit to disperse the contents and help dissolve the agarose. Note whether undissolved agarose particles are still apparent or if the agarose has dissolved. Stop heating the solution as soon as the agarose is dissolved.

   The best gels are made from agarose that has NOT been overcooked.

SAFETY NOTE: The agarose solution will be very HOT when you remove it from the microwave! Fold several paper towels and wrap them around the neck of the flask when you handle it. Be particularly careful of steam escaping from the flask. If you do happen to spill some hot agarose on your skin, wash it immediately with cold water and alert your TA.

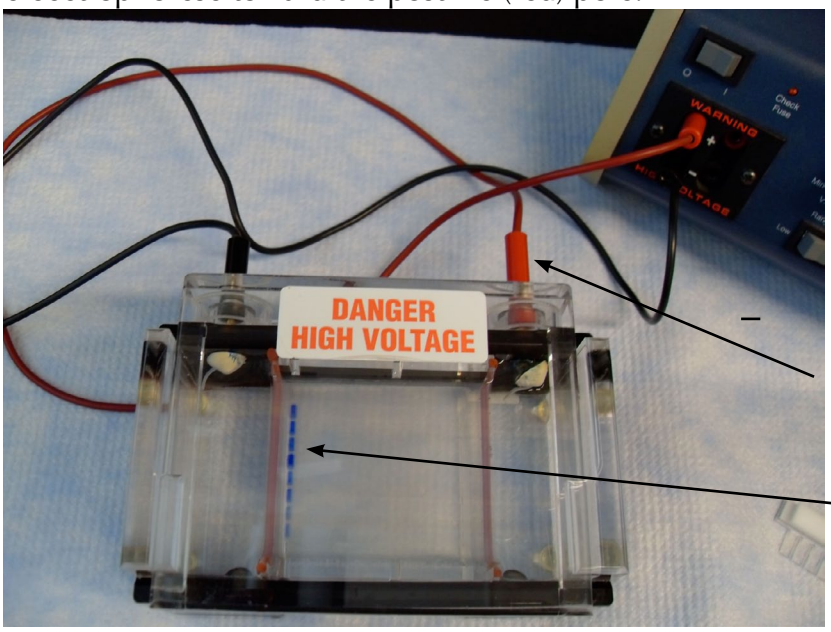
5. Allow the agarose solution to cool until you can comfortably touch the flask with your hands. The gel apparatuses are made of plexiglass, and they can become warped by solutions that are too hot. (Although we aren't incorporating ethidium bromide into our gels, this would be the time it was added.) Pour the gel. Place the sample comb in place. Do not move the casting platform until the gel sets. You will know that the gel has set when it becomes opaque. Allow the gel to cure for an additional 30 minutes or more after it has set.

Prepare the gel and the samples for electrophoresis

1. Mix each sample with 0.2 vol of sample buffer. For example, 1 µL sample buffer would be mixed with 5 µL sample. The sample buffer includes dyes that can be used to track the migration of samples as well as glycerol, which makes the samples dense enough to sink into the wells.The bromophenol blue and xylene cyanol dyes in the sample buffer migrate with "apparent" molecular weights of ~5 kb and ~0.5 kb, respectively. Convenient sample volumes are 5-15 µl, depending on which comb you use to form the well.
2. Obtain a DNA standard(s) for the gel. Each gel should include at least one lane with a DNA size marker. We usually use Standard III, a *Eco*RI/*Hind*III digest of λ phage, as a marker. Because the fragments are present in equimolar amounts, this standard can be used to estimate both the sizes and quantities of DNA fragments in a sample. The standard is diluted so that 5 µl, a convenient volume for your gel, contains 500 ng DNA. If you need added resolution in any size range, check to see if other markers are available that would provide greater resolution in that size range.
3. When the gel has set, carefully remove the comb. The holes left in the gel when the comb is removed will be the sample wells.
4. Orient the gel in the electrophoresis tank such that the wells are oriented toward the black (negative) electrode. Remember that DNA fragments will move toward the red (positive) electrode.
5. Carefully add 0.5 X TBE running buffer to the electrophoresis apparatus until it covers the gel and the buffer level is a few millimeters above the surface of the gel. If there is already running buffer in the apparatus, carefully submerge your gel until it rests on the platform. *These gels are commonly referred to as submarine gels for good reason!*

Proper set-up of an agarose gel

The gel is submerged in running buffer. Samples will elecctrophorese toward the postivie (red) pole.



Red electrical leads are attached to the positive pole of the power pack.

Samples are loaded into wells.

Run the agarose gel.

1. In your notebook, list the order in which you will be loading samples on to the gel. Be sure to include at least one lane for size standards.
2. Set a P20 pipette to deliver the total volume of your samples. Draw up the sample into the micropipette S L O W L Y. Make sure that there is not a bubble in the tip. It is important to avoid air bubbles when loading gel lanes, because bubbles may cause some of the sample to drift up out of the well. Note that you can use the same micropipette tips for all your experimental samples, but use a separate tip for the size standards. If you want, you can rinse out the tip between samples by drawing the running buffer into the tip a few times. Load one sample to each well.
3. Load 5 µL DNA size standards to one lane of the gel.
4. Place the lid on the electrophoresis tank and connect the electrodes to the power supply (black-to-black and red-to-red). Run the gel at a constant voltage of 100 V for ~35-40 minutes.
5. Stop the gel by turning off the power when the bromophenol blue (BPB) dye is ~1-2 cm from the end of the gel. Do NOT allow the BPB to run off the gel, because you will also lose small DNA fragments in your samples.

Analyze the gel.

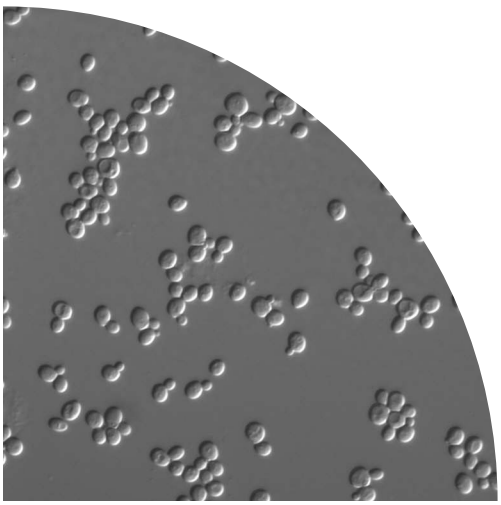
1. (This step is not necessary if interacalating dyes have been incorporated into the gel.) Stain the gel with ethidium bromide. Place the gel in a small tray with 100 mL deionized water. Add ethidium bromide to a final concentration of 0.5 µg/mL. ***Note: Ethidium bromide stocks are light sensitive, so they are stored in the dark.***
2. Place the gel on the transilluminator and photograph the gel. If the staining is very weak or a high background makes it difficult to visualize bands, you could place the gel back in the ethidium bromide solution for a longer time. If the background is too high, destain the gel by replacing the ethidium bromide solution with deionized water and gently rock the tray for ~30 minutes. The ethidium bromide solution and wash solutions should be disposed of in the appropriate hazardous waste container.
3. Construct a standard curve using data from the photograph. Measure the distance that each fragment migrated from the sample well. Make a table in your notebook with the size of each DNA standard and the distance that each fragment migrated on the gel. Plot the $\log_{10}$(fragment length) of each standard on the y-axis and the distance that each standard migrated on the x-axis.
4. Determine the approximate length of the DNA fragments in your samples by interpolating on the standard curve.

21

Preparing Yeast Extracts

Objectives

- Understand the experimental challenges to preparing cell extracts.
- Understand why detergents are used to prepare cell extracts.
- Know how to prepare a yeast cell extract that can be analyzed on SDS-PAGE gels.



Cells express thousands of proteins, which provide the molecular basis for cell function. Investigators analyze cell proteins to understand the pathways that underlie cell phenotypes. In this lab, you'll learn the experimental challenges that arise when preparing extracts from yeast as well as a simple protocol for preparing extracts that can be used for SDS-PAGE.

Background

Proteins comprise about half of the dry weight of most cells and include the myriad of structural proteins, catalysts, receptors and signaling proteins responsible for cell function. To understand cell function, scientists often want to analyze the protein composition of cells. Protein analysis begins with the preparation of a cell extract, ideally under conditions that minimize protein degradation. Preparing good cell extracts is something of an art, and many factors need to be considered during the design of an extraction protocol. In this course, we will be analyzing protein function in yeast. An average haploid yeast cell contains ~6 pg protein (Sherman, 2002) that perform the wide range of functions required for cell viability. Although yeast cells have many advantages for genetic studies, they are notoriously difficult to use for biochemical studies. Nonetheless, scientists have been able to develop procedures for extracting yeast proteins that circumvent many of these experimental barriers.

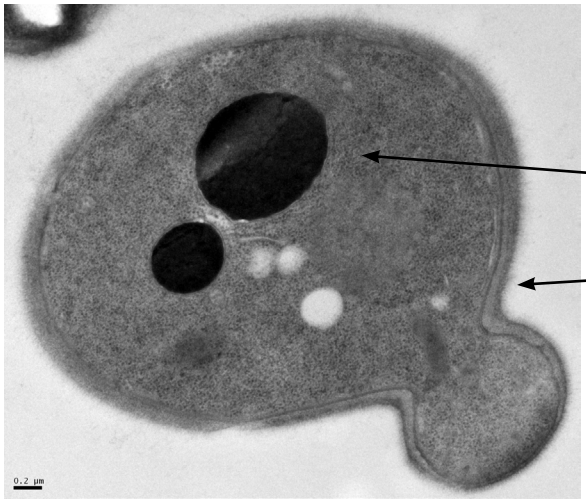


Figure 1.
Electron micrograph of an *S. cerevisiae* cell.

Vacuole contains many proteases.

Polysaccharide-based wall surrounds the cell.

Image by Christopher Bruser
Reproduced with permission
www.celllibrary.org
CIL10452

The first consideration in designing an extraction procedure is the compartmentalization of cells. All cells are surrounded by a plasma membrane and eukaryotic cells contain additional membranes that surround organelles. Many cells, including yeast (Figure 1), possess extracellular cell walls based on polysaccharides that protect cells against mechanical and osmotic forces. These cell walls appear as "fuzziness" surrounding the cell in electron micrographs. Cell extraction procedures begin with the disruption of the plasma membrane by mechanical and/or chemical treatments. Mechanical disruption of yeast cells must be fairly vigorous because of their tough cell walls. Mechanical methods commonly used to disrupt yeast include sonication, high pressure, and "beating" with glass beads. These vigorous treatments run the risk of denaturing proteins because of the heat and foaming generated during the processes.

Chemical treatments offer a gentler alternative to mechanical disruption for preparing extracts. Chemical extraction procedures frequently include detergents that solubilize membrane lipids, thereby allowing proteins to diffuse out of the cell. Most detergents do not discriminate

between intracellular and plasma membranes, so a detergent extract usually contains proteins from multiple organelles as well as cytoplasmic proteins. Investigators are able to choose from detergents that vary in their physical properties and in their denaturing/nondenaturing effects on proteins.

When preparing extracts, care must be taken to maintain the structural integrity of cell proteins. This necessitates inactivating the complement of proteases, which are responsible for normal turnover of proteins in cells. Cells contain proteases with a range of specificities, which become problematic during a protein extraction. Disruption of cell membranes releases proteases from compartments such as lysosomes, providing them access to cytoplasmic proteins. Cell disruption also interferes with more subtle biochemical controls on protease activity. Consequently, investigators often include protease inhibitors in the solutions they use to prepare extracts. Yeast are notoriously rich in proteases. In the intact cell, many of these proteases are located in the prominent yeast vacuole, which is analogous to the mammalian lysosome. The protocol that we'll use in this course (Amberg et al., 2005) relies on chemical disruption and rapid protein denaturation to denature the yeast proteases. The goal is to rapidly denature the proteases before they are able to digest other cell protines.. These extracts are suitable for SDS-PAGE and western blot analysis.

Protocol

Preparing Yeast Cell Extracts

This protocol has been adapted for yeast cells grown in liquid cultures. The procedure works well for quantities of cells up to 2.5 OD_{600}. Sterile technique is not required for the extractions.

Prepare cells for the extraction

1. Grow the cells under various conditions as instructed.
2. Determine the cell concentration. Transfer 100 µL of each cell culture to 900 µL deionized water and measure the OD_{600} of cultures in the spectrophotometer. Calculate the volume of the culture that you will need to obtain the number of cells equivalent to 2.5 OD_{600} units.
3. Collect the cells by centrifugation using either the table top centrifuge or the microcentrifuge. (This latter may require multiple spins in the microcentrifuge.) Your goal is to have 2.5 OD_{600} from each culture in a separate, *clearly labeled* 1.5ml microcentrifuge tubes. Collect the cells by centrifugation for 1 minute at top speed- in a microcentrifuge. Decant and discard the supernatant.
4. Rinse the cells. Add 1 mL deionized water to each tube. Resuspend the cell pellet by gently drawing the cells in and out of a micropipette tip, taking care to prevent premature lysis of the cells. This rinse step removes proteins from the culture medium that may be present in the cell pellet.
5. Collect the cells by centrifuging the suspension for 1 minute at top speed. Decant and discard the supernatant.
6. Resuspend the cells in 100 µL deionized water.

Prepare the protein extract

1. Transfer the cells to an appropriately LABELED 1.5 ml screw-cap microcentrifuge tube.
2. Add 100 µL of 0.2N NaOH to each tube, and incubate the cells for 5 minutes at room temperature. (The addition of sodium hydroxide does not lyse the cells, but it makes them more permeable and more fragile.)
3. Pellet the cells again in the microcentrifuge and remove the supernatant.
4. Resuspend the cells in 50 µl 2 X SDS-PAGE sample buffer*. IMMEDIATELY place the tubes in a boiling water bath. Leave the cells in the water bath for 3 minutes. This treatment effectively denatures the proteins. Yeast cells contain many proteases that could degrade other cellular proteins, so it's important that the proteases are denatured before they have the chance to attack other cellular proteins.

Note: The 2 X SDS-PAGE sample buffer contains 2-mercaptoethanol (also known as beta-mercaptoethanol, or BME). Use appropriate caution and work quickly when handling this reagent. BME is a volatile reagent with a strong odor reminiscent of rotten fish.

5. Remove insoluble cellular debris from the protein extract by centrifuging the extract for 2-3 minutes in the microcentrifuge. The insoluble residue in the pellet contains the cell wall, DNA and some insoluble proteins. Carefully remove the supernatant, which contains the proteins to a new, properly labeled microcentrifuge tube.
6. Store the samples in the freezer for future use.

*2 X SDS-PAGE sample buffer consists of:

120 mM Tris/HCl, pH 6.8
10% glycerol (glycerol is significantly more dense than water)
4% SDS
8% 2-mercaptoethanol
0.004% bromophenol blue (a tracking dye for electrophoresis)

References

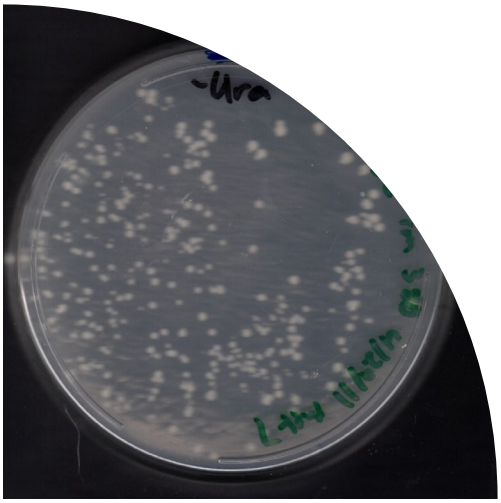
Amberg, D.C., Burke, D.J., and Strathern, J.N. (2005) Methods in Yeast Genetics (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). P. 121.

Sherman, F. (2002) Getting started with yeast. *Methods Enzymol.* **350**, 3-41.

22 Yeast Transformation

Objectives

- Understand the principles of microbial transformation
- Know how competent cells are prepared and used for yeast transformation
- Understand the differences associated with using plasmids and linear pieces of DNA for transformations
- Understand how selection is used in transformations



Techniques for transforming microbial organisms with foreign DNA are essential in modern molecular biology. Transformation allows researchers to move DNA between organisms and to construct novel genotypes. This chapter will introduce you to the techniques and challenges involved in transforming intact yeast cells.

Background

Transformation refers to the uptake of DNA by a cell, causing a change in its phenotype. Naturally-occurring transformation was first described in 1928 by Frederick Griffith (Griffith, 1928), who described a heat-stable "transforming principle" from virulent *Streptococcus pneumoniae* that could transform non-virulent *S. pneumoniae* to an encapsulated, virulent form. The "transforming principle" was subsequently identified as DNA by Avery and colleagues in 1944 (Avery *et al.*, 1944). Since then, transformation has become an indispensable tool in the molecular biology laboratory. The physical basis for transformation is still poorly understood, but researchers have empirically developed conditions that give fairly consistent transformation in the lab. Reliable transformation techniques have been developed for bacteria and many eukaryotes, ranging from yeast to mammalian cells. In the laboratory, cells that take up DNA are referred to as competent cells.

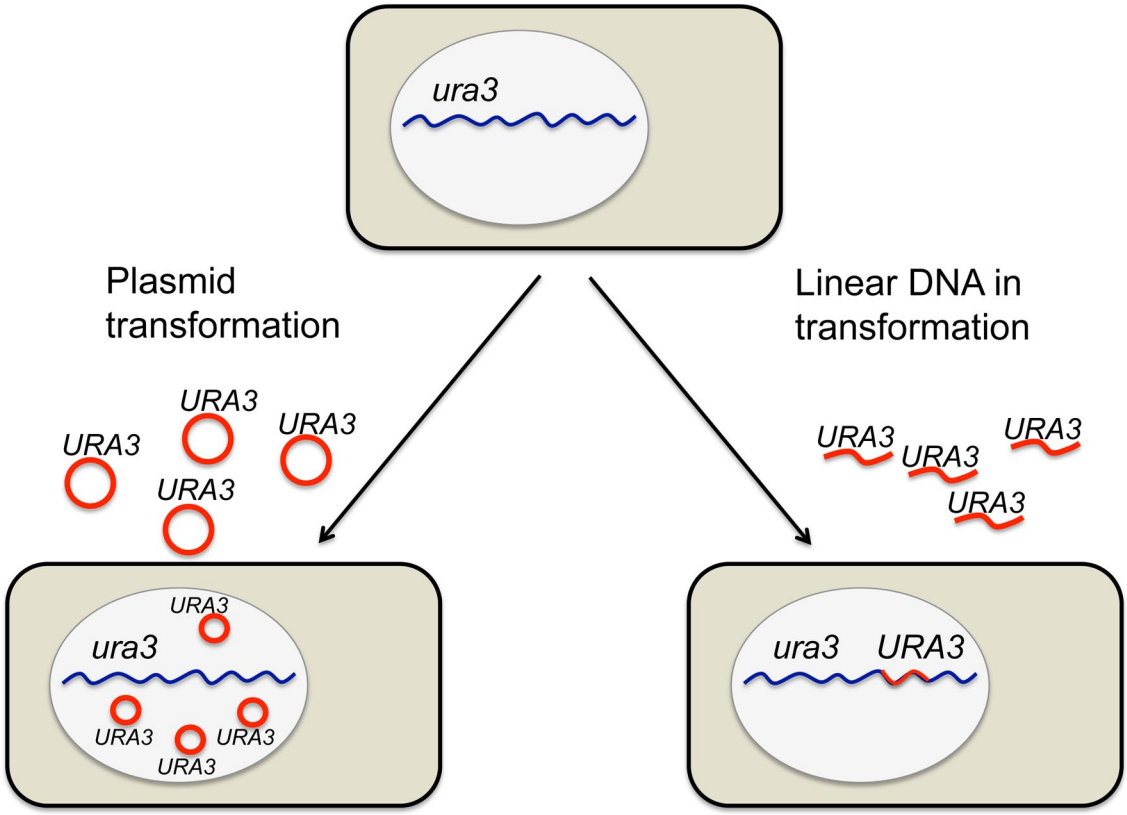
Competent cells are required for transformation

The challenge in laboratory transformation is devise conditions under which DNA will pass across the cell wall and plasma membrane of living cells, which are normally impermeable to DNA. Very few cells are naturally competent, or able to take up DNA. Consequently, researchers use a variety of chemical treatments to render cells competent. In general, these chemical treatments have some kind of destabilizing effect on the plasma membrane. The introduction of DNA into these competent cells can be further encouraged by a physical stress, such as a pulse of electric current or a rapid temperature change. Transformation is not a very efficient process, but because large numbers of microorganisms can be cultured in the laboratory, useful numbers of transformants can be obtained with most strains.

Techniques for yeast transformation are now standard in the laboratory. Depending on the details of the experimental procedure, reactions can yield as many as 10^6 transformants per μg DNA. The structure of the DNA used for transformation greatly affects the transformation efficiency. Transformation efficiencies are considerably higher with supercoiled plasmid DNA than with linear pieces of DNA. The reasons for this difference in efficiency are not entirely clear and probably arise from several factors. Linear DNA and plasmids may be taken up by cells at different rates. Linear DNA is more susceptible to destruction by nucleases than supercoiled plasmid DNA, which would lower its stability in the yeast cytoplasm. In contrast to plasmids, which have their own origins of replication, linear pieces of DNA also must be integrated into the chromosomal DNA before they can bring about the heritable change in phenotype associated with transformation. The differences between transformation with plasmids and linear pieces of DNA are outlined on the figure on the following page.

The most commonly used yeast transformation methods use a combination of lithium acetate, single-stranded DNA and polyethylene glycol (PEG) to generate competent cells. Although no one knows exactly how these components promote transformation, a number of hypotheses have been advanced. Lithium ions neutralize the negative charges on DNA molecules and the phospholipid bilayer, and they may also generate small holes in the plasma membrane that allow the passage of nucleic acids. Single-stranded DNA acts as a carrier for the DNA to be transformed and it may help to protect the latter from endonucleases. The source of the carrier DNA is unimportant. Since the carrier DNA concentration is considerably higher than that of

the DNA to be introduced into the cell, the carrier DNA is usually isolated from an inexpensive source, such as salmon sperm. It is particularly important that the carrier DNA for transformations be single-stranded. Either boiling (followed by rapid chilling to prevent reannealing of the helix) or base treatment can be used to denature the DNA helix before it is used for transformations. PEG may help bring the DNA into closer apposition with the membrane. PEG is often used to promote membrane fusion and is thought to alter water structure around the plasma membrane.



Transformation of yeast cells with plasmid DNA and linear pieces of DNA
Competent *ura3* yeast cells are transformed with either plasmid DNA with the *URA3* gene (left) or a linear DNA cassette with the *URA3* gene (right). Both transformants are able to grow on media lacking uracil. Strains transformed with plasmids must be continuously grown on media lacking uracil to preserve the plasmid. Strains transformed with linear DNA fragments stably incorporate the *URA3* gene into their chromosomes, so selective pressure is no longer needed.

Transformants are isolated on selective media

The DNA being used for transformation must carry a selectable marker whose presence can be detected by screening. Following the transformation, cells are plated on selective media that will allow transformed, but not untransformed, cells to grow. Consequently, it is critical to know the genotypes of the strains that one uses for transformations. For example, DNA carrying *URA3* genes must be transformed into ura3 mutants. In this case, the Ura3p encoded by the plasmid compensates for the deficiency in the *ura3* gene, an example of plasmid complementation.

Another common approach in transformation experiments is to introduce foreign genes that provide the recipient cells with protection against an environmental stress. The *S. cerevisiae* knockout strains that we are using this semester provide excellent examples of this kind of experiment. The knockout strains were generated as part of a systematic gene deletion project in which every *S. cerevisiae* ORF was replaced with a kanamycin resistance (*KAN*^{R}) gene. In the Saccharomyces Gene Deletion Project, researchers used PCR reactions to prepare transformation cassettes with a kanamycin resistance (*KAN*^{R}) gene for each of the *S. cerevisiae* genes. (See Chapter 15 for additional details.) Transformants were selected on plates containing kanamycin, and strains were established from single colonies.

The structure of the transformed DNA also affects the selection regimen that must be used to maintain the transformed strains. Transformed strains carrying plasmids are not stable under non-selective conditions. These strains must be propagated continually on selective media to maintain the plasmid. In the absence of selection, plasmid-bearing cells generally have a competitive disadvantage compared to cells without plasmids and will ultimately be lost from cultures. By contrast, strains derived by transformation and incorporation of linear pieces of DNA are stable. Selection is only required to establish the strains initially. Once the DNA has become incorporated into the chromosome, the marker is maintained by normal mitotic mechanisms.

Protocol

The following protocol is a slight modification of the "Quick and Dirty" transformation protocol described by Amberg *et al.* (2005). With careful attention to detail and cooperative strains, this procedure can yield thousands of transformants per μg plasmid DNA. Modifications to this method can increase its efficiency by several orders of magnitude. These high efficiency methods should be used to transform cells with linear pieces of DNA (Gietz and Schiestl, 2007).

Prepare a transformation master mix

1. Prepare a transformation master mix. The following ingredients provide enough reagents for five transformation reactions. Combine in a microcentrifuge tube:
   - 100 μL sterile 2 M lithium acetate (freshly prepared)
   - 400 μL sterile 50% PEG-3350
   - 4 μL 2-mercaptoethanol (STINKY!! add this in the fume hood!)

Set up individual transformation reactions - for each transformation:

2. Add 15 μL of the denatured salmon sperm DNA (2 mg/mL) to a labeled microcentrifuge tube.

Note: It is important for the salmon sperm DNA to be single-stranded for this procedure to work well. Boil the DNA for 5 minutes to denature the DNA. Quick chill the DNA by placing it immediately on ice. Keep the DNA on ice until you are ready to use it.

3. Add 5 µL of miniprep plasmid DNA to the microcentrifuge tube.
4. Add 100 µL of transformation mix from step 1 to the microcentrifuge tube. Vortex for 10-15 seconds to mix the contents.
5. Use a sterile toothpick, scrape a large yeast colony (or the equivalent of a "match head" of yeast) from a YPD plate. Transfer the yeast to the microcentrifuge tube containing the transformation/DNA solution (step 4) by twirling the toothpick several times. Be sure that the cells are uniformly suspended before proceeding.
6. Transfer the transformation mixture from step 5 to a sterile round bottom culture tube. (Replace the cap on the microcentrifuge tube and save it for step 8.)

Repeat steps 2-6 for each of your transformation reactions. ***Be sure to include a control with contains no plasmid DNA.***

7. Incubate the transformation mixture at 37˚C with shaking for 30 min.

Plate the transformed cells on selective media (YC-Uracil)

8. Transfer the transformations to sterile microcentrifuge tubes. Centrifuge the cells at 6000-8000 rpm for 15 seconds. Note: the cells are fragile after this chemical treatment and heat shock. Do NOT set the microcentrifuge for its maximum speed.
9. Carefully, remove the supernatant and discard it in the waster container. *GENTLY*, resuspend the cells in 100 µL sterile water by slowly pipetting up and down.
10. Remove 10 µL of the resuspended cells to 90 µL of sterile water in a microcentrifuge tube. This sample will be serially diluted for a spot plate (step 12) that you will use to calculate the transformation efficiency.
11. Plate the remainder of the mixture on a YC-uracil selective plate, using a spreader that you make from a Pasteur pipette. (Your TA will demonstrate how to make the spreader. Protective eyewear should be worn!) The spreader is sterile because of the flame. Touch the spreader to the plate to cool it off before adding the cells to the plate. Incubate the plates at 30˚C until colonies can be detected. The earliest that colonies will be visible is usually 2 days. If the colonies are small, allow them to grow an additional day at 30˚C. Count the number of cells on the plate.

Determine the number of viable cells in the transformation mixture.

12. Prepare a series of 4 additional dilutions of the cells set aside in step 10. Use these dilutions for a spot plate on YPD media. (See Chapter 14 for details on preparing spot plates.) Each row on the plate should contain cells from a different transformation reaction. Incubate the cells at 30˚C or room temperature until individual colonies can be detected. Do not allow the plate to overgrow, because you need to distinguish individual colonies.

Calculate the transformation efficiency

13. Calculate the fraction of cells that were transformed. For each transformation reaction, divide the number of transformed cells (step 11) by the total number of cells in the transformation mixture. Use the spot plate data from step 12, correcting for dilution, to calculate the total number of cells in the transformation mixture.

14. Transformation efficiencies are usually expressed by the number of cells transformed per μg DNA. Use your Nanodrop data to calculate the concentration of DNA used for the transformation. Use the data from step 9 to obtain the number of transformed cells.

References

Amberg, D.C., Burke, D.J., and Strathern, J.N. 2005. Methods in Yeast Genetics. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) Pp. 113-114.

Avery, O. T., MacLeod, C. M., and McCarty, M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation desoxyribonucleic acid fraction isolated from pneumococcus type III. *Journal of Experimental Medicine* **79**: 137–157.

Gietz, R.D., and Shiestl, R.H. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**: 31-34.

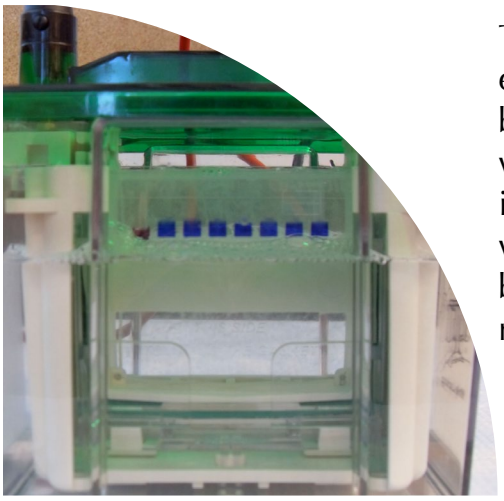
Griffith, F. 1928. The significance of pneumococcal types. *Journal of Hygiene* **27**: 113–159.

23

SDS-PAGE

Objectives

- Understand the theory governing protein separation by SDS-PAGE.
- Know how to cast and run SDS-PAGE gels.
- Understand how stains are used to visualize proteins in gels.
- Know how to estimate the molecular weight of a protein from its migration on SDS-PAGE gels



The sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) system described by Ulrich Laemmli in 1970 provides researchers with a sensitive method for separating proteins in complex samples according to their molecular weights. This chapter discusses the theoretical basis for the technique and provides protocols for running and analyzing SDS-PAGE gels.

Chapter 23

Background

Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) is a widely used method to analyze the proteins in complex extracts. The most commonly used systems are derived from the discontinuous SDS-PAGE system first described by Laemmli (1970). The system actually consists of two gels - a running (resolving) gel in which proteins are resolved on the basis of their molecular weights (MWs) and a stacking gel in which proteins are concentrated prior to entering the running gel. Differences in the compositions of the stacking gel, running gel and electrophoresis buffer produce a system that is capable of finely resolving proteins according to their MWs.

Gel electrophoresis of macromolecules

In gel electrophoresis, an electric field is used to move charged molecules through a matrix of a polymerized substance such as agarose or polyacrylamide. The rates at which individual molecules move through the gel depend on the properties of both the separation system and the molecules themselves. Gel matrices are permeated with networks of pores through which the molecules move. The amount of resistance that the matrix presents to the movement of a molecule depends on the diameter of the pore as well as the size and geometry of the molecule. Researchers can control the size of the pore by adjusting the concentration of gel monomer within a certain range. In general, smaller, more highly charged molecules migrate more rapidly through gels than larger or less charged molecules. The mobility of a molecule is also affected by the buffer system and the strength of the electrophoretic field, i.e. voltage, used for the separation.

We've already discussed the electrophoretic separation of DNA molecules in an earlier chapter. DNA molecules are much larger than protein molecules and they have a uniform charge to mass ratio. Consequently, electrophoretic analysis of DNA molecules is more straightforward than the separation of protein molecules. DNA molecules are usually separated in agarose gels, which have much larger pores than the polyacrylamide gels typically used to separate protein molecules. (However, small DNA molecules, e.g. less than 1000 bp, may be separated in polyacrylamide gels.)

Chemistry of acrylamide polymerization

Proteins are usually separated on polyacrylamide gels formed by the chemical polymerization of acrylamide and a cross-linking reagent (Figure 1), N,N'methylenebisacrylamide. Investigators are able to control the size of the pores in the gel by adjusting the concentration of acrylamide, as well as the ratio of acrylamide to bisacrylamide. Raising either the concentration of acrylamide or bisacrylamide, while holding the other concentration constant, will decrease the pore size of the gel.

Polymerization occurs because of free oxygen radicals that react with the vinyl groups in acrylamide and bisacrylamide. The oxygen radicals are generated from the catalyst, ammonium persulfate (APS). APS is able to generate oxygen radicals in the absence of other catalysts, when it reacts with a second catalyst, N,N,N',N'-tetramethylethylenediamine (TEMED). Aqueous solutions of APS are unstable, but can be stored for about a week in the refrigerator or for months in the freezer.

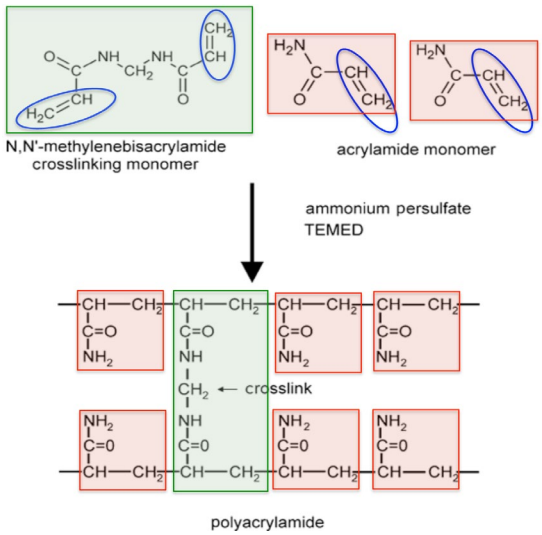


Figure 1.
Acrylamide polymerization

Proteins are denatured prior to electrophoresis

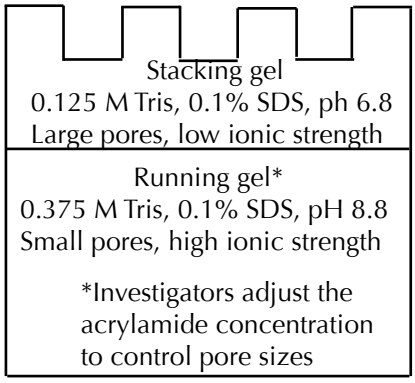
Compared to DNA molecules, proteins are structurally very diverse. In fact, crystallographers have described over 1000 structurally distinct protein folds. Proteins also show tremendous variation in their amino acid compositions and in the distribution of amino acids in the folded structure, features with important implications for electrophoresis. Because the hydrophobic effect is an important contributor to protein folding, hydrophobic amino acids tend to be more frequently found in the interior of proteins than on their surfaces. The surfaces of proteins often have a higher frequency of polar and charged amino acids than the interior of the protein, because these side chains are compatible with the aqueous cytoplasm. Even so, protein surfaces are mosaics with respect to the distribution of charged, polar and hydrophobic groups. Consequently, folded proteins lack the uniform charge to mass ratio that characterizes DNA molecules.

To resolve the proteins in a cell extract, investigators therefore need to reduce the structural complexity of proteins in their samples and to impart a uniform charge to mass ratio to the proteins. In SDS-PAGE, the solution is to denature the proteins by boiling them with the anionic detergent, sodium dodecyl sulfate (SDS) and 2-mercaptoethanol. The combination of heat and detergent is sufficient to break the many noncovalent bonds that stabilize protein folds, and 2-mercaptonethanol breaks any covalent bonds between cysteine residues. Like other detergents, SDS is an amphipathic molecule, consisting of a hydrophobic 12-carbon chain and a hydrophilic sulfate group. The SDS hydrocarbon chain binds to the many hydrophobic groups in proteins, reducing the protein to a random coil, coated with negatively charges along its length. Denatured proteins bind quite a lot of SDS, amounting to ~1.4 g SDS/g protein, or ~one SDS molecule for every two amino acids.

Discontinuities underlie the resolving power of SDS-PAGE gels

The Laemmli SDS-PAGE system can be considered a 3-component system. The stacking and running (resolving) gels have different pore sizes, ionic strength and pH (Figure 2). The third component is a chamber buffer (25 mM Tris, 192 mM glycine,, 0.1% SDS, pH ~8.3) containing large amounts of glycine. The ionization state of the glycine is critical to the

separation. At neutral pH, glycine is a zwitterion, with a negatively charged carboxyl group and a positively charged amino group. The pK_a of the amino group is 9.6, considerably higher than the pH of the chamber buffer. Consequently, very little glycine has a negative charge in the chamber buffer or stacking gel. Significant ionization does not occur until glycine enters the alkaline running gel. Let's follow the progress of protein samples during SDS-PAGE to see how differences in the composition of these three components generate the high resolving power of SDS-PAGE gels.



The sample buffer used for SDS-PAGE contains a tracking dye, bromophenol blue (BPB), which will migrate with the leading edge of the proteins being separated on the gel. The sample buffer also contains glycerol, which allows the protein samples to settle into the bottom of the gel wells. The gel is vertically positioned in the electrophoresis apparatus so that proteins will migrate toward a positive pole near the bottom of the gel. Once a voltage is applied, the chloride ions in the sample buffer and stacking gel begin to move rapidly toward the positive pole, forming the leading edge of a moving front of ions. Glycine molecules have very little charge in the stacking gel, so they migrate very slowly at the rear of the moving ion front. This difference in chloride and glycine mobility sets up a steep voltage gradient in the stacking gel that sweeps along the negatively charged protein-SDS complexes. The large pores of the stacking gel present very little resistance to the movement of protein-SDS complexes, which move swiftly through the gel and "stack up" into a very concentrated region at the interface between the running and stacking gels. Protein-SDS complexes remain concentrated at the interface until the slowly migrating glycine molecules reach the boundary between the two gels.

Dramatic changes occur as the glycine ions enter the running gel. The pH of the running gel is closer to the pKa of the glycine amino groups, so a significant fraction of the glycine molecules assume a negative charge. Negatively charged glycine molecules begin to move at the same rate as the chloride ions, thereby eliminating the voltage difference that controlled protein mobility through the stacking gel. The pores in the running gel are much smaller than those of the stacking gel, however, so the pores present frictional resistance to the migration of proteins. Proteins begin to migrate at different rates, because of the sieving properties of the gel. Smaller protein-SDS complexes migrate more quickly than larger protein-SDS complexes. Within a certain range determined by the porosity of the gel, the migration rate of a protein in the running gel is inversely proportional to the logarithm of its MW.

Proteins are visualized with stains.

With few exceptions, naturally-occurring proteins are invisible on SDS-PAGE gels. Consequently, researchers often use prestained protein standards to monitor the approximate positions of proteins during electrophoresis. Prestained standards are produced by covalently attaching large numbers of chromophores to a protein. The chromophores increase the MW of the protein and produce more diffuse bands on the gel. The diffuseness of the bands reflects variation in the number of dye molecules attached to individual protein molecules.

To visualize the positions of proteins after electrophoresis, investigators stain the gels with various dyes that bind to proteins (Steinberg, 2009). During the staining process, proteins are also "fixed" in the gel, meaning that they become insoluble and unable to diffuse out of the gel. The most commonly used dyes are the related Brilliant Blue R-250 and G-250 dyes, which bind proteins nonspecifically through ionic and Van der Waals interactions. Although Brilliant Blue R-250 staining is slightly more sensitive than G-250 for detecting proteins, G-250 staining procedure is more rapid and does not require destaining with organic solvents. We will stain our gels with the commercial Simply Blue™ reagent, which contains a colloidal suspension of Brilliant Blue G-250. In this procedure, gels are rinsed with water to remove the buffer salts used for electrophoresis and then treated with the colloidal G-250 suspension. Protein bands appear rapidly, and when necessary, the gels can be destained with deionized water to lower the gel background. Brilliant Blue staining intensity is considered a quantitative procedure, because the intensity of a stained band is directly proportional to the amount of protein in a band.

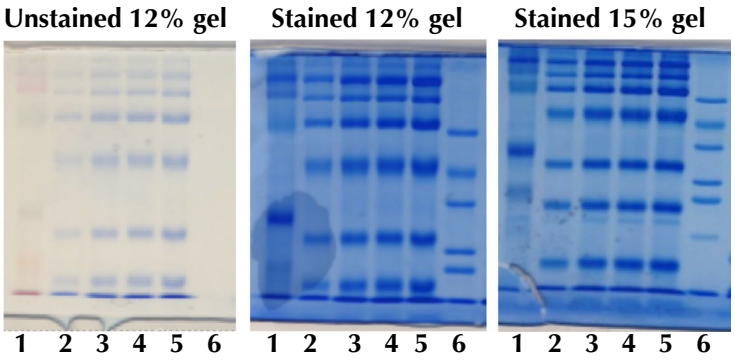


Figure 3. Molecular weight standards
1-Kaleidoscope protein standards (BioRad)
2-6.25 µg prestained protein standard
3-12.5 µg prestained protein standard
4-15 µg prestainded protein standard
5-18.75 µg prestainded protein standard
6-Unstained protein standard

Figure 3 illustrates these points. The same sets of unstained and prestained protein standards were separated on either 12% or 15% SDS-PAGE gels. The prestained standards in lanes 1-5 are visible without staining, but they become much more pronounced after staining. The unstained standard in lane 6 requires staining to become visible, but the bands are much more discrete and will give more reliable values wen calculating MWs of unknown proteins. The data in lanes 2-5 also demonstrate that Brilliant Blue staining is a quantitative procedure, because the intensity of bands in each lane increases in direct proportion to the amount of protein loaded in that lane.

Protein MWs can be estimated from their migration on SDS-PAGE gels

The sizes of proteins in an extract can be calculated by comparing their migration to a set of standard proteins run on the same gel. Researchers select standard proteins that are well-separated on the particular gel that they are running. For example, an investigator using a 7.5% gel will select standards with higher molecular weights than an investigator using a 15% gel, which is better suited to the analysis of small proteins. A plot of the $\log_{10}$MW of the standard proteins against the distance that each protein migrated on the gel (right) will give a

straight line in the region where the gel has good resolving power. (Note: MW is not the same as the mass of a protein. MW is a dimensionless term. For example, myoglobin has a mass of 16.7 kDa and a MW of 16,700.) The sizes of unknown proteins can be estimated by interpolating experimental values on a graph of standard proteins. Proteins whose MWs fall out side this range will not be well-resolved on the gel.

Protocol

SAFETY NOTE: Acrylamide and bisacrylamide monomers are neurotoxins. Gloves, goggles and aprons should be used when working with acrylamide.

Casting SDS-PAGE gels

These instructions will make two 12% or 15% SDS-PAGE gels that fit the BioRad Mini Protean system. Volumes can be easily adjusted for other gel systems and other gel concentrations.

Assemble the gel casting apparatus (Figure 4)

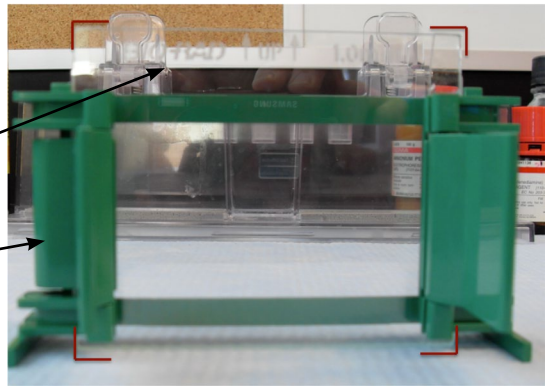
1. Assemble the components that you will need for casting the gel: a tall glass plate with attached 1 mm spacers, a small glass plate, a green casting frame and a casting stand.

Figure 4.
Assembling plates in the casting frame.

Large spacer plate with spacer goes in rear, oriented so the arrows point up and the logo is readable. Small plate is placed in front. (Corners of the plates are marked with bracket marks.)

Green gates of casting frame are open.

Casting frame "feet" and bottom edges of plates are flush against the benchtop.



2. Place the green casting frame on the bench with the green "feet"resting firmly against the bench and the clamps open (perpendicular to the frame) and facing you.
3. Place the two gel plates in the frame. Insert the taller spacer plate with the "UP" arrows up and the spacers facing toward you into the casting frame (the BioRad logo should be facing you). Insert the short glass plate in the front of the casting frame.

4. Secure the plates in the casting frame by pushing the two gates of the frame out to the sides.
5. Clamp the casting frame with glass plates into the casting stand, with the gates of the casting frame facing you.
6. Check to see if the assembled plates in the casting stand are sealed properly by pipetting a small amount of deionized water into the gap between the plates. If the glass plates hold water and don't leak, you are ready to make the gels. Pour the water out by holding the entire casting platform over a waste container and use paper towels or tissues to absorb any residual water. If the gel leaks, disassemble the frame, dry the plates and go back to step 3.

Preparing the running (resolving) and stacking gel solutions

1. Label two 15 mL conical tubes as "Resolving gel" and "Stacking gel".
2. Mix the acrylamide and buffer solutions for each gel as shown in the chart below. The directions are sufficient for a 5 mL resolving gel containing 12% acrylamide and a 5% stacking gel. Mix the ingredients gently, trying not to introduce air. Oxygen inhibits polymerization of acrylamide gels.

NOTE: catalysts should NOT be added to the mixture until you are ready to pour the gels.

Reagent	Resolving gel	Stacking gel
Deionized water	2.62 mL	1.4 mL
30% acrylamide:bis-acrylamide (29:1)	3.0 mL	0.42 mL
1.5 M Tris-HCl, 0.4% SDS, pH 8.8	1.875 mL	-----
0.5 M Tris-HCl, 0.4% SDS, pH 6.8	----	0.675 mL

Pour the resolving gel

NOTE: Be prepared! You will need to work quickly after adding the second catalyst - don't let the gel polymerize in your transfer pipette or the test tube!

1. To the resolving gel mixture, add 50 µL of a 10% ammonium persulfate (APS) solution. Gently mix the solution, trying not to introduce air. Oxygen inhibits acrylamide polymerization.
2. Add 5 µL of TEMED catalyst. Once again, gently mix in the catalyst trying not to introduce air bubbles.

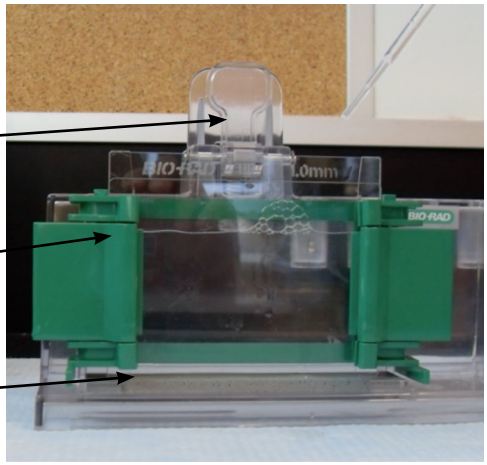
   CAUTION: TEMED has an unpleasant odor. Cover the tube immediately after you aliquot this reagent.
3. Working quickly, use a plastic transfer pipette to fill the space between the two plates until the resolving gel solution reaches a height just above the green clamps on the gel casting frame (Figure 5). Draw up any remaining acrylamide into the transfer pipet. You will know that the acrylamide has polymerized when you can no longer push the acrylamide out of the pipet.

Figure 5. Casting the SDS-PAGE gel.

Casting stand and plates are clamped into the casting stand

Fill plates with running gel solution to the top of the closed green gates of the casting frame.

Bottom edges of plates are flush against the grey pad of the casting stand.



4. Using a transfer pipet, add deionized water so that it gently flows across the surface of the polyacrylamide mixture. The water layer ensures that the polyacrylamide gel will have a level surface once it polymerizes.
5. Allow the gel to polymerize, which takes ~10-15 minutes. You'll note that the interface between the polyacrylamide and water overlay disappears as the gel polymerizes. A sharp new interface will form between the two layers, indicating that polymerization is complete. (You can also check the remaining polyacrylamide in the transfer pipette.)
6. When polymerization is complete, remove the water from the top of the resolving gel by tilting the gel to the side and using a paper towel or lab tissue to wick out the water.

NOTE: Polymerized gels can be saved for later use. If your group will be saving a gel, leave the water on top of the running gel. Remove the gel from the casting frame and wrap it in plastic wrap. Store the gel right-side-up in the refrigerator or cold room.

Pouring the stacking gel

1. Add 20 µL 10% APS and 5 µL TEMED to the stacking gel acrylamide mixture. Mix the contents by gently inverting the tube twice.
2. Use a transfer pipette to pipette the stacking gel on top of the resolving gel between the two glass plates. Add enough stacking solution until it just reaches the top of the small plate.
3. Carefully, but quickly, lower the comb into position, being careful not to introduce air bubbles. (The Bio Rad logo on the comb should be facing you.) Adding the comb will force some solution out of the gel, but this is fine. If air bubbles become trapped below the comb, remove the comb and reposition it.
4. Remove the comb when the gel has polymerized (10-15 minutes). You will be able to see faint schlieren lines (indicators of a change in refractive index) at the top of the spacer gel. You can also check the remaining solution in the 15 mL tube for polymerization.

Running SDS-PAGE gels

Setting up the electrophoresis apparatus

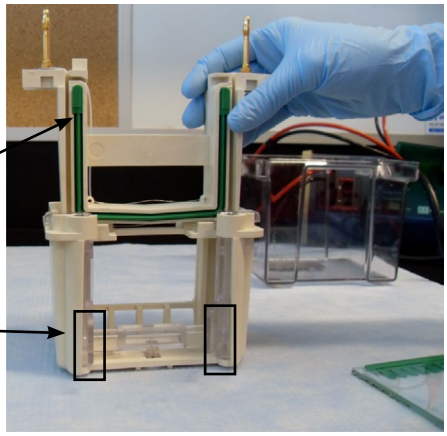
1. After polymerization of the stacking gel, remove the gel cassette sandwich from the casting stand.
2. Remove the casting frame from the gel cassette sandwich and place the sandwich against the gasket on one side of the electrode assembly, with the short plate facing inward (comb facing inward). Place a second gel cassette or a buffer dam against the gasket in the other side of the electrode assembly.
3. Slide the gel cassette sandwich and electrode assembly into the clamping frame (Figure 6).

Figure 6. Electrode assembly

One gel is positioned on each side of the electrode assembly

Top of short plate fits snugly into notch in the green gaskey of the electrode assembly

Electrode assembly with two gels is lowered into the clamping frame and clear plastic gates (rectangles) of frame are closed



4. Press down on the electrode assembly while closing the two levers of the clamping frame. Ensure you have and inner and outer chamber by either using two gels or a gel dam.
5. Lower the chamber into the electrophoresis tank.
6. Fill the inner chamber between the two gels with SDS-PAGE running buffer. This forms the upper chamber for electrophoresis.
7. Add SDS-PAGE running buffer to the outer (lower) chamber until the level is high enough to cover the platinum wire in the electrode assembly.

Loading and running samples on the SDSPAGE gel

1. Carefully remove the comb from the spacer gel. If you wish, place the sample loading guide in between the two gels in the electrode assembly before loading your samples.
2. Using gel loading micropipette tips (tips have very long, thin points and fit P20s or P200s), load up to 15 µL of sample into each well. Load 5 µL of a molecular weight standard into one lane of the gel. Load samples slowly and allow the samples to settle evenly on the bottom of the well.

NOTE: Be sure to record the order of samples loaded onto the gel.

3. Connect the tank to the power supply. Fit the tank cover onto the electrodes protruding up from the electrode assembly. Insert the electrical leads into the power supply outlets (connect black to black and red to red).
5. Turn on the power supply. Run the gel at a constant voltage of 120-150 V. Run the gel until the blue dye front nearly reaches the bottom of the gel. This may take between 40-45 min.

Staining SDS-PAGE gels

1. Turn off the power supply.
2. Remove the gel apparatus from the tank.
   NOTE: Do NOT discard the buffer. Running buffer may be used multiple times.
3. Open the clamping frame and remove the gel cassette sandwich. Carefully, pry the two plates apart with a spatula. With the spatula, remove the lower right corner of the gel as an orientation marker. Be sure to indicate in your lab notebook whether the notched corner corresponds to lane 1 or lane 10 of the gel. You may also remove the stacking gel with the spatula, if you desire.
4. Place the gel in a small plastic tray and lab the tray with you initials on a piece of tape. Fill the tray about halfway with deionized water. Gently free the gel from the glass plate into the water. The gel should move freely in the water.
5. Drain the water from the gel and then refill the tray about half-full. Rock the gel again for 5 minutes. Repeat this wash step one more time.
6. Drain the water from the gel and add ~20 mL of Simply Blue. Cover the gel container with plasatic wrap and rock overnight.
7. In the morning, drain the Simply Blue stain into an appropriately labeled waste container in the hood of the lab room. Destain the gel by filling the container about half full with deionized water.
8. When individual bands are detectable, record your data. You may photograph the gel with your cell phone camera. Alternatively, place the gel in a clear plastic page protector and scan the gel.
9. After recording the data, dispose of the gel in the Biohazard waste container.

References

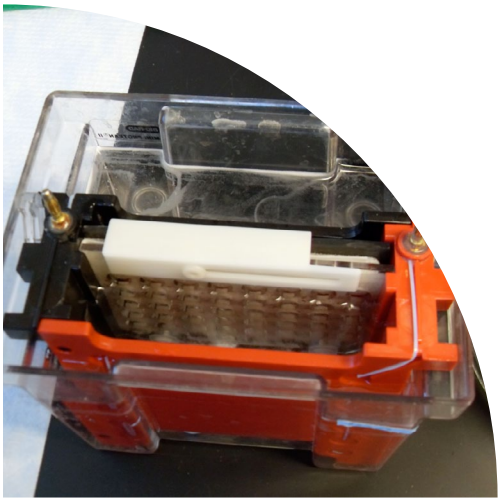
Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.

Steinberg, T.H. 2009. Protein gel staining methods: An introduction and overview. *Methods Enzymol.* **463**: 542-563.

24 Western Blots

Objectives

- Understand how monoclonal and polyclonal antibodies are produced.
- Understand the structural features of antibody molecules.
- Learn how to prepare a western blot.
- Understand how primary and secondary antibodies are used to detect proteins in western blots



Western blots are one of the most widely used techniques in cell biology. In a western blot, investigators take advantage of the exquisite sensitivity of antibodies to identify proteins of interest in complex samples. In this chapter, you'll learn about the different kinds of antibodies used in western blots and how to use western blots to detect proteins in cell extracts.

Background

Western blots provide a method to find the proverbial "needle in a haystack." A typical cell expresses thousands of different proteins, and it is often difficult to detect changes in expression of your favorite protein (Yfp) without a probe that is capable of discriminating the Yfp against a large background of unrelated cellular proteins. Fortunately, antibodies provide highly specific molecular probes that can be used to detect the expression of proteins on western blots. To appreciate the sensitivity of western blots, it's helpful to have some understanding of the generation of antibody molecules during immune responses and an appreciation of antibody structure. (Disclaimer: The following paragraphs provide a highly abbreviated overview of antibodies and one segment of the complex vertebrate immune system. The Department offers an immunology course that will introduce you to the finer details of this fascinating system.)

Antibodies are produced by lymphocytes in response to antigens

Antibodies are produced by vertebrates with adaptive immune systems capable of responding to foreign antigens. In the initial stages of the immune response, small numbers of immature B lymphocytes are able to bind foreign antigen molecules through antibodies that serve as receptors on the lymphocyte surface. Antigen binding stimulates the lymphocytes to proliferate and to differentiate into mature, antibody-secreting lymphocytes. Antigens act as selective agents. The lymphocytes that bind the antigen with the highest affinity receive the greatest growth signal and proliferate most rapidly, because a higher fraction of their surface antibodies are bound to antigen at any one time.

In the initial stages of an immune response, the antibodies on the surfaces of B lymphocytes bind antigen weakly. An amazing transformation occurs as B lymphocytes mature in response to antigen. Antigen binding stimulates responding lymphocytes to rearrange segments of their antibody-encoding genes, producing new potential antigen-binding sites. The scope of antibody diversity is immense - vertebrates are capable of producing billions of antibody molecules with distinct specificities. Most rearrangements are unproductive, but some rearrangements generate antibodies with greater affinity for the antigen. In the latter stages of differentiation, somatic hypermutation further increases the range of potential antibody sequences. Mature B lymphocytes, also known as plasma cells, have survived the selection process. Each plasma cell secretes a single antibody with high affinity for antigen. Plasma cells are virtual antibody factories that can be identified in electron micrographs by their extensive rough endoplasmic reticulum. (See the inside cover for a graphical depiction of this process.)

Polyclonal vs. monoclonal antibodies

An animal's response to antigen is ***polyclonal***, because antigens stimulate the proliferation of multiple lymphocyte clones, each of which produces a different antibody to the antigen. Consequently, serum collected from an immunized animal contains a mixture of antibodies that bind an antigen at distinct sites. These antibody-binding sites on antigens are known as ***epitopes***.

In contrast to polyclonal antisera, which are harvested from the serum of living animals, ***monoclonal*** antibodies are produced from cultured hybridoma cells. Hybridoma cells are generated by fusing lymphocytes from an immunized animal, most commonly a mouse, with myeloma cells capable of dividing indefinitely in culture. Standard culture techniques are used to

isolate individual hybridoma cell lines from the fusions, each of which secretes an antibody that recognizes a single epitope on an antigen. Hybridoma technology has revolutionized biomedical research since its description (Kohler and Milstein, 1975), because monoclonal antibodies recognize well-defined epitopes and because monoclonal antibodies can be produced indefinitely by cultured hybridoma cells. Investigators often use both monoclonal and polyclonal antibodies at different steps in western blots.

Antibodies bind antigens with high specificity

Myeloma cells have provided important tools for understanding the structure of antibodies, more properly described as ***immunoglobin*** molecules. Myelomas are cancers that arise from antibody-secreting plasma cells. Because myeloma cells secrete large quantities of antibodies with a single specificity, researchers have been able to harvest the antibody molecules and to prepare crystals for X-ray diffraction. Based on a large number of crystallographic studies, we now understand the basic architecture of antibody molecules.

The antibodies secreted by plasma cells are immunoglobins of the IgG class. Immunoglobins are named for their heavy chains, which dictate their role in the immune response to antigens. IgG molecules are tetramers consisting of two identical light chains and two identical gamma heavy chains (Figure 1). Both the light and heavy chains are comprised of smaller immunoglobin (Ig) domains, in which two beta sheets form a kind of sandwich held together by several disulfide bonds. Light chains contain two Ig domains, while heavy chains contain four Ig domains. The Ig domains closest to the N-termini of the heavy and light chains are hypervariable domains that have undergone rearrangement during lymphocyte maturation. These hypervariable domains form the antigen-binding site on the antibody. The other IgG domains in the light and heavy chains are constant domains also found in other IgG molecules.

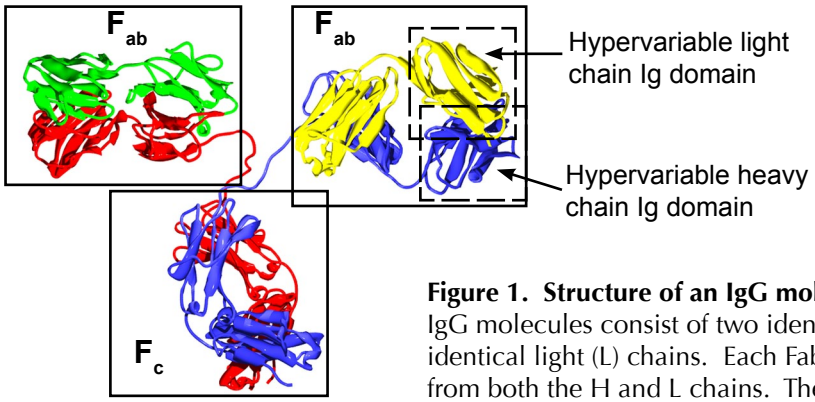


Figure 1. Structure of an IgG molecule
IgG molecules consist of two identical heavy (H) chains and two identical light (L) chains. Each Fab fragment contains two Ig domains from both the H and L chains. The N-terminal hypervariable domains form the antigen-binding sites. Fc fragments are formed by two Ig domains from each of the H chains. This figure is derived from PDB entry 1IGT (Harris *et al.*, 1997).

An IgG tetramer has a Y-shaped appearance. The two arms of the Y, which contain identical antigen binding sites are referred to as F_{ab} fragments, or antibody-binding fragments. The stem of the Y is referred to as the F_c fragment, or crystallizable fragment. All IgGs of the same subclass, e.g. IgG2, have identical F_c fragments.

Western blot overview

In a western blot procedure, proteins are first separated on an SDS-PAGE gel and then transferred to a membrane. This membrane replica is treated with antibodies that specifically recognize a protein or epitope of interest. Additional processing steps generate a signal at the position of the bound antibody. A typical western blot consists of multiple steps, which are discussed in greater detail below:

- Electrophoretic transfer of proteins from an SDS-PAGE gel to a membrane
- Blocking of nonspecific protein binding sites on transfer membranes
- Incubation of the membrane with a primary antibody specific for the epitope of interest
- Incubation with a secondary antibody specific for the primary antibody
- Detection of the primary antibody-secondary antibody complex

Depending on the sensitivity required and the reagents that are available, some of these steps may be combined.

Electrophoretic transfer of proteins from an SDS-PAGE gel to a membrane

The first step in a western blot is to generate a replica of the SDS-PAGE gel by transferring proteins electrophoretically to a synthetic membrane. Theoretically, antibody visualization could be done with a polyacrylamide gel, but membranes offer some distinct technical advantages. Polyacrylamide gels are mechanically fragile and can be easily torn during prolonged manipulations. In addition, the thickness of a polyacrylamide gel, which is 1 mm in our case, is greater than that of a membrane. Antibodies must also diffuse through the pores of the gel before they are able to bind their target molecule. By contrast, proteins are bound to the surface of the membrane, so antibodies don't need to diffuse between pores as they seek out their target epitopes.

The membranes used in western blots have a high protein binding capacity. A number of different types of membranes, including nitrocellulose, nylon and PVDF, are used for western blots. Today, the most commonly used transfer membranes are composed of polyvinylidine fluoride, or PVDF. (PVDF is similar to Teflon™.) PVDF is hydrophobic and doesn't wet properly with water. Therefore, PVDF membranes are first wet with methanol, then rinsed with deionized water, and finally rinsed with transfer buffer. PVDF membranes can be prepared while the SDS-PAGE is running and stored in transfer buffer until they are used. They must not be allowed to dry out during the transfer and immunoblot procedures. If they do dry out, they must be re-wet with methanol and rinsed with water before proceeding.

During the transfer process, the gel and membrane are placed directly against each within a "sandwich" of filter papers and foam pads, as described below. All of the filter papers and fiber pads are pre-wet with transfer buffer before they transfer cassette is assembled. The membrane is cut to just about the same size as the gel. During the transfer, current should flow evenly across the entire surface area of the gel. It's important, therefore, that air bubbles are not trapped between the gel and membrane.

With the miniblot system, proteins are transferred for either an hour at 100 volts or overnight at 20 volts. Considerable heat is generated in the process. Therefore, gels transferred at 100 volts are done with an ice pack in the transfer chamber, because rapidly cooling is needed. (Most of the ice will melt during the transfer.) Gel transfers at 20 volts are usually carried out

in a cold room. After the transfer is complete, membranes with the transferred proteins can be allowed to dry out and stored for later visualization with antibodies.

Blocking of non-specific protein binding sites on membranes

The transfer membranes used in western blots are designed to have a high protein binding capacity. The membranes also bind proteins nonspecifically, which is why they work so well for transfers. Before the membranes are incubated with specific (and expensive) antibodies, they must therefore be pretreated with blocking solutions that contain high concentrations of abundant (and cheap) proteins. If the transfer membranes are not adequately blocked before the antibody is applied, the nonspecific sites on the membranes will absorb some of the primary and secondary antibodies. Antibody binding to these nonspecific sites both reduces the amount of antibody available to bind the target proteins and increases the blot background. Effective blocking of these nonspecific sites increases both specific antibody binding (the signal) and the signal to noise ratio. Blocking proteins are also included in the antibody solutions used in western blots. Because the concentrations of blocking proteins are orders of magnitude greater than the concentrations of antibodies, they will out-compete the antibodies for binding to the many nonspecific sites on the membranes. The antibodies have a much higher affinity for their specific targets than for nonspecific sites on the membrane, so they will concentrate at sites with the target proteins.

Casein proteins from milk are a popular blocking reagent. Caseins are available in various degrees of purity. When a target protein is expressed at a very low level, it's best to use a more purified formulation of caseins. For most purposes, however, nonfat dried milk (direct from the grocery store!) provides a very adequate and inexpensive casein preparation. In addition to high concentrations of proteins, blocking solutions commonly contain a non-ionic detergent. Nonionic detergents reduce nonspecific binding to membranes, but they do not denature proteins. Our blocking solution consists of 5% nonfat dried milk dissolved in a buffer (TBS-T: Tris-buffered saline) with 0.05% Tween-20, a nonionic detergent.

The timing of the blocking step is not particularly critical. Nonspecific sites on the membrane are rapidly saturated with proteins, but it is a good idea to incubate the membranes with blocking solution for at least one hour. Blots can be left in blocking solution for prolonged periods of time, if necessary. In this latter case, blots should be incubated in the cold to prevent microbial growth on the proteins in the blocking solution.

Primary antibody binding

Either polyclonal or monoclonal antibodies can be used as the primary antibody on western blots. Antibodies can be directed toward a naturally-occurring protein or toward an epitope that has been added to a naturally-occurring protein using recombinant DNA technology. Increasingly, researchers are introducing epitopes into proteins to facilitate their detection on western blots, because antibodies against naturally-occurring proteins are expensive and time-consuming to prepare. By contrast, an antibody directed against an epitope can be used to detect many different proteins carrying that same epitope. In our western blots, we'll use monoclonal antibodies directed against viral epitopes that were added to the C-termini of Met proteins during the cloning process. The viruses that produce the epitopes are not indigenous to yeast. Virologists have previously shown that these particular epitopes, which are short amino acid sequences in viral coat proteins, elicit strong immune responses during viral infection. Commercial suppliers

have taken advantage of this information to produce very potent antibodies directed these amino acid sequences. The pBG1805-based vectors that we are using to overexpress *S. cerevisiae* proteins encodes a hemagglutinin (HA) epitope tag from human influenza virus (Sleigh et al., 1981). The pYES2.1-based plasmids that we will use to overexpress *S. pombe* proteins contains a V5 epitope tag from simian virus 5 (Southern et al., 1991).

An important consideration in western blots is the concentration of antibody solution required to detect the protein of interest. Antibody production ultimately depends on the immune response of an animal, which can show significant variability in the strength of the reaction and the spectrum of antibodies produced in response to the antigen. In addition, commercial antibody preparations contain proteins other than antibodies. Polyclonal antibodies are partially purified from animal serum and contain the whole spectrum of immunoglobins in the animal's blood. Monoclonal antibodies are purified from the culture medium of hybridoma cells and may also contain other proteins. Consequently, antibody suppliers measure the effective antibody concentration, or titer, of each antibody preparation. Researchers then use this information to plan their western blots. The primary antibodies that we are using in our experiments are monoclonal antibodies that will be diluted several thousand-fold. Because the primary antibody is monoclonal, a maximum of one antibody molecule will bind to each epitope-tagged protein. It is unlikely that all of the immobilized epitope-tagged proteins will bind antibody, since some are probably be inaccessible to the antibodies on the membranes.

Secondary antibody binding

Polyclonal antisera are commonly used as the source of secondary antibodies in western blotting. The antisera are prepared by injecting a second species with Fc fragments of IgGs from the species used to generate the primary antibody. Antisera designed to be used on western blots are then covalently linked to a reagent that aids in visualizing bound antibody. In some cases, secondary antibodies are bound to a fluorescent dye, so that blots can be analyzed directly with a fluorescent scanner. In other cases, the antibodies are linked to an enzyme, such as horseradish peroxidase or alkaline phosphatase, that can be used for histochemical detection of bound antibody in an additional step. The secondary antibody-binding step plays an important role in amplifying the signal on western blots. Because the antibody is polyclonal, secondary antibodies will bind to multiple sites in the Fc fragment of the primary antibody. If an enzyme is attached to the secondary antibody, the histochemical detection step will provide additional amplification of the signal.

Immunochemical detection of bound antibody

In this final step of the procedure, the western blot is incubated with substrates for the enzyme that has been conjugated to the secondary antibody. Because enzymes typically convert thousands of substrates to products every minute, this step can provide tremendous sensitivity to western blots. Horseradish peroxidase (HRP) and alkaline phosphatase (AP) are the two enzymes most commonly conjugated to secondary antibodies. Both are hardy enzymes with high turnover numbers that are relatively resistant to denaturation. HRP catalyzes the oxidation of various substrates by hydrogen peroxide into products that precipitate at the reaction site. AP removes phosphate groups from its substrates, converting them into insoluble products.

A variety of chromogenic and chemiluminescent substrates are available for both HRP and AP. Researchers consider sensitivity, cost and convenience in their choice of substrates. In general, chromogenic substates that generate visible reaction products are less expensive than chemiluminescent substrates, and they also require no special equipment to record the data. These procedures are easily carried out in the laboratory. On the other hand, procedures using chromogenic substrates are less sensitive than procedures using chemiluminescent substrates, which require X-ray film or specialized densitometers to detect reaction products.

In our experiments, we will use HRP-conjugated secondary antibodies with chromogenic substrates. HRP reacts with the two substrates, hydrogen peroxide and 3,3',5,5' - tetramethylbenzidine (TMB), to generate a dark blue-grey reaction product that precipitates at the reaction site. The amount of reaction product at a particular site on the transfer membrane is directly proportional to the number of HRP molecules bound to the same site. Colored reaction product accumulates with time until the reaction is stopped by washing away unreacted substrate. The reaction should be terminated before nonspecific antibody binding becomes problematic.

Protocol

The following pages describe a general protocol for performing western blots using PVDF membranes and the BioRad Mini-Protean apparatus. You may receive additional instructions that change some of the details involved with individual steps, particularly with respect to the timing. Be sure to incorporate those details in your western blot planning.

Separate proteins on an SDS-PAGE gel

Refer to Chapter 23 for specific instructions on running SDS-PAGE gels.

1. Separate the proteins that will be analyzed on western blots by SDS-PAGE.
2. Remove the electrode apparatus and holder from the tank, and remove the gel from the holder. Do not remove the gel from the plates until you are ready to assemble the transfer cassette (see below).
3. Pour as much of the SDS running buffer as possible back into the SDS running buffer flask, so it can be used again. Dispose of any remaining buffer down the sink. Rinse out the buffer tank with deioinized water to remove residual SDS, which can interfere with the transfer process.

Prepare the transfer membrane

NOTE: DO NOT touch transfer membranes with your fingers. Wear gloves and use filter forceps when you handle transfer membranes.

1. While the gel is running, gather the PVDF membrane and four pieces of thick filter paper, such as Whatman 3MM™. The PVDF membrane and filter papers should be cut to a size that is slightly larger than the SDS-PAGE gel. You will also need a transfer cassette and two fiber pads.
2. Prepare the PVDF membrane. Using pencil, place an orientation mark in a corner of the PVDF membrane for later identification. Wet the membrane by placing it in a small tray containing methanol for ~30-60 seconds with gentle agitation.

3. Dispose of the methanol in the waste container and add deionized water to the tray. Gently agitate for ~1 minute.
4. Replace the deioinized water with transfer buffer. Store the membrane in transfer buffer until you are ready to start the transfer.

Assemble the transfer cassette

1. Fill the buffer tank about half way to the top with transfer buffer. Wet the fiber sponges and filter papers for the transfer cassette by placing them into the transfer buffer in the tank. (Transfer buffer contains Tris and glycine, as well as 10% methanol. Unlike SDS-PAGE running buffer, the transfer buffer does not contain SDS.)
2. Using a spatula or a green plastic wedge, remove the small glass plate from the gel. The gel will remain attached to the large glass plate. With a spatula, remove the lower right corner of the gel to serve as an orientation mark. (This correponds to the first lane of your gel.)
3. Assemble the transfer cassette as shown in the Figure 2. Be sure that all parts of the transfer "sandwich" remain moist at all times.

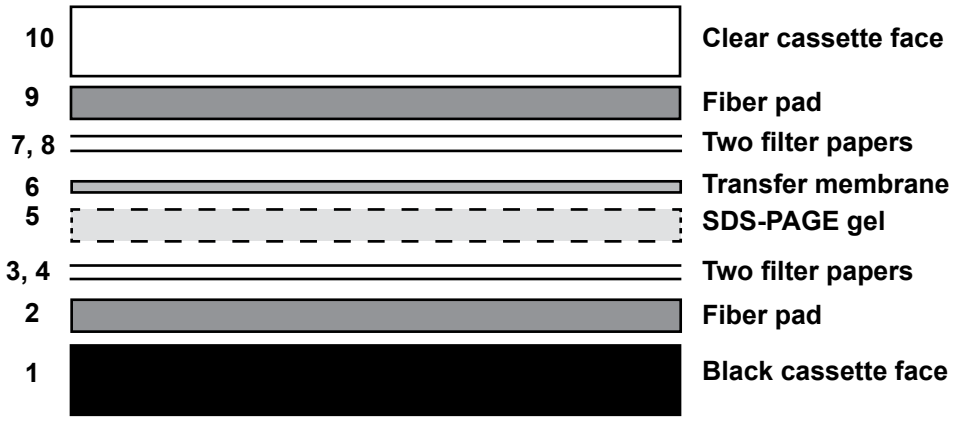


Figure 2. Assembly of the transfer cassette.

- Place a wet fiber pad on top of the black cassette face.
- Add two pieces of filter paper.
- Position the gel on top of the filter paper while it is still attached to the glass plate. Use a spatula to *carefully* release the gel from the plate. You may find it easier to remove the gel by beginning at the bottom edge near the dye front.
- Place the PVDF membrane blot on top of the gel. Orient the gel so that the pencil mark on the membrane corresponds to the clipped corner of the gel. Using a pencil or spatula, carefully roll out any trapped air bubbles between the blot and the gel.
- Add the remaining filter paper(s) and the fiber pad.
- Fold the clear cassette face over the gel assembly and slide the clamp into place.

Electrophoretic protein transfer

1. Place the transfer sandwich into the cassette holder with the black face of the transfer cassette aligned with black side of the cassette holder and the clear face aligned with red side of the cassette holder (Figure 3). NOTE: Each cassette holder can hold two transfer cassettes.
2. Place the cassette holder and assembled cassettes into the electrophoresis tank. If the transfer will be done at 100V (see step 5 below), the electrophoresis tank should contain an ice pack.
3. Fill the electrophoresis tank to the top with transfer buffer.
4. Place lid on tank by aligning black with black and red with red.

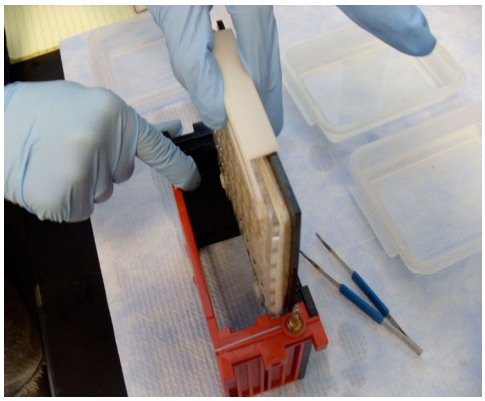


Figure 3.
Inserting the transfer cassette into the cassette holder.

5. Run the transfer at 100V for 1 hour at room temperature or at 20 V overnight in the cold room.
6. When the transfer is complete, remove the transfer cassette from the tank. Pour the transfer buffer back into its original bottle so that it can be reused.
7. Disassemble the transfer cassette. Depending on your schedule:

- If you will be continuing with the western procedure, skip the rehydration step (step 1) below and continue with the blocking step (step 2). ***Be careful that the membrane remains moist!***
- If you will be processing the membrane at a later time, allow the membrane to dry out. Wrap the membrane in plastic wrap and save it for a later lab period.

Immunoblotting - This is a multi-day procedure. Timing may vary for different classes.

Membranes are rehydrated and treated with blocking reagents

1. Wearing GLOVES, unwrap the dry blot from the plastic wrap. Use the Kaleidoscope standards to identify the side of the membrane to which the proteins are bound. Submerge the membrane in methanol with this side facing up. Gently agitate the membrane by hand rocking for 30-60 seconds until the membrane has been uniformly wet with methanol. Decant the methanol into the appropriate container and fill the tray half way full with deionized water. Gently agitate the membrane for an additional minute.
2. Decant the water and replace it with sufficient TBS-T (Tris buffered-saline containing 0.05% Tween 20) to cover the blot. Place the blot on a rocking platform. Equilibrate the blot in TBS-T for 5 minutes with slow rocking. At the end of 5 minutes, drain the TBS-T into the waste container.
3. Pour 50 ml of blocking solution (5% nonfat milk in TBS-T) onto the blot and cover the tray with a small piece of plastic wrap.
4. Place the tray on a rocking platform in the cold room (Higgins 377). The blot should float freely in the tray so that both sides are washed. Incubate the blot for at least an hour or up to 24 hours at 4˚C.

Membranes are washed and incubated with primary antibody (~30 minutes)

1. Locate your blot in the cold room and bring it back to the lab room.
2. Remove the plastic wrap from the container holding the blot and pour off the blocking solution in a waste container. SAVE the plasticwrap! You will need it to cover the container again!

3. With a graduated cylinder, measure 30 ml of TBS-T solution. Pour the TBS-T solution onto the blot and place the container on the rocking platform. Rock for 5 minutes.
4. Pour off the TBS-T and add a fresh 30 ml of TBS-T. Wash, with rocking, for 5 minutes.
5. Repeat step 4, for a total of three washes.
6. Pour off the TBS-T. Add 15 mL of primary antibody diluted in blocking buffer.
7. Cover the container with the same piece of plastic wrap and place the tray on the rocking platform in the 4˚C cold room. Make sure that the blot floats freely in the tray and that the standards are on the top face of the blot. Incubate overnight at 4˚C with slow rocking. ***NOTE: The timing of this step is the most critical in the procedure. Shortening the incubation time with primary antibody may reduce the sensitivity of the western blot.***

Secondary antibody binding and detection (1.5-2 hours)

1. Locate your blot in the cold room and bring it to your lab classroom.
2. Carefully drain the antibody from the blot into the test tube marked "Used primary antibody". (Antibodies are expensive. Fortunately, the solutions can be re-used.)
3. Fill the tray with the blot about half-full with TBS-T. Place the tray on a rocking platform and wash the membrane for 5 minutes to remove unbound primary antibody. Drain the TBS-T when the wash is complete.
4. Repeat step 3 twice, for a total of three washes.
5. Incubate the membrane with 15 ml of secondary antibody solution for 1 hour with gentle rocking at room temperature. The secondary antibody, which is conjuated to horseradish peroxidase (HRP), has been diluted in blocking solution.
6. Carefully drain the antibody from the blot into the test tube marked "Used secondary antibody."
7. Wash the membrane 3 times for 5 minutes each with TBS-T, as in step 3.
8. Drain the TBS-T from the blot. Using a P1000 micropipette, cover the blot with 1 mL of 3,3'5,5'-tetramethyl benzidine (TMB), a colorigenic substrate for HRP. Let the color continue to develop until distinct bands are apparent. Bands will probably become apparent within minutes. Do not allow the blot to over-develop, when nonspecific bands become apparent.
9. Stop color development by diluting the substrate with an excess of deionized water. Drain the diluted substrate into the waste container.
10. Allow the blot to dry on a piece of filter paper. Record your data with the scanner.

References

Harris, L.J., Larson, S.B., Hasel, K.W., and McPherson, A. 1997. Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* **36**: 1581-1597.

Kohler, G., and Milstein, C. 1975. Continuous cultures of fused cells secreting antibody of predefinied specificity. *Nature* **256**: 495-497.

Sleigh, M.J., Both, G.W., Underwood, P.A., and Bender, V.J. 1981. Antigenic drift in the hemagglutinin of the Hong Kong influenza subtype: Correlation of amino acid changes with alterations in viral antigenicity. *J. Virol.* **37**: 845-853.

Southern, J.A., Young, D.F., Heaney, F., Baumgartner, W.K., and Randall, R.E. 1971. Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *J. Gen. Virol.* **72**:1551-1557.

25 Bioinformatics Tools

Objectives

- Learn the basic principles of the BLAST algorithms
- Understand how BLASTP parameters are adjusted
- Learn how the BLASTP algorithm is used to study protein evolution
- Understand the statistical basis for the BLOSUM62 matrix

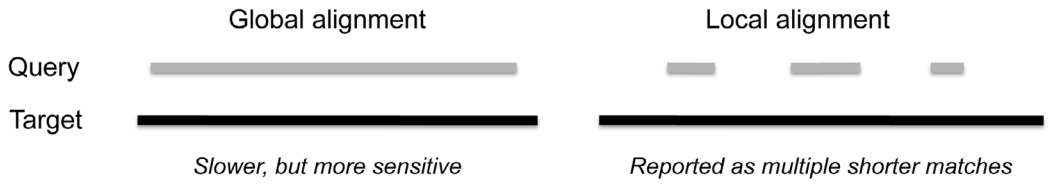
With the rapid advances in DNA sequencing technologies, data pours into the network of international databases. Bioninformatics is a growing research area in biology that develops tools for searching the databases. In this chapter, you will learn about the BLAST algorithms used to search protein and nucleotide sequence databases.

Background

Major advances in DNA sequencing technologies over the past few decades have generated tremendous amounts of sequence information from thousands of different organisms. Fortunately, advances in bioinformatics have occurred at the same time, giving scientists tools to search and analyze these sequences. Throughout this course, we will be interested in comparing gene and protein sequences to the many records in the databases. Databases searches begin with a query sequence. Online bioinformatics tools then take that query sequence and align the sequence against the huge number of records in databases, retrieving records with good matches. In this chapter, we will discuss some of the algorithms that are used in these searches.

BLAST algorithms

There are many different algorithms for searching databases, but BLAST algorithms are some of the most commonly used. BLAST is an acronym for Basic Local Alignment Search Tool, which aptly describes its function. The original BLAST algorithms were developed and published by a group of investigators in 1990 (Altschul et al., 1990). Since that time, investigators have refined and diversified the original BLAST programs so that a whole family of BLAST programs is currently available. BLAST programs are widely used because they are able to quickly search through large amounts of sequence data. The key to BLAST's speed is its use of local alignments that serve as seeds for more extensive alignments. Other algorithms perform global sequence alignments that match the entire sequence of a query sequence against database information. Global sequence alignments are more sensitive than BLAST in finding related sequences (see the figure below), but they require considerably more processing time. BLAST searches are quite adequate for our purposes, and we will use the online BLAST tools available at the NCBI site throughout this course.

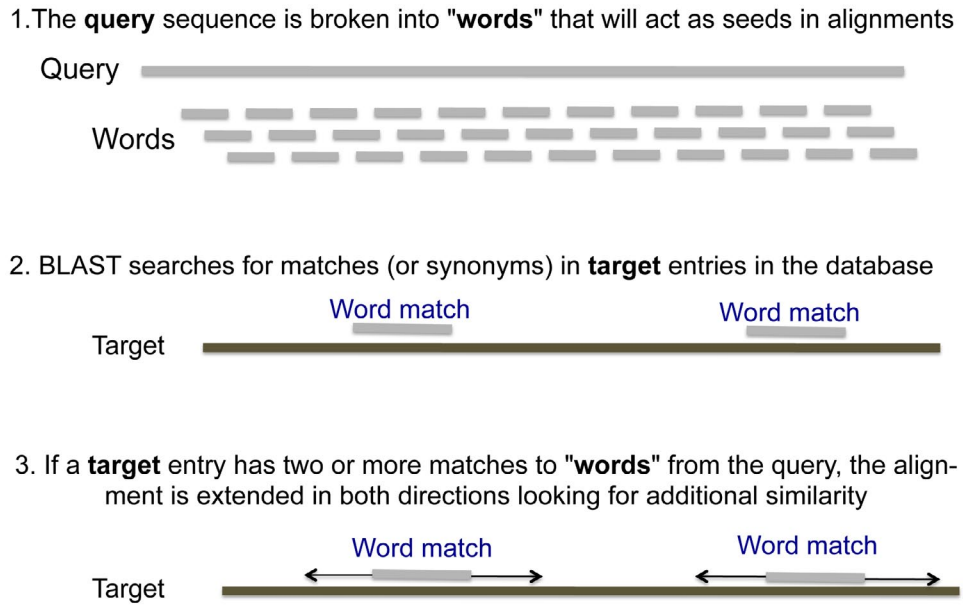


Global vs. local sequence alignments
Global sequence alignments can detect matches that would be missed by local alignment algorithms, because they use the entire sequence as the query.

Either nucleotide or protein sequences can be used as the query sequence for BLAST searches. The BLASTN and BLASTP algorithms, which are designed for nucleotide and protein queries respectively, use different scoring matrices and search parameters. Both BLAST algorithms begin by breaking down query sequence into short "words" and assigning numerical values to the words. A scoring matrix is used to assign numerical values to the words. Words and synonyms above a threshold value are then used to search databases. The default word size for BLASTN is 28 nucleotides, while the default word size for BLASTP is 3 amino acids, because of the stronger evolutionary constraints on protein sequences than nucleotide sequences.

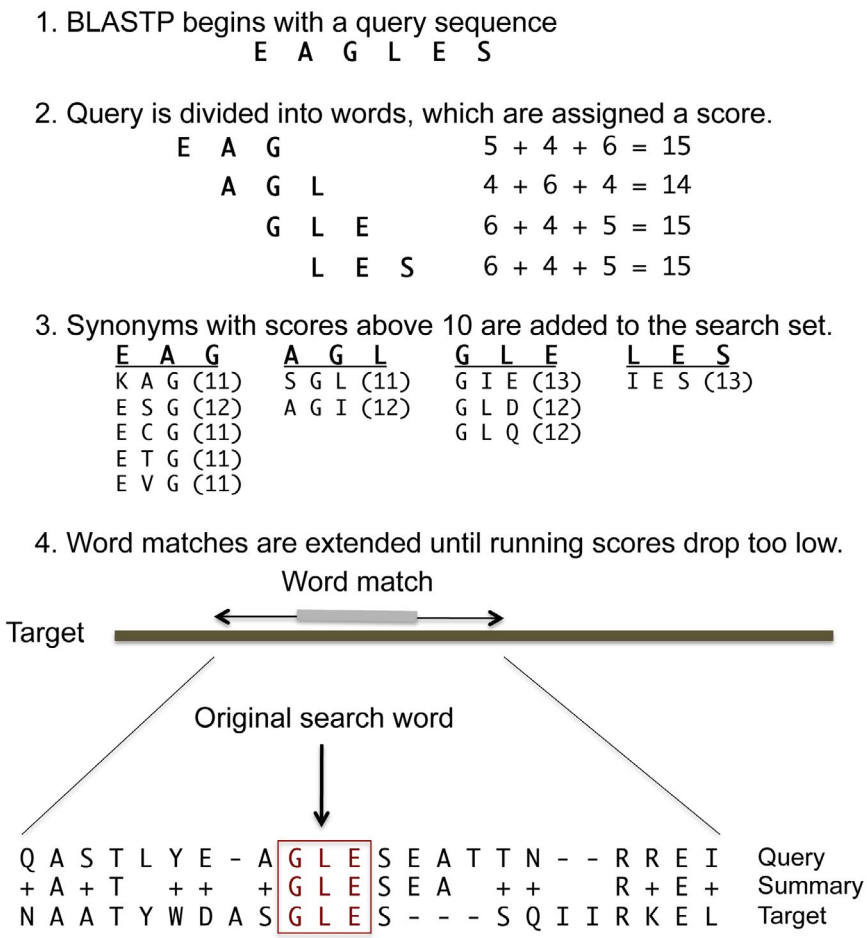
Why do you think that this ?

An outline of the BLAST process is shown below:



Overview of the strategy used in BLAST algorithms
BLASTN and BLASTP use a rolling window to break down a query sequence into words and word synonyms that form a search set. At least two words or synonyms in the search set must match a target sequence in the database, for that sequence to be reported in the results.

We will primarily be using BLASTP in this course, so let's take a closer look at how BLASTP works. In BLASTP, a sequence in broken into all possible 3-letter words that are placed into a search set. A numerical score for each word is assigned by adding the up values for amino acids from a scoring matrix. The user has an opportunity to choose from a variety of scoring matrices, depending on the extent of homology that the investigator considers significant. In the BLASTP example shown on the next page, we have used the BLOSUM62 matrix, which is the default matrix used by the BLASTP tool at NCBI. (The BLOSUM62 matrix is described in greater detail in the following section.) With BLOSUM62, the minimum score for any word is 12. BLASTP next scans through potential synonyms that differ from the words at one position, and those synonyms that exceed a defined threshold value are included in the search set as well. The threshold value can be set by the user. NCBI BLASTP uses a default threshold of 10 for synonyms. Using this search set, BLAST rapidly scans a database and identifies sequences in that database that match at least two words or synonyms in the search set. These sequences are set aside for the next part of the BLASTP process, when these short matches serve as seeds for longer alignments extending in both directions from the original match. BLAST keeps a running raw score as it extends the matches. Each new amino acid either increases or decreases the raw score. Penalties are assigned for mismatches and for gaps between the two alignments. Again, users can adjust the gap penalties, which are usually very severe. In the NCBI default settings, the existence of a gap brings a penalty of 11, which increases by 1 for each missing amino acid. Once the score falls below a set level, the alignment ceases. Raw scores are then converted into bit scores by correcting for the scoring matrix used in the search and the size of the database search space.



Overview of the BLASTP process.
The query sequence EAGLES into broken into three-letter words or synonyms that are used as a search set against records in a protein or translated nucleotide database. See the text for additional details.

The output data from BLASTP includes a table with the bit scores for each alignment as well as its E-value, or "expect score". The E-value indicates the number of alignments with the particular bit score that would be expected to occur solely by chance in the search space. Alignments with the highest bit scores (and lowest E-values) are listed at the top of the table. For perfect or nearly perfect matches, the E-value is reported as zero - there is essentially no possibility that the match occurs randomly. The E-value takes into account both the length of the match and the size of the database that was surveyed. The longer the alignment, and/or the larger the database search space, the less likely that a particular alignment occurs strictly by chance. An E-value of 1 does not necessarily imply, however, that the match has no biological relevance. More nuanced analysis may simply be needed to evaluate the match.

BLOSUM62 scoring matrix

The results obtained in a BLASTP search depend on the scoring matrix used to assign numerical values to different words. A variety of matrices are available, whose utility depends on whether the user is comparing more highly divergent or less divergent sequences. The BLOSUM62 matrix is used as the default scoring matrix for BLASTP. The BLOSUM62 matrix was developed from statistical analyses of the existing protein sequences in databases (Henikoff and Henikoff, 1992). Investigators determined the frequencies of all 210 possible amino acid substitutions in conserved blocks within protein families. The sequences chosen for the BLOSUM62 matrix were more than 62% identical. Thus, the matrix is based on ***observed*** amino acid changes in homologous proteins. The BLOSUM62 score for a particular substitution is a log-odds score that provides a measure of the biological probability of a substitution relative to the chance probability of the substitution. For a substitution of amino acid *i* for amino acid *j*, the score is expressed:

$$S_{ij} = (1/\lambda)\ log\left(\frac{p_{ij}}{q_i q_j}\right)$$

where p_{ij} is the observed frequency of the substitution in homologous proteins, and q_i and q_j are the frequencies of amino acids i and j in the database. The term (1/λ) is a scaling factor used to generate integral values in the matrix. A score of zero indicates that the observed frequency of the substitution is equal to the probability that the substitution occurs by chance.

The BLOSUM62 matrix is shown below. The matrix is consistent with strong evolutionary pressure to conserve protein function. As expected, the most common substitution for any amino acid is itself. Overall, positive scores are less common than negative scores, suggesting that most substitutions have a negative effect on protein function. The most highly conserved amino acids are cysteine, tryptophan and histidine, which have the highest scores. Interestingly, these latter amino acids have unique chemistries and often play important structural or catalytic roles in proteins.

BLOSUM 62 scoring matrix

Ala	**A**	4																		
Arg	**R**	-1	5																	
Asn	**N**	-2	0	6																
Asp	**D**	-2	-2	1	6															
Cys	**C**	0	-3	-3	-3	9														
Gln	**Q**	-1	1	0	0	-3	5													
Glu	**E**	-1	0	0	2	-4	2	5												
Gly	**G**	0	-2	0	-1	-3	-2	-2	6											
His	**H**	-2	0	1	-1	-3	0	0	-2	8										
Ile	**I**	-1	-3	-3	-3	-1	-3	-3	-4	-3	4									
Leu	**L**	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4								
Lys	**K**	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5							
Met	**M**	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5						
Phe	**F**	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6					
Pro	**P**	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7				
Ser	**S**	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4			
Thr	**T**	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5		
Trp	**W**	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	
Tyr	**Y**	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7
Val	**V**	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4
		A	**R**	**N**	**D**	**C**	**Q**	**E**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**

Henikoff, S. and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**:10915-10919.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool *J. Mol. Biol.* **214**: 403-410. PMID 2231712

Henikoff, S., and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**: 10915-10919.