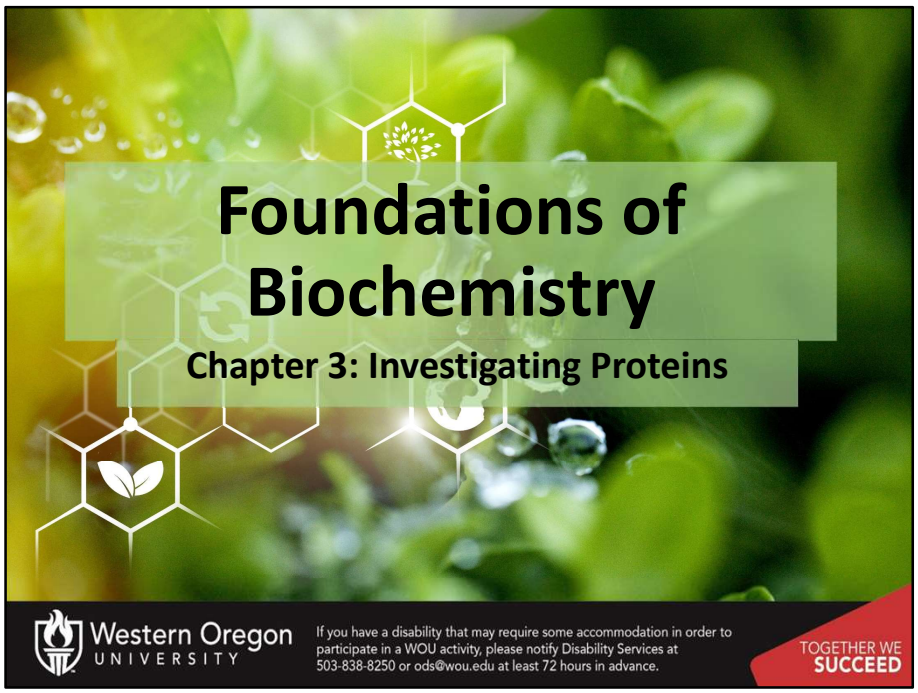
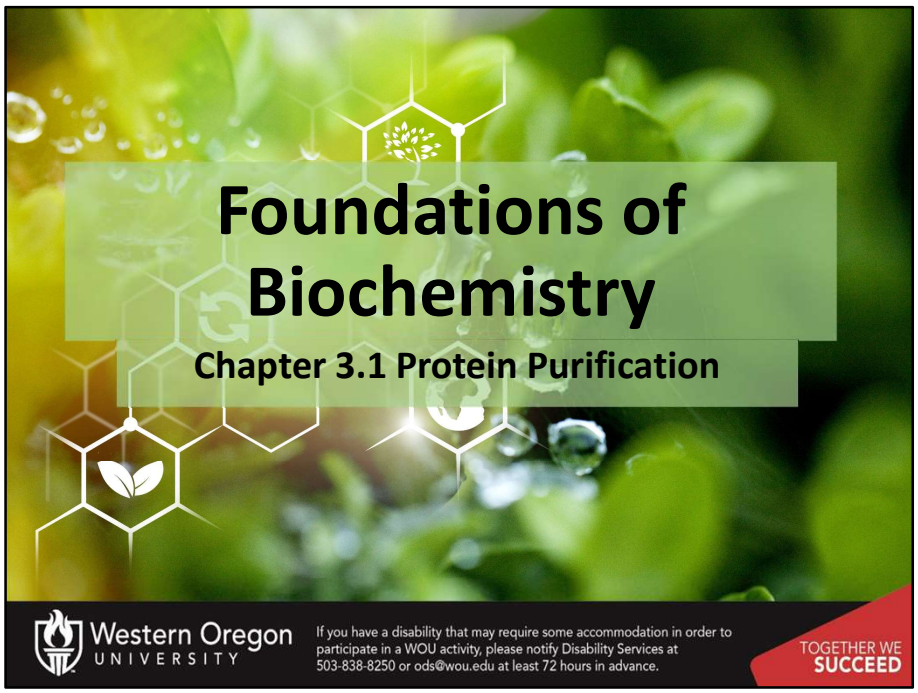


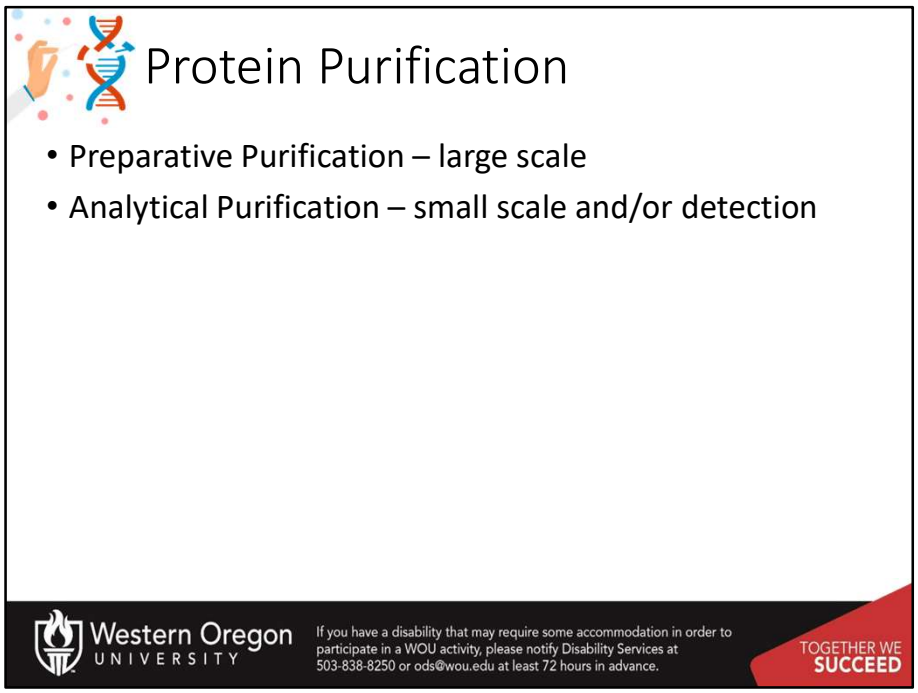
Welcome to our lecture on Investing Proteins. Within this series of lectures, we will discuss common techniques used by biochemists to study protein characteristics, structure, and function.

Key Topics

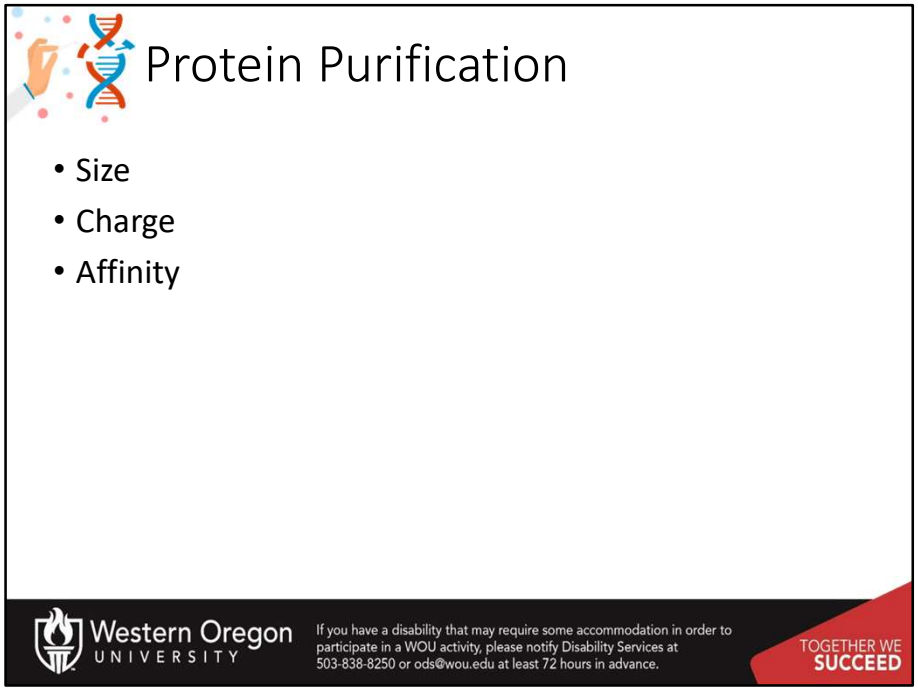
- 3.1 Protein Purification
- 3.2 Protein Identification and Visualization
- 3.3 Protein Synthesis and Sequencing
- 3.4 Protein Structure Elucidation
- 3.5 Proteome Analysis



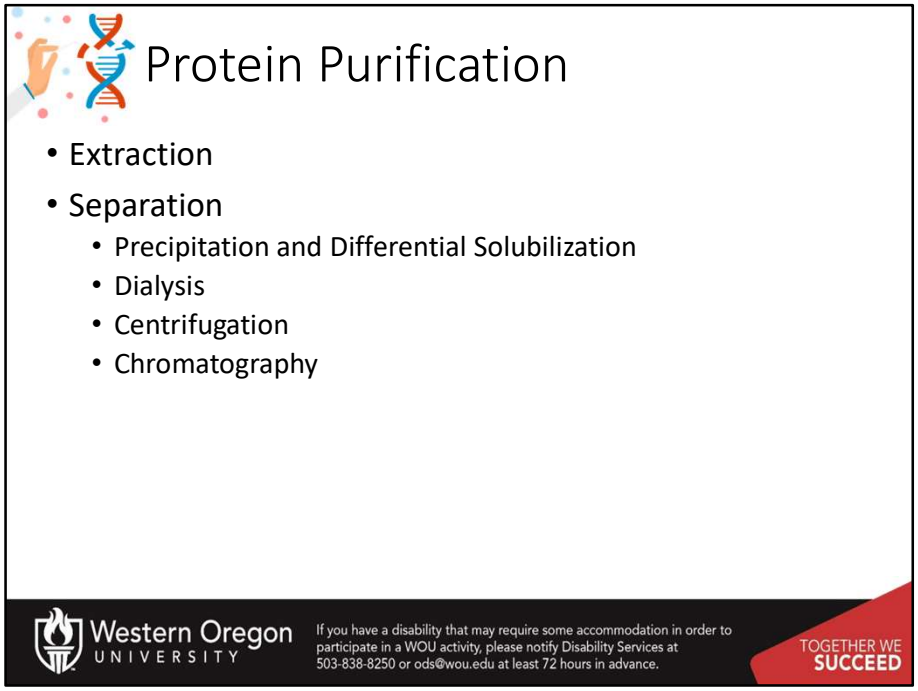
To begin with, we will first look at how proteins are purified from cellular mixtures, so that they can be individually studied.



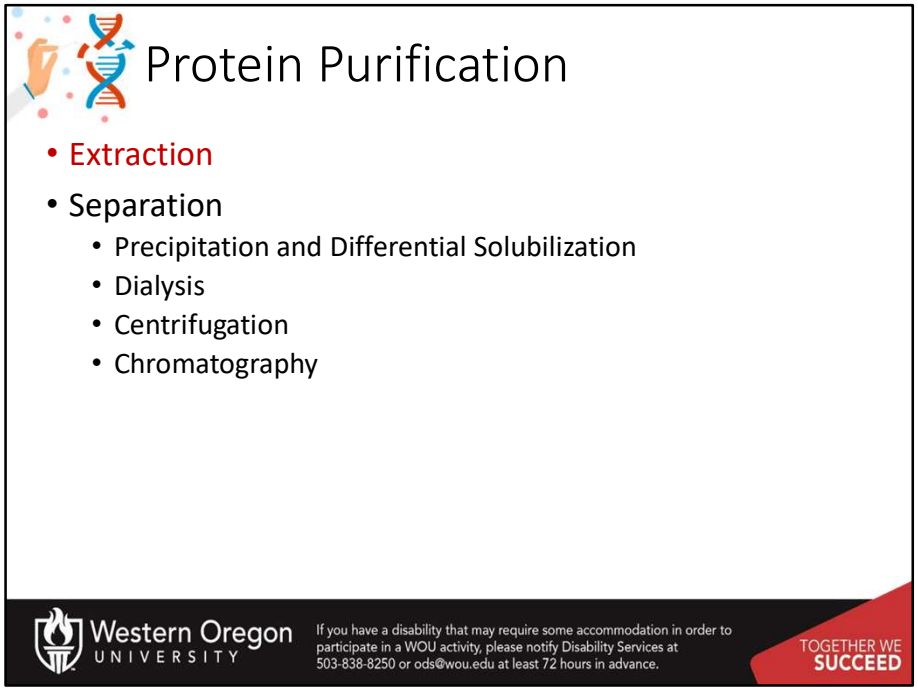
The methods of protein purification are chosen and used depending on the need of the purification. Ie whether you need a lot of the protein of interest or only a small amount. **Preparative purifications** aim to produce a relatively large quantity of purified proteins for subsequent use. Examples include the preparation of commercial products such as enzymes (e.g. lactase), nutritional proteins (e.g. soy protein isolate), and certain biopharmaceuticals (e.g. insulin). High amounts of protein are also often required for structural analysis of the protein. **Analytical purification** produces a relatively small amount of a protein for a variety of research or analytical purposes, including identification, quantification, and studies of the protein's post-translational modifications and function. Sometimes analytical purification can also be used for protein structural analysis, depending on the technique used.



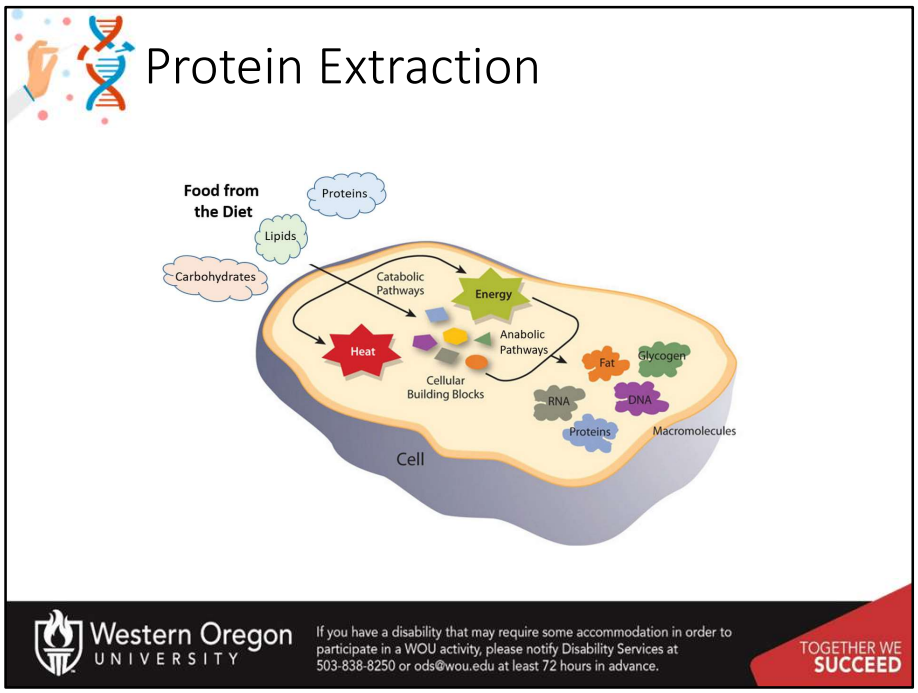
Any protein purification scheme needs to use techniques of separation. How do you purify proteins from complex mixtures? First, let's think about what protein features and characteristics can be used, that will allow differential separation. Proteins differ in the fundamental areas of size, charge, and affinity for other molecules. Thus, these characteristics are exploited when developing purification schemes.



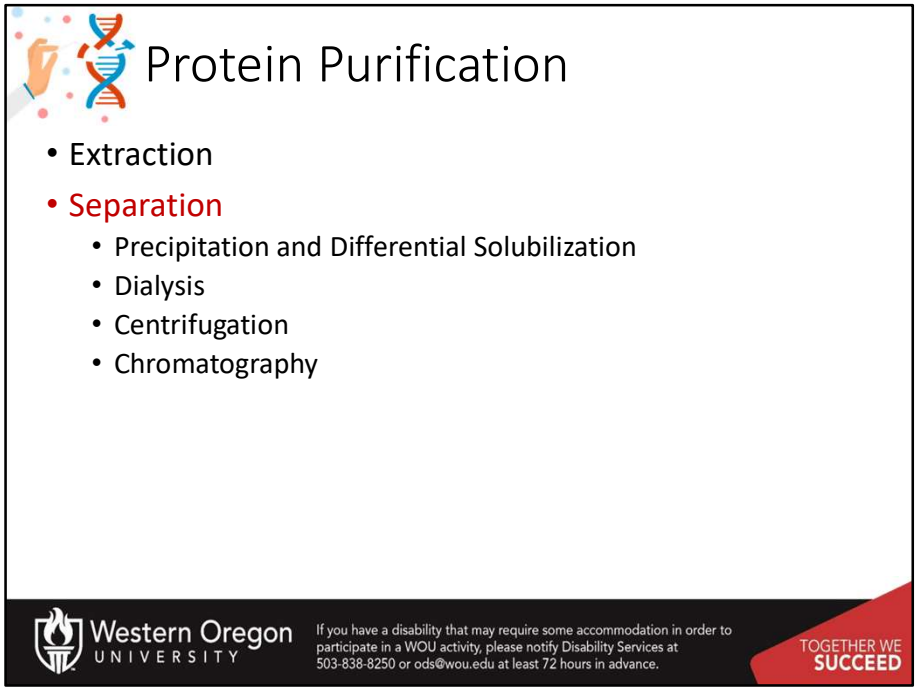
Exploitation of these characteristics of protein differences has given rise to four main protein separation techniques for the large scale preparation of proteins. These are precipitation and differential solubilization, dialysis, centrifugation, and chromatography. With chromatography being the most varied of these techniques. We will take a closer look at all of these techniques.



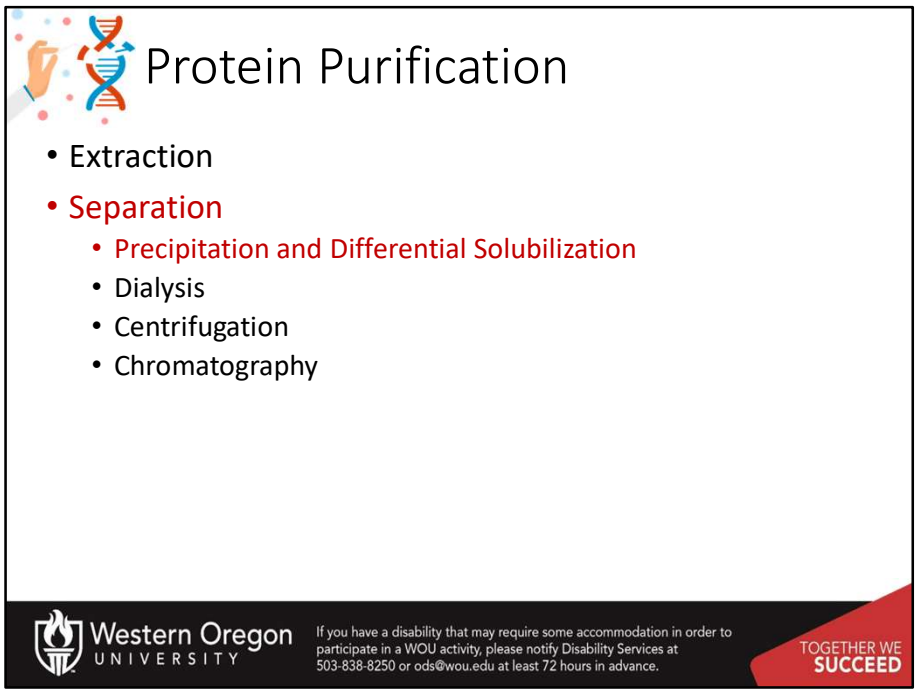
But first they all must begin with extraction, as proteins are housed within cells, that make up more complex tissues.



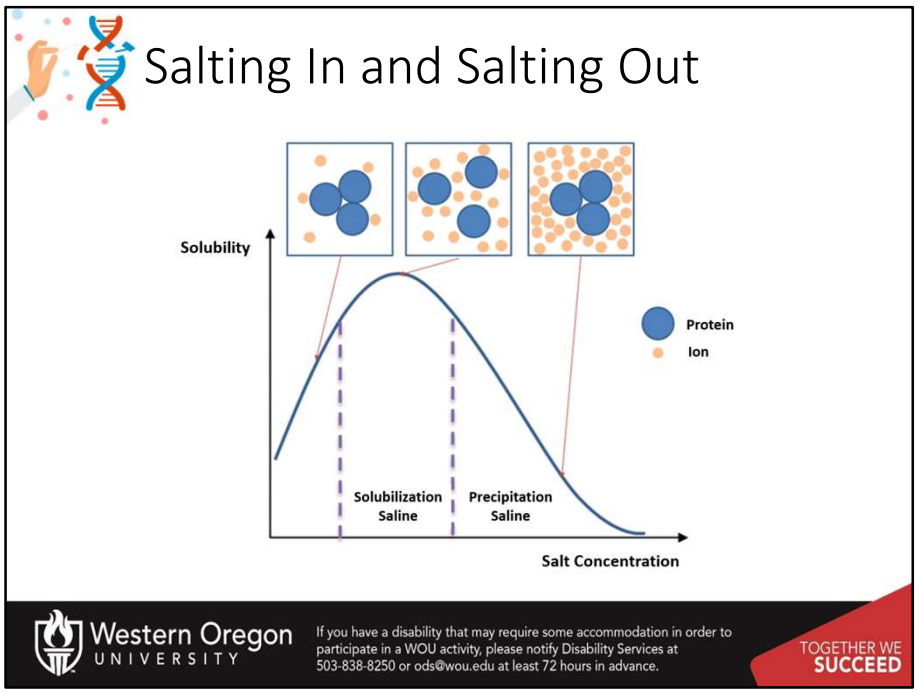
As we saw in Chapter 1, cells are complex places made up of all four major macromolecules (lipids, carbohydrates, proteins and nucleic acids. To purify proteins, the first step must be to isolate the protein fraction of the cellular components away from all of the other components. This can be done by first breaking the cell open. This can be done by freeze-thawing cycles, sonication (or the use of sound waves to shatter the plasma membrane and/or cell wall structure), use of high pressure, grinding, or detergents. During this process, proteases that are normally contained in specialized compartments within the cells will be released. Thus, it is also necessary to inactivate proteases within the extract, and to keep the mixture cold to preserve the shape and biological activity of the proteins.



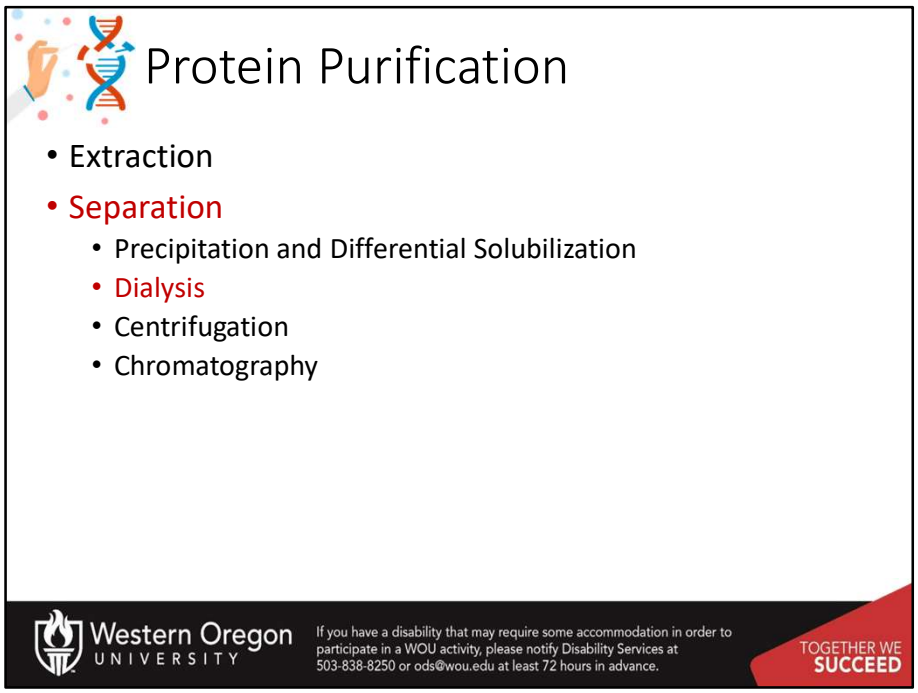
Now we can move on to discuss large scale protein separation techniques.



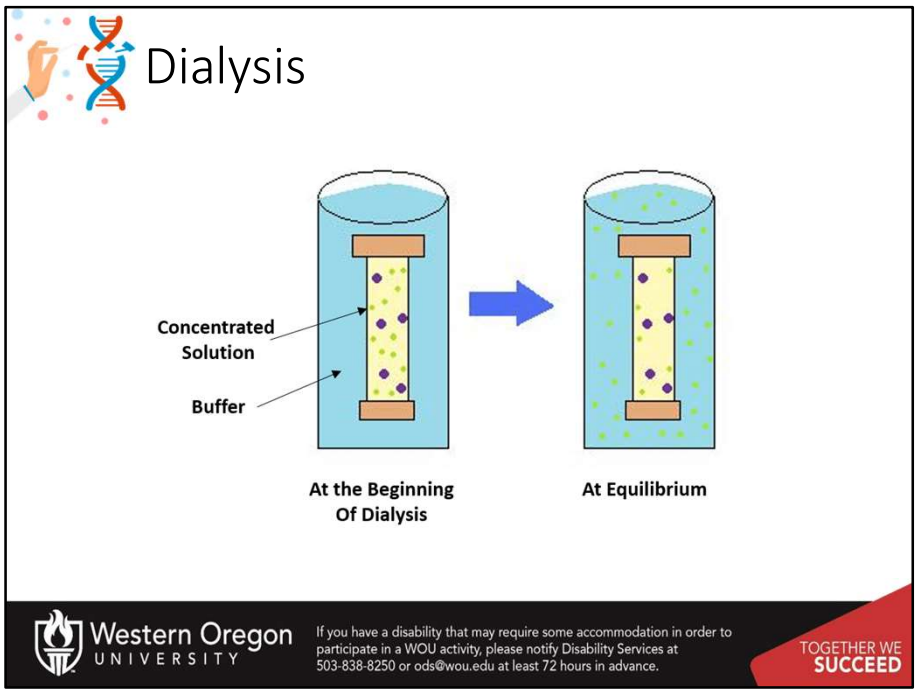
The first one involves protein precipitation and differential solubilization.



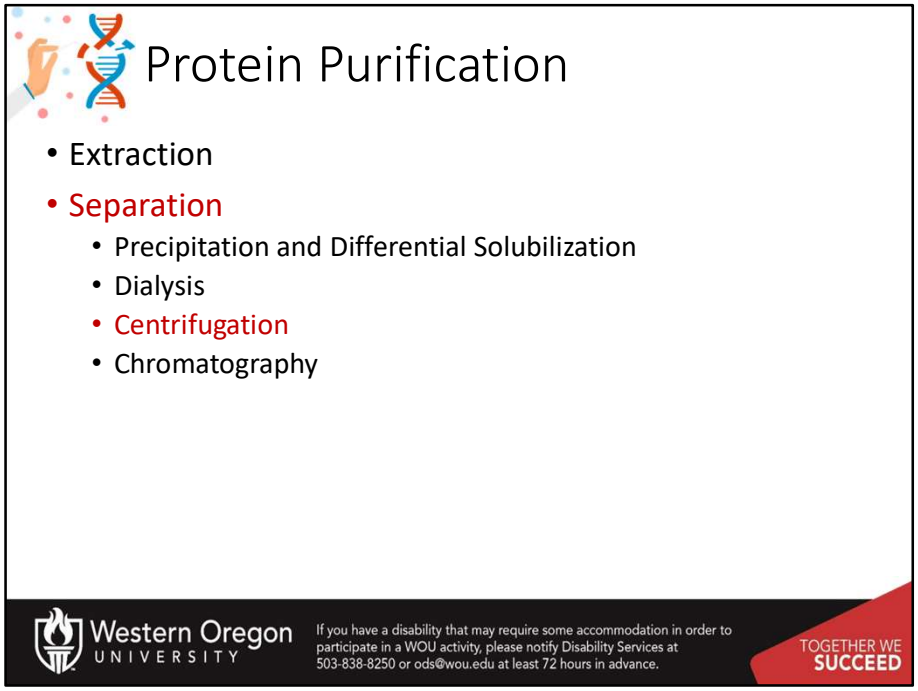
In bulk protein purification, a common first step to isolate proteins is precipitation using a salt such as ammonium sulfate $(NH_4)_2SO_4$. This process is called ***Salting In*** or ***Salting Out.*** This is performed by adding increasing amounts of ammonium sulfate and collecting the different fractions of precipitate protein. Ammonium sulfate is often used as it is highly soluble in water, has relative freedom from temperature effects and typically is not harmful to most proteins.



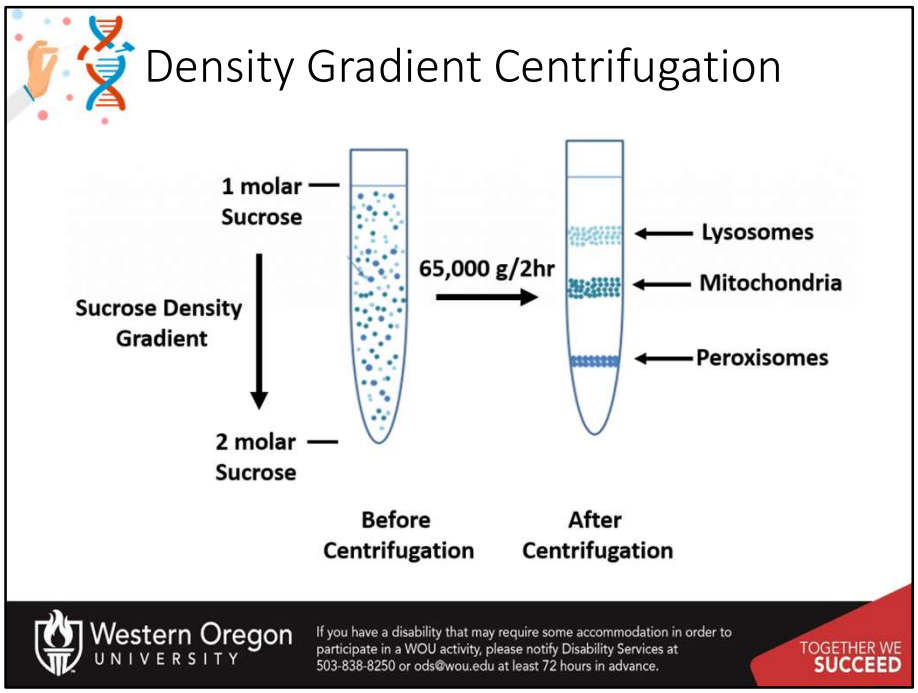
Once proteins have been partially purified by salting in and out, the salt must be removed from the solution to help avoid altering the biological function of the protein by disrupting normal intermolecular interactions. This can be done by dialysis



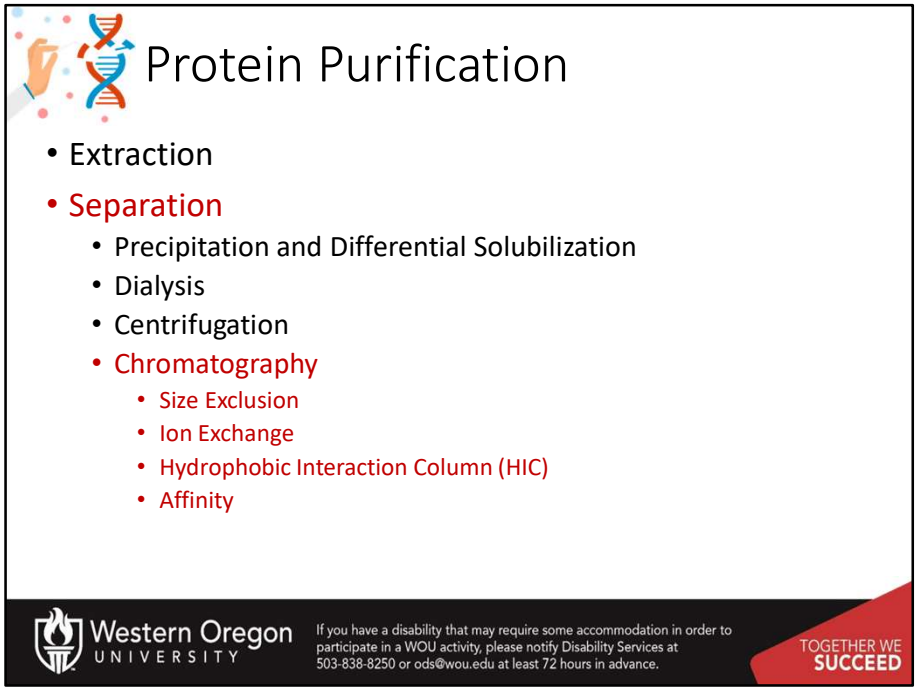
The process of dialysis separates dissolved molecules by their size. The biological sample is placed inside a closed membrane, where the protein of interest is too large to pass through the pores of the membrane, but through which smaller ions can easily pass. As the solution comes to equilibrium, the ions become evenly distributed throughout the entire solution, while the protein remains concentrated in the membrane. This reduces the overall salt concentration of the suspension.



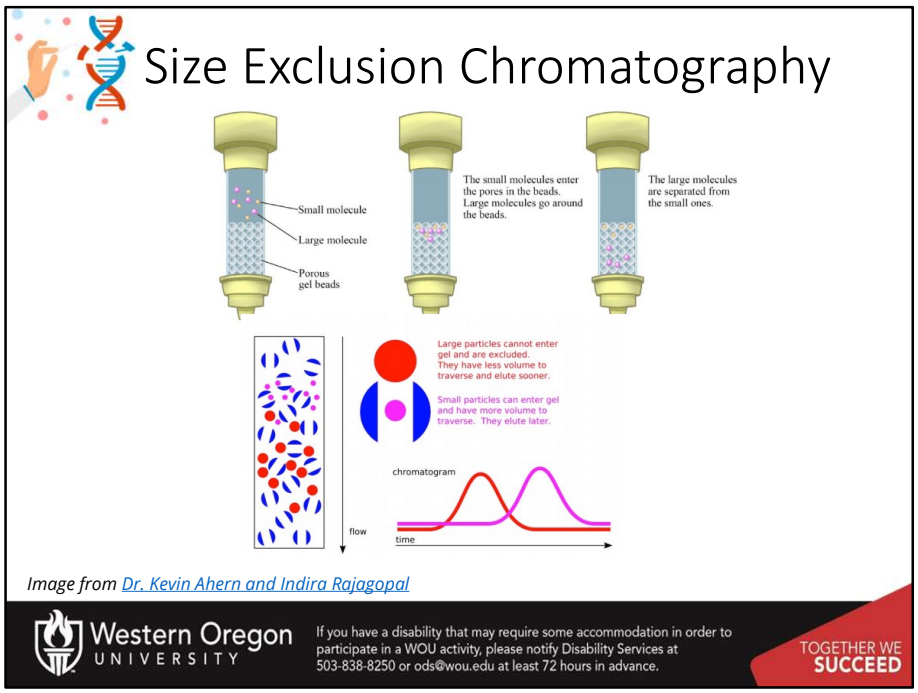
Centrifugation is a process that uses centrifugal force to separate mixtures of particles of varying masses or densities suspended in a liquid. When a vessel (typically a tube or bottle) containing a mixture of proteins or other particulate matter, such as bacterial cells, is rotated at high speeds, the inertia of each particle yields a force in the direction of the particles velocity that is proportional to its mass. The tendency of a given particle to move through the liquid because of this force is offset by the resistance the liquid exerts on the particle. The net effect of "spinning" the sample in a centrifuge is that massive, small, and dense particles move outward faster than less massive particles or particles with more "drag" in the liquid. When suspensions of particles are "spun" in a centrifuge, a "pellet" may form at the bottom of the vessel that is enriched for the most massive particles with low drag in the liquid. Non-compacted particles remain mostly in the liquid called "supernatant" and can be removed from the vessel thereby separating the supernatant from the pellet.



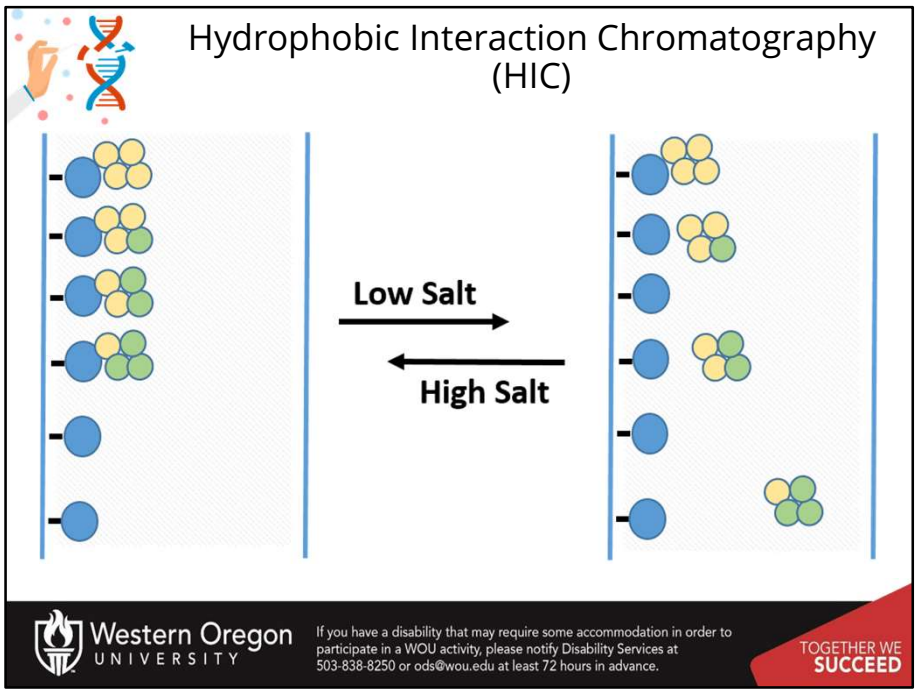
A more specialized centrifugation technique is known as density gradient centrifugation. In this technique a solute that won't damage the cellular components to be studied, is used to create a layered density gradient during the centrifugation. Sucrose and other proprietary solutes, such as Percoll are often used for this purpose. A protein sample is then layered on top of the gradient and spun at high speeds in an ultracentrifuge. This causes heavy macromolecules to migrate towards the bottom of the tube faster than lighter material. A properly designed sucrose gradient will counteract the increasing centrifugal force so the particles move in close proportion to the time they have been in the centrifugal field. Samples separated by these gradients are referred to as "rate zonal" centrifugations. After separating the protein/particles, the gradient is then fractionated and collected. This can be done easily by putting a pin hole in the bottom of the plastic centrifuge tube and collecting fractions as they drip out from the bottom.



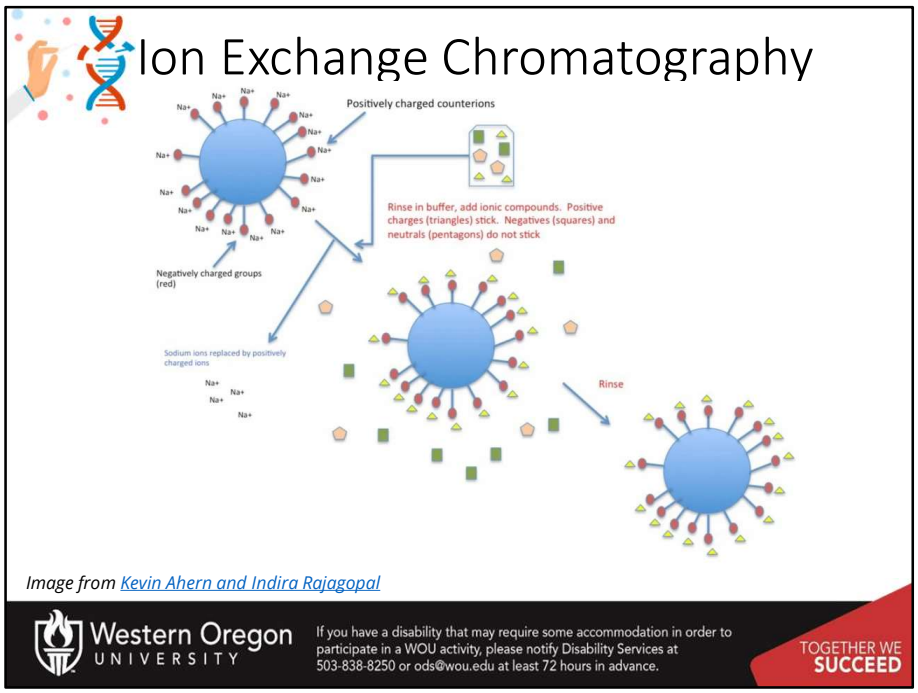
Methods of chromatography are also commonly used to separate complex mixtures of proteins. These methodologies are often employed as a second step in the purification pathway after a salting out or centrifugal step. Column chromatography can be used to separate larger quantities of proteins and can use the properties of charge, size, and affinity depending on the column matrix utilized. There are four main types that you should become familiar with: Size Exclusion, Ion Exchange, Hydrophobic interaction, and Affinity. If your proteins sequence is not known and only the biological activity has been identified, column chromatography is an experimental way to mediate separation. After each step, fractions must be collected and tested, to identify where the protein has eluted. Positive fractions can then be pooled together and analyzed further for purity. Additional purification steps may be required to reach a desired protein purity.



The first type of column matrix that we will discuss, allows separation of proteins based on their size. This is known as size exclusion chromatography (or Gel Filtration Chromatography). This technique involves the use of beads that have tiny "tunnels" in them that each have a precise size. The size is referred to as an "exclusion limit," which means that molecules above a certain molecular weight will not fit into the tunnels. Molecules with sizes larger than the exclusion limit do not enter the tunnels and pass through the column relatively quickly by making their way between the beads. Smaller molecules, which can enter the tunnels, do so, and thus, have a longer path that they take in passing through the column. Because of this, molecules larger than the exclusion limit will leave the column earlier, while smaller molecules that pass through the beads will elute from the column later.

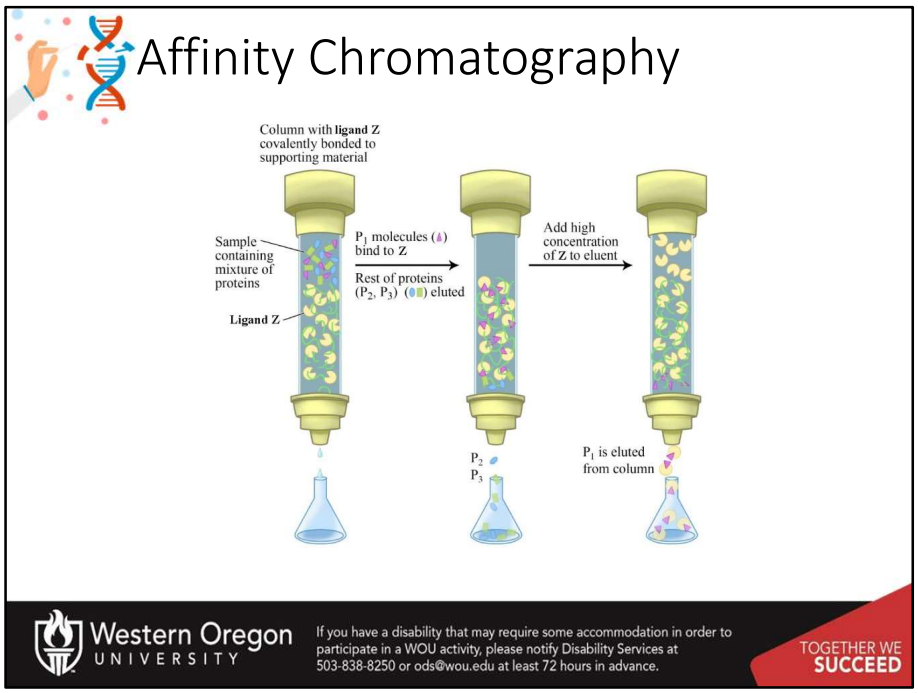


The next system is hydrophobic interaction chromatography or HIC. HIC media is amphipathic, with both hydrophobic and hydrophilic regions, allowing for separation of proteins based on their surface hydrophobicity. The column matrix, shown in blue has a hydrophobic ligand covalently attached. In high salt conditions, proteins will bind to the matrix with differing affinity, with more hydrophobic proteins (shown in yellow) binding more tightly than more hydrophilic proteins (shown in green) When the salt concentration is decreased, proteins that are more hydrophilic will be released first, followed more hydrophobic proteins.

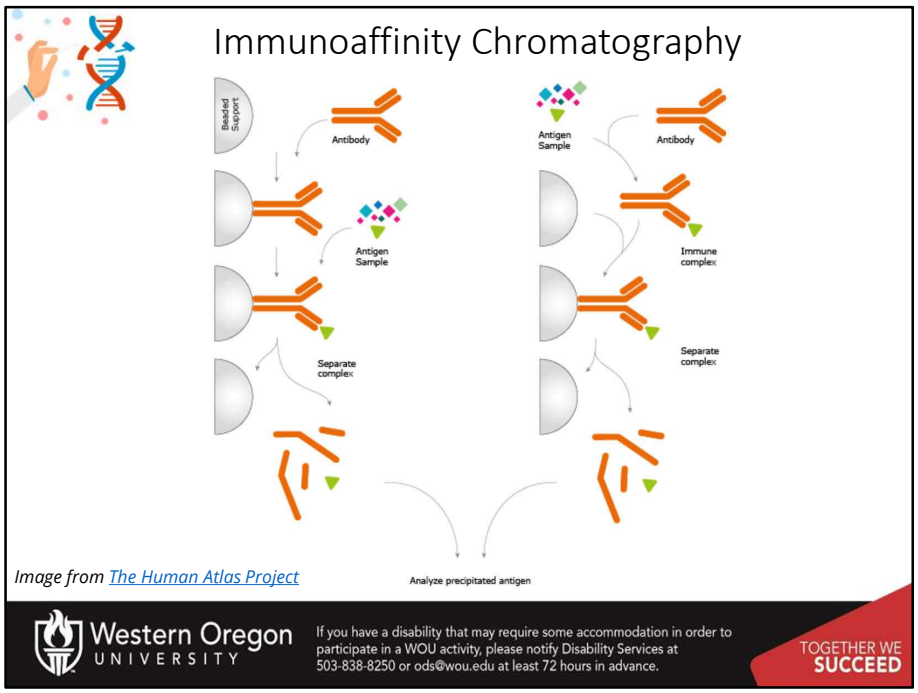


Ion exchange chromatography separates compounds according to the nature and degree of their ionic charge. The column to be used is selected according to its type and strength of charge. ***Anion exchange resins*** have a positive charge and are used to retain and separate negatively charged compounds (anions), while ***cation exchange resins*** have a negative charge and are used to separate positively charged molecules (cations). The diagram above shows a cation exchange column. Within this system, the beads, shown in blue, have negatively charged particles attached to them (shown in red). The column is then washed with a salt solution (such as sodium chloride). The positive sodium ions will stick to the negative beads through an ionic interaction. The protein solution is then applied to the column, and positively charged amino acid residues will displace the sodium ions and cause proteins to stick to the beads. The more positive the protein, the more it will stick to the beads. The proteins essentially 'exchange' places with the positive sodium ions (hence the name ion exchange). Once attached, the column is rinsed to wash away neutral and negatively charged proteins. The positively charged proteins are then eluted by adding higher concentrations of the counter ion (in this case sodium). Anion exchange columns would do the opposite…they would attract negatively charged proteins. If you know the protein charge, you can choose the best ion exchange method. However, if you don't know the charge of your protein of interest all

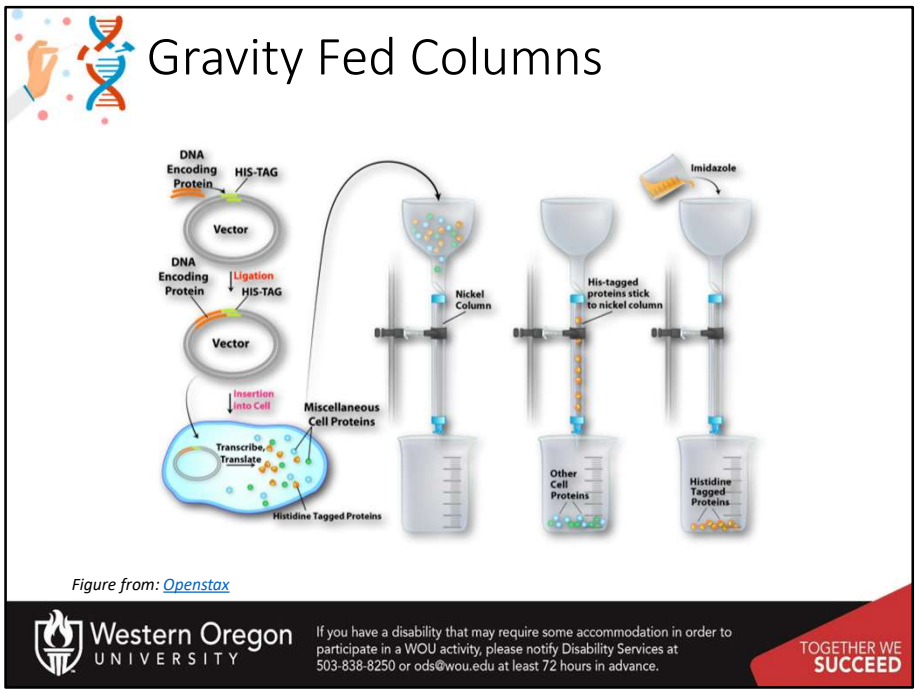
material coming off the column should be collected as fractions and tested. For example, if your protein is neutral, it will come out in the first column wash and won't stick to the column, so it's important to collect everything that comes off the column to find out where your protein elutes.



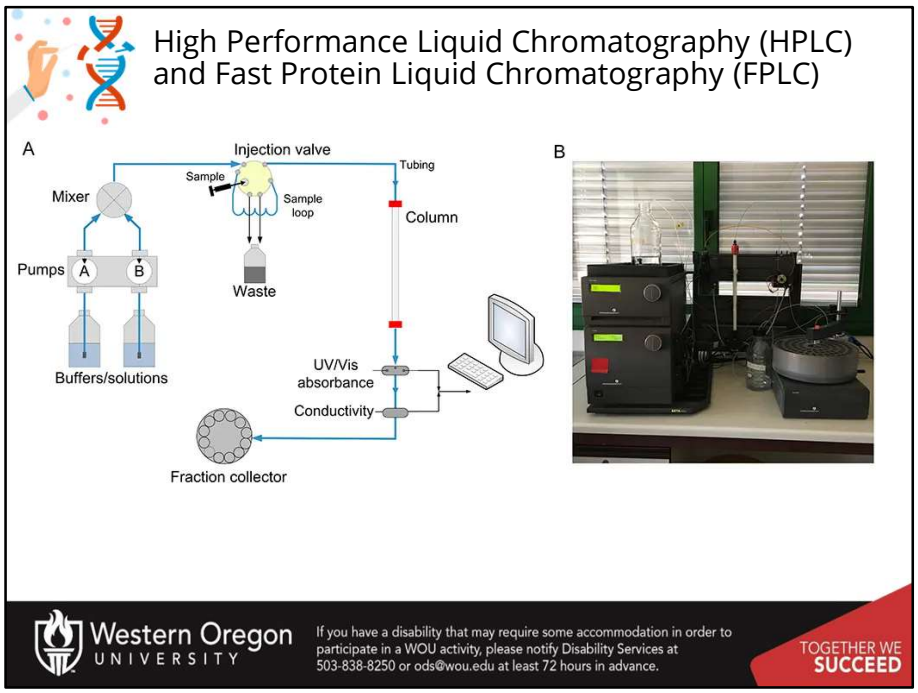
Affinity chromatography is useful if you know that your protein of interest binds with a specific small molecule or ligand. For example, if you know your protein binds with glucose, you could use a column matrix that has been crosslinked with glucose molecules to purify your protein. Your protein of interest would preferentially bind to the glucose, while other proteins flow through the column. The protein would then be eluted from the column by adding free glucose to the elution buffer. The free glucose would compete for the protein binding causing it to elute. Affinity chromatography can yield highly purified protein in a single step. However, some proteins may bind so tightly to their ligand that they cannot be efficiently eluted once they have bound to the affinity column.



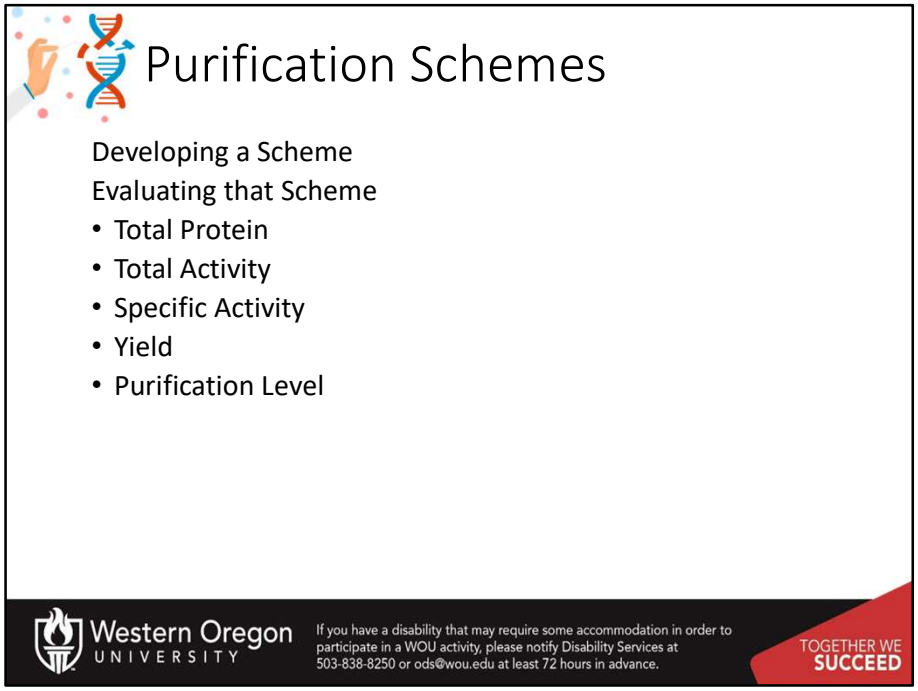
A special type of affinity chromatography is Immunoaffinity chromatography. This technique uses the specific binding of an antibody with its antigen (target molecule that the antibody will bind with selectively) to purify the protein of interest. The procedure involves immobilizing an antibody to a solid substrate (e.g. a porous bead or a membrane), which then selectively binds the target, while everything else flows through. The antibody can be attached to the column beads prior to the interaction with the target protein or after it has bound with the antibody. The column is then washed, and target protein can be eluted by changing the pH or the salinity.



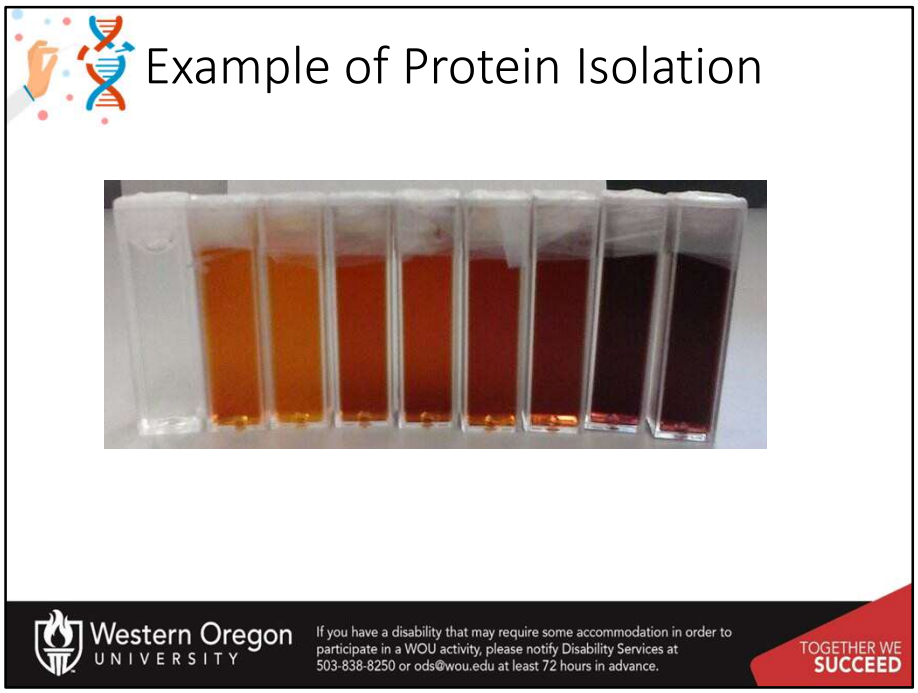
The matrices that we've talked about thus far for Size Exclusion, HIC, Ion Exchange, and Affinity chromatography can be used in different types of columns. The simplest are gravity fed columns where the solutions are loaded onto the column and allowed to flow through where fractions can be collected. The resolution of gravity fed columns is often quite low (ie you will likely get a diffuse band that isn't well separated from other molecules that have similar characteristics). However, large quantities of extract can often be used in this type of column and it may be a great first or second step in a purification scheme.



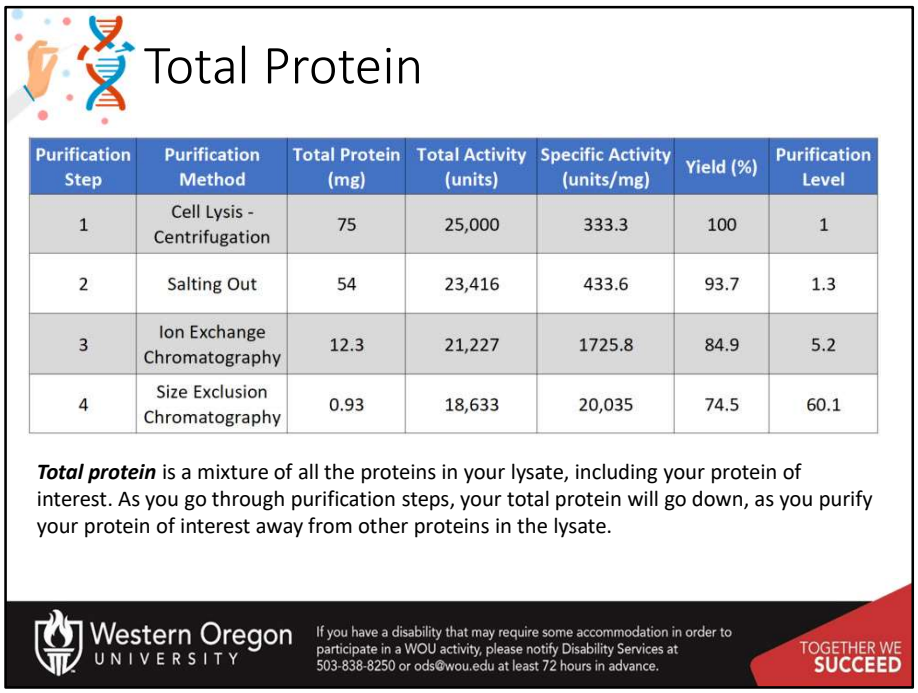
High performance liquid chromatography or high pressure liquid chromatography (HPLC) is a form of chromatography applying high pressure to drive the solutes through the column faster. This means that diffusion is limited and the resolution is improved. The most common form is “reversed phase” HPLC, where the column material is hydrophobic, but other types of resins can be used as well. The high pressure system, however, can often lead to the denaturation of proteins and is a drawback for HPLC for protein purification. To help remedy this ***Fast protein liquid chromatography (FPLC)*** is often used. FPLC systems use low pressure with an aqueous buffer solution. This technique will often increase protein resolution and better separation without causing denaturation. The set up for both HPLC and FPLC is very similar. The system shown here is an FPLC system. The column used is closed, and attached to tubing that is connected with a pump. The pump is used to force liquid under pressure through the column. The protein sample is injected onto the column in a sample loop. A common detection system used UV absorption to monitor protein elution from the column and fractions are collected. The FPLC columns can be glass, as this can withstand low pressure systems. HPLC columns are steel and can withstand much higher pressures.



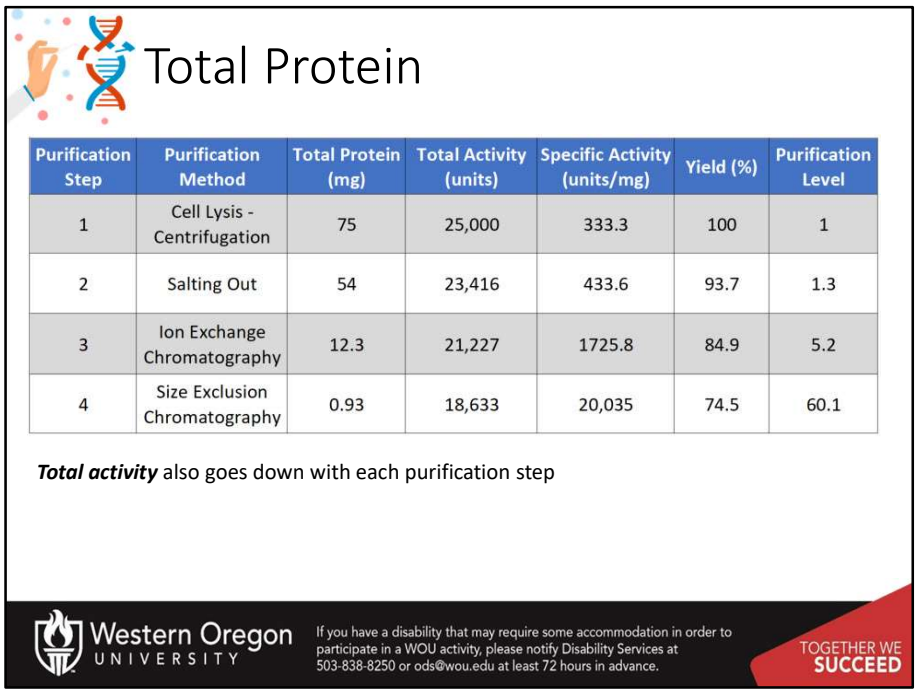
During the purification of an unknown protein, it is necessary to have a quantitative system to determine how much protein has been purified, what concentration the protein represents from the original mixture, how biologically active the purified protein is, and the overall purity of the protein. This will help guide and optimize the purification method being developed. Note that this is a very experimental process and that researchers need to try different types of protein purification strategies to optimize their protocol. A good evaluation scheme will take into account the total protein, total activity, specific activity, yield, and the purification level.



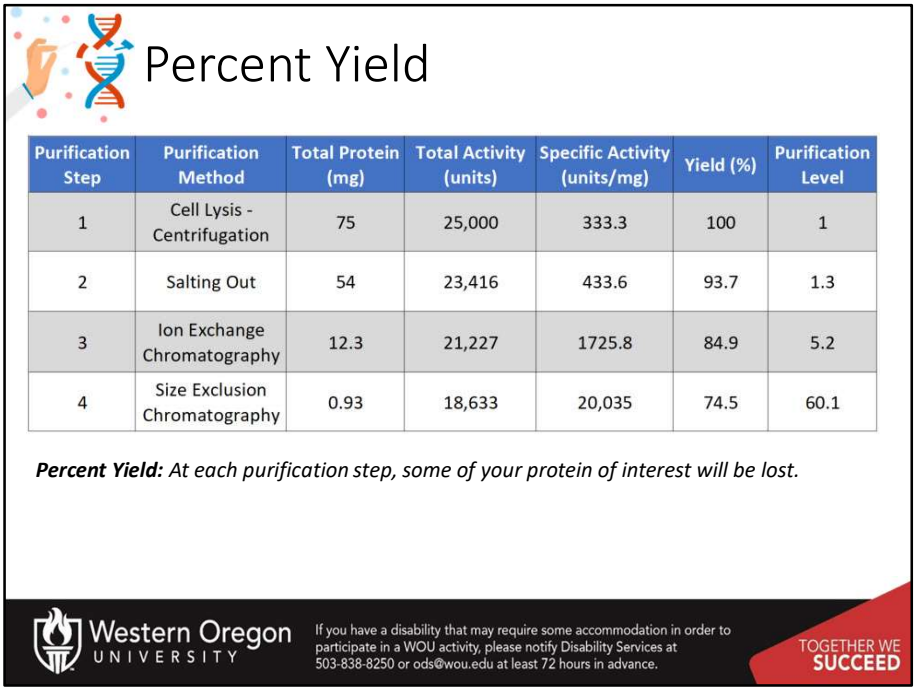
Pretend you are a researcher that wants to isolate a novel, unknown protein from a bacterial culture. You grow 500 ml of the bacteria overnight at 37^{o}C and harvest the bacteria by centrifugation. You remove the culture broth and retain the bacterial pellet. You then lyse the bacteria using freeze/thaw in 10 mL of reaction buffer. You then centrifuge the lysed bacteria to remove the insoluble materials and retain the supernatent that contains the soluble proteins. Your protein of interest has a biological activity that you can measure using a simple assay that causes a color change in the reaction mixture. You also note that this reaction rate increases with increasing concentrations of your protein supernatant. This biological assay can then be used to help you purify the protein responsible for this activity



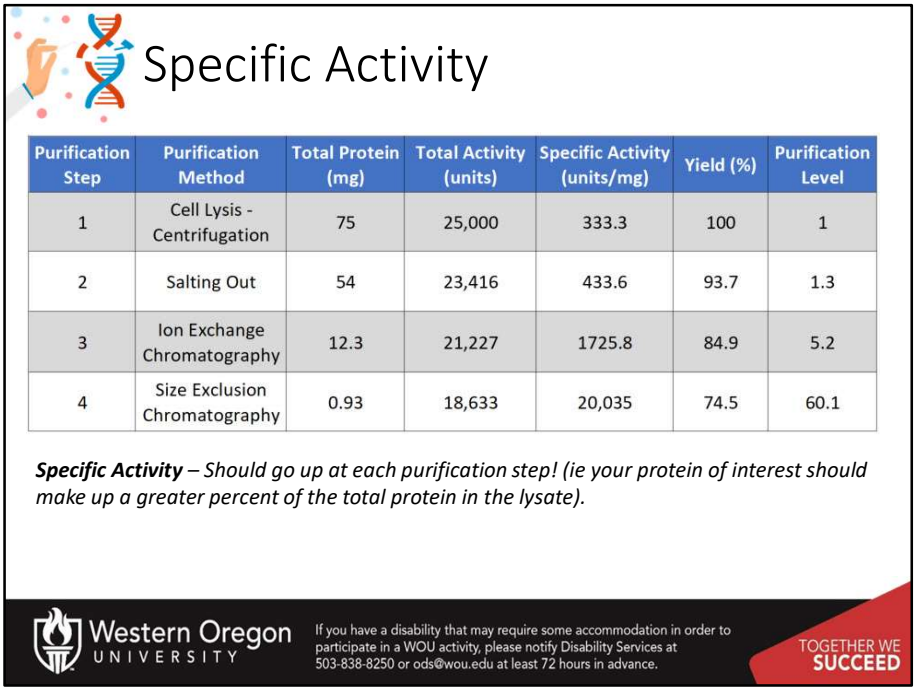
At this point, you can measure your baseline concentrations for the first purification level (bacterial lysis and removal of insoluble proteins and other cellular debris by centrifugation). ***Total Protein*** is calculated by measuring the concentration in a fraction of your sample, and then multiplying that by the total volume of your sample. In this case, you are starting with 10 mL of supernatent. In a typical assay to measure protein concentration, you will use 50 – 200 µL of sample to determine the protein concentration. For example, if you calculate that there is 7.5 µg/µL in your initial assay, you would need to convert that value into mg/mL and then multiply it by 10 mL for a total of 75 mg of protein in 10 mL of supernatant



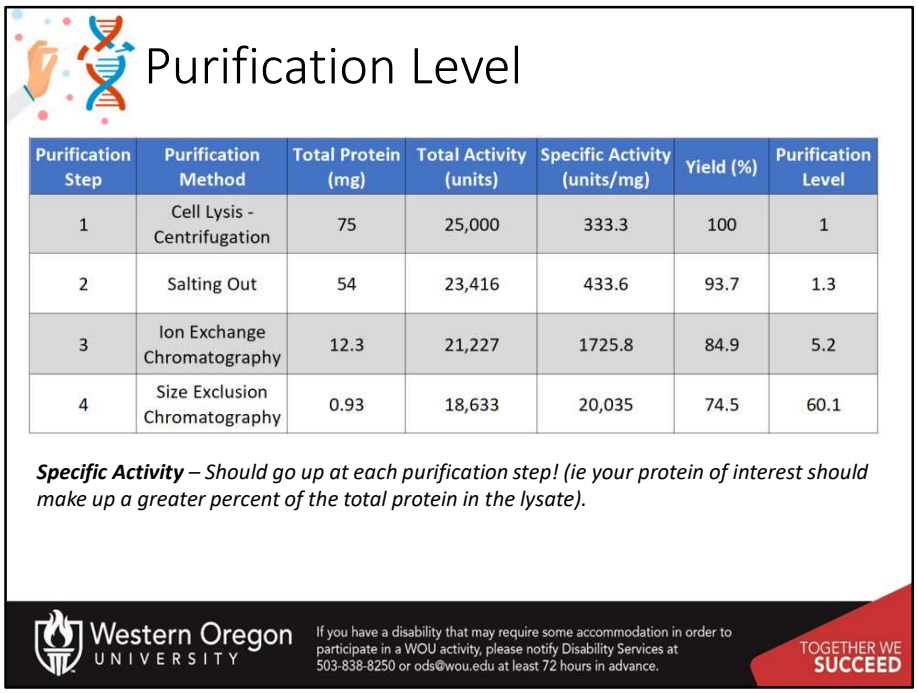
Total Activity also goes down with each purification step, as some of your protein of interest is also lost at each purification step, because (1) some protein will stick to the test tubes and glassware, (2) some protein won't bind with 100% efficiency to your column matrix, (3) some protein may bind too tightly to be removed from the column matrix during elution, and (4) some protein may be denatured or degraded during the purification process. Total activity will be measured in the biological activity assay that is specific for your protein of interest and is usually presented in units. 1 U (µmol/min) is defined as the amount of the **enzyme** that catalyzes the conversion of one micromole of substrate per minute under the specified conditions of the assay method.



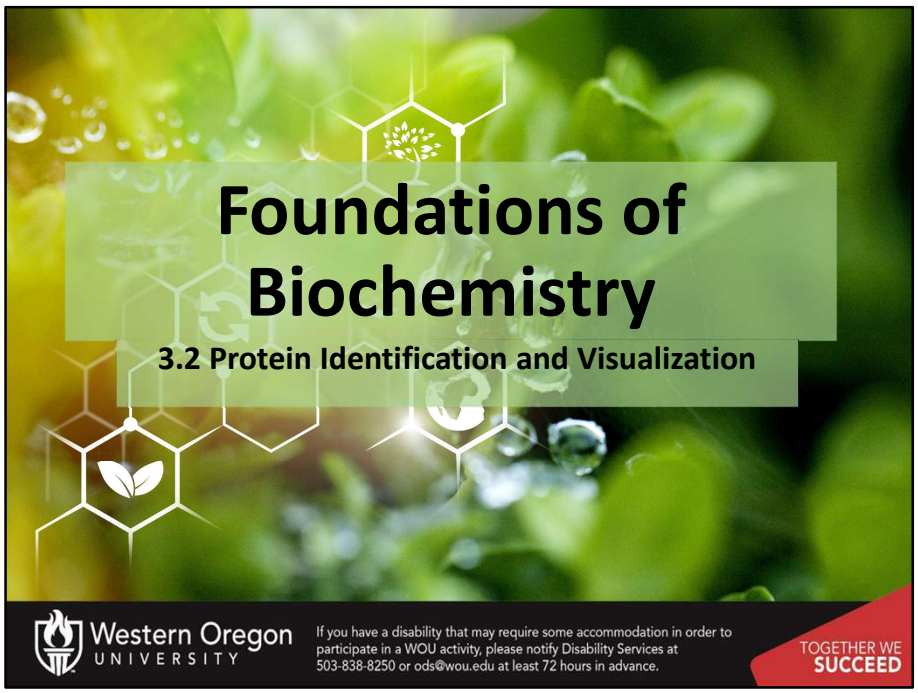
The amount of your protein of interest that is lost is represented within the overall ***percent yield*** for each purification step. The ***percent yield*** is calculated from using the Total Activity values. For example, the percent yield in the first step represent 100% of your protein sample, and in this case has a Total Activity of 25,000 U. The percent yield in Step 2 = 93.7%. This was calculated by dividing the Total Activity in Step 2 (23,416 U) by the Total Activity in Step 1 (25,000 U) multiplied by 100. Each subsequent Total Activity from further purification steps will then be evaluated against the Total Activity in the first step to calculate the ***percent yield.*** If the ***percent yield*** is too low alternative purification methods should be explored.



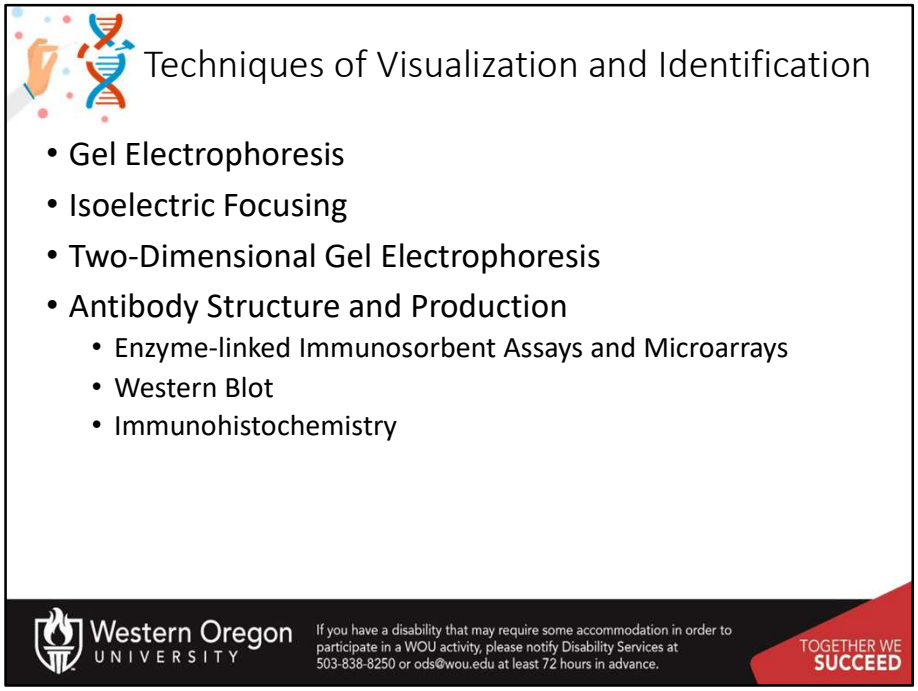
Specific Activity is measured by dividing the Total Activity by the Total Protein.



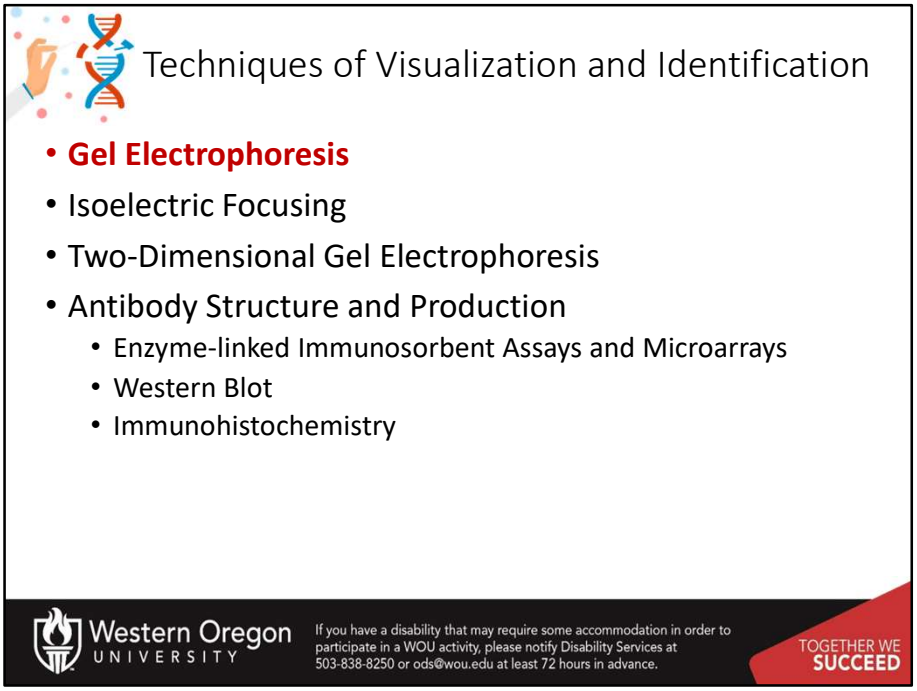
Purification level evaluates the purity of the protein of interest by dividing the specific activity calculated after each purification step by the specific activity of the first purification step. Thus, the first step always has a value of 1. Overall, the fold increase in ***purification level*** should increase exponentially during the purification process. The fold purification that you will wish to achieve will vary depending on what you intend to use your protein for (ie. How pure do you need it? If it will be used in medicine, you might need highly purified samples). It will also vary depending on how concentrated your protein of interest was in the initial sample (ie. If your protein of interest only makes up 0.1% of the lysate, you will need more purification steps to achieve greater purity, that a sample where 10% of your lysate is your protein of interest.) In the next section, we will explore techniques for protein visualization and identification, they are often combined with and used during protein purification processes, as well.



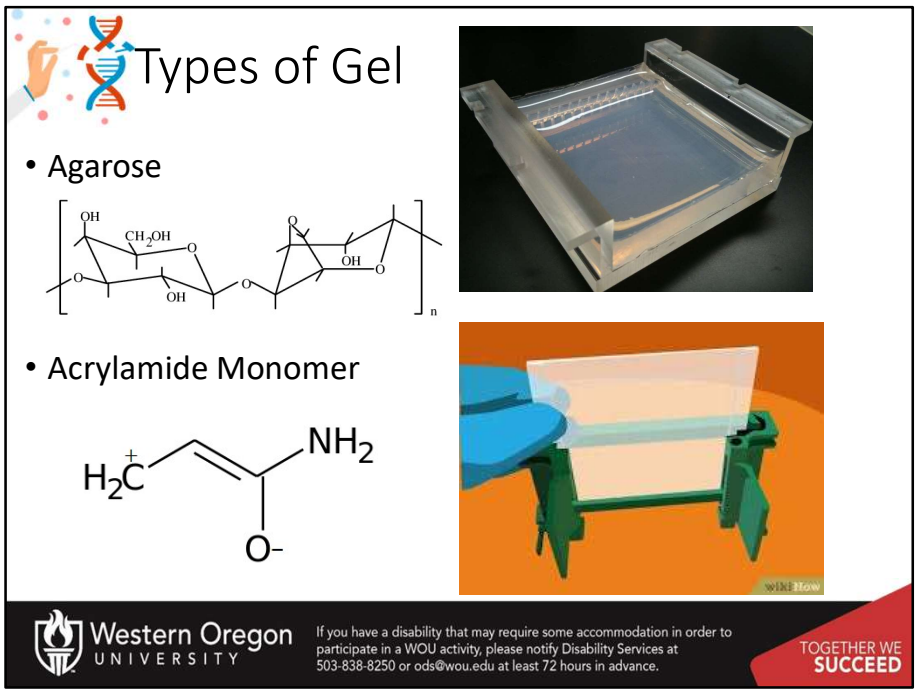
In this section, we will discuss techniques used to identify and visualize proteins. We'll break this section into two parts, as it is quite long!



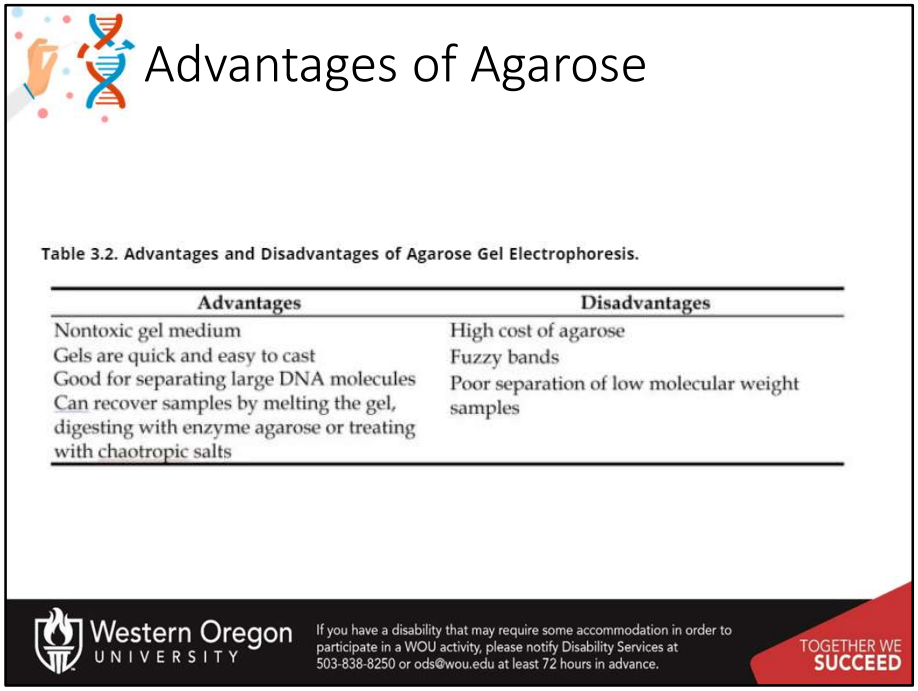
In part 1, we will discuss the first three techniques, which include Gel Electrophoresis, isoelectric focusing, and 2-D Gel Electrophoresis. Part 2 will discuss Antibody-aided assays such as ELISA, Western Blots and Immunohistochemistry.



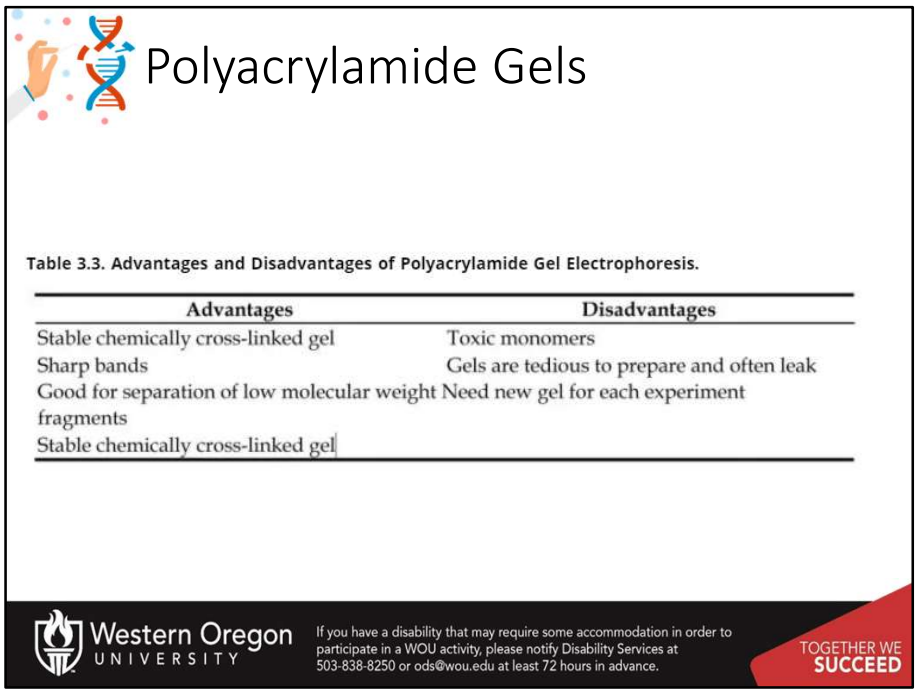
The first one we will cover is gel electrophoresis, where samples of interest are moved through a porous gel matrix under an electrical current.



Agarose is a long, linear polysaccharide compound that is primarily isolated from seaweed, whereas acrylamides are small amide containing compounds that can form crosslinks when exposed to a polymerizing agent, creating a porous polyacrylamide matrix. Both are used in biochemistry. Polyacrylamide gels are more often used for protein analysis, whereas agarose gels are more often used for DNA/RNA analysis.

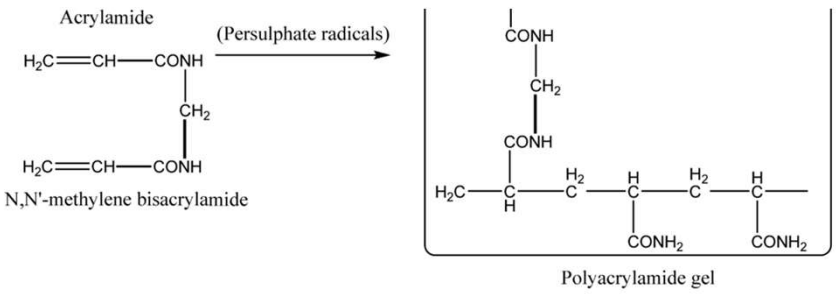


Agarose gels are really good for resolving large molecules, which make it a good medium for the large polymers of DNA. They are also easier to cast than polyacrylamide gels and are nontoxic. Macromolecules can also be recovered from this gel matrix and used in subsequent experiments. Resolution is generally lower than polyacrylamide gels and there is poor separation in lower molecular weight samples. Agarose is also pretty expensive (especially some of the low melt varieties).

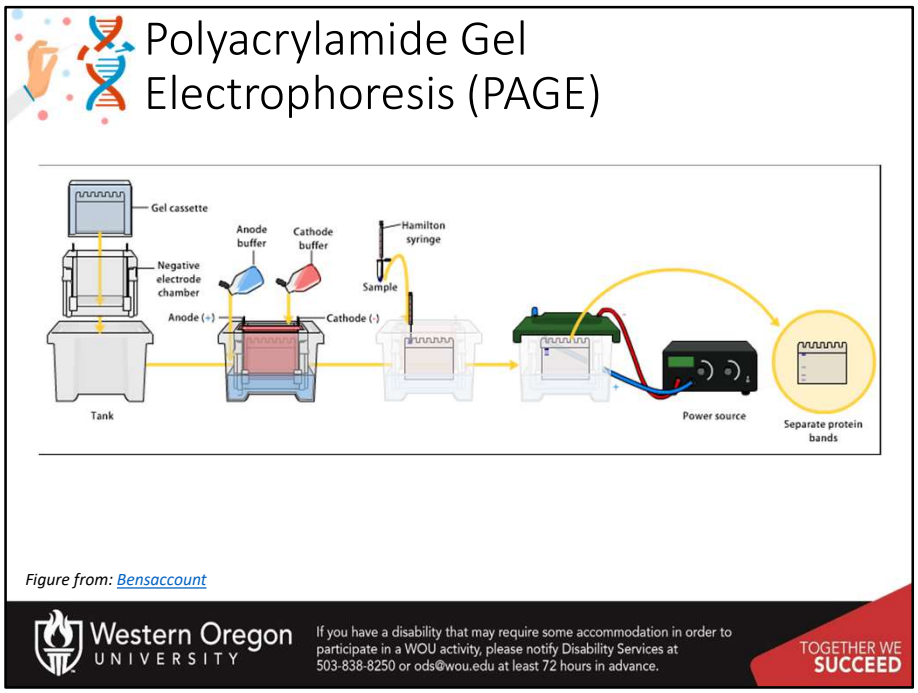


Polyacrylamide gels have good resolving power and are the matrix of choice for protein analysis. It forms a stable crosslinked matrix. Care must be taken when using the unpolymerized monomers of acrylamide, especially in the solid form. It is a neurotoxin and can submilate (go from a solid to a gaseous form directly) which makes inhaling it a real possibility! It is often sold in a liquid form to help avoid toxicity, although it can still cause loss of peripheral neuron function if the solution contacts the skin in repeated exposures. The gels are much more tedious to make and require a polymerizing agent to start the crosslinking reaction.

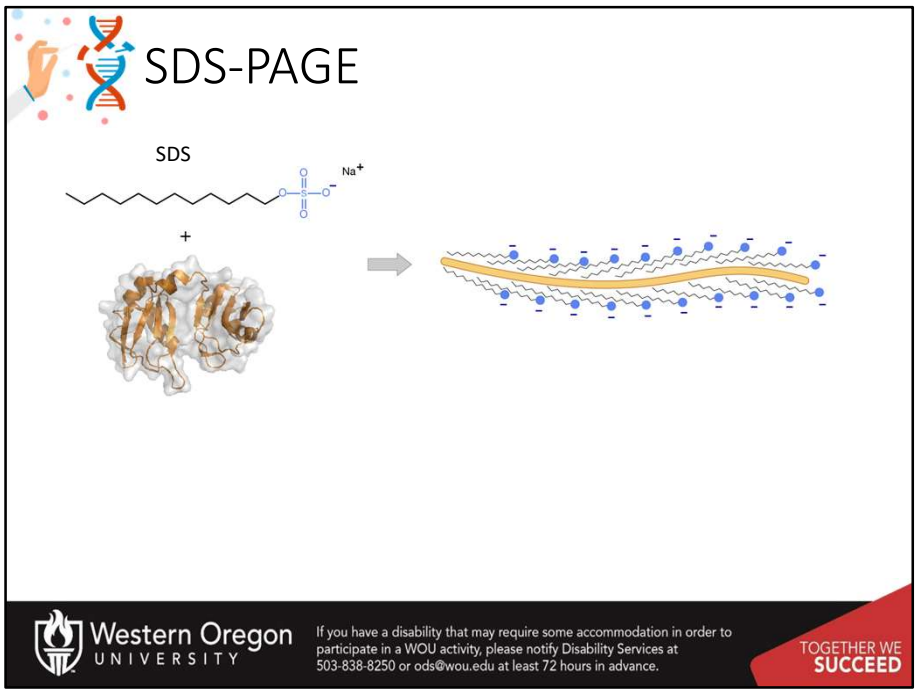
Acrylamide Crosslinking Reaction



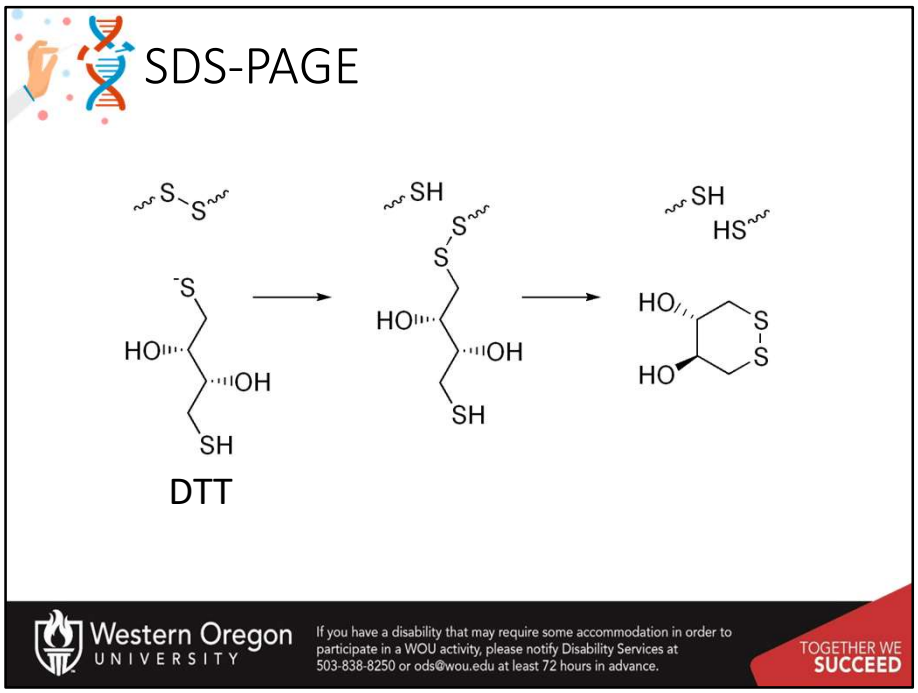
Persulfate radicals are used as the polymerizing agent that initiate acrylamide crosslinking. This forms the porous gel matrix that the protein molecules will move through. We will focus on polyacrylamide gels for the remainder of the protein chapter, but will return to agarose gels when we talk more about DNA.



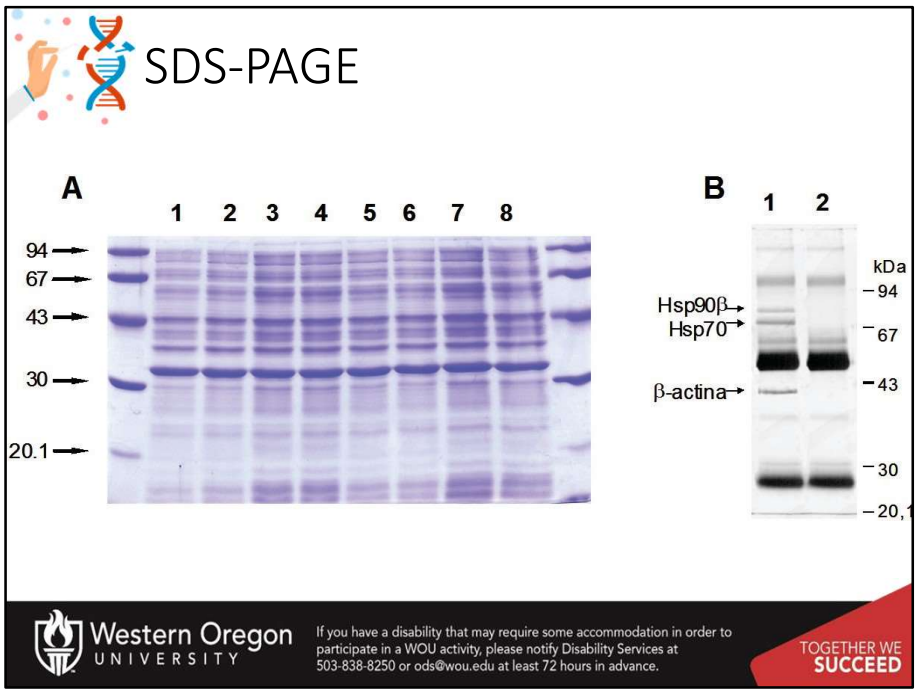
polyacrylamide gel electrophoresis (PAGE) is undoubtedly one of the most widely used techniques to characterize complex protein mixtures. It is a fast, inexpensive method that only requires micrograms quantities of protein. The gels are usually run in a vertical fashion with an upper and lower buffer chamber where buffer contacts the gel. An electrical current is then applied to cause the movement of the proteins in the mixture. Proteins do not have a predictable structure as nucleic acids, and thus their rates of migration are not similar to each other. Furthermore, they will not migrate when applying an electromotive force, when the pH of the system is the same as isoelectric point. PAGE gels that are run in this fashion are called ***Native PAGE***, as the proteins are still folded in their native state found *in vivo*. In this situation, proteins migrate according to their charge, size and shape.



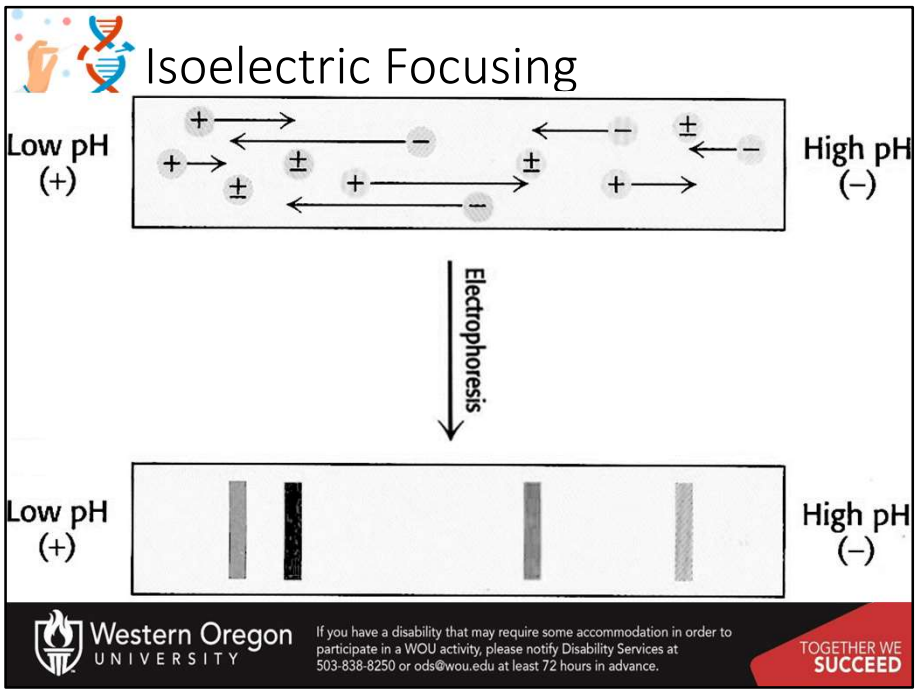
Alternatively, proteins may be denatured prior to electrophoresis. The most common way to denature the proteins is by adding a detergent such as sodium dodecyl sulfate (SDS). This not only denatures the proteins, but it also coats the protein with a negative charge, such that all of the proteins will run towards the positive lead when placed into an electric field. This type of electrophoresis is referred to as ***SDS-PAGE*** and separates proteins exclusively according to molecular weight.



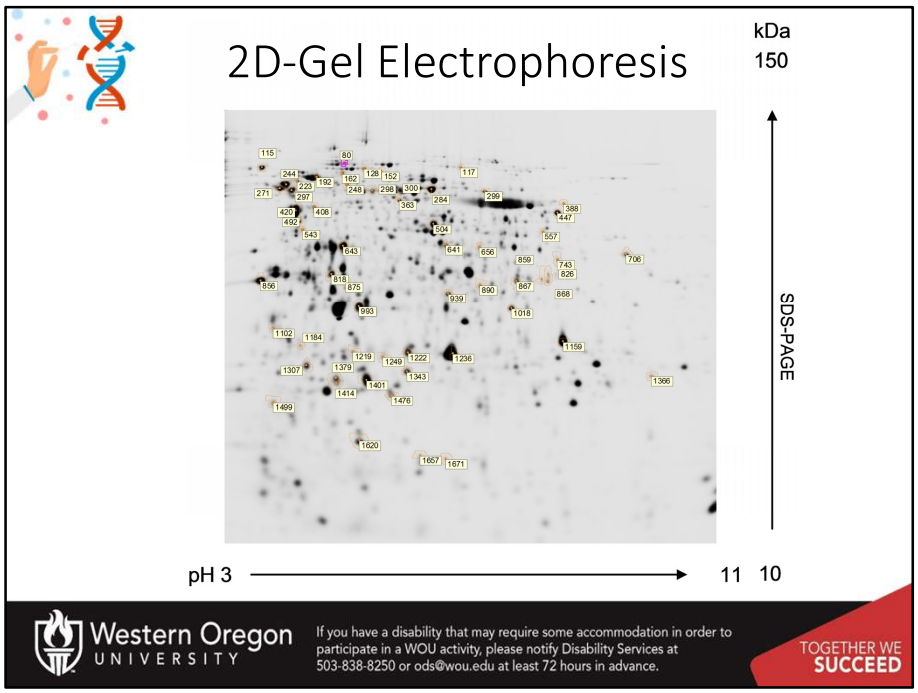
A reducing agent that breaks disulfide bonds, such as dithiothreitol (DTT) is often added to the loading buffer as well, causing proteins to fully denature and dissociate into the monomer subunits. This ensures that the proteins migrate through the gel in direct relation to their size, rather than by charge or shape. Notice that the DTT is oxidized in the process.



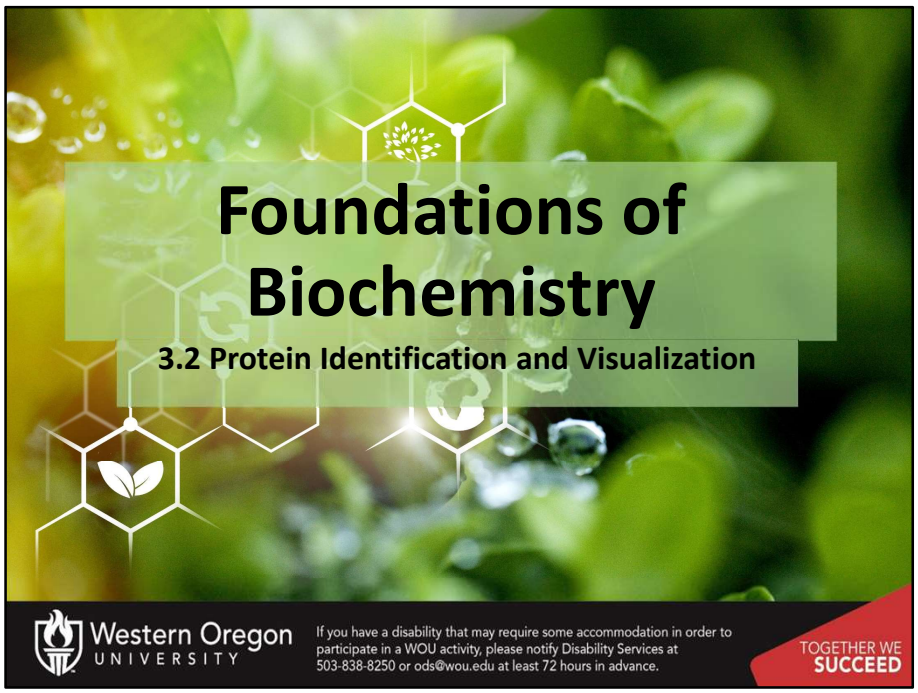
Proteins separated on a polyacrylamide gel can be detected by various methods, for instance dyes (like Coomassie Blue) and silver staining. Proteins can also be labeled with radioactive tags that can be detected.



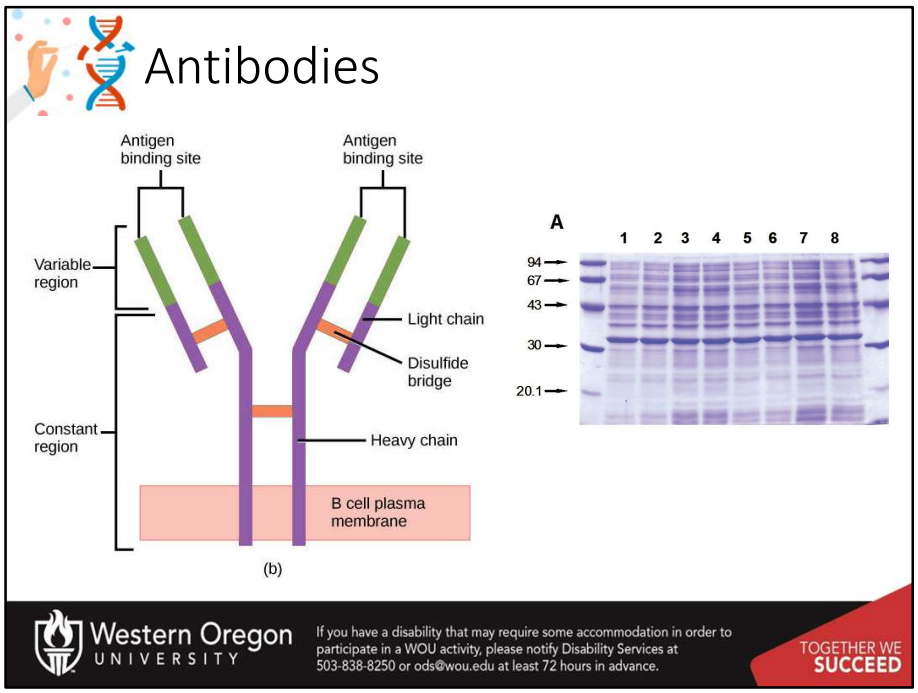
Isoelectric Focusing is a technique is based on the movement of molecules in a pH gradient. The gels are poured such that the gel region near the anode (+) is acidic and near the cathode (-) is alkaline. Between them is a pH gradient between the two extremes, such that the proteins will migrate in the gel to their isoelectric point. Once they reach their isoelectric point, they will be neutral and won't be attracted to the (+) or (-) leads.



SDS-PAGE and Isoelectric focusing can be combined, such that protein mixtures can be separated in these two dimensions on a single gel. The first dimension separation is done by isoelectric focusing, where the proteins are maintained in their native shape and display their native isoelectric point. The sample is loaded in the lower left corner and separated in this direction first. The gel is then equilibrated in a solution containing SDS and DTT to cause protein denaturation. The gel is then run in the second direction causing the denatured proteins to separate by size.

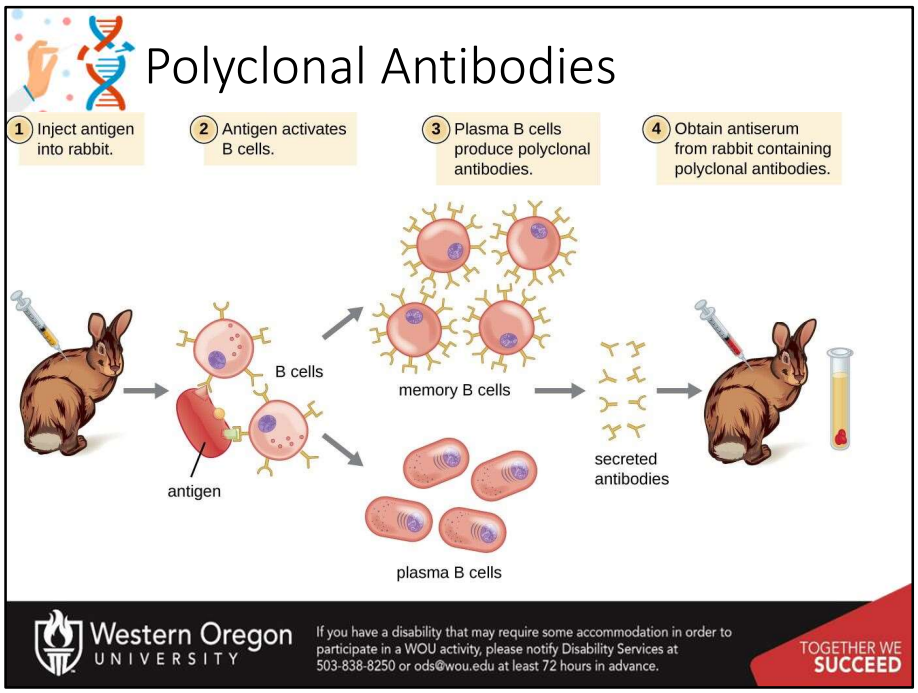


Welcome to the second lecture on Chapter 3.2. In this section, we will focus on the production and use of antibodies in protein detection and identification.

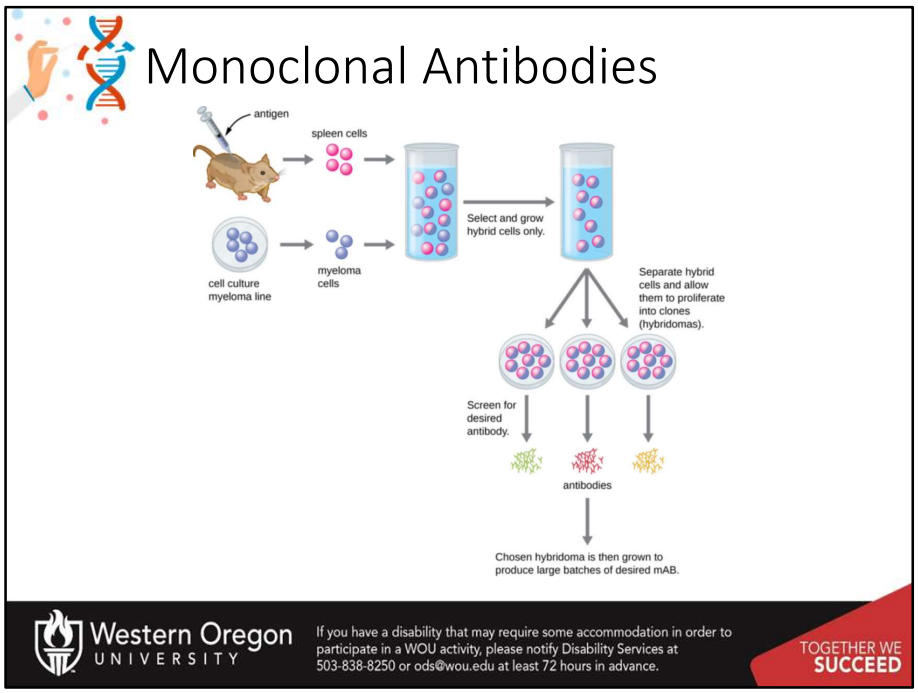


The use of antibodies can extremely enhance our abilities to visualize a single protein within a mixture. Think about our SDS-PAGE gel from the following slide. It would be really hard to visually see a single protein from this complicated mixture and be confident that we are in fact, viewing our protein of interest. If we can use antibodies to highlight where our specific protein is at within a mixture, it can be a powerful tool to help us understand the structure and function of individual proteins within a system.

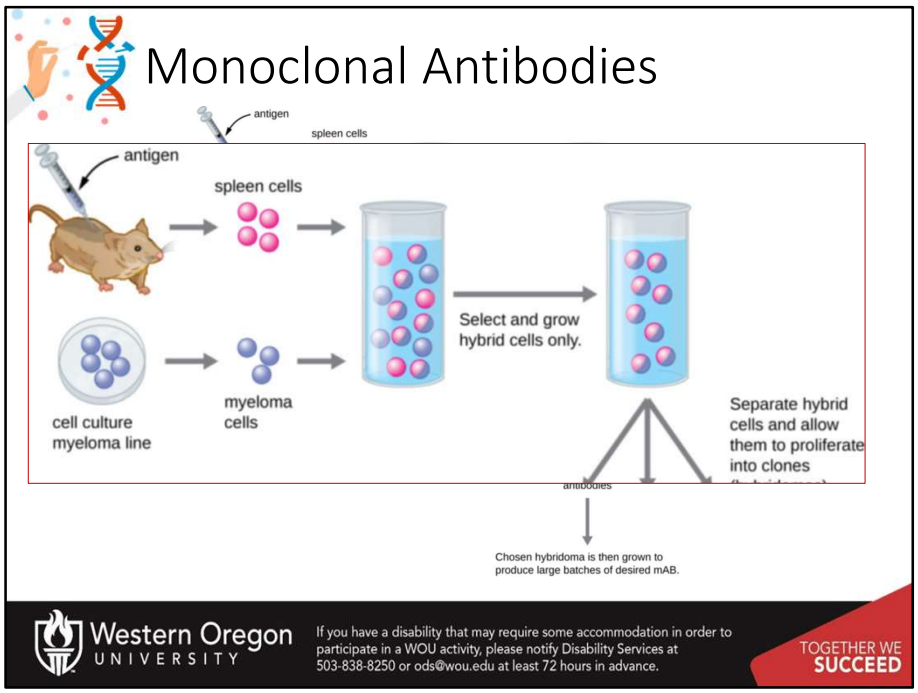
There are several different types of antibodies that can be produced, in vivo, but the ones most commonly used in biochemical studies are the Immunoglobulin type G (IgG) antibodies, as the one shown here. An IgG antibody molecule is comprised of four polypeptides: two identical heavy chains (large peptide units) that are partially bound to each other in a "Y" formation, which are flanked by two identical light chains (small peptide units). Note that each IgG antibody contains two antigen binding sites at the ends of the Y. Antibodies can be raised to specific proteins of interest, by exposing an organism to a non-native protein causing them to have an immune response.



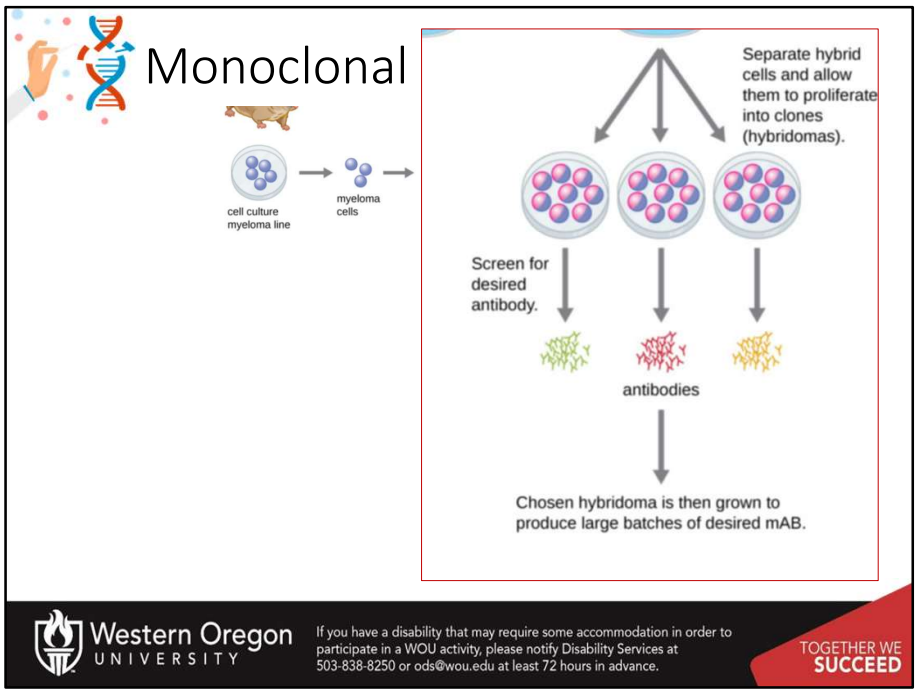
For example, if a rabbit is injected with a human protein, that protein will behave as an antigen and cause the rabbit's immune system to generate antibodies that recognize the foreign protein. The rabbit will then generate memory B-cells that can produce antibodies against the protein of interest. Further challenging the rabbit with second, third and fourth exposures to the protein, will increase the amount of free IgG antibodies that are then secreted from the memory cells. Collection of blood from the animal enables the purification of the secreted antibodies. Because most antigens are complex structures with multiple epitopes, they result in the production of multiple antibodies in the lab animal. Thus, the sera produced is said to contain ***polyclonal antibodies,*** or a pool of antibodies that recognize different epitopes on the same protein. Polyclonal antibodies are very useful for many types of studies, however, some studies require that there is only a single or monoclonal antibody that recognizes a single epitope on the protein. It is very difficult to separate polyclonal antibodies from one another in a pool of antibodies. Thus, monoclonal antibodies must be produced in a different way.



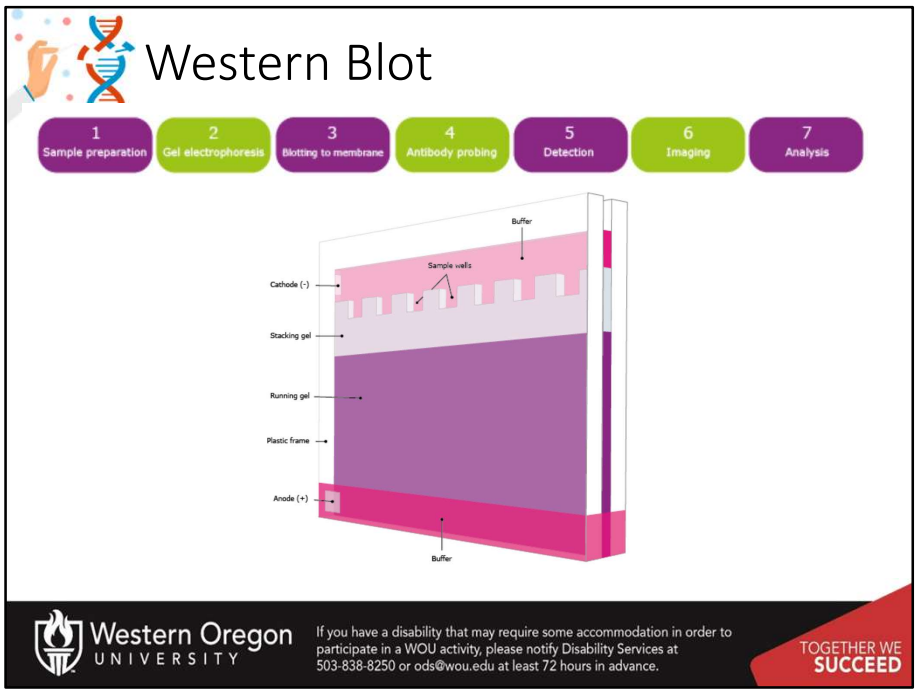
Monoclonal antibodies are more complicated to produce.



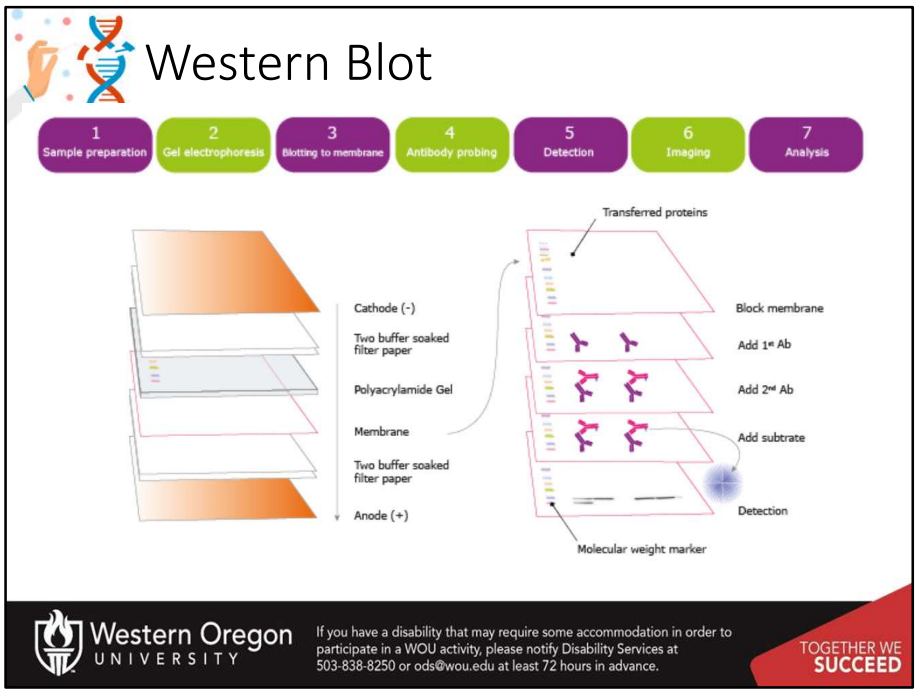
In this system a mouse is repeatedly exposed to an antigen (or protein of interest in this case) to cause an antibody response. This step is similar to what is seen in the production of polyclonal antibodies. However, at this stage, the animal is sacrificed and the spleen is harvested. The spleen cells are hybridized with a myeloma tumor cell line that is immortal and will grow indefinitely in culture (which is unlike the normal spleen cells harvested from the mouse). A special selection medium is used that will only allow the hybrid cells to grow. These cells are called hybridoma cells.



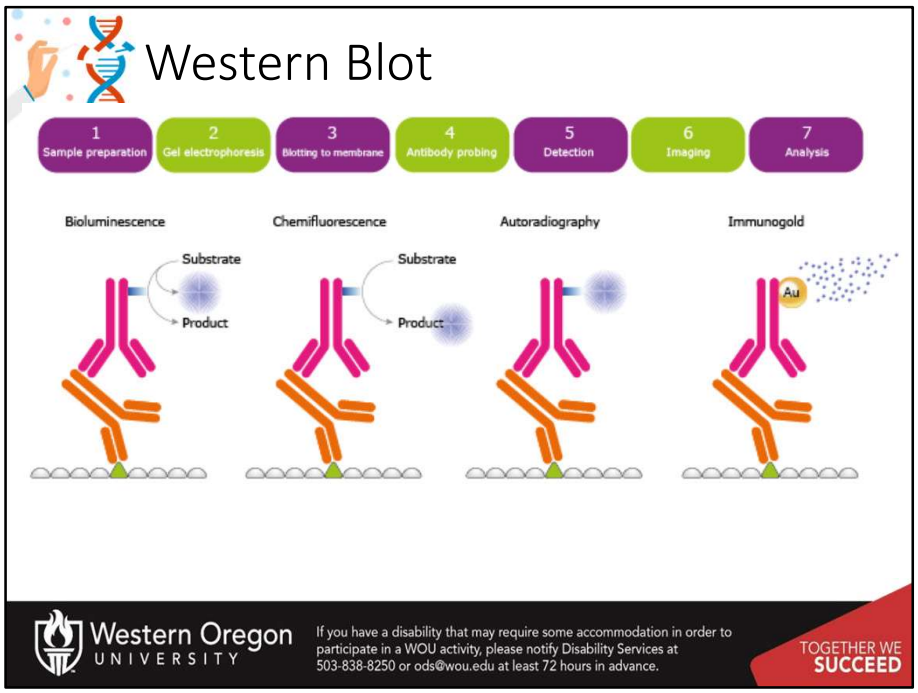
The hybridoma mixture contains polyclonal antibodies that recognize different epitopes of our protein of interest. Thus, to separate these cells, the cells are diluted so that only a single cell will make it into each culture vessel (usually several 96 well plates are used for this step). The single cells are allowed to grow out and each will produce a single antibody with a specific epitope for the protein of interest. This is called a monoclonal antibody. The nice thing about this type of antibody is that it can be cultured and produced indefinitely, whereas polyclonal antibody production and composition can vary in between different host organisms, and even within the same host at different times.



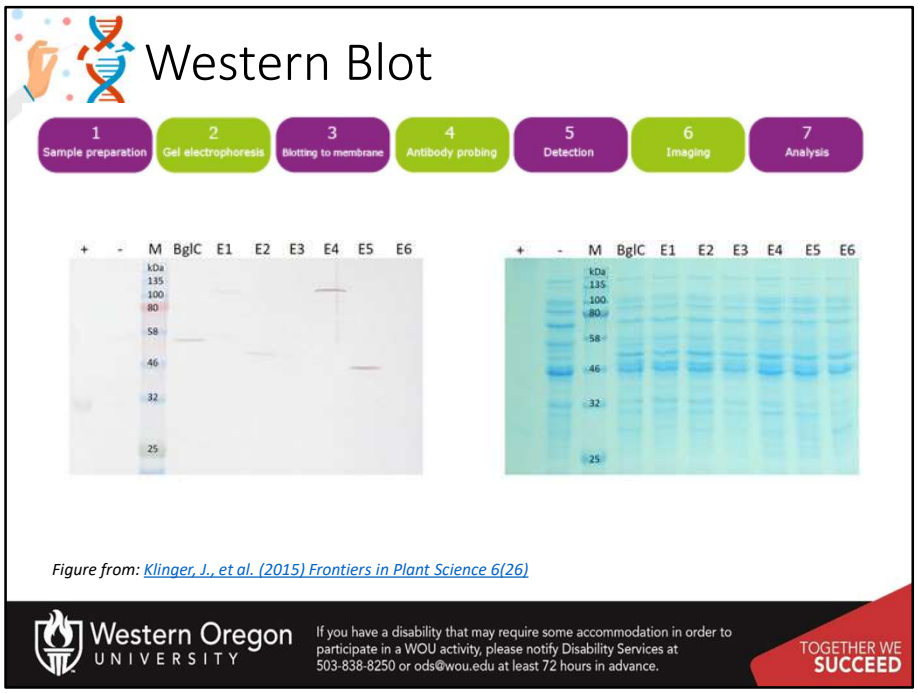
Antibodies can be combined with our SDS-PAGE experiments to produce a new kind of experiment called the Western Blot. There are seven major steps to the Western Blot. The first is sample preparation. This will be the same as for a normal SDS-PAGE Gel. A protein lysate is prepared according to a purification scheme. A small sample is then denatured by heating, and addition of SDS and DTT, to ensure that the protein will migrate in the gel based on size. The samples are loaded into the wells of the gel. A ladder of known protein sizes is also loaded into one of the wells. The proteins, that are coated with the negatively charged SDS particles will run towards the anode (+) charge.



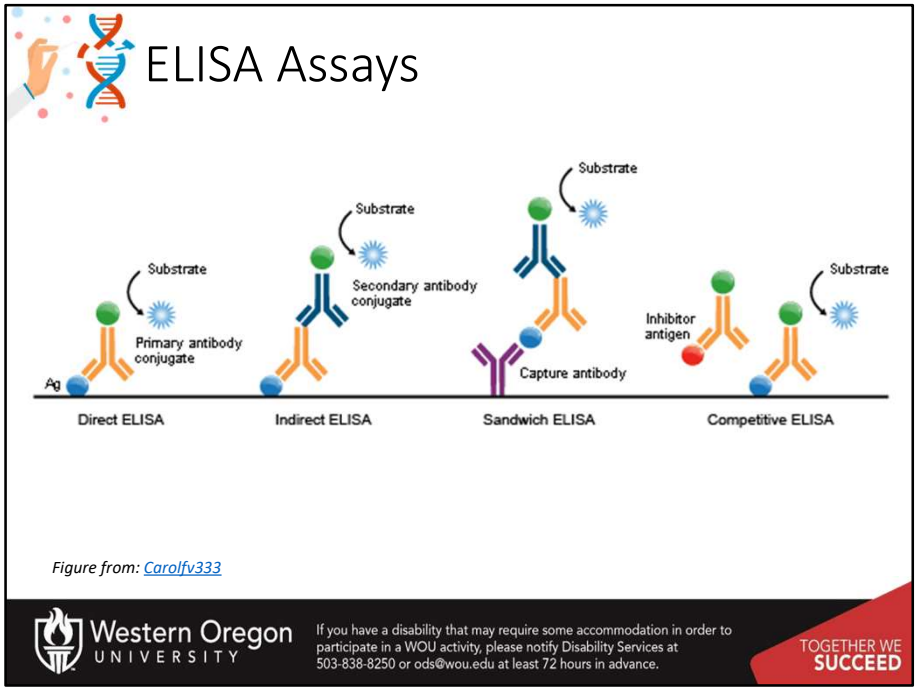
Since the gel is very thin and hard to work with, after it has been used to separate the samples, the proteins are then transferred from the gel to a hydrophobic membrane. This is process is called Blotting. Essentially a sandwich is made with the gel and membrane at the center, and then surrounded by filter paper. It is then exposed to an electric current, so that the proteins will leave the gel and get stuck in the membrane. The hydrophobic nature of the membrane keeps the proteins from moving beyond the membrane and into the filter paper. The proteins are then fixed to the membrane by exposing it to UV-light. The UV-light causes covalent cross-linking to occur between the proteins and the membrane. You can now rigorously wash the membrane and the proteins will not be dislodged.



In steps 4-7, antibodies will be used to detect the protein of interest. First, the blot must be pretreated with nonspecific proteins to block non-specific protein binding sites inherent in the membrane. Nonfat Dry Milk is often used for this purpose. Then a primary antibody, that is specific for your protein of interest, is incubated with the blot. It will bind to the blot only where your protein of interest is found. The blot is washed gently to remove any unbound primary antibody, and then a secondary antibody is added to the blot. This antibody is specific for the heavy chains of antibodies, and will bind to the heavy chain regions of your protein-specific antibody. This secondary antibody is also bound to a detection molecule. This detection molecule can have bioluminescence, chemifluorescence, radioactivity, or colored properties that will make the detection of its location possible. The blot is then exposed to film to detect the signal.



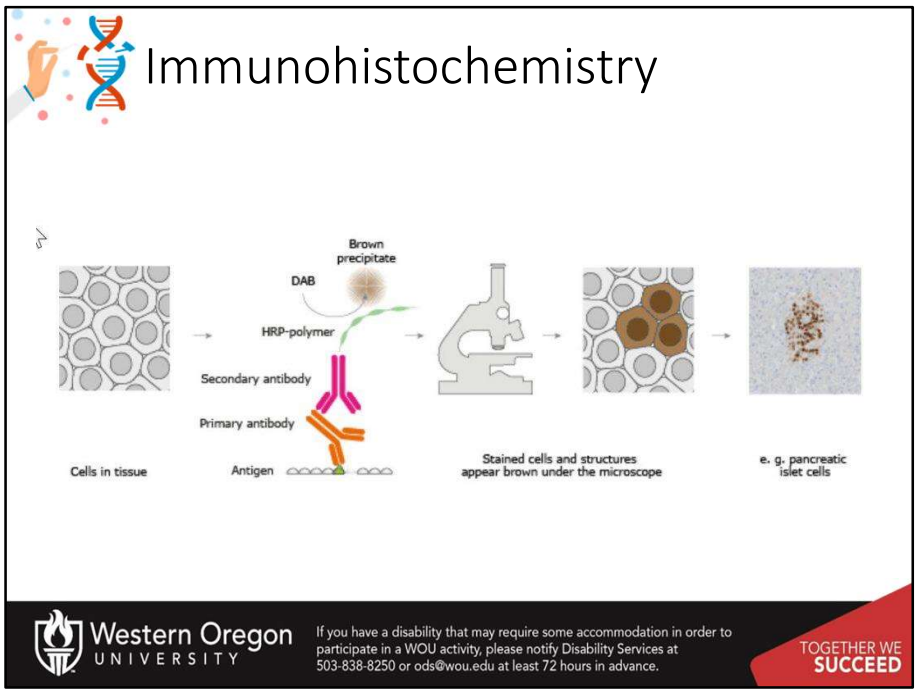
The film on the left shows the results of a Western Blot, and the right shows the blot itself, stained with Coomassie Blue. You can see that the sensitivity of the antibodies can very cleanly detect down to 0.1 ng of protein and allow visibility of the protein in a crowded blot with a dense mixture of proteins. Coomassie staining, on the otherhand, can only detect 100 ng or above.



Enzyme-linked Immunosorbant Assays (ELISAs) offer another method of using antibodies to detect low levels of specific proteins within a lysate. Typically this type of assay is done in a 96-well plate format and uses very small quantities of protein to mediate detection. ELISAs can be run in different ways. In the first two methods, the direct and indirect methods, protein lysates are coated and cross-linked onto the bottom of 96 well plates. In the ***direct ELISA*** a single antibody that is specific for the protein of interest and also has a chemiluminescence marker attached to the heavy chain is incubated with the sample. If the protein of interest is present, the antibody will attach. The well is then washed to get rid of unbound and non-specifically bound antibody, and then the detection solution is applied and sample are monitored using UV/Vis Spectrophotometry. The ***direct ELISA*** is the fastest protocol, but can often suffer from high background. An ***indirect ELISA*** can help decrease background. This method is the same as the direct ELISA, except that the chemiluminescence detection system is attached to a secondary antibody that recognizes the heavy chain of the antibody that recognizes your protein of interest. There are also ***sandwich ELISAs,*** where the primary antibody that recognizes your protein of interest is coated and cross-linked with the 96 well plate. The protein lysate is then incubated with the antibody where it will bind tightly with the antibody. Another primary antibody that recognizes a different epitope of the target molecule is then incubated with the mixture. Wash steps remove any unbound protein or antibody in between each incubation. Then a secondary antibody is used to detect the presence of the primary antibody. ***Sandwich***

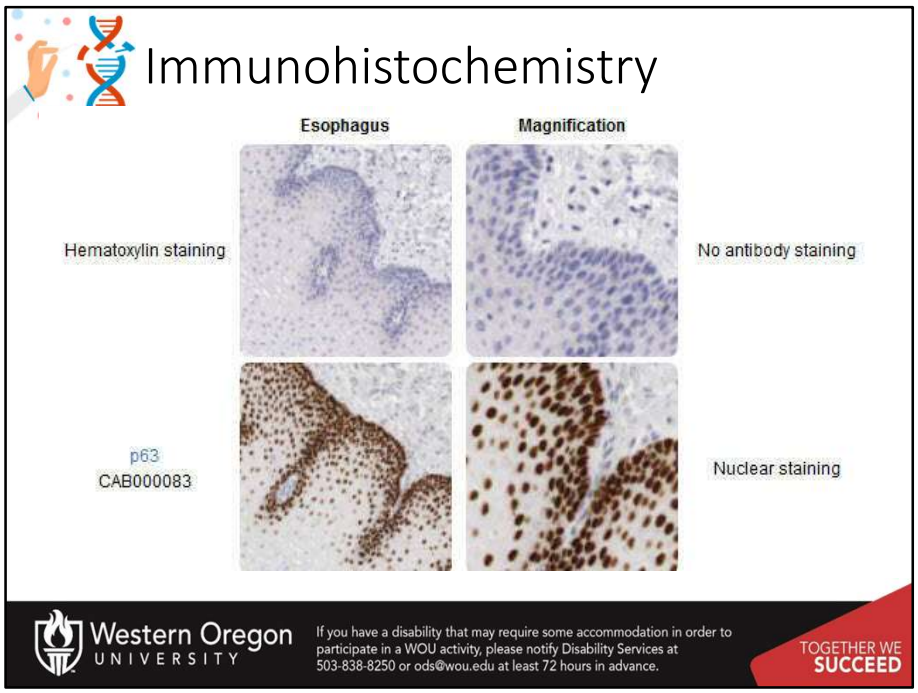
ELISAs have high specificity, but require two primary antibodies that bind different epitopes of the protein. This can be hard to find. The protocol is also much longer with more incubation steps. Competitive ELISAs use a known amount of antigen to coat and bind the 96 well plate. Primary antibody + your protein lysate of interest is mixed together and then added into the well. The more antigen that is present within your sample, the less antibody will be free to bind the known concentration of antigen within the well. If you do a series of experiments like this with different concentrations of your protein lysate, it can be a useful way to help quantitate the amount of your antigen present within the protein lysate.

Here is an example of what a competitive ELISA might look like. You can see that no protein lysate is present on the lefthand side and there is a lot of blue color created as the antibody binds to the known quantity of antigen that is coating the well. As higher concentrations of your protein lyate are added to the mixture, there is less antibody present to bind to the antigen (protein) crosslinked to the well plate. It is effectively being competed away by your protein of interest. This could be useful to detect if protein expression levels change in your cells of interest during different treatments. Perhaps before and after a stress event like exposure to UV light.

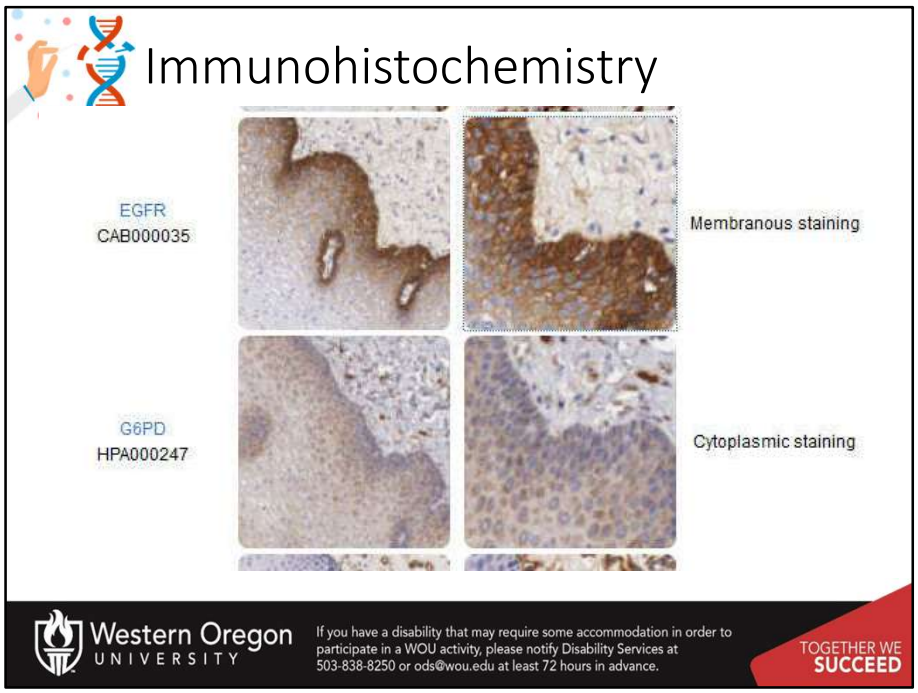


Another way to use antibodies in biochemistry is as a tool to help figure out where proteins are localized within a cell or a tissue. For this type of analysis, very thin slices of tissue are prepared and mounted onto glass slides where they are crosslinked or fixed onto the slide. The cells can be permeabilized by treating them with alcohol or other solutions and then incubated with the primary antibody for your protein of interest. The slide is then washed to remove unbound antibody and is then incubated with a secondary antibody that is linked with a bioluminescent or chemiluminescent marker. The samples are then visualized microscopically for detection.

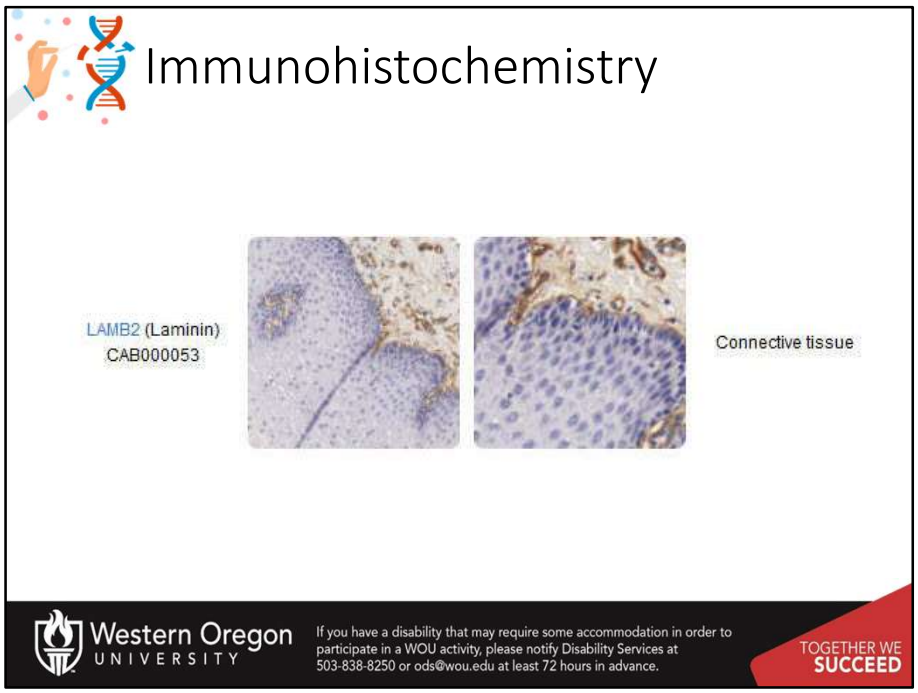
Here are some different examples of staining patterns that can be observed.



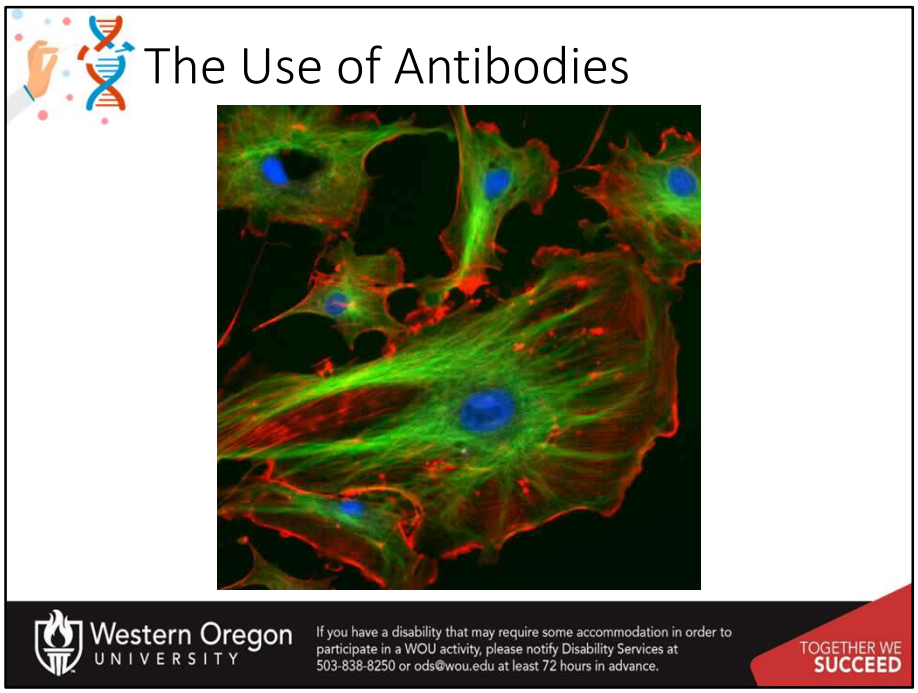
Here are some different examples of staining patterns that can be observed. The top diagram has no staining, whereas the bottom diagram shows nuclear staining, at both lower and higher magnification.



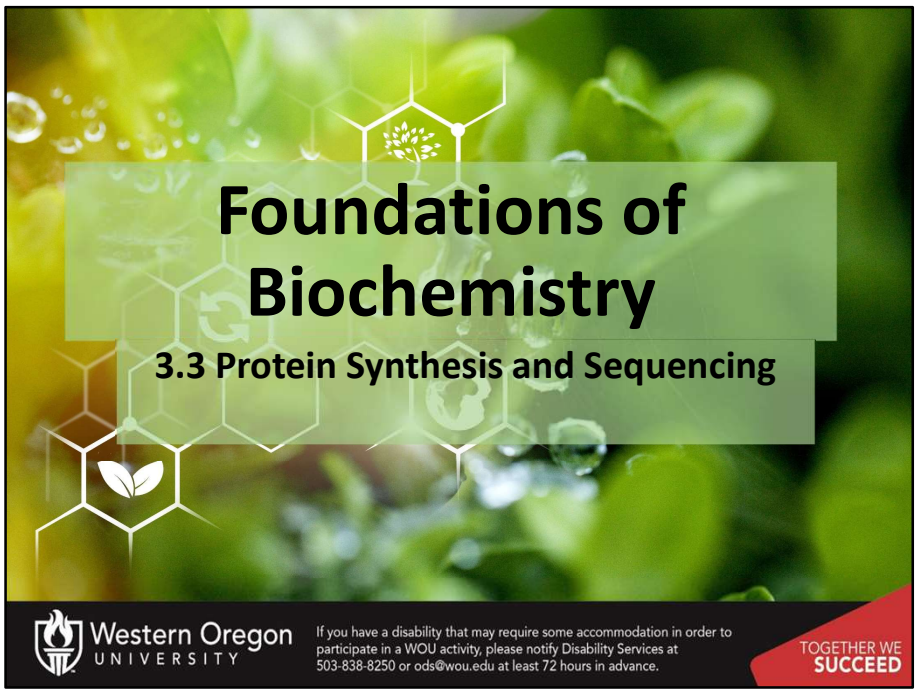
Here are some different examples of staining patterns that can be observed. This one shows membranous staining vs cytoplasmic staining.



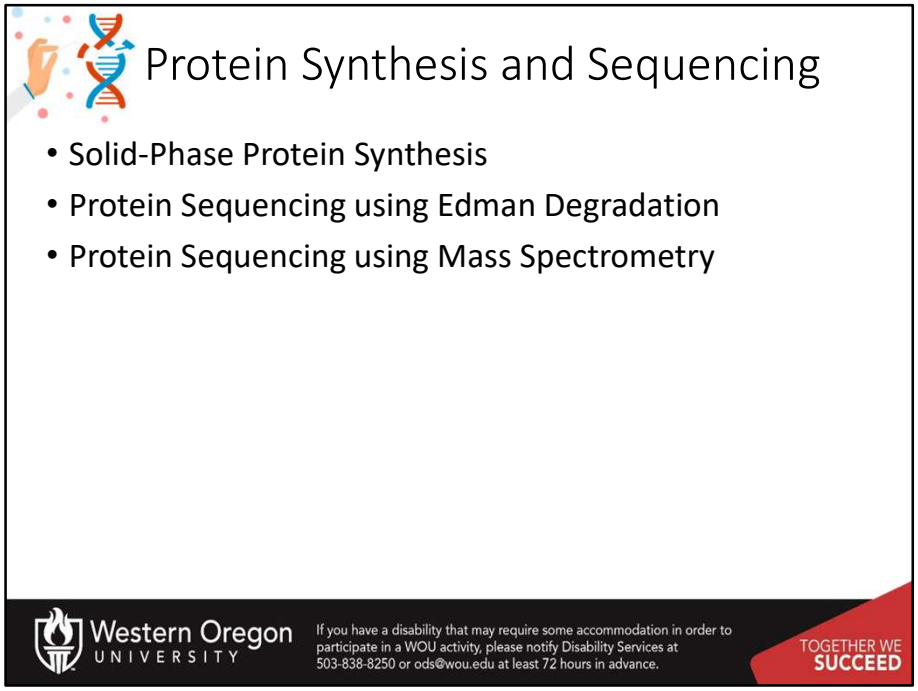
Here are some different examples of staining patterns that can be observed. And this one shows differential staining only within the connective tissue of this sample and not the epithelial tissue.



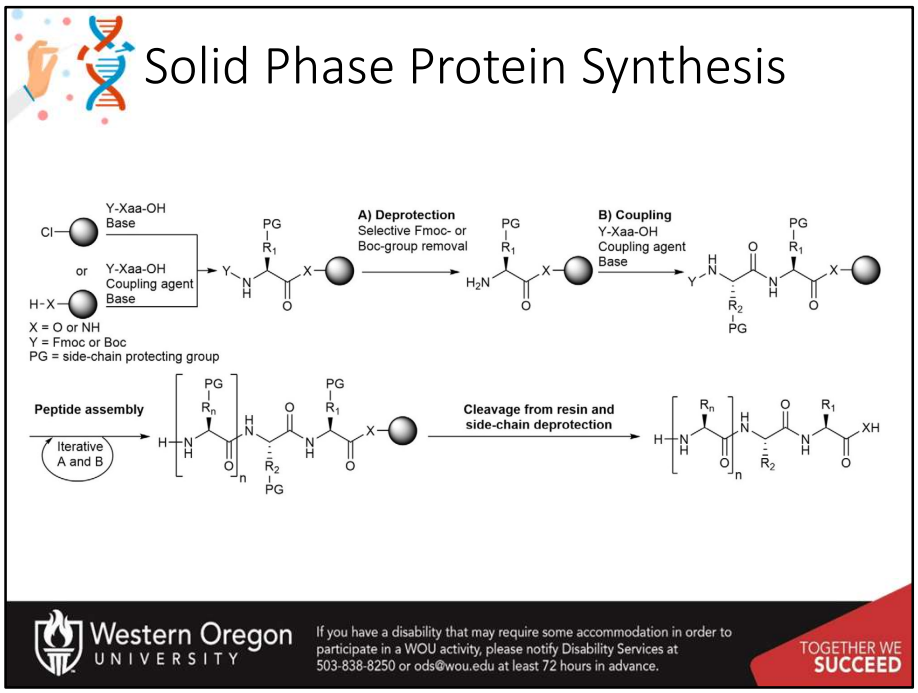
Overall, the use of antibodies within biochemical research has opened up a whole world of cellular exploration that wasn't possible prior to their development. In the next section, we will take a look on methods for sequencing and synthesizing proteins.



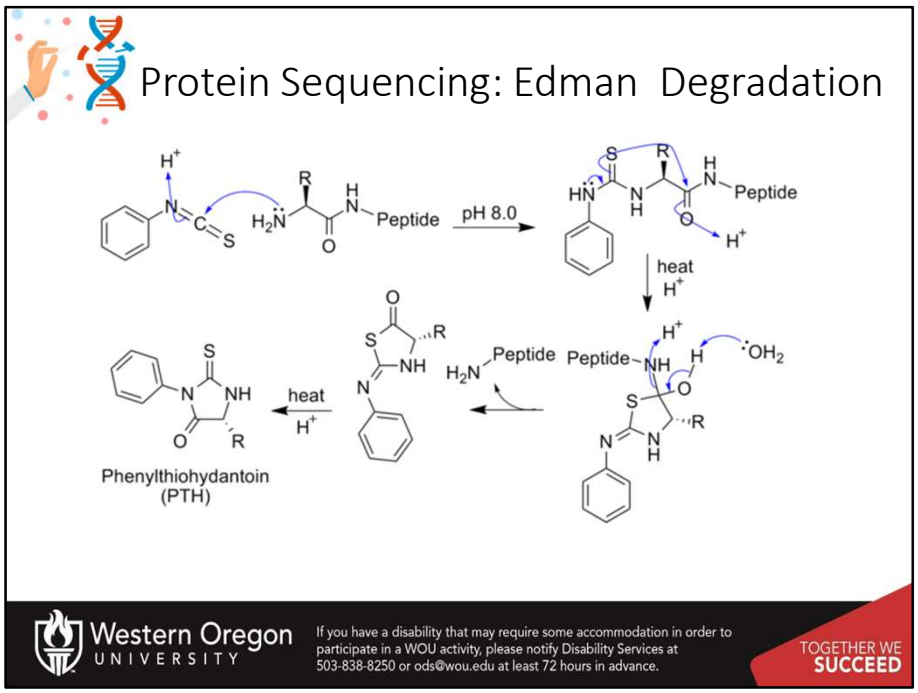
In this section, we will focus on techniques used for protein synthesis and sequencing.



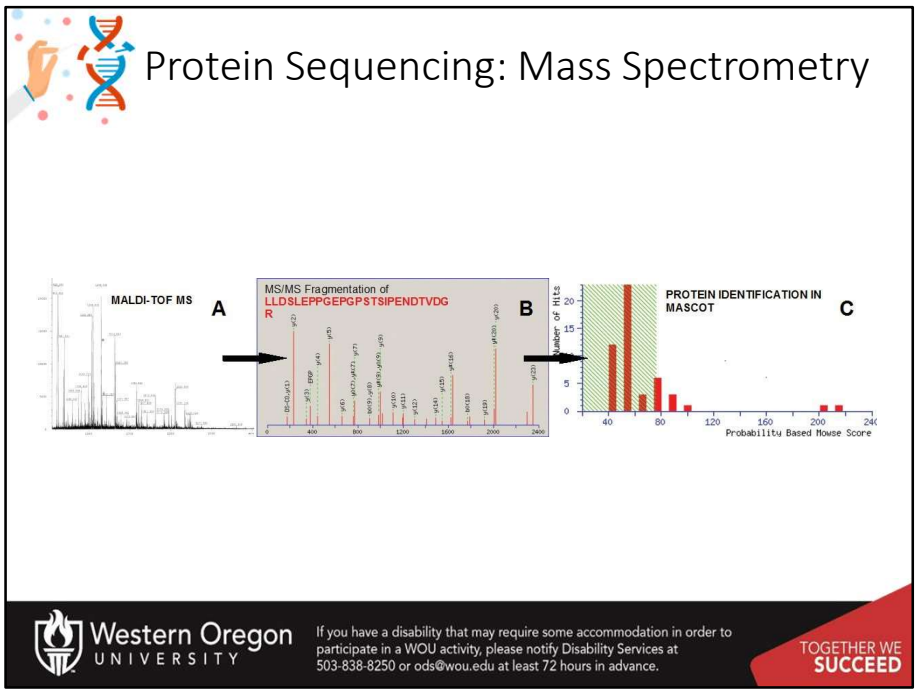
There is one major synthesis technique used (Solid Phase Synthesis) and two methods for protein sequencing (Edman Degradation and Mass Spectrometry).



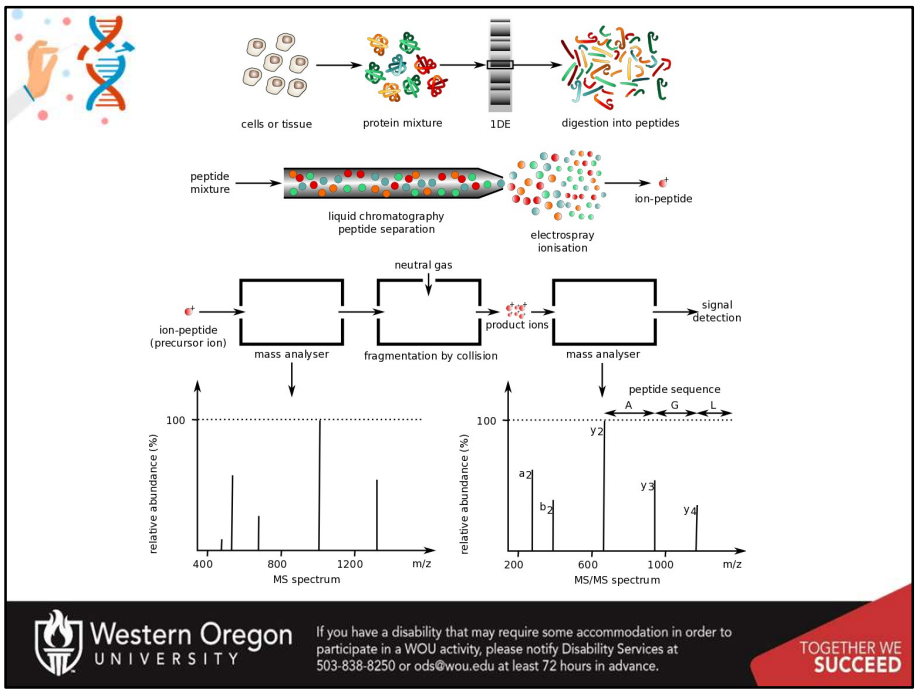
Peptides are chemically synthesized by the condensation reaction of the carboxyl group of one amino acid to the amino group of another. Protecting group strategies are usually necessary to prevent undesirable side reactions with the various amino acid side chains. In solid-phase synthesis, the first amino acid has its N-terminal protected to avoid unwanted reactions, and then the carboxyl terminal is reacted with the solid support of a column to covalently crosslink it with the column beads. Excess amino acid is then washed from the column. The protective group on the amine terminal (labeled Y here) is then removed. The second amino acid to be added to the chain has its N-group protected with the Y-group, and is then added to the column where the amide bond is formed. The column is washed to remove any unbound amino acids and the Y-group is then removed from the nascent peptide. The process is then repeated until the desired length of the peptide is formed. Usually peptide synthesis is limited to a maximum of about 70 residues before the solid support becomes unstable.



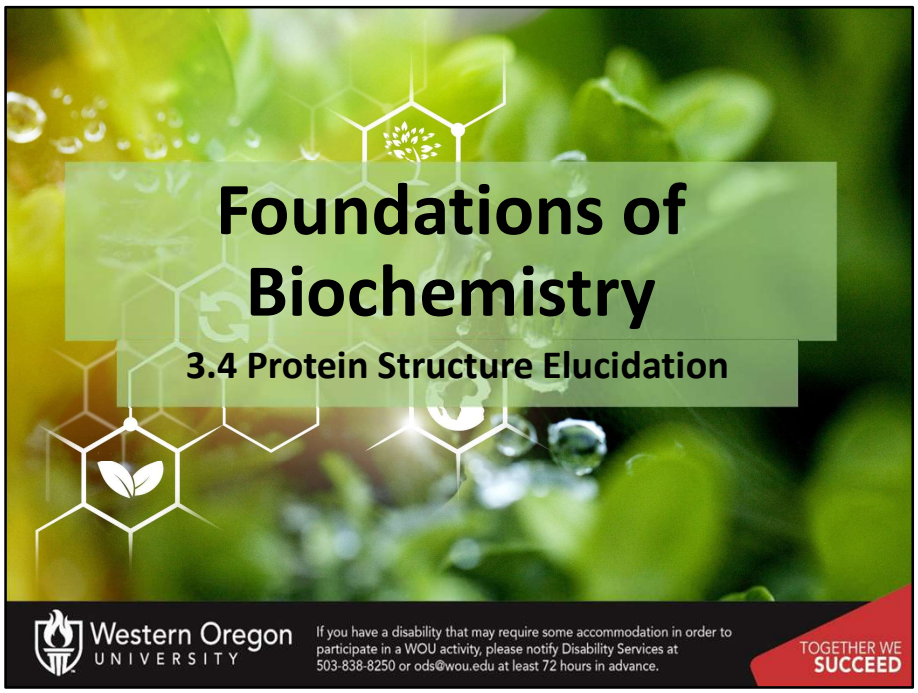
Edman degradation is a method used to sequence the amino-terminal residues of a protein. It can be useful if you have isolated an unknown protein and want to work backwards to identify the gene that encodes it. With this method it is possible to sequence the first 50 – 60 residues. In this method, Phenyl isothiocyanate is reacted with an uncharged N-terminal amino group, under mildly alkaline conditions, to form a cyclical *phenylthiocarbamoyl* derivative. Then, under acidic conditions, this derivative of the terminal amino acid is cleaved as a thiazolinone derivative. The thiazolinone amino acid is then selectively extracted into an organic solvent and treated with acid to form the more stable derivative that can be identified by using chromatography or electrophoresis. This procedure can then be repeated again to identify the next amino acid.



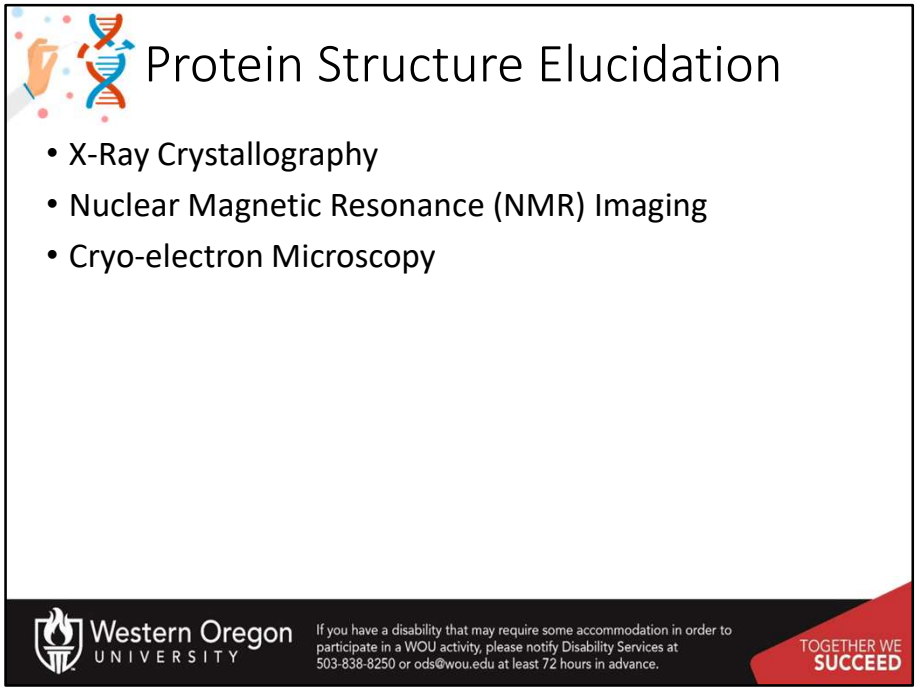
Mass spectrometry offers another way to fully sequence proteins. This can be accomplished with with a top down approach, in which proteins are analyzed intact, or the proteins can be digested first using known proteases and the fragments then analyzed by mass spectrometry. For example trypsin cleaves proteins on the carboxyl-side of Lysine or Arginine (except when they are followed by Proline). Thus, digesting a protein and then running the mass spectrum can create a peptide mass fingerprint that is characteristic of that protein, as shown in figure A. Each peak can then be analyzed further by MS-MS to reveal the sequence of individual fragments (Figure B). This information can then be used to search within the spectral data base, MASCOT to aid in the identification of the protein (Figure C).



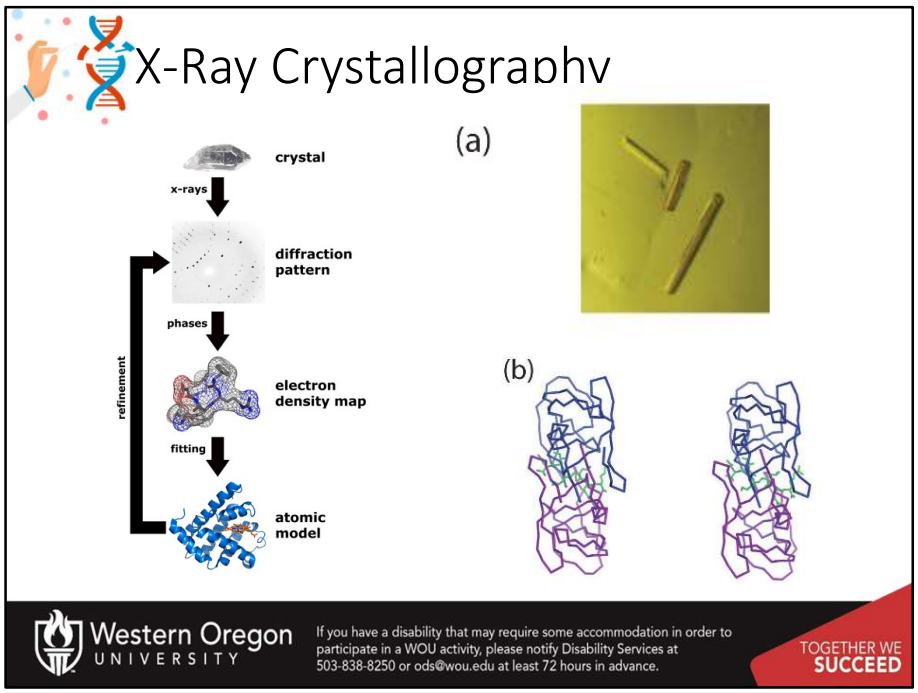
This diagram gives a better representation of the experimental procedure used in protein sequencing via mass spectrometry. First a protein lysate is created from a tissue or cell culture sample. The protein mixture can be separated using SDS-PAGE. A single protein band can be isolated and then digested using trypsin. The fragments can then be analyzed by mass spectrometry to create a fingerprint of the protein. Each peak can then be further analyzed by another round of mass spectrometry to reveal the amino acid sequence of the fragment. In the next section, we will discuss methods for elucidating the 3D-structures and folding patterns of proteins.



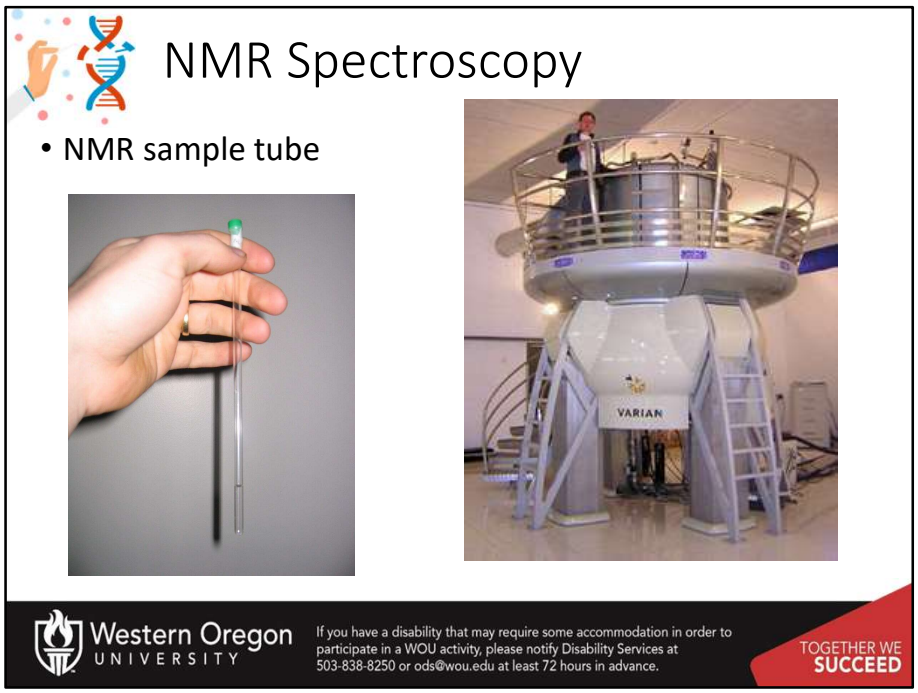
Welcome to our lecture on Protein Structure Elucidation.



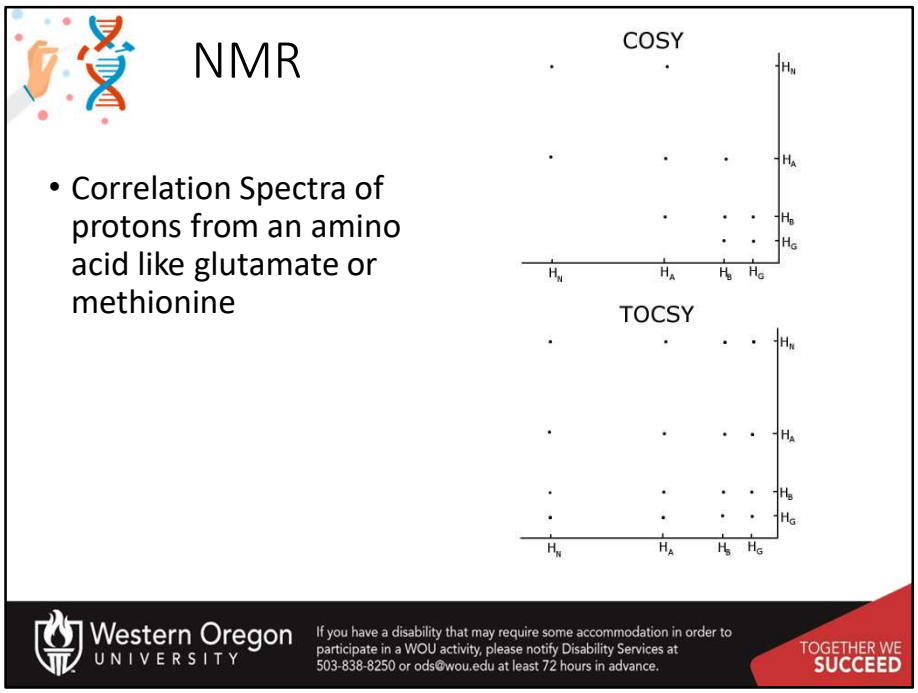
Currently there are three major methodologies for analyzing the 3D-structure of proteins, including X-ray crystallography, nuclear magnetic resonance (NMR) imaging, and cryo-electron microscopy. All of these methodologies are incredibly complex and intricate to learn. We will just take a look at the basics of each and discuss some of their pros and cons.



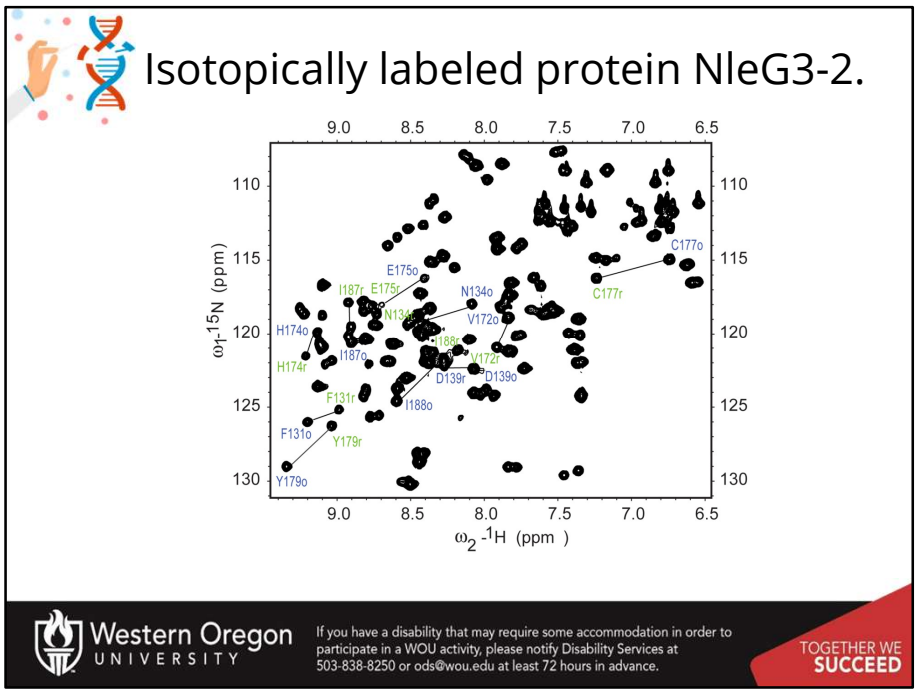
With X-ray crystallography, the process begins by crystallizing a protein of interest. Crystallization of the protein causes all the protein atoms to be orientated in a fixed way with respect to one another while still maintaining their biologically active conformations. This fixed crystal structure is a requirement for X-ray diffraction. To form protein crystals, the protein solution must be supersaturated to start forming nuclei. It is a difficult and often timely process to find the right parameters for crystal growth and formation. Once they have formed, they can be bombarded with x-rays and the diffraction patterns recorded. The structure is then recreated from the diffraction map using Fourier Transformation. High quality structures have 2Å resolution. A stereoimage of the HIV Protease with its inhibitor is shown in (B) on the right. This structure was solved by X-ray crystallography. By analyzing the structure, researchers were able to start designing new protease inhibitors



NMR spectroscopy is another method commonly used to help elucidate the 3D structure of proteins. Currently most samples are examined in a solution in water, but methods are being developed to also work with solid samples. Data collection relies on placing the sample inside a powerful magnet, sending radio frequency signals through the sample, and measuring the absorption of those signals. Depending on the environment of atoms within the protein, the nuclei of individual atoms will absorb different frequencies of radio signals. Furthermore, the absorption signals of different nuclei may be perturbed by adjacent nuclei. This information can be used to determine the distance between nuclei. These distances in turn can be used to determine the overall structure of the protein.



With unlabelled protein the usual procedure is to record a set of two dimensional homonuclear nuclear magnetic resonance experiments through correlation spectroscopy (COSY), of which several types include conventional correlation spectroscopy, *total correlation* spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). The COSY and TOCSY spectra above are representative of an amino acid such as glutamate or methionine showing proton correlations. The TOCSY shows off diagonal crosspeaks between all protons in the spectrum, but the COSY only has crosspeaks between neighbors. The complexity of the spectra dramatically increases when analyzing a molecule the size of protein



If recombinant proteins can be produced, the resulting protein can be labelled with Nitrogen-15 or with Carbon-13 to allow for more detailed experimentation, such as heteronuclear single quantum coherence spectroscopy (HSQC). The most commonly performed 15N experiment is the ^{1}H-^{15}N HSQC. Each peak in the spectrum represents a bonded N-H pair, with its two coordinates corresponding to the chemical shifts of each of the H and N atoms

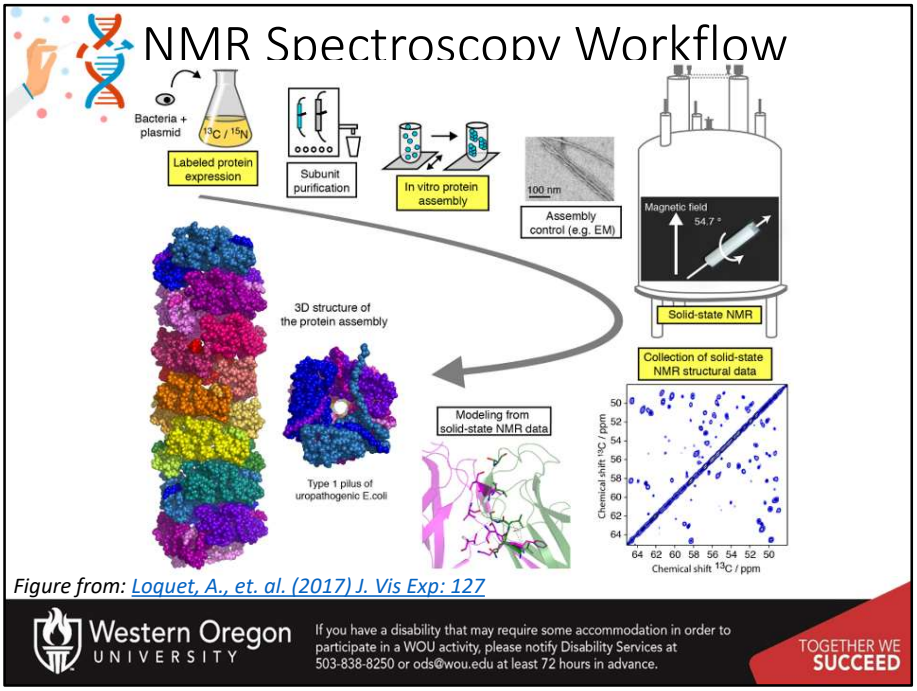
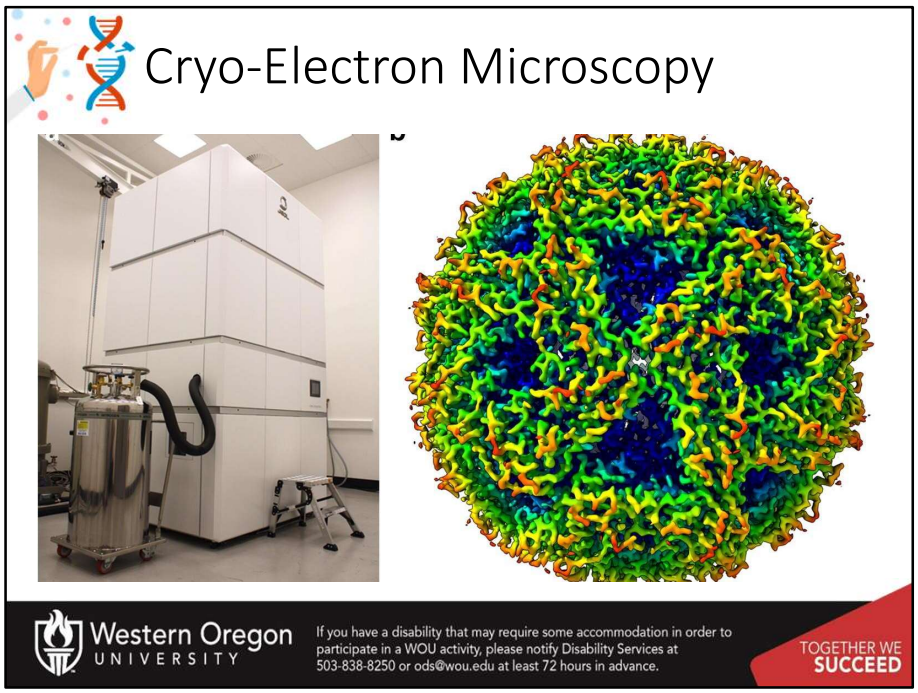
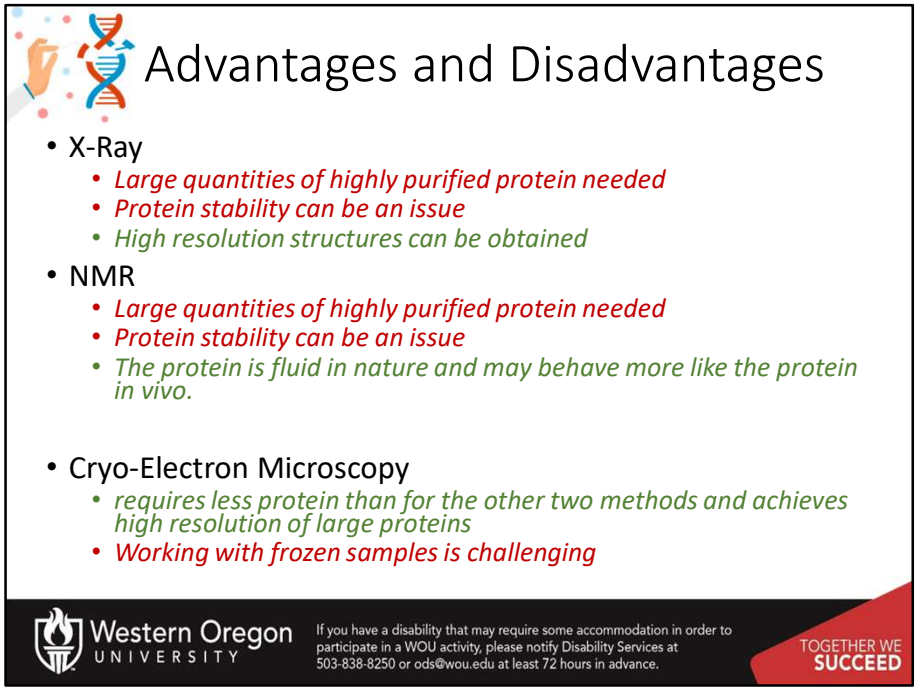


Figure from: Loquet, A., et. al. (2017) J. Vis Exp: 127

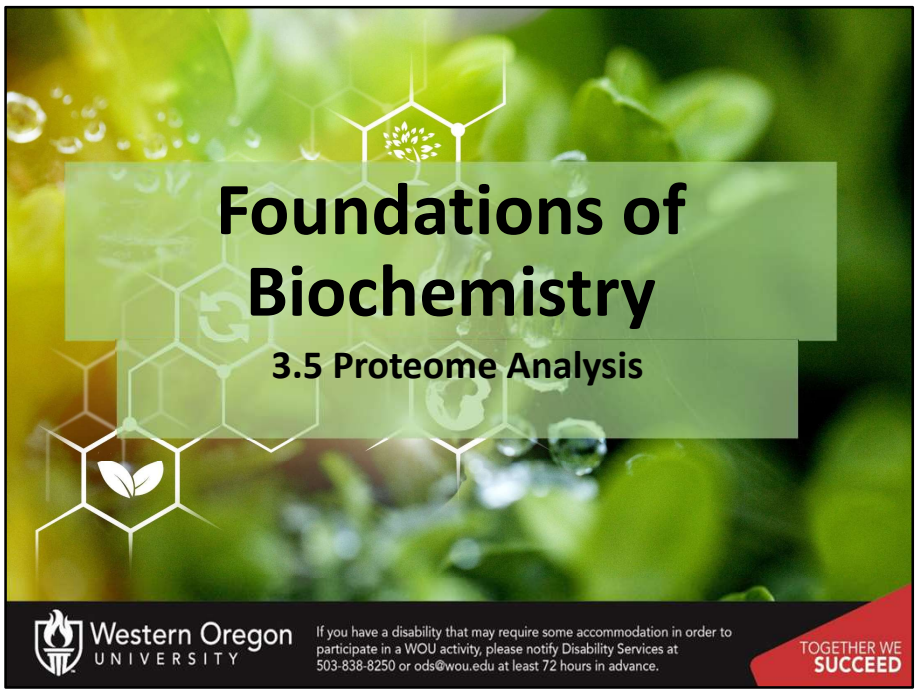
This diagram shows a typical workflow for determining a 3D structure of a protein using NMR. This starts with labelling a recombinant protein with 13C and 15N during the growth of the bacteria. The protein is then purified (and in this case, specific subunits of the protein are expressed separately and reassembled within a test tube. The assembled protein is then analyzed using NMR. Data interpretation allows the recreation of the 3D structure.



Cryogenic-electron microscopy (cryo-EM) has recently emerged as a powerful technique in structural biology that is capable of delivering high-resolution density maps of macromolecular structures. Resolutions approaching 1.5 Å are now possible and the regeneration of very large macromolecular structures are possible. In this technique, protein suspensions are frozen on 3-mm-diameter transmission-electron microscope (TEM) support grids made from a conductive material (e.g. Cu or Au) that are coated with a carbon film with a regular array of perforations 1–2 μm in diameter. A total of 3–5 μl of sample is loaded onto the grid which is then immediately blotted with filter paper with the aim of creating a film of buffer/protein on the grid that, when frozen, will be thin enough for the electron beam to penetrate. Optimising the ice thickness is a vital step in sample preparation as thicker layers of ice increase the probability that the incident electron will undergo multiple scattering events and thereby reduce the image quality. The high resolution of Lumazine Synthase is shown here as an example.



Obtaining pure, highly concentrated protein is a major setback for both X-ray crystallography and NMR spectroscopy, whereas cryo-electron microscopy does not require as much. NMR has some advantages as the protein remains in a fluid system, rather than crystalized or frozen, as in the other two methods. Thus, it may have a more realistic 3D structure. X-ray crystallography and cryo-electron microscopy can produce really high resolution images and cryo-electron microscopy can yield structures for large, macromolecular structures.



The final section of Chapter 3 deals with proteome analysis.

Proteomics deals with the study of all of the proteins expressed within a system, and will vary from cell type to cell type within an organism. The proteome that is expressed within an organism will be a composition of the genetic make up of the organism and the environmental factors that the organism is exposed to. The proteome within an organism can affect physical characteristics and their overall health and robustness. Throughout the remainder of the course, we will spend a lot of time exploring the control of gene expression and protein function.