Fundamentals of Cell Biology

FUNDAMENTALS OF CELL BIOLOGY

LAUREN DALTON AND ROBIN YOUNG

Oregon State University

Corvallis, OR

CONTENTS

ACKNOWLEDGMENTS

This book has existed in some form or another since roughly 2004. It was the brainchild of two pioneering instructors at the Vancouver campus of the University of British Columbia (UBC)—James (Jim) Berger and Ellen Rosenberg, who believed strongly that education should be free and accessible to all. Since then, their original online text has been through many iterations, adapting to the changing body of scientific knowledge and the changing needs of students. In 2010, Robin Young took over curating this material as part of the UBC Vancouver foundational cell biology course, BIOL200. In 2019, the coauthors of this book decided that this work needed to be turned into a more formalized Open Educational Resource so that it would be more widely accessible and then worked to make that happen.

A number of excellent and committed academics have contributed to this work over the past 20 years, including (in no particular order) Nelly Panté, Ljerka Kunst, Ninan Abraham, A. Lacey Samuels, Alicia Mazari, Ken Savage, Sunita Chowrira, Marcia Graves, Megan Barker, Liane Chen, Vivienne Lam, Greg Doheny, and many more. Several researchers at UBC and Oregon State University (OSU) have contributed directly to the content in this newest iteration of the textbook, including Kari van Zee and Rick Cooley, who helped with the addition of the GCE technique, and Caity Smyth, who consulted on the ChIP section. Additionally, feedback from the many teaching assistants, undergraduate peer tutors, and students in UBC's BIOL200 and OSU's BB 314 were essential for us.

We would like to thank our funding sources, the Affordable Learning Grant at Oregon State University and the ALT-2040 Fund at the University of British Columbia's Okanagan campus. This support gave us the resources to have reduced teaching loads, allowed us to hire our illustrator (Heather Ng-Cornish) and our other student worker (Devin Collins), and gave us access to a variety of other support resources that were invaluable in creating this textbook.

This work wouldn't be possible without the amazing support behind the scenes helping to make this book a reality. In particular, the entire team at OSU Open Education, specifically Stefanie Buck, Mark Lane, and our animations team, were invaluable as we navigated this project. We also appreciate the tireless work of our illustrator, Heather Ng-Cornish, who has created a ton of fantastic images. Without her, this project wouldn't have gone anywhere. A note of special thanks also to our student worker, Devin Collins, who helped us stay organized!

Many excellent microscopists graciously shared their work with us to include in this book, including Drs. Yoshi Watanabe and Davis Iritani from the Core Microscopy Facility of Agriculture and Agri-Foods Canada's Summerland Research and Development Centre as well as graduate students Kyle Nguyen from OSU and Shawn Shortill and Lucia Queseda-Ramirez from the Vancouver Campus of UBC. Even more excellent microscopists were open and willing to share their work, even though it didn't all make it into the final cut. We are grateful to them all.

This textbook includes the work of many folks who have graciously offered their work to the world through open creative commons licenses. We value your commitment to sharing and the principles of the creative commons and open licensing. We are also grateful for the feedback from our reviewers, who carefully read over our work and weren't afraid to give us an honest opinion when needed. Finally, we'd like to thank everyone who has had to listen to us go on at length about this textbook,

including our friends and families, as we worked to figure out how to best support the learning of our students. Your insights and support were key to this entire process. We couldn't have done this without any of you.

We dedicate this work to past, present, and future cell biology students. We hope that you find something in this work that excites you and that wherever you go from here, you have gained a tiny appreciation for the beautiful complexity of the cells that make up the living world.

INTRODUCTION AND REVIEW

This textbook was written for two second-year introductory cell biology courses (one at the Okanagan campus of the University of British Columbia, in Canada, and one at Oregon State University, in the USA). Both of these courses have first-year biology and first-year chemistry as their prerequisites. Thus, this textbook is written in a way that assumes preexisting university-level knowledge of many fundamental concepts that are commonly covered in first-year biology and chemistry.

Depending on how you are using the information in this textbook, you may have a different background than what we have assumed in our content. Alternatively, even if you do have the correct background, it could have been a while since you learned this material, and so it might not all be fresh in your mind.

In this section of the book, we have collected a number of videos that discuss specific topics in ways that we feel will help prepare you for what's to come in the rest of the book. This section is in no way comprehensive or exhaustive but serves instead as a starting place for you from which you can go and explore more if you feel you need more information.

We have chosen to link to videos, for the most part, as discussions with our students tell us that videos are often the first place they look over written formats.

CHEMISTRY REVIEW VIDEOS

- The chemistry of water: https://youtu.be/5EZw-I65Q0A
- Chemical bonding: https://h5p.org/h5p/embed/267321
- Polar and nonpolar molecules: https://youtu.be/PVL24HAesnc
- Intermolecular forces video: https://h5p.org/h5p/embed/268232
- Basic principles of thermodynamics: https://www.youtube.com/watch?v=8N1BxHgsoOw
    - Entropy and life: https://youtu.be/a_BKQ_ZPImw
    - Thermodynamics and energy diagrams: https://youtu.be/Ykhn2psFmEM
- Basic of acids/base chemistry: https://youtu.be/mnbS56HQbaU
    - More advanced concepts: Acids and pKas in the context of organic chemistry: https://www.youtube.com/watch?v=BLKqbC_QIZA

BIOLOGY REVIEW VIDEOS

- The basic structures of cells: https://youtu.be/8IlzKri08kk
- Macromolecules
    - Amino acids and proteins: https://youtu.be/wvTv8TqWC48
    - Lipids: https://youtu.be/5BBYBRWzsLA

    - Nucleic acids overview: https://youtu.be/MA-ouz1LtpM
        - Chemistry of DNA: https://youtu.be/o_-6JXLYS-k
        - DNA versus RNA: https://youtu.be/Wdt39RLmE_s
    - Carbohydrates: https://youtu.be/LeOUIXbFyqk
- Transcription, translation, and protein folding overview: https://youtu.be/itsb2SqR-R0
    - Transcription: https://youtu.be/SMtWvDbfHLo
    - Translation: https://youtu.be/TfYf_rPWUdY
- Replication: https://youtu.be/I9ArIJWYZHI
- The cell cycle and mitosis (basics only): https://youtu.be/8uzHTKdv_Sw

We are always looking for new, high-quality material to include in this list that is appropriately licensed to allow us to include it. If you have created a resource that you'd like us to consider, please feel free to let us know.

CHAPTER 1.

VISUALIZING CELLS THROUGH MICROSCOPY

INTRODUCTION

Before we can even begin to discuss cells, it's important to understand that cells are very tiny! Most cells are too small to see with the naked eye. Thus, the beginning of our book must start with a conversation about microscopes and how they work. An understanding of the tools we use to study cells, both their abilities and their limitations, is essential to making sense of the material you're going to be exploring throughout this book. These days, thanks to some pretty impressive advances in computing in the last 50 years or so, there are a great many different types of microscopes that exist. We can't possibly cover all of the different kinds with the time that we have to cover this topic. However, the essence of all of the varied types of microscopes that are used to explore cell biology fit into four major groupings:

1. **Brightfield light microscopy**
2. **Fluorescene light microscopy**
3. **Transmission electron microscopy (TEM)**
4. **Scanning electron microscopy (SEM)**

The capacity to visualize samples in any kind of microscope depends on the particle that is used by that microscope. The first and oldest microscopes were developed about 400 years ago and used visible light from the sun, with glass lenses to magnify the sample. About 80–100 years ago, we discovered that using electrons to image samples greatly increased both the magnification and resolution of the microscope. These days, both light and electron microscopy are still in use. We will look at each type of microscopy in turn and explore their strengths and limitations. An example of each is included in Figure 01-01 for direct comparison in the same organism. We will also show more examples throughout the chapter. Consider how we might identify each type of microscopy and distinguish when we could use one type over the other. However, before we do any of that, we must first discuss the concepts of **magnification** and **resolution**, since all microscopes use both to help us see very small objects, like cells.

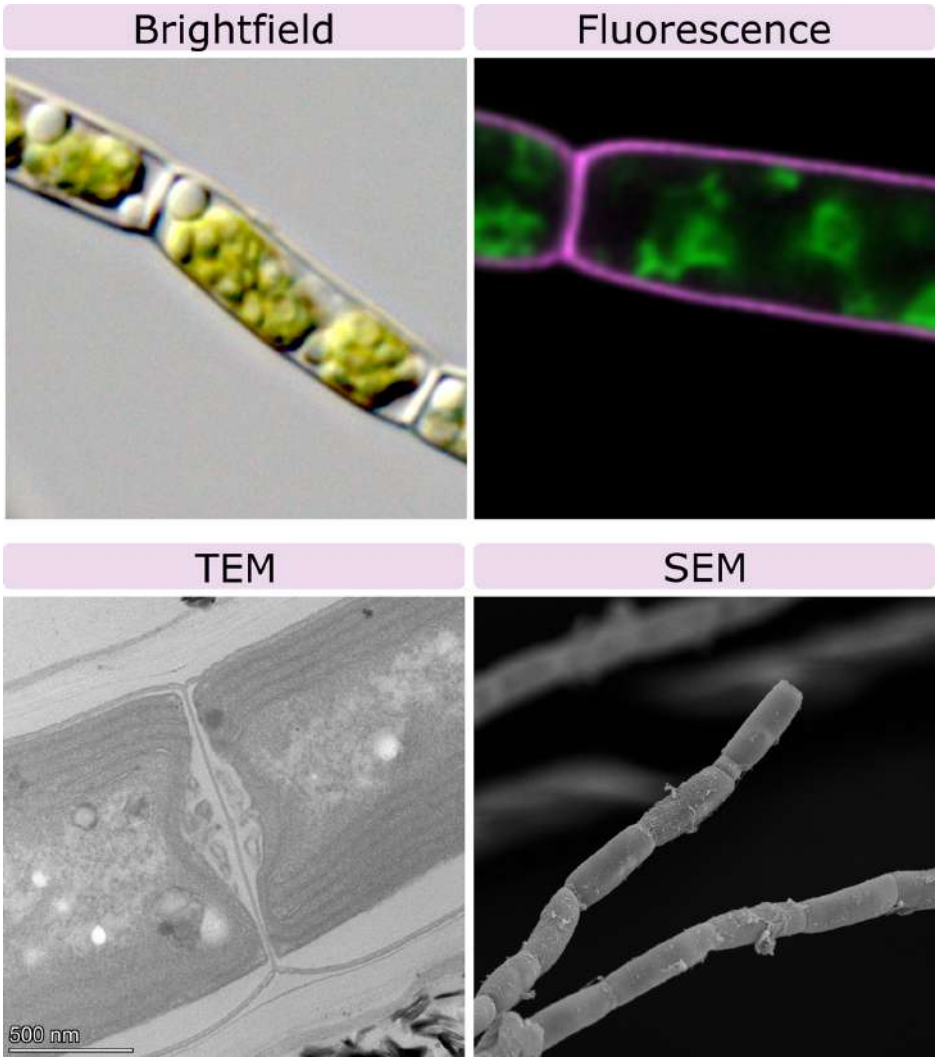


***Figure 01-01**: Examples of the four different types of microscopy, imaging green algae cells (species unknown): brightfield light microscopy, fluorescence light microscopy, transmission electron microscopy (**TEM**), and scanning electron microscopy (**SEM**). An average algal cell is between 2 and 7 µm. All images were collected by Dr. Davis Iritani, Multi-functional Microscopy Technician from the Summerland Research and Development Centre in Kelowna, British Columbia, Canada, and used with permission.*

TOPIC 1.1: MAGNIFICATION AND RESOLUTION

Learning Goals

- Define magnification and resolution.
- Explain how magnification and resolution work together to help us see cellular structures more clearly.

Before we take a deeper dive into the topic of microscopy, it may be useful to revisit the physics behind how microscopes work. It's important to understand the basic principles of the wave/particle properties of visible light, which is one small section of the electromagnetic spectrum. Electrons also have both wave and particle-like properties, so many of the principles that underlie electron microscopy are similar to those that underlie light microscopy. Fluorescence microscopy and scanning electron microscopy (SEM) are derivatives of the other two, in which the source of the particle and/or the way the particle is detected is changed so that a different image is created.

All microscopes have limitations on the size and detail with which they can be used to visualize objects. These limitations are a product of the specific characteristics of the particle being used to create the image (i.e., photons or electrons). The wavelength at which these particles travel has a direct impact on how much detail can be seen in the image generated. When discussing the limitations of microscopes, we generally discuss them in terms of magnification and resolution.

The Merriam-Webster Dictionary defines **magnification** as "the apparent enlargement of an object by an optical instrument." In other words, a microscope allows us to make very tiny things look bigger than they are. The objectives on a standard classroom microscope often tell you their magnifying power so that you have some context for the actual size of the things that you can see using the lens.

On the other hand, **resolution** is defined as the closest spacing of two lines that can be distinguished as separate entities. The resolution of any microscope is limited by the wavelength of the particle being used to illuminate it. In the case of the light microscope, the resolution (d) is

$$\mathrm{d} = 0.61 * \text{wavelength}/\text{numerical aperture of the objective.}$$

If we consider that the shortest wavelength of light that we can see (violet) is about 400 nm, then our calculation for the resolution of a standard light microscope is as follows:

A standard 63× magnification oil-immersion lens (often the highest magnification available in those classroom microscopes) has a numerical aperture of 1.4. So based on this, the resolution of a standard classroom light microscope with that objective would be 0.61*400 nm/1.4, which equals about 175 nm, or roughly 0.2 μm.

This is pretty small. At that range, we can probably see bacteria and maybe even some large viruses (but not all and definitely not well). We also can see some of the larger subcellular organelles in a cell (like the nucleus and mitochondria), but not much of the smaller organelles. Also, any structure that we can see that's close to the limit of resolution of the light microscope won't have much detail visible. Much like if you see someone who is very far away, you may only be able to tell that they are a human, but not if it's someone that you know.

Electrons, on the other hand, have a much shorter wavelength than photons of the visible light spectrum…roughly 200,000 times smaller! Because of that, the resolution limit of an electron microscope is much smaller…there are some kinds of electron microscopes that can view individual atoms! More commonly, in the more widely used electron microscopes, the resolution limit is roughly 0.1 nm. A single double helix of DNA is 2 nm in diameter, and the thickness of the average lipid bilayer is between 4 and 10 nm. We can see both of these in an electron microscope. As such, we can see that this resolution limit is small enough that a lot of the finer structural details inside the cell become visible.

Figure 01-02 gives a rough estimate of what we can see using light or electron microscopy. The website Size of the Universe is also worth exploring. It shows examples that range from the largest thing known (the entire observed universe, or ~10^{27}m) down to the smallest theoretical measurable length in the universe (Planck's length, or 10^{-35}m).

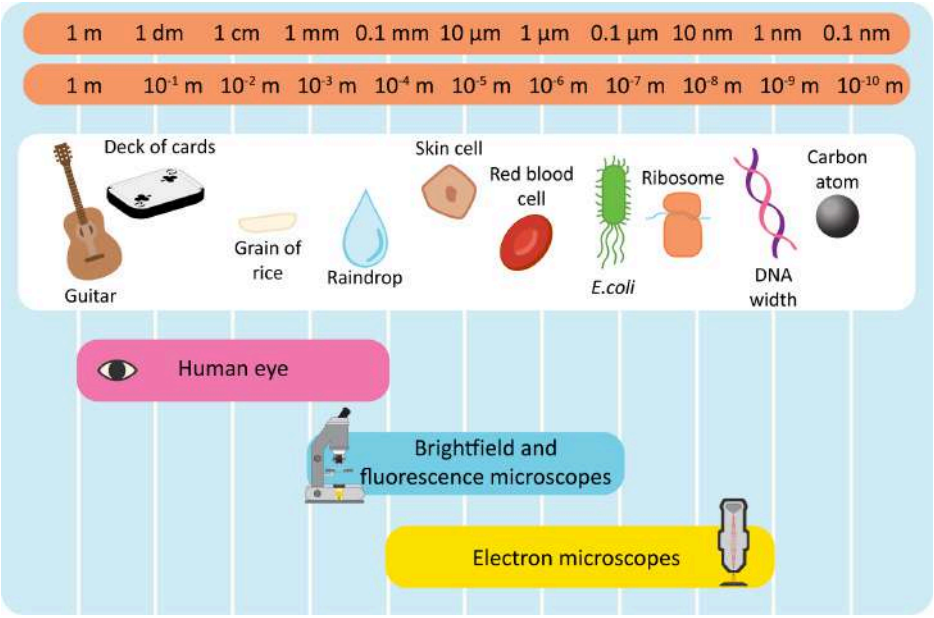


Figure 01-02: Schematic diagram comparing methods to visualize different-sized objects. This image was created by Heather Ng-Cornish and is licensed under a CC BY-SA 4.0 license.

It's important to remember that resolution is a quality that is separate from magnification. The magnification of an image is simply how much we have enlarged the object in order to observe it. For example, if you wanted to blow up a photo to hang on your wall, you would have to increase the magnification in order to do so. If there aren't enough pixels in your image to support the amount that you're magnifying it, then at some point the quality of the image will decrease as you increase the magnification. You can't actually see more detail; you just see the pixels more obviously. The number of pixels the image contains will determine how much detail you can see. Thus, the number of pixels dictates the **resolution** of the image that you can create, especially as you blow the image up larger and larger. Applying this to microscopy, in order to see small structures, you not only have to have a higher magnification in order to zoom in, but *you also need higher resolution* in order to see the details of the smaller structures. Figure 01-03 shows a comparison of microscopy at the same magnification with different resolution to highlight why increased resolution is required with an increase in magnification.

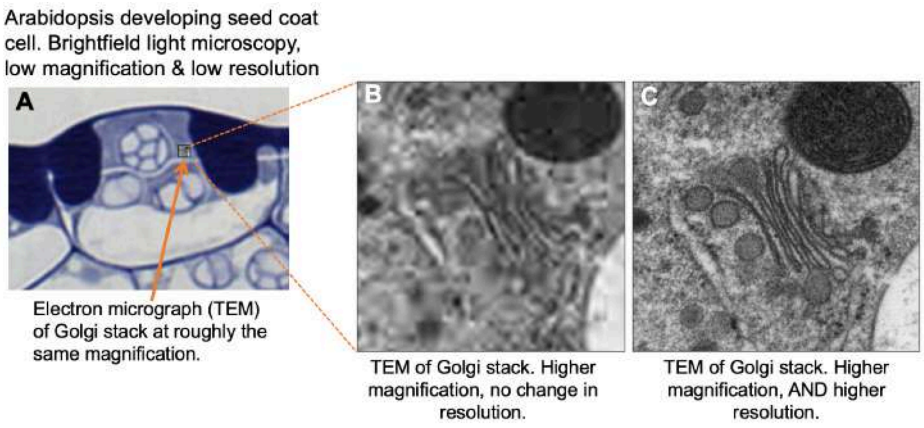


Figure 01-03: Comparison of magnification and resolution on image quality. As a sample is imaged at higher magnifications, we must also increase the resolution so that we can see the finer details of our sample. In all cases, the image shows a seed coat cell of a developing seed of the model plant, Arabidopsis. Image in (A) is a light microscopy image, magnified roughly 1,000×. Samples in (B) and (C) are transmission electron microscopy (TEM) images magnified roughly 24,000×, but they do not have the same resolution. In (B), the resolution was modified to match the resolution of the same region shown in image (A), whereas in (B), the increased resolution of the TEM was left. Microscopy and image creation by Dr. Robin Young and licensed under a CC BY-SA 4.0 license.

With microscopes, as with the photos that you take with your phone, it is important that the resolution increases along with the magnification in order to see the details desired in the image. The kinds of microscopy that we will explore have different resolution limits. Thus, the resolution that is required in order to observe the structure that you're interested in will be a major factor when considering the type of microscopy to use in any given situation.

We will leave the final word on this to the following video, created specifically to help you think about the differences between magnification and resolution and also consider how to interpret images when the microscope requires us to cut the sample into thin sections.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=143#oembed-1

Video 01-01: Introduction to Microscopy, Magnification and Resolution. Created with a UBC Teaching and Learning Enhancement Fund grant to Drs. Sunita Chowrira and Steven Barnes. This video is included with permission.

TOPIC 1.2: LIGHT MICROSCOPY

Learning Goals

- Explain the basics of how a light microscope works, with a particular focus on how we can use that information to interpret micrographs.
- Compare and contrast the different types of light microscopy (i.e., brightfield versus fluorescence) and identify when each type would be useful.
- Identify the major advantages and disadvantages of light microscopy.

The very first microscope was made in the mid-1600s by a British man named Robert Hooke. That microscope had a mirror to direct sunlight into the glass lenses of the microscope so that the sample could be viewed. Microscopes stayed more or less the same for the next 300 years, until the Industrial Revolution of the early 1900s. It is at this point that the advances in technology allowed us to greatly improve how microscopes work. Now there are light microscopes that use various combinations of lasers, lenses, spinning discs, and computer algorithms to push beyond what should be possible given the physical limitations of visible light. We've also made significant advances in sample preparation, which have made additional contributions to what we can visualize. Even though we have incredibly advanced microscopy techniques with which to view cells, their essence remains the same. We will explore the two largest categories of light microscopes: those that collect transmitted light, originating from some kind of light source, to view the sample and those that collect light that is emitted by the sample itself.

Pros to Light Microscopy

Live cells can be viewed and recorded in real time. Specific structures can be labeled for viewing, thus reducing the "noise" in the sample and showing finer details that may otherwise be masked by other components of the sample.

Cons to Light Microscopy

The limit of conventional (brightfield and confocal microscopy) resolution is about 0.2 μm, which restricts the amount of detail we can see of internal cellular structures using light microscopy.

- There is a newer technique, known as **superresolution microscopy**, which is technically able to break this barrier. We will discuss this in more detail later.

Transmitted versus Emitted Light

Light microscopy can be further divided into several subtypes of light microscopy. The most notable of these subdivisions takes into account the types of light that you are viewing when you look at an image (Figure 01-04).

- Transmitted light microscopy (Figure 01-04A) is the traditional light microscopy that you most likely tried in high school biology classes. In this type of microscopy, the sample sits between the light source and the eyepiece, and the light you see is the light that was able to pass through the sample. The simplest version, which is the one we will discuss, is called **brightfield light microscopy**.
- In **emitted light microscopy** (Figure 01-04B), the sample is illuminated by a light source that is off to the side, and then the molecules in the sample get excited by this light and release their own photons. Thus, the light that you view in these kinds of microscopes comes from the samples themselves. This type of microscopy is also known as **fluorescence light microscopy**, which is what we will be calling it after this.

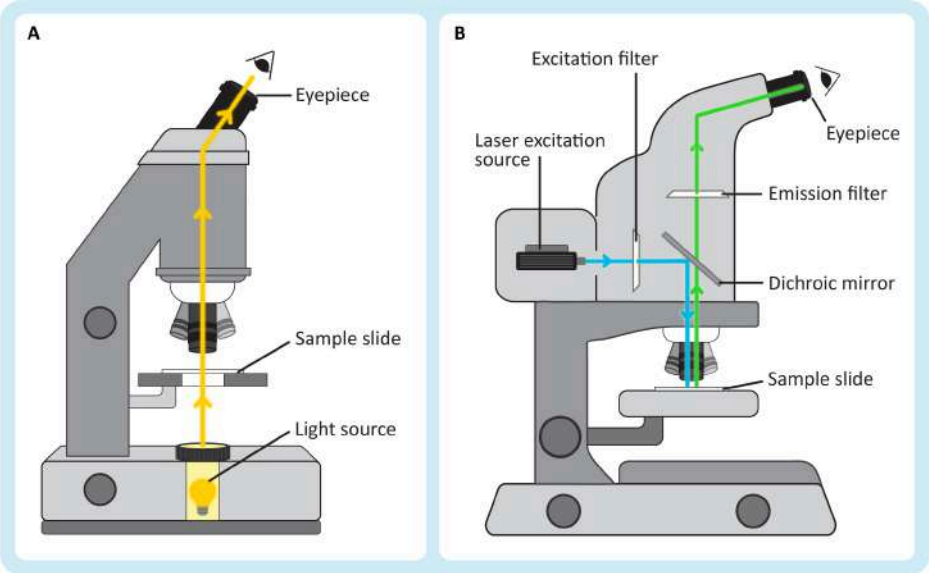


Figure 01-04: Examples of how light is directed through transmitted and emitted light microscopes. (A) A standard desktop brightfield light microscope. Light is transmitted from the light source (at the bottom), through various lenses to condense the light, and then passed through the sample to the eye via the eyepiece at the top. The light you see is the light that was transmitted by the light source but not absorbed by the sample. (B) A standard desktop fluorescence microscope. In this microscope, the light source is off to one side (labeled as the laser excitation source in the image). This light source is either light in a specific color range or, more commonly, a laser of a precise wavelength. In the middle is a special mirror that allows some wavelengths of light to pass but not others (known as a dichroic mirror). The light from the laser hits the mirror and is reflected toward the sample. The photons in the laser light excite photo-activated particles in the sample. The photo-activated particles in the sample will then emit photons of a second wavelength. This light leaves the sample and is able to pass through the dichroic mirror so that it can travel up the objective lens and into the eye of the observer. This image was created by Heather Ng-Cornish and is licensed under a CC BY-SA 4.0 license.

Brightfield and Other Forms of Transmitted Light Microscopy

One of the primary challenges of microscopy is that living cells are fairly transparent. Microscopists use a number of techniques to increase the contrast of their samples in order to see things more clearly. Sometimes we add chemical stains to the sample, such as toluidine blue or hematoxylin and eosin (H&E), which add color contrast (see Figure 01-05). We can also use optical techniques, in which

certain light is included or excluded from viewing by the microscope, to increase the contrast of our samples. An example of an optical technique used in our daily lives is polarized sunglasses. The polarized glasses reduce glare by only allowing light to pass through if it is oriented in a specific way. We can do similar kinds of things with microscopes. Some of the subtypes of transmitted light microscopy are listed below:

- **Brightfield microscopy**: If you have ever used a light microscope in school, this is most likely the kind that you used. It was the first invented, way back in the 1600s, and uses a light source as simple as sunlight. Images from this kind of microscopy usually have a white background, and the sample is expected to be in color when we look through the eyepiece of the microscope. Figure 01-05 shows a few different samples that have been stained and observed using brightfield light microscopy. *This is the type that we focus on in this textbook when discussing transmitted light microscopy.*
- **Phase-contrast microscopy**: In this type of microscopy, shifts in the amplitude and phase of the light are converted into shifts in brightness so that edges and other structures become more visible. Differential interference contrast (DIC) is a variation on this.
- **Darkfield microscopy**: In darkfield microscopy, light is excluded by blocking the center of the beam but not the outer part. This kind of light microscopy has a dark background (unlike other kinds of light microscopy), which again helps the edges of cells and other structures in the sample to stand out.
- **Polarized light microscopy**: This type of microscopy enhances the contrast of specimen by shining light of a particular orientation that causes shadows on structures differentially depending on their composition. This kind of microscopy is especially useful for samples like bone, fibers, or mineral deposits.

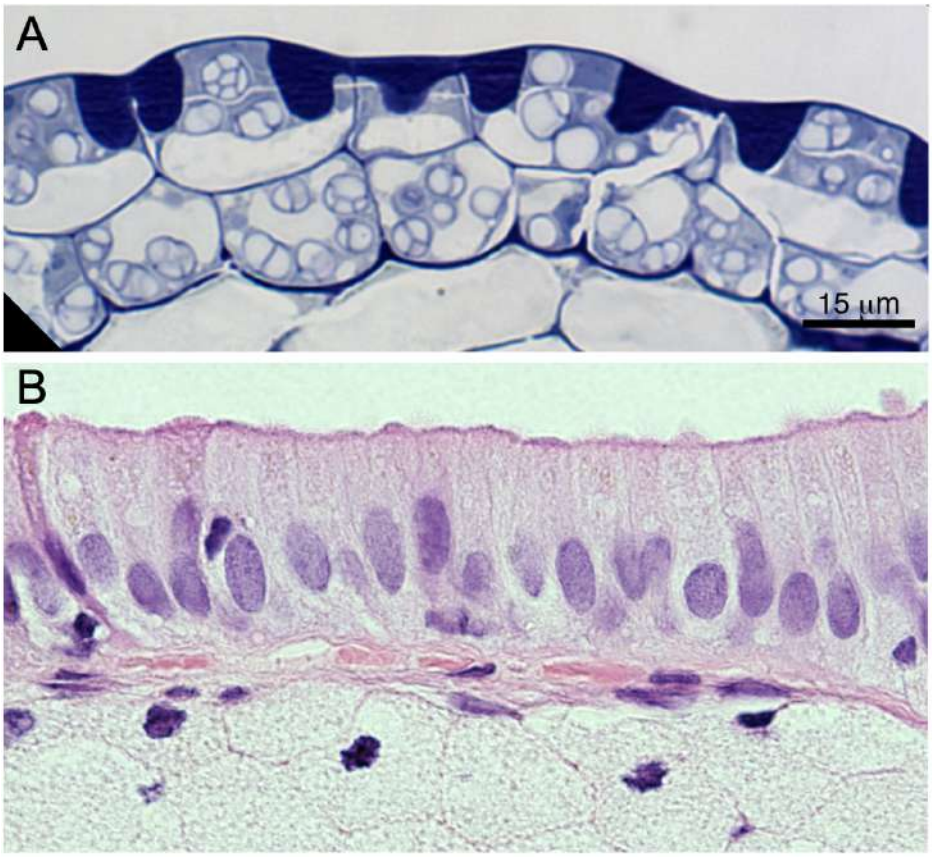


Figure 01-05: Brightfield microscopy in plants and animals. These images are at roughly the same magnification. (A) Developing seed coat of the model plant Arabidopsis thaliana stained with toluidine blue. Microscopy was done by Dr. Robin Young and is licensed under a CC BY-SA 4.0 license. (B) Epithelial cells of a human gall bladder, stained with hematoxylin and eosin (H&E) Attribution: William Karkow (2011), CIL:34859, https://doi.org/10.7295/W9CIL34859. This image is in the public domain.

Fluorescence Light Microscopy

Fluorescence microscopy was first developed in the 1980s and has continued to revolutionize cell biology to this day. Not only does it allow us to view live samples, as other forms of light microscopy do, but it also allows us to label (or tag) specific macromolecules / cell structure so that we can track them within the cell (Figure 01-04B). We tag the structures using fluorescent molecules (called fluorochromes or fluorophores). Then, using our fluorescent microscope, we illuminate the sample with light of a particular wavelength, and the fluorophore responds by emitting light at a second known wavelength that we can then detect. The great advantage of this technique is that only the molecule or structure of interest shows up in the image, and the rest of the sample, which is not emitting light, is dark.

Much like transmitted light microscopy, there are several subtypes of fluorescent light microscopy. Most of these have been developed to increase the resolution of the images that are being viewed. The simplest type of fluorescence microscopy is known as **epifluorescence**, which uses powerful halogen lightbulbs and colored filters to produce light of specific wavelengths. **Confocal laser scanning microscopy** (or confocal for short) is very similar to epifluorescence, except a laser is used as the source of light. Using a laser allows us to focus the light very specifically and then use computer algorithms to remove out-of-focus light. There are also a number of extremely advanced techniques

that allow us to seemingly "break" the laws of physics. Using a combination of computer algorithms and specific image collection protocols, microscopists have been able to resolve structures that should be much too small to see in light microscopy. These techniques are collectively known as **superresolution microscopy**.

Unlike the different kinds of transmitted light microscopy, the images that you create from the different subtypes of fluorescence light microscopy are very similar. It's not always easy to tell what kind of fluorescence microscope was used to create the image just by looking at it. For this reason, we will not differentiate between the different types in this textbook. Instead, we will treat the images that are created using fluorescence microscopy as a single group. Examples of different fluorescence microscopy images are shown in Figure 01-06.

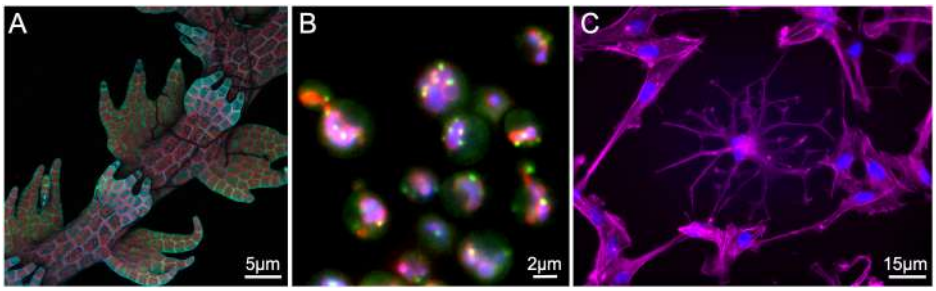


Figure 01-06: Three different fluorescence micrographs. Live imaging of (A) Lepidozia reptans, a small liverwort collected in the Pacific Northwest. This sample is unstained and is imaged using a special fluorescence technique known as two-photon imaging, which amplifies intrinsic fluorescence within the sample. Image collected by Dr. Robin Young. (B) Baker's yeast (Saccharomyces cerevisiae) showing a vacuolar hydrolase (magenta) overlapping with Rab5 (green) signal at endosomes. The yeast vacuole was stained in blue. Image collected by Shawn Shortill, a PhD student at the University of British Columbia in Vancouver, British Columbia, Canada. (C) Human astrocytes showing a nucleus stained with blue and the actin cytoskeleton labeled in red. Image collected by Kyle Nguyen, a graduate student at Oregon State University in Corvallis, Oregon, USA. This image is shared under a CC BY-SA 4.0 license.

Sample Preparation for Fluorescence Microscopy

There are several approaches to sample preparation in fluorescence microscopy. It depends on both the scientific question you're trying to answer and the fluorescent stains you have available. There are a number of chemical stains that can be added to live cells that latch onto specific components of the cell and then fluoresce. This is known as **direct staining**. An example of this is DAPI, which fluoresces when bound to double-stranded DNA. This is great when you're interested in exploring a sample for which a fluorescent live-cell stain exists. Sadly, that is not always the case.

Another option is **immunolabeling**. In this case, antibodies bound to a fluorophore that have been specifically created to bind to your protein or structure of interest are used to label the cell. If your structure/molecule of interest is on the cell surface, then this works just fine. However, if your structure is in the interior of the cell, then the antibody will have difficulty accessing it. In this case, the cell is usually treated with fixatives, which kill it but also hold everything in place. Then the membrane can be disrupted just enough to allow the antibodies to access the interior of the cell.

One of the most important advances in fluorescence microscopy was the development of genetic engineering protocols so that one could add a fluorescent tag to any protein in the cell. There are a number of naturally occurring proteins that fluoresce and are used by organisms to produce

bioluminescence. Some of the fluorescent proteins have been isolated and attached to other, nonfluorescent proteins. The most commonly used tag is known as **green fluorescent protein (GFP)**, which was isolated from bioluminescent jellyfish. As its name states, GFP is a protein that fluoresces in the green range of visible light (~550 nm) when it is illuminated with blue light (~450 nm). The genetically engineered protein usually functions like the normal protein, except it now fluoresces its location under a particular wavelength of light. Thus, we can now track structures that include the fluorescent protein.

GFP was first isolated in the 1980s, but since then, we have genetically modified it to create versions in virtually every wavelength of the visible spectrum. We've even been able to produce forms that are split in half and only fluoresce when the two halves come together so as to identify when two proteins of interest interact with each other.

TOPIC 1.3: ELECTRON MICROSCOPY

Learning Goals

- Explain the basics of how the two major types of electron microscopes work, with a particular focus on how we can use that information to interpret micrographs.
- Compare and contrast the different types of electron microscopy (i.e., transmission versus scanning) and identify when each type would be useful.
- Identify the major advantages and disadvantages of the different kinds of electron microscopy.

Electron microscopy works in a very similar way to light microscopy, except that it uses a beam of electrons instead of light to image the sample. Since electrons can easily be scattered by air molecules (which would destroy our ability to use them for imaging), almost all electron microscopy is done in a vacuum, including both **transmission electron microscopy (TEM)** and **scanning electron microscopy (SEM)**. The "electron gun" is the source of electrons. Electricity is sent through a metal filament (often made of tungsten) until it releases electrons into the vacuum of the microscope. Electromagnetic lenses (similar in function to the optical lenses of the light microscope) are used to control and focus the electron beam so that the sample can be imaged.

Again, like both brightfield and fluorescence microscopy, there are several types of electron microscopy that exist—more than we can reasonably discuss here. Here we focus on the two types that are most commonly used to study biological materials.

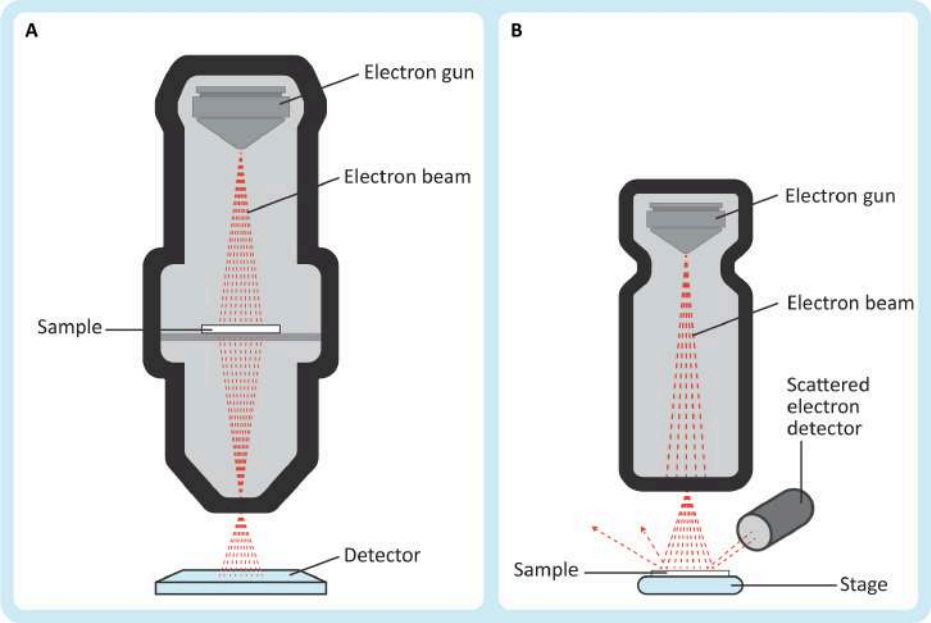


Figure 01-07: The path of electrons through different electron microscopes. (A) Transmission electron microscopy. (B) Scanning electron microscope. In each case, the electron "gun" is the source of electrons. The detector is used to capture the electrons that will be used to create the image, which will be viewed on a computer screen attached to the microscope. This image was created by Heather Ng-Cornish and is licensed under a CC BY-SA 4.0 license.

Resolution

Electrons have shorter wavelengths than photons in the visible light spectrum. For the electron microscope, an accelerating voltage of 100 kV produces a beam wavelength of close to 0.004 nm. The numerical aperture of the instrument is close to 0.012, so take a minute and calculate the resolution of the electron microscope using the equation from earlier in this chapter. (*Hint: It's 0.2 nm.*) That's a 1,000× smaller resolution than we calculated for the light microscopes.

Pros and Cons to Electron Microscopy

Pros	*Cons*
• TEM and SEM have good resolution at a size range that is important for cell studies (200 nm to 0.2 nm); size from organelles to macromolecules.	• Samples must be able to withstand a vacuum and electron bombardment, so elaborate specimen preparation is required. This can take quite a long time to do well, depending on your samples and your protocol. Some protocols are weeks or even months long, while others can be completed in an afternoon. • Also, since organic material is made primarily of carbon (an atom that has a lot of empty space in it), heavy metal stains must be used to increase the contrast in the samples so that the cells can be viewed. These heavy metal stains tend to be quite toxic and must be handled carefully.

the limits of sem and tem

Both SEM and TEM have limits on what can be viewed in the microscope, which will limit its usefulness. We will look at each type a little bit more in the next section.

- Thin sections imaged by TEM can have quite a lot of detail in them, but since they are very thin slices of a 3D object, they are missing an entire dimension. Trying to imagine what a 3D object looks like from a 2D image can be challenging. There are types of electron microscopy that allow us to digitally reconstruct 3D images from several serial sections of a sample. This is known as **electron tomography**, but unfortunately, how it works is beyond the scope of this textbook.
- In SEM, the primary focus is the surface of the sample. This may mean the exterior of the organism (if it's small enough) or even the cell, but it doesn't have to. Samples can be frozen and then cracked open and prepared to view the interior surfaces, but anything that isn't very close to the exposed surface will not be easily viewed.

Each of these types of electron microscopy functions quite differently, so we must also look at each one in more detail as well.

Transmission Electron Microscopy (TEM)

In many ways, TEM is analogous to traditional brightfield light microscopy, but with much greater capabilities with respect to magnification and resolution. The sample sits between the electron source and the detector (Figure 01-08A) so that the beam of electrons passes through the sample on its way to the detector. Just like in brightfield light micrographs, the empty spaces will show up as white in the image. Unlike light microscopy, there is no color in electron micrographs. Instead, the areas where the electrons were absorbed by the sample will be gray or black (depending on how "electron dense" your sample is in that region).

In TEM, the electron beam passes through the specimen. Electrons cannot penetrate very deeply, so the specimen must be extremely thin. The specimens are typically 50–100 nm sections of cells that have first been stained with heavy metals and then embedded in plastic resin. As you can imagine, all of this sample processing, along with the fact that the inside of a TEM is a vacuum, means that TEM cannot be used with living material. Instead, samples are carefully preserved, prior to embedding in epoxy resins, by either flash freezing (also known as **cryofixation**) or using chemical fixatives. If we do this carefully enough, we can capture cellular events that were happening at the moment of fixation. Being gentle with samples so as to avoid disturbing the location of each and every molecule while also removing water and replacing it with epoxy resin is very difficult, which is why this process can take days, weeks, or even months.

The electrons that have passed through the specimen are collected by an electron detector attached to a digital camera. An example is shown in Figure 01-08, where some of the subcellular structures in a plant cell have been labeled for you.

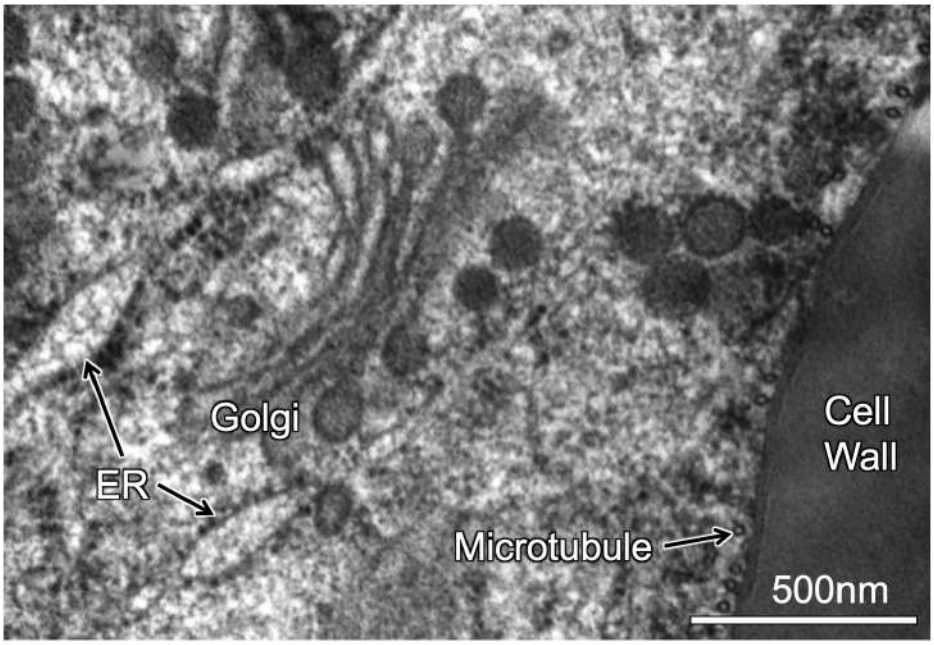


Figure 01-08: TEM micrograph showing a 70,000× magnified portion of the cytoplasm of a plant cell. The ER, a Golgi stack, and some vesicles are visible. The structures are near the edge of the cell, as a cell wall is visible on the right-hand side of the image. Microtubules can be viewed lined up in cross section, just underneath the plasma membrane. Image created by Dr. Robin Young, shared under a CC BY-SA 4.0 license.

Much like in light microscopy, samples must be stained so that the structures are visible. As mentioned, we use heavy metal stains for electron microscopy. This is because of how electrons are able to move in heavy metal atoms. Heavy metals like lead (Pb), uranium (U), osmium (Os), gold (Au), and silver (Ag) are considered to be more "electron dense" and thus will be better able to block the electron beam than the atoms biological samples are made of, primarily carbon (C), oxygen (O), and hydrogen (H). We are also able to use immunolabeling in TEM in order to identify the location of specific proteins. However, we use colloidal gold particles of a specific size, rather than any kind of fluorescent tag, that are attached to antibodies that recognize our structure or protein of interest.

Negative staining is another technique that is most commonly used to examine very small objects, such as bacteria, viruses, and even individual proteins. In this technique the sample is suspended in an electron-dense stain that does not penetrate the sample but rather surrounds it so that it is outlined and all of the cracks and crevices can be observed.

Scanning Electron Microscopy (SEM)

In SEM, the specimen is usually dried and coated with a very thin layer of gold, platinum, or another heavy metal. The electron beam is "scanned" across the surface of the specimen. As the beam of electrons hits the specimen, secondary electrons are scattered from the surface of the specimen (Figure 01-08B). These are collected by a detector that builds an image electronically based on

electron intensity (from white to black). This technique produces images of surfaces only. Since this microscope also uses electrons to image, the magnification and resolution are similar to TEM.

While samples are usually dead, when imaged with SEM, they do not have to be embedded in plastic like they are for TEM. Panels B and C of Figure 01-09 are of material that was cryofixed and imaged in the frozen state.

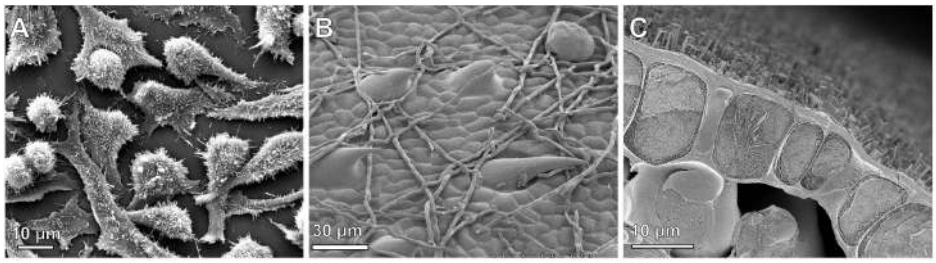


Figure 01-09: Three different SEM micrographs. (A) Cells derived from the Henrietta Lacks (HeLa) cell line on an unknown substrate. Attribution: Anna Baccei and Marian Rice (2011), CIL:12597. https://doi.org/10.7295/W9CIL12597. CC BY-3.0 license. (B) Powdery mildew filaments growing on the surface of cannabis. Image taken by Lucia Queseda-Ramirez, graduate student at UBC Vancouver. Used with permission. (C). Cross section of the leaf of Arabidopsis thaliana showing the wax crystals of the cuticle on the leaf surface. Image taken by A. Lacey Samuels, UBC Vancouver. Used with permission.

TOPIC 1.4: PUTTING IT INTO PRACTICE: DETERMINING THE TYPE OF MICROSCOPY FROM THE MICROGRAPH

Learning Goals

- Interpret the results in experiments using microscopy based on scale, magnification, resolution, and plane of section.
- Identify the type of microscopy that is best suited to detect and study cellular components based on their size and the functional aspect being studied.

One of the major skills you should be practicing as you learn about cell biology is how to interpret micrographs. This is not a trivial skill to learn. However, the very first step in this process is being able to identify the type of microscopy that was used based solely on the image.

When you look at a micrograph of any kind, there are a number of questions that you should be asking yourself, including the following:

- What is the magnification/scale of the image? Can I find a scale bar or other information to help me tell?
- Does the image look three-dimensional? Is there depth to the image?
- Is there a black background? Does the sample look like only some of the structures in the

sample are visible?

- What types of structures do I think I can see? Does that match with the magnification/ resolution I would expect based on the type of microscopy?
- Is it in color? Does the color look like it might be computer-generated after the fact?
- Is it moving? Does it look like it could have been imaged live?
- Finally, what other information do I have about this image? Is there a figure legend or other text associated with the image that provides insight? It is quite common to place additional information in the text associated with the image, as we have done in this textbook chapter. Any information that you have about an image is there for you to use, so don't forget to use it!

There is no single question in this list that can determine definitively what kind of microscopy was used. However, a combination of answers could help you make a reasonable decision, or at least a solid guess, based on the evidence. And every time you look at a new image, you're going to have to figure it out again.

Light versus Electron Microscopy and the Question of Scale

While the size of what you're looking at can be an important indicator of what type of microscopy you're looking at, it's also important to realize that more than one type of microscopy can often be used at any given magnification. We have repeated Figure 01-02 below because it does such a good job of visualizing the overlap and also to facilitate the discussion we're about to have. It's important to know what you can reasonably see in each type of microscopy so that you can make appropriate predictions about the information within the micrographs that you're interpreting.

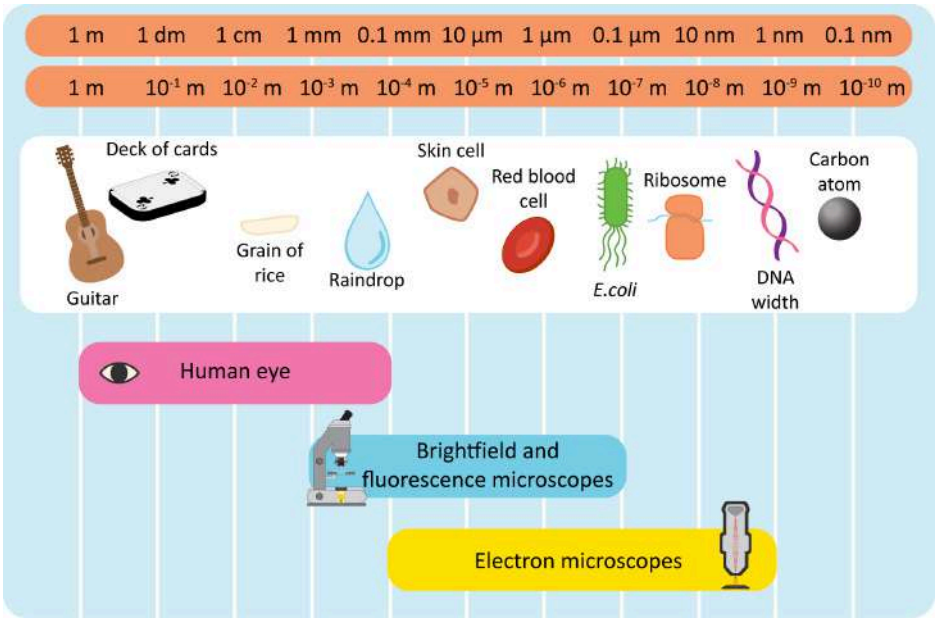


Figure 01-02 (repeated): Schematic diagram comparing methods to visualize different-sized objects. This image was created by Heather Ng-Cornish and is licensed under a CC BY-SA 4.0 license.

It is common when learning to identify microscopy from images to assume that anything imaged at

a lower magnification is taken with a light microscope and anything at a higher magnification must be electron microscopy. However, that is not necessarily the case. Notice in Figure 01-02 that there is quite a lot of overlap in the range of light and electron microscopy. Other than the most extreme ends of very low magnification (like whole tissues, with many cells visible) and very high magnifications (like individual ribosomes), either type of microscopy can be used. This means that you need to take into consideration *more than just the magnification*. As we mentioned at the beginning of this chapter, resolution is also extremely important. For example,

- If you are using a light microscope at very high magnifications, you may be able to see bacteria (like *E. coli* as shown in the figure), but you won't be able to see much of the details of the cytosol.
- On the other hand, at the same magnification, an electron microscope would likely allow you to see much more fine detail. You would likely be able to identify the location of the nucleoid and see some of the details of the flagella and their attachment to the plasma membrane.

Brightfield versus Transmission Electron Microscopy (TEM)

We have also found that it can be quite challenging for students to differentiate between a lower-magnification TEM image and a higher-magnification brightfield image. Video 01-01, posted earlier in this chapter, does an excellent job of explaining this, so we would encourage you to watch that video if you haven't already. However, here are some things to consider that might help you to tell the difference:

- Brightfield microscopy allows for much lower magnification images than TEM does. At best, a low-magnification TEM image would only allow you to see two or three cells, so if you can see more than that, it's probably brightfield.
- Similarly, TEM allows for much higher magnification images, so if you're seeing only a small subsection of the cell's cytoplasm in great detail, it's probably TEM.
- If you can see a whole cell, or even a large part of a cell, it could be either TEM or brightfield, so here you need to think about resolution. TEM will show much more of the finer details of the cytoplasm, and many of the organelles will likely be identifiable. Brightfield is more likely to show you a bunch of fuzzy blobs inside the cell, and you may not be able to tell what they all are.
- Any technique that uses light microscopy has the potential to produce color images, so that could be something to consider as well, but cautiously. Remember that color is exceptionally easy to manipulate in digital images, and it can be added and removed at will. Also, if you are working to identify micrographs on a paper exam, it will most likely be printed in black and white, so you may or may not have the information about color in the image you are analyzing.

Scanning Electron Microscopy (SEM) versus Fluorescence Microscopy

With the advances in fluorescence microscopy and the ability to digitally reconstruct a stack of images into a single 3D image, it can also be challenging to tell the difference between images created by these two types of microscopes.

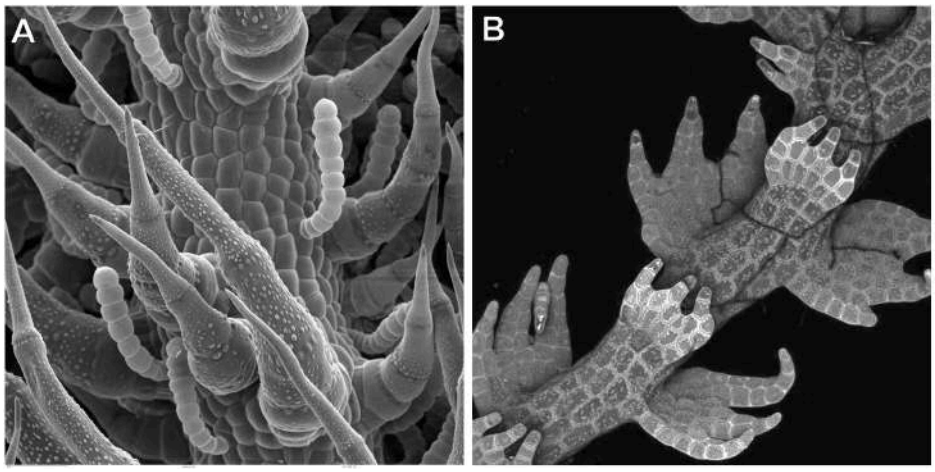


Figure 01-10: SEM versus TEM. (A) The lower surface of a sunflower leaf (Helianthus annuus) imaged by SEM. Attribution: Louisa Howard (2012), CIL:39342, https://doi.org/10.7295/W9CIL39342. This image is in the public domain. (B) 3D reconstruction of the structure of the little fingers liverwort (Lepidozia reptans). This image was collected using fluorescence light microscopy by Dr. Robin Young and licensed under a CC BY-SA 4.0 license.

The key difference is that in SEM, you can only ever see the surface of the sample. (Compare Panels A and B from Figure 01-10.) In Figure 01-10A, we have an SEM image, which shows the surface structure of a leaf. Only the surface of the leaf cells is visible and nothing else. On the other hand, in fluorescence microscopy, the sample itself is emitting the light that you are viewing. This means that anything that is emitting light will be visible, whether it's an internal structure or a surface element. In Figure 01-12B, it looks as if the cells are little compartments with dots in them. The dots are the chloroplasts inside the cell, emitting light. So even though this image looks 3D, we can still see beyond the surface of the sample, into the interior of the cells. Sometimes in SEM, the tissue is cracked open, so the inside of the cells is visible. (See Figure 01-12C for an example.) However, it is still very clear that it is the surface of the sample that's being viewed in SEM, even though that surface is from the interior of the cell.

One additional thing to look for is the "quality" of the blackness of the background (when visible). In fluorescence, since the light is being emitted by the sample, the background will be very black, with no defining features. In SEM, however, the electrons are bounced off the surface of the sample, so there's more likely to be some depth or texture to the background, even when it's very dark/black. This may be subtle, but when visible it gives a useful clue.

Of course, none of these tricks work 100% of the time. The technology we have developed for imaging is very good, and the differences between them are getting smaller. There are times when it can be impossible to tell. In those cases, there are still options. In these cases, we encourage you to take your best guess based on the evidence you have. Sometimes that is all you can really do, and then you can see where the rest of the conversation takes you.

CHAPTER SUMMARY

Microscopes are arguably the most essential tool to study cell biology. Virtually everything we study in cell biology is far too small to see with the naked eye. There have been some pretty significant advances in microscope technology, especially in the last 100 years. Here we break down the different forms of microscopy into the following:

1. emitted light microscopy, represented by the brightfield light microscope
2. transmitted light microscopy, represented by the fluorescent light microscope
3. transmission electron microscopy (TEM)
4. scanning electron microscopy (SEM)

Each one has different capacities in terms of magnification and resolution, both of which are key to determining what you should expect to see with each of these microscopes. It's important to remember that the bulk of cell biology is within a size range that allows both light and electron microscopy to be used, so it's important to consider how much detail you can see (i.e., the resolution), as well as the magnification and other parameters, when trying to decide how a sample was imaged.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

1. What determines the resolution of a microscope? Why is resolution more important at higher magnifications?
2. Compare and contrast the four different kinds of microscopy covered in this chapter (brightfield, fluorescence, transmission electron, and scanning electron microscopy). As you do this comparison, consider the following parameters:
    a. Particle used to image (i.e., photons versus electrons)
    b. Path of particle through the microscope to produce the image (Can you make a rough sketch?)
    c. Sample preparation (Should samples be alive or dead? Stained or unstained? Any other unique characteristics of each one?)
    d. Major advantages and disadvantages of each kind of microscopy
    e. Tips and tricks for recognizing images produced by each type of microscopy that will help you on your exams
3. Practice interpreting micrographs! Find a really good website that showcases micrographs that were imaged using a variety of microscopes (an exceptionally good one is www.cellimagelibrary.org, but there are others as well). Then use that website to practice interpreting micrographs and identifying cellular features. Try to find examples of the same structure that has been imaged using different types of microscopes and using different staining techniques.
4. Identify the type of microscopy that is best suited to detect and study cellular components based on their size and functional aspect being studied.
5. How do different staining techniques (i.e., immunolabeling, negative staining, etc.) help you

visualize your sample?

6. What are the advantages of using green fluorescent protein (GFP) instead of more traditional immunofluorescent staining techniques? How does it work?

CHAPTER 2.

BIOLOGICAL MEMBRANES

INTRODUCTION

Now that you've had a chance to better understand microscopes, we now turn to learning about the makeup of the cells themselves. In this chapter, we will start by exploring the composition and function of biological membranes. This will give us the opportunity to review what we know about macromolecules, specifically lipids and proteins and, to a lesser extent, carbohydrates. After that, we will discuss the function of the plasma membrane and how the structure of those macromolecules contributes to the function of the plasma membrane. Finally, we will discuss cell surface–adjacent structures, like the extracellular matrix or plant cell wall, that are found just outside of the plasma membrane.

TOPIC 2.1: THE CHEMICAL FEATURES OF BIOLOGICAL MEMBRANES

Learning Goals

- List the four primary features of a biological membrane and explain why they are important for cellular function.
- Explain how the chemical composition of a membrane (including lipids, carbohydrates, and proteins) contributes to its function.
- Explain how thermodynamics and the hydrophobic effect hold membranes together and selectively exclude some molecules but not others.

All membranes in the cell share some common features, despite the fact that they each have a different composition, fluidity, and permeability. In this topic, we will discuss the common features of all membranes and touch briefly on which macromolecules contribute to specific membrane functions.

It's worth noting that a **biological membrane** is not the same thing as a **phospholipid bilayer**, despite the fact that these terms sometimes are used interchangeably. A phospholipid bilayer is made only of phospholipids and nothing else, whereas a biological membrane will have many types of lipids in it (including glycolipids and cholesterol) as well as proteins. We'll see examples of this later in the topic.

Membranes Are Barriers That Define Compartments

One of the most basic features of cells is that they are separated from the environment by a membrane.

Not only do the membranes create a barrier that separates the inside of the cell from the exterior, but they also function in a way that gives the cell a great deal of control over what can enter or exit it. This creates an environment inside the cell that is different from the outside (Figure 02-01).

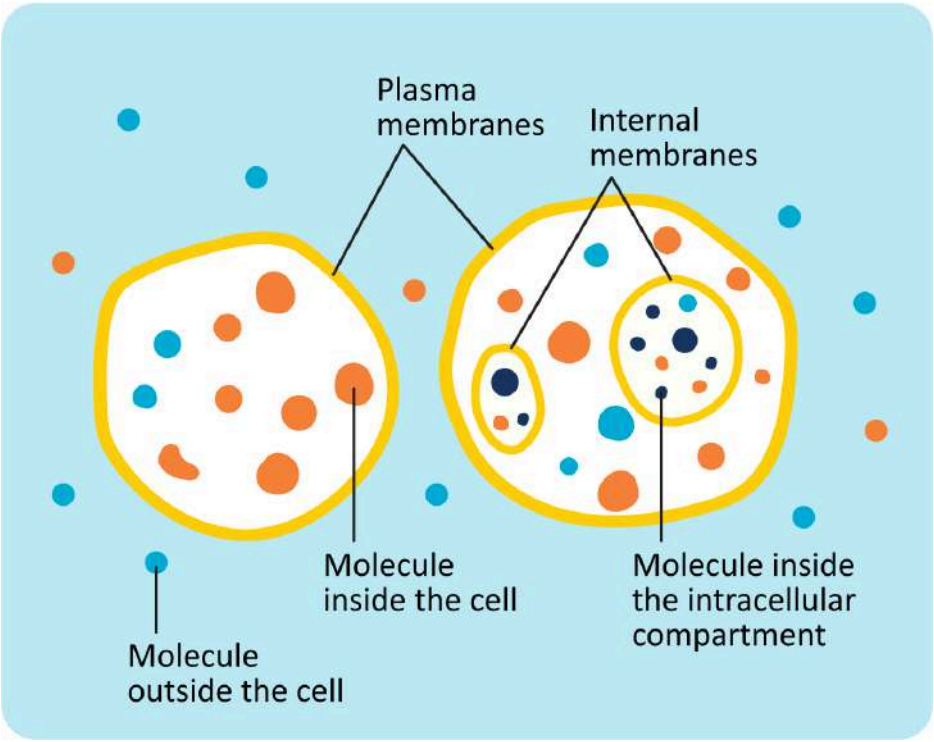


Figure 02-01: Biological membranes act as barriers and separate molecules (represented as colored blobs) into different compartments. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Molecules are able to move around freely (diffuse) within each cellular compartment. However, it is more challenging for some molecules to cross membranes. This is because biological membranes have specific chemical properties, which determine when and how molecules cross the membrane. Thus, membranes are a key feature that allows cells to maintain compartments with a distinctive chemical composition.

Like everything in science, the idea that biological membranes contain both lipids and proteins took time for cell biologists to accept. Sometime in the 1970s, the current concept of the membrane was developed and was known as the "fluid mosaic model." The "fluid" part refers to the lipids being able to move around within the layers of the membrane, and the "mosaic" part refers to the fact that proteins are expected to be scattered across and throughout the membrane. Since then, further research and experimentation have expanded and refined our understanding of how membranes work. In this part of the chapter, we will explore four general features of biological membranes:

1. The membrane is a bilayer, made up of lipids and proteins.
2. The membrane is selectively permeable.
3. The membrane is organized but fluid.
4. The membrane is asymmetric.

We will look at each of these in a little more detail. Then in Topics 2.2 and 2.3, we will take a more in-depth look at the relationship between lipids and membrane fluidity as well as the organization and function of membrane proteins.

General Features of Biological Membranes

The Membrane Is a Bilayer Made Up of Both Lipids and Proteins

The earliest work on membranes was done on plasma membranes from red blood cells. This is because red blood cells are easy to collect in relatively large quantities. In addition, in mammals, they lack a nucleus and internal membranes. So they contain mostly "pure" plasma membrane without the need to experimentally separate other membrane types during experimental isolation. To isolate these membranes, cells are burst open and the membranes are gently washed to remove cytoplasmic debris. Lipids and proteins are separated by treatment with strong detergents or organic solvents (e.g., ether) and characterized. Next, we'll take a look at the chemical properties of lipids and proteins separately to discuss how they each contribute to the structure of the biological membrane.

1a. The Lipid Component of Membranes: Formation of Lipid Bilayers

(For a complete review of the basic structure and formation of lipid molecules as well as a description of **polar** and **nonpolar** molecules, please see the introduction.)

The main lipid components of membranes are **phospholipids**, a family of molecules with a polar phosphate head group and two fatty acid tails, each connected to an arm of a glycerol via an alcohol residue (Figure 02-02A). However, they are almost never alone in a biological membrane. Other lipids (glycolipids, sphingolipids, and cholesterol) as well as a great many proteins are found in every biological membrane. The chemical properties of phospholipids allow them to adopt a **bilayer** shape, a key component to a biological membrane.

In water, fats and oils will form into large droplets of jumbled molecules that do not mix with the water (think salad dressing). However, membrane lipids are able to form sheets because of two very important characteristics:

- They are **amphipathic** (i.e., there is a polar "head" region, which can freely form hydrogen bonds with water, and a nonpolar "tail" region, which cannot). Some sources refer to polar molecules/regions as **hydrophilic**, which translates to "water loving." Conversely, the nonpolar molecules/regions are sometimes referred to as **hydrophobic**, which translates to "water fearing," but as these terms are misleading and stray from referencing the chemical properties of the molecules, we don't find them to be as useful. We prefer the terms **polar** and **nonpolar** when describing the head and tail regions of the phospholipids, as they are more scientifically accurate. However, both will be used in this textbook.
- They form a roughly cylindrical shape, which tends to stack together well (much like cans of soda; see Figure 02-02B).

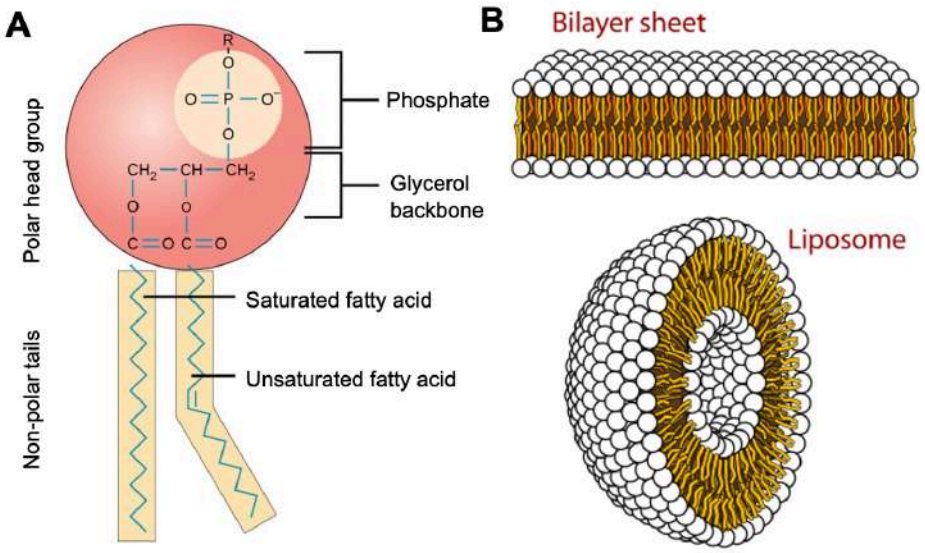


Figure 02-02: Phospholipids form membranes. (A) Structural features of an individual phospholipid, showing the polar head group (which also contains the phosphate ion, the namesake of the molecule) and the nonpolar hydrocarbon tails. Often at least one of the tails will be unsaturated and have one or more double bonds. Figure 05 01 02.jpg by CNX Biology Textbook is shared under a CC BY 4.0 International license. (B) Orientation of phospholipids in a bilayer and liposome structure, where the head group faces out and the lipid tails associate in the center of the bilayer. Phospholipids aqueous solution structures.svg by LadyofHats. This image is in the public domain.

Phospholipids, such as the one shown in Figure 02-02A, come together to form a bilayer—two sheets of phospholipids with their polar head groups oriented outward. Each layer of the bilayer is called a **leaflet**. Since the nonpolar tails of the phospholipids face inward, away from the water, they are sequestered away and do not interact directly with the water molecules (Figure 02-03). This orientation allows for the maximum freedom of the surrounding water. Because of the stable membrane conformation, if a tear or a hole is created in a lipid bilayer or a biological membrane, it rapidly seals up again.

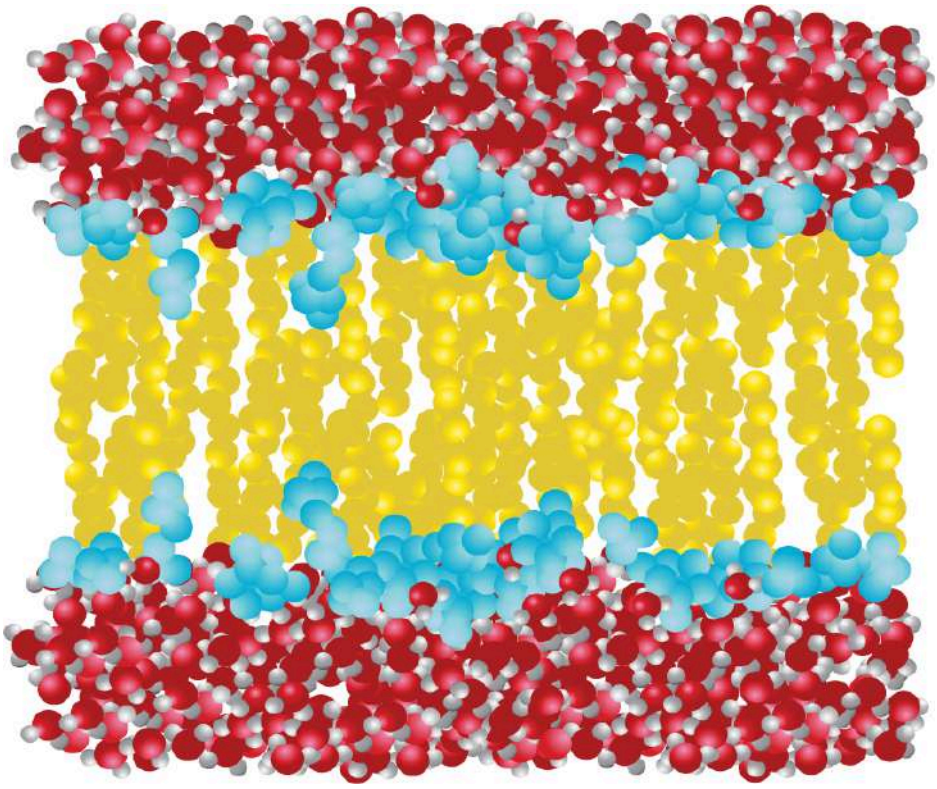


Figure 02-03: Molecular representation of the lipid bilayer. Water molecules (red/white) cluster near the phospholipid head groups (blue). The yellow lipid tails congregate in the center of the structure. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

When phospholipids are mixed with water, they will spontaneously form into spherical bodies called **liposomes** that have water on the inside as well as on the outside (Figure 02-02B). Liposomes are commonly used for the targeted oral delivery of drugs and other agents in medical treatment. Liposomes are the simplest version of a cell membrane, so they are able to fuse with a real cell membrane to release its contents directly into the cell. This is useful if the compound inside the liposome would not easily pass through a membrane under normal circumstances. A great example of this technology in action is the RNA vaccines developed to target the SARS-CoV-2 virus (the cause of the COVID-19 pandemic). These RNA vaccines use an outer coating called a lipid nanoparticle, which is a combination of phospholipids, cholesterol, and other compounds that are designed to help contain and stabilize the RNA vaccine and ease its entry into the cell.

Thermodynamics

This text assumes that you have preexisting knowledge of thermodynamics from first-year chemistry. If you need a quick refresher, there are videos in the introduction

Thermodynamics Drive the Formation of the Lipid Bilayer: The Hydrophobic Effect

Phospholipids form stable bilayers in an aqueous environment due to thermodynamics. The

spontaneous formation of a lipid bilayer from lipids in an aqueous solution, simulated in Video 02-01, shows the power of thermodynamics in action. The laws of thermodynamics explain why it is more energetically favorable for phospholipids to clump together and form a bilayer in water than it is for them to remain dispersed.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=145#oembed-1

Video 02-01: Computer simulations of molecules in solution over time. Fatty acid tails (green) spontaneously assemble in a tail-to-tail fashion with the head groups interacting with the water molecules (gray).

The phenomenon where nonpolar molecules in an aqueous environment clump together is called the **hydrophobic effect**. Clumping together is not favorable because the nonpolar groups attract each other per se, but because, when clumped together, nonpolar molecules do *not* have to interact as much with water, which is polar. When the nonpolar material clumps together, water and other polar molecules are freer to move and, more importantly, hydrogen bond with each other. This freedom of motion for the water allows for the overall reaction to be spontaneous.

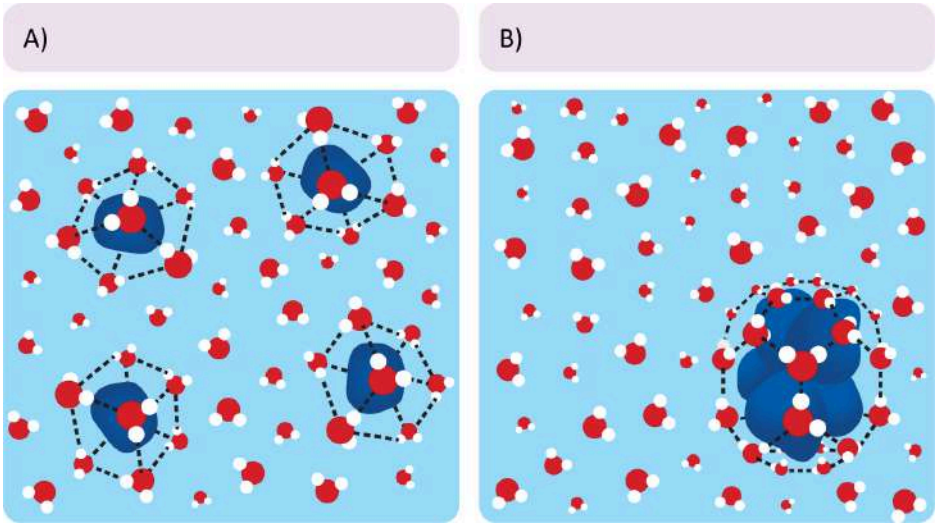


Figure 02-04: Nonpolar molecules clump together in aqueous solution. (A) several nonpolar molecules are in aqueous solution. Water molecules around a nonpolar particle form a cage-like structure, which reduces their motional freedom. (B) The nonpolar molecules are clumped together. This allows more water molecules to be free and engaged in the cage-like hydration shell outside of the nonpolar particle. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

While we don't want to get too far into the equations of thermodynamics, here is a brief description

of how thermodynamics favors the hydrophobic effect. You may want to return to your general chemistry notes for this (or refer back to the information in the introduction).

Recall from general chemistry that the **free energy equation** is

$$\Delta G = \Delta H - T\Delta S,$$

where ΔG is the Gibb's free energy change, ΔH is the change in enthalpy (often described as internal energy or bond energy of the molecule), ΔS is the change in entropy (which is often thought of as motional freedom), and T is temperature.

As a reminder, *reactions are energetically favorable when ΔG is negative, which means that energy will be released.* This can be accomplished by *decreasing* the enthalpy or by *increasing* entropy. In the case of the formation of the lipid bilayer (and any other clustering of nonpolar molecules in water), it is the entropy of the water molecules surrounding the nonpolar groups that is changed when lipids assemble into bilayers. Two points are critical here:

- Nonpolar particles force the surrounding water molecules into an energetically unfavorable configuration that distorts the normal hydrogen-bonded structure of water (Figure 02-04). This cage-like structure restricts the movement of the water surrounding the nonpolar particle. Clumping the nonpolar molecules together reduces the surface area that is required for the nonpolar molecules to interact with the surrounding water.
- In the case of phospholipids, when the nonpolar lipid tails are sequestered away from the surrounding aqueous environment, they don't come into contact with water at all. Thus, the water molecules are not constrained and are able to freely hydrogen-bond with the polar head groups and/or other polar molecules. This net reduction of the number of "constrained" water molecules provides a large *increase* in entropy. Based on our equation above, all other things being equal, that increase in entropy can be enough to result in a negative change in free energy. Thus, it requires less overall energy to hold phospholipids in a bilayer formation, making it "energetically favorable."

1b. The Protein Component of Membranes

The second major component of membranes is proteins, which should come as no surprise. Virtually everything in the cell is either made *of* proteins or made *by* proteins, so the importance of understanding how the structure of proteins impacts their function cannot be overstated.

We expect that you have learned about proteins in your general biology classes, so we approach this topic from that perspective. (*If you need a refresher, we suggest looking at the material in the introduction.*) However, even if you have learned about proteins before, there is much more to learn. Proteins are important enough that you should expect to cover them in several courses, with increasing levels of detail and complexity. In this textbook we will be exploring protein structure and function, and how it specifically relates to membrane structure and function, in Topic 2.3.

General Features of Biological Membranes

The Membrane Is Selectively Permeable

We say that membranes are **selectively permeable** because experiments show that they allow some molecules to pass through while excluding others. However, it has also been shown that the properties of different membranes vary…a lot! Some are very permeable to ions and water, while others allow almost nothing through.

The chemical composition of the membrane, specifically the lipids that make up the bilayer, is a large factor in the membrane's capacity to be selectively permeable. Due to the primarily nonpolar environment inside the core of the bilayer, it is extremely unfavorable for water or other polar molecules to spend any time in there; thus, the interior of the membrane is almost entirely a water-free zone. This does not necessarily mean that small polar molecules (like water) never enter the core of the bilayer, but if they do, they exit again quickly. This is the basis for the reduced permeability of membranes for polar molecules.

To further explore selective permeability, we will look at the simplest example—a synthetic bilayer made entirely of phospholipids. In biological membranes, there are a variety of **transport proteins**, which allow molecules to pass through the membrane. Looking first at a pure phospholipid bilayer can help us understand why some molecules need transporters, whereas others do not.

- Depending on the properties of the molecule in question (size, polarity), different molecules will have more or less difficulty crossing through that nonpolar portion of the membrane. Small nonpolar molecules (like oxygen, carbon dioxide, and nitric oxide) can cross a lipid bilayer easily, without any help.
- Polar molecules that are very small (like water) are also able to cross bilayers, but not very well. As such, the process is slower, and many membranes have transport proteins to facilitate passage when necessary.
- Larger molecules (amino acids, nucleotides, and glucose) will struggle to cross the bilayer as a result of their size. These molecules can also be polar, which can add to the challenge.
- Charged molecules, regardless of size, will be completely unable to pass through the bilayer without help. This is the basis for many of the electrochemical gradients that the cell uses to drive work in the cell.

Most biologically important molecules are either too big (like glucose) or too charged (like ions) to pass through a membrane spontaneously. Transport proteins help these molecules to go across a membrane even if their chemical properties prevent them from diffusing spontaneously. The selective transport of molecules may or may not require the input of energy, like ATP, or rely on concentration gradients. The Amoeba Sisters have an excellent review video (Video 02-02) that explores the selective permeability of the membrane and how transporters regulate molecules going in and out of the cell.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=145#oembed-2

Video 02-02: Review of cellular transport across membranes.

Having control over which molecules are able to cross a given membrane is a key factor in how cells and organelles function. For example, the mitochondria's capacity to generate ATP is almost entirely dependent on the formation of a concentration gradient (discussed in Video 02-02, above).

Concentration gradients are used extensively by the cell to do work. They are analogous to a hydroelectric dam. A dam holds water in an upstream reservoir and controls the flow of water through turbines to create energy. In this analogy, the membrane is the dam, and specific ions are held at different concentrations on either side. As the ions naturally "flow" from the side of the membrane that has a high concentration to the side that has a low concentration, energy is released. The release of the molecules down the gradient is often used to "power" other cellular processes and thus is a key part in understanding advanced physiology. In this textbook, due to the metabolic pathways that we focus on, we mostly discuss the formation of proton gradients (i.e., H^+ ions), but other electrochemical gradients are equally important to proper cellular function.

General Features of Biological Membranes

The Membrane Is Organized but Fluid

Lipids have a key role to play in the organization and fluidity of a membrane (a key concept). Proteins in the membrane, while also organized, may or may not be fluid, depending on whether they are anchored to internal or external structures. First, we will explore how we define fluidity in this context.

There are two kinds of fluidity in membranes that must be considered:

- lateral motion of lipids within a single leaflet of the bilayer and
- overall "stiffness" of the membrane.

Let's look at each of these briefly in turn.

Lateral Motion of Lipids within a Single Leaflet Is Far More Favorable Than Movement between Leaflets of the Bilayer

The hydrophobic effect keeps the nonpolar portions of membrane lipids in the center of the membrane and is the reason membranes form spontaneously. As long as the nonpolar portions of the phospholipids stay in the interior of the membrane, their movement is not restricted (Figure 02-05). Within a membrane or bilayer, individual phospholipid molecules can spin rapidly on their axis and/or diffuse laterally within the plane of the bilayer. On the other hand, they very *rarely* flip-flop across the bilayer, as it would require the polar head group to pass into the nonpolar center of the membrane.

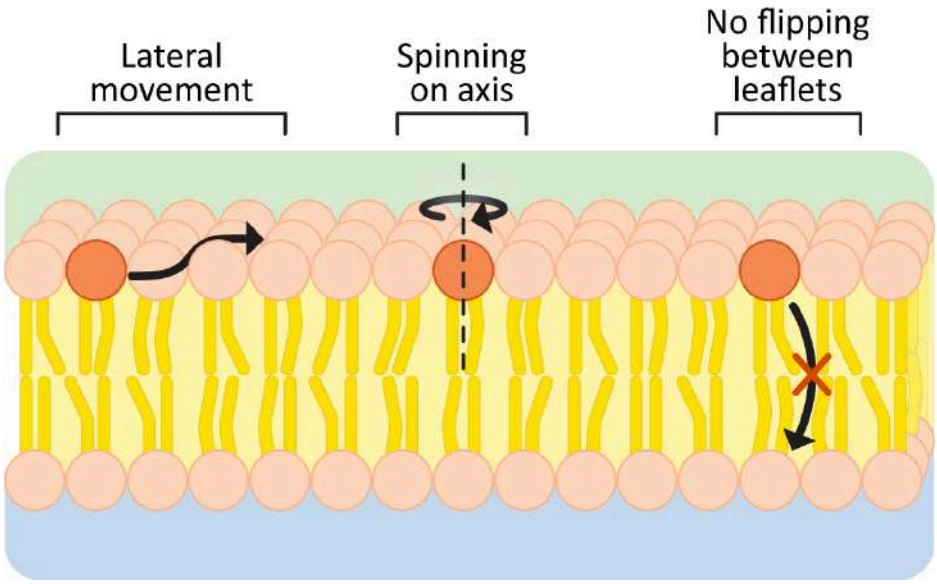


Figure 02-05: Membranes are two-dimensional fluids. Each leaflet of the bilayer is separate from the other. The lipids within the leaflet can move in a variety of ways (shown). The one thing that they can't do well is flip from one side of the membrane to the other. Flipping is not energetically favorable. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

To further illustrate this point, here are some data to convince you:

- the frequency of lateral shifts of a phospholipid within the same layer is roughly 10^6/sec, but
- the frequency of flips from one leaflet to the other is closer to 10^{-5}/sec.

Thus, flips between bilayer leaflets are about 10^{11} times *less frequent* than lateral movements within a phospholipid layer. Based on these numbers, an individual phospholipid would "flip" to the other leaflet about once every 28 hours on average. Thus, when we discuss membrane fluidity, we are generally referring to the lateral mobility of membrane components within a single leaflet of the membrane. It is also why we often refer to membranes as *two-dimensional fluids*.

Functional biological membranes require some movement of their lipids. The movement of lipids allows for rapid resealing of membranes in response to small holes or tears. However, too much fluidity can be damaging. A membrane that allows too much motion might have trouble keeping everything in its correct location. Think about a boiling pot of water, where it's boiling so hard that water is spilling out everywhere. Membranes with too much fluidity may become holey or lose vital parts, which will make it challenging for the membrane to maintain its functionality.

Membrane Composition Also Determines How "Stiff" a Membrane Is

Membrane stiffness refers to how pliable or bendable the membrane is. It is a separate but equally important component of membrane fluidity, which is determined by the composition of the membranes. If membrane lipids pack too close together, the membrane will freeze in place, and the result will be a membrane that is too rigid to adapt as the cell moves and changes. On the other hand, a membrane that is too pliable will struggle to maintain the shapes required. Thus, the cell needs to manage the fluidity of its membranes in more ways than one. The cell manages this by controlling the precise lipid composition of the membrane. Remember that within each grouping of membrane

lipids (phospholipids, glycolipids, sphingolipids) is a large family of similar molecules. By changing which phospholipids (or sphingolipids, etc.) and how much cholesterol is in the membrane, the cell can maintain its membranes within the correct range. That way they will be fluid enough to allow movement and bending but not so fluid that the cell struggles to control function. In Topic 2.2, we will look more deeply at how this fine balance is maintained by the cell in all kinds of environmental conditions.

General Features of Biological Membranes

The Membrane Is Asymmetric

Since membranes are exposed to different compartments and different environments on either side, it stands to reason that one would not expect the membrane, or its composition, to be identical on either side as well. This phenomenon is referred to as membrane asymmetry. There is always an inside that faces the cytoplasm (cytoplasmic side) and an outside facing the interior of an organelle or, in the case of the plasma membrane, the external environment (extracellular side). Both the lipids and the proteins play an important role in membrane asymmetry.

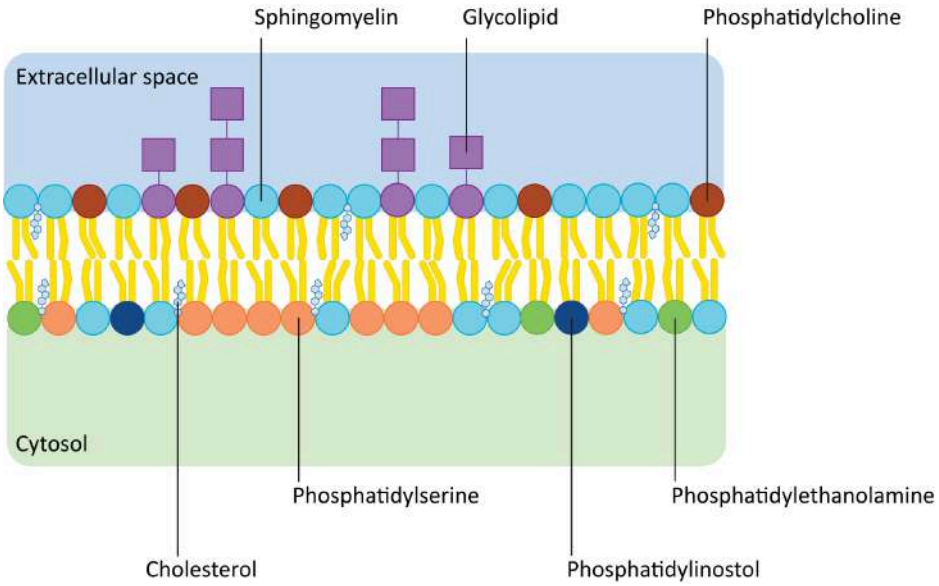


Figure 02-06: The composition of lipids in the biological membrane is distinct and asymmetric. This schematic of a plasma membrane clearly shows differences in the lipids on each side of the membrane. On the exterior side, which faces the extracellular space, we see a combination of glycolipids, phospholipids, and sphingolipids. On the side of the membrane that faces the cytosol, we see no glycolipids and very few sphingolipids (if any). Instead, we see an increase in specific types of phospholipids. Each membrane will have a slightly different composition in the cell, and different cell types will also have a different composition. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Figure 02-06 shows how the composition of lipids typically differs in the two leaflets of the plasma membrane. Remember that the term **phospholipid** refers to a whole family of lipids, so even though leaflets are made of phospholipids, the exact composition may not be the same. In addition, there are other kinds of lipids (e.g., sphingolipids and glycolipids) that might be enriched in one side of a membrane versus the other. For example, glycolipids are almost never found on the cytosolic side of the plasma membrane. Glycolipids face the cell exterior, as they play key roles in extracellular signaling and cell identity. Within the cell, organelle identity is also signaled by specific membrane

lipids known as phosphatidylinositols, or **PIPs** for short. PIPs face the cytosolic side of the membrane and are used extensively by signaling and vesicle transport machinery.

The proteins in a membrane are also asymmetric. The parts that are sticking out of each side of the membrane will naturally be different to accommodate their specific functions. For example, a receptor protein will need to have the signal-binding site on the exterior surface of the membrane where the signaling molecule is present. Thus, each membrane protein needs to be inserted into the membrane in a very specific way to preserve proper function. Thus, the two sides of a membrane are not interchangeable, and the cell must keep them properly oriented at all times. We will discuss how this orientation is maintained in specific processes, like vesicle trafficking, in Chapter 4.

For Many Membranes, Carbohydrates Are Also Major Contributors to Membrane Asymmetry

At this stage, it is essential that we take a moment to point out the role of carbohydrates, the fourth major biological macromolecule. While they sometimes get overlooked in introductory cell biology courses—simply because there are so many other things to cover—they are absolutely essential to proper cellular function. The plasma membrane of animal cells is usually completely covered with carbohydrates—most commonly attached to proteins but also to lipids. This coating (called the **glycocalyx**) plays a key role in cellular identity and signaling. Plants and fungi have so many carbohydrates surrounding them that they form a structure known as a **cell wall**. We will take a moment at the end of this chapter to explore both the **extracellular matrix** and plant cell wall in a little more detail.

The Origin of Membrane Asymmetry

While it's easy to understand the need for membrane asymmetry, it is much more complex to understand how asymmetry is created by the cell in the first place. In order to understand that, we must explore how membrane lipids are made by the cell.

All membrane lipids are first synthesized in the cytosol and then inserted into the cytosolic face of the smooth ER. This means that all new lipids are being added to *a single leaflet* of the ER membrane. As you can well imagine, this causes pressure on the ER membrane, as one leaflet is growing while the other is not. If left this way, the pressure would eventually cause a potentially catastrophic rearrangement as the membrane returns to a more thermodynamically favorable conformation. To avoid this kind of event, the cell uses proteins known as **scramblases** to rebalance the lipids in each leaflet (Figure 02-07A). This is done nonselectively—the scramblase simply equilibrates the phospholipids on either side of the membrane regardless of type. Later, in other regions of the ER, but more commonly in the Golgi, the membrane is organized more precisely. Enzymes called **flippases** move specific phospholipids unidirectionally from the exterior leaflet to the interior one so that membrane asymmetry is properly established, and the lipids are placed where they need to be for function. Both enzymes use a similar mechanism—the enzymes are big enough to span the membrane, and the polar head group is slid through a slot in the enzyme so it can cross the nonpolar part of the bilayer (the tails stick out and are left free to the environment; see Figure 02-07). Thus, they can overcome the energetically unfavorable parts of the reaction and flip lipids from one leaflet to another.

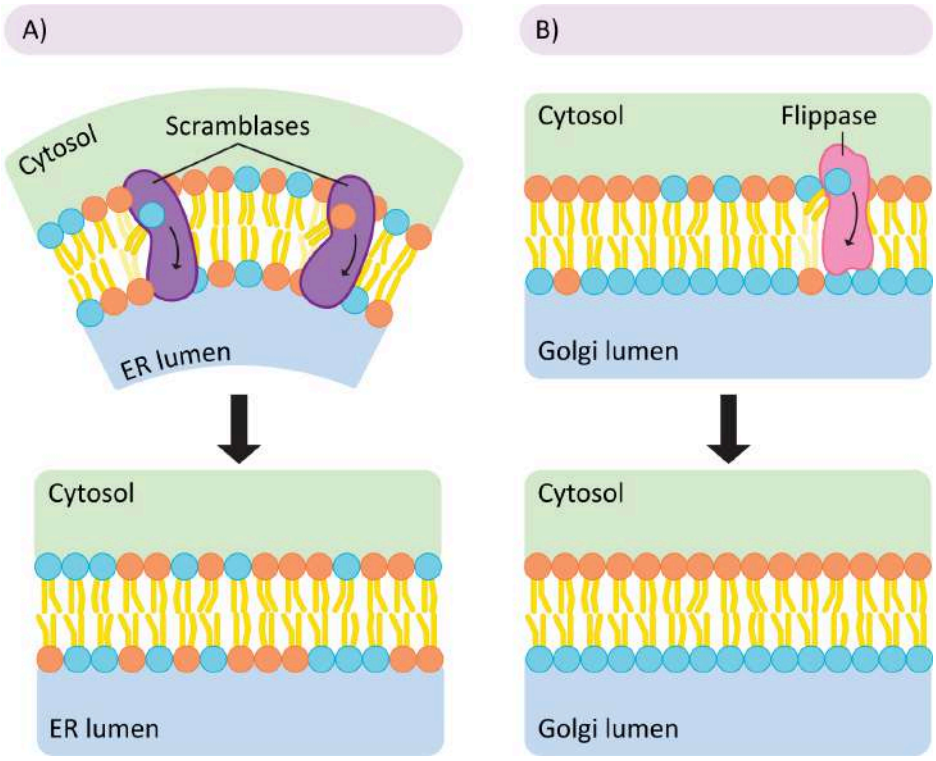


Figure 02-07: Flippases and scramblases flip lipids from one leaflet to another. (A) Scramblases are nonspecific and can help relieve the pressure buildup from adding lipids to one leaflet only in the Golgi. (B) Flippases flip specific lipids to establish asymmetry of lipid types in the biological membrane. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

TOPIC 2.2: MAINTAINING FLUIDITY IN THE MEMBRANE

Learning Goals

- Describe how the lipid composition can influence membrane fluidity.
- Explain how membrane fluidity is maintained under different environmental conditions, like temperature.
- Explain how fluorescence recovery after photobleaching (FRAP) works, and interpret outputs from these experiments.
- Compare and contrast membrane-adjacent extracellular components in plants (i.e., the plant cell wall) and animals (i.e., the extracellular matrix).

In the first topic of this chapter, we focused on the major characteristics of biological membranes as a whole, and how the chemistry of the membrane can contribute to its properties. Here we'll go into more depth about how the fluidity of a given membrane is the product of its precise lipid composition. We will also learn about an experimental technique known as **fluorescence recovery**

after photobleaching (FRAP), which allows us to assess the mobility of molecules within the membrane.

in depth
Membranes Are Organized and Fluid

When we think of the term *membrane fluidity*, the first thought might be that this simply indicates how quickly any given molecule can move from point A to point B within the membrane. However, this is only one aspect of membrane fluidity. As we mentioned in the previous topic, the *stiffness* of the membrane must also be considered.

Membrane stiffness describes how easily the membrane bends (compared to how easily or quickly the molecules within the membrane can move around). Membranes vary in how stiff they are and how they respond to their shape being distorted. For example, red blood cells are elastic, which means they can be distorted but bounce back to their original shape. On the other hand, white blood cells are much more rigid and don't deform as easily. As an analogy, a red blood cell is like a plastic bag, and a white blood cell is more like a cardboard box. This is the result in differences in both the lipids and proteins that are associated with each cell's plasma membrane. The composition of the membrane is key for the cell's capacity to function in its environment.

Membrane Fluidity Is Heavily Influenced by the Environment and Must Be Maintained at a Constant Level

Organisms live in all kinds of environments, from the hottest regions of the Sahara to the coldest regions of the Arctic and everything in between. They also live in regions of extreme pressure, like the deep sea, which can impact both the freezing and boiling point of water. In all of these extreme environments, cells must control the chemistry of their membranes so that they remain fluid and functional. For example, membranes must not freeze when exposed to low temperatures (or high pressure) or lose their integrity at higher temperatures. Any organism that lacks a mechanism to control the fluidity of its membranes will die if/when environmental conditions change.

There are two major approaches that organisms take to controlling the fluidity of their membranes:

Approaches to Membrane Fluidity

Endotherms		*Ectotherms*	
Pros	Cons	Pros	Cons
• As the core body temperature is controlled, *endotherms* can handle rapid changes in external temperature. This also means that the cells spend less energy synthesizing membrane lipids in response to temperature changes.	• They must be careful to maintain that core temperature in all situations, which can be challenging depending on the environment. Small changes in core temperature can have big effects. Examples of this that humans experience are hypothermia and fever.	• *Ectotherms* can live in a wide variety of environments as they are more adaptable to extreme conditions.	• Production of temperature flexible membranes takes time. If the environmental changes happen too quickly, the organism may not be able to adapt.

This section naturally focuses on exotherm membranes, whose membrane composition changes more in response to the environment. One good example of membrane composition being used as an environmental adaptation is cold-hardiness in plants. As the temperature lowers in the fall and winter, membrane lipids become increasingly unsaturated. This lowers the melting/freezing point of the membrane, which helps it maintain flexibility, and retains the mobility of the membrane components. This also explains why organisms might be fragile with respect to sudden, unexpected changes in temperature. In the example of the plant, a rapid drop in temperature at the wrong time (like a cold snap in early autumn) could result in the death of the plant, as they wouldn't have the time required to synthesize new membrane lipids (and other antifreeze proteins, which help avoid freezing).

Membranes That Are Unable to Adapt Chemically Are Susceptible to Failure in Nonoptimal Temperatures

The capacity of a membrane to remain fluid at any given temperature is primarily dependent on its capacity to maintain just the "right" amount of entropy in its phospholipid tails. If they pack too close together, they will freeze, and if they have too much motional freedom, they will fail in other ways.

Note

Increased pressure has more or less the same effect as a decrease in temperature. Both temperature and pressure reduce the kinetic energy of the molecules, making it harder for them to partake in random kinetic motion.

When a membrane (or anything else) freezes, what really happens is that as the temperature lowers, the kinetic energy of the molecules is also lower, so they become sluggish and do not move around as freely. At some point, the kinetic energy of the phospholipids is so low that they get drawn into the transient *induced dipole–induced dipole interactions* that continue to happen between neighbors and can no longer break away. Eventually, the phospholipids will get "trapped" by the intermolecular forces being exerted on them by their neighbors, and the lipids will freeze in place. In chemistry terms we say that the membrane is in "gel phase." You can see the result of this any time you pull a piece of frozen meat from the freezer. It is quite stiff and inflexible compared to its room-temperature counterpart.

At the other extreme, when the temperature is too high for the membrane, the molecules have quite a lot of kinetic energy. If the kinetic energy is too high, then the intermolecular forces that hold the membrane together are too easily broken, and molecules can break free all together. When this occurs, important complexes will come apart, ions and other molecules may be able to slip through the chaotic membrane lipids, and the membrane itself could just fall apart.

Maintaining Fluidity—Membrane Composition

Cells have three different ways that they can modulate in order to influence the fluidity of a given membrane.

1. Change the Degree of Unsaturation of the Hydrocarbon Chains

Increasing the unsaturation of the lipid fatty acid tails increases fluidity of the membrane. **Unsaturated lipids** contain one or more double bonds between carbons in the hydrocarbon chain (compared to **saturated lipids** that have only single bonds in their hydrocarbon chains). The double bonds change the shape of the tail because they create bends, or "kinks," in the hydrocarbon chain (Figure 02-08). A kinked tail will not be able to pack as tightly with its neighbors. Imagine that on a packed subway everyone stands wide-legged, with their elbows sticking out. The net effect would be that each person would take up more space, and fewer people would be able to get into each car. The net effect of unsaturation is essentially the same—less tight packing, which effectively reduces the number of induced dipole–induced dipole bonds between lipids in the membrane. The result is that the lipids are able to maintain their mobility even at the reduced kinetic energies of lower temperature.

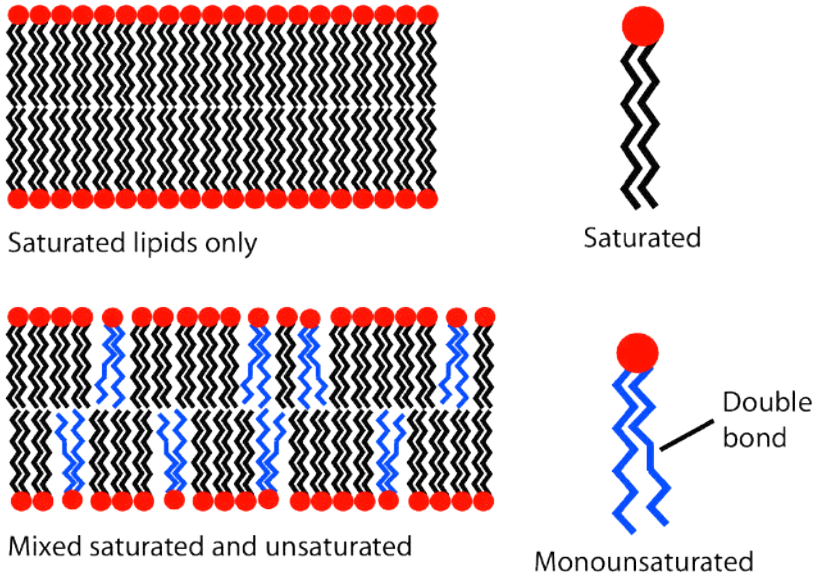


Figure 02-08: Depiction of how unsaturated phospholipids can impact the packing of lipids in a bilayer. Lipid unsaturation effect.svg by MDougM. This image is in the public domain.

2. Change the Number of Carbon Atoms in the Fatty Acid Tails

Shortening the fatty acid tails of a phospholipid will have a similar effect to using unsaturated fatty acid tails in that it decreases the number of transient bonds that can form between neighbor phospholipids. Longer hydrocarbon chains have more atoms in them; thus, there is the potential for more bonds to form. By shortening the hydrocarbon chains, the cell reduces the total number of atoms that can get involved in bonding at once, thus reducing the amount of energy that can be used to hold neighbor phospholipids together. Therefore, membrane lipids with shorter hydrocarbon chains will tend to have more mobility at lower temperatures. Conversely, regions of membranes with longer fatty acid tails can be used to reduce overall fluidity in that region and to hold specific proteins within that region.

3. Change the Sterol Content of the Membrane

Cholesterol is a type of sterol (ringed lipid) that sits in the spaces between fatty acid chains

(Figure 02-09). At first glance, based on what we have previously learned about creating space to reduce bonding, one would expect that this would be beneficial at low temperatures, as it would reduce the interactions between the fatty acid tails of the phospholipids. This is exactly true. At lower temperatures, *increasing* the amount of cholesterol will also increase membrane fluidity. But in this case, the story doesn't end there.

Paradoxically, cholesterol also has an effect at higher temperatures. However, in this case, increasing cholesterol content will *reduce* the fluidity of the membrane. It does this by filling the space in between lipids that is created when lipids have high kinetic energy. They also have a stiff ring structure, which also contributes to the stiffness of the membrane. Cholesterol at high temperatures *increases* the number of intermolecular forces between neighbors, which is important to stabilize the lipids in the membrane.

Think of cholesterol as a fluidity buffer. It helps keep the membrane within a specific range of fluidity so that it doesn't get too rigid or too fluid. Modulating the cholesterol can help control fluidity at many temperatures, which makes it an important membrane lipid in all environments. (*Note: Cholesterol is unique to animals. However, sterols are an important component of all eukaryotic membranes. Even though plants do not make cholesterol, they do use other phytosterols for the same purpose.)*

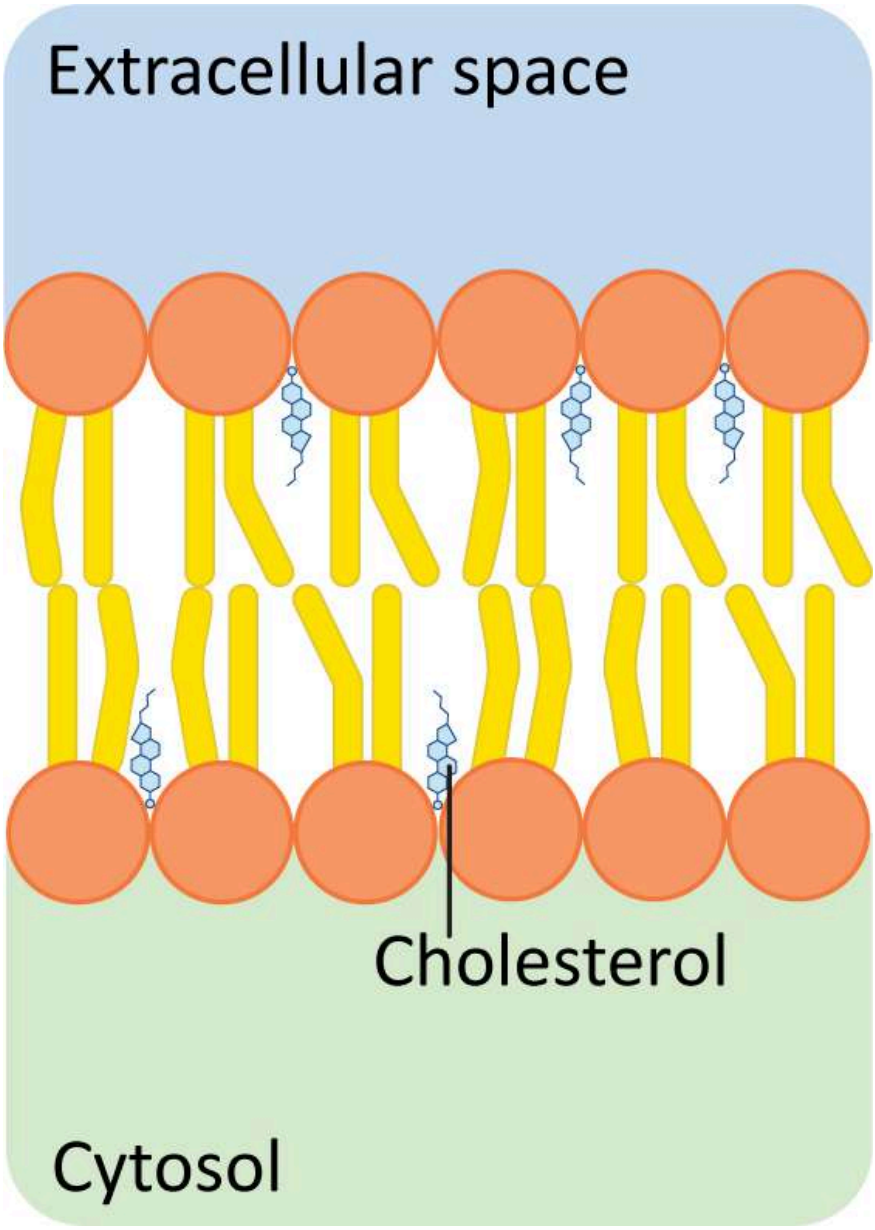


Figure 02-09: Cholesterol fits between the other lipids, inside the hydrophobic region of a phospholipid bilayer. Because cholesterol has a very small hydroxyl group as its only polar component, the majority of the molecule sits in the hydrophobic tail region. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

"Real" Membranes Contain Regions with Varying Fluidity

While it might be tempting to think of the cell membrane as a homogenous thing that is the same across the entire cell, there is significant evidence that this is not true. Both the lipid and protein components of any given membrane will be different in different regions, as different functions will need to occur. For example, **lipid rafts** are distinct regions of the membrane that contain a higher concentration of sphingolipids with longer fatty acid tails and cholesterol (Figure 02-09). The increased length of the sphingolipid tails in that region helps hold the lipids together and also increases the thickness of the membrane in these regions. Increased cholesterol in the region offsets the tendency toward decreased membrane fluidity that would result from the longer tails. Specific

proteins are preferentially drawn to these regions based on their chemical features (e.g., longer membrane-crossing domains). The specific clustering of proteins in lipid rafts can be used to keep proteins together that all work toward the same function (e.g., signaling or vesicle trafficking). A diagram of a lipid raft is shown in Figure 02-10.

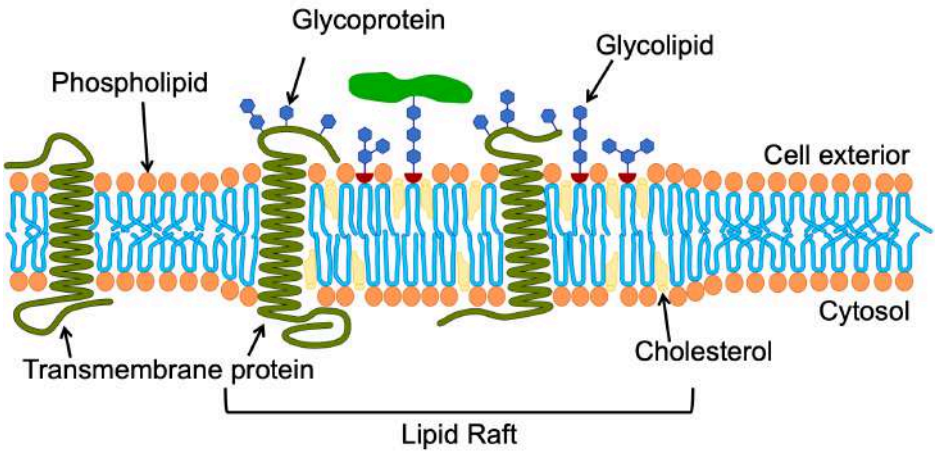


Figure 02-10: Organization of a lipid raft. Lipid raft organisation scheme.svg by WarX. This image is in the public domain.

Studying Cells: Fluorescence Recovery after Photobleaching (a.k.a. FRAP)

One of the experimental tools that is used to examine how cellular components move is a fluorescence microscopy technique known as FRAP. While it can be used to study any number of cellular functions that involve movement, in this case we will focus on how it can be used to track movement within a membrane.

In this technique, scientists fluorescently label a specific membrane component (usually a lipid or protein). Then the high-powered laser of a confocal microscope is used to **photobleach** the fluorescent tag, meaning that the fluorescence of the tag is extinguished without damaging the protein or lipid to which it is attached. A small, specific area of the cell is bleached, and then we track how long it takes for unbleached molecules to return to the bleached area. There is an excellent explanation of this in Video 02-03.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=145#oembed-3

Video 02-03: Explanation of the technical aspects of fluorescence recovery after photobleaching and how to interpret the results of FRAP experiments.

The video does a better job of explaining, so we think you should go and watch that instead of trying

to imagine this graph in your head.
The video also shows us the kinds of results we can expect from these experiments. To summarize, graphs are generated from these data, with the fluorescence intensity in your region of interest on the Y-axis and time on the X-axis. Based on the amount that the fluorescence "recovers"—meaning the fluorescent intensity in that region increases again—researchers can compare how mobile a particular lipid or protein is within the membrane. We call this graph a **recovery curve**, as it shows how the fluorescence "recovers" in the region of interest over time.

TOPIC 2.3: STRUCTURE AND FUNCTION OF MEMBRANE PROTEINS

Learning Goals

- Explain how the primary sequence and the environment of a protein influence its final 3D structure, specifically with respect to the different types of intermolecular forces that it will form with itself and its environment.
- Distinguish between integral and peripheral proteins with regard to their solubility properties, structure, and manner of attachment to membranes.
- Describe some roles of glycoproteins within a biological membrane.
- Describe the structure and use of plasma membrane–adjacent structures like the animal cell extracellular matrix and the plant cell wall.

Earlier in this chapter, we learned that a biological membrane is different from a phospholipid bilayer. A biological membrane also has proteins embedded in it that will change its properties and add functions. In this topic, we will first review the formation of proteins from amino acids, then discuss the characteristics of membrane proteins. We will also examine the unique properties of the plasma membrane as the barrier that separates the cellular contents from the outside world. We end with a brief discussion of what's on the outside of a cell—namely, the extracellular matrix (animals) or a cell wall (plants, algae, and fungi).

The protein content of a membrane can range from ~25% (in the case of the myelin sheath cells of neurons) to nearly 80% (for the mitochondrial inner membrane), with more typical membranes consisting of protein and lipids in an approximately equal ratio (50:50) by weight.

Because of the associated proteins, the thickness of biological membranes is almost always thicker than that of a simple lipid bilayer. Biological membranes are typically 6.5 to 10 nm thick. A lipid bilayer without proteins is about 5.5 nm thick.

Brief Review of Amino Acids and the Chemistry of Protein Folding

We assume you have preexisting knowledge about amino acids and proteins, so we are approaching this as a review. As always, we encourage you to explore the resources in the introduction if you need a refresher.

There are 20 different types of amino acids that are used to form proteins. Each amino acid contains

an amino group (NH_3^+) and a carboxylic acid group (COO^-). In between is a carbon atom connected to a variable chemical structure called the **R group or side chain**. To make a protein, the amino group from one amino acid will covalently bond with the carboxylic acid group of another amino acid to create a *peptide bond*. This creates a repeating N–C–C pattern when the amino acids are strung together. This repeating N–C–C is called the **protein backbone**.

The sequence of the amino acids in a polypeptide is based on the sequence of codons that are read from the mRNA by the ribosome during translation. We call the order of the amino acids in the polypeptide chain the **primary structure** or primary sequence.

There are many, many versions of side chain charts on the internet for you to find, each designed with a slightly different learner in mind, which is why we have not included one here.

The R groups are what give each amino acid its distinct chemical properties. Some R groups are large, and some are small; some are negatively charged at physiological pH, and some are positively charged. It's worth taking some time to look at the chemical structure of the R groups and start investigating how these chemical groups impact protein structure/function.

The order of the amino acids in the primary structure plays a crucial role in determining folding and function of a protein. A given protein will fold in a very specific way depending on the molecular interactions of the amino acids in its primary structure. A variety of **intermolecular forces** (i.e., hydrogen bonds, ionic bonds, and induced dipole–induced dipole/van der Waal forces) contribute to the final 3D fold of the protein. Proteins will, many times, spontaneously fold into the shape that is the most stable and requires the least amount of energy to maintain. Other times, **chaperone proteins** will help facilitate the proper folding of a protein. We will learn about chaperones in more detail in Chapter 4.

When discussing protein folding, we split the different interactions into four categories, which we call the "levels" of protein folding. It's important to remember that these do *not* happen sequentially. Instead, they happen more or less all at once in different parts of the protein as it folds into its final 3D shape. These levels are as follows:

- **Primary structure** forms when amino acids are covalently bound together by peptide bonds. The order of the amino acids is the key feature of the primary structure.
- **Secondary structure** forms when the backbone interacts with itself via hydrogen bonds. It forms a repeating local pattern, commonly with nearby amino acids. Examples of this structure include **alpha helices** and **beta sheets**.
- **Tertiary structure** forms when R groups get involved. They interact with either other R groups or the backbone, usually via intermolecular forces like the ones mentioned above.
- **Quaternary structure** forms when different polypeptide chains come together to form a protein (or protein complex). All of the same intermolecular forces get used as we would expect in tertiary structure.

Disulfide bridges are the only other covalent bond used in folding aside from the peptide bond that creates the primary structure. Disulfide bridges form between two cysteine residues. They can play a role in tertiary or quaternary structure but are far less common than the other bonds. To compare, there are usually only a handful of disulfide bridges in a protein, if any, whereas there may be hundreds, or even thousands, of hydrogen bonds. As a result, we tend to discuss disulfide bridges much less often, but that doesn't mean that they should be forgotten.

The chemical environment (especially pH) also plays a very important role in protein folding. A soluble protein in the cytosol will experience different intermolecular forces than a protein that is embedded in a membrane. In each case, the protein will respond to its folding environment. Unlike the soluble proteins of the cytosol, many membrane proteins must be able to embed themselves in the lipid bilayer, including the nonpolar tail region. This requirement impacts how the protein folds and which amino acids of its primary structure are expected to be on the exterior in each region.

Integral and Peripheral Membrane Proteins

Proteins can be associated with the membrane in multiple ways. At the most fundamental level, proteins may

- exist solely on the surface of a membrane, living peripherally on one side or the other, or
- extend into the nonpolar tail region of the membrane in some way so that they are integrated into the membrane itself.

This may sound like a very simple thing to differentiate: Do they exist on the surface or extend into the membrane itself? However, it isn't always as easy as it sounds. For example, how do we categorize a protein that sits on the surface but also has a covalently linked lipid tail that extends into the membrane? (*Answer: We consider it to be integral.)*

Since proteins in membranes are too small for us to see, the categorization of proteins has historically been based on experimental evidence. So if we include the experimental evidence in our definition, we come up with the following:

- **Integral membrane proteins** are proteins that *cannot be removed from the membrane without destroying the membrane completely*. Usually, this requires the use of strong detergents, which disrupt the structure of the membrane so that proteins can be removed. Only proteins that extend into the membrane in some way will require this level of disruption to extract.
- On the other hand, **peripheral proteins** are much easier to remove. Usually, a simple ionic salt wash is enough to dislodge these proteins from the membrane, as the intermolecular forces are not as strong and can be more easily disrupted.

Table 02-01: The difference between integral and peripheral membrane proteins

INTEGRAL VS. PERIPHERAL MEMBRANE PROTEINS

Criterion	*Integral proteins*	*Peripheral proteins*
Requirements for removal from the membrane	Strong detergents or organic solvents used to break up hydrophobic associations	Mild detergents, high salt concentrations, and/or metal chelating agents used to remove
Association with lipids after extraction	Usually remain partially associated with lipids after removal	Usually completely free of lipids once removed
Solubility after extraction	Usually insoluble or clumped together in neutral aqueous solutions (e.g., water)	Soluble and dispersed in neutral aqueous solutions

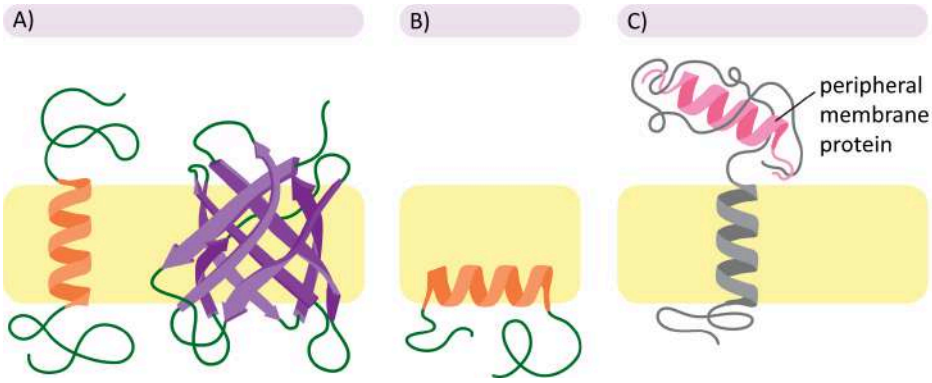


Figure 02-11: Examples of membrane proteins. (A) Integral transmembrane proteins. The protein on the left is a single alpha-helical region spanning the membrane. The protein on the right shows a beta barrel configuration. (B) An integral monolayer-associated protein. The alpha helix associates only with one leaflet of the membrane bilayer. (C) Peripheral membrane protein. The protein highlighted in pink is noncovalently attached to a membrane-anchored protein (gray). This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Figure 02-11 also helps illustrate the different subcategories of membrane proteins:

- **Transmembrane proteins** (Panel A): These integral membrane proteins cross the entire membrane and stick out on either side into the cytosol. These proteins are held in place primarily through the hydrophobic effect. More on that in the section below.
- **Monolayer-associated proteins** (Panel B): This is also a type of integral membrane protein. Although they are not considered to be very common, they still occur rarely. In this case, the protein is held in place by an alpha helix that is amphipathic (i.e., nonpolar on one side only). We will not be discussing this type further.
- **Peripheral membrane proteins** (Panel C): These proteins are most commonly (but not always) attached to the membrane through associations with integral membrane proteins. These linkages are most often via ionic or hydrogen bonding or some combination of several types of intermolecular forces. The association of peripheral proteins with one side of the membrane or the other further contributes to the asymmetry of the membrane.

Transmembrane Proteins Require Specific Secondary Structure to Be Able to Pass through the Membrane

One of the biggest challenges faced when proteins must pass through a membrane is how to deal with their backbones. Unlike the composition of R groups, the backbone of the polypeptide chain is always polar. The H, N, and O atoms that make up the peptide bond are electronegative and capable of forming hydrogen bonds. As such, they face a thermodynamic challenge when required to pass through the hydrophobic center of the lipid bilayer. So how is this addressed in the cell?

The answer lies in their secondary structure. As you may recall, secondary structures are defined as local, repeating structures that are formed via hydrogen bonding of backbone atoms to other backbone atoms. These repeating structures usually have a twofold effect:

- they allow the protein backbone to form hydrogen bonds with itself, which is thermodynamically stable, and
- they usually push the R groups outward, where they are available to interact directly with the environment. This leaves the protein backbone in the center, sequestered away from the

nonpolar lipid environment.

The two secondary structures we mentioned earlier are the alpha helix and the beta sheet. Both of these structures are very commonly found in transmembrane proteins. However, the details of how they form are naturally going to be different.

Alpha Helices in Transmembrane Proteins

Many transmembrane proteins have one or more alpha helices in their transmembrane domains. These alpha helices are stretches of about 20+ nonpolar amino acids (depending on the width of the membrane). Remember that in the alpha-helical arrangement, the amino acid R groups extend outward, and the backbone is in the center of the helix. By forming an alpha helix, the nonpolar R groups will shield the polar polypeptide backbone from the nonpolar environment in the center of a lipid bilayer, thus creating thermodynamic stability. Figure 02-12 shows an alpha helix from the side, embedded in a membrane. You can see the backbone structure represented by the purple ribbon winding up through the membrane. Hydrogen bonds between the atoms of the backbone (shown as dashed lines between the ribbon) hold the shape of the helix. The R groups of each amino acid in the polypeptide chain extend outward from the polypeptide backbone to interact with the tails of the phospholipids.

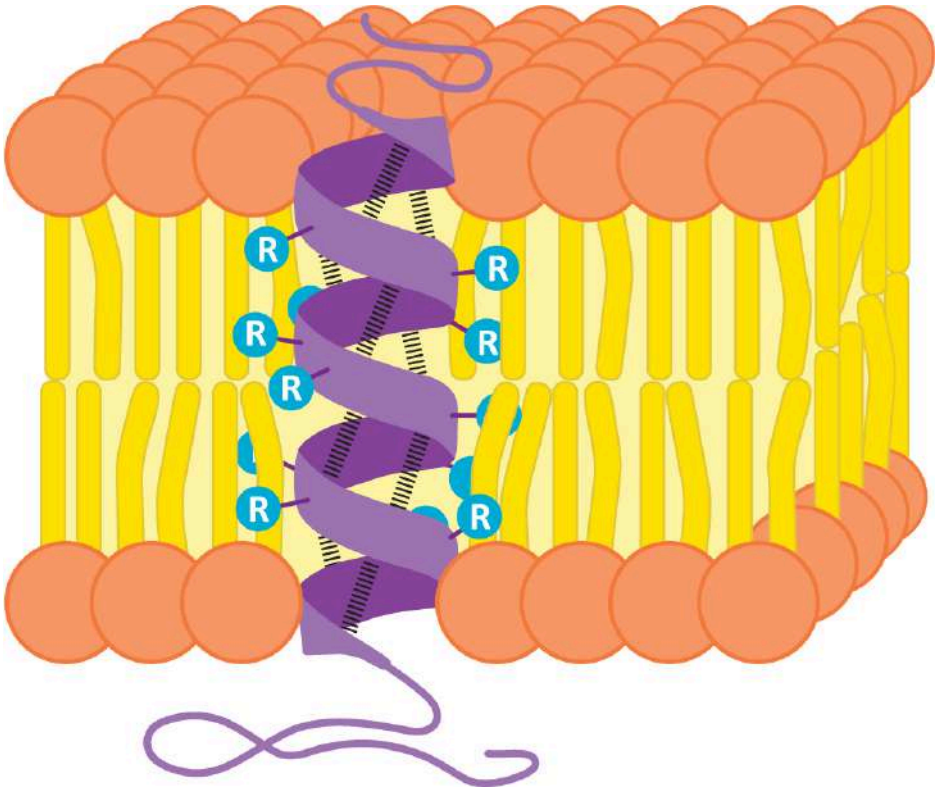


Figure 02-12: Integral membrane protein with a single alpha helix crossing the membrane. The polypeptide backbone curls into a helix (purple ribbon), while the R groups (blue circles) of each amino acid in the chain extend into the nonpolar portion of the membrane. Dashed lines represent the hydrogen bonds. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Since transmembrane proteins must be able to exist in both the nonpolar center of the membrane

and the aqueous parts of the cell, it stands to reason that these proteins are amphipathic like other components of the biological membranes. The R groups of amino acids within the nonpolar portion of the membrane are nonpolar, whereas in other regions that interact with the aqueous environments, the amino acids will primarily be polar.

It is important to note that *a single alpha helix does not form a channel, so it cannot allow anything to pass through it.* The molecules of the backbone completely fill the space inside the helix (Figure 02-13). Alpha helices can be used to form the pores and channels that control the entry/exit of many molecules in/out of the cell. Because a single helix cannot act as a channel, several membrane-spanning alpha helices must instead cluster in a roughly circular arrangement through the membrane.

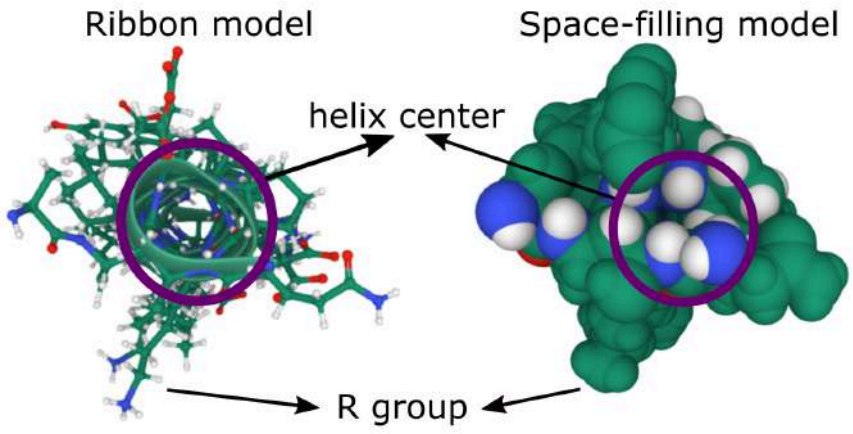


Figure 02-13: Different top-down views of a single alpha helix. Ribbon models of helices (shown on the left, with the R groups sticking out the "ribbon" path of the backbone) can give a false illusion that there is space inside the helix large enough for molecules to pass through. The space-filling model confirms that the internal area of the helix is filled by the space taken up by the atoms of the peptide backbone. This image is a derivative of 5EH6 created with NGL viewer by Dr. Lauren Dalton and is shared under a CC BY-SA 4.0 license.

Aquaporin is an excellent example of a transmembrane protein that is made of several alpha helices, which are used to create a central channel (Figure 02-14). Aquaporin is nonselective and allows water and other small solutes to pass through the membrane. Note that the amino acids lining the interior of the channel will need to be polar and possibly charged in order to interact with the water and solutes that are expected to pass through. On the other hand, the exterior of aquaporin will have a large strip that will directly interact with the phospholipid tails of the membrane; thus, those regions will need to be nonpolar.

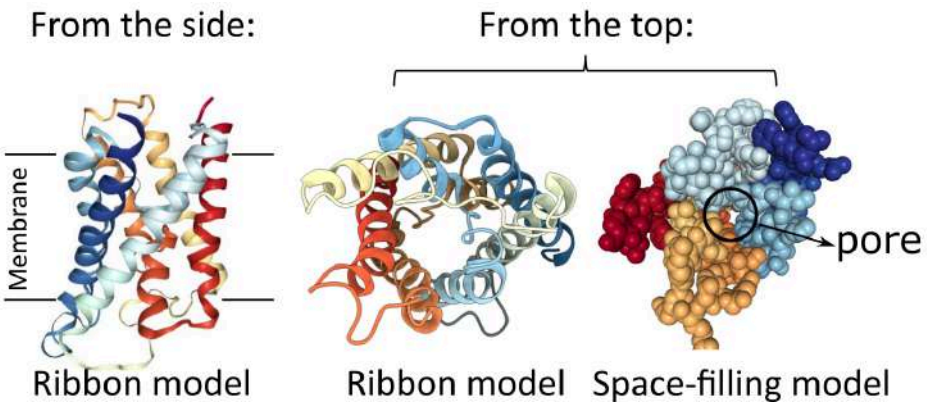


Figure 02-14: The protein structure of the water channel protein aquaporin, shown in multiple formats and angles. For simplicity, the R groups are not shown in the image, but they extend outward from the helical ribbon as we saw in Figure 02-13. The side view illustrates the orientation of the helical bundles to form a column through the membrane. The top views illustrate the circular arrangement that is made with enough room to form a small pore through the membrane. This image is a derivative of 1H6I created with NGL viewer by Dr. Lauren Dalton and is shared under a CC BY-SA 4.0 license.

Beta Barrels in Transmembrane Proteins

While a single alpha helix is stable enough to be used as the sole membrane-spanning structure for a transmembrane protein, a beta sheet is not as well designed for this purpose. The edges of the sheet will have exposed backbone molecules that will not easily interact with the nonpolar portion of the bilayer. However, a beta sheet can circularize itself by hydrogen bonding the ends of the sheet together (Figure 02-15). This means that the backbone's bonding requirements are met within the nonpolar region of the membrane, and we once again have an ideal scenario for a thermodynamically stable structure that can pass through a membrane. We call these circularized beta sheet structures **beta barrels**.

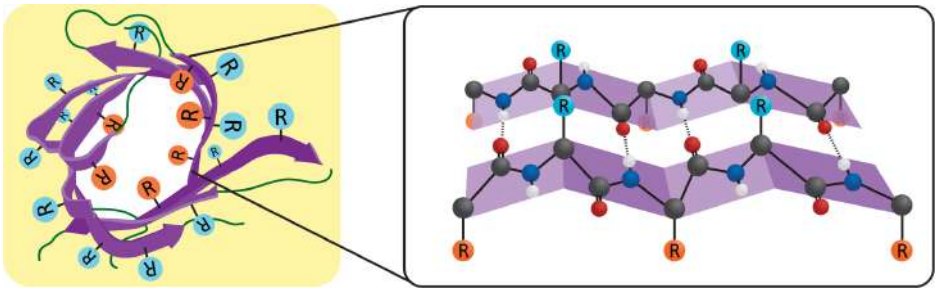


Figure 02-15: Beta sheet structure within a beta barrel. The image on the left is a beta barrel embedded in a membrane. The pore formed in the center is lined with polar R groups (orange), while the membrane side is lined with nonpolar R groups (blue). The box on the right is a magnified view to show further detail of the beta barrel section highlighting the peptide backbone interactions and their relationship to the R groups in a beta barrel. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Like the alpha helix, beta sheets and beta barrels require a specific arrangement of their R groups in order to form. Unlike the alpha helix, the R groups stick out perpendicular to the face of the beta sheet in an alternating pattern (Figure 02-15). This means that if we were to unfold the beta barrel and examine the primary sequence, we'd likely see a series of amino acids with alternating properties (i.e.,

nonpolar, polar, nonpolar, etc.) because one amino acid with an R group facing the lipid environment would be next to an amino acid with an R group facing the aqueous pore environment.

We can see this in action by examining the beta barrel structure of a real bacterial protein called outer membrane protein G (Figure 02-16). Once again, we see that this particular secondary structure, the beta barrel, precisely suits the function of the protein as a transporter. This protein is used to take up large carbohydrates. Thus, the outside of the beta barrel must be able to interact with the lipid environment and help hold the protein in the membrane, whereas the inside of the barrel must create enough space to allow specific molecules to pass through the membrane. In addition, the amino acids of the center channel must be polar as well to interact with the carbohydrates that are transported. In summary, the properties of the amino acids must match the chemical environment where they reside and function.

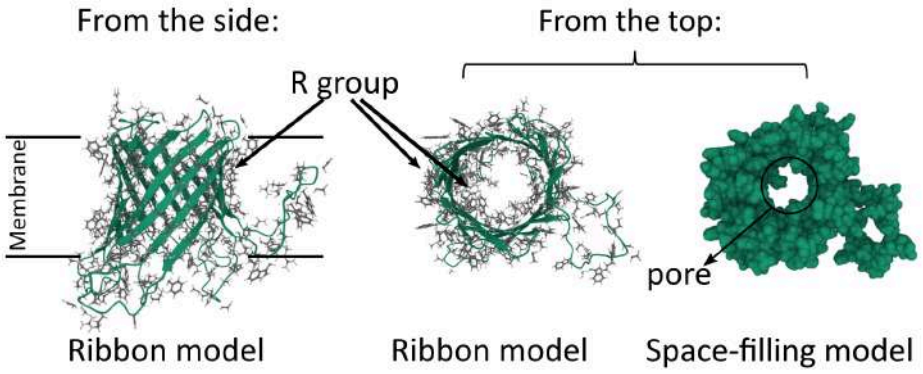


Figure 02-16: Multiple views of bacterial outer membrane protein G. The side view illustrates the orientation of the beta sheets to form a barrel structure through the membrane. The top views illustrate the circular arrangement that is made with enough room to form a small pore through the membrane. Note that R groups are removed from this image for simplicity. This image is a derivative of 2JQY created with NGL viewer by Dr. Lauren Dalton and is shared under a CC BY-SA 4.0 license.

Function of Membrane Proteins

Membrane proteins carry out many different functions in the cell. It is important to remember that membranes in the cell have *different* sets of proteins in them, as they each must carry different functions. As such, the protein composition of the endoplasmic reticulum (ER) membrane is different from, say, the plasma membrane.

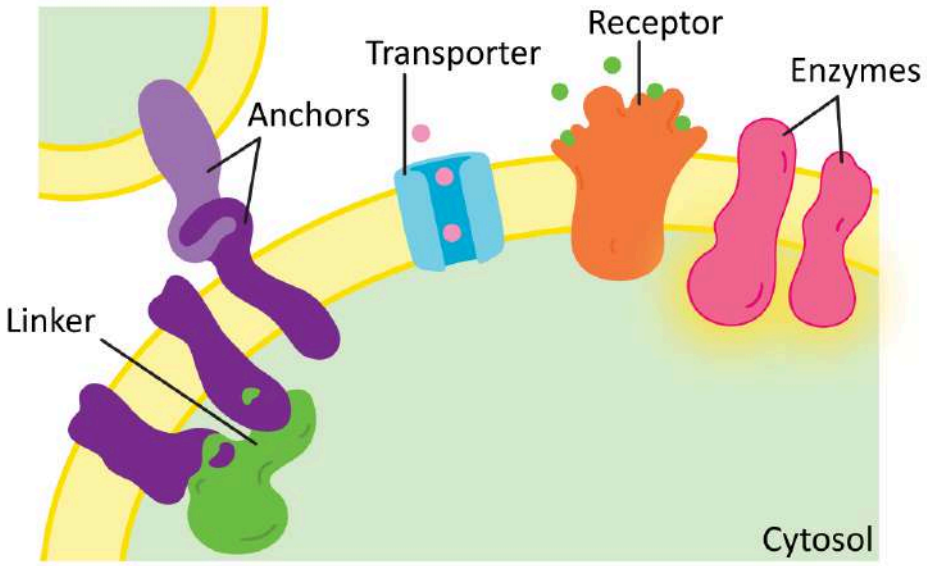


Figure 02-17: Examples of different functions of membrane proteins. Here we see linker proteins, anchors, transporters, receptors, and other enzymes that are embedded in the plasma membrane of a cell. See text for the details of the function of each of these. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

While there are many different functions for membrane proteins, for the most part they'll fall into one of the following categories (Figure 02-17):

- *Structural proteins, such as linkers and anchors*—Anchors help attach the membrane to organelles, the extracellular matrix, or even other cells, whereas linkers help connect several proteins in the membrane and can help provide shape. We will see many examples of structural proteins as we discuss cellular function.
- *Transporters*—These proteins mediate the transport of different types of molecules across the membrane in either direction. We saw examples of these in our earlier discussion of alpha helices and beta barrels. We will look at some examples, such as proton pumps and the ATP synthase.
- *Enzymes*—Many membrane proteins have enzymatic activity for a whole variety of cellular functions. We will explore many examples, including synthesizing or modifying enzymes, flippases, scramblases (shown earlier, in Figure 02-07), or kinases.
- *Receptors*—Receptors are key for the cell to be able to sense and respond to its environment. Receptors extend across the membrane and bind to small molecules or other proteins on the outside of the cell and in response initiate a chain of events leading to transmission of a signal inside the cell. Many receptors are also enzymes. We will look at receptors in more detail when we discuss cell signaling.

The Plasma Membrane: An Example of a Real Biological Membrane

The **plasma membrane** is the membrane that surrounds the cell and is the first point of contact between the cell and its environment. Red blood cells were one of the first cells to have their plasma membranes studied and as such have one of the most well-characterized plasma membranes. In most cell types, the plasma membrane is supported and shaped by a network of proteins. In red blood cells,

most of the protein network is inside the cell, just underneath the plasma membrane. However, in other cell types, this network can also be on the outside (via connections to the extracellular matrix or cell wall) or on both sides of the membrane.

Figure 02-18 shows a diagram of the plasma membrane of a red blood cell. In it, we see that the lipid bilayer of the plasma membrane is attached to an internal meshwork of spectrin protein filaments. These filaments provide a framework that supports the membrane and gives it its elasticity. You can also see that lipids and proteins on the extracellular side of the membrane are connected to carbohydrates (as indicated by the glyco- prefix). Since red blood cells move through the bloodstream, spectrin does not make any permanent connections with external structures.

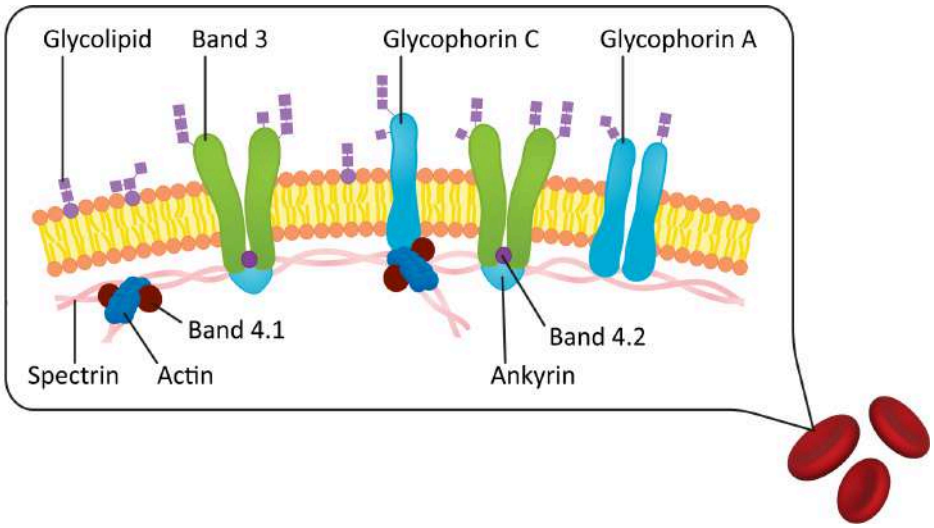


Figure 02-18: The plasma membrane of a red blood cell. There are a variety of different proteins found in the plasma membrane of the red blood cell. Carbohydrate groups (purple squares) are found on the exterior side of the membrane and contribute to cell identity. On the inside, we see the actin network, connected to spectrin and ankyrin. Together, these proteins help give the red blood cell its characteristic concave shape. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Plasma Membrane Carbohydrate Groups Are Found on the Outside of the Cell

Something that is important to point out in the plasma membrane figures in this chapter (see Figures 02-06, 02-10, and 02-18 as examples) is that the exterior of the plasma membrane is usually covered in a coating of carbohydrates. These carbohydrates are most commonly in the form of glycolipids and glycoproteins that are integrated directly into the plasma membrane. The carbohydrate component of the plasma membrane has a very important function. In biological systems, cells have an identity (and can be recognized) on the basis of the configuration of carbohydrate molecules on the surface of the membrane.

A classic example of this is your blood type. The cell surface polysaccharides carried by your red blood cells are genetically determined so that your body knows which cells belong to you and which are foreign. The ability to identify foreign cells is absolutely vital for your body so that it can recognize pathogens and other invaders. This also influences our ability to carry out blood transfusions when needed. Your "blood type" makes direct reference to the polysaccharides carried on the surface of your cells. The ABO system is the most well known; however, there are other cell surface signals that the body uses to identify which blood belongs to you, such as the Rhesus (Rh) factor. If you receive a

transfusion of blood that contains the incorrect polysaccharide markers on the cell surface, your body will produce antibodies to attack the blood and destroy it.

In some cell types (such as bacteria but also many eukaryotic cells), this carbohydrate coating on the plasma membrane is complex enough that it has its own name: the **glycocalyx**. Most examples of eukaryotic cells that have a glycocalyx are found in animals. This includes many epithelial cells (like the cells lining the gut and our blood vessels). Another really great example of a glycocalyx, which you can actually see with your naked eye, is the slimy coating on fish. The polysaccharide coating is found on virtually all fish, including the sockeye salmon shown in Figure 02-19. It plays multiple roles, including protection, cell-to-cell recognition, and even immune functions.

Figure 02-19: A male (bottom) and female (top) sockeye salmon (Oncorhynchus nerka). The glycocalyx is on the exterior of the body and adds to the shininess of the fish seen here. It feels slimy if you pick up a live or recently dead fish. This image originates from the National Digital Library of the United States Fish and Wildlife Service and is in the public domain.

Plasma Membrane–Adjacent Structures in Animal and Plant Cells

Cells in tissues are surrounded by a matrix of protein, polysaccharide, and fluid. The composition of the extracellular environment varies widely in different tissue types as well as in different organisms from different kingdoms. However, a common theme is that there are quite a lot of carbohydrates. The carbohydrates may or may not be associated with proteins, but collectively they form a three-dimensional network that connects cells together and provides a sort of hydration layer to trap water near the cells. Here we will highlight some of the key differences between this external environment of plant and animal cells.

In both plants and animals, the extracellular macromolecules are synthesized and secreted by the cells that live within them through a process called exocytosis, which we will cover in Chapter 4.

The Extracellular Matrix of Animal Cells

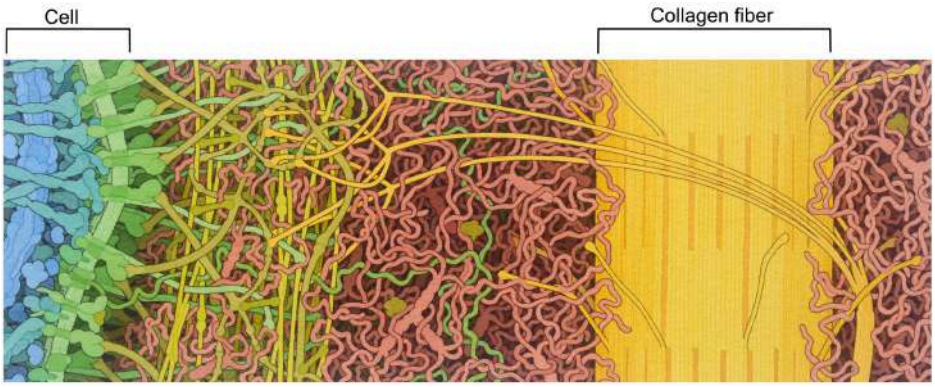


Figure 02-20: Artist's rendition of the mammalian extracellular matrix. Illustration by David S. Goodsell, RCSB Protein Data Bank, https://doi.org/10.2210/rcsb_pdb/goodsell-gallery-033. Shared under a CC BY 4.0 license.

The **extracellular matrix (ECM)** of animal cells is made primarily of **proteoglycans**. These are very similar to glycoproteins, except there is quite a lot more polysaccharide attached to a relatively small protein. The proteins and polysaccharides of the extracellular matrix are all connected to each other in a large 3D network, making it difficult to tell where one molecule ends and another begins. Polysaccharide chains known as **hyaluronic acid** are a major component of the animal extracellular matrix. Hyaluronic acid is important due to its gel-like properties. It helps trap water, which makes the entire ECM look and feel a bit like Jell-O. Thus, polysaccharides like hyaluronic acid help the extracellular matrix remain hydrated and, as such, resist compression. For this reason, it is a large component of the cartilage in our joints.

The most abundant protein in the mammalian extracellular matrix (not to mention the most abundant protein in the entire human body) is a very large protein called **collagen** (Figure 02-20). This protein is made of three intertwined polypeptide chains, which makes collagen very strong. Collagen provides structural support and helps with things like wound healing.

The Plant Cell Wall

The plant cell wall is also made mostly of carbohydrates. Interestingly, it contains many fewer proteins and is more rigid than an animal extracellular matrix. There is a cell wall that surrounds every single cell in a plant. So while it can be thought of as a type of extracellular matrix for the plant, its role is more complex than that of the animal extracellular matrix. It provides the structural support for the *whole* plant, similar to the role of the skeleton in animals.

Like the animal extracellular matrix, the plant cell wall is synthesized and secreted by the cell. Most of the polysaccharide components are synthesized by the Golgi apparatus; however, there is one notable exception. Cellulose is a very strong fiber of crystallized glucose chains that is synthesized in a unique structure called a **rosette** (Figure 02-21). This rosette is embedded in the plasma membrane, and it moves through the membrane as cellulose is synthesized. Microtubules have a role to play in the synthesis of cellulose and, as such, have a very different organization than we see in animal cells.

While there are a few proteoglycans and glycoproteins in the plant cell wall, the complex, branched polysaccharides are really the key players in both its structure and its function. In addition to cellulose, there are other long polysaccharide chains, such as hemicellulose (used to connect the cellulose

together) and pectin. Pectin acts in a similar way to hyaluronic acid in that it helps trap water and create a gel matrix. Incidentally, this is also why we use pectin to make jam.

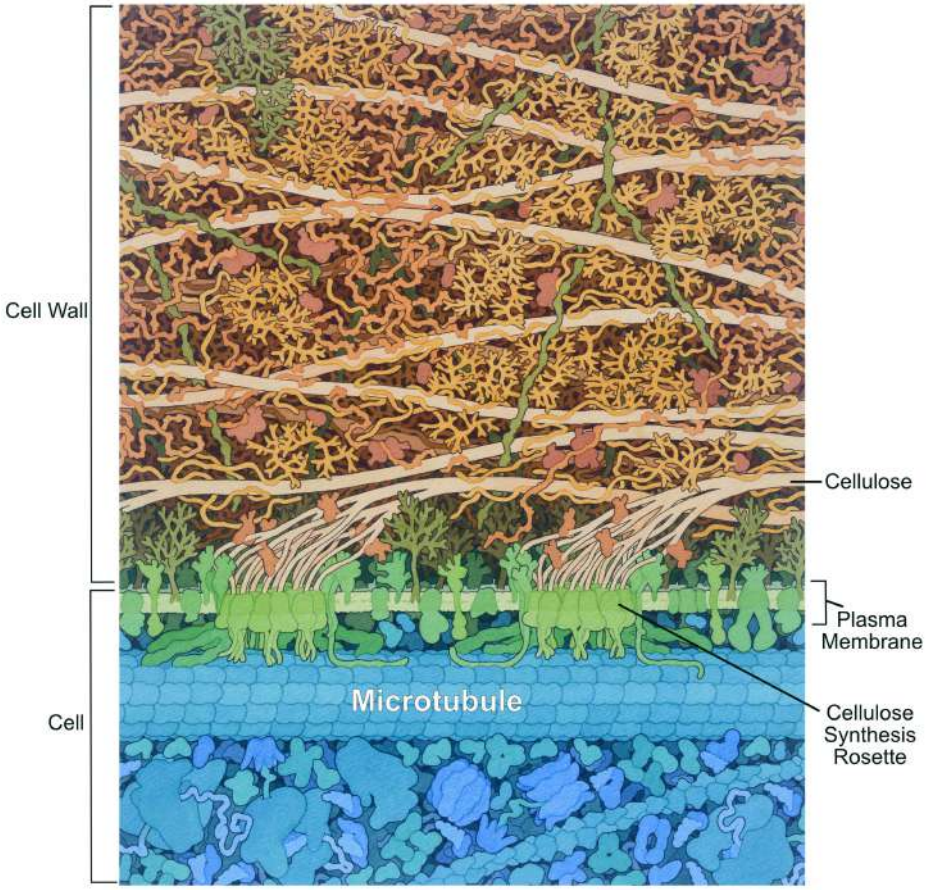


Figure 02-21: Artist's rendition of the plant cell wall and its underlying cell. Illustration by David S. Goodsell, RCSB Protein Data Bank, https://doi.org/10.2210/rcsb_pdb/goodsell-gallery-029. Shared under a CC BY 4.0 license.

TOPIC 2.4: PUTTING IT INTO PRACTICE DETERMINING PROTEIN LOCATION AND ORIENTATION

Learning Goals

- Explore the two most common experimental approaches to studying membrane proteins by analyzing data from representative examples:
  - hydropathy plots as an example of the bioinformatics approach to studying proteins
  - gel electrophoresis and SDS-PAGE as an example of a biochemical approach to studying proteins

Biological research changed quite drastically in 2000, when the human genome was completely sequenced (see the original publication of the human genome). Since then, our access to genetic information from a variety of species has increased much faster than we could have possibly imagined. This has fundamentally changed how we do research and has revolutionized how we explore cell biology. We now do much initial work **in silico** (meaning inside the computer) to first make solid predictions about protein function based on recognizable patterns from their primary sequence. To test the predictions made in *in silico* experiments, we turn to laboratory techniques that manipulate the DNA and proteins directly. In this section, we will explore one bioinformatic technique and one laboratory technique as examples of the many ways that scientists explore biological questions about membranes and their proteins.

Technique 1: Bioinformatics Approaches and the Hydropathy Plot

The amount of DNA sequence information that is openly accessible on the web continues to increase. The data sets can be quite large and can be difficult to handle without the help of computer programs and algorithms to organize, align, and make predictions about structure, function, and subcellular location of unknown proteins. The field of computer-based analysis of DNA and proteins has become so important that it has its own name—we call this **bioinformatics**.

Protein folding is subject to the same underlying chemistry and thermodynamics as everything else in the universe. This means that as long as we know the primary sequence of a protein, we can make predictions about structure, which can also tell us important information about protein function. These days we use computer algorithms and modeling software to learn as much as we can about a protein before we start to experiment in the lab. In some cases, lab experimentation on a protein is not possible, so computer predictions are our only option. The number of online apps and tools that are available (many for free) also continues to increase as biological computer scientists continue to build tools to answer questions. There are far too many of these tools to reasonably cover in any detail, so we'll focus on one of the most fundamental: the **hydropathy plot**.

A hydropathy plot is a bioinformatic tool that analyzes the sequence of amino acids in a protein and looks for specific patterns. Hydropathy plots can go by a few different names in the wider world (hydropathy index, hydrophilicity or hydrophobicity plots, etc.), but they all search for the same thing.

Specifically, hydropathy plots predict *alpha-helical transmembrane domains*. To do this, they look for a linear stretch of amino acids that are "hydrophobic enough" to allow them to exist stably within the confines of the nonpolar portion of the lipid bilayer. Usually, this equates to around 18–20 nonpolar amino acids in a row in the primary sequence. This is the approximate number of amino acids required to span the length of a typical membrane bilayer when adopting the alpha-helical arrangement. Video 02-04 explains hydropathy plots (called both hydropathy scales and indexes in this video) in more detail.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=145#oembed-4

Video 02-04: How to read a hydropathy plot.

The output of the computer analysis of the primary sequence takes the form of a graph (Figure 02-22), where

- the x-axis is the linear sequence of the protein, listed from N-terminus to C-terminus, and
- the y-axis is a number known as the **hydropathy index**.

The hydropathy index is a measure of the average hydrophobicity of a given amino acid residue (**hydropathy score**) *and* the average hydropathy score of a certain number of adjacent amino acids on either side combined into one value. The hydropathy score of each amino acid is a constant that has been determined based on computer-generated and experimental data. You can think of it as a number that measures how difficult it would be to immerse that amino acid into water. A higher hydropathy score indicates that it takes more energy to immerse that amino acid in water, which indicates it is more nonpolar than other amino acids.

At each point on the graph that the line is above 0, the amino acid and its neighbors are considered to be nonpolar, and below 0, the amino acid and its neighbors are polar on average. If the peak is high enough and wide enough, then a transmembrane alpha helix becomes a reasonable prediction. In this textbook, the "threshold hydrophobicity" is marked on the graph by a dashed line, and anything that crosses that line is considered to be "hydrophobic enough" to be a potential transmembrane region. Note that only major peaks are counted, not minor ones.

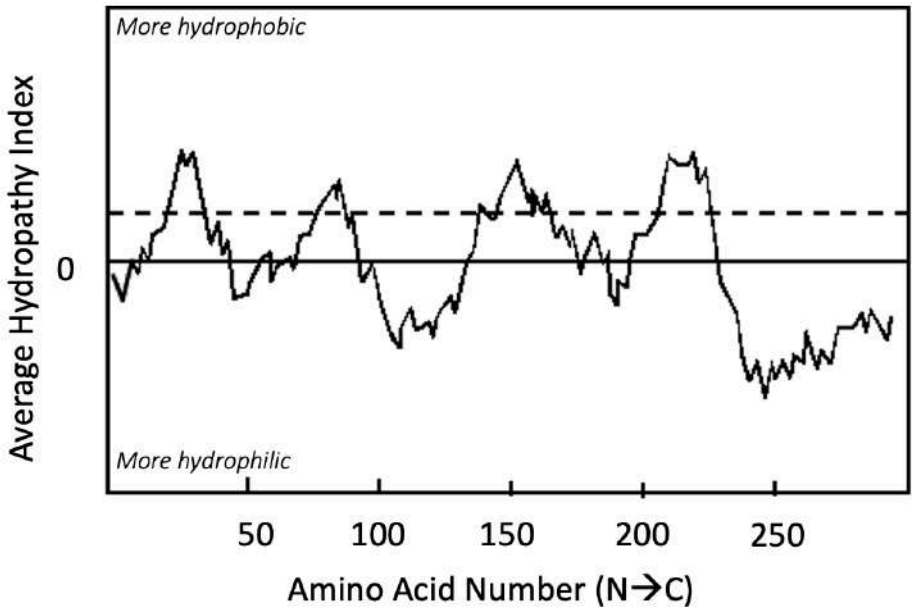


Figure 02-22: A representative graph showing the hydropathy plot of an unknown protein. See if you can figure out whether this protein has transmembrane domains and, if so, how many. (Hint: The answer is 4! Can you find the peaks? Are there also minor peaks that you should be ignoring?) Graph created by Robin Young.

It is worth emphasizing that the hydropathy plot can only make *predictions* about protein structure. In order to confirm the prediction, you would be expected to follow up your bioinformatic analysis with additional experimentation. It is entirely possible that even though this protein is predicted to be transmembrane, the reality could be quite different. The "transmembrane region" predicted

here could actually be a region sequestered inside of a soluble protein, or it may have a completely different function entirely. A good scientist will always follow up computer-based predictions with experiments to prove that the computer predictions are true. We will see examples where hydropathy plot predictions are, in fact, false when we discuss the endomembrane system in Chapter 4.

Technique 2: Biochemical Approach: Gel Electrophoresis

Despite the many advantages of bioinformatics approaches, there are also disadvantages that must be considered. For one thing, bioinformatics can only make predictions, and those predictions are limited to what we can learn from the amino acid, DNA, or RNA sequence. At some point, the scientist must test their bioinformatic predictions experimentally. One of the most commonly used techniques in cell and molecular biology is known as **gel electrophoresis**.

Gel electrophoresis is a way to separate macromolecules based on chemical features like size and charge. We can use gel electrophoresis to separate protein, DNA, or RNA (but not usually all at the same time!). In this technique, molecules are inserted into a gel matrix (not dissimilar from Jell-O but made from a different substance) that has pores of particular size. This gel matrix holds the sample and provides a filter to separate the molecules. The term *electrophoresis* refers to the electric field that is used to separate the molecules. Since the molecules carry a charge, the electric current causes them to move toward one side of the gel (see Video 02-05 below).

There are many different types of gel electrophoresis techniques available. Each one is named slightly differently to represent its unique features. Here's a short list of the most common types:

- **SDS-PAGE** (*S*odium *D*odecyl *S*ulfate *P*olyacrylamide *G*el *E*lectrophoresis) is used to separate proteins based almost exclusively on their molecular weight. This technique involves dissolving and denaturing proteins with the detergent SDS. This detergent adds a strong negative charge to the protein equal to the number of amino acids. This causes it to migrate to the positive pole during electrophoresis based on the size of the protein.
- **Agarose gels** can be used to separate nucleic acids. Nucleic acids naturally contain a strong negative charge on each residue, so detergents like SDS are not required. This consistent charge/mass ratio allows these molecules to travel to the positive electrode and separate based on size.
- **Native gels** do not denature molecules in contrast to SDS-PAGE. Thus, this technique exploits the natural charge carried by proteins, so the size, shape, and natural charge of the folded molecule also impact its movement in the gel.
- **2D gels** separate molecules based on molecular weight *and* isoelectric point.

It's worth noting that gel electrophoresis on its own doesn't tell us much. It is merely the tool we use to separate the molecules in our sample. It is important to know how the samples are treated before they are separated on the gel. In addition, in many samples, there are lots of proteins mixed together. Thus, often an additional technique is required afterward SDS-PAGE to highlight and/or identify the proteins of interest in the gel (i.e., staining, antibody labeling, mass spectrometry, etc.). We will first show an example where we can use gel electrophoresis to study proteins using SDS-PAGE. Afterward we will briefly touch on how it can be used for studying nucleic acids in preparation for Chapter 3.

SDS-PAGE is one of the more commonly used types of gel electrophoresis. Its name comes, in part, from the molecules used in the technique. A polyacrylamide gel is a specific type of gel that can be used to separate proteins primarily. SDS is a detergent that helps denature and separate the proteins in your sample. First, a brief video (Video 02-05) shows how the process works.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=145#oembed-5

Video 02-05: Unnarrated animation outlining the steps of the SDS-PAGE method.

As you can see, protein samples are first treated with the heat and the detergent SDS (Figure 02-23). The heat helps the proteins to unfold and become linear. The SDS then evenly coats the proteins, which helps them maintain their unfolded state and also adds a uniformly negative charge. Thus, all proteins end up completely unfolded, reducing any variation of travel based on the shape/compactness of their normal folding pattern. This makes it so that proteins are separated by their size and not some other characteristic.

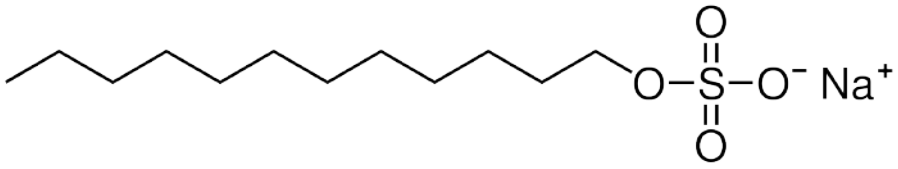


Figure 02-23: Chemical structure of sodium dodecyl sulfate (SDS). Image by Calvero is in the public domain.

Once the samples are heated and treated with SDS, they are added to the gel matrix using small wells that are set near the negative electrode (i.e., the *cathode*). The electric current is then turned on, and the proteins travel toward the positive electrode (i.e., the *anode*) at the other end of the gel. As the current runs through the gel, the proteins move slowly toward the positive electrode. Proteins that are larger will take longer to move through the gel matrix, and smaller proteins will travel farther through the gel. This way, the proteins in the sample are separated over time, with the larger proteins staying closer to the site or origin and small proteins moving farther away.

As mentioned at the start, SDS-PAGE on its own doesn't tell us much. It is merely a way to separate the proteins in a sample by molecular weight. The experiment you want to do will always include treatments of your samples *before* you do SDS-PAGE. The animation below (Video 02-06) shows an example of an experiment that uses SDS-PAGE to determine the type of membrane protein found in a sample of plasma membrane.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=145#oembed-6

Video 02-06: Animation explaining an SDS-PAGE experiment, from pretreatment of samples all the way to obtaining the results. Can you interpret the gel based on what you have learned?

DNA Gel Electrophoresis

While this chapter is not about nucleic acids and DNA, we wanted to add a brief discussion of how gel electrophoresis can be used to study nucleic acids, for two reasons:

1. This is, by far, the most common usage of gel electrophoresis in modern times.
2. The next chapter of this textbook looks at DNA and genomes in more detail, and you will need to understand how DNA gels work in order to be prepared for the techniques covered in Chapter 3.

Gel electrophoresis for separating DNA by size follows many of the same principles as SDS-PAGE, but with some differences:

- SDS is not required, as the phosphates in the backbone of DNA create a uniform negative charge on the molecule that is proportional to its size.
- The DNA gets loaded into a gel matrix made of agarose instead of polyacrylamide. DNA and RNA can be quite large—too large for a polyacrylamide matrix. The wider pore sizes in an agarose matrix accommodate the larger molecules.
  Agarose is a naturally occurring polysaccharide derived from sea kelp.

Like SDS-PAGE, an electric current is run through the gel once the samples are loaded, and the DNA will travel toward the positive terminal. As a result, just like in SDS-PAGE, the DNA will separate based on the size of the molecule such that bigger molecules will stay closer to the origin and smaller particles will move farther away. Video 02-07 is an excellent explanation of gel electrophoresis using agarose gels for DNA. Compare this technique to what you learned about SDS-PAGE above.

As you explore the techniques in Chapter 3, don't forget to come back here if you need to remember exactly how gel electrophoresis works.

One or more interactive elements has been excluded from this version of the text. You can view them online

here: https://open.oregonstate.education/cellbiology/?p=145#oembed-7

Video 02-07: Using DNA gel electrophoresis in an experiment.

CHAPTER SUMMARY

In this chapter, we've explored the structure and function of membranes. To do this, we first needed to explore some of the characteristics of the lipids and proteins that the membrane is made of. We identified four major characteristics of membranes:

1. The membrane is a bilayer made up of lipids and proteins.
2. The membrane is selectively permeable.
3. The membrane is organized but fluid.
4. The membrane is asymmetric.

We also learned that the fluidity of the membrane is dependent on its lipid composition—namely, the length and degree of unsaturation of the phospholipid tails and the cholesterol content. After that we discussed the difference between integral and peripheral proteins and how the secondary structure of membrane proteins contributes to its ability to span the entire membrane, including the hydrophobic portion of the membrane. We rounded out our discussion of membranes by briefly exploring the plasma membrane, which is a unique membrane in the cell, as it is in contact with the external environment. Outside of the plasma membrane is an additional structure known as the extracellular matrix in animal cells and the cell wall in plants and fungi. While these two have some similar roles, they are not the same.

Finally, we explored several experimental techniques throughout this chapter that can help us learn about membranes: fluorescence recovery after photobleaching (FRAP) helps us measure the movement of membrane components; hydropathy plots are a bioinformatic tool that predicts membrane-spanning alpha helices from the primary sequence of a protein, and SDS-PAGE and other kinds of gel electrophoresis can be used to learn more about DNA, RNA, and protein.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

Topic 2.1: The Chemical Features of Biological Membranes

1. For the following molecules, examine the molecular structures and label the following regions (note that they may not all exist in each of the molecules: phosphatidylcholine, cholesterol, generic glycolipid, sodium dodecyl sulfate [SDS]):

    a. the polar region (differentiate between charged and uncharged molecules of the polar region) and the nonpolar region

    b. a region that would be stiff and inflexible

    c. a glycerol residue (Some molecules have a serine residue instead; does yours?)

    d. a region that could easily have C=C bonds added (How would that affect the structure of that region?)

2. For the molecules you found in Question 1, identify the lipids that would be able to form lipid bilayers on their own. Use the details of the structures that you drew to explain why or why not.
3. Define the four levels of protein folding and explain how each one is stabilized.
4. Self-assembly of macromolecules is an important concept. What do you think that means?
5. Are disulfide bridges covalent or noncovalent interactions? How do they form? Why are they considered to be uncommon in the cytosol?
6. Find images of all 20 amino acids (don't memorize them!) online. Critically assess the structures and identify the following:

    a. The portion of the amino acid that is common to all of them.

    b. The portion that is unique to each amino acid, known as the R group or side chain.

    c. The functional groups that will form the peptide bond. Will anything be lost/gained during that reaction? How do these parts relate to the "N" and "C" terminus of a protein?

    d. The next questions are specifically for the R groups. Based solely on their structure, identify the following:

        - Any R groups that you would expect to have *acidic* properties. Will they gain or lose a proton during that reaction?
        - Any R groups that you would expect to have *basic* properties. Will they gain or lose a proton during that reaction?
        - Any R groups that would not be able to form H-bonds with water.
        - R groups that would form H-bonds but would not be acidic or basic.
        - R groups that could interact ionically with their neighbors.
        - Any R groups you would consider to be "big" or "small" relative to the others.

Topic 2.2: The Lipid Bilayer

1. What are the major differences between a synthetic phospholipid bilayer and a biological membrane?
2. How do cells adjust their membrane composition to maintain fluidity of their lipid bilayers in varying conditions?
3. Despite appearances, cholesterol cannot form bilayers on its own. Use the structure of the molecule to explain why.
4. Explain how the structure of phospholipids is the basis of the major properties of the bilayers that they form: physical form of the bilayer, self-sealing property, selective permeability, and fluidity of the bilayer.
5. Why are lateral movements of phospholipids in a bilayer so much easier than "flips" from one leaflet to the other? Explain why this means that we sometimes call a biological membrane a "two-dimensional fluid."
6. Describe how fluorescence recovery after photobleaching (FRAP) works and list the types of scientific questions that can be answered using this technique.

Topic 2.3: Membrane Proteins

1. What is the difference between integral and peripheral membrane proteins? Discuss and compare the different strategies for association of proteins with membranes.
2. What is a domain in a protein? How does it relate to structure and/or function?
3. Compare the hydrophobic forces that hold a membrane protein in the lipid bilayer to those that help proteins fold into a unique three-dimensional structure.
4. How do cells restrict the movement of membrane proteins? Outline the different strategies and provide brief examples.
5. What feature of a membrane's structure allows it to have separate identities and functions on each side?

Topic 2.4: Experimental Techniques in Membrane Biology

1. What is a hydropathy plot, and how does it help predict transmembrane regions?
2. Explain why hydropathy plots can only make predictions about transmembrane alpha helices and not beta barrels.
3. Explain why experimental techniques are needed to verify predictions from hydropathy plots.
4. Why is SDS needed for sample preparation in SDS-PAGE experiments?
5. What types of questions can be answered using SDS-PAGE?
6. How is DNA gel electrophoresis different from SDS-PAGE on proteins?

CHAPTER 3.

DNA, CHROMOSOMES, AND THE INTERPHASE NUCLEUS

INTRODUCTION

The structure and function of the interphase nucleus are fascinating and are a topic with a very long history. As one of the largest organelles in the eukaryotic cell, the nucleus was identified and characterized hundreds of years before we knew what DNA was. And yet the nucleus has also been a bit of a mystery. The DNA that is housed in the nucleus is so essential to function, it can be difficult to study. Perturbations of the nucleus can easily kill the cell, which does not help us learn about how it works. In recent years, advances in microscopy technology and bioinformatics have given us new options for studying the architecture of the nucleus and how the cell manages and controls all of that DNA.

We'll start with a discussion of how the DNA of the genome is organized and contained within the nucleus during interphase. Then we'll look at how the cell regulates gene expression through a combination of managing access to DNA by controlling chromatin structure, the use of transcription factors, and other mechanisms. Finally, we'll end by looking at the structure of the *rest* of the nucleus (i.e., the membrane, pores, and other components that help create this protective compartment around the genome) as well as how the cell controls what enters/leaves the nucleus in order to protect the DNA.

TOPIC 3.1: CHROMATIN AND CHROMOSOMES

Learning Goals

- Explain how intermolecular forces between specific proteins and DNA help form nucleosomes, chromatin loops, and ultimately interphase chromosomes.
- Compare and contrast euchromatin and heterochromatin, and explain how histone modification can act as a catalyst for chromatin remodeling to convert from one form to the other.

terminology check
Chromatin, Chromatid, and Chromosomes

Before we get started, it is absolutely vital that we briefly review what you already know about the structure of the eukaryotic genome from previous courses. The terminology is very confusing, as many of the terms are similar, so it's worth taking a moment to go over it. Note that quite a few of these terms will come up again in the chapter, and we will explain them in more detail. However, we

feel that an overview that helps you connect the terms together will be beneficial for you to refer to later.

The eukaryotic genome is made up of a number of **chromosomes**. In **diploid** organisms such as ourselves, there are two copies of each chromosome (for comparison, **haploid** organisms only have one copy of each chromosome). As a reminder, humans have 23 pairs of chromosomes. These "pairs" of chromosomes are similar in that they have all of the same genes in the same order, but they often carry different versions of the genes, which are known as **alleles**. One set of 23 chromosomes comes from our egg-bearing biological parent, and the other set we acquire from our sperm-bearing biological parent.

During most of the cell's life, each of these chromosomes will be made of a single **chromatid**, and that chromatid will exist as **chromatin**. Chromatin is a complex of DNA and proteins that helps keep the DNA organized inside the nucleus. If the cell plans to undergo meiosis or mitosis, then the DNA will be replicated so that each chromosome now is composed of two identical **sister chromatids**, which are exact copies of each other and are connected to each other via the **centromere**. In **interphase**, chromatin is in its more relaxed form, which allows access to the DNA. Despite this, not all of the DNA is equally relaxed: chromatin that is less condensed, allowing the genes in that area to be expressed, is called **euchromatin**. In contrast, **heterochromatin** is a form of chromatin that is less active and somewhat more compact.

Just prior to mitosis or meiosis, all nuclear function is shut down, and the chromatin takes on its most condensed conformation to form the characteristic **mitotic chromosomes** that we imagine in our heads when we think of chromosomes. They look like tiny fuzzy Xs and are often what we draw when asked to draw a chromosome. Interestingly, despite the mitotic chromosome being the image we associate with chromosomes in our head, the chromosomes only look like this during an extremely narrow window of time. The vast majority of the time, they exist as decondensed and unreplicated chromatin. It is the interphase form of chromatin that is our focus in this chapter.

Video 03-01 is an excellent video that helps clarify this terminology. We encourage you to click the link and watch the video.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=147#oembed-1

Video 03-01: The difference between chromatin, chromosomes, and other terminology of the cell cycle.

Chromatin Is Formed from DNA and Histones

As you can imagine, DNA is very small. The diameter of the DNA double helix is roughly 2 nm.

DNA

If you need a refresher on the structure of DNA, we suggest returning to the introduction and viewing some of the linked videos.

On the other hand, if you were to take all of the DNA in a human cell and line it up end to end, it would be over 7 feet long! That's taller than the average doorframe! Since the average nucleus in a human cell is only 6 μm in diameter (that's 6/1,000 of a millimeter!), the question of how the cell packs all of that DNA into such a small space arises. Not only must the DNA be packed into a tiny space, but it must be extremely organized so that each gene can be accessed quickly and accurately.

The first step to this organization is the formation of **chromatin**. Chromatin is formed soon after replication, when the DNA is carefully folded and organized, with the help of proteins that are ideally suited to this particular job. Much of the DNA will be packed up so that it cannot easily be accessed. However, when gene expression is required, specific regions of the packed DNA will be loosened (by shifting or removing some of the packing proteins) so that transcription factors, RNA polymerase, and other expression machinery can bind to the DNA and transcribe it.

Chromatin consists of DNA combined with two classes of proteins, which are known as **histones** and **nonhistone chromatin-associated proteins**. Here we will focus on the histones.

Histones are a set of proteins that interact strongly, but reversibly, with DNA. They are found in all eukaryotes, and even Archaea, but not bacteria. The functional importance of histones is reflected in how well conserved they are in different species. In fact, they are thought to be some of the most highly conserved proteins in all eukaryotes! They are considered to be "basic" proteins due to the overall positive charge (due to high **pKa**) they carry in their amino acid sequence.

pKa

pKa is a term you should have learned in general chemistry. See the introduction if you need a refresher on this topic.

This overall positive charge attracts the DNA due to the negative charge that is carried on the DNA backbone. There are five major types of histones that are used to help pack the DNA and produce chromatin:

- H2A, H2B, H3, and H4 are called the **core histones**. They interact strongly with each other to form a core complex. DNA wraps around the outside of this core protein complex.
- Histone H1 is a unique histone that binds to the outside of the nucleosome and helps pack the

nucleosomes together to tightly pack the DNA.

Most organisms have several different genes to represent each of the histone variants (H1, H2A, H2B, etc.). This allows for some mixing and matching of proteins in the histone core. As a result, the histone core can increase or decrease its affinity for the DNA, which, in turn, can alter the tightness of the packing of the DNA. As we will see throughout this chapter, how tightly packed the chromatin is has a significant impact on how and when specific genes are expressed. We'll discuss this idea in more detail later.

The Core Histones

The core histones come together in specific pairs to form a larger complex called an **octamer**. In total, there are two copies each of histone H2A, H2B, H3, and H4 in the octamer. They come together in a very precise arrangement (Figure 03-01). The core histone octamer interacts with the DNA in such a way that the DNA wraps around the outside of the octamer, similar to the way that thread wraps around a spool. The histone octamer with the DNA wrapped around it is known as the **nucleosome** (sometimes also called the nucleosome core). In between each of these nucleosomes, there is a stretch of DNA that we call **linker DNA**.

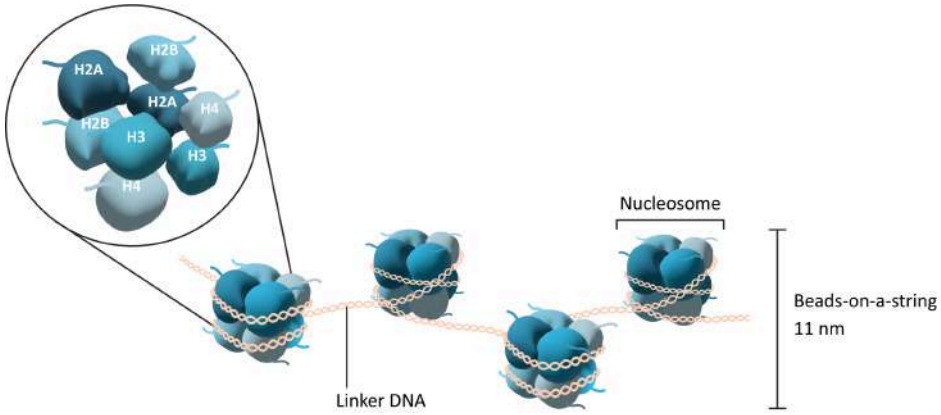


Figure 03-01: Nucleosomes wrapped around DNA. Two copies of the four core histone proteins H2A, H2B, H3, and H4 come together to form a nucleosome. DNA is wrapped around the outside of the protein octamer with a stretch of DNA called the linker DNA that connects the nucleosomes to each other. The tails of the core histones can be observed in this figure. This basic structure is approximately 11 nm in diameter. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Importantly, each of the core histone subunits has a short "tail" that sticks out and remains accessible even when the DNA is bound. This tail is a key feature of the histone, as it is an important site for modification and regulation of the histones as well as of chromatin structure more generally. We will see how these tails influence function in a moment.

The Linker Histone

In addition to the histone proteins in the core, a fifth histone family exists known as histone H1. Once again, there are several members of this family, each with their own specific use. The role of H1 is different from that of the other histones. H1 does not form part of the nucleosome core, but rather it sits on the surface of the nucleosome, on top of the DNA, and helps keep it in place. It also helps pull

in the linker DNA so that the chromatin is more tightly packed (Figure 03-02). Interestingly, the H1 histone is considered to be a highly dynamic protein in that it both spends most of its time bound to DNA and also shuffles around to different parts of the chromatin at a very high rate. The reasons for this are not entirely clear, as research on H1 in chromatin packing is ongoing.

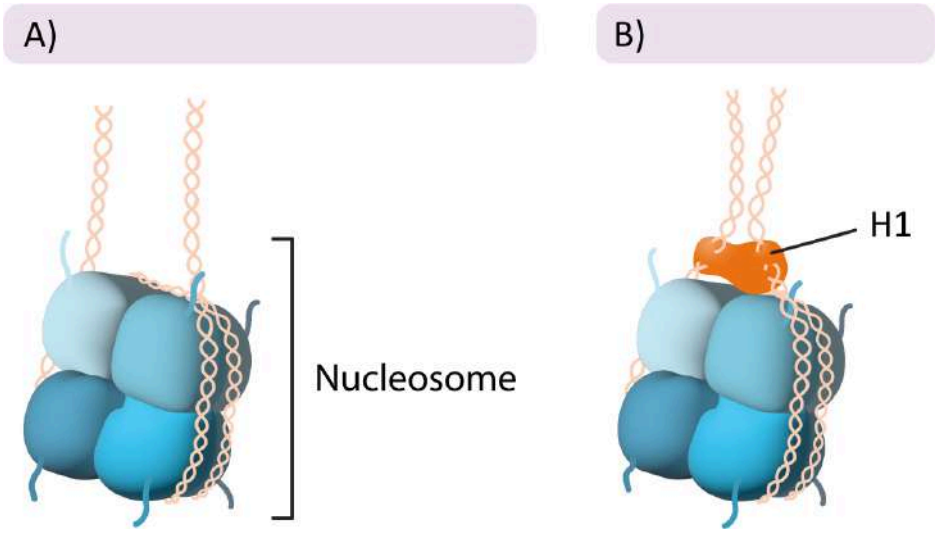


Figure 03-02: The position of the H1 histone on the nucleosome. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

DNA Packing in the Interphase Nucleus

The interphase nucleus is an extremely organized place. To fit all of that DNA into the nucleus in a way that allows efficient access to the required genes is no easy task. The chromatin helps with the packing and organization of the nucleus. Assembly of the histones and DNA into chromatin is very precise. We usually discuss chromatin formation as "levels" of packing of the DNA. These are as follows:

1. The initial association of the DNA with the histone octamers to form what we call the "beads-on-a-string" structure (Figure 03-01).
2. The nucleosomes pulled together to form a more tightly wound form of chromatin, called the chromatin fiber or 30 nm fiber (Figure 03-03).
3. Higher-order packing to form the most condensed forms, used for mitosis and meiosis (more on this in Chapter 8).

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=147#oembed-2

Video 03-02: How DNA is packaged. This video shows various levels of DNA packaging.

First Level of DNA Packing: "Beads-on-a-String," or the 11 nm Fiber

In humans, the amount of DNA associated with the nucleosome core is about 146 base pairs, with ~30–50 base pairs of DNA between nucleosomes. In this first level of packing, histone H1 is not present, the linker DNA is extended, and the nucleosomes are more distant from each other (Figure 03-01). This arrangement of DNA and histones has been given many names over the years:

- Most commonly it is called the "beads-on-a-string" model.
- Historically, it was also called the type A fiber, but that name is less common these days.
- We also call it the 10 or 11 nm fiber due to its measured diameter from transmission electron microscopy (TEM) images.

The addition of nucleosomes significantly reduces the length of the DNA. In fact, the original DNA is between 5.6 and 7 times longer than the beads-on-a-string version of the DNA. While the nucleosome formation is considered to be quite stable, it must also allow for changes and rearrangements so that the underlying genes can be accessed when needed. As such, nucleosomes can undergo a number of modifications to facilitate gene expression. For example, nucleosomes can be shifted along the DNA in a process known as **nucleosome sliding**. This is achieved by a group of proteins called **chromatin-remodeling complexes**, which we will discuss later in this chapter.

Nucleosomes are formed as soon as the DNA is replicated, using a combination of preexisting histones from the old DNA strand and newly synthesized subunits. The preexisting histones will already carry specific modifications on their histone tails, which can influence the structure and function of the newly synthesized DNA. Since the newly synthesized DNA is most likely destined to be passed on to a new cell via mitosis or meiosis, this helps, in part, to create a chemical "memory" for the chromatin that is passed on to the new cell.

Second Level of DNA Packing: The Chromatin Fiber (Sometimes Called the 30 nm Fiber)

Re: Chromatin Fiber

This terminology is obviously messy, as scientists both refer to the general structure of DNA + protein as *chromatin* but also have specifically given the 30 nm fiber the name *"chromatin"* (short for *interphase chromatin*). We will use the term *"chromatin fiber"* when talking about the 30 nm fiber to stay somewhat in line with current naming conventions in the literature and also to distinguish from the broader definition of chromatin (used in textbooks like this), which refers to any association of histone and DNA.

Further condensation of the nucleosomes occurs using the linker histone, H1. As mentioned earlier, H1 sits on the outside of the nucleosome and helps hold the DNA in place (see Figure 03-02). As it binds to the outside of the nucleosome, it also pulls adjacent nucleosomes together, thus forcing

the 11 nm fiber into a loose spiral, which can be observed in Figure 03-03 and Video 03-02. This is the form of DNA fiber that we will call the **chromatin fiber**, but again it has several names, including interphase chromatin (often shortened to just chromatin), the 30 nm fiber (due to its average diameter), and the type B fiber (again, this name is no longer very common). This fiber has a packing ratio of ~50:1 (meaning that the original DNA strand is roughly 50 times longer than the packed DNA!), and it is only about 30% DNA—the rest is composed of packing proteins. The chromatin fiber is the form of DNA that is found in the nucleus throughout interphase.

An active interphase genome will naturally show variation in how the genome is packed in different regions—some sections will be packed away tightly (like structural components and genes that are not currently being expressed), and other sections will be more open so that gene expression can take place. This means that even though we discuss this chromatin fiber as if it is always the same, for the purposes of explaining how packing works, we must also remember that the levels of DNA packing are a little more nuanced. We will explain more when we discuss **euchromatin** and **heterochromatin**.

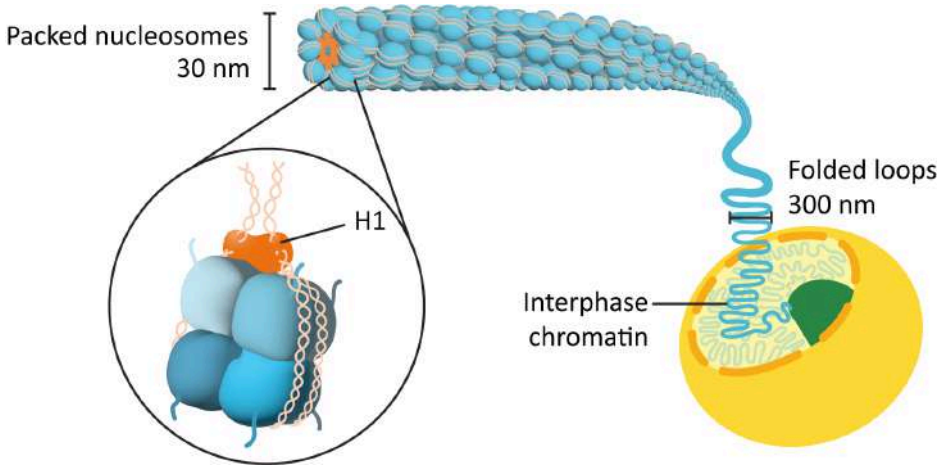


Figure 03-03: DNA in the interphase nucleus is organized and packaged. First, the DNA is wrapped around core histones to form nucleosomes. H1 then helps loop nucleosomes together into a fiber, which then can be further looped and packaged inside the nucleus in a highly organized manner. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Higher-Order Packing

Even though interphase chromatin is well packed compared to the original DNA strand, this is not the end of it. In interphase, when the genome is active, additional packing of the genome must be done in such a way that the genes present on the DNA are taken into account. This is required so that each gene can be easily accessed when needed. As an analogy, you would not store your bike in a box at the back of your basement, behind many other boxes, if you're going to use it every day. That would not be efficient at all. Generally, you store your bike in a way that's accessible so that it is ready and available when you want to ride it, like in the garage or some other accessible area in or around your home.

Much like your bike, the genome is also packed in such a way that the genes are easy to access when needed. The cell uses a variety of **structural maintenance of chromosome**, or SMC, complexes to help loop and organize the chromatin. While these complexes are made of several proteins, one of the most famous ones is **cohesin**. Cohesin was first discovered in mitosis, as it is used to hold sister

chromatids together (which will be explained in Chapter 8). We are now learning that cohesin has an important role to play in interphase as well (see Skibbens, 2019 for a recent review article). Cohesins are thought to bind to the DNA in a sequence-specific way and then help with the formation of loops that contain one or several genes within them. Loops can then be brought together into either actively expressing **topologically associated domains (TADs)** or genetically inactive regions that are part of the three-dimensional organization of the genome.

During mitosis (and meiosis, which we do not discuss in this textbook), we see the most extreme levels of DNA packing. Each of the chromosomes of the cell must condense itself into the tightest conformation possible and then have its sister chromatids separated into two newly forming daughter cells. Mitotic chromosomes are between 20,000 and 50,000 times shorter than the original DNA strand. No gene expression can happen during this time, which puts the cell at risk, so mitosis is completed as quickly and efficiently as possible. Again, the SMC complexes get involved. In addition to the cohesin and SMC complexes, we also have another protein type called **condensin**. We will explore the details of how chromosomes prepare for mitosis in detail in Chapter 8.

Euchromatin, Heterochromatin, and the Organization of the Nuclear Genome

There is a lot going on in the interphase nucleus. Some genes are being actively expressed, while others are being actively repressed and put away. Some parts of the genome don't carry genes at all but instead are important structural regions that are needed to protect the DNA and help with mitosis. All of this is housed in an extremely tiny cellular compartment, the nucleus. If the cell undergoes DNA replication, then that tiny compartment gets even more cramped. On top of that, Eukaryotic genomes tend to be quite a bit larger than their prokaryotic counterparts. The largest genome discovered so far is that of the herbaceous plant *Paris japonica*, which has a whopping 150 Gbp of DNA in its genome! For comparison, the human genome is only 3.2 Gbp. All of this is to say that it is absolutely vital that the contents of the nucleus remain as organized as possible at all times.

Euchromatin versus Heterochromatin

There are a couple of ways the cell manages to maximize space in the nucleus, the first of which is to ensure that only the DNA that is currently needed is unpacked enough to allow for gene expression, and everything else is tightly packed away. This leads to differences in the packing of interphase chromatin in different regions of the nucleus. The packing differences can actually be pronounced enough that it changes how the chromatin looks in electron microscopy. Scientists named these different forms of chromatin **euchromatin** and **heterochromatin**, based on how they looked in an electron microscope, when they were first observed in the 1940s and 1950s. At this point, we did not have a clear understanding of the role of euchromatin and heterochromatin in the cell. The TEM micrograph in Figure 03-04 shows regions of darkly staining material and lightly staining material inside the nucleus. The darkly stained material is the heterochromatin, and the lightly stained material is euchromatin.

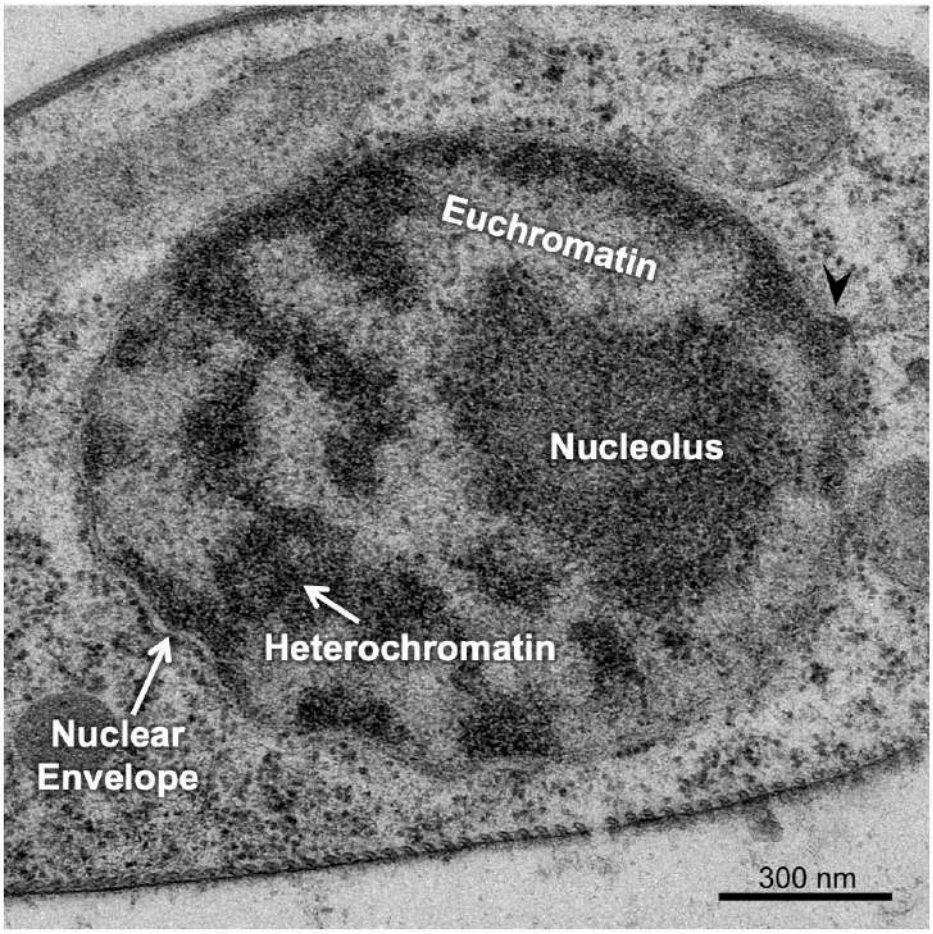


Figure 03-04: TEM micrograph showing the nucleus of the parasitic protozoan Trypanosoma brucei. Major structures of the nucleus are labeled, including the difference between euchromatin and heterochromatin. Attribution: Johanna Höög, Dimitra Panagaki, Jacob Croft (2020), CIL:50940, Trypanosoma (Trypanozoon) brucei, T. brucei cells. CIL. Dataset. https://doi.org/10.7295/W9CIL50940. Licensed under CC-BY 3.0.

Since its identification using electron microscopy, we have learned quite a bit more about the structure and function of heterochromatin and euchromatin, though there is still much work to be done. Here's what we know so far:

- **Heterochromatin** is the more tightly packed of the two forms of chromatin. It is the form we described earlier that has the H1 histone bound to it so that the nucleosomes form a spiral and pack together tightly (see Figure 03-03). Heterochromatin does not allow proteins like transcription factors or polymerases to access the DNA. As a result, in these regions *no gene expression can take place*. At any given moment, most of the chromatin in a cell is in the form of heterochromatin. However, we differentiate between different "types" of heterochromatin:
    - **Constitutive heterochromatin** is found in regions of the DNA that are structural, such as telomeres and/or centromeres. These regions of the DNA never really need to be unpacked, as there are no genes there. Thus, the chromatin stays tightly packed up all of the time so that it takes up less space.
    - **Facultative heterochromatin**, on the other hand, is found in parts of the genome where genes do exist, but they are not currently needed by the cell. These parts are also

tightly packed, but if the cell requires one of the genes in this region, it will unpack the DNA to allow for transcription to take place. Thus, these regions may be more dynamic, packing and unpacking as required.

- **Euchromatin** is the less-condensed form of chromatin. These regions may have some or all of the histones removed so that the DNA can be accessed. Active transcription is very likely taking place in these regions as well as other forms of gene regulation. When the genes in these regions are no longer required, they will be packed back up into facultative heterochromatin until they are needed next. These regions of the DNA are considered to be very active and dynamic.

As can be seen in Figure 03-05, any given chromosome can have both euchromatin and heterochromatin existing in distinct regions. The placement of these regions can also change over the life of the cell depending on the types of genes or structural elements located within a particular chromosome region. Based on Figure 03-05, try to identify which of the heterochromatin areas you would expect to be constitutive and facultative.

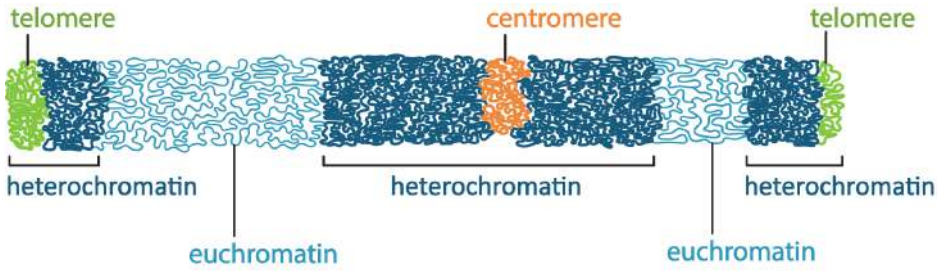


Figure 03-05: Schematic showing structural elements found in a typical chromosome. Every chromosome contains structural elements such as the telomere and centromere that remain packaged into heterochromatin. In addition, genetic regions can be either loosely packed as euchromatin or more densely packed as heterochromatin. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

There is one more thing to note before we move on. While all cells will have regions of euchromatin and heterochromatin within the nucleus, the placement of those regions is not always the same from one cell to another. In any given cell, at any given moment, it will be expressing a specific subset of genes. The subset of genes being expressed may or may not be the same as a different cell. This is especially true of cells in different tissue types. A cell of the pancreas synthesizing and secreting digestive enzymes will be expressing a very different set of genes than a neuronal cell, for example. In addition, cells will change the genes that they need to express over time. There are a number of genes that are only turned on during embryonic development and then get turned off. Mitosis also requires a specific set of genes that must be expressed to prepare for mitosis and then be turned off again. All of this has the potential to result in shifts in the parts of the genome that are more or less accessible, which, in turn, will change the regions that are packed as euchromatin or as facultative heterochromatin (since structural regions don't have genes, constitutive heterochromatin is less likely to change from cell to cell).

Three-Dimensional Organization of the Genome within the Nucleus

It's worth taking a moment to stop and consider, once again, the incredibly complex job of the nucleus

as the home for the cell's DNA. As an example, the human genome consists of 23 chromosomal pairs, which include 21,000 protein coding genes (protein coding genes are thought to make up ~1.5% of the human genome), additional important noncoding regions, and 3.2 Gbp of DNA. To translate this into terms that are easier to comprehend, the average human chromosome is a piece of DNA that's about 5 cm long (and 2 nm in diameter). If we do the math, our 46 chromosomes equate to over 2 m (or well over 7 feet) of DNA that gets stuffed into each nucleus. And of course, after DNA replication, that number doubles. That's a lot of DNA to protect and organize!

It stands to reason that the nucleus would be an incredibly organized space when you think about it in those terms. However, science has had a difficult time unraveling the mysteries of the nucleus, especially the arrangement of the complete chromosomes within the three-dimensional space of the nucleus. Since the nucleus is large enough to be visible with a light microscope, we've been able to observe it for a long time. It was first "discovered" and named in the 1830s, and we started to see initial inklings of nuclear organization. With only transmitted forms of light microscopy, however, our view was limited to more obvious structural features. For example, the nucleolus was identified in 1836. Also, the changes that occur during mitosis showed us that the genome is split into many paired chromosomes. The invention of TEM in the 1940s allowed us to see heterochromatin and euchromatin, but it wasn't yet clear how that was related to genome organization. It wasn't until the 1980s, when fluorescence microscopy was invented, that we were able to identify that the various chromosomes that make up the genome exist in discrete territories within the nucleus. The early 2000s brought not only the complete sequencing of the human genome but additional technical advances that allowed us to explore the physical 3D arrangement of the genome as it exists within the nucleus. Technical advances such as these have really blown this area of research wide open, and we are learning more every day about how the nucleus manages such vast amounts of DNA. In this section, we'll try to summarize what we know about the spatial organization of the nucleus *so far*, but you should expect that the science will advance past this more quickly than we'll be able to add to / update the material in this textbook.

When we discuss the organization of the genome, we start from the naked DNA strand. This means everything we've learned so far contributes to how the DNA is organized within the interphase nucleus. To summarize briefly,

- the negatively charged DNA associates with the positively charged core histone complexes to form nucleosomes (Figure 03-01),
- the H1 histone helps further pack the DNA to form chromatin (Figures 03-02 and 03-03),
- the chromatin is looped with the help of cohesins and the SMC complexes, and
- the loops are brought together to form TADs.

At this point, we're going to look at the effects of chromatin looping and the formation of TADs on the organization of the genome in more detail.

As mentioned earlier, fluorescence microscopy showed us that each of the chromosomes of the genome exists in its own discrete space within the nucleus, known as a **chromosome territory**. We also know that each chromosome will consist of a combination of euchromatin and heterochromatin (see Figure 03-05) that needs to be functionally organized. To this end, each chromosome is organized physically into what is known as **A/B compartments**. They break down into the following:

- The A compartment, which tends to be located closer to the center of the nucleus, contains many more genes than the B compartment. All of the genes actively undergoing transcription, or those waiting for their turn to be actively transcribed, will be found in the A compartment. TADs are primarily observed in the A compartment as a result. Also, the A compartment will be made primarily of euchromatin.
- Conversely, the B compartment tends to contain mostly constitutive heterochromatin and genes that have been inactivated in a more permanent way (as they will never be needed). The DNA in the B compartment is also more likely to make physical connections to the nuclear envelope (via the **nuclear lamins**), which will be discussed later in this chapter.

The DNA in each of the compartments also tends to physically interact more with the DNA within the same compartment rather than the DNA in a different compartment. Figure 03-06 summarizes the spatial organization of the nucleus that we have described here.

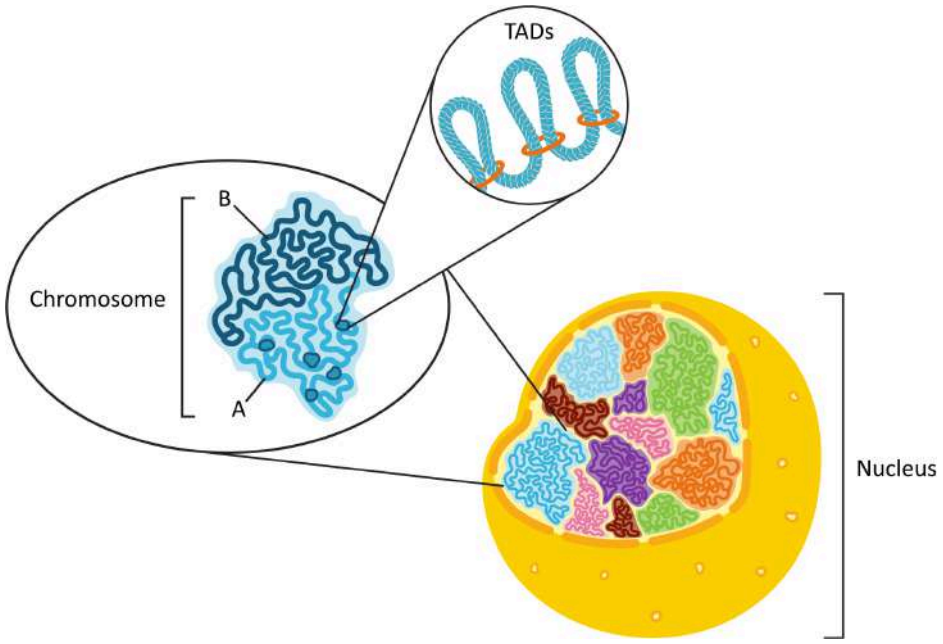


Figure 03-06: Chromosomal organization within the nucleus. Chromosomes inhabit specific spaces within the nucleus, which are indicated by the different colors inside the nucleus in the image. Further, each chromosome is organized into A and B compartments depending on their transcriptional activity. A is more transcriptionally active than B, and TADs can be observed. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

On top of all the organization already described here, there are also a number of **nuclear bodies** that can be observed within the nucleus. The largest and most well known of these is the **nucleolus**, which we will discuss in detail later in this chapter, but there are other nuclear bodies that have been identified as well. Examples include Cajal bodies, PML bodies, speckles, paraspeckles, PIKA bodies, and more. While the exact function of each of these nuclear bodies is still unclear, they are generally thought to be involved in specialized nuclear function (ribosome production, replication, transcription, splicing, repair, etc.). Usually, these nuclear bodies are identified via fluorescence light microscopy, whereas the compartmentalization described above was discovered via other means. As a result, it is also unclear exactly how these bodies correlate with what we have learned about the spatial

organization of the nucleus so far, but it is still worth remembering that they exist. If nothing else, this helps remind us of how much we have left to learn about cells and their function.

TOPIC 3.2: REGULATION OF GENE EXPRESSION

Learning Goals

- Discuss the different types of DNA and histone modifications and their roles in chromatin remodeling and regulation of gene expression.
- Distinguish between the different types of transcription factors (e.g., basal, activators, and repressors) and explain how their interaction with specific regulatory regions of DNA can influence transcription.
- Discuss the mechanism and specificity of mRNA splicing events and explain how alternative splicing increases the diversity of protein products encoded from a single gene.
- Explain how mRNA processing events (e.g., cap, tail, splicing sites) are used to identify mature, functional RNA ready for export.
- Explain how chromatin immunoprecipitation (ChIP) can be used to answer scientific questions about genome structure and regulation of transcription.

Introduction

The evolution of Eukaryotes brought with it many changes to organismal form and function. As a general rule, eukaryotic cells are larger and more complex than either bacteria or Archaea. Eukaryotes are also frequently multicellular, which creates options for cell and tissue specialization that are not possible in a single-celled organism. The increase in complexity that comes with multicellularity also required an evolution of the eukaryotic genome. As a result, the eukaryotic genome tends to be quite a bit larger than its bacterial counterparts. While there are surely a number of reasons for this, one reason may be that a larger, multicellular organism is simply going to require more genes to run it than a smaller, single-celled organism. Specialized cells and tissues will require different subsets of genes to be active to support their needs. Multicellular organisms are also more complex to build compared to single-celled organisms, so development will require a number of specific genes that are dedicated to that purpose and then are no longer required. Together, these both point to the requirement for more extensive and nuanced regulation of gene expression compared to our bacterial and Archaeal counterparts.

In a multicellular organism, it is very likely that any particular cell will carry more than a few genes in their genome that they will never need to express due to their particular specialization. Even more genes will only be needed in specific situations or as a result of specific environmental and/or developmental cues. The result of this is that the eukaryotic cell must have the ability to precisely decide which genes to express when and to turn off all the genes that are not required at that time.

It's worth remembering that genes and their gene products are heavily regulated at *every stage* of their life cycles. The cell determines not only when they will be transcribed and translated but also

at what speed this will happen and for how long. Once the proteins have been synthesized, they continue to be regulated through chemical modifications, such as phosphorylation, cleavage, and so on. Even the decision of when to destroy a protein is one that is highly controlled. Figure 03-07 does an excellent job of showing how genes and proteins are regulated. Some examples from this figure to highlight include the following:

- *Transcriptional control* determines when and how often genes are transcribed, whereas
- *RNA processing control* determines which combinations of introns/exons are produced, so different proteins can be made from the same gene.
- Once the processed mRNA leaves the nucleus, the cell continues to regulate the gene products by controlling
    - when and how translation happens,
    - when the mRNA is degraded,
    - what kinds of posttranslational modifications take place, and ultimately,
    - when the protein is tagged for destruction.

In this topic, we focus on the ways that genes can be regulated within the nucleus only—in other words, control of when and how transcription happens and how the RNA is processed prior to export into the cytosol. In later chapters, we will see some examples of posttranslational control mechanisms.

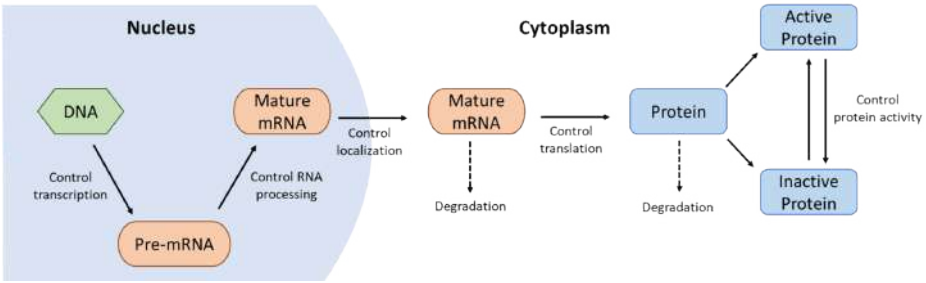


Figure 03-07: Regulation of genes and gene product flow chart. Boxes indicate the type of biomolecule (DNA in hexagons, RNA in ovals, and proteins in rounded boxes), and the labeling refers to the type of regulation that can occur in that step. This image was created by Dr. Lauren Dalton and is shared under a CC BY-SA 4.0 license.

Transcriptional Control: Chromatin Remodeling Allows Access to the Gene

As we've alluded to more than once in this chapter, one of the ways that eukaryotic cells deal with the overwhelming amount of DNA in their nuclei is by keeping anything they're not currently using packed away tightly in the form of heterochromatin. As a result, there are many gene-coding regions of the DNA that will undergo rounds of packing and unpacking as the needs of the cell change. Changing the packing level of the chromatin not only saves space but can be used by the cell as a form of regulation. By packing DNA tightly (or not, as the case may be), the cell can influence how *accessible* genes are to the transcription machinery. **Epigenetics** is the study of how gene expression can be regulated at the chromatin level. This kind of regulation can be so powerful that it can sometimes be inherited from your parents and also passed on to your own offspring.

While constitutive heterochromatin almost never decondenses, facultative heterochromatin is much more likely to undergo a transition to euchromatin so that genes in the area can be transcribed. This is facilitated by a combination of proteins known as **histone-modifying enzymes** and **chromatin-remodeling complexes**. Usually, the existing histone proteins, within the chromatin, are chemically modified first by the modifying enzymes, which then allows the chromatin-remodeling complexes to bind and do their work. Ultimately, the result is that the packing of the DNA in that region is changed in some way.

Modification of Histone Tails Regulates Chromatin Packing

Histones are a key component of how chromatin structure is managed. More specifically, the tails of the histones are vital to this process.

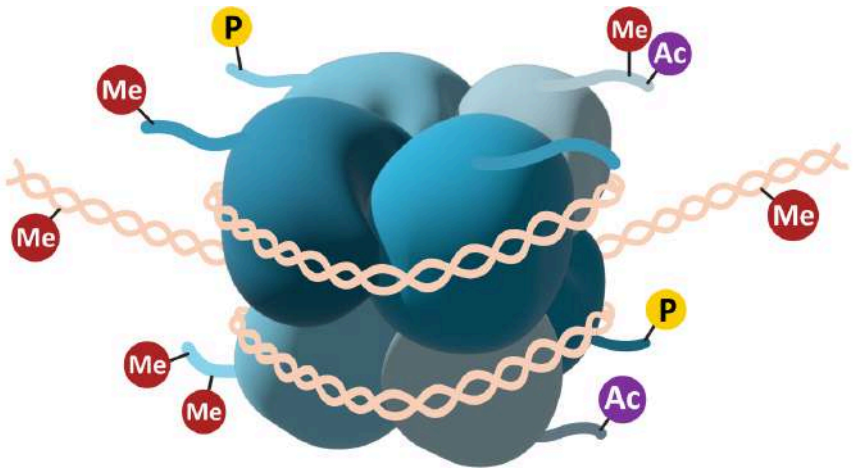


Figure 03-08: Histone tails can be chemically modified, which will alter their function. Chemical functional groups, such as phosphorylation (P), acetylation (Ac), and methylation (Me) can be added to specific amino acids of the histone tails. These posttranslational modifications alter the stability of the histone subunits with one another and with DNA. Tails can carry no modifications, one, or multiple. DNA itself can also be chemically modified (almost always via methylation, as shown), which will also have an impact on packing and/or gene expression. Most commonly, DNA is methylated on cytosines, but adenines have also been shown to allow methylation as well. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Each of the eight core histones has a short "tail" that can be accessed by **histone-modifying enzymes** inside the nucleus. These tails are usually found at the C-terminus of the primary sequence of the core histone and can be modified by the addition of a variety of chemical functional groups. The most common modifications include **acetylation** (Ac), **methylation** (Me), and/or **phosphorylation** (P; see Figure 03-08). The functional groups are added to specific amino acid side chains in the histone tail. Lysines and arginines can be methylated or acetylated, whereas phosphorylation is usually done on serines, threonines, and/or tyrosines. Each of these has different effects on the histones, which, in turn, will impact the availability of the genes in that region. Acetylation is *usually* associated with an increase in gene expression—the changes in electrical charge that are the result of acetylation will reduce the ability of the histone to interact efficiently with the negatively charged DNA. Methylation can result in either an increase or a decrease in regulation depending on the location of the methyl group on the histone tail. Methylation works by creating docking sites for other proteins, which is why its impact is variable. Phosphorylation is not as common on histones, though it is a very common

way to control other proteins in the cell (we'll see examples of this in later chapters). Phosphorylation also changes the charge of the histone, so it can work similarly to acetylation. Phosphorylation of histones has specifically been shown to play a role in DNA repair mechanisms and the extreme DNA packing that is required during mitosis and meiosis.

Chromatin-Remodeling Complexes

In order to actually pack and unpack the DNA, the histones need to be shifted around or even removed so that the DNA can be accessed. Chromatin-remodeling complexes use ATP to drive reactions that affect nucleosome location and/or structure (Figure 03-09). There are a few different ways that the chromatin-remodeling complexes can interact with the nucleosomes, including the following:

1. **Nucleosome sliding**: In this, the nucleosome is not removed but merely shuffled along the DNA. This can be used to expose a nearby regulatory sequence, for example, without opening up the DNA too much.
2. **Nucleosome eviction**: Sometimes the DNA needs to be opened up more, so one or more histone cores will be removed entirely to allow better access to the DNA.

---

3. *Histone Variants*

There's an online database of all of the histones and their variants across species at Histone DB 2.0—with Variants. Go and explore for yourself!

Histone exchange: In this scenario, one or more subunits of a histone core are removed and replaced with a different histone variant. For example, the H2A subunit is commonly replaced with a different histone known as H2A.Z in sites where transcription or DNA repair is required. Histone H2A is thought to have the highest number of variants. The most common H3 variant is H3.3. H2B and H4 have very few variants that have been discovered, if any at all.

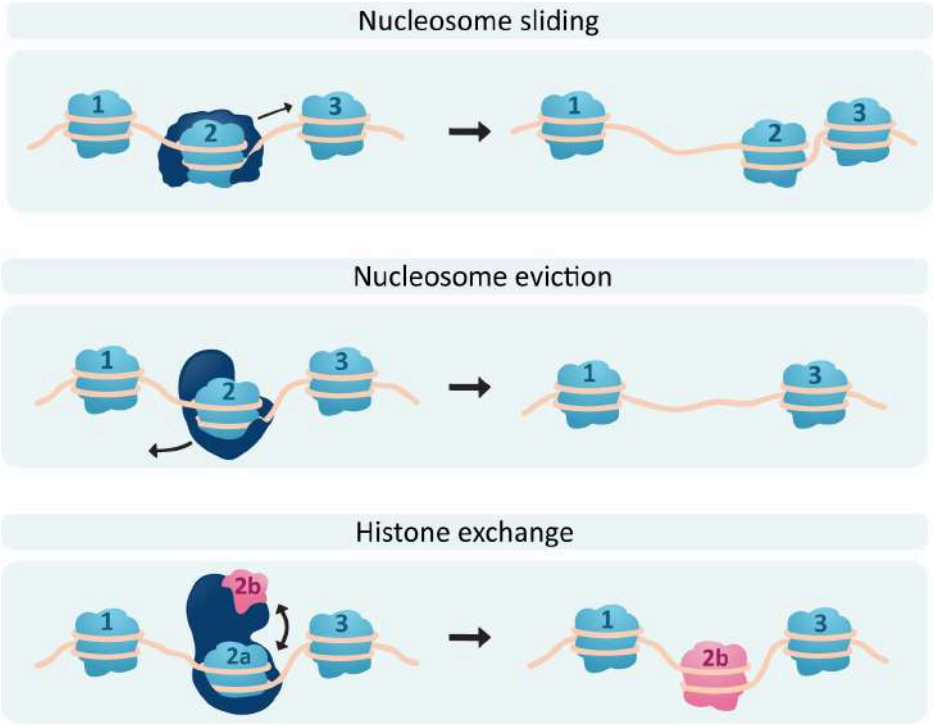


Figure 03-09: Chromatin remodeling can occur via a few major methods. Nucleosome sliding occurs where a histone-modifying enzyme slides the nucleosomes along DNA to a new location. The enzyme binds and physically moves the nucleosome, revealing a new space on the DNA. Nucleosome eviction is when the nucleosome is removed from the DNA entirely, allowing for new access to a DNA region. Histone exchange is a process in which an enzyme removes a nucleosome and replaces it with a nucleosome with different histone subtypes within it. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Of course, if chromatin can be opened up, it can also be repacked and made inaccessible, so each of these processes will also have the capacity to be reversed. Video 03-03 is an excellent video that illustrates the effects of chromatin packing and remodeling using the inactivation of one of the X chromosomes in genetically female cells as an example. Since genetic females have two copies of the X chromosome, and genetic males only have one, it is normal for one of the two X chromosomes in genetic females to be inactivated by the cell via condensing it into heterochromatin. The video is about 11 minutes, but it's worth taking the time to watch. (Note: The video uses somewhat outdated and noninclusive terminology surrounding gender assignments. Despite this, the descriptions of epigenetics and DNA packaging are quite good.)

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=147#oembed-3

Video 03-03: The definition of epigenetics and the process by which an X chromosome in XX individuals is packaged and silenced.

Transcriptional Control: Transcription Factors Regulate Gene Expression at the DNA Level

Histone and chromatin remodeling can allow or inhibit access to genes, thereby allowing or inhibiting transcription. However, this level of control is not always nuanced enough for the needs of the average eukaryotic cell. **Transcription factors** provide an important additional level of control. Not only are *general* transcription factors required to allow the RNA polymerase to bind to the DNA for initiation, but additional gene-specific transcription factors can *enhance* or *inhibit* transcription. We will look at the role of transcription factors in more detail, but first we must review the structure of a gene and the mRNA that gets transcribed.

The Structure of a Eukaryotic Gene

Figure 03-10 shows the key structural elements of a gene (also known as a transcriptional unit). While this figure shows the features of a protein coding gene (noted due to the highlighted mRNA at the bottom of the image), the elements found on the DNA are going to be common to all genes, whether they ultimately code for protein or not.

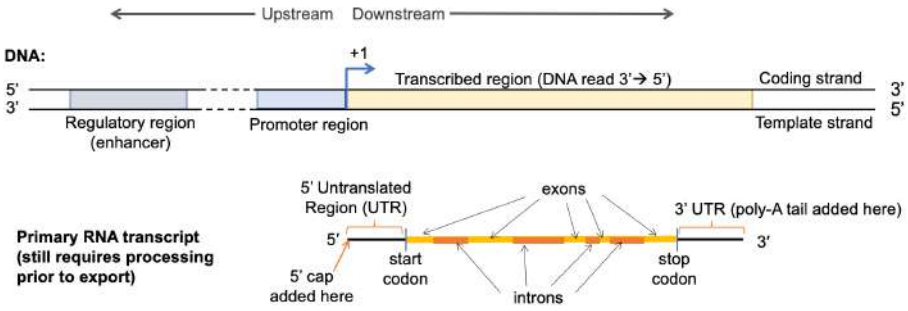


Figure 03-10: The structure of a eukaryotic gene (top) and its relationship to the transcribed mRNA (bottom). This image was created by Dr. Robin Young and is shared under a CC BY-SA 4.0 license.

There are several key DNA sequences shown in Figure 03-10 that need to be highlighted. First, all genes have two major regions: the **regulatory region**, in which all regulatory sequences reside, including the site where the promoter is located, and the **transcribed region**, from which the gene transcript (mRNA, tRNA, rRNA) is derived. The regulatory region can be further broken into the following:

- **Regulatory DNA sequence**—These sequences are used to control when and how much transcription takes place. In this example, we can see a regulatory region that would *enhance* (i.e., promote and/or increase transcription), but there are other regulatory regions that would *suppress* transcription (i.e., reduce or inhibit) and also regions that are required but are considered *neutral* (i.e., neither enhances nor suppresses).
    - Regulatory regions such as these can be located hundreds or even thousands of base pairs away on the linear DNA. The looping of the DNA into topologically associated domains (TADs, described in the previous section) will bring these sections together in the 3D space of the nucleus.
- **Promoter**—This is the binding site for RNA polymerase and other factors involved in the initiation of transcription. It is usually directly adjacent to the transcription start site.

- Other regulatory proteins are also required at the promoter region to help the RNA polymerase bind properly. These proteins are known as **general transcription factors**, as they are required for all transcription. Like all transcription factors, the binding of the protein to the DNA is sequence specific.

The transcribed region begins at the point where the first nucleotide from the DNA is read to create RNA and ends at the site of the last nucleotide that is transcribed into RNA. This region includes the following:

- **Transcription start site**—Also known as the +1 site. This is the first nucleotide from the DNA template that actually gets transcribed into RNA by the polymerase. So it's where transcription "begins."
- Transcription must also be stopped (**transcription stop site**) when the transcript is complete; however, it is not entirely clear how this happens in Eukaryotes. It is likely that different organisms and different RNA polymerases (Eukaryotes can have as many as five different polymerases) will use different mechanisms for termination of transcription.

One of the challenges when discussing genes is that genes are only found on DNA, and DNA is a double helix. As such, there are always two DNA strands present within the gene. However, only one of the strands is ever used for transcription, and for that specific gene, it is always the same strand of DNA that gets used for this purpose. This often creates confusion when discussing genes, especially for students…how do we identify which strand is which? Scientists use the following terms and conventions to make things clearer:

- The **template strand** is the DNA strand that the polymerase will physically bind to and use as the *template* to transcribe the RNA.
    - The template strand is obviously very important to the cell. Interestingly, geneticists and molecular biologists don't discuss this strand very often. This is because the sequence on this strand is complementary to the RNA sequence. When working with genetic sequences on computers (i.e., bioinformatics), this can be quite confusing to read and make sense of.
- Instead, scientists use the sequence on the *other* DNA strand. This is called the **coding strand**, as it has the same sequence as the RNA (except, of course, that it has base T in the DNA instead of the U in RNA). This strand is much easier for molecular biologists to refer to when doing genetic research. We say that it "carries the code" that is in the same direction that the RNA will be read to translate the protein.
    - By convention, we *always* refer to the coding strand in the 5′ to 3′ direction (if needed, refer to the introduction links to review the terminology for gene directionality as well as Figure 03-10), which is the same direction in which the ribosome reads the mRNA during translation. Again, this simplifies our work as scientists when analyzing the extremely long sequences found within genes and genomes.
    - We refer to all parts of the gene *in relation to the coding strand*, since the code is the same as the one that we're interested in (i.e., the one that is used to translate the protein). This is key and should be remembered, as it will help you remain oriented as we talk

about this topic.

- As mentioned above, the +1 site is where the RNA polymerase will begin transcribing the DNA. It is used as a point of reference within the gene to help us orient ourselves with the rest of the chromosome. Anything that is on the 5' end of the +1 site on the coding strand is said to be **upstream**, and anything on the 3' side of the +1 site is said to be **downstream** with respect to that gene.

The product of transcription is an RNA transcript. In Figure 3-10 above, we have shown a gene that codes for a protein, so the RNA produced from that gene is known as **messenger RNA (mRNA)**. MRNA also has several key sections you should know:

- **5'** and **3' untranslated regions (UTRs)—**These are the sequences upstream and downstream of the protein coding region on the RNA. They are used, in part, to help make sure the ribosome can hold onto the RNA well enough to completely translate the **protein coding region**. They also have a role to play in RNA stability.
- The protein coding region is the section that will eventually become the protein once the mRNA is exported to the cytosol and combined with the ribosome.
- **Translation start site—**This is the site of the start codon, which will initiate translation. The start codon marks the beginning of the protein coding region. Note that *the start codon is not in the same spot as the +1 site on the DNA*. The 5' UTR comes first, and the start codon is downstream of that.
- **Translation stop site** (stop codon)**—**This is where the stop codon is located and also marks the end of the protein coding region. Once read, the ribosome is released from the RNA, and translation is terminated.
- In between the start and stop codon, there are **introns** and **exons**. We will discuss both a little later, but in essence, introns are *intervening sequences*, which will be removed before the mRNA is mature and ready to be sent to the cytosol for translation. Exons are left in the sequence so that they can be *expressed*.

One final thing to remember here is that RNA does not *always* code for proteins. There are many genes in the genome that code for RNA only (i.e., the RNA is transcribed but will *not* be translated by the ribosome). For example, the building of proteins also requires **ribosomal RNA (rRNA)**, which is what the catalytic regions of the ribosome are made of, and **transfer RNA (tRNA)**, which is covalently bound to the amino acids and will be used to translate the mRNA codons into a sequence of amino acids. In addition to these three types, a number of additional forms of RNA have also been discovered in recent years, and virtually none of them code for protein. RNA that is not going to be used as mRNA will not have a translation start/stop site, nor will it have introns or exons (though they may have sections that are removed). Additionally, RNA that is not going to be used for building proteins should generally not be discussed in terms of codons, as that is a language for translation. As little as 1–2% of the human genome is thought to actually code for proteins. Interestingly, 26% of the human genome is thought to be introns, which are removed from coding genes as mRNA is processed, as we'll see in a later section of this topic.

Transcription Factors Control When and How Transcription Happens

Transcription requires a number of different proteins, in addition to the RNA polymerase, to bind to the DNA. Figure 03-11 shows a eukaryotic transcription complex that has assembled on the DNA in the regulatory region of the gene and is ready to begin transcribing DNA into RNA. In this figure, we see a number of key components:

- Chromatin-remodeling complexes help shift or remove nucleosomes to allow access to the DNA.
- General transcription factors help the RNA polymerase to bind, and other transcription regulators determine when gene expression is activated, and to what level.
- Mediator is a large protein that can act as a hub to bind general transcription factors as well as other transcription regulators together. This is helpful, since some of the regulatory DNA they bind to can be far away on the linear strand.

These all assemble within a topologically associated domain (TAD) and form a multipart complex to initiate transcription.

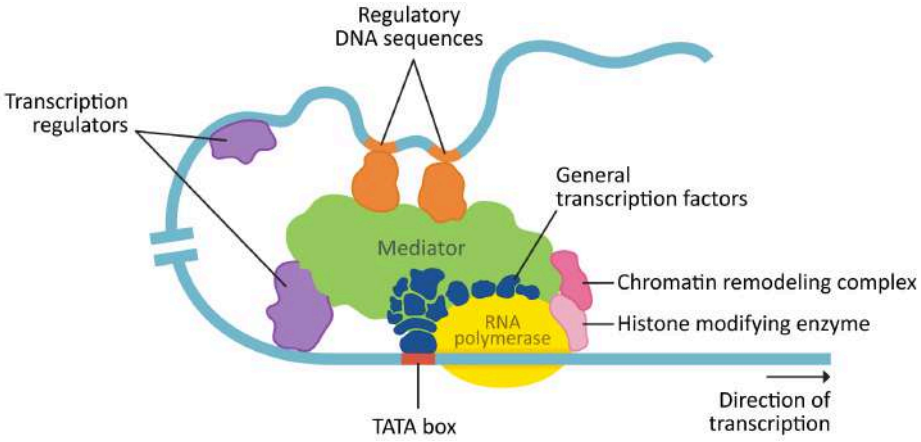


Figure 03-11: The Eukaryotic transcription complex. The mediator acts as a hub to assemble a variety of factors to initiate transcription, including general transcription factors, chromatin-remodeling complexes, histone-modifying enzymes, and additional transcription regulators like enhancers/repressors. Note the "gap" in the DNA on the left-hand side (identified by -||-), which indicates that some regulatory regions could be very far away from the +1 site—sometimes as much as thousands of bases away. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Gene expression can be controlled by a number of different types of transcription regulators:

- **Activator proteins** bind to **enhancer regions** on the DNA.
- **Repressor proteins** bind to suppressor regions on the DNA.
- **Cofactors** work together with other regulatory proteins to change the transcriptional response of the gene.
- **Histone-modifying enzymes** will chemically modify the histones to facilitate additional changes to the chromatin.

- **Chromatin-remodeling complexes** bind to nucleosomes (using the tail modifications created by the histone-modifying enzymes) and help open up the DNA and make it accessible for transcription. It is common for these complexes to also have enzymatic activity.

While some of these regulatory proteins will be present in all genes, each gene has its own unique set of regulatory sequences by which it is controlled and may only include a subset of the ones in the list. These sequences are often spread over hundreds to thousands of base pairs, and they accomplish very complex regulatory tasks. The following are some examples:

- different genes can be transcribed at different rates,
- the same gene can be transcribed at different rates in different tissues,
- the same gene can be transcribed at different rates at different times during development in the same tissue, and
- some genes will not be transcribed at all, as they are not required in that particular cell type or at that stage in development.

The binding of different combinations of transcription regulators in different tissues and during different times in development is what allows such flexibility in the expression of eukaryotic genes. This concept is often referred to as *combinatorial control*, and it's a very powerful way to produce all of the nuanced transcriptional responses required to make the average multicellular organism (such as ourselves) continue to function properly throughout its life-span.

Posttranscriptional Control: mRNA Processing

The physical separation of transcription (in the nucleus) and translation (in the cytosol) in Eukaryotes has created space for increased flexibility in gene expression, as well as providing additional protection from mutation. Thus, in Eukaryotes, the RNA that is synthesized by the RNA polymerase is often referred to in textbooks such as this as the *pre-RNA* or the *primary transcript*. Once transcription is complete, the pre-RNA will be further modified to prepare it for the next stage of its journey, which often includes export from the nucleus. RNA processing is extremely complex and an area of active research. As is so often the case in cell biology, we are only just beginning to understand what the cell can do.

General Principles of Transcript Processing

The first and most important concept to remember is that *all transcripts—mRNA, rRNA, tRNA, and all other forms of noncoding (nc)RNA—are processed in the nucleus* before they are exported to the cytoplasm. Each type of RNA will require unique processing steps. First, some general information about RNA transcript processing:

- RNA processing is carried out by proteins (and RNA) that bind to, and modify, the transcripts directly.
- Virtually all processing signals are encoded into the primary sequence of the RNA transcripts themselves. This is a concept we have seen before when referring to protein processing and folding.

- Processing may include any of the following modifications, depending on the class of RNA:
    - addition of sequences (e.g., 5′ cap and poly[A] tail in mRNA)
    - cleavage of the transcript into several pieces (rRNA)
    - removal of some sequences (all classes)
    - splicing (i.e., removal of sequences by cleavage followed by rejoining remaining RNA fragments back together; this is how introns are removed from mRNA)

Note that in Eukaryotes, there are between three (in mammals) and five (in plants) different RNA polymerases. In addition to their role in transcribing the RNA, the polymerases are also often involved in the first step(s) of posttranscriptional processing of RNA. In this textbook, we are focusing on the synthesis of mRNA, which is the job of RNA Polymerase II (RNA Pol II). RNA Pol II is itself a rather large protein complex (17 subunits). However, once it combines with the various transcription factors and other required proteins and enzymes, it's considered to be one of the larger structures within the cell and is roughly the same size as the ribosome. Many of the proteins responsible for RNA processing hitch a ride on the RNA polymerase, which allows them easy access to the transcript once transcription begins. This also results in some, but not all, of the "posttranscriptional" processing reactions happening *at the same time that transcription is taking place*.

Again, there are many different classes of RNA, each with its own processing requirements. To simplify the rest of this section, we will focus solely on the processing of mRNA. Remember that other types of RNA (mRNA, tRNA, rRNA, and other ncRNA) will have their own unique steps to complete before they are considered *mature RNA*.

mRNA Processing

There are three major processing events that are required before a pre-mRNA is considered to be *mature* and ready for export (Figure 03-12):

1. **RNA capping** at the 5' end of the RNA. The 5' cap consists of a modified guanosine (G) with an extra methyl group attached to it, which is joined to the initial 5′ nucleotide of the nascent RNA using a triphosphate linkage. This is added by enzymes that are part of the RNA polymerase complex right at the start of transcription, when the transcript is still only about 25 nucleotides long.
2. **Polyadenylation**. A poly(A) (polyadenylic acid) tail of about 100–200 adenylic acid (A) residues is added near the 3′ end of the primary transcript. There is a specific base sequence (AATAA) in the 3' end of mRNA that acts as the signal site. That sequence is recognized by a specific endonuclease (i.e., an enzyme that cuts nucleic acids). The endonuclease cuts the transcript 20–30 bases downstream of the recognition sequence and then adds the A residues.
3. **Splicing**. During splicing, portions of the coding region of the mRNA transcript are removed. This will be discussed in more detail below.

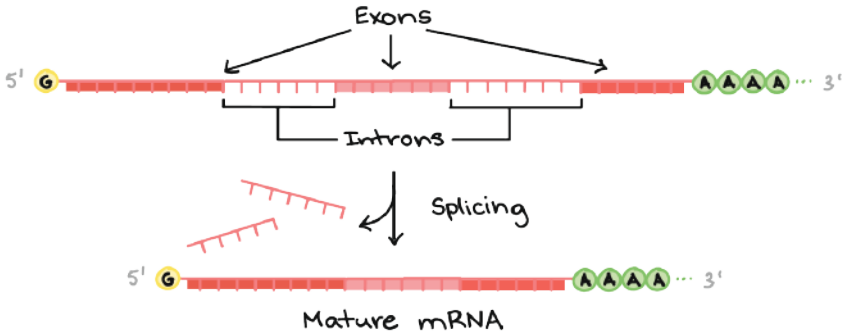


Figure 03-12: Posttranscriptional RNA processing. Exons are removed and the ends joined together to make a mature mRNA product that can be exported to the cytoplasm. Image created by Khan Academy and licensed under CC BY-NC-SA 3.0. Note: All Khan Academy content is available for free at www.khanacademy.org.

The roles of RNA capping and polyadenylation are similar; they both serve to increase stability of the final mRNA molecule and to identify it as a completed, mature transcript that is ready to be exported out of the nucleus. Nuclear export proteins will need to bind to these regions of the mRNA transcript in order to facilitate mRNA export for the nucleus. (Nuclear export will be discussed in more detail later in this chapter.)

As mentioned above, mRNA splicing is when a portion of the RNA is excised from the coding region of the transcript, leaving behind a shorter mRNA that will be used for translation. A typical coding region in a primary mRNA transcript will include the following:

1. **Introns** (which stands for *intervening sequences*) are noncoding RNA segments that are recognized and removed from the primary transcript. Usually, 75–80% of the initial primary mRNA transcript is lost as a result of splicing. In some cases, it has been shown to be as much as 95%.
    - Just because introns are noncoding (i.e., not translated into protein), this does not mean that introns do not carry important information. Often regulatory sequences are found in the DNA within intron regions. These may regulate the gene in which they sit, but they may also regulate other genes that are upstream or downstream of that site.
2. **Exons** (which stands for *expressed sequences*) are the coding sequences that are left behind in the transcript. They contain the sequence that codes for the protein and are destined for export to the cytoplasm.
    - A gene could have many exons (some genes have more than 50!) that are joined together to produce a processed transcript.

In order for the cell to remove introns and then join the remaining exons together, the cellular machinery once again looks for cues within the mRNA sequence itself. Analysis of exon/intron boundaries and intron sequences reveals the following common features, as shown in Figure 03-13:

- The bolded sequence indicates the nucleotides present at the intron/exon boundary.
- Other nucleotides in the vicinity that are important for establishing the intron/exon boundary are also indicated. The letters in the sequence represent the following:

- A, G, C, and U are the nucleotides. Note that since this is RNA, uracil is present and not thymine.
- R stands for either A or G.
- Y stands for either C or U.
- N stands for any nucleotide (A, C, G, or U).
- The dashed line indicates nucleotide sequences of varying lengths that are not key to removal of the intron.

The A in the center of intron 1 is the site where the **lariat loop** will be joined. (More details on this in Figure 03-14.)

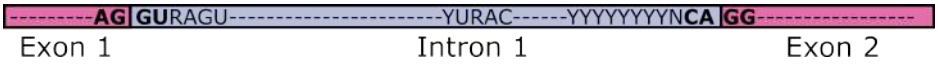


Figure 03-13: Intron/exon boundaries are identified by specific nucleotide sequences. The pink labeled sections reference two exons that are to be bound together, while the blue is the intronic section that is to be removed in the splicing process. The bolded letters indicate nucleotides on the very edge of the intron/exon boundary. Other nucleotides are also labeled. This image was created by Dr. Lauren Dalton and is shared under a CC BY-SA 4.0 license.

The sequences identified in Figure 03-13 are *required* for proper intron excision. They are considered to be almost universal. Interestingly, despite the fact that almost all splicing is thought to use the same sequences, this process is still extremely complex, and researchers don't entirely understand it.

Cellular Function: Mechanism of RNA Splicing

It is interesting that the required sequences for splicing are quite short (compared to the length of the genes themselves) and have a lot of variation built into them, and yet intron removal is an extremely precise process. The general mechanism of splicing is described in Figure 03-14.

Prior to describing the process of splicing, we will first explore the proteins that do this work, as they are unique. A large complex known as the **spliceosome** does the work of binding to the ends of the introns, cutting them out, and then rejoining the ends of the exons. This complex combines both proteins and special protein-RNA complexes called **snRNPs** (*s*mall *n*uclear *r*ibo*n*ucleo*p*roteins), pronounced "snurps." SnRNPs are enzymes that contain a small RNA molecule that is complementary to the recognition sequences at the intron-exon junction. The RNA molecule within the snRNP helps make sure that the binding is precise. The rest of the proteins in the spliceosome help with the other aspects of its function (described below). There are as many as 5 different snRNPs and over 200 proteins that could be a part of the spliceosome. As such, specific parts of the spliceosome can be changed (i.e., proteins, RNA, or both that can be swapped in/out), which adds additional layers of specificity to the process.

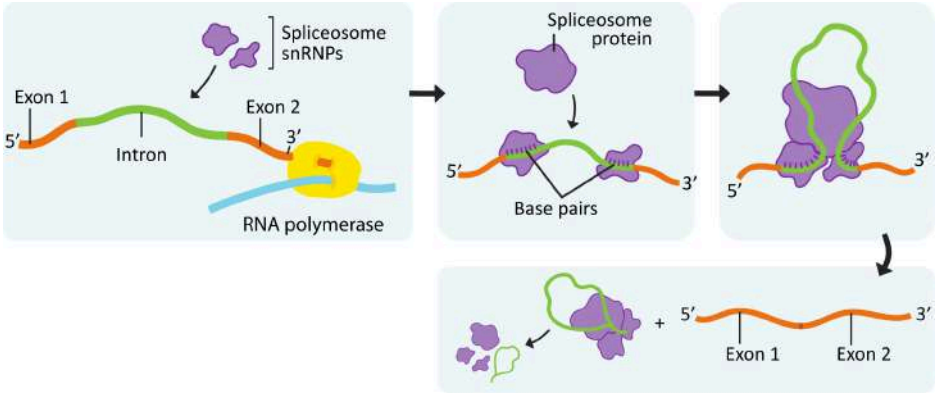


Figure 03-14: mRNA splicing mechanistic diagram. snRNPs bind to the intron/exon boundaries in the first panel. This recruits the other spliceosome proteins in the second panel. The spliceosome loops the intron and finally removes the intron and joins the exons in the final step. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The key steps in splicing, shown in Figure 03-14, are as follows:

1. The initial formation of the spliceosome begins when the RNA portions of the snRNPs recognize the intron/exon junctions and base pair with them.
    - Some of the enzymes involved in splicing are transferred from the RNA polymerase complex that formed to initiate transcription (the others are available within the nucleus).
    - The components assemble spontaneously at any site that carries the proper sequences.
    - Different snRNPs are required for different parts of the process, and as such, the sequence of the mRNA is constantly checked and rechecked as new snRNPs must bind to join the spliceosome. This is thought to be one way that the precision of the splice sites is maintained.
2. Once the components have assembled at the intron/exon junctions, other proteins and snRNPs arrive and interact to bring the two ends of the intron together to form the complete spliceosome so that splicing can begin.
    - The 5' boundary of the intron is cut.
    - Then the 5' end of the intron is bonded to the 3′ hydroxyl of one of the nucleotides near the 3' end of the intron to form a structure known as the **lariat loop**.
3. The second cut in the transcript occurs at the right edge of the intron, and the two exons are joined together.
    - Interestingly, while introns seem to have a lot of variation in their length, exons tend to be more uniform in size. This is thought to contribute to the ability of the spliceosome to determine which parts are exons and which are introns.

While the assembly of the spliceosome begins during transcription, the actual splicing may not occur until after transcription has ended. As a result, *there is no guarantee that introns are removed in the order that they appear on the transcript*.

In addition to Figure 03-14, there is an excellent video (Video 03-04) produced by the DNA Learning Centre that highlights this mechanism very well from a conceptual perspective.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=147#oembed-4

Video 03-04: RNA splicing animation.

Alternative Splicing Increases the Number of Proteins Possible from a Single Gene

In Eukaryotes, virtually all protein coding genes are made of a combination of introns and exons. There are thought to be several advantages to this, including the fact that it may protect against mutations impacting protein sequence. Another advantage, for which we see the evidence in many genomes, is the ability to produce multiple different variations of a single protein, all of which can be transcribed using the same gene. This is known as **alternative splicing**, and it is a relatively common occurrence in Eukaryotes—about 95% of all human protein coding genes are thought to be involved in alternative splicing. Simply by changing what is recognized as an "intron" and what is recognized as an "exon," the cell can produce a different final product. These differences in splicing patterns are often to produce tissue- or developmental stage–specific protein variants.

The most common form of alternative splicing is known as **exon skipping**, in which one of the exons gets treated as part of an intron and is removed. Exon skipping is illustrated in Figure 03-15 below. However, there are other common patterns as well (though we do not have time to discuss them in this text).

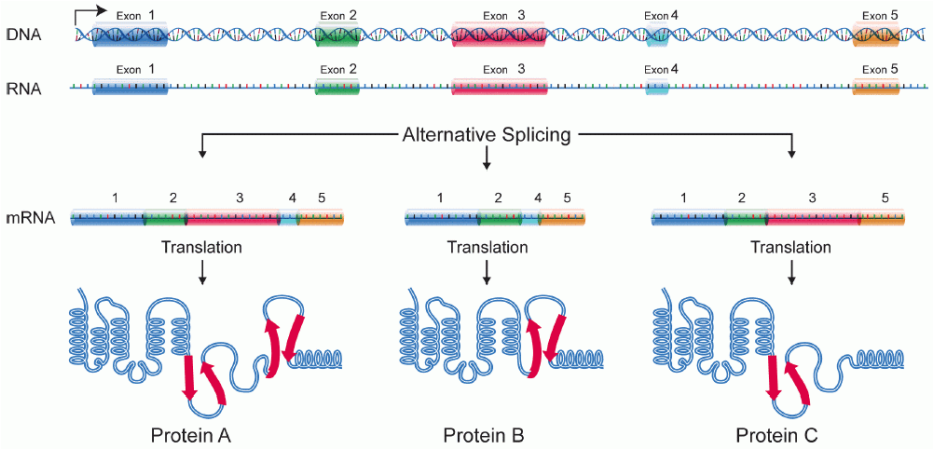


Figure 03-15: Examples of alternative splicing by the National Human Genome Research Institute. Protein A expresses all of the exons. In protein B, exon 3 is skipped, whereas in protein C, exon 4 is skipped. As a result, proteins B and C are each missing a domain. This image is in the public domain.

Final Thoughts on Splicing

Experiments show that mutations at intron/exon junctions often result in changes in splicing patterns. This lends weight to the idea that splicing is a precise process that requires specific sequences to function. On the other hand, evidence also shows that the spliceosome is capable of adapting as needed. Generally, the best possible splice site tends to take precedence over other options. However, if one (or several) alternative splice sites are available, the spliceosome can take advantage of those sites as well. Choosing the "correct" binding site could take a little extra time. Indeed, the components of the spliceosome have been shown to assemble *co-transcriptionally* (i.e., during transcription) but are sometimes delayed in initiating the process of splicing. It should also be noted that despite the consensus that splicing is considered to be an extremely precise process when analyzed in a test tube, it is unknown how accurate it is in the cellular context. This is because any improperly processed RNA transcripts in the nucleus are immediately degraded, which makes it impossible to measure the error rate on the process in a live cell.

To further complicate matters…

- In addition to alternative splicing sites, it is also possible for a gene to have alternative cleavage and poly(A) addition sites.
- Some genes will have two or even more promoters, each of which leads to the production of a different initial exon. This is referred to as "promoter choice."
    - Usually, different promoters are active in different tissues or developmental stages. How might this occur? This is a challenging phenomenon.
- If that weren't complicated enough, there are even some cases where exons from two separate gene transcripts are spliced together to produce a completely new mRNA (this is known as trans-splicing).
    - The benefit of this is largely unknown, but one theory is that it can improve the efficiency of translation.

While scientists cannot currently answer all of these questions, the fact that the cell has so much flexibility in RNA processing is, in itself, astounding.

Studying Cells: ChIP to Investigate How Histone Modifications Impact Transcription of Specific Genes—a Case Study

There are a variety of techniques used to study the structure of the genome, the level of compaction, and the degree of transcriptional activity within areas of the genome. One such technique commonly used is called **ChIP**, which stands for *ch*romatin *i*mmuno*p*recipitation. While this may sound like a complex technique, it builds on what you have learned in this course so far.

Topic 3.1 in this chapter introduced you to all of the ways that the cell can control access to the DNA in a particular region of the genome, while Topic 3.2 was focused on the various ways the expression of a particular gene within the genome can be controlled, both before and after transcription. In a live cell, these two methods of genetic control would work in tandem to determine when and how genes are expressed. Thus, scientists have worked to find ways to study them together within the context of a real cell.

Since the genome is quite large, making it rather unwieldy to study in its entirety, ChIP gives us

options that help break down the genome into more manageable bits and allow us to look at *both* the genetic (i.e., transcription factors) and epigenetic controls (i.e., histone modifications, etc.) that are being used by the cell.

Figure 03-16 shows a schematic of how this technique works:

1. In essence, reversible fixatives are used to physically cross-link all proteins that are bound to the DNA (e.g., histones or transcription factors) at a specific moment in time. Since this initial step is done in a live cell, the entire genome can be fixed at once.
2. Then the DNA is broken apart into smaller, more manageable fragments (~500 bp each) using either mechanical stress or enzymes that digest chromatin.
3. Then antibodies are added that bind to specific proteins we know in order to purify and concentrate them through a process known as **immunoprecipitation**. In essence, the antibody is attached to a glass or latex bead and "precipitated out of solution" using a centrifuge. The result is that any chromatin fragments with your protein in them bind to the antibody and can be removed from the rest of the solution.
4. Next, the fixative is removed and the purified chromatin bits are separated into DNA and associated proteins. Both DNA and proteins can be analyzed to look for patterns and themes of interest.

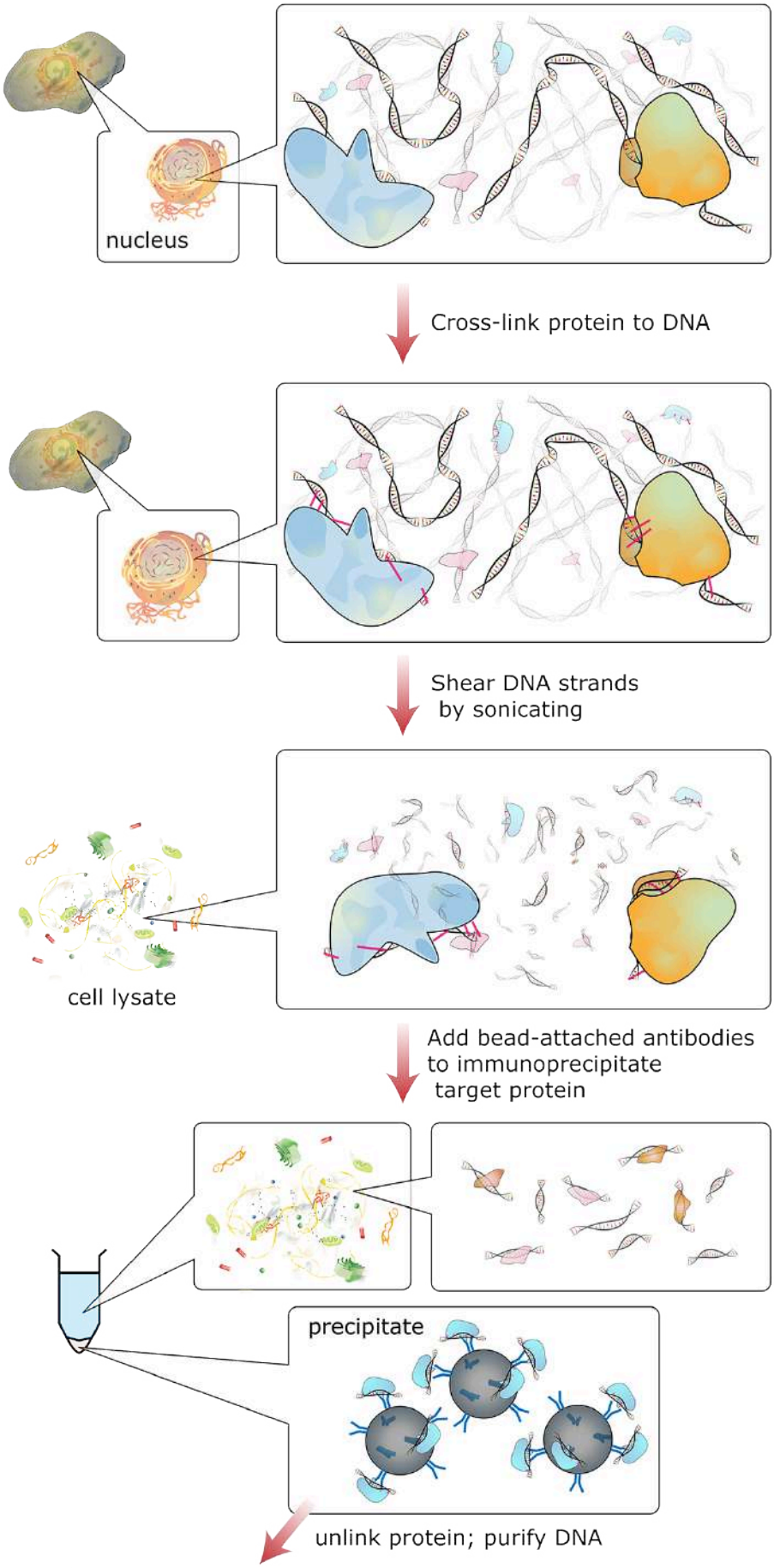

nucleus
Cross-link protein to DNA
Shear DNA strands
by sonicating
cell lysate
Add bead-attached antibodies
to immunoprecipitate
target protein
precipitate
unlink protein; purify DNA

Figure 03-16: Schematic of chromatin immunoprecipitation workflow. Proteins are linked to one another and the DNA, then purified, separated, and analyzed. This image "ChIP-Seq Workflow" by Jkwchui is shared under a CC BY-SA 3.0 license.

ChIP experiments can help answer a variety of different scientific questions related to chromatin structure and the regulation of gene expression. For example,

- *What proteins interact with specific histones?* To answer this, after purification of the chromatin, using an antibody to the specific histone in question, you could then explore what additional proteins were brought along with the histone, as that would imply they were physically interacting at the time of initial fixation.
- *Do environmental conditions change the expression level of your favorite gene?* To answer this question, you might grow cells in different environmental conditions and then look for a specific DNA sequence that represents your gene. Depending on the proteins associated with the fragments bearing your sequence of interest, you may be able to identify the level of chromatin packing (i.e., heterochromatin-like or euchromatin-like), which could provide valuable insight to when the gene is expressed.
- *What are all the genes targeted by a transcription factor?* If you used an antibody to your transcription factor for the immunoprecipitation, you might then choose to sequence all the DNA fragments that were cross-linked to that protein. This genome-wide approach is called **ChIP-seq**.

As you can see, there are a variety of options you can explore after you have a purified sample of chromatin fragments. Indeed, there are now many different types of ChIP experiments that can be conducted. We will not go over all of these, as there are too many variations to consider at this point. Instead, we'll look at one example of ChIP in action in order to explore what kinds of information we can learn when combining ChIP with other techniques, like gel electrophoresis (which you learned about in Chapter 2).

Case Study: Comparing Variants of Core Histone H2A

This case study focuses on a recent research paper written by PhD student Hilary Brewis and her colleagues from the University of British Columbia. Brewis et al. studied a histone core protein variant called H2A.Z in the budding yeast *Saccharomyces cerevisiae* (Brewis et al., 2021). To understand the research findings, we must first "set the stage." Earlier in this topic (3.2), we discussed the concept of histone exchange as a way to remodel chromatin, consequently aiding in the regulation of gene expression (revisit Figure 03-09 and associated text). H2A is the "original" histone added to the DNA during replication; however, H2A.Z is often exchanged for H2A later in particular regions and/or functional scenarios that are not entirely understood at this time. Histone H2A.Z is added into the histone core by a protein complex known as SWR1-C.

Histone H2A.Z and H2A have about 60% of their protein sequence that is identical. While this number is considered extremely high, it is clear that the 40% difference is enough to make them distinct from each other. Nucleosomes that contain H2A.Z instead of H2A are known to interact with different proteins and react differently to cellular signals for DNA compaction/decompaction. Thus, the differences in the amino acid sequences of the two histones are key to the proper function of both

H2A and H2A.Z. Brewis and colleagues wanted to explore the details of these amino acid differences as a way to further explore histone function.

The first step is to explore the amino acid sequences using bioinformatics. This analysis told Brewis et al. that the differences in the sequences fall into nine distinct regions of the histone primary sequence. However, the role of each of these regions, and how each of the regions contribute to the overall function of H2A, or H2A.Z, was still not known. To test which region was necessary, they genetically engineered nine H2A variants, each of which has one of their nine distinct regions swapped for the H2A.Z version of the sequence. They then could test which of the H2A.Z-specific functions the genetically modified H2A had "picked up" as a result of the sequence swap. Finally, using ChIP, they assessed if any of the engineered H2A proteins (with pieces of H2A.Z swapped in) had functions that are usually associated with H2A.Z only, not H2A. In this case study, we will focus on a single component of Brewis et al.'s work—namely, how to make it so that H2A can interact with the H2A.Z-associated complex, SWR1-C.

In the first set of experiments, Brewis et al. used ChIP to extract and purify wild-type and genetically modified H2A proteins from the chromatin of the yeast cells. Once the chromatin fragments were purified, the DNA was separated from the associated proteins. Brewis et al. then performed SDS-PAGE (see Chapter 2 if you don't remember this technique) on the isolated proteins and probed to see if SWR1-C was one of the proteins that ChIP pulled out.

Figure 03-17 is a schematic representation of what they found.

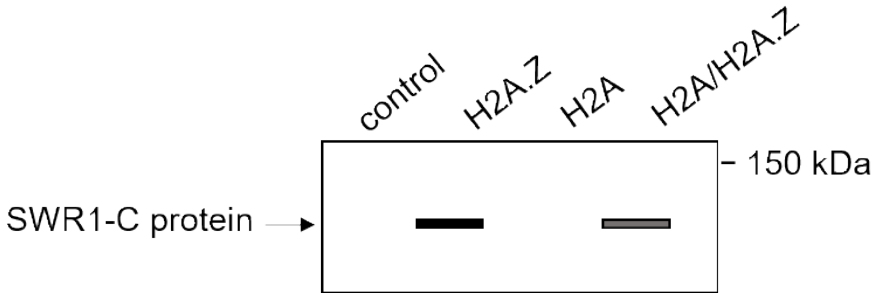


Figure 03-17: Results of an SDS-PAGE experiment that found that SWR1-C can only bind to H2A when it is modified to contain specific sequences of the H2A.Z protein. The presence of the band in the lanes for H2A.Z and the H2A/H2A.Z construct (but not the other lanes) shows that they both interact with the SWR1-C protein, while normal H2A does not. The control lane should not contain any chromatin, and thus no protein should show up in this lane. These data are adapted from Brewis et al. (2021). The figure by Dr. Lauren Dalton is shared under a CC BY-SA 4.0 license.

You can see in Figure 03-17 that there is a band visible in the H2A.Z lane and in the genetically engineered H2A/H2A.Z hybrid protein lane. This means that when they purified H2A.Z and the modified version H2A, they found that a SWR1-C protein was cross-linked to those proteins. It did not bind to the negative control or to the original H2A protein. This tells us that the SWR1-C was only able to bind to proteins that carried specific H2A.Z sequences. From this ChIP and SDS-PAGE experiment, they were able to specifically identify a function for one of the nine variable regions in the H2A.Z proteins. This tells us that there is a specific amino acid sequence required in order for SWR1-C to be able to bind to a histone, and any histone with that complex will be able to interact with SWR1-C. We don't yet know what the H2A version of this sequence does, but since SWR1-C is involved in chromatin remodeling, it may give us some ideas as to what types of functions to explore for H2A's region.

The ChIP that we explored was trying to answer two questions about SWR1-C binding to H2A.Z:

1. What is *necessary* within the amino acid sequence of H2A.Z to allow SWR1-C to bind?
2. Are those amino acids *sufficient*, or is something else also required in order for SWR1-C to bind?

These are very common questions to ask when designing experiments in cell and molecular biology. Many experiments are designed specifically to ask one or both of these questions. If we continue to tease apart the previously described ChIP experiment, we can begin to understand the logic to the experiments that were performed:

- At the beginning, we knew that H2A was lacking a function that H2A.Z was capable of (i.e., binding to SWR1-C).
- By genetically modifying known variable sequences in the H2A protein to match the H2A.Z sequence, Brewis et al. were trying to give the H2A protein a new function that it doesn't normally have. This is what's known as a **gain-of-function experiment**.
    - Gain-of-function experiments are used to explore the absolute minimum requirements needed for a particular function (i.e., what is *sufficient*). Since H2A is not normally capable of the same functions as H2A.Z, we know that any modifications we make are responsible for the gained functionality. If the modifications we make are small enough, then we can tell exactly the minimum amount of change that is needed to make the new function happen.
- Another way to address this question of what's required for SWR1-C binding would have been to mutate H2A.Z in specific known ways and watch for when it loses its ability to bind to SWR1-C. This kind of experiment is known as a **loss-of-function experiment**.
    - Loss-of-function experiments help us determine what aspects of the system are required (i.e., *necessary*). Carefully introducing mutations into H2A.Z in different parts of the amino acid sequence would tell us which sequence(s) are *necessary* for proper H2A.Z function. It won't tell us whether the thing we mutated is the only thing required (i.e., whether it's sufficient), as there may be other sequences that are also required that we have not mutated.

What is necessary versus sufficient in a biological system are common questions that scientists try to answer. To further understand these concepts, we have included these two short videos (Videos 03-05 and 03-06). They cover the same information, more or less, but in slightly different ways.

One or more interactive elements has been excluded from this version of the text. You can view them online

here: https://open.oregonstate.education/cellbiology/?p=147#oembed-5

Video 03-05: The difference between necessary and sufficient using statements about burgers with cheese.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=147#oembed-6

Video 03-06: The difference between necessary and sufficient conditions.

Scientists spend a lot of time trying to refine the question that their experiment is trying to answer and also ensuring that there are no other factors that could be influencing the result. *Controls* are used to rule out other options as well as to ensure that any assumptions we have made are appropriate. We expect that the concept of controls is not new to you, especially since there have been many discussions of experiments that use controls in earlier chapters of this text.

TOPIC 3.3: THE INTERPHASE NUCLEUS—STRUCTURE, FUNCTION, AND PROTEIN IMPORT

Learning Goals

- Describe the structure of the interphase nucleus and identify the structural elements in different kinds of microscopy.
- Describe the nuclear pore complex (NPC) and explain how it controls access to the interior of the nucleus.
- Explain how proteins are actively transported into the nucleus, including the roles of the nuclear localization signal (NLS) nuclear transport receptors and the NPC itself.
- Using experimental evidence from fluorescence microscopy, discuss how the primary sequence of a protein contains all of the information to determine whether a protein is imported into the nucleus.

As we have discussed already in this chapter, the existence of the nucleus in eukaryotic cells is key to

their success. Not only does the physical separation of transcription from translation allow space for RNA processing, but the capacity to have a highly organized genome within the nucleus contributes to the efficiency of gene regulation. What we have not yet discussed is exactly how the structure of the interphase nucleus contributes to its ability to house, organize, and protect the DNA for which it is responsible. This is a very important job, as damage to the DNA could easily result in the death of the cell. As such, the structure of the nucleus is designed to maximize its protective power and aid in the management and organization of these oversized Eukaryotic genomes. As we work our way through the final topic in this chapter, we will look at the role of each of the structural elements of the interphase nucleus (as shown in Figure 03-18) and discuss their functions. We will end this chapter by exploring how the cell controls what is able to enter and exit the nucleus through the nuclear pores.

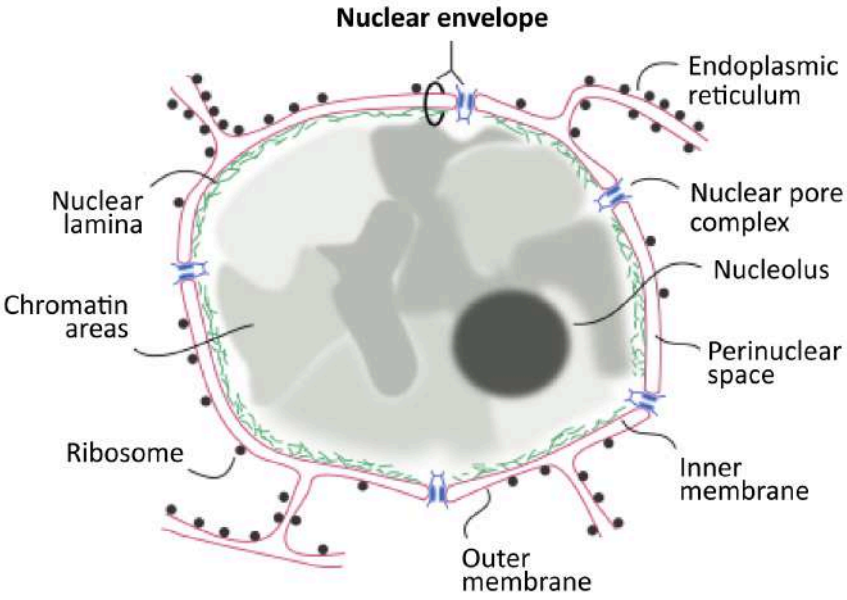


Figure 03-18: Labeled diagram of the nucleus. Modified (labels updated) from Atlas of Plant and Animal Histology by Manuel Megias Pachero, Pilar Molist Garcia, and Manuel Angel Pombal Diego and shared under a CC-BY-NC-SA 3.0 license.

The Nuclear Envelope

The very first thing we should point out is that the **nuclear envelope** is a *double* membrane that surrounds the contents of the nucleus (Figures 03-18 and 03-19). It consists of an inner and an outer membrane, with the **perinuclear space** in between. The outer nuclear membrane is continuous with the endoplasmic reticulum (ER). Thus, the perinuclear space is also continuous with the ER lumen. The cytoplasm and **nucleoplasm** (i.e., the fluid within the nucleus) are connected through the **nuclear pores**.

Since the outer membrane of the nuclear envelope is continuous with the ER, the outer (cytoplasmic) surface of the nuclear envelope can become studded with ribosomes, much like the rough endoplasmic reticulum (rER; Figures 03-19 and 03-20).

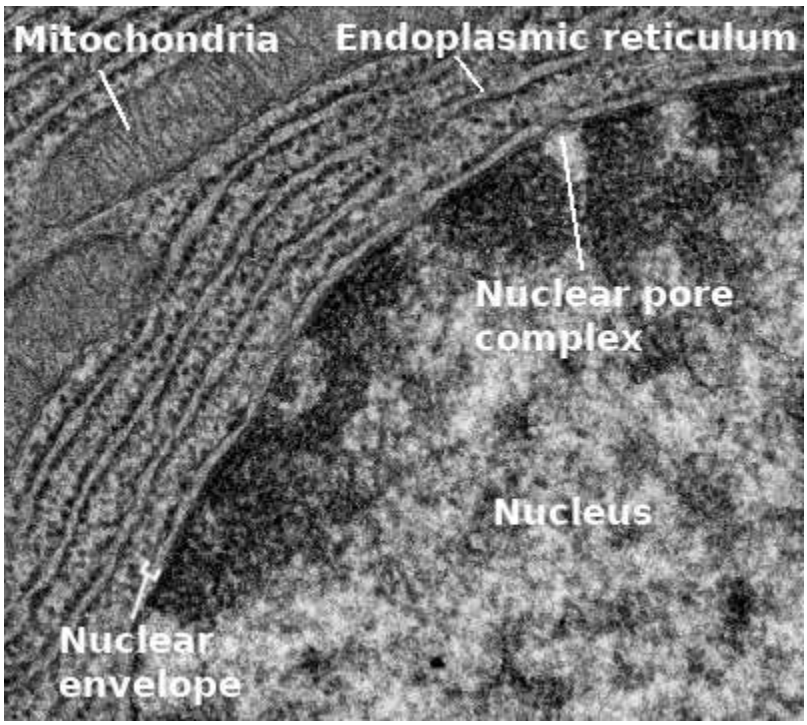


Figure 03-19: TEM micrograph of the nucleus of a cell showing the double membrane of the nuclear envelope as well as other cellular structures. Original image from Atlas of Plant and Animal Histology by Manuel Megias Pachero, Pilar Molist Garcia, and Manuel Angel Pombal Diego. CC-BY-NC-SA 3.0 license.

The Nuclear Lamina

The inner membrane of the nuclear envelope has a meshwork of fibrous protein under it known as the **nuclear lamina**. The role of the nuclear lamina is to shape and support the nuclear envelope. Interestingly, the proteins that form the nuclear lamina have an intriguing evolutionary history and vary somewhat in the different biological kingdoms. In many animals, this meshwork is composed of proteins known as **nuclear lamins**, which are part of a larger family of filamentous proteins known as intermediate filaments, which we will discuss in more detail in Chapter 6. Vertebrates have the largest variety of nuclear lamins, whereas invertebrates have a reduced subset and are considered to be more evolutionarily "primitive." In plants, algae, and other protists, the proteins that make up the nuclear lamina are different. Some of them *may* be ancestors to the animal lamins, but most are not genetically related.

The nuclear lamina is attached to transmembrane proteins embedded in the nuclear envelope, the nuclear pores, and chromatin, so it plays a key role in holding all of the parts of the nucleus together.

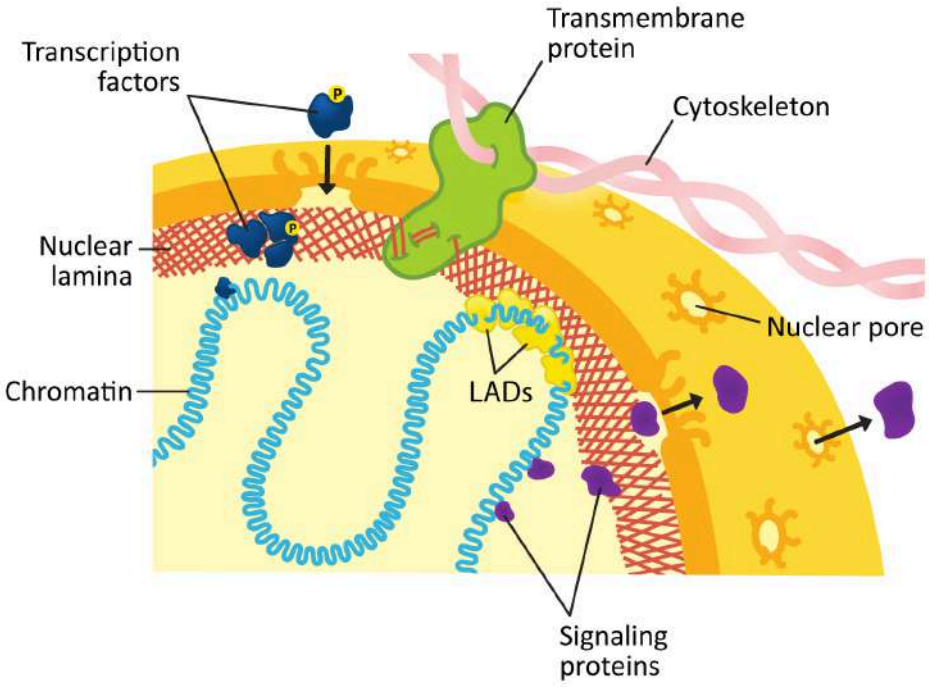


Figure 03-20: The nuclear lamina and its role in the nucleus. The nuclear lamina is a meshwork of protein (red) that is directly adjacent to the inner membrane of the nuclear envelope (yellow). A variety of proteins interact with the lamina to aid in the functions of the nucleus. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Figure 03-20 highlights several roles played by the nuclear lamina in the function of the lamina. These include the following:

- First, the nuclear lamina is a meshwork of proteins directly adjacent to the nuclear envelope. One of its more important functions is to provide structural integrity and support to the nucleus. It helps shape the nucleus as well as protect its contents against whatever might be happening in the rest of the cell.
- It is also heavily involved in organization of the genome, which was described in more detail in Topic 3.1. Chromosomes are attached to the nuclear **lamina via lamin-associated domains (LADs)**, which help form the **chromosome territories**. As a general rule, the LADs tend to form in structural regions of the chromosomes, such as in the telomeres, so that they don't interfere with gene expression.
- The nuclear lamina helps form direct connections to the rest of the cytoplasm that is outside the nucleus through a variety of transmembrane proteins that cross both membranes of the nuclear envelope. Often, they are connected to the cytoskeleton outside of the nucleus and the nuclear lamina inside. These connections are key to maintaining proper positioning of the nucleus as well as transporting it to a new location in the cell if/when necessary.
- The nuclear lamina can also get involved in signaling and regulating gene expression to some extent, as it is able to bind to and sequester proteins that have entered the nucleus. For example,
    - Transcription factors can bind to the nuclear lamina in certain situations, which will stop them from binding to the exposed DNA of the euchromatin farther inside the

nucleus.

- Additionally, signaling proteins that have entered the nucleus can use the nuclear lamin as a platform on which to assemble into complexes.

- Finally, the nuclear lamina plays a key role in the breakdown of the nucleus during mitosis, which is our next focus.

The Nuclear Lamina and Mitosis

During mitosis, the nuclear lamina has an integral role in the breakdown of the nuclear envelope so that the chromosomes can be released and the mitotic spindle can form. This happens when the lamins are phosphorylated (i.e., they have a phosphate group covalently attached), which causes a slight conformational change in the lamins. The result of this tiny change is that the entire network of nuclear lamins is destabilized and the meshwork of the nuclear lamina breaks down. Since the lamins are attached to the nuclear envelope via both the nuclear pores and additional transmembrane proteins, this breakdown tears apart the nuclear envelope as well, and the entire structure disintegrates (Figure 03-21).

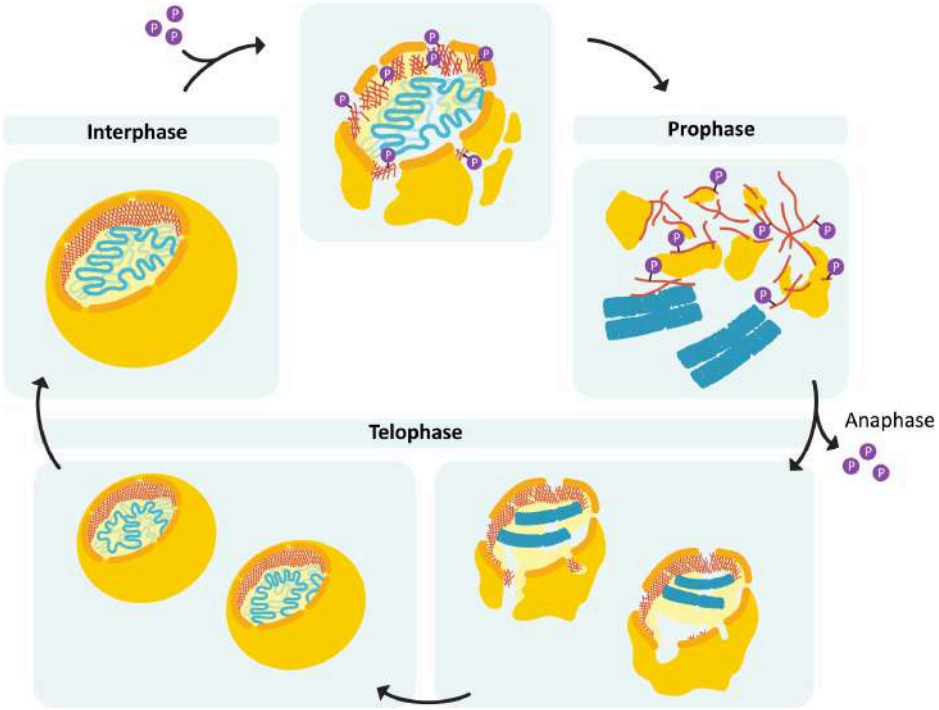


Figure 03-21: The breakdown and reformation of the nuclear lamina during mitosis. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Once mitosis is over and the nuclear envelope must be reformed, the cell removes the phosphate group from the lamins, which allows them to reform the nuclear lamina (Figure 03-21). Since they are attached to both the chromosomes and the nuclear envelope, mitosis ends with all of the components of the nucleus in their proper place, sequestered away from the rest of the cytoplasm.

The Nucleolus

As you already know from Topic 3.1, the interphase nucleus is a highly organized place.

Chromosomes are maintained in their own discrete territories and are further organized into actively transcribing regions and inactive domains (also known as the A/B compartments). On top of that, fluorescence and electron microscopy have identified several specific regions within the nucleus that appear to be associated with particular functions, such as transcription and splicing. One of these regions, the **nucleolus**, deserves further discussion, as it is both the largest and most well-studied subcompartment of the nucleus. Being so large, the nucleolus was first discovered in the 1830s. At that point, all we really knew about it was that cells without a nucleolus did not survive. It wasn't until the 1960s that we were able to figure out the role of the nucleolus in the cell. The nucleolus is generally identified easily in light and electron microscopy as a large, densely staining region that is at or near the center of the nucleus (Figure 03-22). All functional cells appear to have at least one nucleolus with their nuclei, and there are some examples of cells having more than one.

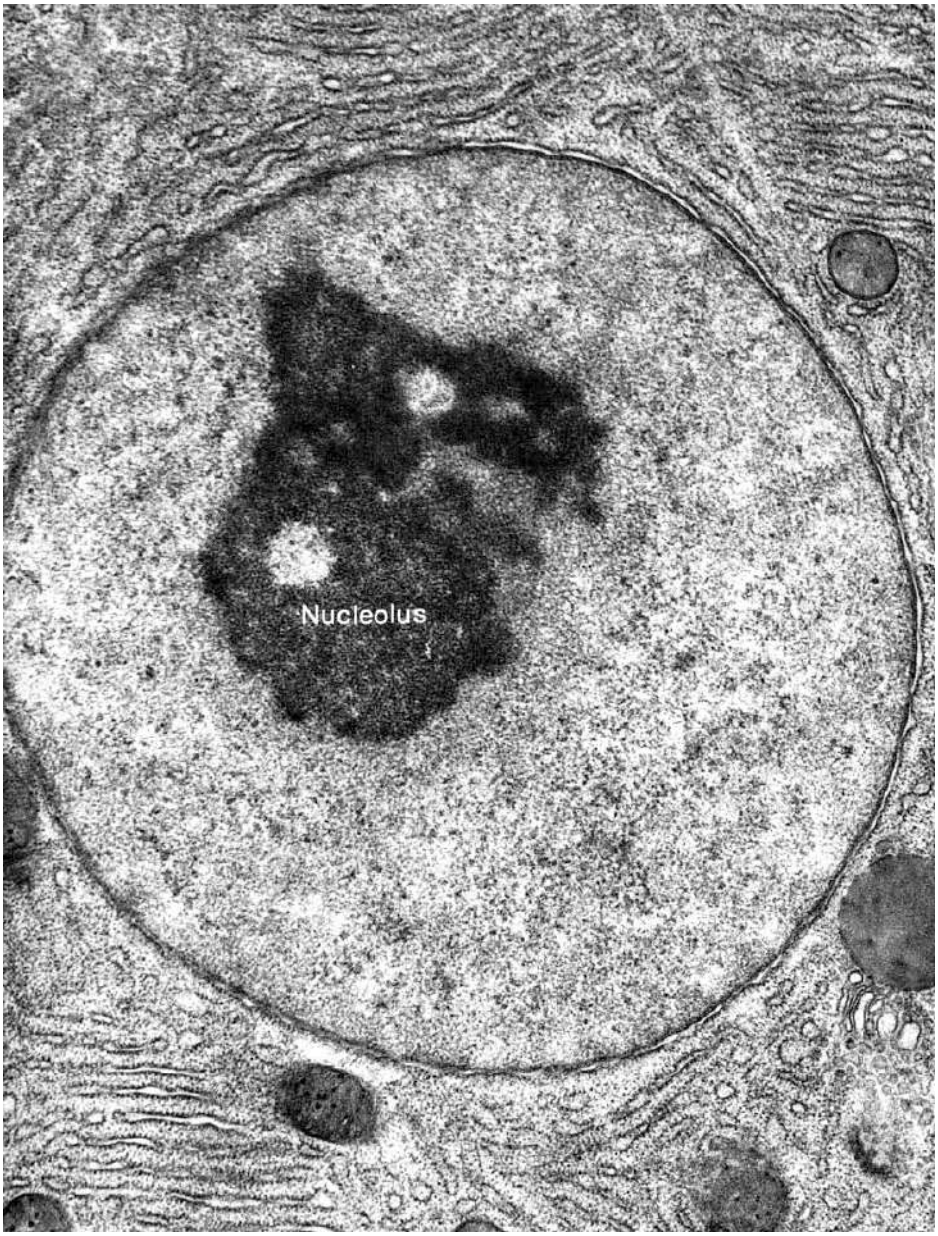


Figure 03-22: Electron micrograph of the nucleolus in a pancreatic acinar cell. Original image from the Cell Image Library, Don W. Fawcett (2011), CIL:10974. Dataset. https://doi.org/10.7295/W9CIL10974. CC-BY-NC-ND 3.0 license.

Note

You may remember from first-year biology that the large and small subunits of the ribosome only come together in the cytosol when translation begins.

The primary purpose of the nucleolus is to synthesize all of the **ribosomes** the cell needs to continue to function. Ribosomes are a type of **ribonucleoprotein**, which means that they are made of both protein and RNA (called ribosomal RNA, or rRNA). The ribosomal proteins are made in the cytosol

using mRNA that was transcribed inside the nucleus (but not in the nucleolus). Once the mRNA has been translated in the cytosol, the ribosomal proteins are then imported into the nucleus via the nuclear pores and assembled with rRNA in the nucleolus to become the large and small subunits of the ribosome.

As an aside, the ribosome is an important example of a **ribozyme**, which means that it is an RNA molecule that has enzymatic activity. The proteins of the ribosome are mainly structural and help hold the rRNA in the right shape to allow for the catalysis of the peptide bond. While ribozymes are not thought to be as common as protein-based enzymes, the functions that they carry out are absolutely vital to cellular survival. In this chapter alone, we have seen two examples of ribozymes: the ribosome and the spliceosome.

Since the cell makes all of its proteins using ribosomes and is almost always in the process of synthesizing hundreds, possibly thousands, of proteins, new ribosomes are *always* required. Thus, the nucleolus is a large, active region of the nucleus. Within the human genome, there are five separate chromosomal pairs that have rRNA genes on them. These regions are known as **nucleolus organizing regions (NORs)**. The DNA in these regions is usually referred to as rDNA because it codes for rRNA. The rRNA genes in the rDNA exist as tandem repeats, which means that there are several copies of the gene in a row. This helps increase the rate at which rRNA can be transcribed. In addition, the nucleolus has structural proteins that interact with the rDNA to collect those regions of the different chromosomes into a single area of the nucleus, thus creating the nucleolus.

During mitosis, the nucleolus must be disassembled so that the separate chromosomes can condense. After mitosis, the new nucleus will begin by building several small nucleoli. As time passes, these small nucleoli fuse to produce a single large nucleolus, which we can see in Figure 03-22. It's important to note that there is no membrane that surrounds the nucleolus (or any of the other discrete regions of the nucleus).

Nuclear Pores and Protein Sorting

The last structural element of the nucleus that we must discuss is the **nuclear pore**. The nuclear pore is the gatekeeper for the nucleus, as it controls all traffic to and from the nucleus. The pores are the primary point of contact between the interior of the nucleus and the rest of the cell. There are as many as 1,000 nuclear pores on the average vertebrate nucleus, and each pore facilitates roughly 1,000 transport events per second! That's a lot of traffic! Since the cell's only copy of its DNA is housed inside the nucleus, the ability of the nuclear pore to carry out its function accurately is essential.

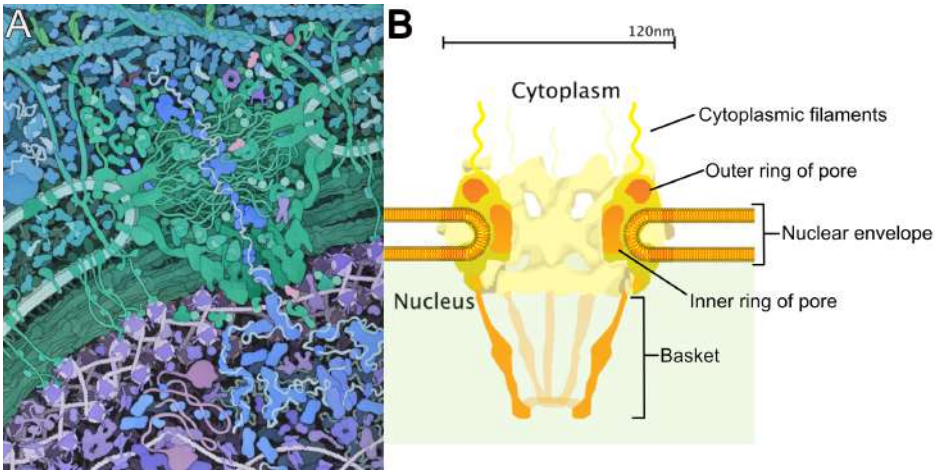


Figure 03-23: Artistic renderings of the nuclear pore based on current molecular data. (A) A nuclear pore is shown in cross section (green) with a molecule of RNA in transport through the pore (light blue) attached to export proteins (darker blue). Modified version of original illustration by David S. Goodsell, RCSB Protein Data Bank, https://doi.org/10.2210/rcsb_pdb/goodsell-gallery-041. Shared under a CC BY-4.0 license. (B) Labeled cross section of a nuclear pore. Adapted from "Side-view Diagram of a Nuclear Pore" by Mike Jones, shared under a CC BY 2.5 license. Image compiled by Dr. Robin Young.

The nuclear pore is one of the largest, most complex structures in the cell. In humans, it is 120 nm in diameter, and it has a mass of roughly 124 megadaltons (MDa)! There are over 450 proteins in each nuclear pore, composed of 34 distinct types of protein. The nuclear pore spans *both* membranes of the nuclear envelope (Figure 03-23). The body of the nuclear pore is a series of large protein rings embedded in the nuclear envelope at a point at which the outer and inner membrane of the nuclear envelope fuse. On the cytoplasmic face of the pore, there are a number of fibrils that stretch out into the cytosol (Figure 03-23B). On the interior of the nucleus is a structure known as the basket, which is shaped roughly like a basketball hoop. A June 2022 issue of the journal *Science* explored the structure of the nuclear pore in detail and had an absolutely spectacular image of the nuclear pore on the cover. It shows the proteins of the pore at near-atomic resolution. Take a look!

The interior of the pore is rather large by cellular standards. In humans, it's about 5.2 nm, but in some species, it's twice that size. There are a number of polypeptide strands that extend out into the interior of the channel (Figure 03-23A). These strands have regions that are rich in phenylalanine and glycine and as such are known as *FG-repeats*, based on the one-letter code for these amino acids. The polypeptide strands act as a diffusion barrier and help control what passes through the nuclear pore.

Mechanisms of Transport through the Pore—Diffusion versus Active Transport

As we mentioned already, the primary point of entry into the nucleus is via the nuclear pore. However, that doesn't mean that everything is transported through the pore in the same way. In fact, there are two possible mechanisms that are commonly used. The specific method a given molecule will use is mostly dependent on the size of the molecule.

1. *Small, water-soluble molecules* can diffuse through the center of the pore without help. The maximum size for diffusion is 30–60 kDa, depending on the organism. This size limit would allow the diffusion of very small molecules, like water, ions, ATP, and other nucleotides, and even some smaller proteins (but not very many).

2. *Molecules that are larger than the diffusion limit* have a much harder time passing through the pore on their own. As such, the cell takes a more active role in managing import of the larger proteins. The cell is selective so that only proteins that *need* to enter the nucleus are allowed to do so. This kind of transport is generally active in that energy will be consumed to facilitate the process.

Examples of Proteins That Enter/Exit the Nucleus

What goes in...

- Histone proteins. One million histones are needed every three minutes during DNA replication for the formation of new nucleosomes. This is about 100 histone molecules per pore per minute that must be actively imported into the nucleus.
- Polymerases and other enzymes required for replication or transcription.
- Transcription factors and other proteins required for the regulation of transcription.
- Ribosomal proteins and other proteins (including spliceosome snRNPs) that form a complex with newly formed transcripts. Approximately 240 ribosomal proteins must be passed through each pore per minute.

What goes out...

Note

The pore structure allows passage of the ribosomal subunits, but not fully assembled ribosomes. Remember that *there is no protein synthesis in the nucleus.*

- Ribosomal subunits—three large and three small subunits per pore per minute.
- RNA-protein complexes. RNA cannot exit the nucleus on its own. Thus, it is always complexed with proteins, which carry the export signal in their primary sequence. This helps control which RNA can leave the nucleus and which must remain.
    - For example, only fully processed and spliced mRNA transcripts are allowed out of the nucleus. The bits that were spliced out are not released but will instead get degraded.
- Transcription factors may also need to exit the nucleus once their job is done. Many transcription factors regularly shuttle between the nucleus and cytosol.

The idea of active, controlled transport through the nuclear pore brings us to a very important theme in cell biology, which is that *movement of proteins out of the cytosol into organelles is a strictly controlled process*. We need to take a step back and address this key concept before we continue to discuss nuclear import/export.

Protein Trafficking to the Organelles Requires Targeting Signals

Virtually all protein synthesis begins in the cytosol. However, not all proteins are made to function in the cytosol. Some proteins are made to function inside an organelle (like, for example, histones, which need to be inside the nucleus). Still others are designed to be used outside of the cell (like collagen, for example). If all protein synthesis begins in the cytosol, then, at some point, the protein will need to be transported out of the cytosol and sent to its final destination before it can become functional. Think of this in the same way that not all of the parts of a car or computer are built at the site where the final product is assembled. Various components are made by different companies and then must be transported to the factory for final assembly.

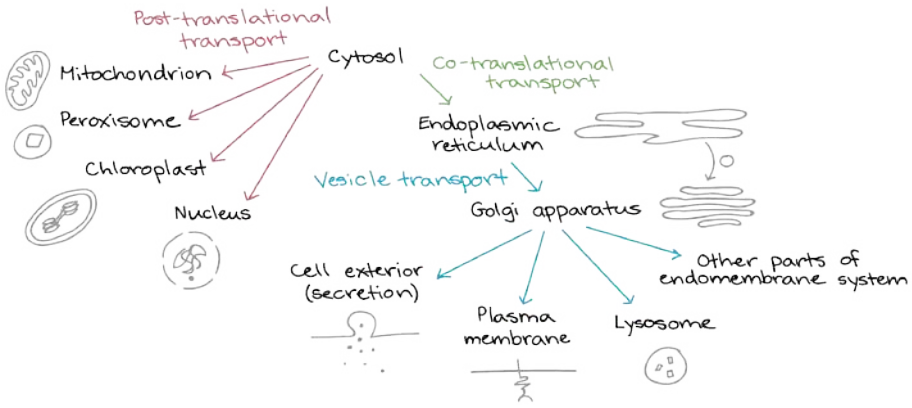


Figure 03-24: Flow chart of protein trafficking from the cytosol. From Khan Academy and shared under a CC BY-NC-SA 3.0 license. Note: All Khan Academy content is available for free at www.khanacademy.org.

The cell needs a mechanism to control all of this traffic (Figure 03-24) and to ensure that the correct proteins are sent to the correct organelles. If, for example, a digestive enzyme destined for lysosome were to end up in the nucleus by accident, it could be a disaster, as the DNA would be in grave danger. To control traffic to the organelles, the cell uses a series of **targeting sequences** that are a part of the primary amino acid sequence of the protein.

Every protein that is targeted to a specific site within the cell must have a destination-specific targeting sequence, or code, associated with it. Additionally, there must also be some sort of specific receptor for the destination it's trying to get to. A good analogy is an instant message, text, or email. An electronic message must be directed to the proper person, so it has a unique identifier (i.e., an email address, phone number, or other electronic handle/username) that must be used, otherwise the intended recipient will not get your message. You may be able to use multiple methods to send your message to the intended recipient, but each identifier corresponds to a single recipient only. This same principal applies to subcellular targeting of proteins.

As you know already, synthesis of all proteins, regardless of their final destination, starts in the cytosol of the cell. (The exception is a small number of chloroplast and mitochondrial proteins, which are synthesized directly inside those organelles. There is a really cool reason for this, which we will discuss in more detail in Chapter 5.) Any differences in the final destinations of proteins are a consequence of targeting signals contained within the primary amino acid sequence of the protein itself. Sometimes, if a protein needs to travel through more than one organelle or is targeted only to a subregion of an organelle, it will actually require multiple sequences.

Import into the nucleus is done via the nuclear pores. Remember that very small, water-soluble molecules may diffuse freely through the pore, but larger molecules (which will mostly be proteins) will require help entering the nucleus. This is where trafficking and targeting sequences come in.

The targeting sequences that scientists have identified for each organelle are generally considered to be **consensus sequences**. This means that while there is some variation in the sequence from protein to protein, the consensus sequence is the most common amino acid sequence that has been identified that works. Mutations in the DNA that encodes the targeting sequence of the protein often result in the protein being mislocalized and thus unable to perform their function in the cell. This shows that these sequences are *necessary* for a protein to be properly localized.

On the other hand, mutations in the DNA encoding the genes for the protein machinery (such as the nuclear pore itself) that control transport into or out of a given organelle are usually lethal, as the transport of all proteins using that machinery will be affected. As a result, the impact will be much greater if the protein machinery is affected by mutation versus if the targeting sequence of an individual protein is mutated.

For proteins to get imported into the nucleus, the targeting sequence includes a series of basic (i.e., positively charged) amino acids on the surface of the folded protein. The consensus amino acid sequence that is most commonly accepted is "*-Lys-Lys-Lys-Arg-Lys-*." We often refer to this by the one-letter code for the amino acids of the sequence, which is *KKKRK*. It is also known as a **nuclear localization sequence**, or NLS for short. Additionally, the NLS is often near, but not at, the N-terminal end of the molecule.

The NLS must be on the surface of the 3D protein in order for the nuclear import machinery to access and identify it for import. There is a proline just prior to the NLS, which is used to produce a bend in the polypeptide and allows the NLS to lay on the surface of the protein. This tells us that proteins are imported into the nucleus after translation, in a fully folded state. This is not true of all organellar import, as you will see in later chapters. For example, in some organelles, proteins are inserted before translation is fully complete (e.g., ER insertion), whereas in others, proteins are inserted after translation but before folding (e.g., mitochondrial and chloroplast insertion).

Nuclear import of proteins carrying an NLS follows a similar pattern of import that is common to virtually all organelles, which means that there are some commonalities in the machinery that will be involved as well. These import processes include the following features:

1. *A protein bearing the correct targeting sequence.* In the case of the nucleus, we need an NLS on the surface of the folded protein.
2. *An import receptor, which helps identify the specific targeting sequence that we need.* In this case, the protein in question is called a **nuclear import receptor (NIR)**. Most commonly, the proteins that carry out this role are known as **importins**.
3. *Some kind of* **translocation channel** *that will allow the protein to cross the membrane an enter the organelle.* The nuclear pore is used for this purpose in the nucleus.
4. *Some form of energy consumption.* This part is quite different for different organelles. In the case of the nucleus, we will need to explore the **Ran cycle** a bit more closely to see how it helps out with nuclear import.

Since we have already discussed the structure of the nuclear pore and of the NLS itself, we'll now focus on the Ran cycle and the way that the NIR gets involved in the process.

Nuclear Import/Export and the Ran Cycle

It is a common theme that some kind of energy input is required for unidirectional transport across membranes. However, the details of where and how that energy is used are different from one organelle to the next. In the case of the nucleus, a special protein called Ran is used. Ran is a small protein that can bind to, and hydrolyze, GTP. It changes conformation based on whether it is bound to GTP (before hydrolyzation) or GDP (after hydrolyzation). This means that it can act as a **molecular switch**, as it is able to change states easily depending on whether it's bound to GTP or GDP. Usually, these switches have an "on" state and an "off" state, much like a light switch. Molecular switches are very common in cell biology, and we will see examples of them in many different chapters of this text.

In the case of the Ran protein, what is most interesting about it is that not only does it switch between the GTP-bound version (Ran-GTP) and the GDP-bound version (Ran-GDP), but it also changes location in the cell depending on which form it is currently in. Ran-GDP can only get converted into Ran-GTP in the nucleus, as the protein required to facilitate this switch (called Ran-GEF) remains bound to chromatin inside the nucleus. On the other hand, its GTPase activity can only be activated in the cytosol, which means that in order for the protein to cycle between these two states, it must also cycle between the two compartments.

Ran gets involved in *nuclear import* from inside the nucleus, where Ran-GTP helps the NIR to release the newly imported protein (Figure 03-25). The NIR, which is still bound to Ran-GTP, now goes back through the pore and out into the cytosol. At that point, the GTP on Ran hydrolyzes and converts to Ran-GDP, which causes the release of the NIR in the cytosol.

Similarly, in *nuclear export* (which is a different process than import, which we will discuss after this), Ran-GTP binds to the **nuclear export receptor (NER)** as well as the cargo protein, and the entire complex moves through the pore together (Figure 03-26). Again, once in the cytosol, Ran-GTP is converted to Ran-GDP, and the NER and its cargo are both released.

Once in the cytosol, in its GDP-bound form, Ran is capable of diffusing back through the nuclear pore on its own, without help from other proteins. After it returns to the nucleus, it gets converted to Ran-GTP and the cycle continues.

The Steps of Nuclear Import

The active transport of proteins *into* the nucleus happens as follows (Figure 03-25):

1. A protein destined for the nucleus will have a preexisting NLS on its surface. The NLS must be accessible for import to be initiated.
2. The NLS region on the surface of the nuclear protein binds to importin, which is an NIR, and the two proteins form a complex.
3. The protein-importin complex binds first to the cytosolic fibril of the nuclear pore. It then works its way through the center of the pore. The FG-repeats mentioned earlier are key to this process, as they interact directly with the NIR carrying the cargo to facilitate transfer.
4. Once inside the nucleus, the Ran-GTP binds to importin, and the cargo is released inside the nucleus.

5. The Ran-GTP-importin complex returns to the cytosol, the GTP on Ran is hydrolyzed to GDP, and Ran dissociates from the importin.
6. Ran-GDP can then return to the nucleus to be converted back to Ran-GTP so that it is ready to help with the next nuclear import event.

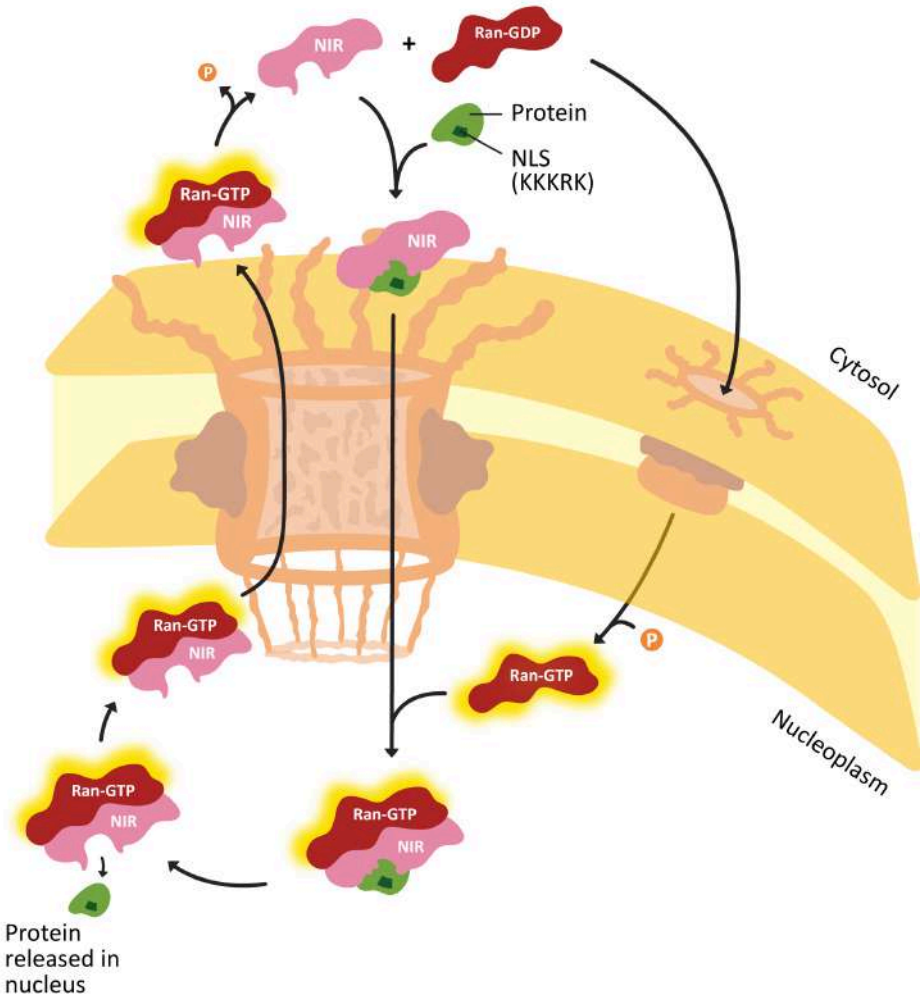


Figure 03-25: Protein import into the nucleus. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The cargo protein can remain in the nucleus for as long as is required. It may be that the protein is a transcription factor that goes through multiple rounds of nuclear import and export. Even if its role is to stay in the nucleus permanently, if the cell goes through mitosis, the nuclear proteins will be released back to the cytosol, as the nuclear envelope breaks down at that time. As such, nuclear proteins will keep their NLS as a permanent feature on their surface. This is different from other protein targeting to other organelles, where the targeting sequence is cleaved after import. We will see these later in the text.

Video 03-07 shows import of proteins into the nucleus. See how many of the different players you can identify. At a few points during the animation, it takes the viewpoint of the molecule, like you are the one traveling through the pore. It can be a bit disorienting if you aren't prepared.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=147#oembed-7

Video 03-07: Molecular animation of protein import into the nucleus.

Nuclear Export

Nuclear export follows a similar pattern to nuclear import, but with some differences in the details (Figure 03-26). For example,

- The **nuclear export signal (NES)** is still an amino acid sequence found on the surface of the protein, but it is not KKKRK. It is a different sequence, which we will not discuss in detail here.
- **Exportins** act as the nuclear export receptor. They are closely related to importins and together form a larger family of proteins known as the **karyopherins**. But their roles are not interchangeable.
- As mentioned earlier, Ran-GTP facilitates export as well as cargo release in the cytosol.

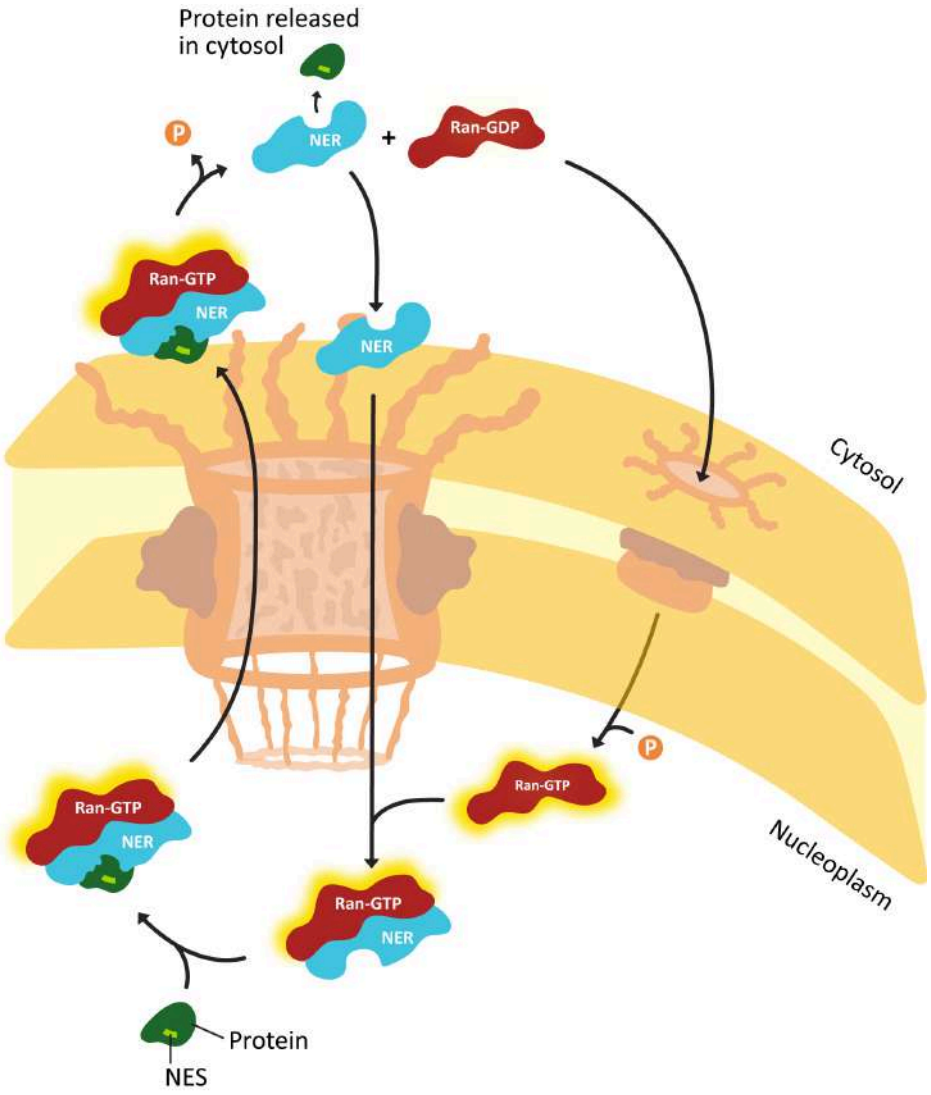


Figure 03-26: Protein export from the nucleus. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

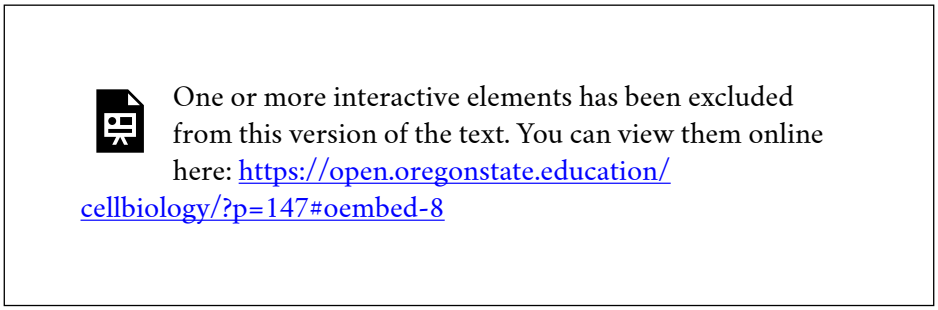
Another major difference is that proteins are the primary cargo for import into the nucleus. However, both proteins and RNA get exported. This is a unique challenge, as the targeting sequences used for import and export are *amino acid sequences*, not nucleotide sequences. So how do we export RNA when it is not capable of carrying the targeting sequence?

The answer is rather clever in that it not only facilitates export of RNA, but it ensures only fully processed, mature RNA is exported and nothing else. In essence, most mature RNA binds to proteins prior to export, and the proteins carry the export sequence. More specifically,

- **Transfer RNAs (tRNAs)** are small enough that they can diffuse, so export signals are not required. In the cytosol, they interact with the tRNA aminoacyltransferase protein and have their amino acid added at the proper site.
- **Ribosomal RNA (rRNA)** is exported only after it is assembled with proteins to form the large and small subunits of the ribosome, so the proteins in the complex can carry the export sequence.
- For **messenger RNA (mRNA)**, the export proteins bind to specific regions of the mRNA, such

as the 5' cap, the poly(A) tail, or the sites where splicing has occurred. All of these are hallmarks of a mature mRNA that will not exist in immature forms or discarded mRNA. These proteins that have bound to the mature RNA will contain nuclear export signals and will facilitate mRNA export from the nucleus.

Video 03-08 provides a visual walk-through of the export process.



Video 03-08: Protein export through the nuclear pore complex.

CHAPTER SUMMARY

As we have seen in this chapter, inside the nucleus, a variety of complex regulatory mechanisms are at play. I hope that you have learned that the nucleus is more than a passive organelle that houses the cell's DNA. Instead, dynamic regulation takes place to regulate compaction of DNA, access to DNA, chromosomal structure, and transport of proteins in and out, as well as synthesis, processing, and preparation of RNA.

Specifically, we discussed that because DNA is so long compared to the nucleus where it is housed, the DNA has to be compacted to be able to fit within the bounds of the nucleus. This is achieved with various packing proteins, which will pack and unpack the DNA as needed.

But just as it needs to be packed to fit, we also discussed how DNA needs to be accessed for transcription of gene products. These gene products rely on transcription factors to fine tune the degree to which they are transcribed into RNA. We discussed some of the processing steps that occur for mRNA prior to export into the cytosol. In particular, splicing allows for multiple protein products that could result from a single gene. Further, we learned about the structure of the interphase nucleus with its nuclear lamina meshwork that provides structural support and protein anchor points as well as the structure of the nuclear import channels, which regulate the import and export of molecules into the nucleus.

Finally, we explored experimental design and a method to study the chromatin structure within the genome using a process known as chromatin immunoprecipitation (ChIP). We also discussed the questions we ask when designing experiments to determine whether the molecular machinery involved is necessary and how to determine whether the machinery in question is sufficient or if something else is required for proper function.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

Topic 3.1: Chromatin and Chromosomes

1. Review of nucleic acid structure: Find structures of the five major nucleic acids online. Identify the following features on the structures and then use this structure as the basis for an argument to explain why base pairing is so precise.
    a. The ribose sugar. What is the major difference between this sugar in DNA and RNA?
    b. The phosphate group(s). Identify the bond broken when a nucleotide triphosphate is incorporated into DNA or RNA.
    c. The purine or pyrimidine ring. How do these structures contribute to base stacking and 3D structure of the final molecule?
    d. The H-bonds involved in base pairing.
    e. The 3' and the 5' ends of a strand of nucleic acids.
2. RNA is single stranded, but the bases are still capable of base pairing. Explain why this is an advantage for the cell.
3. DNA and RNA are said to be polar because of the presence of the 3' and 5' end specifically rather than due to any charges that might be present. Explain why this is so.
4. Explain how proteins and DNA interact to form chromosomes, starting with the 2 nm naked DNA molecule.
5. Explain how the properties of the R groups of the amino acids in histones must be arranged in order for them to be able to form the nucleosome core (i.e., they must be able to interact specifically with each other) and for them to promote interactions with the acidic backbone of the DNA.
6. What is the difference between chromatin, chromatids, and chromosomes? How are they related to each other?
7. What are the major differences between chromosomes in interphase versus during mitosis?
8. Can you list the different ways that the genome is organized three-dimensionally within the interphase nucleus? How are they related to each other?
9. What is the difference between euchromatin and heterochromatin?

Topic 3.2: Regulation of Gene Expression

1. Where and how are histones most commonly modified? What is the result?
2. How does the 3D arrangement of the genome contribute to the fact that Eukaryotic genes

can have regulatory regions that are thousands of base pairs away from the gene that they regulate?

3. List the ways that chromatin can be remodeled via modifications to individual nucleosomes.
4. What are transcription factors? How do they bind to DNA?
5. What's the difference between a basal transcription factor, an enhancer, and a suppressor?
6. Define the following: 5' cap, intron, exon, spliceosome, snRNPs, polyadenylation site, alternate splicing, exon choice, and promoter choice.
7. snRNPs play a critical role in the intron excision process. What is this, and why is it important that the process of intron excision be absolutely precise?
8. Alternative splicing patterns are usually tissue specific. Promoter choice is also often region, stage, or tissue specific. Explain how eukaryotes use these strategies to their advantage.
9. What is ChIP, and how can it be used to learn about chromatin structure?
10. What is the difference between something being necessary to a particular function and it being sufficient? Is it possible to be necessary but not sufficient? What about being sufficient but not necessary?

Topic 3.3: The Interphase Nucleus—Structure, Function, and Protein Import

1. Discuss the relationship between the nuclear envelope and the endoplasmic reticulum.
2. What is the functional role of the nuclear lamina?
3. One of the key regulatory proteins that initiates the cell division process phosphorylates the proteins of the nuclear lamina. When this happens, the proteins dissociate. From what you already know of the events of mitosis (from general biology), what does this tell you about the role of the nuclear lamina?
4. Describe the three major structural elements of the nucleolus.
5. Sketch a nucleus on your paper. Draw and label the following:
    a. nucleolus
    b. nuclear envelope
    c. nuclear pores
    d. nuclear lamina
    e. heterochromatin
    f. euchromatin
6. Compare and contrast nuclear import and export.
7. The Ran cycle can be somewhat confusing to make sense of. Can you put together how Ran can be involved in *both* nuclear import and export? It may help to sketch out the processes.
8. Why can't RNA carry a nuclear export signal (NES), which is necessary for export from the nucleus? How does the cell then export RNAs if they don't carry the signal?

REFERENCES

Brewis, H. T., Wang, A. Y., Gaub, A., Lau, J. J., Stirling, P. C., & Kobor, M. S. (2021). What makes a histone variant a variant: Changing H2A to become H2A.Z. *PLOS Genetics*, *17*(12), e1009950. https://doi.org/10.1371/journal.pgen.1009950

Skibbens, R. V. (2019). Condensins and cohesins—one of these things is not like the other! *Journal of Cell Science*, *132*(3), jcs220491. https://doi.org/10.1242/jcs.220491

CHAPTER 4.

THE ENDOMEMBRANE SYSTEM

INTRODUCTION

Since the authors of this textbook both did research in the area of endomembrane biology, we must admit that we are somewhat partial to this topic. That being said, the endomembrane system is arguably the most interesting and complex of all of the cellular structures we cover. It's also quite challenging to study, as its function is essential to cell survival. Anything in the cell that is "essential" is also difficult to study, as the cell dies quickly if that function is disrupted.

There are several membrane-bound organelles that work together to form the endomembrane system: the endoplasmic reticulum (or ER for short), the Golgi apparatus (named after a person, so its name is capitalized), endosomes, lysosomes (called vacuoles in some taxa), and the vesicles that traffic cargo from one organelle to another. This system is extensive, and it is somewhat messy to differentiate one part from another. In this chapter we do our best to create order from the chaos and highlight how we study the endomembrane system as well as what we don't yet know.

TOPIC 4.1: THE ENDOPLASMIC RETICULUM AS THE GATEWAY TO THE ENDOMEMBRANE SYSTEM

Learning Goals

- Explain the structural and functional relationships between the different compartments of the endomembrane system and identify them on different micrographs.
- Describe how proteins are targeted and imported into the endoplasmic reticulum and compare these mechanisms to protein targeting and import into the nucleus.
- Predict the targeting sequences required to insert a protein into the endoplasmic reticulum (ER) membrane in any orientation and predict protein topology from a corresponding domain map.
- Discuss the role of chaperones in protein folding and the role of the proteasome in the Unfolded Protein Response.
- Describe general principles behind protein glycosylation in the ER and Golgi.

Introduction to the Endomembrane System

The endomembrane system consists of the **endoplasmic reticulum (ER)**, **Golgi apparatus**, **lysosomes**, and **endosomes**. These compartments are involved in a great many cellular functions,

including the processing of proteins that are destined for export from the cell, dealing with proteins that have been brought in from the outside of the cell, lipid synthesis, and a variety of signaling events.

The endomembrane system has a number of compartments, and cargo travels from one compartment to the next using smaller, membrane-bound structures known as **vesicles**. We say that the "start" of the endomembrane system is the ER, as this is the point of entry for newly synthesized proteins (Figure 04-01). Once proteins enter the ER they never return to the cytosol; they are carried by vesicle transport to the other compartments of the system. This flow of vesicles is highly regulated. The "end" of the endomembrane is usually considered to be the cell exterior, as proteins that pas through the entire length of this system, without being diverted, will eventually undergo **secretion**, ending up in the extracellular space, or embedded in the plasma membrane.

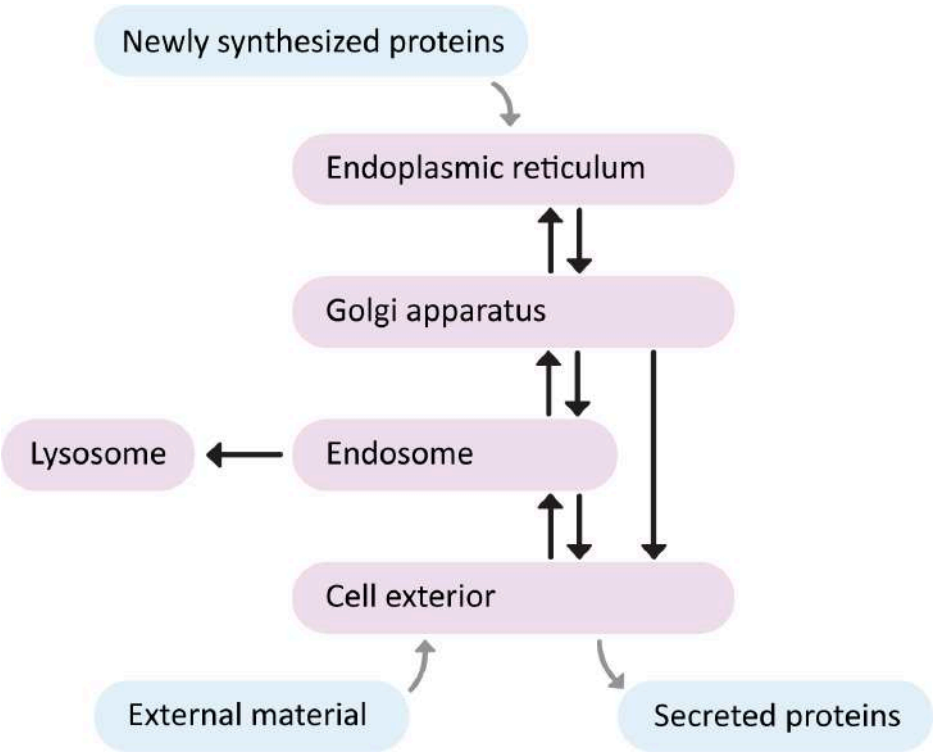


Figure 04-01: The compartments of the endomembrane system, and the flow of traffic between them. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

There are three major pathways to travel through the endomembrane system. See if you can trace each one in Figure 04-01.

Re: The Nuclear Envelope

It is sometimes said that the *nuclear envelope* is also a part of the endomembrane system. This is because the outer membrane of the nuclear envelope is continuous with the ER. This means that the space inside the ER, and the space between the membranes of the nuclear envelope (called the *perinuclear space*) is continuous with the ER lumen. In this textbook we don't really discuss the

nuclear envelope as part of the endomembrane system, even though it is continuous and may get involved in some of the same functions that the ER does. This is mostly to simplify language, as we discuss the various structures and their function.

1. The **secretory pathway**, which is also often called the "default pathway" through the organelles. It is the path that most newly synthesized ER-targeted proteins will take as they eventually exit the cell. Proteins start in the ER, move to the Golgi and then finally moved to the plasma membrane.
2. The **lysosomal pathway**, which is the path that newly synthesized digestive enzymes will take as they move to their eventual destination (the lysosome). Proteins start in ER, transported to the Golgi, delivered to the endosomes when they finally go to the lysosome.
3. The **endocytic pathway**, is the path inward from the cellular exterior for proteins and substances that are unable to cross the membrane through diffusion or by transport protein. Material brought in via endocytosis will travel to endosomes and likely end up in the lysosome, where they can be degraded, and their components recycled into building block molecules. Molecules entering the cell this way are said to have undergone **endocytosis**.

Proteins that enter this interconnected set of organelles can only enter at the "start" of the pathway. So newly synthesized proteins that are destined for the secretory or lysosomal pathways must always enter at the ER. Material coming into the cell from the extracellular space via endocytosis begins at the plasma membrane. Additionally, proteins that enter the endomembrane system always travel the same route. They don't skip compartments, or return to the cytosol (unless they are misfolded and need to be destroyed…but more on that later…).

In this topic, we have two main ideas to discuss:

1. How proteins are modified after translation. This is called **protein processing**.
    - Since protein processing begins in the cytosol and (sometimes) ends in the endomembrane system, we're going to talk about this first, and then get into the details of how the endomembrane system works.
2. How newly synthesized proteins enter the ER, thus gaining access to the rest of the endomembrane system.

Protein Processing in the Endomembrane System

As you (hopefully!) remember, protein translation happens in the cytosol. (As always, the introduction is there for you to review should you need it.) Once translated, *all proteins are processed*. Processing can include any number of events, such as the following:

- Folding: *All* proteins are folded to generate their 3D structure; thus, *all* proteins are processed in at least this one way. They will form the three (or four) levels of protein structure we mentioned in Chapter 2. The final folded 3D structure is based on the primary sequence of the protein.
- Removal of the first methionine at the N-terminus. Remember that the codon that serves as the "start" codon is also the codon that codes for methionine. In many proteins this

methionine is removed once the protein is translated.

- Methylation
- Phosphorylation
- Acetylation
- Formation of disulfide bridges (most common in secreted proteins)
- **Glycosylation** (only in proteins destined for the cell exterior or the lysosome)
- Cleavage and more…

We will discuss some of these in more detail, but not all. We will focus first on protein folding and how it is managed by the cell. Then we'll look briefly at how and where disulfide bridges form. Later in the chapter, we'll look at how proteins are glycosylated, since that process is primarily the job of the Golgi apparatus.

Protein Folding

As the newly synthesized polypeptide emerges from the ribosome, it will immediately begin to fold based on its primary structure. Protein folding is a spontaneous process in which the polypeptide will take on the lowest energy conformation possible. This you should already know from our discussion of membrane proteins in Chapter 2. This is not to say that the protein will spontaneously "find" the proper conformation. Folding is extremely complex, and we continue to learn about it. What we do know boils down to this:

- Proteins can successfully fold a lot of different ways. They do not follow the same path each time.
- It is a process that can go wrong very easily. Not every translated protein ends up properly folded. Misfolded proteins are usually destroyed by the cell.

Figure 04-02, below, perfectly illustrates this concept. In it, we see many of the different steps a protein can take to get from "unfolded" to "properly folded." Note that this image excludes all of the ways that a protein could fold and end up at a "dead end," where it would not be able to find the appropriate final conformation. There are just as many, if not more, ways a protein can misfold. Since misfolded proteins are destroyed by the cell, it is unclear exactly how often it happens, but some estimates suggest that as many as 50% of proteins misfold!

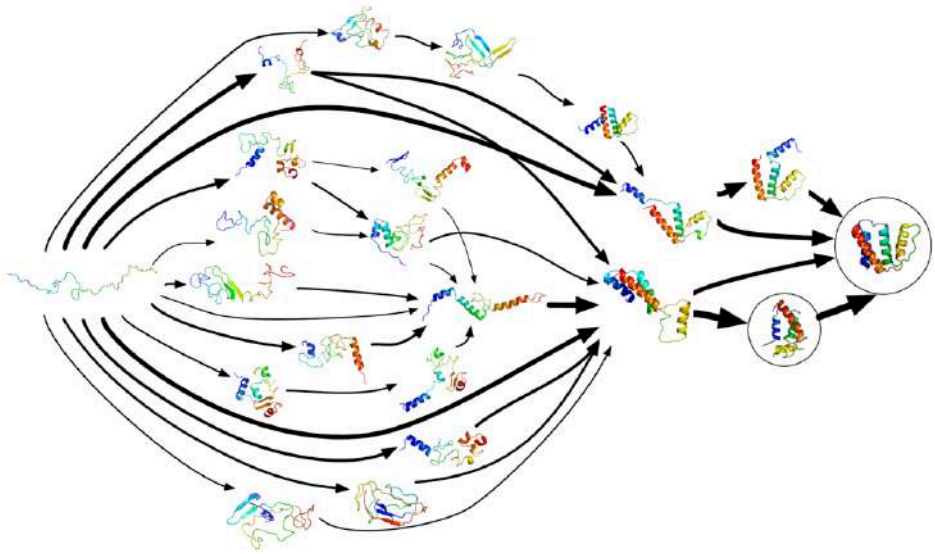


Figure 04-02: Some of the different paths a protein can take to become a fully and correctly folded protein. It's worth noting that this image does not include any of the possible "dead-end" paths, where the protein is unsuccessful in reaching the proper final fold. "ACBP MSM from Folding@home" by Vincent Voelz is shared under a CC-BY-SA 3.0 license and is based on work published by Voelz et al. (2012).

Chaperone Proteins Help Ensure Proteins Fold Properly

The capacity for a protein to spontaneously fold correctly is important for more reasons than you might expect. Improperly folded proteins can be very detrimental to the cell. Not only would the proteins be unable to perform their function, but they might be insoluble and form large aggregates in the cell. In fact, misfolded proteins are a common source of disease, including Alzheimer's, type 2 diabetes, cystic fibrosis, Parkinson's, and prion diseases.

To ensure that the highest number of translated polypeptides manage to successfully fold, **chaperone proteins** often aid in the process (Figure 04-03 and Video 04-01). The new polypeptide forms a complex with chaperones that facilitate folding. Different chaperones help in different ways. Some may help simply by binding to specific regions of the polypeptide (Figure 04-03B) to prevent them from folding too early. This is how a chaperone known as *HSP70* works. Others, such as *HSP60 (or chaperonin)*, act as a chamber, providing a protected space away from the rest of the cytosol, where the protein can fold in isolation (Figure 04-03A).

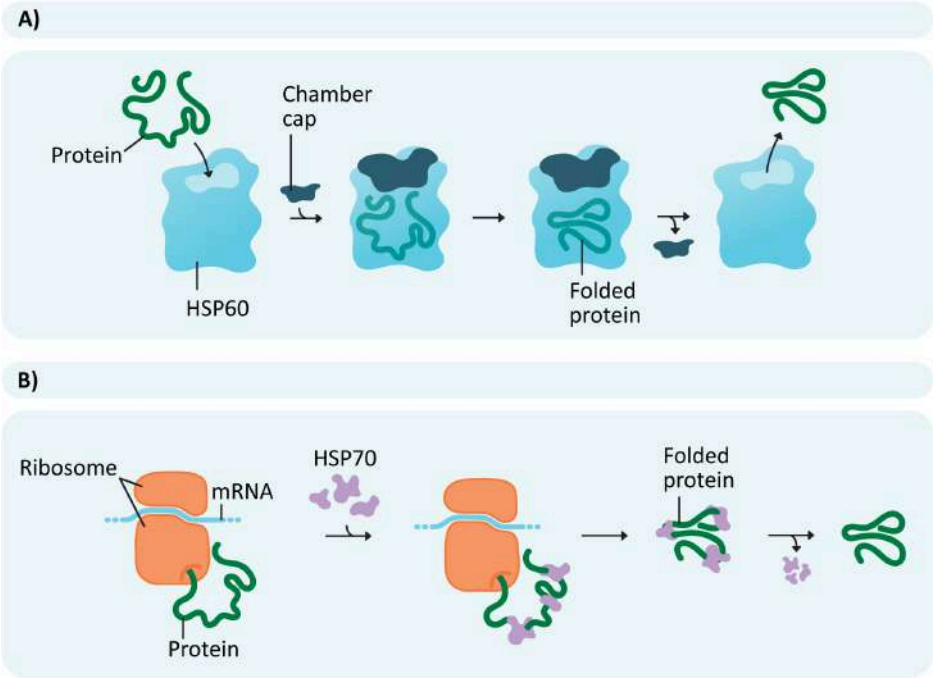


Figure 04-03: Chaperone proteins help proteins fold. (A) HSP60-style chaperone proteins form a chamber, creating a specialized environment for client proteins to fold. The protein enters the chamber, and a chamber cap is placed. When the protein is done folding, it is released. (B) HSP70-style chaperone proteins bind to client proteins as they are synthesized, allowing certain areas to wait before being folded. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-1

Video 04-01: Chaperones and misfolded proteins. The process of protein folding and the role of molecular chaperones HSP60 and HSP70 are discussed.

Misfolded Proteins Are Sent to the Proteosome

It is vital that proteins either fold correctly or get disposed of properly when they misfold. The most obvious reason for this is that misfolded proteins are unlikely to function correctly, and this is true. However, it is equally important that these dysfunctional proteins get sent to the proteasome for degradation. The reason is that a misfolded protein is more likely to form complexes with other misfolded proteins nearby. This is due to exposed nonpolar regions on misfolded proteins that would otherwise be hidden away inside the protein. Misfolded proteins often form large clumps, or aggregates, inside the cell. The result of this is a giant blockage of protein inside the cell that gums up the works and inhibits function. If the aggregate is big enough or sticks around for a long time, it

can actually kill the cell. There are a number of diseases in which protein aggregates accumulate in cells. For example, in Alzheimer's disease and other forms of dementia, large protein aggregates are observed inside neuronal cells. One hypothesis is that these aggregates disrupt cellular function, and the cells die as a result. In sickle-cell anemia, the hemoglobin proteins misfold due to a mutation and form large clumps that ultimately destroy red blood cells.

The **proteosome** is a large complex in the cytosol of the cell whose sole function is to destroy proteins that are damaged, misfolded, or no longer needed. It is one of the larger protein complexes inside the cell.

Want to compare the size of the proteosome to other cellular proteins? *Follow this link to the Protein Data Bank interactive map!*

How does the proteosome work? Proteins that need to be degraded are tagged with a small peptide tag known as **ubiquitin** (Figure 04-04 and Video 04-02). The ubiquitin tag is recognized by other proteins, whose task is to deliver the ubiquitinated protein to the proteosome. Once there, the protein is threaded into the interior of the cylindrical center of the proteasome, and the peptide bonds are broken in a chemical process known as **proteolysis**. That way, the amino acids in the protein are kept and can be recycled by the cell. The ubiquitin is also removed and recycled.

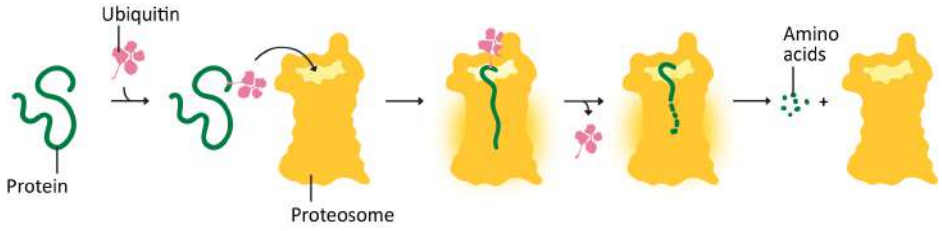


Figure 04-04: Degradation of ubiquitin-tagged proteins by the proteosome. Ubiquitin attaches to proteins, which sends them to the proteasome. The protein enters the chamber of the proteasome and is digested. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-2

Video 04-02: The proteasome: the cell's trash processor in action

Protein Folding and the Formation of Disulfide Bridges

Disulfide bridges are covalent bonds that form between the sulfurs of cysteines in an amino acid chain. Usually, the cysteines are far away from each other in the linear polypeptide chain but are brought together during folding. Enzymes usually facilitate the formation of disulfide bridges (also known as S-S bonds). These bonds help stabilize the 3D structure of a protein and/or help hold

different subunits together. We touched on them briefly in Topic 2.3 as well, when discussing protein folding.

Interestingly, disulfide bridges only form under very specific conditions. They can only be formed under oxidizing conditions within the cell. (*Hint: Remember your general chemistry on redox?*) In reducing environments, they are unstable and tend to fall apart. What is the most interesting about this is that the cytosol tends to be a more reducing environment, which means that disulfide bridges are not easily formed or maintained there. On the other hand, the interior of the ER, known as the **ER lumen**, and the extracellular space tend to be oxidizing environments. This means that proteins that have moved through the ER during folding are far more likely to have disulfide bridges in them. We'll see more on which proteins are folded in the ER and why a little bit later in this topic, but in a nutshell, proteins that are destined for any compartments of the endomembrane system or proteins that are targeted to the plasma membrane and/or the exterior of the cell are the ones that enter and do at least some of their folding in the ER.

The ER Is the Point of Entry into the Endomembrane System for Newly Synthesized Proteins

The ER consists of flattened membrane sacs, known as **cisternae**, and tubules. It is directly connected to the outer membrane of the nuclear envelope, but unlike most of the cartoons of the ER found in textbooks like this, it stretches throughout the entire cell. There is no part of the cell that is far from the ER. The ER has a number of different regions in it that are all connected together (Figure 04-05):

- **Rough ER (rER)**. The cytosolic surface of the rER membranes has docked ribosomes that are synthesizing proteins for import into the ER. This is the site of synthesis for proteins destined for secretion, lysosomes, or membranes.
- **Smooth ER (sER)**. This is continuous with rough ER, as shown in the photo below, and is the site of lipid and steroid synthesis. As mentioned in Chapter 2, new lipids and membranes are made in the sER.
- At the site where these two types of ER meet, there is a third form of ER known as **transitional ER (tER)**. This is the site where vesicles usually form and newly synthesized proteins exit the ER in order to move on to the next destination (i.e., the Golgi apparatus). As such, this region is also known as an **ER exit site**.

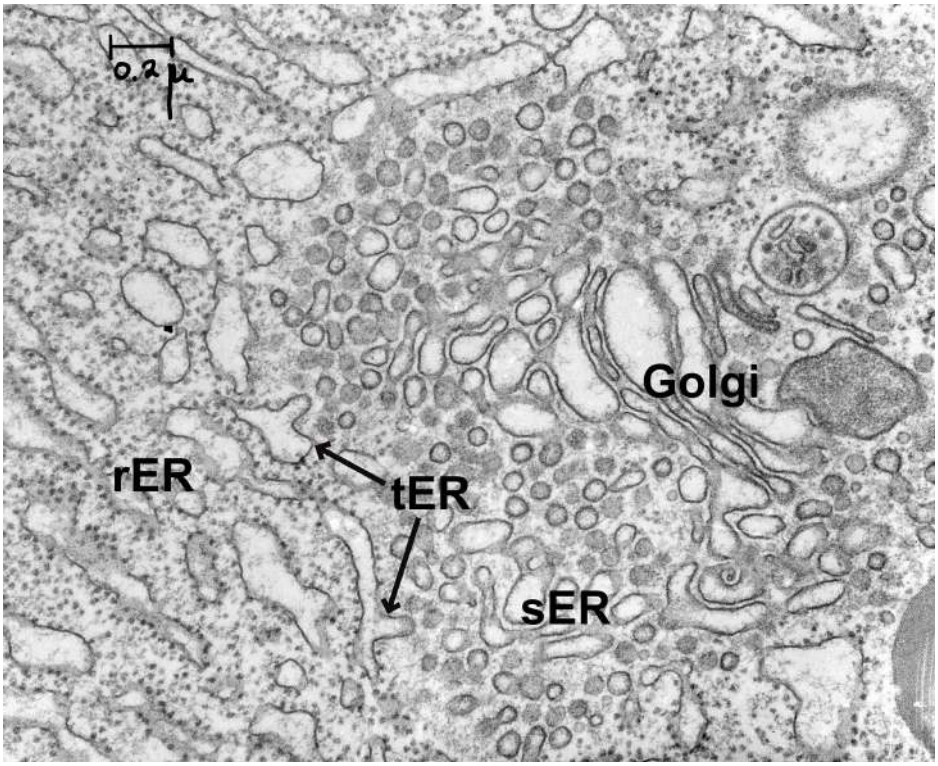


Figure 04-05: Electron micrograph showing the ER and Golgi in a pancreatic cell of a guinea pig. The rough ER (rER) can be identified by its "rough" appearance due to the ribosomes docked on the surface of the membrane. The smooth ER (sER) has no ribosomes but still looks like it should be a bunch of interconnected tubules that have been cut in cross section. The point at which they join and the ER transitions from rough to smooth is known as transitional ER (tER) and is an exit point for vesicles leaving the ER. The Golgi apparatus of the cell is also visible here and can be identified by the fact that it resembles larger sacs that are somewhat flattened. Original micrograph: George E. Palade (2012), CIL:37239, Cavia porcellus, pancreatic cell. CIL. Dataset. https://doi.org/10.7295/W9CIL37239. Labeled by Robin Young. Licensed by CC BY-NC-SA 3.0.

There are a couple of things to note about the traffic between the ER and the cytosol:

- The ER is the point of entry into the entire endomembrane system for newly synthesized proteins, so there is a lot of traffic heading from the cytosol to the ER.
- On the other hand, once a protein enters the ER, it does not return to the cytosol.
- Proteins that need to leave the ER to travel to the next compartment in the endomembrane system (which is the Golgi) will get packaged into vesicles, which will bud from the tER, and then fuse with the Golgi membrane. This way the protein can travel between compartments without returning to the cytosol.

Misfolded Proteins Leave the ER and Are Degraded in the Cytosol

There is one notable exception to the flow of protein traffic from the ER to the rest of the endomembrane system. Misfolded proteins that cannot be saved by the chaperone proteins must be sent to the proteasome for degradation. Since the proteasome is in the cytosol, this is the one instance when a protein will be transferred back across the ER membrane to the cytosol.

Just like in the cytosol, an accumulation of misfolded proteins in the ER lumen triggers the production of chaperone proteins and the expansion of ER, which can help reduce misfolding and

aggregation. This is known as the **unfolded protein response (UPR)**. Dealing with improperly folded proteins is thought to be a massive undertaking in the cell, which gets much worse when mutations exist that result in proteins that don't fold efficiently.

Many proteins with multiple subunits, such as antibodies, are assembled in the ER. If these proteins are not properly assembled (e.g., via the formation of disulfide bridges), they will also trigger the UPR to address it.

Cells make lots of mistakes in the assembly of proteins. Proteins are made of hundreds, if not thousands, of amino acids, making them easy to misfold or misassemble. A huge part of the job of the ER is to ensure that properly folded proteins move on and misfolded ones are dealt with efficiently and don't build up inside the cell and block traffic.

Entry/Exit from the ER Is Strictly Controlled

Just like the nucleus, the cell controls precisely what is allowed to enter and exit the ER. However, since the role of the ER is different from that of the nucleus and its structure far more intricate, controlling access is also more complex. Like the nucleus, only proteins with the proper targeting sequence will be allowed to enter the ER. Of those that are allowed to enter, some will become residents of the ER, while others will simply be passing through on their way to other destinations farther along in the secretory or lysosomal pathways.

Proteins Are Inserted into the ER Co-translationally

This is our second organelle in which we discuss protein import. Our first example was nuclear import (in Chapter 3). In the next chapter (Chapter 5), we will see how proteins enter the mitochondria and chloroplasts. In the case of the endomembrane system, there is more than one process to consider, as several organelles are included. Each compartment will have its own unique sequences, which must be properly read at the correct moment in the pathway. The additional sequences will be discussed later, when we explore the function of each of the compartments.

The first thing you must know is that *translation of all proteins, including ER proteins, begins on free ribosomes in the cytosol* (Figure 04-06). Whether or not a ribosome settles on the rER or stays in the cytosol to complete translation is dependent entirely on the protein it is currently translating. A ribosome that is currently translating a protein destined for the ER will attach itself to the surface of the rER during translation through a series of cues found in the amino acid sequence of the new protein. It is this field of ribosomes actively translating on the surface that gives the ER it its "rough" appearance. Smooth ER (sER) does not participate in protein synthesis (sER is important in lipid synthesis as well as metabolizing many toxic chemicals, like ethanol).

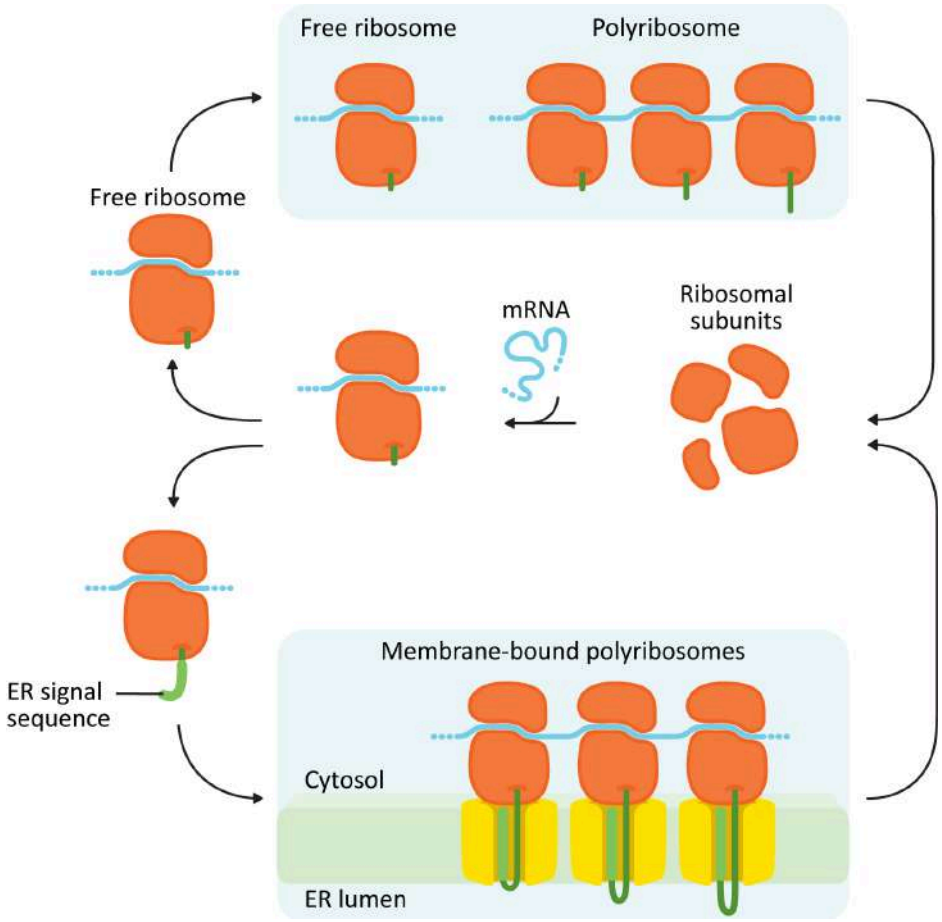


Figure 04-06: All proteins are translated by the same pool of ribosomes in the cytosol. As the ribosomes bind to the mRNA (blue thread), they begin translation (emerging green line). If they contain an ER insertion sequence, they are taken to the ER membrane. However, if no such sequence is present, the ribosomes continue to translate the protein in the cytosol. Once translation is finished, the ribosomes separate and are free in the cytosol. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Structurally, the proteins entering the ER fall into two major categories:

1. **Soluble proteins**: These proteins are completely translocated across the ER membrane into the ER lumen. They do not contain any membrane-bound portions.
2. **Integral membrane proteins**: These proteins are only partially translocated into the ER and end up getting "stuck" with part of the protein embedded inside the membrane. These proteins may be destined for the ER, membranes of another organelle (Golgi, lysosomes, or endosomes), or the plasma membrane. Once a protein is inserted into a membrane, it cannot be removed.

The difference between the formation of a soluble protein and a membrane-bound protein is due to the number and placement of the **ER insertion sequences**. These sequences are used to identify when the ribosome should dock on the ER membrane and also which regions should become transmembrane domains. Here are some things to know about ER insertion sequences:

- Unlike the nuclear localization sequence (NLS), the specific order of amino acids in an insertion sequence is not as important as the chemical properties of the amino acids within

the sequence. In all cases, the ER insertion sequence is about *8–10 nonpolar amino acids in a row*.

- These ER insertion sequences go by a variety of names, depending once again on their location in the protein.
    - If the sequence is directly at the N-terminus, it is called the *signal sequence*, the *signal peptide*, or an **N-terminal START sequence**.
    - If the sequence is anywhere else within the primary sequence, it will be called an **internal START** or **STOP transfer sequence**, depending on how it aligns with the other ER sequences in the polypeptide.
- Despite the different names and different locations, the sequence itself remains the same. It is still 8–10 nonpolar amino acids, as mentioned above. The location of these targeting sequences within the polypeptide chain determines whether a protein is soluble or an integral membrane protein and also impacts the structure and orientations within the membrane.

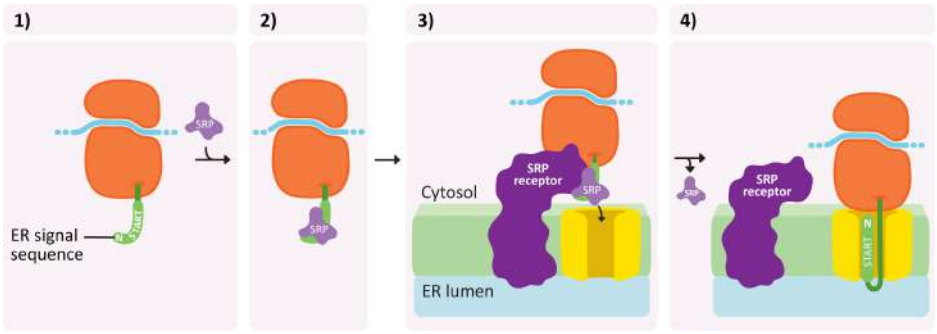


Figure 04-07: Insertion of a protein into the ER. Insertion requires an ER targeting sequence, the signal recognition particle (SRP) receptor, and translocation channel. The protein is recognized by the sequence and brought to the translocation channel to be inserted into the ER lumen. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Insertion into the ER takes place in a number of steps and is illustrated in Figure 04-07:

1. An ER insertion sequence is translated and almost immediately recognized by a ribonucleoprotein known as the **signal recognition particle (or SRP)**.
2. The SRP binds to this sequence and inhibits translation.
3. The entire complex (ribosome + mRNA + partially translated protein) is brought to the ER and binds to a special SRP receptor protein in the ER membrane. The ribosome becomes attached to a **translocation channel** for the newly synthesized polypeptide. This attachment is facilitated by the SRP receptor and requires GTP as an energy source.
4. As the ribosome becomes attached, the SRP is removed and translation resumes, but now the new protein is being pushed through the translocation channel into the ER lumen.

Video 04-03 does an excellent job of showing the ribosome docks onto the ER at the molecular level. However, it has no narration, so you're going to have to consider when each step is happening on your own.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-3

Video 04-03: Initiation of protein insertion into the ER via the translocation channel.

Different Types of Protein Insertion into the ER Membrane

Since both soluble and membrane-bound proteins are inserted in this way, using these same insertion sequences, it stands to reason that we need to explore a few different scenarios. In all cases, the location and order of the various ER insertion signals will determine whether a protein is soluble or membrane bound as well as how many times it passes through the membrane. Let's explore how this works.

Soluble Proteins

Soluble proteins need only a *single* transfer sequence, and it is always found at the *N-terminus* (Figure 04-08). After the signal sequence is recognized, the ribosome docks, and the polypeptide is threaded through the translocation channel into the lumen of the ER as it is synthesized. The signal sequence remains embedded in the membrane and is later cleaved off by a protein called the **signal peptidase**. Once that happens, the new protein is free and soluble in the ER lumen.

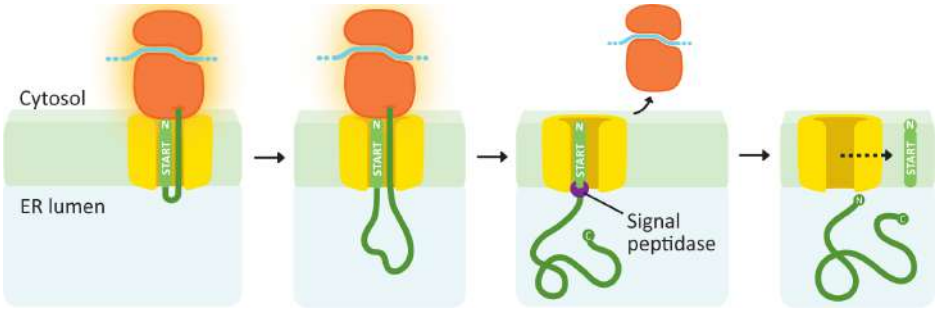


Figure 04-08: Insertion of a soluble protein through the ER membrane. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Membrane Proteins

A key point in the production of membrane proteins is that the orientation of a protein in the membrane is established when it is first inserted into the membrane, during translation. The orientation of the protein persists throughout its life-span, even as it is shuttled from one compartment to the next. That is, *the cytosolic side of the protein remains on the cytosolic side of the membrane throughout its entire life*. More on this later.

As membrane proteins are being translated, ribosome docking and co-translational insertion will not begin until an ER insertion sequence is encountered. Thus, the very first insertion sequence

encountered is called the "START transfer" sequence. The first START may be at the N-terminus, as we saw in the previous example, but it doesn't have to be.

The new protein will continue to be translated into the ER until a stop codon is reached (which ends translation) or a second insertion sequence is encountered. The second insertion sequence serves as a "STOP transfer" signal, which will close the translocation channel, release the ribosome, and stop co-translational insertion. When the START and STOP sequences are inside the amino acid sequence (i.e., not at the N-terminus), they serve as transmembrane domains for the growing protein. If the ribosome is still translating the protein after the STOP is encountered, it will remain tethered to the ER via the translating protein until either a new START sequence is met or translation ends. This process can be repeated many times in a single protein for as many ER insertion sequences as exist in the primary structure.

To further clarify how the ER insertion sequences are used to create membrane proteins, we have provided a number of examples.

Example #1

A single internal START sequence produces a protein with one transmembrane domain and the N-terminus on the cytosolic side.

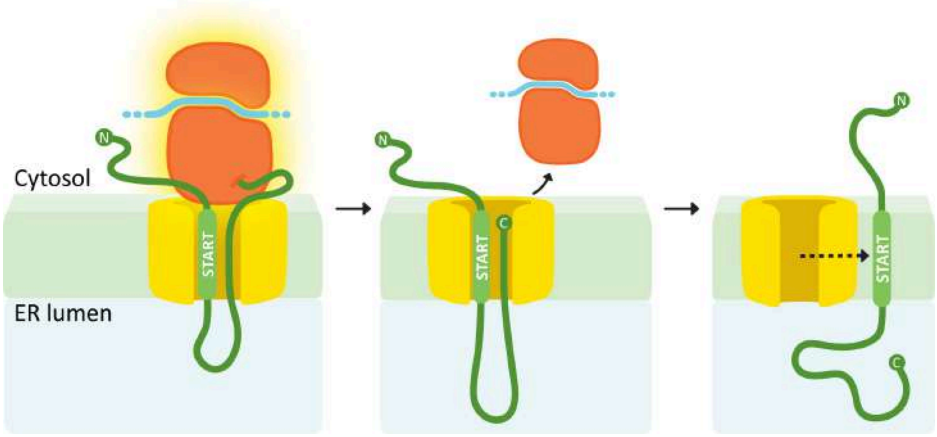


Figure 04-09: Insertion of a protein with one internal START sequence. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

In this example (see Figure 04-09), the internal ER insertion sequence is recognized partway through protein synthesis, and the ribosome is brought to the translocation channel. That sequence is the START transfer, and everything after it is threaded through the translocation channel. Once translation is finished, the ribosome is released. The translocation channel opens and allows the protein to diffuse laterally into the membrane. The internal START transfer sequence is *not* cleaved and instead remains embedded in the membrane, becoming the transmembrane domain for the protein. As discussed in Chapter 2, the transmembrane domain of the protein holds the protein in the membrane because of the very strong association between the nonpolar amino acids in this region and the nonpolar lipid tails in the lipid bilayer. The N-terminus remains in the cytosol, while the C-terminus ends up inside the ER lumen.

Example #2

An N-terminal START transfer sequence followed by a STOP transfer creates a single-pass protein with an N-terminus in the lumen and a C-terminus in the cytosol.

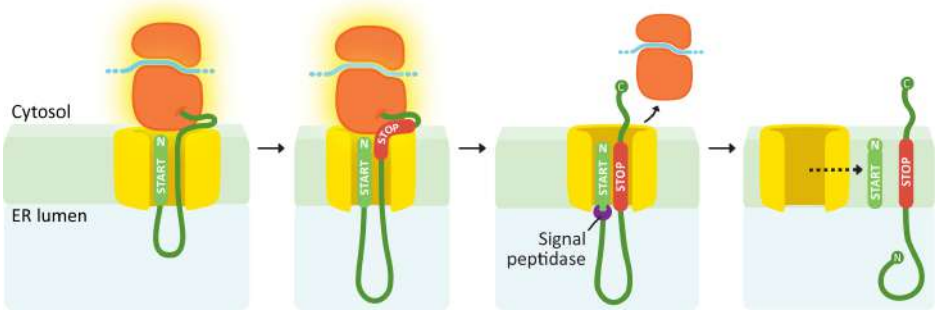


Figure 04-10: Insertion of a protein with an N-terminal START and a STOP insertion sequence. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Similar to the soluble protein example (Figure 04-08), this next example protein has an N-terminal START sequence (Figure 04-10). This targets it immediately to the ER membrane, and the growing polypeptide is threaded through the channel from the very beginning of translation. At some point during translation, it runs into another ER insertion sequence, which will act as a STOP transfer sequence. When the STOP transfer sequence is encountered, it causes the translocation channel to stop threading polypeptide through the translocation channel into the ER lumen. The ribosome undocks from the membrane but is still tethered via the translating protein. Since there are no other insertion sequences in this protein, the ribosome will complete translation undocked but tethered. Since the START sequence is N-terminal, it will be cleaved off, which results in the N-terminus of the protein being free in the ER lumen, while the C-terminus remains in the cytosol.

Example #3

An internal START followed by a STOP creates a double pass membrane protein with both an N-terminus and a C-terminus in the cytosol.

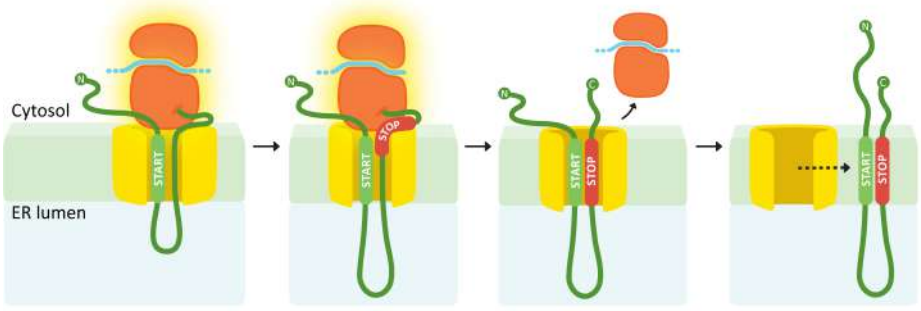


Figure 04-11: A protein with two transmembrane domains that has been inserted into the membrane. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

In this example (Figure 04-11), the first insertion sequence encountered is in the middle of the polypeptide, so it will be an internal START. Again, the internal START sequence initiates ribosome

docking, and the polypeptide begins to thread through the translocation channel. This continues until the second insertion sequence is encountered (a STOP transfer sequence). The STOP sequence ends insertion through the translocation channel, and the ribosome once again completes translation undocked but tethered to the ER. Once translation is complete, the START and STOP are released from the translocation channel and diffuse laterally into the membrane. In this case, both insertion sequences are retained and become transmembrane domains. Both the N- and C-termini of the resulting protein will be in the cytosol.

Example #4

A protein with many membrane-spanning regions.

Using the three previous scenarios as a foundation, you can create a protein with any number of transmembrane domains simply by adding more insertion sequences, which become alternating START and STOP transfer sequences. A START transfer sequence will bring the ribosome to the translocation channel to thread the growing polypeptide into the lumen, and the STOP transfer sequence will release the ribosome, thereby ending the continued use of the translocation channel. The internal START and STOP transfer sequences will each become transmembrane domains, whereas an N-terminal START (if present) will get cleaved. In this way, the protein is essentially stitched into the membrane, and the ribosome is bound and released a number of times. Our final example (Figure 04-12) has six transmembrane regions with the N- and C-termini in the cytosol.

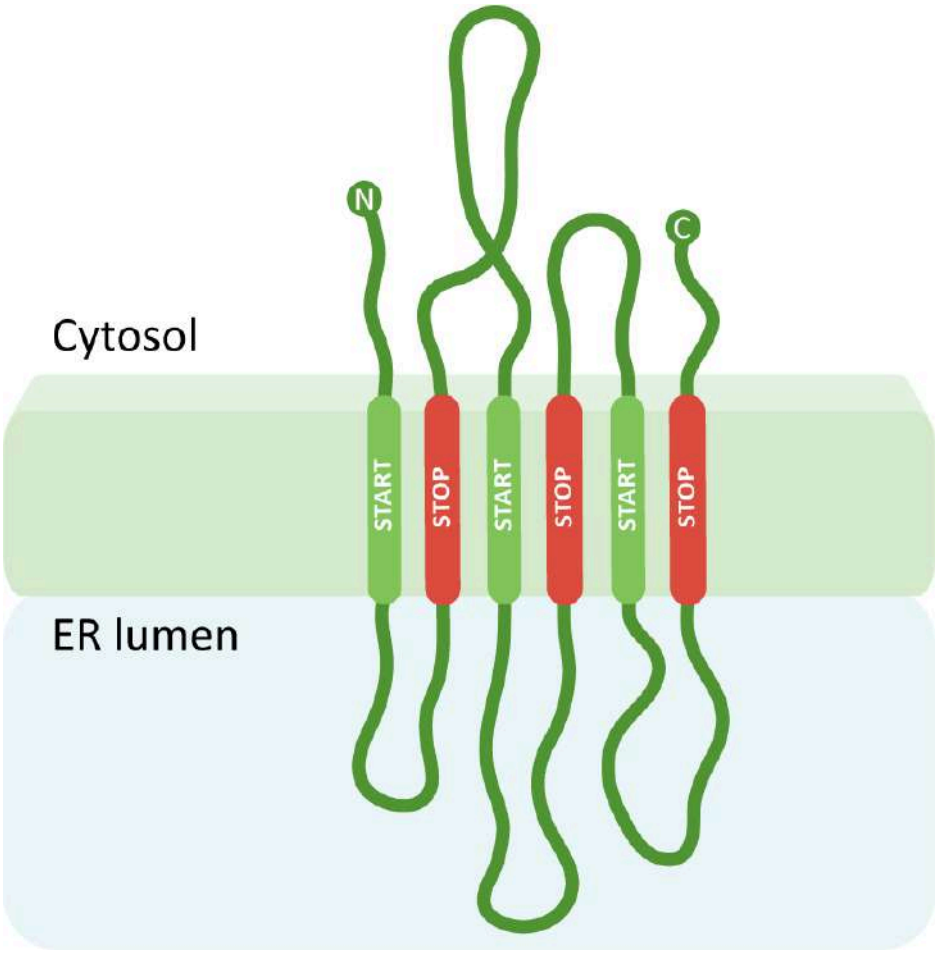


Figure 04-12: Multispanning membrane protein with START and STOP sequences noted. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

We hope that going through these examples has helped make sense of how protein ER insertion sequences work. To briefly summarize, we want you to take away the following the key points from this discussion of protein insertion into the ER:

- The ER insertion sequence is always 8–10 nonpolar amino acids. The difference between how each of the sequences is treated by the cell is due entirely to its *location* within the primary sequence of the growing polypeptide.
- There are two major categories of hydrophobic signals used in the insertion of membrane proteins. All of these are membrane-crossing domains:

    1. *START transfer sequences*. There are two kinds of start transfer sequences:
        - *N-terminal START transfer sequence.* Like its name says, this sequence is at the N-terminus of the protein. It remains in the membrane during translation and is cleaved off of the protein by the signal peptidase. This is also called the signal peptide or signal sequence.
        - *Internal START transfer sequence*. Similar to a signal sequence but located internally (i.e., not at the N-terminal end of the protein). It also binds to the SRP and initiates transfer into the ER. Unlike the signal sequence, it is not cleaved

after transfer of the protein.

2. *STOP transfer sequence*. A STOP transfer is never the only signal in a polypeptide chain. It follows either an N-terminal or an internal START transfer sequence. The STOP transfer signal is a membrane-crossing domain. It remains in the membrane. The sequence is not cleaved.

- The orientation of the protein in the membrane is completely dependent on whether there is an N-terminal START sequence or not. If the protein has an N-terminal START, then the N-terminus of the protein will be in the ER lumen. If the first ER insertion sequence encountered is internal, then the N-terminus will remain on the cytosolic side of the ER membrane.
- The number of transmembrane domains a given protein will have will be equivalent to the number of *internal* transfer sequences (i.e., all STARTS and STOPS that are anywhere other than at the N-terminus).

Video 04-04 is an animation that does an excellent job of showing the details of protein insertion into the ER. We find that students really benefit from seeing this concept in action.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-4

Video 04-04: Several scenarios of protein import into the ER. Basic principles and several example scenarios are shown.

ER Resident Proteins Require an Additional Signal to Stay in the ER

ER resident proteins are proteins that are retained in the ER, as this is where they function. The chaperone proteins that are sometimes required inside the ER to help with protein folding are a great example of an ER resident. These proteins are inserted into the ER as usual, using the ER insertion sequences in their primary sequence, but also carry an **ER retention signal** called the **KDEL** (or HDEL in some species) sequence. It is named the KDEL sequence because this is the one-letter code for each of the amino acids in the sequence. It is located in the primary sequence of the protein, always at the C-terminus. The KDEL sequence ensures that any ER resident proteins that might accidentally get packaged into vesicles and shipped to the Golgi will be captured and sent back to the ER.

Protein Insertion and Membrane Asymmetry

As you may recall from Chapter 2, membranes are *asymmetric*, meaning that the cytosolic side is different from the noncytosolic side. Part of this membrane asymmetry is provided by the location and orientation of the proteins in the membrane. The initial insertion of these proteins into the ER

membrane is a big part of how that asymmetry is established. Once that protein is inserted into the membrane, it can't be removed, so it needs to be done right.

It is also very important that proteins remain properly oriented in the membrane as they move through the organelles of the endomembrane system toward their final destination. The use of membrane-bound **vesicles** is ideal for this, as the method of budding and fusion of vesicles maintains the orientation of the protein in the membrane while also eliminating any possibility of release into the cytosol (Figure 04-13). In all cases, and at every step, the cytosolic side of the membrane faces the cytosol. This is a key concept in understanding the endomembrane system. Check out the protein orientation in Figure 04-13 and notice that the part of the protein pointing into the cytosol is always the same part even when it gets to the donor compartment.

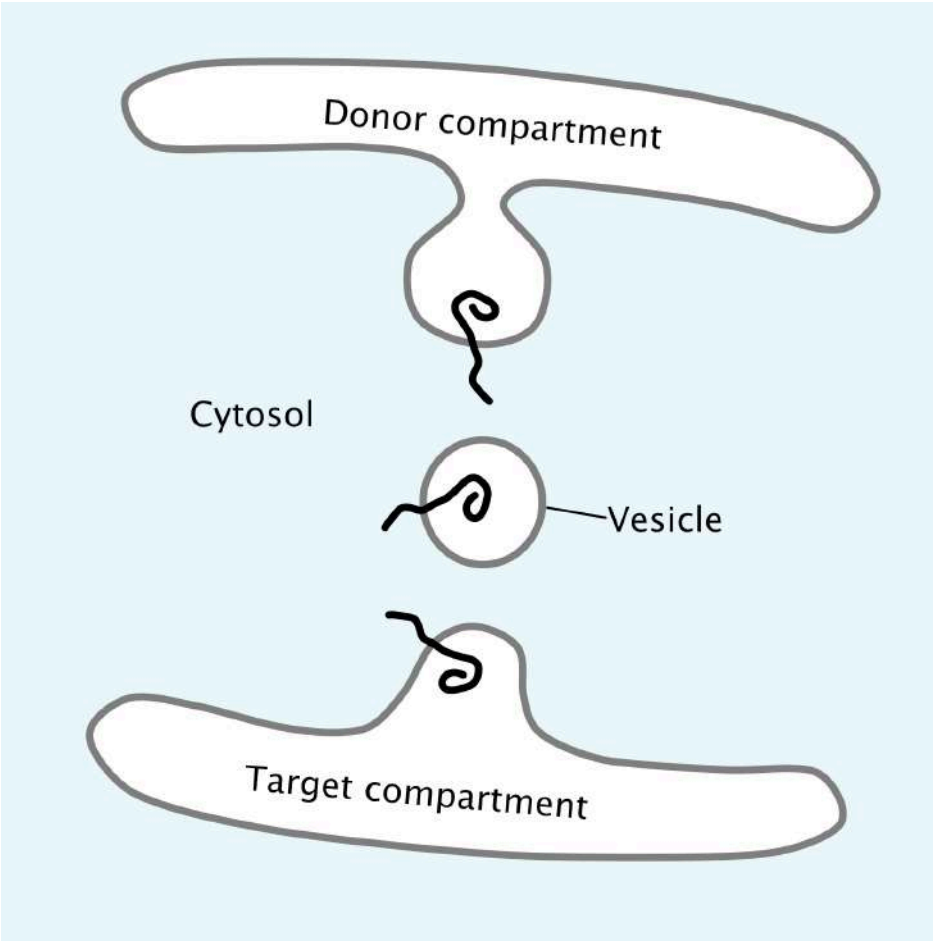


Figure 04-13: Vesicular traffic maintains the orientation of proteins in the membrane. This image was created by Dr. Robin Young and is shared under a CC BY-SA 4.0 license.

putting it into practice
Mapping Protein Domains

As we learned in Chapter 2, bioinformatics is a useful tool to identify known patterns within a protein sequence. We also explored how hydropathy plots allow us to identify potential transmembrane domains based on the amino acid sequence of the protein. If we add that to the information you've learned about so far in this chapter, we get a clearer picture of how we can learn about proteins using bioinformatics. You now know that when a hydropathy plot identifies hydrophobic sequences,

it's also identifying the ER insertion sequences that exist within the amino acid sequence. Hopefully, this shows you how the signatures of many functions of proteins exist in their amino acid sequences. Other sequences, such as the NLS or the KDEL, would also be easily recognized by an algorithm that was created to search for them. Import sequences are a great example, but they are by no means the only patterns or motifs that we can identify in a sequence that tell us something about protein function. Regions of a protein that have a specific function attributed to them are called **protein domains**.

With all of the different patterns and motifs within proteins, scientists need a system to keep track of which motifs exist in any given protein based on the bioinformatic analysis. Visual representations of the linear protein chain are used to map signature motifs within proteins. These are known as **domain maps**. Figure 04-14, below, shows how we could collect the hydropathy plot data from a given protein (Glycophorin A) and translate them into a domain map that shows how these regions function. Glycophorin A is a membrane protein found in red blood cells (see Figure 02-18, found in Topic 2.3).

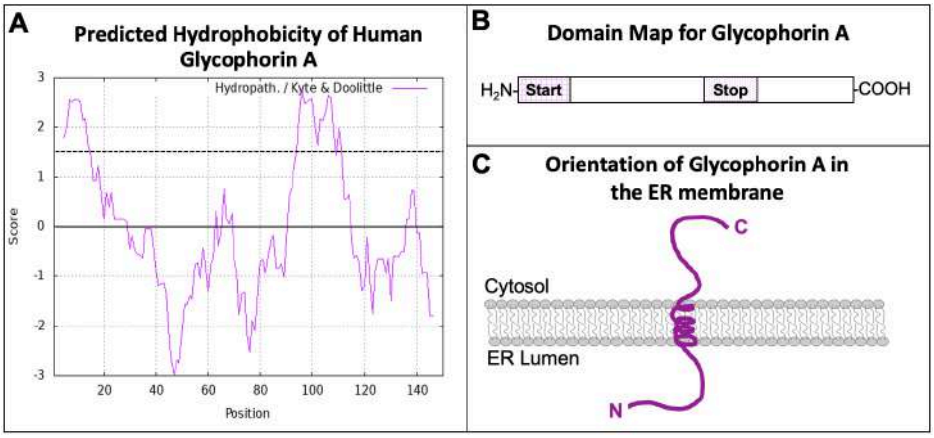


Figure 04-14: Domain mapping of Glycophorin A using information from its hydropathy plot to identify ER insertion sequences. (A) Kyle and Doolittle hydropathy plot generated using Expasy Protscale, based on Glycophorin A sequence data from UniProt. The y-axis indicates the hydrophobicity score (see Chapter 2 for tips on interpreting this graph). The x-axis indicates the amino acid number. (B) Domain map for Glycophorin A based on hydropathy plot data in A. (C). Diagram of final protein and how it will be inserted into the ER membrane during co-translational insertion. Since this protein has no other identified targeting sequences, its final destination will be the plasma membrane. This image was created by Dr. Robin Young and is shared under a CC BY-SA 4.0 license.

The hydropathy plot of Glycophorin A (Figure 04-14A) shows two regions within the amino acid sequence that have high hydrophobicity. These two regions are the two ER insertion sequences. Using the information in this hydropathy plot, we can draw a domain map for the protein (Figure 04-14B). The first hydrophobic region is directly at the N-terminus, which makes it an N-terminal START sequence. The second is in the interior of the protein, and it follows a START, so it is a STOP transfer sequence, and it will become a transmembrane domain. From the domain map that we have created, we can also draw the predicted protein and its orientation in the ER membrane (Figure 04-14C). Since the orientation of the protein is preserved as it moves through the endomembrane system (Figure 04-13), you should also be able to predict the orientation of the protein in its final destination, which would be the plasma membrane in the case of Glycophorin A. The portion of the protein that is

currently in the ER lumen will eventually be placed on the exterior side of the plasma membrane. The cytosolic side of the protein always remains on the cytosolic side.

For this example, we have focused specifically on mapping ER insertion sequences. However, it is important to remember that there are lots of things that can be mapped in this way. For example, you can map all of the different localization sequences (i.e., NLS, NES, ER insertion, KDEL, etc.) onto domain maps. If the protein was an enzyme with an active site that could be identified, then that could also be mapped. Many secreted proteins also have sites where they get glycosylated. Glycophorin A (the example from Figure 04-14) is glycosylated multiple times. If you were to follow the link to the bioinformatic protein data, the glycosylation sites are also identified and mapped. Maps of this kind are very important for helping us visualize and understand the structure and function of the proteins we study. You will likely have the opportunity to practice domain mapping as we continue to work through the various organelles and how proteins are targeted to them.

TOPIC 4.2: VESICLE TRANSPORT

Learning Goals

- List the four stages of vesicle transport, and list the protein machinery involved at each step.
- Explain how vesicle coats facilitate cargo loading and vesicle budding.
- Discuss the importance of maintaining specificity in both docking and fusion of vesicles at their target compartments and how Rabs, tethers, and SNAREs facilitate this process.

Introduction

As we mentioned in the previous topic, the ER is the point of entry for all newly synthesized proteins that either reside in the endomembrane system or are destined to be secreted. In addition, once proteins enter the endomembrane system, they do not return to the cytosol. This poses a unique challenge for the cell in that the organelles of the endomembrane system are not physically connected to each other, and yet they must transfer cargo from one organelle to the next without allowing cargo to "escape" back to the cytosol. As such, **vesicles** are used to shuttle cargo around the cell.

Vesicles carrying cargo bud from one of the membrane-bound organelles of the endomembrane system (ER, Golgi, endosome, lysosome, etc.). They then move through the cytosol to fuse with the next organelle in the endomembrane system or the plasma membrane (Figure 04-15). In this topic, we will examine the details of how vesicles form, travel, and then fuse with the membrane at their target destination.

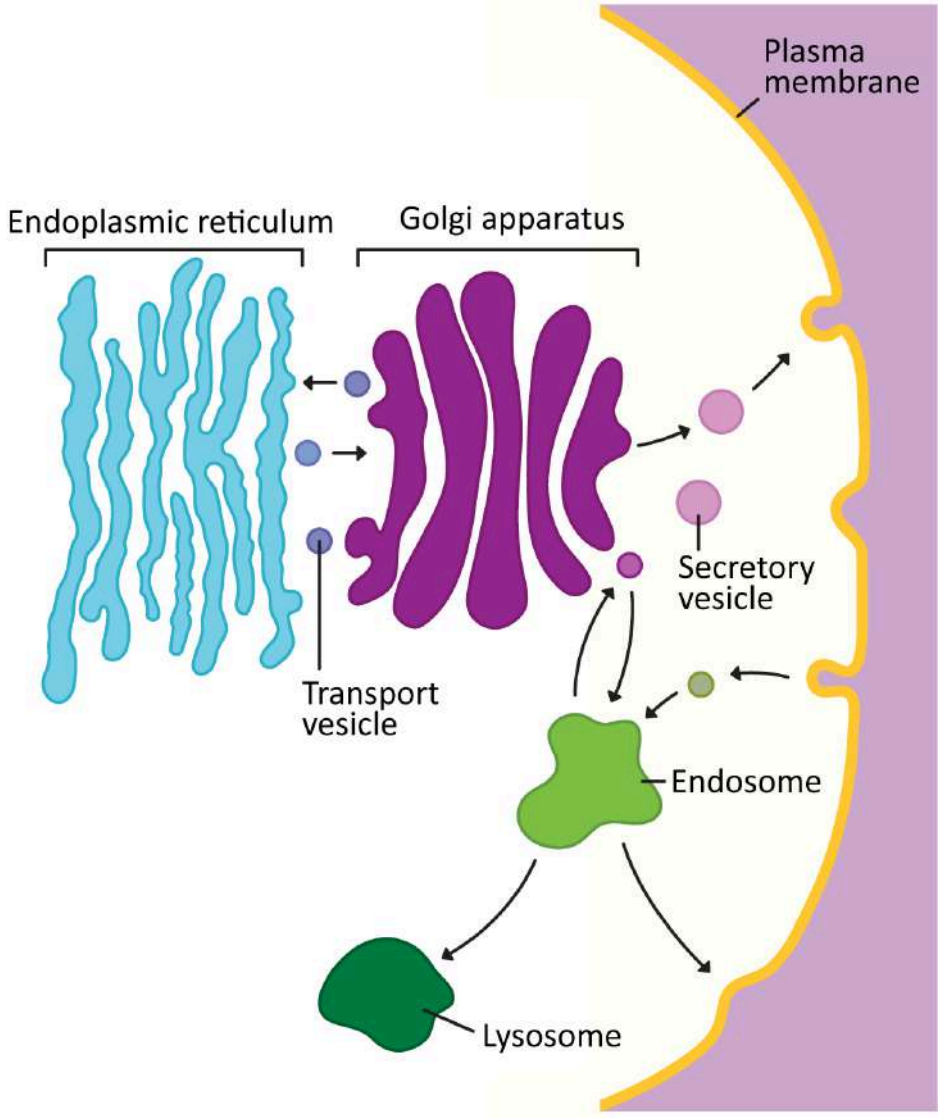


Figure 04-15: The movement of cargo through the endomembrane system. Each arrow indicates a vesicular transport path between those organelles. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

It's important here for us to take a moment and point out that vesicle trafficking is extremely complex! There's a lot we know about how vesicle trafficking works, but there's also a lot that we *don't* know. Vesicles are produced by *all* of the organelles of the endomembrane system. They all have specific cargoes and specific destinations. Some organelles (like the Golgi) are producing separate vesicles that will be sent to any number of different locations. The cargo is sometimes quite large, making it very difficult to put in a vesicle (collagen is a great example). And yet it, too, must get packaged properly and sent to its destination. For each new insight we gain about vesicle transport, we also find out that there are exceptions to what we thought were "the rules." This situation can become overwhelmingly complex if we also consider all of the specialized cell and tissue types, differences between organisms, and even differences between biological kingdoms and how their endomembrane systems have evolved to address their unique needs.

Despite the complexity, in this section, we provide you with a framework to understand the parts of vesicle traffic that apply in *most* situations. This will hopefully help you learn what to look for when encountering a new trafficking scenario.

All of these pathways through the endomembrane system rely heavily on vesicles to transport their cargo from one stage in the pathway to the next. Since the three major pathways (see Figure 04-01 to remind you of what they are) are also always running in tandem, we begin to see why traffic through the endomembrane system is so complex.

Vesicle Traffic

The Basics

The basic principle of vesicle trafficking is that vesicles must take the correct cargo from the donor compartment and deliver it to the correct target compartment. This means that

- the correct cargo must *get into* the correct vesicle, and
- the correct vesicle must *get sent to* the correct destination.

A good analogy to consider here is that of an Amazon order (or any other online/mail order). When an online order is placed, we expect a couple of things to happen. The first one is that the product that was ordered will be properly packaged with the rest of the things that were ordered and prepared for shipping. This implies that no one else's order will be included in this package. The second is that the package will then be delivered to the correct address, which is identified by the original order form. If you have ever had an online order that was shipped incorrectly, you have an understanding of the importance of doing this correctly the first time.

Vesicle trafficking works in a similar way. Instead of mailing addresses, targeting sequences are used. These sequences determine which type of vesicle the protein will enter so that it can be sent to the correct destination. The exception is when the cargo is destined to be secreted via the **secretory pathway**. Since secretion is considered the "default pathway" of the endomembrane system, proteins that are in the ER that don't have other targeting sequences will get scooped up into vesicles in a nonspecific way and sent along on their journey. We'll see more about the different pathways through the ER later in this chapter.

Vesicle traffic is controlled by the protein machinery that helps the vesicle form, travel, and fuse with the target membrane. It is this protein machinery that is the focus of this topic.

In all vesicle traffic, there are four main stages to the process (see Figure 04-16):

1. **vesicle budding**
2. **vesicle transport**
3. **vesicle docking**
4. **vesicle fusion**

It is sometimes said that there is a fifth stage, which is *resetting the system* so that the next round of trafficking can happen. Repeated rounds of vesicle traffic require that the trafficking machinery gets recycled back to its original location so that new vesicles can be formed and targeted. We will focus primarily on the first four stages of this process and only mention system reset briefly at the end of this topic.

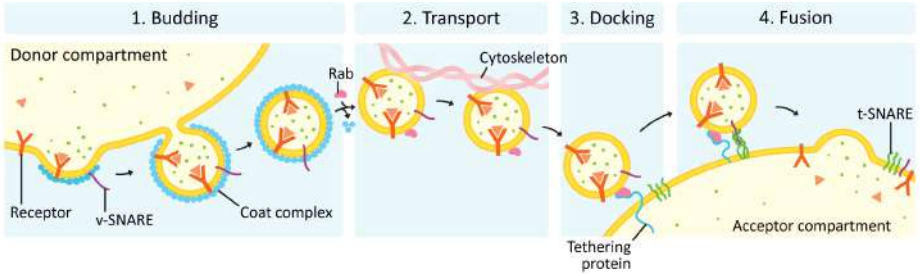


Figure 04-16: Stages of vesicle budding, transport, docking, and fusion. (1) Budding: Membranes deform toward the cytosol and pinch off, becoming free floating vesicles. (2) Transport: These vesicles are carried toward the target location, often using the cytoskeleton. (3) Docking: Vesicles are brought closer to the target membrane. (4) Fusion: The vesicle fuses with the target membrane, and any soluble components are released into the lumen of the acceptor (target) compartment. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Each of these stages *must* be carried out correctly. There is no margin for error in vesicle traffic due to the potentially destructive nature of the cargo that is being transported. Remember that proper function of the endomembrane system is absolutely essential to survival, which means that errors may very well kill the cell. In this material, we will focus on budding, docking, and fusion. Transport of vesicles is carried out using a combination of the cytoskeleton (which will be discussed in Chapter 6) and sometimes short-range diffusion (which we are not going to discuss further).

Step 1. Vesicle Formation and Cargo Selection (a.k.a. Budding)

It is important to understand that membranes do not curve on their own. Energy input is required to change the shape of a membrane. As such, it is thought that virtually all vesicles require some kind of protein machinery to help them form. However, the exact protein machinery used in the formation of many types of vesicles has yet to be discovered. We still have a great deal to learn about vesicle traffic despite over 100 years of research on the endomembrane system.

While there is a variety of protein machinery used to help promote the formation of vesicles, we can identify specific elements that are very common when looking at different ways that vesicles can form (Figure 04-17). All of the vesicles that we are going to study require the following:

- **Cargo receptor**: Proteins destined for particular target organelles must be collected together in one spot so that a vesicle can be formed around them. This is accomplished by binding of the proteins to receptors in the membrane that recognize them. This helps create the specificity required for protein transport. From our online order analogy, the cargo receptor is what picks up the merchandise and puts it in the correct box and also helps ensure that no other merchandise ends up in the box by accident.
- **Adaptor protein**: The adaptor protein acts as a bridge between the cargo receptor and the coat protein. Usually there are more adaptors than coat proteins, so this allows for another level of specificity. Different adaptors can bind to different receptors but still form a vesicle using the same coat proteins.
- **Coat protein**: The coat protein provides the structure to support the bending membrane while the vesicle is being formed. There are many components in a vesicle that help determine the direction of transport, but specific coats are generally associated with particular pathways and directions of transport. We'll see examples of this as we move through this topic.

- **GTPase**: A GTPase is a protein that binds GTP and hydrolyzes it. The energy released when the GTP is hydrolyzed can be used at any part of budding, though it appears to be primarily associated with the release of the vesicle coat after its job is done.

It's also important to know that the lipid composition at the site of vesicle budding is just as important as the protein machinery. Membrane lipids found in **ER exit sites**, and other regions from which vesicles bud, are important in helping to assemble protein machinery, allowing the membranes to curve, and even as identity markers. So even though we will not discuss the lipids much in this chapter, you should know that the lipids are equally important to proper formation and function of vesicles.

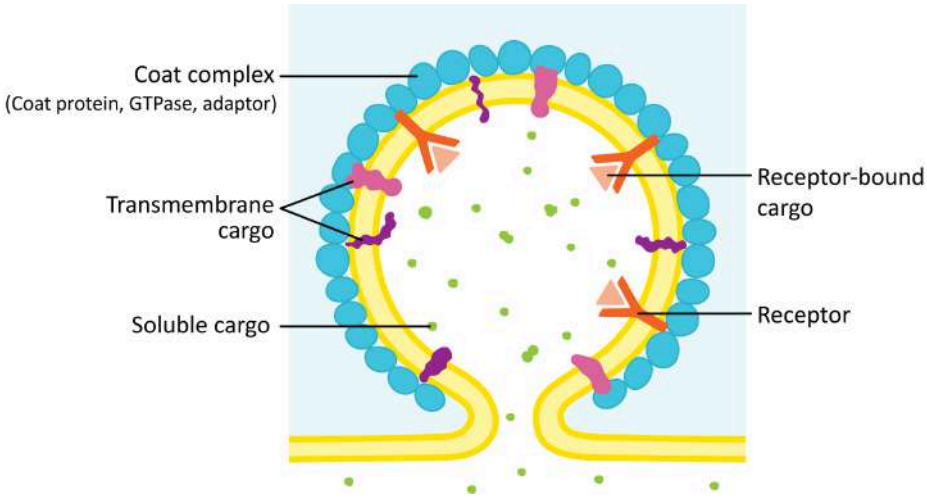


Figure 04-17: The protein machinery required for vesicle budding. A coat complex containing a coat protein, adaptor protein, and a small GTPase help stabilize the forming bud. Cargo is loaded into the forming vesicle. Some molecules are brought into the vesicle by bulk flow, while others are specifically recruited via interactions with a receptor or the coat adaptor. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Steps in the Formation of a Vesicle

1. **Cargo proteins** are held in place by transmembrane **cargo receptor proteins**.
2. An **adaptor protein** is a protein that binds to the cytosolic side of the receptor as well as a cargo receptor protein.
3. A small GTP-binding protein (called a **GTPase**) acts as a regulatory unit to determine exactly when the adaptor and coat proteins are allowed to bind to the receptor. It is in its active form when bound to GTP. Over time, the GTP gets hydrolyzed to GDP, which inactivates it, and the GTPase falls off the membrane.
4. Proteins involved in flipping lipids across the membrane (related to the flippases and scramblases of Chapter 2) initiate the curvature, which is the start of a vesicle bud.
5. **Coat proteins** assemble into a lattice or network on the cytosolic side of the forming bud to help stabilize and extend this curvature (budding) until a vesicle is formed.
6. To release the vesicle, the forming vesicle bud needs to be cut off from the donor membrane (a

process known as **scission**). This action may be part of the function of the coat protein, or it may be done by a separate protein.

Once the vesicle has budded off, the coat is usually removed, the vesicle can be transported to the target organelle, and the naked vesicle fuses with its target membrane.

An Introduction to the Vesicle Coat Proteins

All coat proteins have two primary functions:

1. To stabilize membrane curvature in the forming bud
2. To capture cargo molecules for transport

There are at least *six* types of vesicles where the vesicle-forming machinery is currently known and many more types of vesicles where we are still unsure how they are formed. Of the six known types, this textbook explores three of them in more detail, as they bear the most similarities to each other and are the most well characterized. More specifically, we will be looking at the three major classes of coat proteins and their role in vesicle formation. The three types of coat proteins we will study are COPI, COPII, and clathrin. Each of these types of coats tends to be used in different locations in the endomembrane system and/or for different purposes. They are also considered to be the "canonical coats," probably because they were discovered first but also because they appear to be the most widely used by the cell.

1. **Clathrin** coats were the first coat proteins to be discovered and as such are the ones that we know the most about. They are most commonly used between the **trans Golgi network (TGN)** and endosomes, lysosomes, and the plasma membrane. However, even in these sites, they aren't always used. Clathrin coats have been shown to be important in specific kinds of traffic:
    - carrying newly synthesized lysosomal enzymes from the TGN,
    - bringing in molecules from the extracellular space via receptor-mediated endocytosis (and possibly some other endocytosis as well),
    - recycling unnecessary membranes from secretory vesicles as they mature, and
    - sending proteins out of the cell using the secretory pathway in highly regulated circumstances (like in neurons, where vesicles full of neurotransmitters are sent down to the axon terminus).
2. **COPI coats** tend to be used for **retrograde traffic**, which means traffic that goes "backward" through the endomembrane system. Specifically, COPI coats are used for either Golgi-to-ER traffic or moving cargo backward through the cisternae of the Golgi. Some evidence suggests they can also be in other areas, but only in very specific situations. COP stands for "coatamer protein."
3. **COPII coats** are most commonly used for **anterograde traffic**, which means forward traffic. Specifically, they're used for ER-to-Golgi traffic. They also have the ability to make vesicles of different sizes and shapes, unlike some of the other coat proteins, so they are used for the

secretion of large molecules (like collagen, for example). As a result, COPII can also be found at the TGN, but they are not as common as clathrin. COPI and COPII coats were named in the order in which they were discovered.

In a nutshell, clathrin is primarily used in vesicles originating from the TGN and from the cell surface. COPI and COPII are mostly used as the coat proteins for vesicles originating from the ER and within the Golgi. Table 04-01 summarizes some of the structural details of these three coats as they relate to the coat components described in Figure 04-17.

Table 04-01: Comparison of the components of the three "canonical coats" covered in this textbook

COMPARING VESICLE COATINGS

Coat	Coat components	Adaptor	GTPase	Scission protein
Clathrin	Clathrin triskelion complex	Adaptins (AP1, AP2, or AP3) or GGA complexes	Arf (for GGAs only)	Dynamin
COPI	Sec13/31 complex	Sec23/24 complex	Sar1	Included in coat
COPII	7 subunits in complex	← Included in coat	Arf	Included in coat

Clathrin-Coated Vesicles: A Case Study

Since clathrin is the most extensively studied of the vesicle coats, we will use it as a case study to look at the finer details of how vesicle coats form. The clathrin cage is made from two different kinds of proteins known as the *heavy* and *light* chains. These chains come together in a specific arrangement to form a **triskelion**, which can be seen in Figure 04-18.

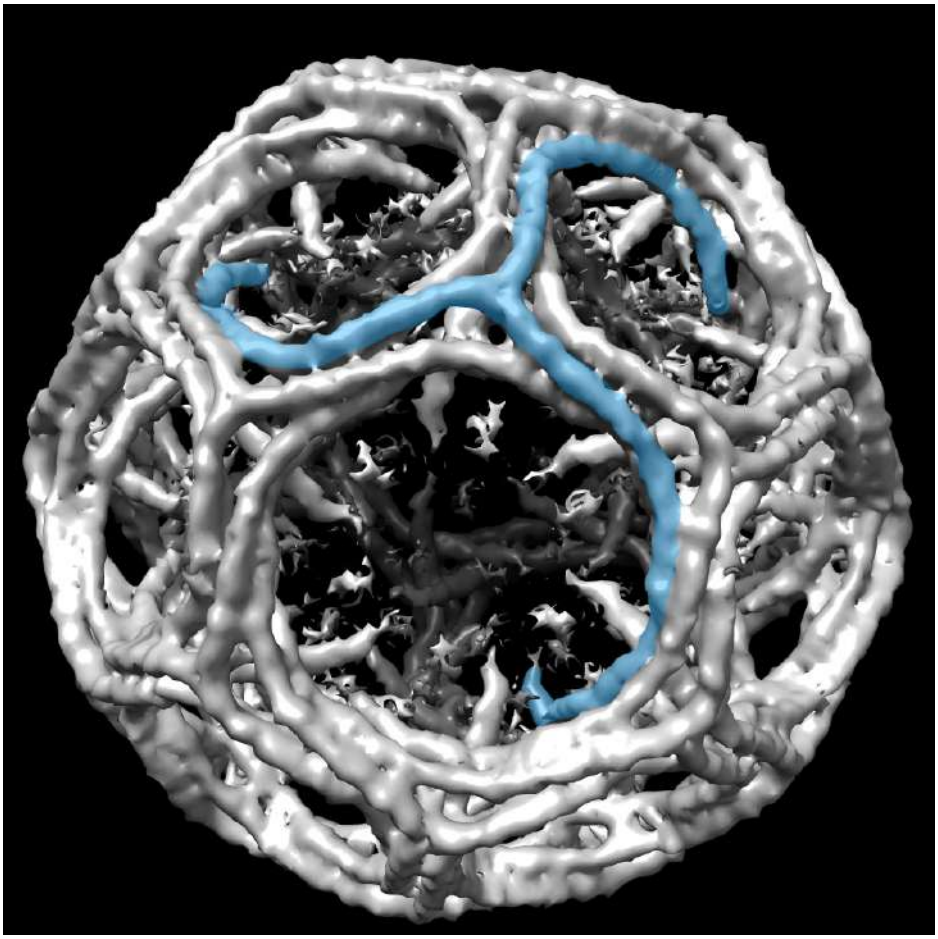


Figure 04-18: Cryo-electron microscopy of clathrin coat. This coat is made of 36 triskelions and is approximately 90 nm in diameter. A single triskelion is highlighted in blue. "Clathrin cage viewed by cryoelectron microscopy" by Mazuraan is shared under a CC BY-SA 4.0 International license.

The triskelions bind to the adaptor (usually adaptin for clathrin-coated vesicles) and assemble to form the coat. As they assemble, they pull the membrane with them to form a ball. While the lattice helps shape the vesicle, it cannot cut the vesicle off from the donor membrane. A separate scission protein called **dynamin** is used for this purpose. Dynamin has a spiral shape and wraps around the stalk formed by the budding vesicle. Dynamin binds GTP and hydrolyzes it, which allows the protein to constrict around the membrane stalk until the membrane splits and the vesicle is released.

Clathrin works primarily to mediate traffic between the TGN, the endosomes, lysosome, and the plasma membrane. It's important to note that not *all* traffic in this area uses clathrin, but most of the traffic that clathrin mediates is somewhere in this region. For example, clathrin is used in certain kinds of **endocytosis**, which means that the cargo is coming in from the cell exterior. It's also used to mediate the movement of newly synthesized lysosomal proteins as they move from the TGN to the endosome, which will eventually mature into the lysosome. A notable exception is that clathrin is not usually involved in protein secretion (i.e., the default pathway out of the cell).

COPI and COPII Coats

The process of vesicle budding is essentially the same in COP-coated vesicles as it was in the clathrin-coated vesicles we just saw. In fact, all three follow the same general trajectory that we saw at the start of this section. Video 04-05 (shown below) is an excellent molecular animation of COPII vesicle

formation. The different structures of the coats make for differences in the geometry of the cage that is formed and the size of the vesicle (Figure 04-19). Additionally, COPI and COPII coats both are able to promote scission themselves. As such, no additional scission protein (like dynamin) is required. Instead, a small GTPase is used as part of the coat and helps the coat release when its job is done. For COPI, the GTPase used is known as Arf, and for COPII, it's called Sar1.

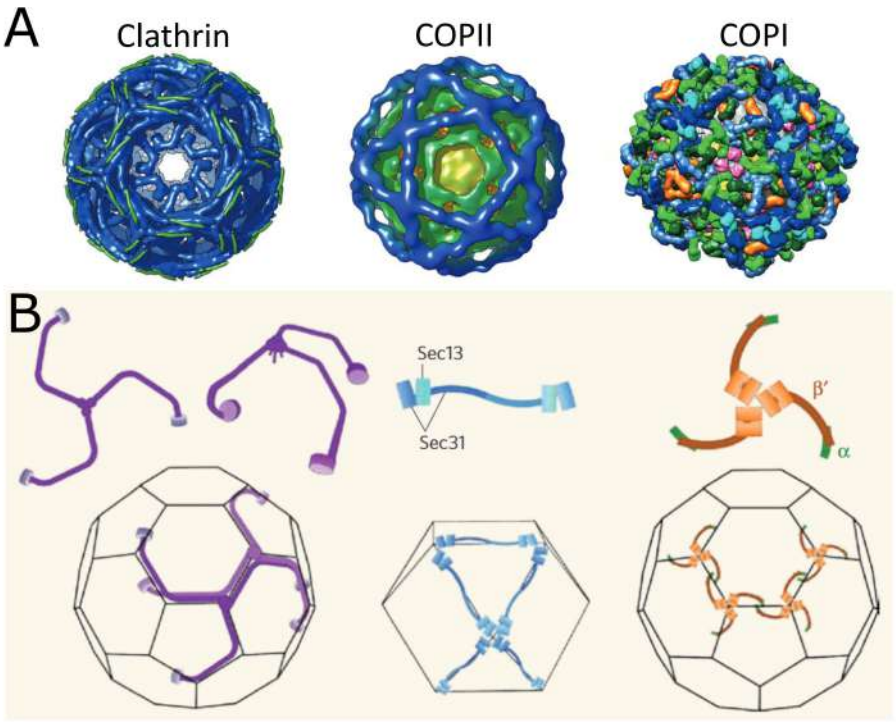


Figure 04-19: Molecular structure of the different coat proteins. (A) Molecular models from Noble and Stagg (2015). (B) Schematic coat proteins overlaying a sketched cage structure from Harrison and Kirchhausen (2010). Average diameters for all vesicle coats shown are roughly similar, ranging from 60 nm (COPI) and 90 nm (clathrin), with COPII somewhere in the middle. Panel (A) is from Noble, A. J., & Stagg, S. M. (2015). COPI gets a fancy new coat: An interconnected scaffolding of proteins bends the membrane to form vesicles. Science, Volume 349, Issue 6244, July 2015, Pages 142–143. http://www.jstor.org/stable/24748469. Reprinted with permission from AAAS. Please do not redistribute without permission of the rights holder. Panel (B) is included under fair use as described in the CBPFUOER. Please do not redistribute without permission of the rights holder.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-5

Video 04-05: COPII vesicle formation. A host of molecular machinery comes together to form the vesicle as well as recruit the cargo into the forming bud.

COPI and COPII mostly mediate traffic between the ER and Golgi and travel in opposite directions (Figure 04-20). This was mentioned briefly above, but this list gives you more detail:

- COPII vesicles move in the forward direction from the ER to the Golgi (i.e., anterograde traffic). As such, it usually is transporting newly synthesized proteins that need to leave the ER on their journey through the endomembrane system.
- COPI vesicles, on the other hand, mediate retrograde traffic back to the ER from the Golgi. This will include proteins that need to be returned to the ER and that are residents of that compartment. This is where the **ER retention signal** (also known as **KDEL**, first discussed in Topic 4.1) becomes important.
    - The KDEL is recognized by the **KDEL receptor**, a transmembrane protein that lives in the Golgi where vesicles from the ER fuse.
    - The KDEL receptor binds to the KDEL sequence, which is a part of the ER resident protein primary sequence.
    - When the KDEL receptor binds to its cargo, a conformation change happens that exposes a COPI binding site on the cytosolic side of the protein.
    - Once the COPI binding site is exposed, the COPI coat can assemble so that the KDEL receptor, with its cargo, can be packaged into vesicles and sent back to the ER.
    - Once its cargo is released, the empty KDEL receptor is returned to the Golgi via COPII-coated vesicles.

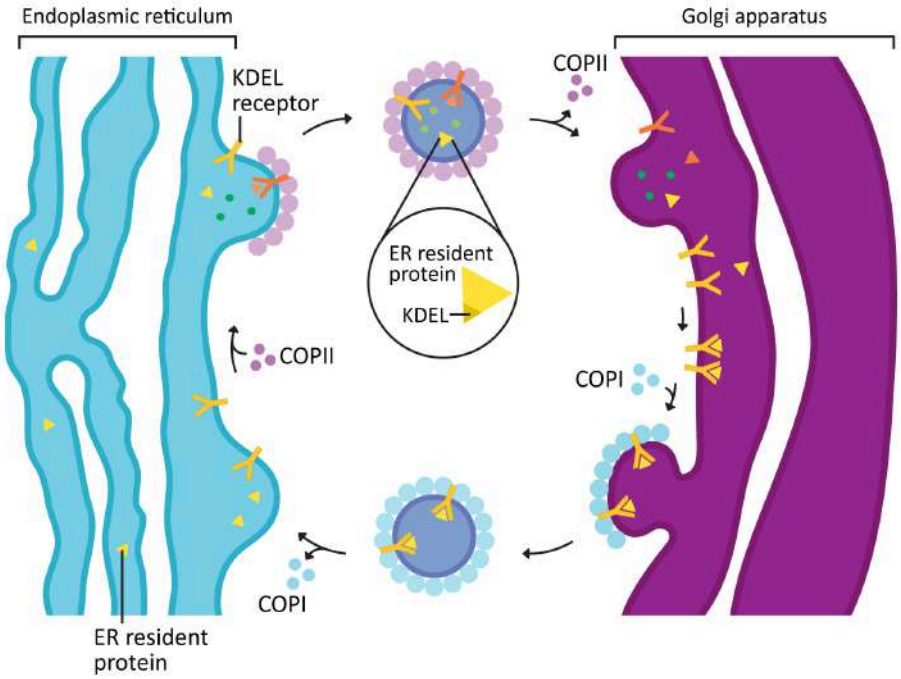


Figure 04-20: Traffic between the ER and Golgi is managed by COPI- and COPII-coated vesicles. Proteins bound for the Golgi are packaged into COPII-coated vesicles. On occasion, ER resident proteins also get trafficked to the Golgi, where they are recognized by the KDEL amino acid sequence and packaged into COPI vesicles that return to the ER. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Interestingly, COPI and COPII coats both have a unique additional feature, which we really only discovered in the last 10–20 years. The COPI and COPII cages are somewhat flexible in how they are arranged on the membrane. As such, they can produce larger "vesicles" that can accommodate cargo that would not otherwise fit. For example, collagen is one of the most abundant proteins in the extracellular matrix of animals, which we were introduced to briefly at the end of Chapter 2 (see Figure 02-20). Collagen is synthesized by the cell in the same way that other secreted proteins are: it is co-translationally inserted into the ER lumen, travels through the endomembrane system, and then is packaged into vesicles at the TGN to be sent to the plasma membrane for secretion. Collagen is a very large, fibrous protein, so it doesn't fit into traditional COPII vesicles. However, COPII is able to rearrange itself into a giant tubule that is large enough to accommodate collagen so that it can be sent out to the extracellular matrix.

Other "Coat" Proteins

As mentioned earlier, our focus in this textbook is on the more traditional vesicle coats that do the bulk of the work in the endomembrane system. However, there are a number of other vesicle "coats" that are involved in the formation of vesicles that are often ignored in introductory cell biology texts such as these. We will list them here, and what they do, so that you have a more complete perspective of vesicle traffic. Also keep in mind that this list might not be complete. Like so many aspects of cell biology, this is an area of active research, and there's always more to know!

- *Caveolin* helps create vesicles during endocytosis. Whereas clathrin-coated vesicles are usually used for **receptor-mediated endocytosis**, caveolin is used for a type of endocytosis known as **pinocytosis**. We'll learn more about endocytosis (but not caveolin) in the next topic in this

chapter.

- The *SNX/retromer complex* is primarily involved in retrograde traffic from the endosome to the TGN. In general, it forms large tubules that help in cargo selection. Then the vesicles are budded off the end of the tubule.
- *ESCRT* is pretty cool, as it is used to bud vesicles in a way that is opposite from most other coats. Instead of budding vesicles into the cytosol, they bud them *away* from the cytosolic compartment. This is usually used to push vesicles into the endosome so that membrane proteins can be fully degraded by the lysosome. Unfortunately, ESCRT is a common target for hijacking by membrane-bound viruses to help them bud out of cells so that they can go move to a new cell and restart the infection cycle. Both HIV-1 and the Ebola virus are known to do this.

While their mechanism of function is different from the more canonical coats, a few truths still hold. These protein complexes help bend the membrane and pinch off the new vesicle. The retromer and ESCRT also have mechanisms to choose cargo, whereas caveolin is considered nonspecific.

Step 2. Transport of Vesicles from the Donor Compartment to the Target Compartment

There are two possible modes of transport for a vesicle:

1. Over short distances, vesicles can move by diffusion. This is thought to be quite common in mitosis of plant cells, especially when a new cell wall needs to be rapidly secreted between the two new cells. It is also known as *bulk flow.*
2. Over longer distances, vesicles move along cytoskeletal tracks (usually microtubules but also actin) and are moved by motor proteins (kinesins or dynein for microtubules, myosin for actin).

As mentioned before, we will look at how the cytoskeleton works in Chapter 6, so we direct you there to understand how this process might work.

Steps 3 and 4. Targeting of Vesicles (a.k.a. Docking and Fusion)

Like vesicle budding, docking and fusion must be specific. Errors cannot happen or the cell might die. As such, there is additional machinery involved in making sure that docking and fusion happen accurately and efficiently at each of the different target membranes.

3. Vesicle Docking

Vesicle docking is a way of bringing a vesicle in close to the target membrane to place it in the perfect position for the final fusion step (see Figure 04-21). There are two main categories of proteins involved in this process: Rabs and tethers.

- **Rabs** are small GTPases that sit on the vesicle surface when activated and help identify the vesicle as one that is headed to a particular target membrane.
    - Rabs are said to be used for "membrane identity." This means that each organelle in the endomembrane system has its own sets of Rabs. If a vesicle buds from the ER, for example, any ER Rabs will be lost, and a vesicle Rab will dock. Once the vesicle fuses

with the target compartment, the vesicle Rab will fall off, and a new Rab will take its place.

    - Rabs are part of a larger "superfamily" of small GTPases whose members you will meet over and over again in cell biology. Ran, which you may remember from nuclear import in Chapter 3, is also part of this family, as well as Arf1 and Sar1, which help with vesicle coat assembly and disassembly in COPI and COPII, respectively. Even tubulin (used to make microtubules) is a distant relative.

- **Tethers** are a much more diverse group of proteins, usually with very little genetic similarity between them. However, their role is the same in all cases: to bind to the vesicle Rab and help capture the vesicle from the cytosol and pull it in closer to the membrane of the target compartment.

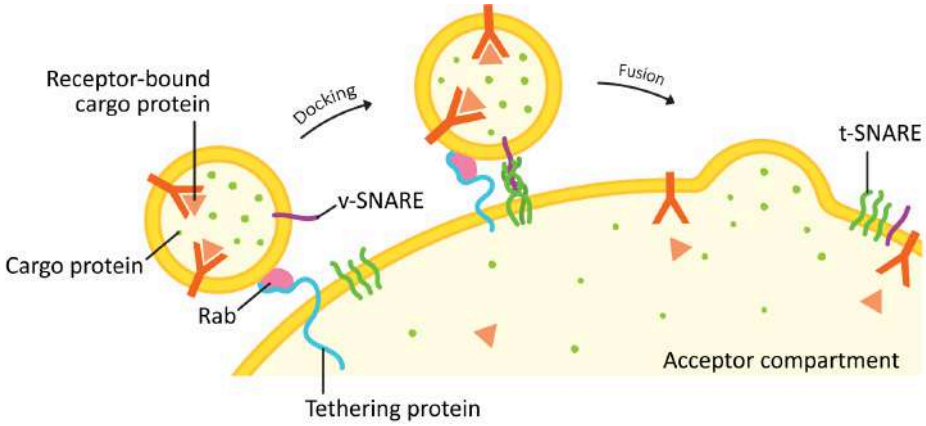


Figure 04-21: Docking and fusion. Tethers bind to Rabs to help bring the vesicle close to the target membrane (docking). When close enough, SNARE proteins interact to fuse the membranes together. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

4. Fusion with the Target Compartment

Once the Rabs and tethers have done their work, the final step of this process is mediated by a family of proteins known as **SNAREs**. The word *SNARE* stands for *SN*AP *re*ceptor (SNAP itself is an acronym that stands for synaptosomal-associated protein). For this reason, we always write the name of the protein in all capital letters as SNARE.

SNAREs are generally categorized into two major groups: **vesicle-SNAREs (v-SNAREs)** and **target-SNAREs (t-SNAREs)**. This grouping is based on their location in the cell. The v-SNARE is embedded in the vesicle membrane, and the t-SNARE is embedded in the target membrane.

Structurally, SNAREs fall into two major categories: Q- and R-SNAREs. In order for membrane fusion to occur via SNAREs, one R-SNARE and 3 Q-SNAREs must be present. Some act as v-SNAREs, whereas others will be t-SNAREs.

Once the vesicle has been tethered, the four SNARE coils interact and "zipper" together (Figures 04-21 and 04-22). This pulls the vesicle in tightly enough that all of the water molecules get pushed out of the way, and the two membranes can interact directly. Once that happens, the membrane lipids can intermingle, and the membrane will fuse. Both v-SNAREs and t-SNAREs must be present for this to occur. This binding is very specific…not any old v- or t-SNARE will do.

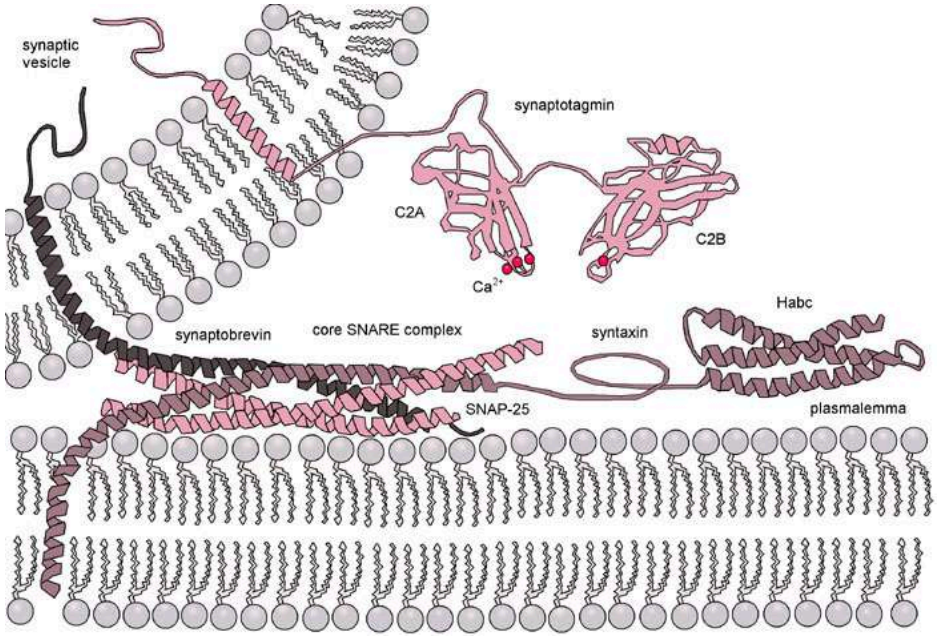


Figure 04-22: The mechanism of neurotransmitter release at the synapse. The four SNARE coils can be observed, zippered together as part of the core SNARE complex. Three proteins are a part of the SNARE complex: synaptobrevin (a v-SNARE), syntaxin, and SNAP-25 (both of which are t-SNAREs). Note that SNAP-25 has two SNARE coils. The rest of the protein machinery here is specific to neurotransmitter release and not necessarily present at other vesicle fusion sites. "Molecular machinery driving vesicle fusion in neuromediator release" by Danko Dimchev Georgiev, M.D. is shared under an unported CC-BY-SA 3.0 license.

A further illustration of vesicle fusion can be seen in the Video 04-06. At the end of the video, a protein comes in and uncoils the four SNAREs from each other. This is known as *resetting the system*, which was mentioned way back at the start of this topic. Separating the SNAREs after fusion is vital so that they are available for the next vesicle that docks.

In addition to uncoiling the SNAREs, resetting the system involves transporting the v-SNAREs back to their original target compartment so that they can be used in another round of vesicle fusion. Transporting SNAREs back is not trivial, as you need to make sure that they don't accidentally get used in transport. Usually, the SNAREs that are traveling as cargo get covered up by a regulatory protein (called n-Sec1) so that they can't get in the way. But that story is mostly beyond the scope of this textbook.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-6

Video 04-06: From neuronal communication to SNARE recycle. Vesicle fusion requires the interaction of SNARE proteins. These coil together once they interact to bring the vesicle close to the target membrane. SNARE complexes are specific between donor

compartment and target membrane. At the end, the SNAREs are uncoiled, and the v-SNARE is sent back to its original compartment so that it can participate in further rounds of fusion.

TOPIC 4.3: THE GOLGI APPARATUS THE HEART OF THE ENDOMEMBRANE SYSTEM

Learning Goals

- List the major function(s) of the Golgi.
- Explain how the structural compartmentalization of the Golgi creates different environments and how this contributes to Golgi function.
- Compare and contrast the Golgi structure and function in animals, plants, and fungi.
- Describe the sequence of events occurring during protein glycosylation in the ER and Golgi.

Introduction

As material travels through the endomembrane system, it will move from one compartment to the next, using vesicles to traffic between compartments. There is an order to the flow of traffic within this system. Newly synthesized proteins always start their journey at the ER, which was our focus at the start of this chapter. Then they are packaged into vesicles and are sent to the Golgi apparatus. From there, they may take different paths, depending on their role and destination. Some proteins will be sent to the endosome and then the lysosome, while others will head directly for the cell exterior. Even cargo that has entered the cell via the endocytic pathway (Figure 04-01) sometimes comes as far into the cell as the Golgi before being redirected to another destination.

Just like the ER, and the other organelles of the cell, the Golgi apparatus has multiple roles. It is often considered to be the heart of the endomembrane system, as it receives cargo from all directions (ER, plasma membrane, endosomes), modifies it, and then repackages it into new vesicles to be sent to new destinations. It is also the site of all of the major cellular work involving polysaccharides. While some of the work starts or ends elsewhere in the cell, the Golgi is the site of virtually all of the known glycosyltransferases (i.e., enzymes capable of covalently attaching sugars to each other). Not only is this important for the production of glycoproteins (i.e., proteins with sugars attached), but it is vital for the production of the cell walls in plants, algae, and fungi. In this topic, we will specifically focus on the structure and function of the Golgi.

Structure of the Golgi

The Golgi was named after an Italian researcher by the name of Camillo Golgi in the late 1800s. He discovered the Golgi during his exploration of the central nervous system by microscopy. (It is for this reason that we always write the name of the organelle with a capital *G*.)

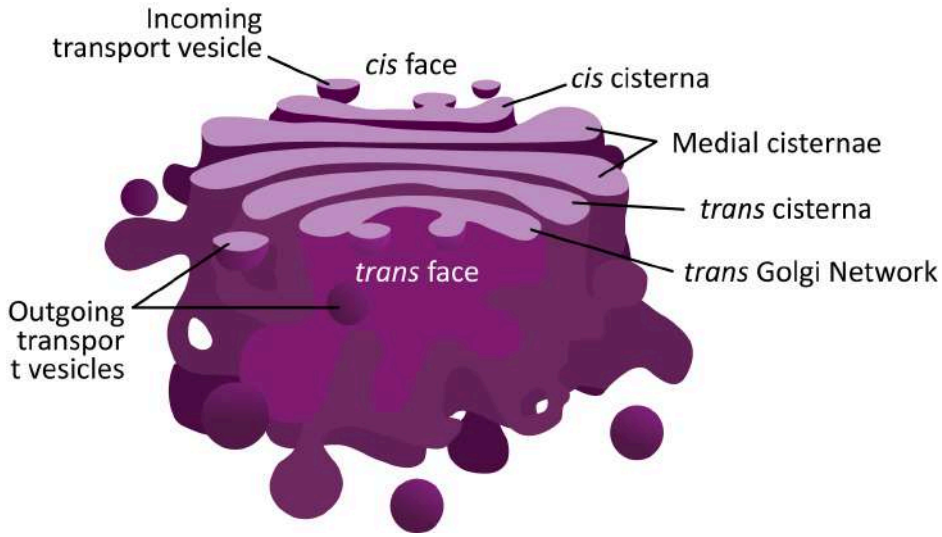


Figure 04-23: The structure of a typical Golgi stack. Vesicles from the ER fuse to form the cis cisterna. This side of the Golgi is labeled the cis face. On the opposing side, vesicles break off from the trans Golgi network to other cellular destination. This is called the trans face of the Golgi. The direction of movement within the Golgi is from cis to trans. Heather Ng-Cornish, modified "Diagram of the Golgi apparatus" by Kelvinsong, shared under a CC BY 4.0 license.

Figure 04-23 showcases the common structural features of the Golgi and highlights many of the new terms you will need to know to discuss this organelle. The following are some important points to note:

- Each of the "pancakes" of the Golgi is called a **cisterna** (plural *cisternae*).
- The interior space of each cisterna is known as the **lumen**. You have seen this term before, as it is also used to identify the space inside the ER, called the ER lumen.
- The Golgi is considered to be a "polar organelle" in that cargo enters at one end (the *cis* face, for newly synthesized proteins) and exits at the other end (the *trans* face). As such, we identify each of the cisternae by their location within the stack:
    - The *cis* cisternae are the ones nearest the *cis* face, which receive vesicles from the ER.
    - The medial cisternae are in the middle of the stack. There can be more than one.
    - The *trans* cisterna is closer to the *trans* face of the Golgi stack.
    - The *trans* Golgi network (TGN) is the *trans* most cisterna of the sac. Often it is more convoluted, as vesicles will be budding and fusing at this site. We'll hear more about the TGN later.

The structure of the Golgi apparatus is both surprisingly variable between different kingdoms/species and yet also very similar. For example, in all eukaryotes, the Golgi is made up of a series of flattened sacs that work together like an assembly line, moving cargo through them in order. However, the location and arrangement of those sacs can differ.

- In animals, the cisternae of the Golgi are large and are located centrally, near the nucleus (thus we say they have a **perinuclear** position). Usually, there is only one, or maybe two, Golgi in an animal cell.

- In plants, on the other hand, Golgi stacks are very small and tumble through the cytoplasm, using a combination of cytoplasmic currents and the actin cytoskeleton to help them move through the cell. There can be hundreds of these tiny Golgi stacks in a plant cell.
- In yeast and other fungi, The Golgi arrangement is quite variable. In some fungi, it is more "animallike," while in others, it is more "plantlike." In still others, it is different yet again. In the yeast *Saccharomyces cerevisiae*, the individual cisternae of the Golgi move around the cell independently of each other.

In this chapter, we will primarily focus on the features of the Golgi that are consistent across kingdoms, with only a few highlights of the unique characteristics specific to particular cell lineages.

While the cisternae do look like flattened sacs in cross section (Figure 04-23), there is more complexity to the structure. Vesicles are regularly budding and fusing with the margins of the cisternae as material is moved forward or backward within the Golgi stack. As such, if you were to look at a single Golgi cisterna from the top, it would look a bit more like a "paint splat" than a nice round pancake. This can be observed in the electron micrograph below (Figure 04-24).

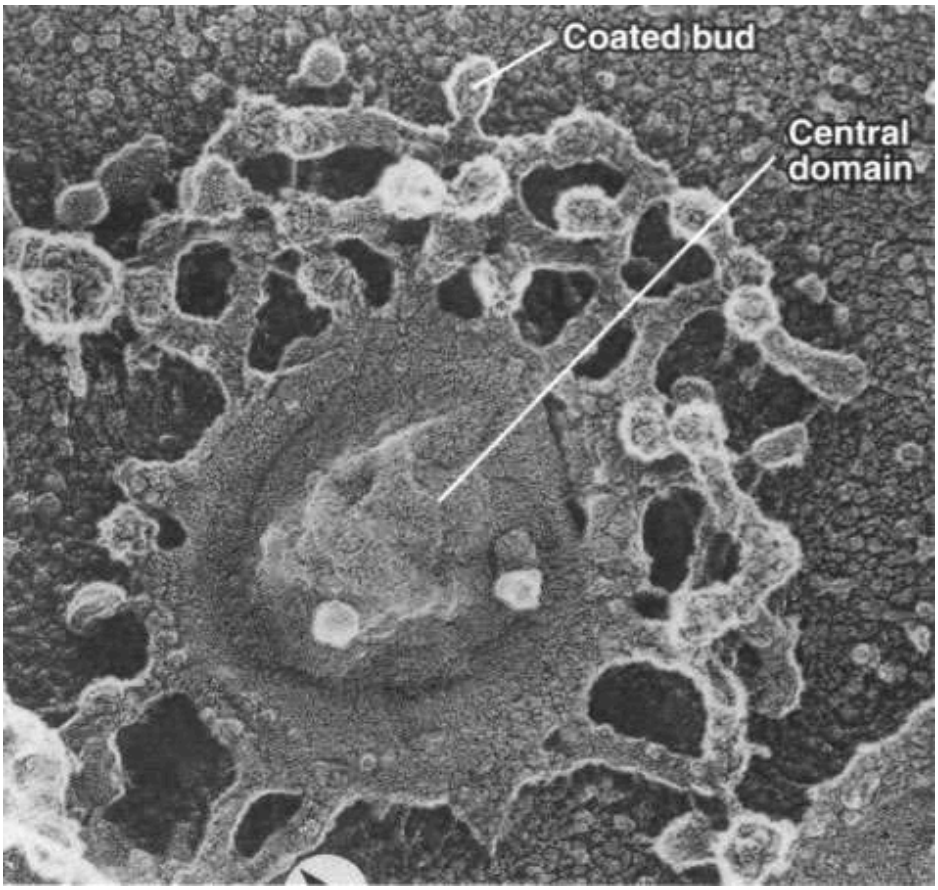


Figure 04-24: Labeled electron micrograph of a Golgi showing the shape of a single cisterna from the top. Reprinted from Trends in Cell Biology, Volume 5, Issue 8, Peggy J. Weidman, Anterograde transport through the Golgi complex: do Golgi tubules hold the key?, 302-305, copyright 1995, with permission from Elsevier. Please do not redistribute without permission of the rights holder.

Localization of Resident Proteins Can Be Quite Specific

Within the Golgi there is a net flow of newly synthesized proteins (i.e., the cargo) from *cis* to *trans*. In addition, each part of the organelle has specific proteins that are residents in that region. For example, there are specific proteins that are located in the *cis* cisterna of the Golgi that are required to remain there in order to carry out their function. The same is true for the enzymes that are found in the medial cisternae, the *trans* cisterna, and also for those that locate to the TGN. This arrangement can be visualized exceptionally well using immunolabeling and a technique known as electron tomography (Figure 04-25). Tomography is a specialized technique that allows us to create 3D models from the sample sections we normally use in an electron microscope. In each of the panels of Figure 04-25, we see an example of a 3D reconstruction of a Golgi stack with different resident enzymes that have been labeled using gold particles attached to an antibody. The colored dots (yellow in A and C; red in B) represent that gold particle label.

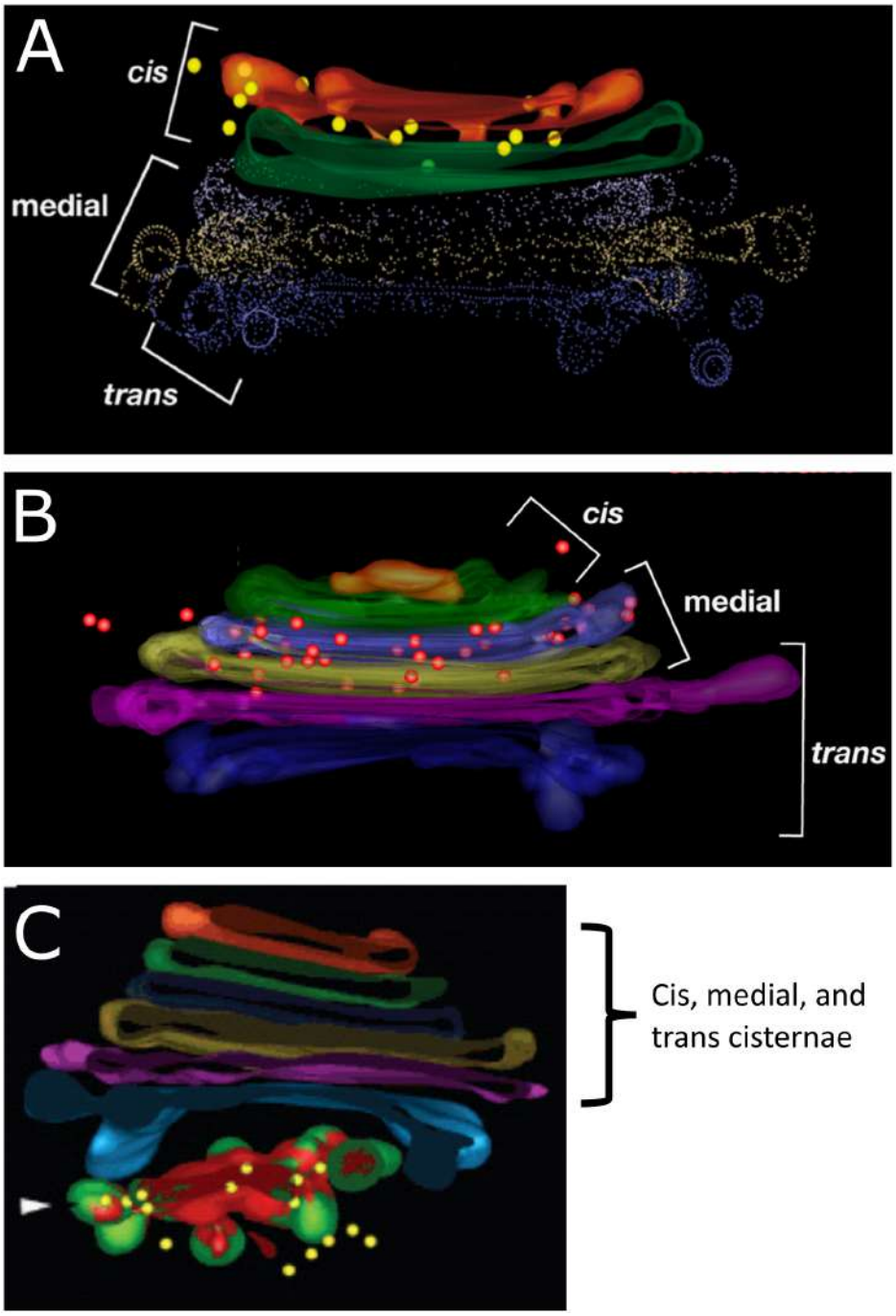


Figure 04-25: Golgi resident proteins localize to specific cisternae. A 3D model showing the location of all of the gold particles as yellow (panel A and C) or red (panel B) dots. These indicate the location of proteins: (A) PDI7, a protein with an unknown function in the cis Golgi cisternae; (B) ManI, which aids in specific glycosylation events in the medial Golgi; and (C) a Rab protein needed for vesicle trafficking called RabA4b in the TGN. This figure includes modified versions of original data found in (A) Yuen et al. (2017) and (C) Kang et al. (2011). Panel (A) is included under fair use as described in the CBPFUOER. Please do not redistribute without permission of the rights holder. Panel (B) is used with permission of Oxford University Press, from Staehlin & Kang (2008) Plant Physiology, 147(4), 1454-1468, https://doi.org/10.1104/pp.108.120618; permission conveyed through Copyright Clearance Center, Inc. Please do not redistribute without permission of the rights holder. Panel (C) is used with permission of Wiley and Sons, from Kang et al. (2010) Traffic, 12(3), 313-329, https://doi.org/10.1111/j.1600-0854.2010.01146.x; permission conveyed through Copyright Clearance Center, Inc. Please do not redistribute without permission of the

rights holder. This figure was compiled by Dr. Lauren Dalton.

Just like we've seen with other organelles, the ability of the proteins of the Golgi to localize so specifically can be traced back to the order of amino acids found within their primary sequences. There are specific sequences that are used to allow proteins to become part of a COPI- or COPII-coated vesicle and other sequences that allow proteins to be retained as a resident in the *cis*, medial, or *trans* Golgi (instead of moving through the cisternae and out the other side). As we learn more about the Golgi, we find that the specific localization of individual resident proteins might be even more precise, not only living within a single cisterna but also located in a specific region of that cisterna (inner versus outer ring). While you are not expected to know the details about the specific composition of these targeting sequences, it is important that you recognize that the ability of these proteins to localize so precisely is based on their amino acid sequence and can be traced right back to the DNA code they were made from.

Golgi Dynamics: How to Get from One Cisterna to the Next

One of the biggest challenges that the Golgi faces is that some proteins are "transient" and must move through the Golgi (i.e., from *cis* to *trans*), whereas other proteins are "resident" and need to stay where they are inside a given cisterna. As If that weren't challenging enough, vesicles constantly fuse and bud from the different cisternae, adding and removing proteins and lipids from the membrane. So how does the Golgi maintain its resident proteins in the proper location while allowing other cargo to be pushed along through the system? Scientists have been trying to solve this dilemma for almost as long as the detailed structure of the Golgi has been known.

Over the years, a number of ideas have been put forth to explain this. As we learn more, different ideas have gained and lost favor. Currently, the most widely discussed models for transport through the Golgi are called the **cisternal maturation model** and the **vesicular transport model** (sometimes called the *stable cisternae model*). See Video 04-07 for an excellent explanation of the two.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-7

Video 04-07: Two models describe how the Golgi resident proteins stay in the correct cisternal compartment while other proteins transit through the Golgi. Note that the video discusses the ERGIC (endoplasmic reticulum–Golgi intermediate compartment), which was briefly mentioned above but is not otherwise discussed in this text.

While the two models are present as separate possibilities, the current thinking is that the method by which proteins are moved and processed in the Golgi is probably more complex than any one model predicts (not really a surprise). The most widely accepted model right now is the cisternal maturation model (Figure 04-26), even though there is clear evidence that at least some things move via vesicular transport. The cisternal maturation model assumes that new *cis* cisternae are continually formed by

the fusion of vesicles that are flowing to the *cis* Golgi network from the ER. The previous *cis* most cisterna is now a medial cisterna, as it has been pushed further along in the stack. In this model the *trans* cisterna is the oldest and it started life as a *cis* cisterna. The *trans* cisterna breaks up into tubules and vesicles to form the TGN.

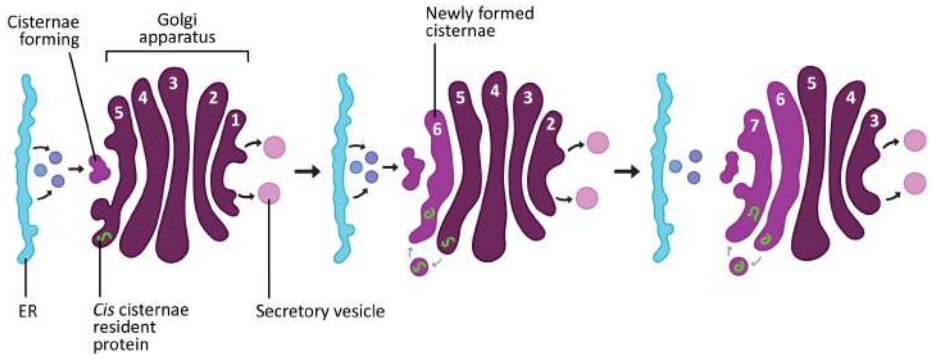


Figure 04-26: Cisternal maturation model. Vesicles coming from the ER merge to form the cis cisterna of the Golgi (lighter colored in image; labeled 6). As time goes on a new cis cisterna will be formed (7) causing 6 to become a medial cisterna. Resident proteins that work in the cis cisterna are transported back to the new cis cisterna (7) via vesicles. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

According to the cisternal maturation model, the proteins that need to pass through the Golgi (like newly synthesized proteins) stay in place inside the cisternae and simply enjoy the ride as they move from *cis* to *trans*. On the other hand, since the different cisternae also have different functions, "becoming" a *medial* cisterna will require that the resident proteins that work in the *cis* cisternae be removed from the new medial cisternae and sent backward to the newly formed *cis* cisternae. This retrograde trafficking is thought to be achieved via COPI vesicles.

Protein Processing in the Golgi: Glycosylation

One of the most common posttranslational modifications of proteins is the addition of polysaccharides. This process is called glycosylation (Figure 04-27). The enzymes that carry out these reactions are located in the lumen of the ER and the Golgi apparatus and not in the cytosol. This means that the proteins being glycosylated are residents of one of the organelles of the endomembrane system, are bound for secretion, or are membrane proteins.

There are two major types of glycosylation: **N-linked glycosylation** and **O-linked glycosylation**. Most introductory cell biology textbooks focus primarily on N-linked glycosylation, as it is believed to be the more common of the two. The biggest difference between the two types of glycosylation is how and where the sugar is attached to the protein. In N-linked glycosylation, the sugars are added to an asparagine (which has the 1-letter code N) within a specific amino acid sequence, whereas in O-linked glycosylation, it is a serine or threonine that gets used to covalently link the sugars (specifically, an oxygen in the R group gets used, hence the name *O-linked*). Otherwise, the process is somewhat similar in each case.

Glycosylation begins in the ER. In the initial step of the process, a prefabricated oligosaccharide "tree" consisting of N-acetyl-glucosamine, several mannose sugars, and a few glucose residues is transferred from a lipid known as dolichol (Figure 04-27, Step 1). Like all other lipids, dolichol is assembled on the cytosolic side of the ER membrane and then flipped to the lumen (cisternal) side

of the membrane by a **flippase** (see Chapter 2, Topic 2.2). After the tree is transferred to the protein, the dolichol is free to be reused by adding a new oligosaccharide tree to it for another round of glycosylation.

After the "tree" is added to the protein, the sugars can be trimmed in the ER before being packaged into vesicles and sent to the Golgi apparatus. It is inside the Golgi where most of the glycosylation work happens. Each **glycosyltransferase** mediates the formation of a single type of bond between specific sugars (i.e., glucose, galactose, mannose, etc.). Thus, sugars are added one by one as the protein moves through the Golgi, like an oligosaccharide assembly line. Different glycosyltransferases reside in the different cisternae, which means that a glycosyltransferase that resides in a later compartment will be unable to add its sugar if a step in an earlier compartment does not happen. The result is an array of possible oligosaccharides that can be built onto the proteins, which can then be used for specific functions later on.

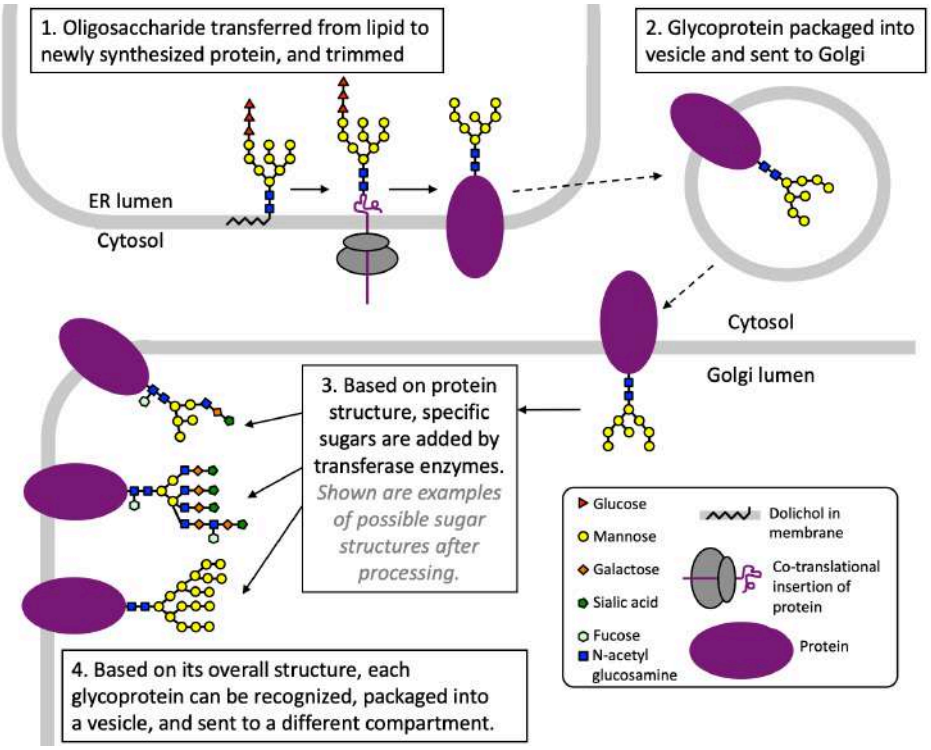


Figure 04-27: The process of glycosylation. (1) It starts in the ER, where a sugar "tree" is transferred from a specific glycolipid (by the name of dolichol) in the ER membrane to the protein as it is co-translationally inserted into the membrane. The sugars are then trimmed in the ER before being packaged into a vesicle and sent to the Golgi (2). (3) Once in the Golgi, the sugar trees are then modified again based on the overall structure of the protein. The result is different glycosylation patterns on different proteins. (4) Once this process is complete, the protein has made it to the trans Golgi network (TGN) and is ready to be packaged into vesicles and sent to the next location on its journey. This image by Dr. Megan Barker is shared under a CC BY-NC-SA 4.0 license.

Cell Wall Production in the Golgi Apparatus

Of the various eukaryotic kingdoms, several of them include cells that are commonly surrounded by a cell wall (specifically, plants, algae, and fungi). Animals do not have cell walls, and protists are extremely variable in whether they have cell walls, not to mention what those walls are made of, which

makes them a topic for a whole course on their own. The cell walls of plants, algae, and fungi are made primarily of polysaccharides. Since the Golgi is the location of virtually all of the glycosyltransferases that create new covalent bonds using sugars, it has a major role to play in the production of cell walls.

Cell walls are made of complex, often highly branched sugar chains that create a gel matrix in which cellulose is embedded. We saw examples of this before, in Chapter 2 (see Figure 02-20). These complex polysaccharides are synthesized one sugar at a time as the molecules move from *cis* to *trans* inside the Golgi, in a way similar to what we just saw for glycoproteins. They are packaged into vesicles in the TGN and then shipped to the plasma membrane, where they are secreted and then incorporated into the existing plant cell wall. The most common cell wall compound made in the Golgi is pectin, which is a highly branched molecule that helps hold all of the parts together.
This is exactly what it does in jam as well! It acts as a gelling agent.
The cellulose is not created in the Golgi but instead uses a special protein structure at the plasma membrane known as a **rosette**.

The *Trans* Golgi Network (TGN) Is a Major Cargo Sorting Center for Secretion

Once the proteins arrive in the TGN, it means that they have been through all of their needed processing and are ready to be sent to the next place / final destination. The primary function of the TGN is to act as a sorting center. Cargo moving through the Golgi will end up at the TGN, and based on the structure of the protein and/or the carbohydrates attached to it, it will be sorted into one of two major routes:

1. The **secretory pathway** (a.k.a. exocytosis): This pathway is for proteins that are destined either to be excreted into the extracellular space or to become an integral membrane protein in the plasma membrane.
2. The **lysosomal pathway**: As its name states, the lysosomal pathway is used for proteins that are residents of the lysosome.

As a result of its role, the TGN is less "pancake-like" than the other cisternae of the Golgi. In electron microscopy, vesicles are often captured in a variety of states of budding, making the TGN somewhat bulbous (Figure 04-23). There is quite a lot of activity at the TGN, as it ensures that each protein gets packaged into the correct transport vesicle so that it can be sent off to the correct destination for its function.

Interestingly, despite the fact that we spent a whole topic examining vesicle formation using coat proteins, our understanding of how cargo gets packaged into vesicles at the TGN is vague. In some cases, the structures produced are too large to be packaged into standard coated vesicles. There have been debates as to whether the TGN is able to peel off from the rest of the Golgi stack and act as its own transport compartment. In addition, recent findings show that both COPI *and* COPII coats are able to produce larger tubules as well as smaller vesicles, though it is still unclear how often this happens. Additionally, there are other cases where there is no evidence for *any* known coat protein being involved. Since the trafficking of vesicles at the TGN is an essential cellular process, it is very difficult to study. There are still a great many unanswered questions in this field.

TOPIC 4.4: PATHWAYS OF THE ENDOMEMBRANE SYSTEM AND THE TECHNIQUES TO STUDY THEM

Learning Goals

- Describe the secretory, lysosomal, and endocytic pathways and trace the path of cargo through each of these pathways.
- Differentiate between the functions of endosomes and lysosomes, and describe the structural relationship between the two compartments.
- Compare and contrast between constitutive and regulated versions of secretion and endocytosis.
- Interpret the results of the different experimental tools studying traffic through the endomembrane system.

After newly synthesized proteins have traveled from the ER to the Golgi, they hit a crossroads at the *trans* Golgi network (TGN), where their paths diverge, depending on destination and function. As such, this is a good time to step back and take a look at the bigger picture of the endomembrane system and how proteins and other cargo travel through it. Thus, in this last topic of the chapter, we will discuss each of the major pathways that cargo takes through the endomembrane system. As is our practice in this textbook, we will end by exploring the experimental techniques scientists use to study how cargo moves through the different pathways.

First, to remind you, we will be looking at three primary pathways through the endomembrane system. They were originally identified way back in Figure 04-01:

1. The **secretory pathway**, which is also often called the "default pathway" through the organelles. It is the path that most newly synthesized ER-targeted proteins will take through the endomembrane system on their way out of the cell.
2. The **lysosomal pathway**, which is the path that newly synthesized digestive enzymes will take to get to the lysosome, which is their final destination.
3. The **endocytic pathway**, which is the path inward from the cell exterior for proteins and molecules brought in by the cell. Material brought in via endocytosis will also likely end up in the lysosome, where the macromolecules are broken up into their basic building blocks to be reused in biosynthetic pathways.

The Secretory Pathway (a.k.a. Exocytosis)

As mentioned above, proteins that are destined to be sent out of the cell (or to live in the plasma membrane) have two different versions of the secretory pathway that they might take (Figure 04-28):

1. In **regulated secretion**, proteins are packaged into vesicles that are stored in the cell until they are secreted in response to a specific signal.
2. In **constitutive secretion**, vesicles continuously form and carry proteins from the Golgi to the

cell surface. This is also sometimes called the *default secretory pathway*, as it requires no special targeting sequence (beyond the ER insertion sequence required to first enter the endomembrane system) for cargo to get secreted via this pathway.

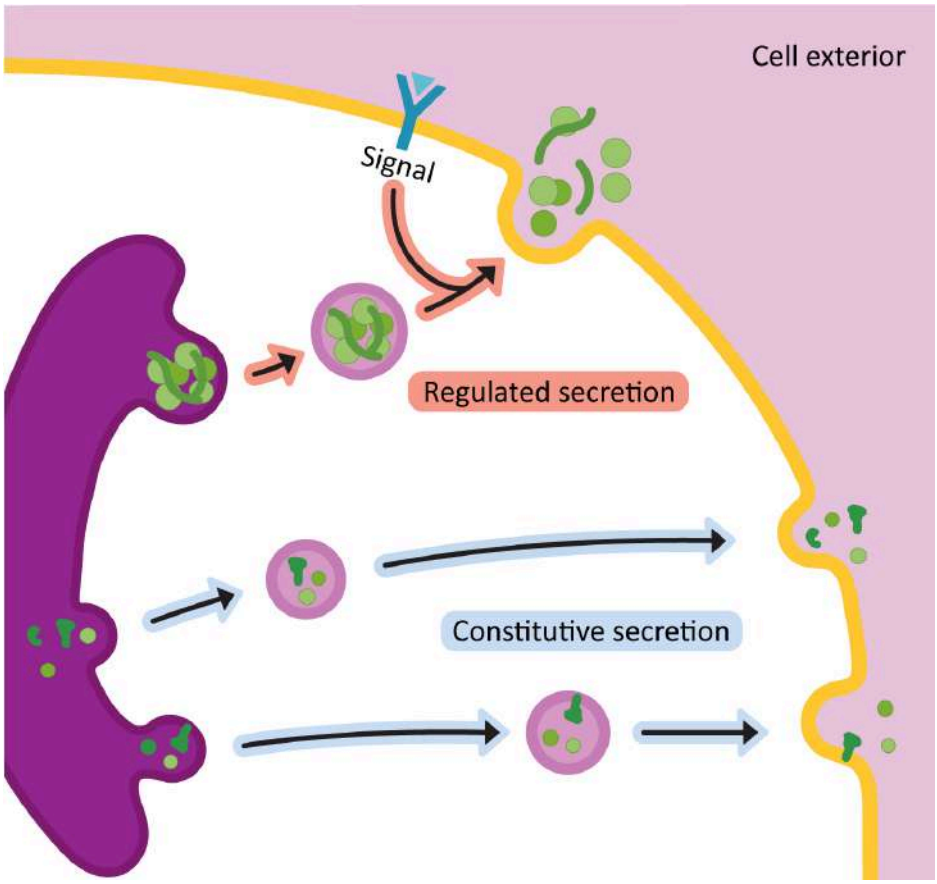


Figure 04-28: Constitutive versus regulated secretion pathways from the Golgi. Constitutive secretion occurs continuously, while regulated secretion occurs in response to a stimulus. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

With two pathways that both lead to the plasma membrane and out of the cell, why might a particular protein be secreted via one pathway or the other? Consider, for a moment, why one protein's exit from the cell might need to be controlled and another one might not. Can you think of examples of proteins that might be better served by traveling through one pathway or the other? These are questions to keep in mind as we explore the different forms of secretion.

Constitutive Secretion

The constitutive exocytic pathway operates continually in all cells and supplies a continuous stream of vesicles containing lipids, proteins, and polysaccharides to the plasma membrane for release outside of the cell. When we think of "secretion," this is often the pathway we are referring to. Because it exists in all cells, and because it appears to require no specific signal for cargo to enter this pathway, we generally consider this to be the *default pathway* through the endomembrane system. A protein that is able to enter the endomembrane system will contain an **ER insertion sequence**. However, with no other targeting sequences, it will eventually get secreted out of the cell. Examples of components that are secreted via the constitutive exocytic pathway are the following:

- extracellular matrix components (in animals),
- cell wall components (in plants and fungi), and
- antibodies in activated, antibody-producing B-cells.

Cells that are currently undergoing a great deal of constitutive secretion often can be identified in transmission electron microscopy (TEM) due to the large amount of ER in the cytosol (especially rough ER).

Depending on the cargo, it is thought that constitutive secretion may, or may not, require coat proteins. Clathrin-coated vesicles have been shown to be involved in some cases, but there are other cases where clathrin is most definitely *not* involved. As we said, sometimes COPI or COPII gets involved, though this seems to be mostly for large cargo (such as collagen and other large extracellular matrix components) that is far too big to fit into a clathrin-coated vesicle. In some cases, we have evidence that the vesicles form and that transport is happening, but we have no understanding of how the vesicles are forming. Thus, this is another area of active and ongoing research.

Regulated Secretion

Like the name says, the release of vesicles carrying cargo in the regulated secretory pathway requires a stimulus from either inside or outside of the cell before the final fusion of vesicles with the plasma membrane is allowed to take place. The vesicles form and then sit in the cytosol until they are ready. As a result, in TEM, a high concentration of vesicles in the cytosol is considered to be a marker of a cell that is involved in regulated secretion.

It is thought that vesicle formation in regulated secretion is aided by the proteins that use this pathway for export. There is evidence that the cargo proteins tend to "clump together" in acidic conditions (we call this **aggregation**, and the "clumps" are known as **protein aggregates**). This is helpful, as the Golgi apparatus tends to become more acidic from the *cis* to *trans* sides of the stack (from a pH of about 6.8 in the *cis* cisternae to as low as 6.3 in the *trans* cisternae). The result is that in the TGN, we find these large protein aggregates, which, if large enough, can then push up against the membrane and help promote curvature. Like the constitutive pathway, none of the known vesicle budding machinery seems to be involved, so these protein aggregates are likely extremely important to the formation of vesicles in regulated secretion.

The regulated secretory pathway is used for products that must be secreted very quickly, but only on demand. For example, insulin is produced by the beta cells of the pancreas and stored in *dense core secretory granules*. When blood sugar increases to a threshold level, insulin-containing secretory granules fuse with the plasma membrane, releasing insulin into the blood. Some other examples of regulated secretion are the release of neurotransmitters in the nervous system and the release of breast milk during infant feeding.

In some cases, regulated secretion allows for further processing of the protein after it leaves the Golgi but before it gets sent out of the cell. For example, many of the secreted digestive enzymes in our gut are further processed after being packaged into vesicles. Often this involves a section of the protein being removed, which activates it. Our previous example, insulin, is another protein that undergoes processing while it is in vesicles, waiting to be secreted.

The Lysosomal Pathway

The other major pathway for proteins to take from the TGN is what we call the **lysosomal pathway**.

The **lysosome** is a specialized compartment that is the site of intracellular digestion. Lysosomes contain a mixture of some 40 different types of digestive enzymes, including those that degrade nucleic acids, proteins, carbohydrates, and lipids.

Like the Golgi, there is quite a lot of variation in the lysosome and endosome structures in different tissues and eukaryotic species. Plants are able to use their **vacuole** as a lysosome instead of having a separate compartment for it. In white blood cells that engulf bacteria as part of our immune response, the lysosomal enzymes are delivered to the large **phagosome** directly, using vesicles, rather than via the mechanism we will describe here.

It is also worth noting that lysosomes are but one of many different membrane-bound compartments that exist in this "post-Golgi" region (meaning the various destinations that the TGN may be sending cargo to *after* it leaves the Golgi). Lysosomes develop from another "post-Golgi" compartment, known as the **endosome**, using a mechanism that is similar to the **cisternal maturation model** of the Golgi. There are several different endosome-like compartments that are recognized, and discussed by scientists, based on their functions. The difference between the various compartments is fuzzy at times, as one compartment can "become" another compartment through the addition and removal of specific resident proteins. For example, we sometimes differentiate endosomes as "early endosomes," which are not as acidic, and "late endosomes," which are more acidic and have more characteristics of the lysosome that they will eventually become.

There are also recycling endosomes, multivesicular bodies, and a compartment called the endo-lysosome. While we don't have time to go into all of the details of each of these compartments, this makes the point that we do our best to give you a set of well-defined "rules" to understand the processes we're exploring. However, we cannot completely remove the complexity, as it underlies everything that we are going to talk about. This function of this part of the cell is a bit fuzzy, somewhat confusing, and also very likely species specific.

Targeting Proteins to the Lysosome

Like all of the other organelles we've explored, in order for a protein to be directed to the lysosome, it must have a specific sorting signal within the chemical structure of the protein that the cell can recognize. Lysosomal targeting is somewhat unique compared to the other targeting signals covered so far, as it is actually a carbohydrate signal that is attached to the protein via **N-linked glycosylation** (described previously in Figure 04-27). The carbohydrate signal used for lysosomal targeting is called **mannose-6-phosphate**, or M6P. As the name implies, the M6P signal is a six-carbon mannose sugar, which has a phosphate group attached at the 6-carbon. The M6P is part of a larger oligosaccharide that was added during N-linked glycosylation and is usually at the end of the chain that is farthest away from the protein it's attached to. These modified mannose sugars can be identified by receptors in the TGN so that they get packaged into vesicles and sent to the endosome (Figure 04-29).

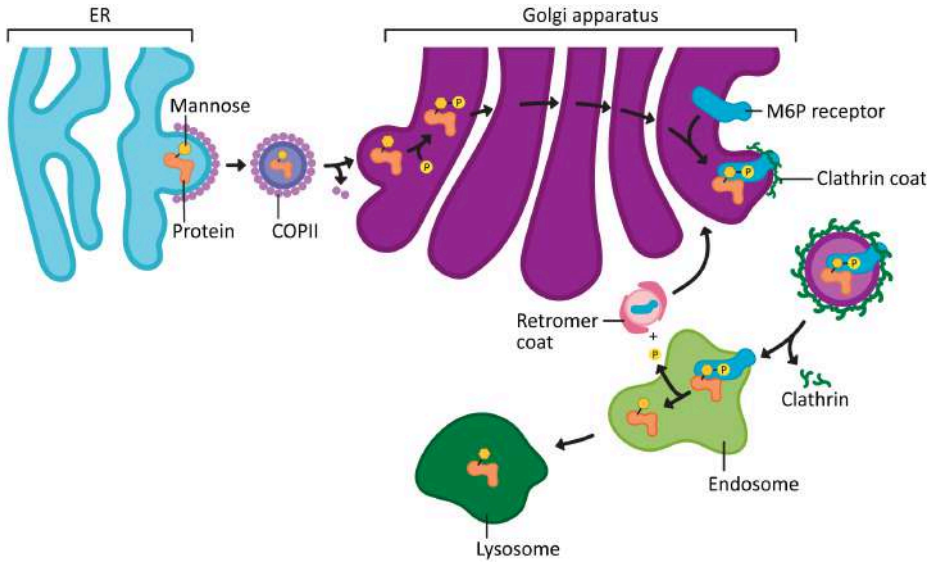


Figure 04-29: Targeting to the lysosome. Proteins destined for their jobs in the lysosome are picked up by the M6P receptor in the TGN and travel in vesicles to the endosome. The pH of the endosome causes the lysosomal resident to separate from the receptor, and the receptor is sent back to the TGN for more rounds of transport. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Here's how it works:

1. N-linked glycosylation occurs as usual, with the oligosaccharide "tree" being added in the ER and trimmed.
2. In the *cis* Golgi, a phosphate is added to the sixth carbon of a specific mannose in the tree. The result of this is twofold:
    a. The M6P tag has been formed, which will be used later by the receptor.
    b. This also stops other glycosyltransferases in the Golgi from further modifying the sugar "tree." That way the M6P tag isn't masked or destroyed.
3. In the TGN, the M6P is exposed so that it can bind to the **mannose-6-phosphate receptor (M6PR)**.
4. The protein-receptor complex gets packaged into clathrin-coated vesicles, and the vesicles are sent from the TGN to the endosome.
5. There, the lysosomal proteins are separated from their receptors (due to the increased acidity in the endosome).
6. The lysosomal proteins stay in the endosome, which will eventually mature into the lysosome, and the now-empty M6P receptors are recycled back to the *trans* Golgi network using a unique type of vesicle called the *retromer*.

Function of the Lysosome

The function of the lysosome is simple: its job is to break down cellular or extracellular material into its basic building blocks. These materials could include large polymers and other components brought in from the exterior via the endocytic pathway. Lysosomes are also involved in the breakdown of bacteria that are engulfed by the cell during a process known as **phagocytosis** as well as in the

destruction of internal components that no longer function (such as a broken-down mitochondria). As a result, in some cases, the lysosome itself is the site of digestion, while in others (such as phagocytosis), the lysosome fuses with another compartment, which releases the digestive enzymes into the new compartment to aid in the digestion of larger structures.

The lysosome is a very acidic compartment. It usually has a pH in the range of 4–5. The acidic pH of the lysosome (and the endosomes they are derived from) is maintained by membrane-bound ATP-driven pumps that move protons into the lysosomal lumen. The lysosome has a single membrane that surrounds it. Additionally, all of the proteins in the lysosome are heavily glycosylated. This makes sense for two reasons:

- First, the targeting signal is a carbohydrate signal (M6P), so the only proteins that can get targeted to the lysosome are, by definition, glycosylated proteins.
- Second, carbohydrate coverings are often used as protection in extreme environments such as this one. For example, your stomach is also an extreme, highly acidic environment that has a very similar function to the lysosome, except on a much larger scale! The cells that make up the stomach lining are protected from harm by a thick layer of mucus. Mucus is simply made up of secreted complex carbohydrates. In the lysosome, instead of secreted carbohydrates, the proteins themselves are heavily glycosylated, which provides protection to the lysosomal membrane from the digestive enzymes within it.

One more thing to reiterate: in many organisms (plants, fungi, and many ciliated protists), there is a large **vacuole** but no obvious lysosomes. This is because vacuoles often perform the role of the lysosome in the cells that have them. There is quite a lot of variation in how this is arranged in organisms from different kingdoms, so we won't say much about it, but it's still important for you to know the relationship between the vacuole and the lysosome.

Lysosomes in TEM

Figure 04-30 shows an electron micrograph of three lysosomes in the cytosol of a ciliated unicellular protist. Lysosomes often contain what looks like strange bits and pieces. This material and the internal structure of the lysosome change over time because the contents are derived from the remains of cellular digestion. Sometimes the bits inside are very dark, and sometimes they're not. This can make lysosomes a little bit difficult to identify in TEM, as their appearance changes depending on what they're digesting.

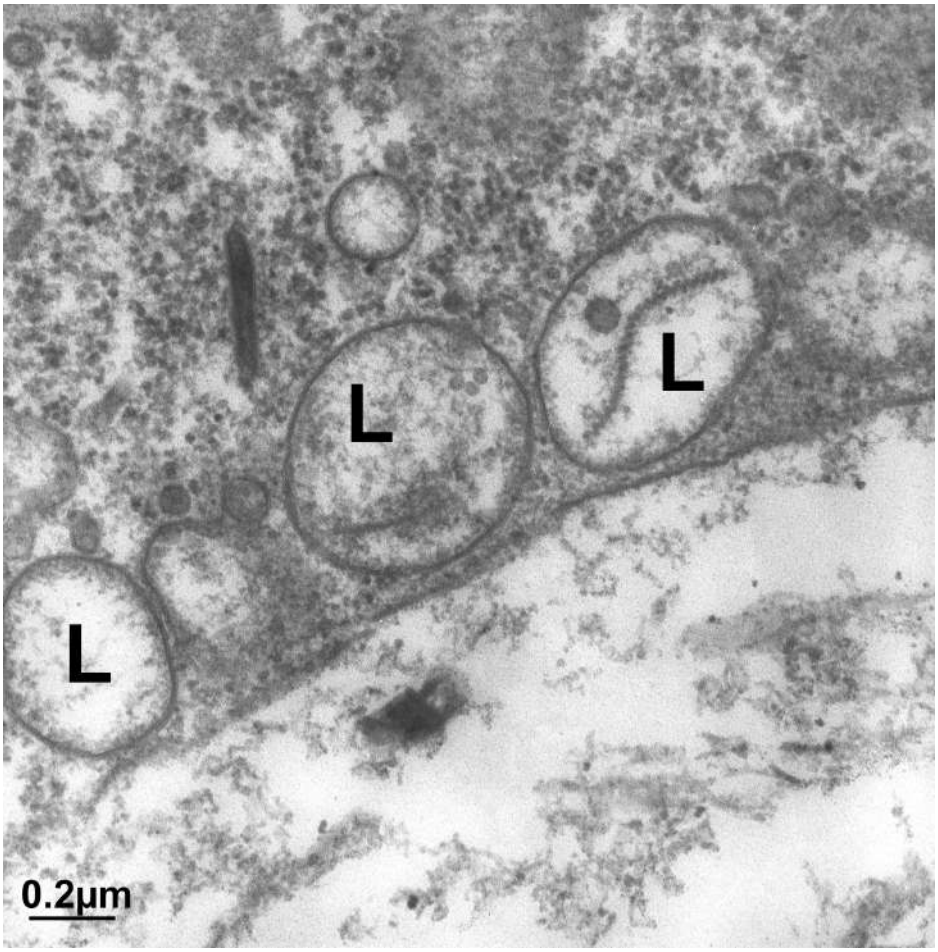


Figure 04-30: A TEM image showing three lysosomes (L) in the cytosol of the single-celled protist Paramecium multimicronucleatum. Image from Richard Allen (University of Hawaii) (2011) CIL:36739, Paramecium multimicronucleatum, cell by organism, eukaryotic cell, Eukaryotic Protist, Ciliated Protist. CIL. Dataset. https://doi.org/doi:10.7295/W9CIL36739. This image is in the public domain.

The Endocytic Pathway (a.k.a. Endocytosis)

The final pathway through the endomembrane system is the endocytic pathway, also known as **endocytosis**. This pathway starts at the plasma membrane instead of the ER. In this pathway, material is brought in from the exterior of the cell to the endosome. From the endosome, it can go several places:

- Proteins and other structures that are destined for degradation will go to the lysosome to be digested.
- Some cell-surface receptors are recycled back out to the plasma membrane after releasing their cargo in the endosome.
- There is some evidence, especially in plants, that some endocytic cargo gets sent to the TGN, where it may be sorted and redirected to another site afterward.

You can trace all these pathways in Figure 04-15, found in Topic 4.2.

Types of Endocytosis

When we discuss the endocytic pathway, we are discussing three distinct forms of internalization (Figure 04-31):

5. **Phagocytosis**: the ingestion of large particles by cells. The particles are taken into **phagosomes** or food vacuoles. In multicellular organisms, this occurs primarily in specialized cells, like macrophages. However, aquatic single-celled protists, such as amoeba or paramecium, are also capable of this.
6. **Pinocytosis**: the ingestion of fluid, ions, and other small particulates via small vesicles <150 nm in diameter. This is carried out by all cells and is believed to be nonspecific.
7. **Receptor-mediated endocytosis**: the ingestion of specific molecules (e.g., cholesterol or iron) that are captured by cell-surface receptors to be packaged into vesicles.

We will look at each of these in turn.

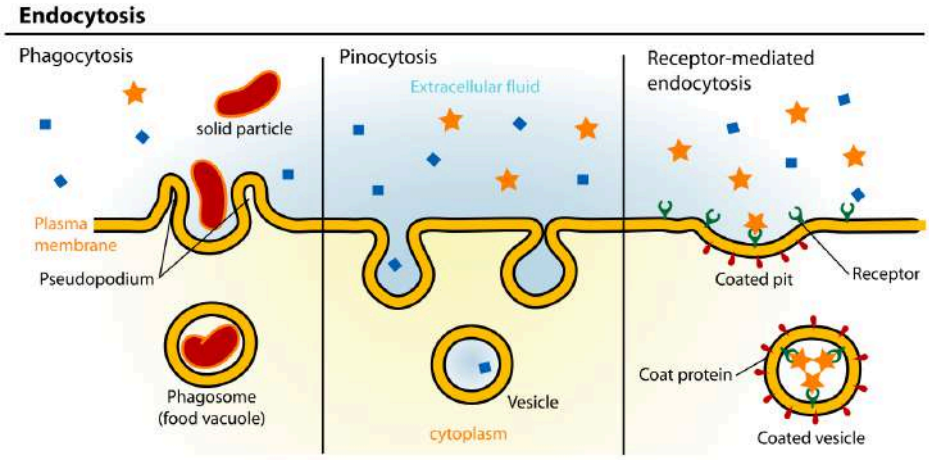


Figure 04-31: Three different kinds of endocytosis. Phagocytosis is ingestion of large particles. Pinocytosis is continuous small-scale endocytosis. Receptor-mediated endocytosis involves a signal to specifically bring in certain molecules. Image by LadyofHats is in the public domain.

8. Phagocytosis

In **phagocytosis**, bacteria or other large particles are brought into the cell using large membrane-bound compartments we call **phagosomes**. Phagocytosis generally occurs as a response to the activation of cell-surface receptors, which results in the opening of calcium channels. The increased internal calcium concentration then signals the actin cytoskeleton to rearrange. The phagosome is then formed. Phagocytic cells play an important role in the removal of foreign organisms (like bacteria) and in the scavenging of dead or damaged cells in the immune system of many animals. Many of our white blood cells are capable of phagocytosis. In single-celled protists, this is also a mechanism to ingest food.

Phagosomes are complicated organelles to produce, as they require a great deal of membrane, which needs to be pulled up and around the structure that is being engulfed. Actin is thought to be heavily involved in this process, as it is a major mechanism used by the cell to push the membrane outward (more on this in Chapter 6), but the mechanism is still under investigation.

Once the phagosome has formed, a lysosome will fuse with it to release digestive enzymes into the compartment, which can then be activated as the compartment acidifies. Interestingly, the acidic phagosome is also a site of active immune warfare in our bodies. Some bacteria are known to use the phagosome to invade their host. They are able to block lysosomal digestion and remain functional inside the phagosome. The bacteria that cause tuberculosis, Legionnaire's disease, and listeria all manipulate the phagosome to protect themselves from the immune system in humans and other mammals.

9. Pinocytosis

This type of endocytosis can also be thought of as constitutive endocytosis, because it is thought to occur all of the time. Pinocytosis is relatively nonselective in what it brings into the cell (water, small ions, and other nutrients). It is also one of the ways that the cell manages to maintain a constant size despite the continuous flow of vesicles to the plasma membrane via secretion. For example, macrophages ingest about 25% of their own volume in fluid each hour and an area of membrane equal to its *entire surface* every 30 minutes. This is offset by the outward flow of vesicles (including membrane, fluid, and cargo) through exocytosis.

Pinocytosis may or may not use clathrin-coated vesicles for transport. More commonly, a special structure called a *caveola* is used. Caveolae (plural) are formed using a different protein complex that includes a protein named caveolin. They're not as well understood as clathrin or the COP coats, so we won't discuss them further.

10. Receptor-Mediated Endocytosis

Receptor-mediated endocytosis is a much more selective process than pinocytosis. It's triggered by the binding of macromolecules to receptor proteins at the cell surface. The increased concentration of bound molecules triggers the formation of a vesicle. Endocytic cargo may be concentrated up to a thousandfold by binding and accumulating on the cell surface before incorporation into vesicles. Cargo molecules bind to receptor molecules that then bind to the adaptor protein (adaptin) and coat protein (clathrin in this case) that result in their concentration in the area of the forming membrane bud that will give rise to the endocytic vesicle.

An excellent example of this is the internalization of cholesterol. As you know from its chemical structure, cholesterol is not very soluble in aqueous environments, such as our blood, and this makes it difficult to transport from our digestive tract (where it's absorbed from our food) to the rest of the cells in our body. As such, the cell packages cholesterol into a structure known as a *low-density lipoprotein (LDL)*. The LDL can travel through the bloodstream and get endocytosed by the cells that require it (Figure 04-32).

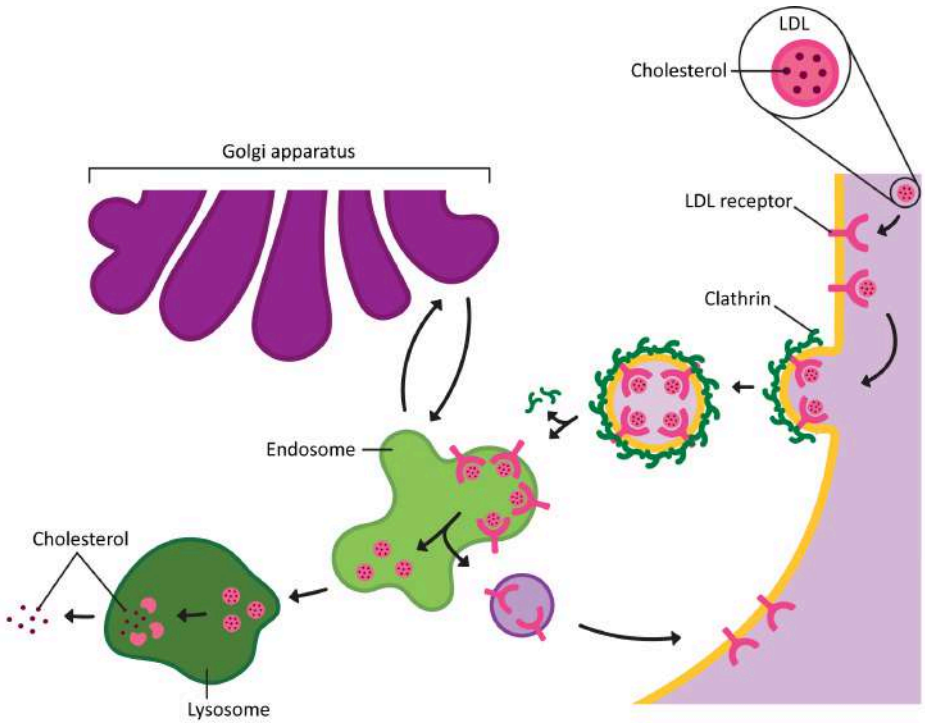


Figure 04-32: Receptor-mediated endocytosis. Cargo (such as the LDL shown here) is brought in from the cell surface by binding to a receptor. The complex is internalized and send to the endosome, where the cargo and receptor are separated. The cargo is then transported to the lysosome where it is broken down into component pieces. In this case, cholesterol is released to be used in cellular processes. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The following are the steps of receptor-mediated endocytosis of the LDL:

1. The LDL complex binds to receptors on the cell surface, which triggers the formation of clathrin-coated vesicles.
2. The vesicles then lose their coats and travel to and fuse with the endosome.
3. In the acidic environment of the endosome, the LDL separates from its receptor.
4. The LDL receptors are returned to the plasma membrane via transport vesicles. Cholesterol is released from LDL in the lysosome and available to use in membranes throughout the cell.

The Endosome Is Also a Sorting Center

The endosome is the second major sorting center in the endomembrane system (the first one being the TGN of the Golgi). It is the junction point between the lysosomal pathway and the endocytic pathway (Figure 04-32). This means that it must be able to differentiate between cargo that is headed to the lysosome and receptors that must be returned to either the TGN or the plasma membrane.

As if that were not complicated enough, endosomes receive receptors and cargo from the plasma membrane that may be headed to several different destinations:

- Many receptors are returned to the same plasma membrane domain from which they came, while their cargo is passed onto the lysosome. The LDL receptor is an example of this (shown in Figure 04-32).

- Some receptors are not recycled but instead sent to the lysosome to be degraded. This can be done with or without their cargo. Many growth hormones and receptors for irreversible developmental cues will follow this route.
- Finally, some receptors, along with their cargo, are transferred to another plasma membrane domain. This allows specific cargo to be passed from one side of the cell to another, in a process known as **transcytosis**. This is the process by which antibodies get passed from parent to child through the breast milk–producing cells when a child nurses.

Whether the cargo has the same fate as the receptor it rode in with or not is, in part, dependent on whether the acidic environment of the endosome causes cargo release. If the cargo is one that gets released by the acidic endosomal environment, then the cargo alone will end up in the lysosome, and the receptor is free to be recycled. Receptors that do not respond to low pH by releasing their cargo are more likely to travel with the cargo to the lysosome and get degraded.

Experimental Methods for Studying the Endomembrane System

As we have said many times in this chapter, the endomembrane system of a cell is essential, which means that the cell cannot survive without it. Essential cellular processes such as this create an interesting conundrum to researchers who wish to study them. One of the most common methods for studying a biological system is to disrupt it and study the effects. But when a system is essential, the cell dies if you disrupt it. Unsurprisingly, it is extremely difficult to learn much from a dead cell. Thus, scientists have had to get creative to find ways to study the secretory system in cells. Their methods have been very innovative and have been vital to our understanding of cell biology.

11. Visual Techniques—Microscopy

One of the most commonly used techniques these days to study how cargo moves through the endomembrane system is live-cell imaging of fluorescently labeled proteins. (*Note: We have already spoken about visual techniques at length back in Chapter 1, so if you need a refresher, you may want to refer to the material found there.*) We can combine live-cell imaging with other techniques described below to track how the system changes when mutations are made or blockages are temporarily induced.

As an example of the power of fluorescent microscopy to teach us about the endomembrane system, Video 04-08 shows the ER (green network) and the Golgi (purple blobs) in a plant leaf cell. In TEM we would see more detail of the structure, but since the samples must be dead and fixed, we cannot learn about the dynamics of these structures. In this live-cell video, we see the Golgi and ER both moving and changing shape dramatically throughout.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=149#oembed-8

Video 04-08: Fluorescent Golgi bodies (purple blobs) moving around a

plant cell. The ER network can be seen in green.

12. Biochemical Techniques

When researchers were first trying to figure out the endomembrane system, a common approach was literally to take it apart to investigate the job of each of these compartments in vitro (i.e., in a test tube). The biochemical techniques used were quite involved and complex. One common way to isolate different parts of the endomembrane system was to use a process called **cell fractionation**. In this process, the cells were literally put in a blender to break them up and then centrifuged at varying speeds, under different conditions, to separate rough ER from smooth ER from Golgi and so on. Researchers could then work with isolated membranes to see what each one was capable of on its own. In biology, isolating your "thing" of interest is a very common way of trying to understand function. Figure 04-33 shows an example of reconstructing an ER from isolated membranes in a test tube for an experiment.

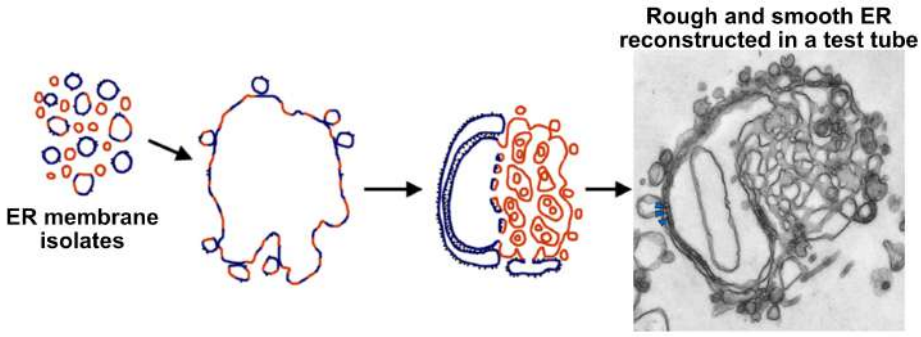


Figure 04-33: Biochemical technique example—isolating ER membranes. Here we see the steps that the isolated membranes of the ER go through as they reform themselves into a complete, interconnected endoplasmic reticulum, with both rough and smooth ER regions. "In vitro reconstruction of ER membranes" by Dr. Robin Young is shared under a CC BY-SA 4.0 license.

13. The Development of Temperature-Sensitive Secretion (sec) Mutants in Yeast and the Rise of Mutational Analysis as the Primary Tool for Studying the Endomembrane System

One of the most common ways for cell biologists to study cellular function is to disrupt a protein's ability to function (through mutation of its gene) and then see what processes are disrupted when that protein is absent. When studying essential processes like secretion and the endomembrane system, this can pose a challenge. Many of the proteins involved are so important that the organism cannot survive at all without it. As you can imagine, this creates difficulties for scientists who want to better understand how the proteins of secretion function. The discovery of **temperature-sensitive mutations** was a turning point in our understanding of many essential cellular processes—so much so that the combination of temperature-sensitive mutations and live-cell imaging techniques is a very large part of how endomembrane research is done today.

A temperature-sensitive mutation is one in which the protein amino acid sequence has changed due to the mutation, but it's not so severe a change that the protein is nonfunctional. The protein is still able to fold more or less correctly. At certain temperatures (known as the *permissive temperature*, usually a few degrees lower), these proteins work just fine. They are stable and are capable of doing their job. However, if the temperature is changed even a few degrees (to what is known as the *restrictive*

temperature), the protein becomes unstable, denatures, and is no longer able to maintain its function. As long as the cell is kept at the restrictive temperature, the protein will not regain its function. If left for long enough, the cell will die. If the cell is moved back to the permissive temperature, in time, the protein will renature, regain function, and thus the regular cellular processes will resume.

An excellent example of a temperature-sensitive mutation is that of the coloring of a Siamese cat (Figure 04-34). Siamese cats very famously have lighter beige hair on their bodies and darker brown hair on their extremities (paws, face, tail). The particular coat coloring of a Siamese cat is the result of a temperature-sensitive mutation in the tyrosinase enzyme, which is one of the many enzymes involved in producing the brown color pigment, melanin. The cat's extremities are slightly cooler, so in these areas, the tyrosinase enzyme is functional, and the dark color is produced. Near the animal's core, the temperature is a little bit higher, so the enzyme does not function, melanin is not produced, and the animal has albino fur.

Figure 04-34: A Siamese cat with its characteristic coat coloring. "Neighbor's Siamese" by Cindy McCravey is shared under a CC-BY 2.0 license.

One of the advantages of this system, from a cell biology perspective, is that a temperature-sensitive mutation is *heritable*, because a cell at a permissive temperature can survive long enough to divide and/or reproduce. This is important in experiments, as it is very challenging to generate these cell lines with mutations in specific genes. Even if one does regenerate the same mutated gene, there is the possibility that the same gene with a different mutation will not show the same effects, so being able to maintain a cell line with a specific mutation in an essential gene like the ones of the endomembrane system allows us to spend more time studying each mutation and thus learn more about cell biology as a whole.

For the endomembrane system, a great deal of work has been done on the budding yeast, *Saccharomyces cerevisiae*. There are multiple reasons why they were an ideal system in which to study secretion:

- They are single-celled Eukaryotes with a very short life cycle (~1–2 hours) that tolerate a wide range of temperatures. This makes them ideal for creating temperature-sensitive mutants and studying the effects.
- They have a small genome with very few genes compared to us (~12 million base pairs and ~6,200 genes). A simpler system is easier to study.

As such, a great deal of work has been done in budding yeast in order to study the endomembrane system and secretion. A great many temperature-sensitive mutants exist that have shed light on many different parts of the protein machinery that drives function. They are called the secretion, or *Sec*, mutants. One such example is Figure 04-35, where we see the difference in one of the mutants, called n-Sec1, as restrictive and permissive temperature.

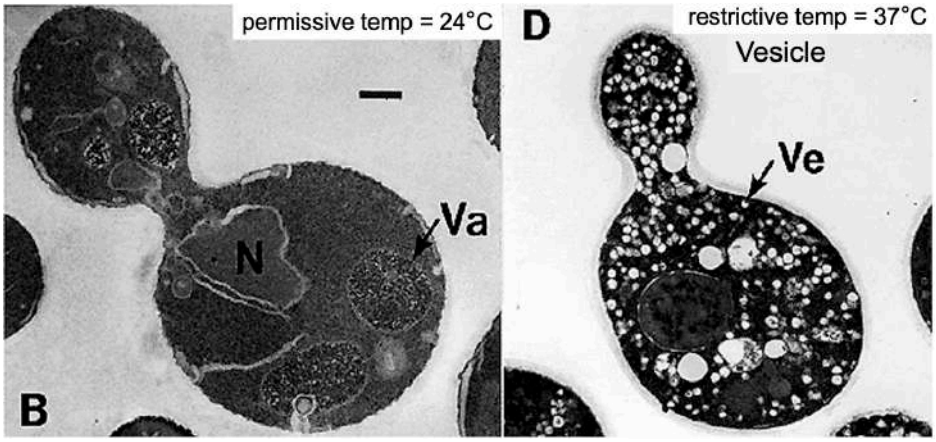


Figure 04-35: The n-Sec1 mutant at permissive (B) and restrictive (D) temperatures. Scale bar in B is representative for both panels and is 0.5 µm. Va: vacuole, Ve: vesicle, N: nucleus. From Novick, P., & Schekman, R. (1979). Secretion and cell-surface growth are blocked in a temperature-sensitive mutant of Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America, 76(4), 1858–1862. https://doi.org/10.1073/pnas.76.4.1858. Reprinted with permission from the authors. Please do not redistribute without permission of the rights holder. Adapted (labels added) by Dr. Robin Young.

CHAPTER SUMMARY

The endomembrane system is a dynamic system involving many organelles, each with a unique job to do. Fittingly, the regulation of this system is equally dynamic and complex. In this chapter, we have discussed the roles of the ER, Golgi, endosomes, and lysosome and how resident and transient proteins are sent to or pass through these different compartments.

The ER is the entry point into the endomembrane, and proteins targeted anywhere in the endomembrane system contain an ER insertion sequence that is part of their primary sequence. Depending on the type and location of these ER targeting sequences, the protein can be either soluble or membrane bound. This starts their journey in the endomembrane system.

Proteins move to different organelles within the endomembrane system via a small membrane-bound vesicle. These vesicles provide specific transport based on the proteins contained in each vesicle and the targeting machinery on the surface.

The Golgi is the protein processing and sorting center. All proteins that enter into the

endomembrane system pass through this hub. It is the primary site of glycosylation, and it is the primary site of production of cell walls in plants, fungi, and algae. Proteins can reach the cell exterior, be sent back to the ER, or be targeted to the endosome/lysosome from the Golgi.

Finally, a secondary sorting hub, known as the endosome, receives cargo from the cell surface and redirects it. Proteins and other molecules can be sorted and returned to the Golgi / cell surface or sent to the lysosome to be degraded.

This is a crucial system that is also quite dynamic, which makes it difficult to study. However, some pivotal experimentation as well as advances in technology have allowed scientists to gain significant insights into this highly intricate system. This chapter simply scratches the surface of the scientific knowledge on the endomembrane system and we hope that you find it as interesting as we do. If not, we hope that you have at least a deeper appreciation for the importance and complexity.

Review Questions

Note on usage of these questions: *Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.*

Topic 4.1: The Endoplasmic Reticulum as the Gateway to the Endomembrane System

1. Define the following terms and determine which ones might be synonyms of each other:
    a. targeting sequence
    b. N-terminal START sequence
    c. signal recognition particle
    d. signal recognition particle receptor
    e. signal peptidase
    f. START transfer sequence
    g. STOP transfer sequence
2. Describe the differences between smooth, rough, and transitional ER.
3. How come disulfide bridges are more common in secreted proteins than in cytosolic proteins?
4. Define the role of chaperones in protein processing. What determines whether a given protein will need chaperones for proper folding?
5. How many ways can a given protein fold? What determines the folding pathway that a given protein takes?
6. What happens to proteins that fail to fold properly in the ER?
7. Explain the idea of a cellular compartment, the ideas of specificity of content and function, and the "sidedness" of membranes. How does the cell maintain the proper "side" of the

membrane exposed to the cytosol in all organelles (including the plasma membrane)?

8. Outline the mechanisms by which proteins are transported to different membrane-bound compartments in the cell.
9. What are the common properties of "START transfer" and "STOP transfer" sequences?
10. What determines whether the N-terminal end of a protein is in the cytosol or inside the ER cisterna?
11. Describe the general steps to protein insertion into the ER.
12. Many membrane proteins contain a number of membrane-crossing domains. How do these proteins get inserted into the membrane?
13. Name a few examples of proteins that might be ER resident proteins and how they are able to stay in the ER.
14. Compare and contrast the mechanism of protein insertion in the ER with nuclear import. Think about the ideas of necessary and sufficient and how they apply to the endomembrane protein targeting system.

Topic 4.2: Vesicle Transport

1. Make a table of the different coat proteins, and include all of the protein components for each one (coat, adaptor, GTPase, and whether it needs a separate scission protein, location in the cell, and/or trafficking pathways)
2. What is the role of clathrin or other coat proteins in the selection of cargo and the formation of membrane vesicles?
3. How do coat proteins facilitate the process of vesicle formation?
4. What sort of forces need to be overcome to form a vesicle?
5. How do Rabs and tethers facilitate vesicle docking?
6. What are SNAREs, and how do they regulate membrane fusion?

Topic 4.3: Pathways through the Endomembrane System

1. Define the following terms and determine which ones might be synonyms of each other:
    a. *cis* cisternae
    b. medial cisternae
    c. *trans* cisternae
    d. *trans* Golgi network
    e. lumen
    f. glycoprotein
    g. proteoglycan
    h. KDEL receptor
2. What portion of a membrane glycoprotein is glycosylated (cytosolic, membrane-crossing domain, or the portion facing the lumen of the ER)?

3. What is the role of glycolipids in the glycosylation of proteins?
4. How do cells "know" where to add oligosaccharides to prospective glycoproteins during the glycosylation process?
5. The lumen of the ER contains a number of enzymes that are involved in protein processing. These proteins are retained in the ER lumen, even though most of the material in the lumen is transferred to the Golgi through vesicle transport. How does this happen?
6. Compare and contrast the structure and function of plant and animal Golgi.
7. Sometimes the enzymes of the ER are accidentally transported to the Golgi. How are these proteins recovered and returned to the ER?
8. Explain the central role of the Golgi in the endomembrane system with respect to protein synthesis, import of material from the cell exterior, vesicle traffic, and cellular digestion.

Topic 4.4: Pathways of the Endomembrane System and the Techniques to Study Them

1. Describe the structure and function of the lysosome. How does it fit into the rest of the endomembrane system?
2. Compare and contrast the regulated and constitutive secretory pathways with regard to the types of material carried and the regulation of secretion.
3. Explain how a protein can continue to be processed even after it gets incorporated into a secretory vesicle.
4. Trace a molecule of a lysosomal protein from its point of synthesis on a ribosome in the cytosol to its final destination in a lysosome. Include everything that must happen to it for successful targeting to the lysosome and where each of those things happen in the cell. Be as specific as you can.
5. Explain how the mannose-6-phosphate receptor is recycled.
6. What determines whether a protein is sorted into the regulated or the constitutive secretory pathway?
7. What is the difference between early and late endosomes?
8. Define receptor-mediated endocytosis and describe the steps. Be as specific as you can.
9. In both endocytosis and secretion, proteins are sorted into vesicles destined for different target compartments. Where does sorting occur in each pathway? What are the similarities and the differences?
10. How does an early endosome "become" a late endosome, and then a lysosome?
11. Do all receptors that get endocytosed get recycled back to the plasma membrane? What happens to the ones that don't?

REFERENCES

Harrison, S., & Kirchhausen, T. (2010, July). Conservation in vesicle coats. *Nature, 466,* 1048–1049. https://doi.org/10.1038/4661048a

Kang, B. H., Nielsen, E., Preuss, M. L., Mastronarde, D., & Staehelin, L. A. (2011). Electron tomography

of RabA4b- and PI-4Kβ1-labeled trans Golgi network compartments in Arabidopsis. *Traffic (Copenhagen, Denmark)*, *12*(3), 313–329. https://doi.org/10.1111/j.1600-0854.2010.01146.x

Noble, A. J., & Stagg, S. M. (2015). COPI gets a fancy new coat: An interconnected scaffolding of proteins bends the membrane to form vesicles. *Science*, *349*(6244), 142–143. http://www.jstor.org/stable/24748469

Novick, P., & Schekman, R. (1979). Secretion and cell-surface growth are blocked in a temperature-sensitive mutant of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(4), 1858–1862. https://doi.org/10.1073/pnas.76.4.1858

Staehelin, L. A., & Kang, B. H. (2008). Nanoscale architecture of endoplasmic reticulum export sites and of Golgi membranes as determined by electron tomography. *Plant Physiology*, *147*(4), 1454–1468, https://doi.org/10.1104/pp.108.120618

Voelz, V. A., Jäger, M., Yao, S., Chen. Y., Zhu, L., Waldauer, S. A., Bowman, G. R., Friedrichs, M., Bakajin, O., Lapidus, L. J., Weiss, S., & Pande, V. S. (2012). Slow unfolded-state structuring in Acyl-CoA binding protein folding revealed by simulation and experiment. *Journal of the American Chemical Society*, *134*(30), 12565–12577. https://doi.org/10.1021/ja302528z

Weidman, P. J. (1995). Anterograde transport through the Golgi complex: Do Golgi tubules hold the key? *Trends in Cell Biology*, *5*(8), 302–305. https://doi.org/10.1016/s0962-8924(00)89046-7

Yuen, C. Y. L., Wang, P., Kang, B. H., Matsumoto, K., & Christopher, D. A. (2017). A non-classical member of the protein disulfide isomerase family, PDI7 of *Arabidopsis thaliana*, localizes to the *cis*-Golgi and endoplasmic reticulum membranes. *Plant and Cell Physiology*, *58*(6), 1103–1117, https://doi.org/10.1093/pcp/pcx057

CHAPTER 5.

MITOCHONDRIA AND CHLOROPLASTS

Structure-Function Relationships

INTRODUCTION

The mitochondrion (plural mitochondria) and the chloroplast are fascinating organelles for a whole variety of reasons. Not only are they our primary energy producers, working together to capture light energy and then transforming it multiple times into various forms of chemical energy that we can use, but they have an origin story that is truly unique from the rest of the eukaryotic cell. On a global scale, mitochondria and chloroplasts are major contributors to the carbon cycle, in which carbon is taken up into the biosphere (through photosynthesis), cycled through multiple organisms as they interact (predation, herbivory, etc.), and then released back into the atmosphere (by cellular respiration) to start the cycle again.

The discussion of mitochondria and chloroplasts can get extremely complex quickly, as their primary function is to transform energy. Thus, much of the focus in introductory textbooks is on the biochemical reactions that take place during the process of photosynthesis and/or cellular respiration. In this textbook, we have chosen to take a different approach. Since this is a cell biology textbook, our focus will be on the structure/function relationships within these organelles rather than the biochemical reactions taking place. We will explore *where* these reactions occur and *how* the structural features of these two organelles support their function as energy producers. There are many other good open educational resources that explore the biochemistry of photosynthesis and cellular respiration. Suggestions include the following:

- Cellular Respiration Resources:
  - Overview of cellular respiration (Amoeba Sisters): https://youtu.be/eJ9Zjc-jdys
  - Molecular model of the mitochondrial electron transport chain and ATP synthase: https://youtu.be/LQmTKxI4Wn4
- Open Textbook Biology 2e chapter on cellular respiration: https://openstax.org/books/biology-2e/pages/7-introduction
- Photosynthesis Resources:
  - Overview of photosynthesis (Amoeba Sisters): https://youtu.be/CMiPYHNNg28
  - Molecular model of photosynthesis: https://youtu.be/jlO8NiPbgrk
  - Open Textbook Biology 2e chapter on photosynthesis: https://openstax.org/books/biology-2e/pages/8-introduction
- Additional Open Biochemistry Textbook Resource on Energy Transformation: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_For_All_(Ahern_Rajagopal_and_Tan)/05%3A_Energy

If information on the biochemistry of these organelles is part of what you need to know, we encourage you to explore these and other options you may be aware of to supplement the information in this chapter.

TOPIC 5.1: EVOLUTIONARY ORIGINS AND PROTEIN TARGETING TO THE MITOCHONDRION AND CHLOROPLAST

Learning Goals

- Explain the evidence for the endosymbiont theory, which states that mitochondria and chloroplasts were once free-living bacteria.
- Compare and contrast the genomes of mitochondria and chloroplasts.
- Explain how proteins that are encoded by the nuclear DNA enter the mitochondria or chloroplasts.
- Compare and contrast protein targeting to mitochondria and chloroplasts with targeting to the nucleus and ER.

Mitochondria and Chloroplasts Evolved from Free-Living Bacteria

There has been a lot of debate over the years about the evolutionary origins of both mitochondria and chloroplasts, not to mention the origin of all of the other internal structures of eukaryotes (Figure 05-01). The fossil record tells us that prokaryotes, such as bacteria, evolved first, which means that somehow a prokaryote had to develop internal structures like the endomembrane system, nucleus, mitochondrion (and sometimes chloroplasts), and many other organelles that have not been observed in modern-day bacteria.

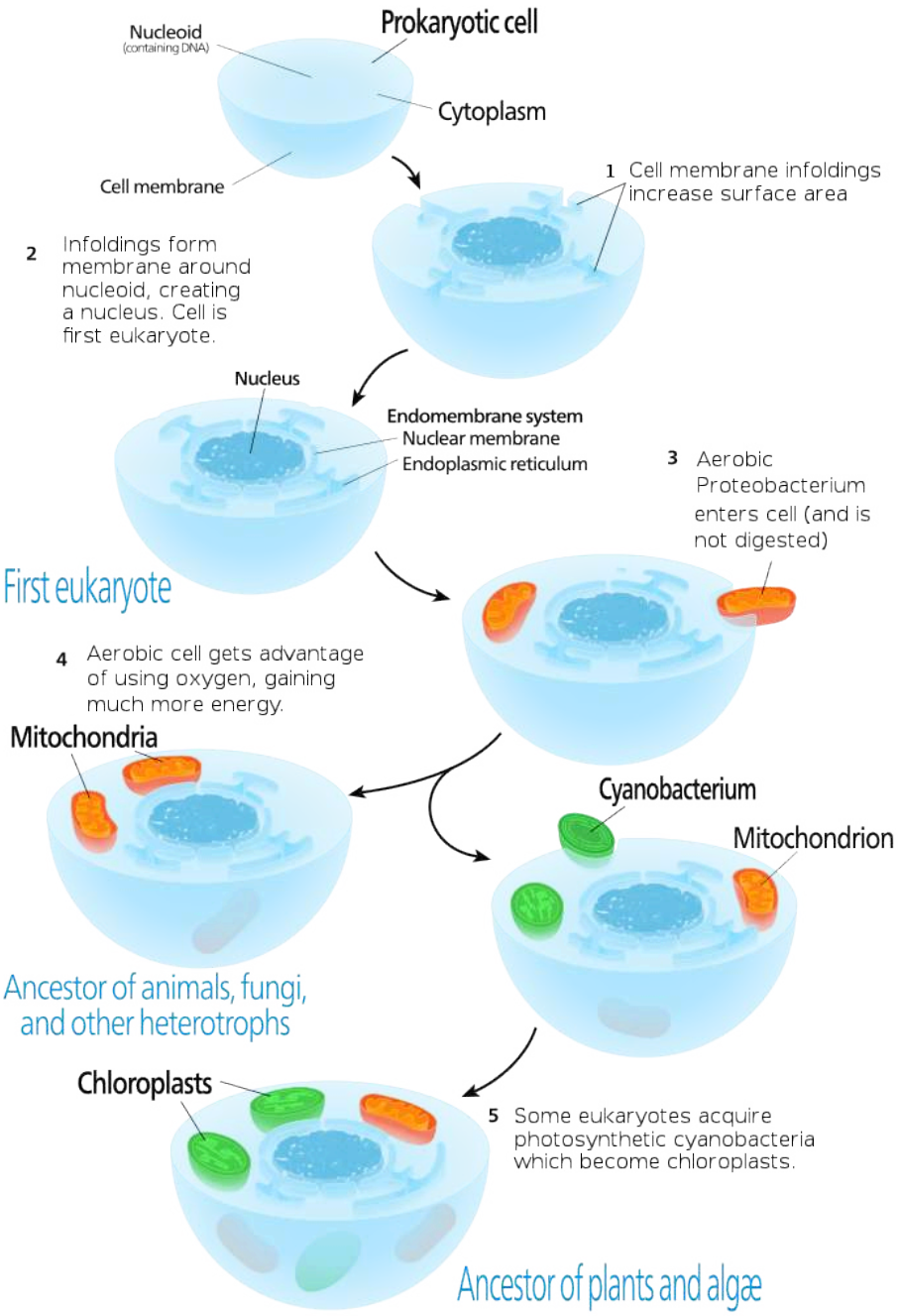


Figure 05-01: A visual representation of the most widely accepted theory about the evolution of Eukaryotes. Note that the "first eukaryote" had mitochondria but then acquired a second organelle, the chloroplast, in order to become photosynthetic. Image by Kelvinsong shared under a CC BY-SA 3.0 license.

A Note on Evolution

We want to briefly acknowledge that this topic, like most in biology, relies heavily on the existence of

evolution as part of the origin of these organelles. Evolution can be a challenging topic for some, depending on both cultural and religious backgrounds. We offer this video, from YouTooBio, which does an excellent job of offering our perspective on this topic. If you are someone who may be feeling conflicted about evolution, we hope that it will help you reduce the conflict you may be feeling so that you can focus on the material at hand.

These days, the most widely accepted theory for the origin of the mitochondria and chloroplasts is a theory known as the **endosymbiont theory**. This theory was first proposed as far back as the late 1800s. It was based primarily on the observation that, much like bacteria, mitochondria and chloroplasts are able to duplicate themselves by binary fission. The endosymbiont theory was not widely accepted until, in the mid-1960s, it was discovered that mitochondria and chloroplasts contained DNA and ribosomes. This meant that they were able to synthesize at least some of their own proteins. This discovery, in addition to several publications by American scientist Dr. Lynn Margulis, who championed the theory, resulted in the eventual acceptance of the idea.

This is a tribute written at the time of Dr. Margulis's death and is an excellent description of her remarkable science.

That is not to say that acceptance was easy. Dr. Margulis's work was rejected and ridiculed in the 1960s, '70s and '80s. However, she persisted with promoting this theory, and her work on this topic is now considered to be revolutionary.

The endosymbiont theory is based on both the structural and genetic similarities among mitochondria, chloroplasts, and bacteria. Video 05-01 provides a simple summary of the principle behind the endosymbiont theory.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=151#oembed-1

Video 05-01: Fundamentals of the endosymbiont theory.

A Summary of the Evidence for the Endosymbiont Theory

There are several compelling arguments for the idea that mitochondria and chloroplasts were once free-living bacteria. These are based on a combination of structural and genetic similarities to specific bacterial species that exist today. (See Figure 05-02 for a structural comparison of a chloroplast and its nearest bacterial ancestor.) It should also be noted that the mitochondria and chloroplasts are thought to have derived from *different* bacterial ancestors at different points in the evolutionary timeline.

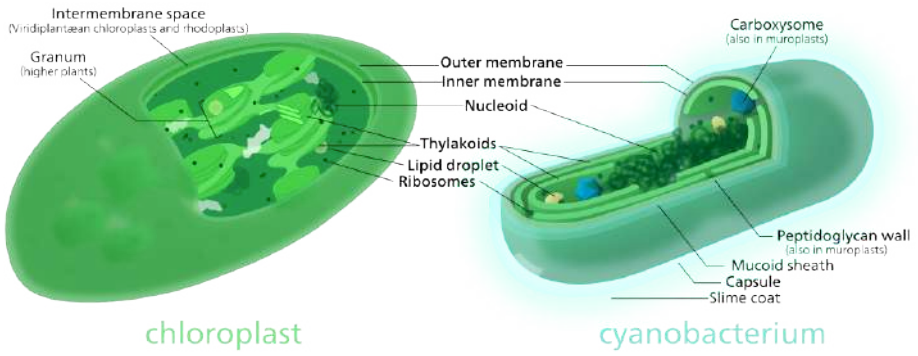


Figure 05-02: Structural comparison between a chloroplast from a flowering plant and present-day cyanobacterium, which is considered to be the closest living relative of the chloroplast. Many of the internal structures in the chloroplast, including thylakoid membranes and ribosomes, are also present in the cyanobacterium. They are also very similar in size (and both are considered to be much larger than mitochondria). The chloroplast also has a single circular chromosome of DNA, just like cyanobacteria; however, the chloroplast one is severely reduced in size and number of genes. "Chloroplast" by Kelvinsong shared under a CC BY-SA 3.0 license.

Below is a list of the most commonly mentioned similarities among mitochondria, chloroplasts, and their closest bacterial relatives, in no particular order:

- Just like bacteria, new mitochondria and chloroplasts can only be produced through **binary fission**.
    - Binary fission is an asexual reproductive process where one cell splits into two cells.
    - Because new chloroplasts and mitochondria can only be created through binary division, a cell's mitochondria or chloroplasts cannot be re-created if they are all lost or destroyed. For example, in some algae, such as *Euglena sp.*, chloroplasts are sometimes destroyed by exposure to specific chemicals or a prolonged absence of light. In such a case, the chloroplasts will not regenerate. However, as long as the cell has access to some kind of external carbohydrate source, they are able to survive and reproduce without them.
- There are several chemical components that are found in mitochondria and/or chloroplasts that are only found in bacteria and not in other parts of eukaryotes. For example,
    - A membrane lipid called *cardiolipin* is exclusively found in the inner mitochondrial membrane and bacterial cell membranes but not other eukaryotic membranes.
    - Proteins translated by mitochondria and chloroplasts use *N-formylmethionine* as the initiating amino acid, as do proteins created by bacteria. Eukaryotes use the unmodified form of methionine as the first amino acid.
- Transport proteins called **porins** are found in the outer membranes of mitochondria and chloroplasts and are also found in bacterial cell membranes.
- Mitochondria and chloroplasts most commonly carry circular DNA molecules that bear striking similarities to bacterial DNA (in both size and structure).
    - Genome comparisons suggest a close relationship between mitochondria and bacterium known as *Rickettsia sp*.
    - Chloroplast genome evolution appears to be slightly more complex (especially in the

algae); however, there are many similarities between the genomes of higher plant chloroplasts and those of cyanobacteria.

Genomic Similarities of Mitochondria and Chloroplasts to Bacteria

---

A Note on Units

Ribosome sizes are measured in *Svedberg [**S**] units*. This is a measure of density, rather than mass, related to their sedimentation rate when centrifuged.

In addition to the structural features described above, the genomes of mitochondria and chloroplasts also have many bacteria-like features as well as similarities in their protein-synthesizing systems. As mentioned, they have small circular DNA genomes, which are roughly comparable reduced versions of bacterial genomes. They synthesize some of the components of the electron transport chains used for oxidative phosphorylation (in mitochondria) and the light-harvesting photosynthetic systems (in chloroplasts). They also have ribosomes that are more similar to bacterial ones than they are to the ones in the cytosol of their own cell. Table 05-01 summarizes some of this information.

Table 05-01: Comparison of some of the properties of genomes and the protein-synthesizing systems of bacteria, mitochondria, and chloroplasts as compared with the eukaryotic system

SOME PROPERTIES OF BACTERIA, MITOCHONDRIA, AND CHLOROPLASTS AS COMPARED TO EUKARYOTIC SYSTEMS

Trait	*Gram-negative bacteria*	*Mitochondria*	*Chloroplasts*	*Eukaryotes*
Ribosomal RNAs	23S, 16S	23S or smaller, 16S or smaller	23S or smaller, 16S or smaller	28S, 5.8S, 18S
Ribosomes	70S total 50S, 30S subunits	70S or smaller	70S or smaller	80S total, 60S, 40S subunits
Genome	Circular DNA, no centromere	Circular DNA, no centromere	Circular DNA, no centromere	Linear DNA with telomeres and centromeres
Number of genes	~5000	5–35 genes*	~120	15,000–100,000
Translation inhibitors	Chloramphenicol, streptomycin	Chloramphenicol, streptomycin	Chloramphenicol, streptomycin	Cycloheximide

*varies among eukaryotic groups

The Mitochondrial Genome

Mitochondrial genomes are usually very small. The human mitochondrial genome is one of the smallest known, at only 16.8 kbp and 13 protein coding genes (plus some tRNA and rRNA genes). Brewer's yeast (*Saccharomyces cerevisiae*) have a somewhat larger mitochondrial genome, at about 86 kbp and 100 protein coding genes. The model flowering plant, *Arabidopsis thaliana*, has a 367 kbp mitochondrial genome, which is more than three times the size but still only includes 86 genes.

By comparison, the genome of modern-day *Rickettsia prowazekii*, a bacteria that is considered to be the closest living relative of mitochondria, is around 1100 kbp and includes 834 genes. Despite the difference in the size between the genomes of mitochondria and *R. prowazeckii*, they are still considered to be extremely similar overall. As an example, they are both somewhat unique, compared to other bacteria, in that they lack genes for glycolysis. As a result, in eukaryotic cells, glycolysis happens only in the cytosol of the eukaryotic cell, and not in the mitochondria.

Some examples of the genes included on mitochondrial genomes are the following:

- specific components of the transcription/translation machinery (tRNAs, ribosomal proteins, and rRNA)
- specific components of the electron transport chain
- specific components of the ATP synthase

In general, the proteins encoded tend to be the hydrophobic subunits of membrane proteins. The proteins are translated on mitochondrial ribosomes and remain in the mitochondrion. Other mitochondrial proteins, no longer synthesized within the mitochondrion, will need to be transcribed in the nucleus, translated in the cytosol of the eukaryote, and then imported into the mitochondrion prior to function. However, that is a topic we will tackle a little later in this chapter.

Of course, there is some diversity among Eukaryotes as to what exactly is contained within their respective mitochondrial genomes. Organisms within the same kingdom tend to be more similar to each other overall. Most mitochondrial genomes are circular, especially within the animals, but there are some that are linear.

The Chloroplast Genome

Chloroplast DNA also forms a circular genome and encodes for a variety of important chloroplast genes, including the following:

- specific components of the transcription/translation machinery (tRNAs, ribosomal proteins, and rRNA)
- specific components of the ATP synthase
- the large subunit of RuBisCO, the key enzyme involved in carbon fixation
- roughly 20 components of the chloroplast electron transport chain, used for photosynthesis

The size of the chloroplast genome is typically in the range of 120 to 160 kb. There seems to be much less variability in the size of the chloroplast genome from species to species compared to the genomes of the mitochondria. This point is interesting, as the origins of the chloroplast appear to be much more complex. Considering that all cells that have a chloroplast also have mitochondria, it is clear that chloroplast endosymbiosis came after the events that resulted in the development of the mitochondrion. However, it is also clear that chloroplasts were acquired by additional means after the initial **primary endosymbiosis** event. **Secondary endosymbiosis** is when a cell engulfs a chloroplast-containing cell, such as a red or green algae, which then will develop into that organism's chloroplast. Secondary endosymbiosis results in some unique features in the chloroplast, which can be identified. For example, a "primary chloroplast" will have two membranes that surround it; a "secondary chloroplast" is expected to have three or more membranes. It may also have other

"leftover" structures in between its membranes that may represent parts of the engulfed cell. Secondary endosymbiosis is virtually unknown in land plants but is quite common in protists, and red and green algae.

What's Missing from Endosymbiont Organelle Genomes?

When we consider the size difference between the genomes of mitochondria and chloroplasts compared to their bacterial relatives, there is a question that arises and must be considered: *What happened to the rest of the genes required for organellar function?*

There are two possible answers based on the evidence we currently have:

1. Some genes were probably lost over time, as their functions were no longer essential to survival in the intracellular environment. Alternately, these functions were essential but were already being carried out by the host cell, so the bacterial version was no longer essential. Genes coding for proteins involved in defense would fit into this category, for example.
2. Other genes from mitochondrial or chloroplast genomes have, over the course of time, been transferred to the eukaryotic nucleus and incorporated into the nuclear genome.

The net result is that the majority of mitochondrial or chloroplast proteins are now encoded by nuclear genes. In the case of mitochondria, only a few highly hydrophobic protein subunits are encoded by mitochondrial genes and synthesized on ribosomes in the mitochondria. This means that almost all proteins in mitochondria or chloroplasts require the same kind of targeting system that we have seen for the nucleus and endomembrane system. The mitochondrial proteins encoded in the nuclear genome will be synthesized on free ribosomes in the cytosol and are then targeted to mitochondria or chloroplasts. We will end this topic by exploring protein import into these organelles.

Structural Features of Modern Mitochondria and Chloroplasts

While the current consensus, based on the evidence described above, is that mitochondria and chloroplasts have bacterial origins, it's also quite clear that they have evolved a great deal since then. The structure of both organelles is reminiscent of their bacterial ancestors, but it is not identical. Figure 05-03, below, highlights the structural characteristics of each of these organelles.

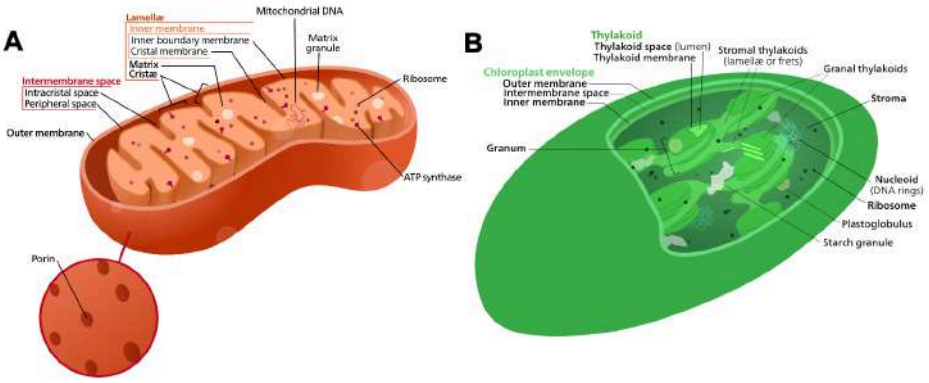


Figure 05-03: Structure of the mitochondrion (A) and the chloroplast (B) of modern Eukaryotes. Both images (A and B) were originally created by Kelvinsong and are shared under a CC BY-SA 3.0 license.

It's worth taking a moment to study each of these images and to refer back to them as needed. In the rest of the chapter, we will be using much of this terminology to describe the location of the proteins and processes of each organelle, so it is important that you are familiar with them. Then in Topic 5.3, we will do a deeper dive into the structure of each of these organelles and how they function.

Protein Import into Mitochondria and Chloroplasts

The import of proteins into the mitochondria and chloroplasts follows the same general pattern that we've already seen in both the nucleus (Chapter 3) and the endomembrane system (Chapter 4). To remind you,

1. The newly translated protein contains a targeting sequence, which identifies that it needs to be relocated.
2. The targeting sequence is recognized by some kind of receptor that helps move the protein to the organelle in question.
3. The protein enters the organelle, usually through some kind of translocation channel or pore.

Considering the structure of mitochondria and chloroplasts, targeting to these organelles can be quite complex. For example, a soluble protein that is destined for the matrix of the mitochondria will require one set of targeting sequences, whereas a protein that is destined to be embedded in the membrane of the cristae will require a different set of targeting sequences. In the chloroplast, there's a third membrane, called the thylakoid membrane. Consequently, targeting to the chloroplast is even more complex than mitochondrial targeting, since there are more membranes, which results in more compartments. It would be impossible for us to cover all of this complexity in a textbook of this kind, so if this is a topic that you are interested in, here are some relatively recent review articles for you to explore: this one is on mitochondrial targeting (Wiedemann & Pfanner, 2017), and this one is about chloroplast targeting (Bölter & Soll, 2016).

There are only two import scenarios that we are going to consider in this section:

1. Nucleus-encoded proteins that are destined for the mitochondrial matrix
2. Nucleus-encoded proteins that are destined for the chloroplast stroma (with some bonus information about how proteins then go to the thylakoid lumen)

Targeting to the Mitochondrial Matrix

Our discussion begins with the targeting sequence used to identify that a protein is destined for the mitochondrial matrix. Like the other targeting sequences we have seen before, it is an *amino acid sequence*, which is included in the primary sequence of the protein. Key features of the targeting sequence include the following:

- It is *always* at the N-terminus of the protein so that it exits the ribosome first during translation.
- It has a very specific amino acid sequence (which we are not going to expect you to know).
- It shapes itself into an alpha helix after it is translated.
- The alpha helix acts exactly like a 3D key, which fits perfectly into a binding pocket on the import receptor protein.

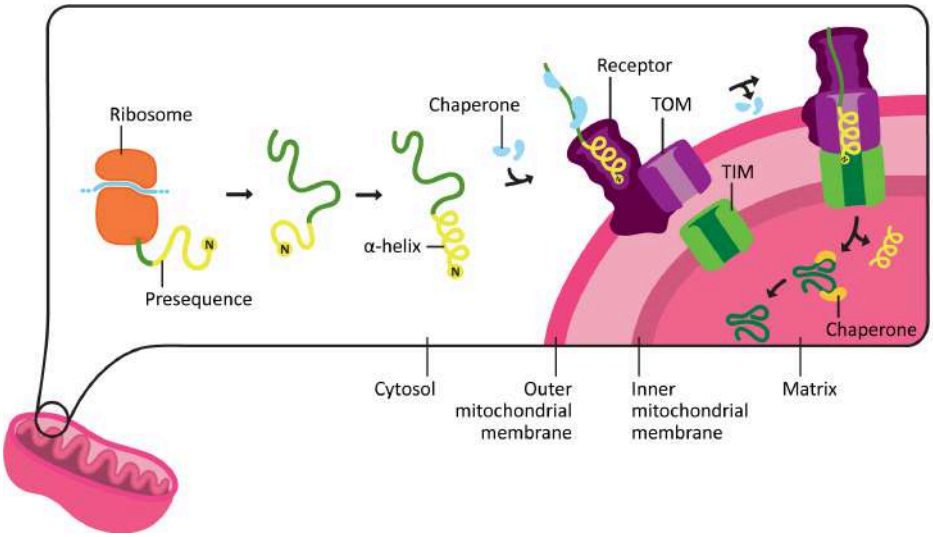


Figure 05-04: Protein targeting and import into the mitochondria. Proteins destined for the mitochondrion are synthesized in the cytosol. The presequence folds into an alpha-helical shape and is recognized by the receptor. The protein is transported through TOM and TIM to reach the mitochondrial matrix, where the presequence is cleaved off the protein. Chaperones aid in this process in various locations to help with folding and unfolding as required. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The import of mitochondrial matrix proteins happens roughly as follows:

1. The protein containing the mitochondrial targeting sequence, known as a **presequence**, is synthesized entirely in the cytoplasm.
    - The presequence is always located at the N-terminus of the protein and contains roughly 15–50 amino acids. This sequence is quite specific to make the correct shape once it is translated.
    - As it emerges from the ribosome during translation, the presequence coils into an *amphipathic alpha helix.* This is an alpha helix that has one side that is polar, and the other is nonpolar. The shape and chemical properties of the presequence acts like a 3D

key.

2. The presequence binds to a receptor on the outer mitochondrial membrane. Chaperones help maintain the protein in an unfolded state until it can be transported into the mitochondria
3. The protein is moved across the outer membrane, and then the inner membrane, through a pair of translocation channels known as **TOM (transporter outer mitochondrial membrane)** and **TIM (transporter inner mitochondrial membrane)**.
    - This process requires energy. A combination of ATP hydrolysis and harnessing the energy already stored in the proton gradient across the inner mitochondrial membrane is used to drive protein import. (We'll talk about the proton gradient more in the next topic.)
4. Chaperones assist with keeping the protein unfolded on the outside of the mitochondria, pulling the protein through the channel, and then folding the protein on the inside of the mitochondria.
5. Once the protein is inside the matrix of the mitochondria, the presequence is *removed* by a proteolytic enzyme known as the **mitochondrial processing peptidase**, or MPP. After the presequence is removed, the protein is further processed and folded into its final mature shape.

Targeting to the Chloroplast Stroma

Chloroplast targeting is very similar to mitochondrial targeting. However, there are some unique features that are worth pointing out (Figure 05-04). For example,

- The chloroplast targeting sequence, known as the **transit peptide**, is also at the N-terminus of the protein, like the one for the mitochondrion. However, *it's not the same sequence*. Plants must send different proteins to their mitochondria and chloroplasts, which means that the targeting sequences for these organelles also need to be different.
    - The sequence of the transit peptide is highly specific, just like we saw in the mitochondria, but may also provide additional locational information to identify exactly *where* in the chloroplast the new protein is headed. For example, in Figure 05-04, we see that the transit peptide is immediately followed by a sequence that indicates the protein belongs in the thylakoid lumen.
    - Like the mitochondrial presequence, it has a region that forms an alpha helix, which can act as a 3D key. Unlike the presequence, there is also a region that is unstructured whose function is not entirely clear but is likely involved in fine-tuning the targeting so that proteins can be sent to specific regions of the chloroplast.
- The transit peptide is recognized by a chloroplast import receptor on the surface of the organelle. Chloroplast proteins headed for the stroma are fed across two translocation channels called **TOC (transporter outer chloroplast membrane)** and **TIC (transporter inner chloroplast membrane)**. This is, once again, an energy-requiring pathway, and ATP is consumed.

- Proteins destined for the thylakoid lumen will end up in the stroma at first (Figure 05-04). However, a second sequence is revealed at the N-terminus when the initial chloroplast targeting sequence is cleaved, which then directs the protein to the thylakoid lumen. We won't be discussing this further, but it is very cool.

As you may have already guessed, the chloroplast transit peptide is also cleaved once the protein has entered the stroma. In this case, the protein responsible for the removal of the transit peptide is called the **stromal processing peptidase**, or SPP.

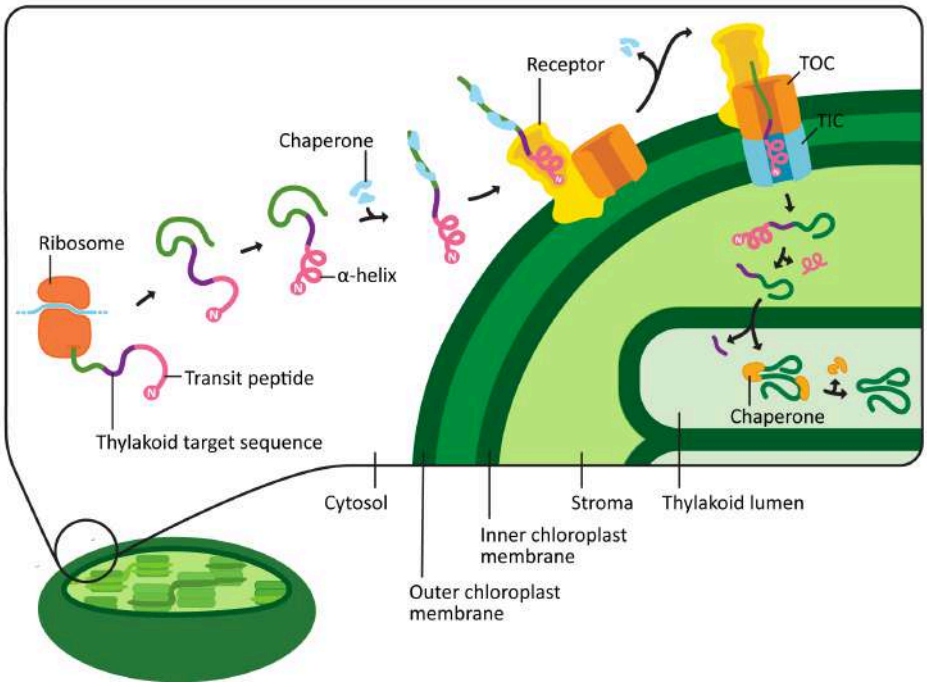


Figure 05-05: Protein targeting and import into the chloroplast. Proteins destined for the chloroplast are synthesized in the cytosol. The transit peptide folds into an alpha-helical shape and is recognized by the receptor connected to TOC. The protein is transported through TOC and TIC to reach the chloroplast stroma, where the transit peptide is cleaved off the protein. Chaperones aid in this process in various locations to help with folding and unfolding as required. An additional targeting sequence is present for proteins that are destined for the thylakoid lumen. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

TOPIC 5.2: FUNCTION OF MITOCHONDRIA AND CHLOROPLASTS ATP PRODUCERS WITH DIFFERING END GOALS

Learning Goals

- Compare and contrast chemiosmotic coupling of proton pumping and ATP formation in mitochondria and chloroplasts, and relate this to the ultimate functional goal of each organelle.

- Discuss the relationship between mitochondria and chloroplasts in all cells.

Introduction

As we have already explored in the previous topic, mitochondria and chloroplasts are thought to have evolved from bacterial ancestors and as such have some characteristics that are more aligned with bacteria than with eukaryotes. One of these similarities is how they produce ATP. However, the *purpose of producing ATP in each of these organelles is completely different.* This is a key point in understanding these organelles.

- The most important purpose of the mitochondria with respect to energy production is *to produce ATP for the rest of the cell*. Almost all known Eukaryotes have mitochondria, including photosynthetic plants and algae.
- On the other hand, the primary purpose of the chloroplast is *to produce sugars via photosynthesis*. While it also produces ATP within the organelle, *all* of the ATP produced in the chloroplast is used to build sugars (this is an energy-intensive process!). In fact, ATP made in the chloroplast is unable to leave the organelle, as it has no membrane transporters capable of ATP export.

The difference in function is extremely important to understanding these two organelles. In this section, we will first look at what's similar in these organelles—namely, how they produce ATP. In the last topic, we will follow up by looking at the unique features of each organelle.

ATP Production in Mitochondria and Chloroplasts Happens via Chemiosmotic Coupling

As you likely know already, ATP is a vitally important biological molecule that is used to provide the chemical energy required to drive the chemical reactions catalyzed by our cells' enzymes. You should also know that the mechanisms for generating ATP are ancient. For the most part, they predate the evolution of Eukaryotes, and if Eukaryotes had not found a way to harness the ATP-synthesizing abilities of bacteria and/or Archaea, they would not have survived.

There are three known ways of producing ATP, which fall into two major categories:

1. **Substrate-level phosphorylation**. This is when a single enzyme is able to produce ATP by swapping an unstable phosphate from the enzyme's substrate to ADP to produce ATP. Glycolysis produces all of its ATP via this method.
2. **Chemiosmotic coupling** (Figure 05-06). In chemiosmotic coupling, a chemical gradient is generated by one set of membrane-bound protein. The gradient is used to help drive ATP production with a different set of membrane-bound proteins. In this way, these reactions are said to be *coupled* to produce ATP. This means that *both* are required for the system to function, and if either one is shut down for any reason, the other one will also stop functioning. In this kind of ATP generation, we see several membrane-bound protein complexes working together to produce ATP at a much higher rate than is possible via substrate-level phosphorylation. There are two main types that exist:

a. **Oxidative phosphorylation**, done by the mitochondria. In this type, the oxidation of complex energy storage molecules (i.e., sugars, lipids, etc.) from food is used to provide energy to produce a proton gradient, which, in turn, is used to drive the synthesis of ATP.

b. **Photophosphorylation**, done by the chloroplast. The premise is almost identical to oxidative phosphorylation, except that in the case of photophosphorylation, the source of energy that produces the proton gradient is the sun.

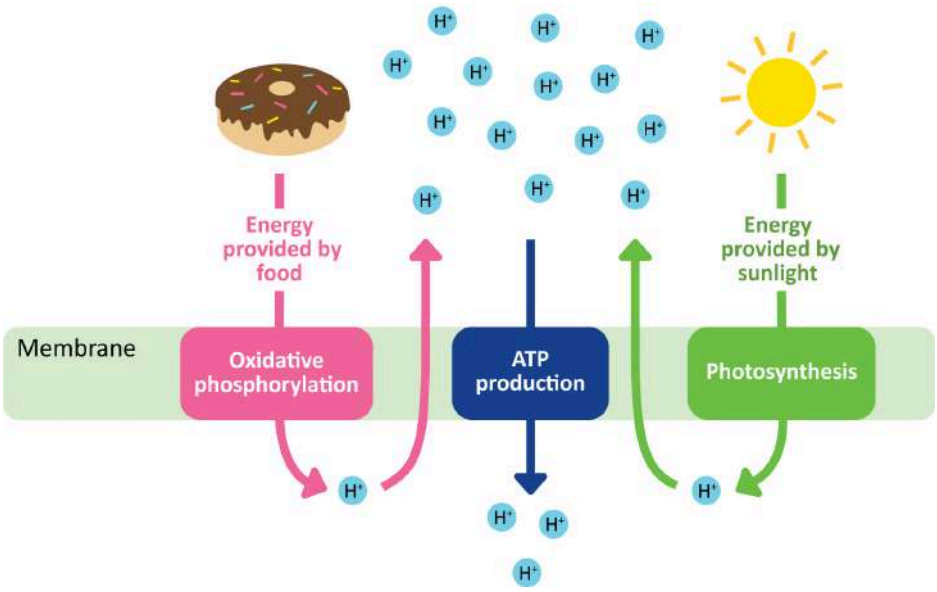


Figure 05-06: Chemiosmotic coupling uses one set of proteins to build a proton gradient and another set of proteins to harness the energy in the proton gradient to produce ATP. This is true in both oxidative phosphorylation (which happens in the mitochondrion) and photophosphorylation (which happens in the chloroplast). This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

How Chemiosmotic Coupling Works

Substrate-level phosphorylation, which is used to make ATP in glycolysis, does not produce enough ATP to meet the energetic needs of most complex organisms (like Eukaryotes). We know this, in part, because there are no known multicellular organisms that rely entirely on glycolysis for their energetic needs. Chemiosmotic coupling is a process that is very ancient and widespread in the tree of life. Many of the proteins involved are highly conserved across kingdoms.

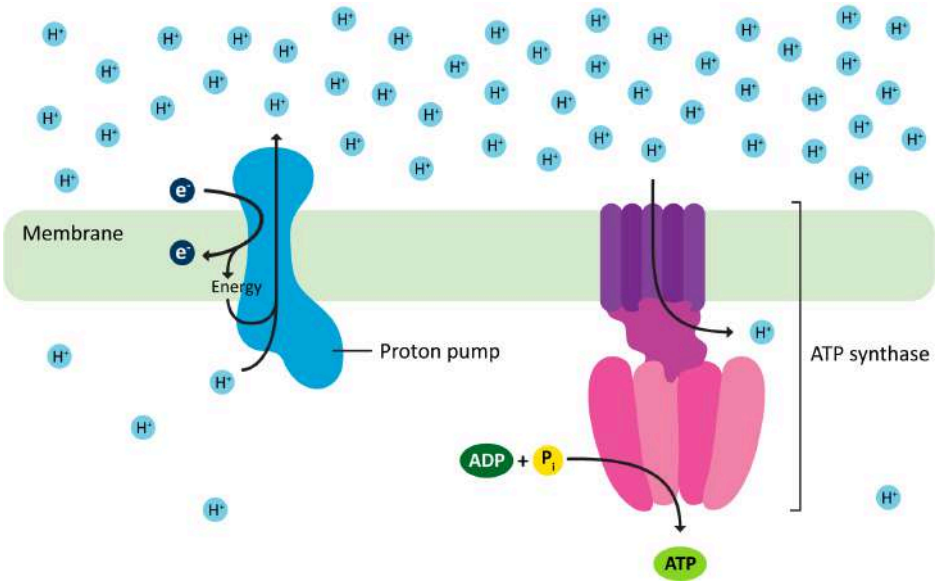


Figure 05-07: Making ATP via chemiosmotic coupling is a two-step process. Protons are pumped to one side of the membrane. This creates a proton gradient. Then protons flow through the ATP synthase and are the driver of ATP production. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The key idea is very simple. It is a two-step process, which is visualized in Figure 05-07:

1. A **proton pump** uses some kind of energy source to actively push protons (H^+) across a membrane against their gradient. This produces an **electrochemical gradient** (meaning that there is both a chemical and charge difference across the membrane).
    - The proton pumps of mitochondria and chloroplasts are each quite complex and consist of several large protein complexes that work together. Due to how they manage to build the proton gradient, we call them **electron transport chains (ETCs)**. Again, while the premise here is similar, *these ETCs are not the same*.
    - In mitochondria, the high-energy products of the **citric acid cycle** are fed into the ETC to drive the pumping of protons across the inner membrane of the mitochondria, whereas in the chloroplast, it is the energy of the sun, captured by chlorophyll, that helps drive the pumping of protons by the chloroplast ETC.
    - The energy that was used to produce the electrochemical gradient is now stored within it as potential energy. We can calculate the energy stored in the gradient by measuring the pH on either side of the membrane and using that to convert to the amount of energy stored. (*If you take biochemistry in the future, you should expect to learn how to do this calculation.*)
2. The energy contained in this electrochemical gradient is used to drive the phosphorylation of ADP to produce ATP. This occurs as protons move through a channel across the membrane that is part of the **ATP synthase** complex.

We call this particular usage of hydrogen ions, in which they are actively being transported against their gradient first so that they can then flow back down their gradient to do work, the **proton-motive force**. It is a crucial part of ATP production in both mitochondria and chloroplasts.

The ATP Synthase

The ATP synthase is key to the production of ATP in both the mitochondrion and the chloroplast. The ATP synthase is a very large protein complex that is embedded in either the inner mitochondrial membrane or the thylakoid membrane, depending on the organelle (Figure 05-08 and Video 05-02). This is the same enzyme used by aerobic bacteria to produce ATP. Versions of this protein are also found in other membranes, like the endomembrane system. However, outside of the mitochondria, they are not used for ATP synthesis but instead break down ATP to actively pump protons across the membrane to create the acidic environment inside of endosomes and lysosomes.

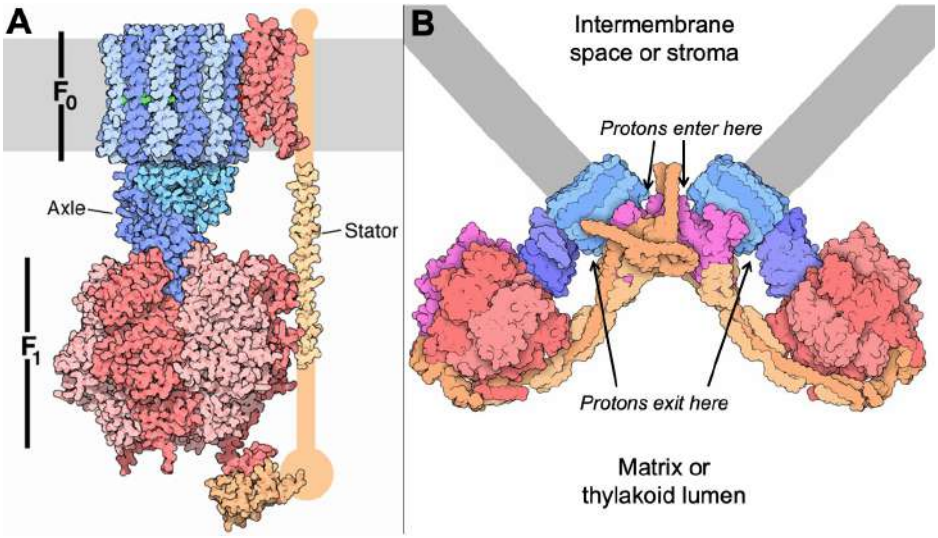


Figure 05-08: Molecular structure of the ATP synthase. (A) The ATP synthase has several structural components. The base, which is used to generate the energy that will be harnessed in ATP, is called the F_0 subunit, whereas the round head that actually makes the ATP is called the F_1 subunit. An axle connects the F_0 to the F_1 and helps move the heads of the F_1 so it can build the ATP. The F_1 subunit is held onto the axle by a long protein called the stator. It stabilizes while also allowing the movement required for ATP production. (B) In the membrane, two ATP synthases dimerize, which causes a kink in the membrane and is thought to contribute to the characteristic folded shape of both cristae and thylakoids. Protons flow through the F_0 subunit, which turns the axle and results in the production of ATP. It is said that this works very similarly to a water wheel. Original images by David S. Goodsell, shared under a CC BY 4.0 license. Image compiled by Dr. Robin Young.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=151#oembed-2

Video 05-02: Molecular representation of ATP synthesis.

The ATP synthase is a truly unique molecular machine. There are very few biological examples of proteins that rotate in this way to create energy. And yet this is a highly effective and elegant way

to generate the energy required to create ATP. The ATP synthase, and the electron transport chain that creates the proton gradient that powers it, is one of the most important factors in the success of Eukaryotes.

Chemiosmotic Coupling in Mitochondria and Chloroplasts Is Arranged Differently

It is important to note the differences in how the proton pumps and ATP synthase are arranged in each of the organelles. Part of the reason that they are this way is due to the bacteria that they evolved from. (Remember from the previous topic that the two organelles are thought to come from separate endosymbiotic events.) This is illustrated in Figure 05-09.

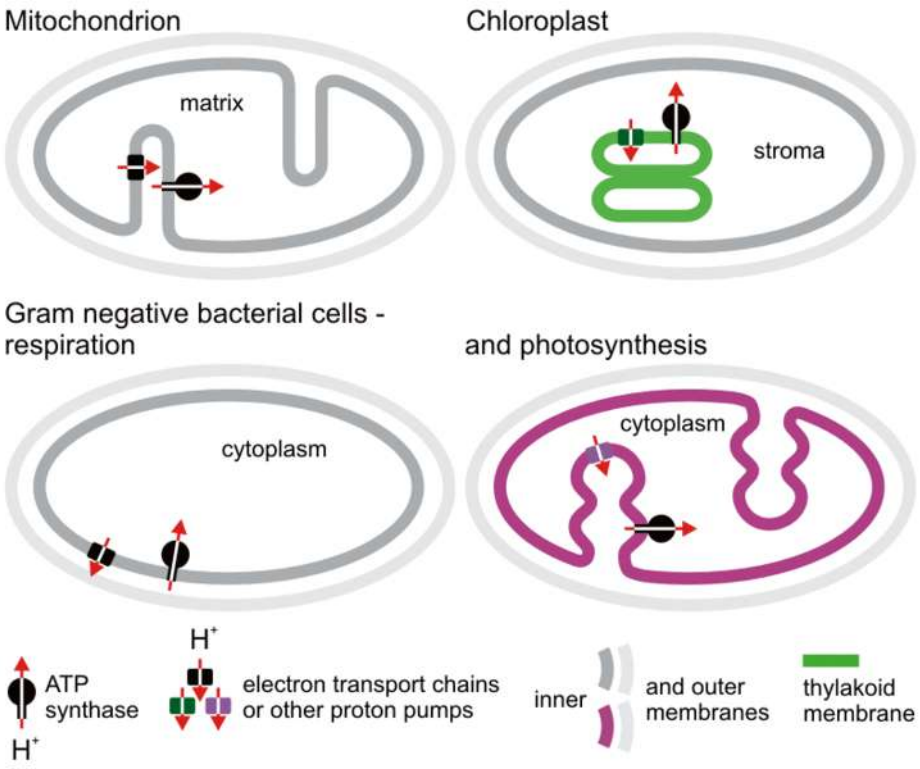


Figure 05-09: Illustrated comparison of the arrangement of the ETC and ATP synthase of mitochondria (left column, top), chloroplasts (right column, top), and their closest living bacterial relatives (bottom). Image by Darekk2 is shared under a CC BY-SA 3.0 license.

It's important to point out a few things from Figure 05-09 that highlight key differences between mitochondria and chloroplasts. Most notably, mitochondria evolved from "simpler" bacteria than chloroplasts did. We see this, in part, in the arrangement of the internal membrane structures in each organelle. The chloroplast ancestor already had inward projections of its membranes, which eventually evolved into the thylakoid membrane. As such, chloroplasts have an additional set of internal membranes that are not present in mitochondria. This has resulted in differences in how the two organelles function:

- In *mitochondria*, the electrochemical gradient is built across the inner mitochondrial membrane (also known as the cristae). The difference in pH across that inner membrane where the ETC and ATP synthase reside is quite small (sometimes less than a pH point). This is because of the following:

- The gradient is produced by pumping the protons *out of* the organelle, so the matrix pH goes up and becomes more basic.
- The exterior membrane is porous (due to **porins** that allow free passage of water and solutes; see Figure 05-10 later in this chapter). So the protons will diffuse out of the intermembrane space into the cytosol over time by way of the porins. This means that even though the ETC is continuously pumping protons into the intermembrane space, there is no real accumulation on ions in the exterior compartment. So the pH of the intermembrane space remains roughly the same as the cytosolic pH.
- Despite this, the small change in pH creates a large enough difference in electrical potential that the ATP synthase can function.

- In *chloroplasts*, the thylakoid membrane is the site of the ETC and the ATP synthase.
  - The protons are pumped into the thylakoid lumen. As such, the pH in that compartment is much lower compared to the stroma. Since the thylakoid lumen is an enclosed space (unlike the intermembrane space of the mitochondria, which is connected to the cytosol), a much more significant pH difference can be created in the chloroplast than the one produced by the mitochondria.
  - Interestingly, the thylakoid membrane is highly permeable to Mg^{2+} and Cl^{-} ions, which pass freely and are in abundance. This effectively destroys the electrical component of this gradient. So virtually all of the energy stored in this gradient comes from the difference in chemical concentration of protons across this membrane.

How Do the Mitochondrion and Chloroplast Work Together?

When undergraduate students are asked to learn about photosynthesis, a common question is often overheard: *"OK, great. Photosynthesis gave us enough oxygen in our atmosphere, but what have plants done for us **lately**? I mean, I don't even have a chloroplast, so why should I care about it?"* Here we present some arguments for you to consider. Hopefully it will convince you that understanding plants and photosynthesis is more important than you may realize.

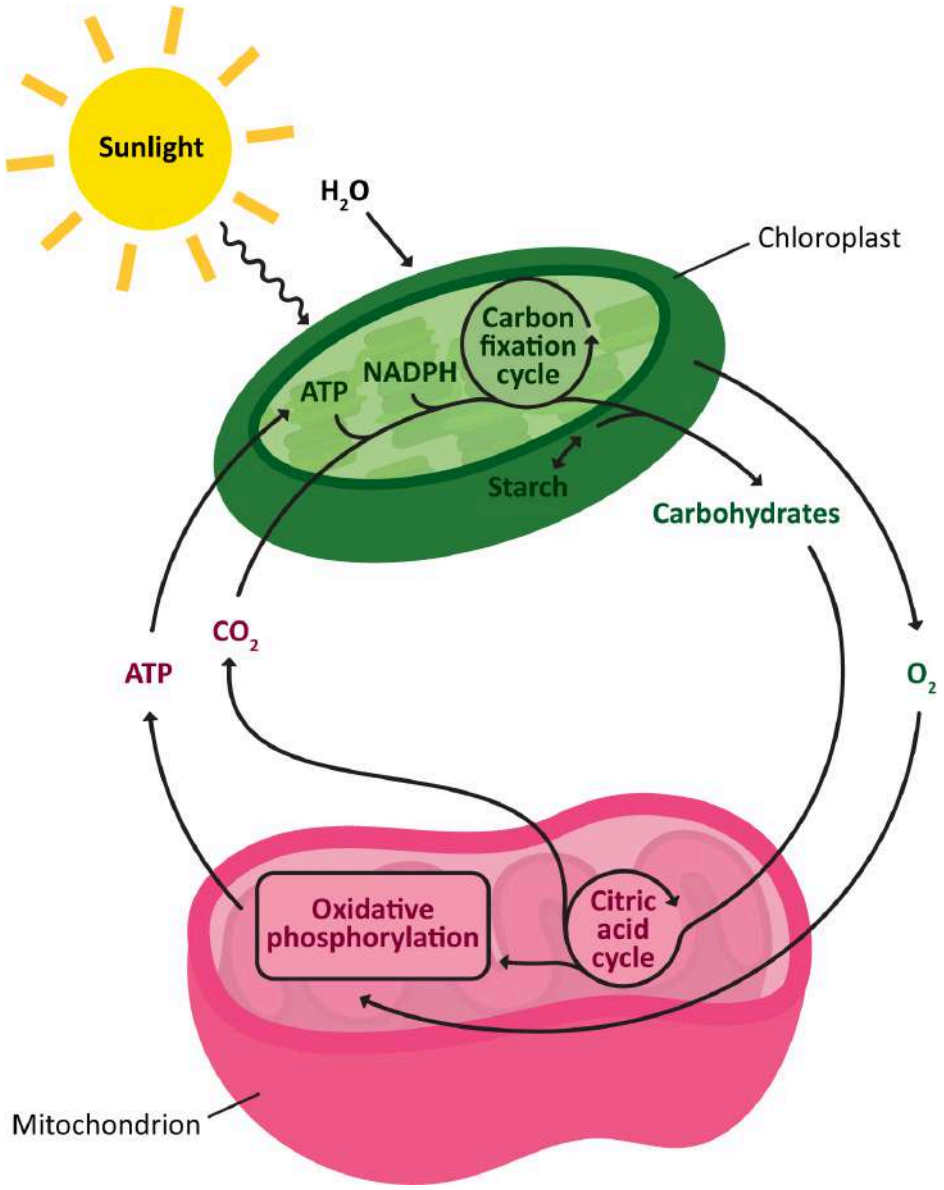


Figure 05-10: The cyclical relationship between mitochondria and chloroplasts. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

First and foremost, regardless of whether you have your own chloroplast or not, mitochondria and chloroplasts *work together*. Each one is required by the other in order to keep the energy (and carbon) flowing through the biosphere. There is a cycle that forms between the chloroplasts of plants and algae and the mitochondria of all living organisms (including those of plants and algae).

- The chloroplast uses carbon dioxide and water to produce oxygen and carbohydrates.
- The oxygen and carbohydrates are then used by the mitochondria, energy is released, carbon dioxide and water are produced, and the cycle continues.

On top of that, even though *you* don't have a chloroplast, you absolutely do require its functions in order to survive. Plants are both **autotrophs** (meaning that they are capable of making everything they need to survive) and **primary producers** (meaning that they sit at the bottom of the food chain, and everything else is dependent on them). We, on the other hand, are **heterotrophs** (meaning that

we cannot produce everything we need to survive) and are at the top of our food chain (meaning very few things prey on us for food). Without the primary producers, the *entire* food chain will fall apart. Without an apex predator like ourselves? Most likely a different apex predator will take our place, and no one will notice we're gone. All of this is to say that we need plants and their ability to photosynthesize far more than they need us.

It is for this reason that cell biology textbooks such as this spend time teaching you about an organelle that you don't even have…but wouldn't it be cool if you did?

TOPIC 5.3: STRUCTURE-FUNCTION RELATIONSHIPS IN MITOCHONDRIA AND CHLOROPLASTS

Learning Goals

- Identify mitochondria and chloroplasts in micrographs and differentiate them from other organelles and from each other.
- Explain how the structure of the mitochondrion contributes to its function as the primary source of ATP in the cell.
- Explain how the structure of the chloroplast contributes to its function as the primary source of sugars for plant cells and the primary site of carbon fixation on the planet.
- Compare and contrast the role of chloroplasts with the role of mitochondria.
- List additional functions of the mitochondrion and the chloroplast.

Topics 5.1 and 5.2 mostly focused on how the mitochondrion and chloroplast are similar. We explored their evolutionary origins, their protein targeting and import, and how they produce ATP. However, at this point in the chapter, it is clear that we also need to spend some time considering what sets each of these organelles apart from the other. We will look at the mitochondria first, followed by the chloroplast.

Mitochondrial Structure

It is unclear how the mitochondria became such an internet sensation, yet somehow it did. So it should come as no surprise to you that the mitochondria is the powerhouse of the cell. However, it's important that you now step back and consider what that actually means from a cellular perspective. The mitochondrion is the primary producer of chemical energy in the cell, mostly in the form of ATP. A single organelle is known as a **mitochondrion**, whereas many of this organelle are known as mitochondria.

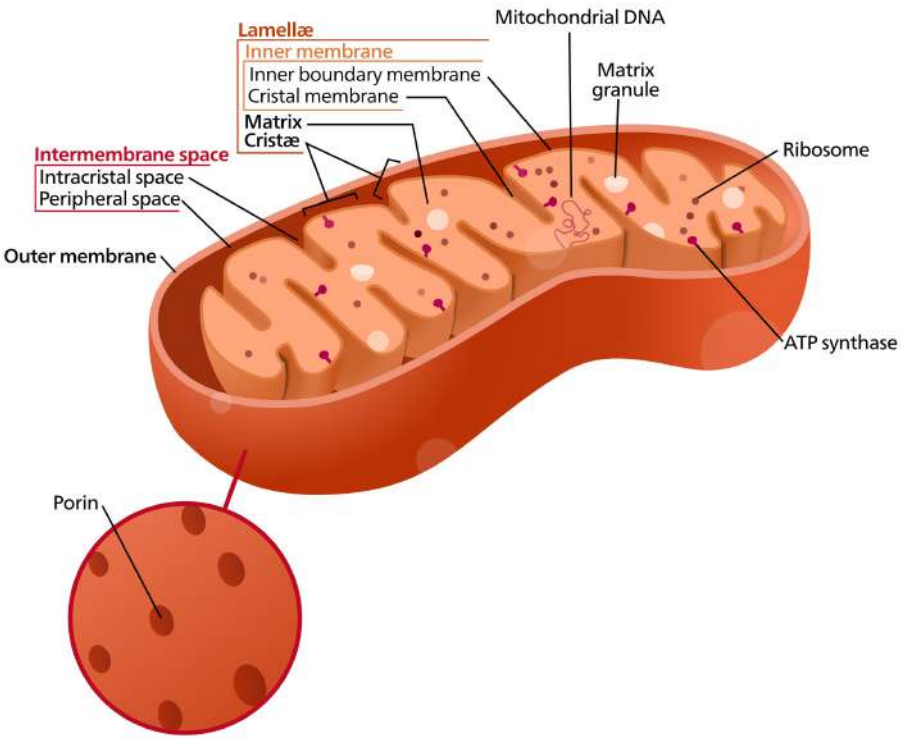


Figure 05-11: Labeled diagram showing the structure of a mitochondrion. Image by Kelvinsong is shared under a CC BY-SA 3.0 license.

The mitochondrion has two exterior membranes that surround it (Figure 05-10). The "cytosol" inside the innermost membrane of the mitochondrion is called the **matrix**. The original bacterial plasma membrane corresponds to the inner mitochondrial membrane. The inner membrane of mitochondria has infoldings or invaginations called **cristae** (crista is the singular form). For the most part, cristae look more or less like you've seen in textbooks; however, that does not have to be the case. They can also be tubular or fingerlike projections into the matrix. As we discussed previously, the curvature of the crista membrane is the result of the dimerization of the ATP synthase, which we discussed earlier (Figure 05-07 and associated text).

The bacterial origins of the mitochondrion, along with transmission electron microscopy (TEM) images showing cross sections of the organelle, tend to give a false impression of the shape of a mitochondrion. Most folks think of the mitochondrion as a small and discrete structure with a very static kidney bean shape. *This is not so.* In fact, mitochondria can be quite versatile and dynamic, constantly fusing and dividing in a cell. Newer data from live-cell imaging show that mitochondria are extremely dynamic and constantly move, fuse, and split apart. They form long, interconnected networks that can extend throughout the cytoplasm, as you can see in Video 05-03 below. (FYI, the mitochondria are the little yellow inchworms. The nuclei are the large cyan balls on the left-hand side. It's unclear what the red is from the information in this video.)

One or more interactive elements has been excluded from this version of the text. You can view them online

here: https://open.oregonstate.education/cellbiology/?p=151#oembed-3

Video 05-03: Live fluorescent microscopy of a cancerous cell showing the mitochondria in yellow, nuclei in cyan, and actin filaments in red.

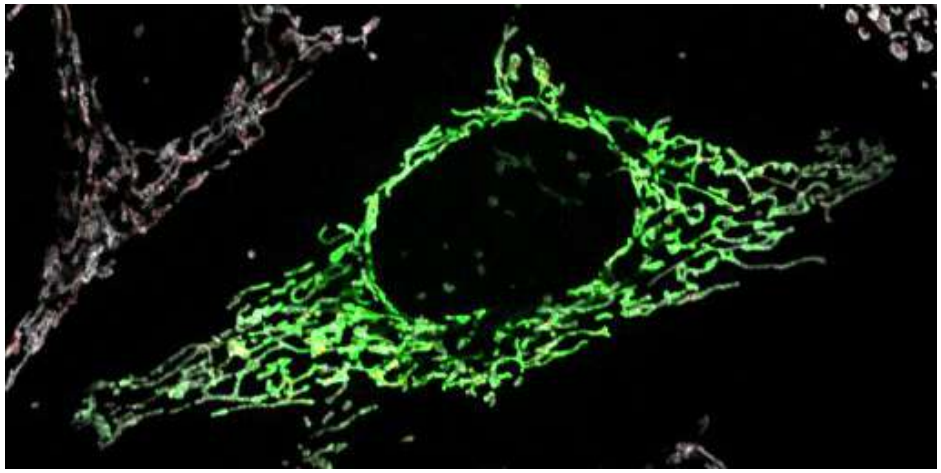


Figure 05-12: Fluorescence micrograph showing the distribution of the mitochondria in an isolated cell derived from the Henrietta Lacks (HeLa) cell line. Image by Seok Min Jin, Michael Lazarou, Chunxin Wang, Lesley A. Kane, Derek P. Narendra, and Richard J. Youle (2011) is housed in the Cell Image Library (CIL:13729) and shared under a CC BY-NC-SA 3.0 license.

What you should be able to see in both Figure 05-12 and Video 05-03 (above) is that the mitochondria are very different in reality than the caricature we have of them in our heads (and in the internet memes). The rise of fluorescence microscopy (and now superresolution microscopy) has really changed our understanding of the structure of these organelles as well as how their dynamic nature might contribute to their function.

Mitochondrial Function

The mitochondrion has several functions, not the least of which is the production of ATP for the rest of the cell. We will not discuss the details of how that happens here, as we've already said everything that we want to say about this in Topic 5.2. Here, we are mostly interested in the structure-function relationships within the mitochondrion. In other words, we expect you to know where all of the steps of ATP generation take place within this organelle. Figure 05-12, below, summarizes the main steps and locates them nicely within the organelle.

Cellular Respiration and Oxidative Phosphorylation

While we are not discussing the process of cellular respiration in all of its biochemical detail, we will start this section by reminding you of the major steps, which are also identified in Figure 05-12:

1. **Glycolysis** happens in the cytosol of the cell. It produces small amounts of ATP, but more importantly, it converts glucose into pyruvate, which is required for the next seps of the

process.

2. Pyruvate is imported into the mitochondrial matrix and is then converted into a molecule called Acetyl-CoA. In the process, some NADH is made.
3. Acetyl-CoA is fed into the **citric acid cycle** (also known as the Krebs cycle), which creates more NADH and also a similar molecule called FADH2. These are both important **electron carriers** and are key to providing the energy required for the active transport of protons across the electron transport chain.
4. The electron transport chain creates the proton gradient, which will then be used to synthesize many ATP, via the ATP synthase that was extensively covered in Topic 5.2.

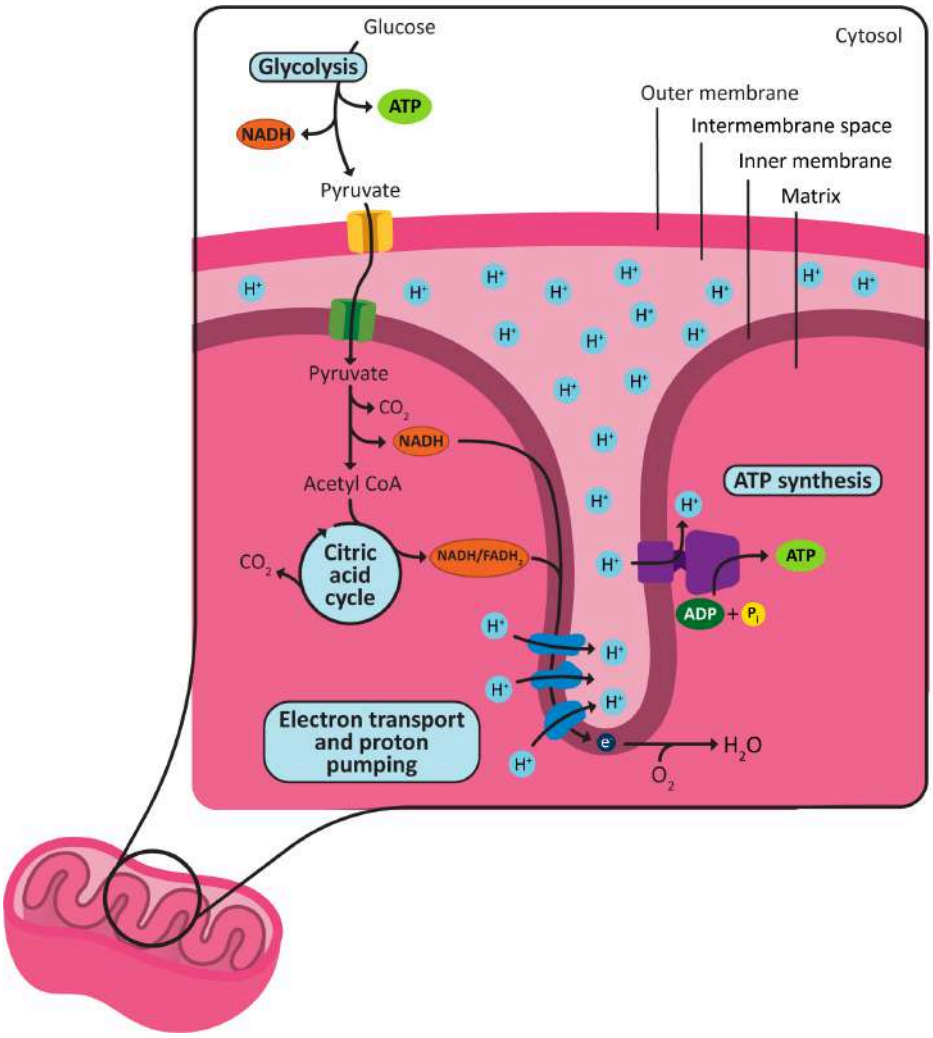


Figure 05-13: The location of the reactions of cellular respiration and the proteins that carry out these reactions. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

There are a couple important details you should make sure that you understand from Figure 05-13:

- The location of each of the metabolic steps of cellular respiration (glycolysis, Acetyl-CoA synthesis, the citric acid cycle, and the ETC and ATP synthesis), including which membrane they are attached to (if any).

- The areas of low and high pH based on the location of the protons in the drawing. Also note where the protons flow (from which structure to which structure) during the process of oxidative phosphorylation.

It is also worth noting where and how the chemiosmotic coupling reactions take place:

- For example, despite the fact that the ATP synthase is embedded in the cristae, we say that ATP production occurs *in* the matrix, as the active site is located on the matrix side of the membrane, and neither ATP nor its original reactants cross the membrane during the catalyzed reaction.
- On the other hand, we say the gradient is formed *across* the inner membrane, as the proteins that do the work transfer the proton from one side of the inner membrane/cristae to the other.

The video below provides some additional information, based on molecular data, about how the mitochondria may function. Not only does it show how the ETC and ATP synthase works, but it also starts with an animation of a protein being imported into the mitochondrial matrix via the TOM/TIM transporters.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=151#oembed-4

Video 05-04: Targeting of proteins to the mitochondria, oxidative phosphorylation, and ATP synthesis.

Other Functions of the Mitochondrion

In introductory textbooks such as this, it is very common to focus almost exclusively on the role the mitochondrion plays in cellular respiration, as this is obviously a very important function that is key to our survival. However, that is not the only function that is carried out by this organelle. It has a number of vital roles that it plays in cellular function. Some of these functions include the following:

- It can help regulate broader cellular processes due to its ability to control access to ATP and energy. Events such as when cell division will take place appear to be linked to a functional mitochondrion actively producing ATP via oxidative phosphorylation.
- It acts as a storage site for calcium ions. Calcium is an important signaling molecule (which will be discussed in Chapter 7). As a result, the mitochondrion plays a key role in several signaling pathways involving calcium.
- It generates heat. This is done at the expense of ATP production, as it runs the electron transport chain, but the ions are allowed to return to the matrix without harnessing the

energy in ATP. When the cell does this, we say that it has *uncoupled* the ETC from the ATP synthase. This is very useful for young infants, who experience more heat loss due to their small size. Running the ETC in this way happens in a type of fat known as brown fat, and it helps keep the baby warm. It is also used by many plants for early spring flowering (to break through the snow) and to attract pollinators.

- It helps with the production of steroids, such as cholesterol, testosterone, and other hormones.
- It is involved in the regulation of programmed cell death, a process in which a cell within a multicellular organism intentionally dies in response to specific cues. This is important in many cellular processes and also in getting rid of severely damaged cells. A good example is the loss of cells connecting the fingers during embryonic development in humans and many other animals. The "webbing" must all be removed, so the cells in this area intentionally die off, and the fingers are released from each other.

The Chloroplast

Chloroplasts are part of a larger family of organelles in plants known as **plastids** (Figure 05-13). There are several different types of plastids in plants, most of which have no photosynthetic activity. They are involved in a variety of functions, including storage; organ coloration; production of sugars, lipids, and nucleotides; and even gravity sensing in roots and shoots.

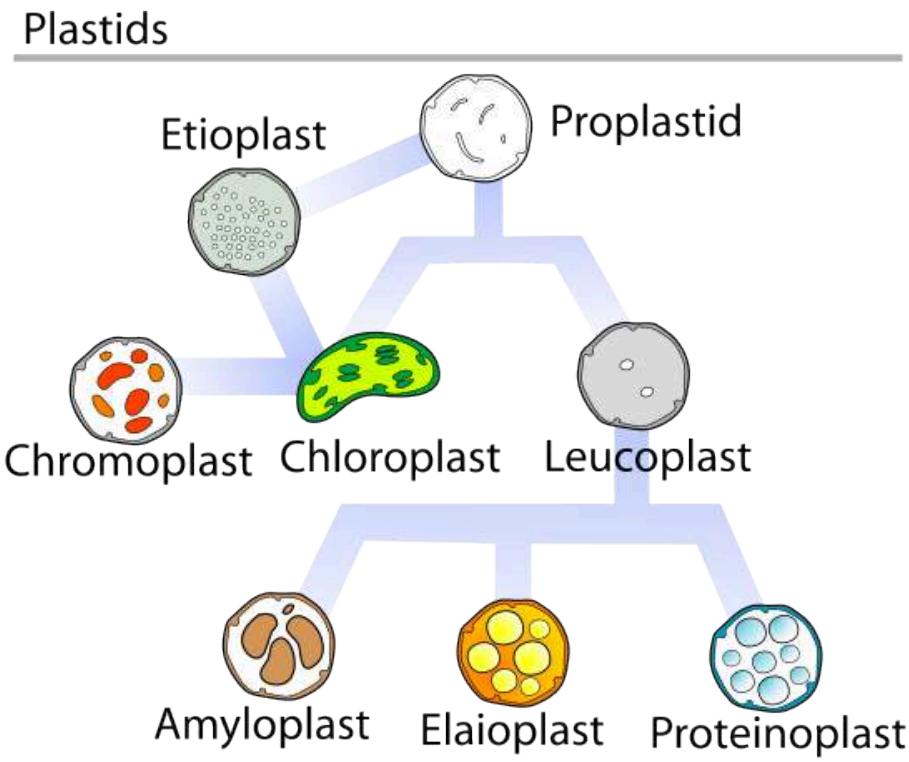


Figure 05-14: A diagram of the relationship among the various plastids in plants. In seeds, all plastids are in the form of a proplastid. As the seed germinates, the shoots and aboveground tissue will develop into etioplasts before they reach the light. Once they break through the soil and are exposed to light, the etioplasts convert into chloroplasts. Below-ground tissue, like roots, will develop into unpigmented leucoplasts, which may then convert into storage plastids such as amyloplasts (to store starch), elaioplasts (to store lipids), or proteinoplasts (to store protein). Chloroplasts can also convert into colored plastids known as chromoplasts. This process can be seen when fruits ripen and change color or when the fall leaves change color before they drop. This image by LadyofHats is in the public domain.

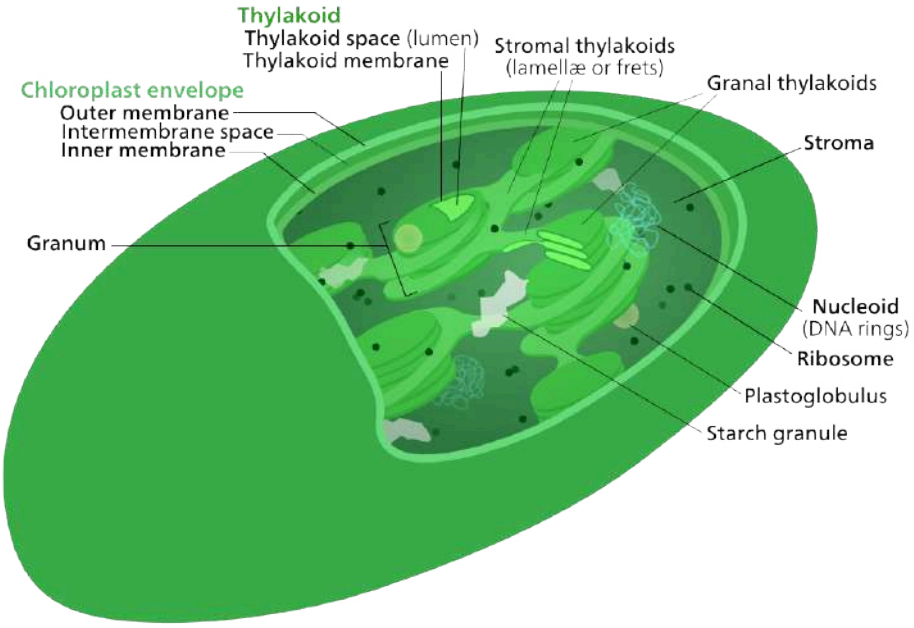


Figure 05-15: Labeled diagram showing the structure of the chloroplast. Image by Kelvinsong is shared under a CC BY-SA 3.0 license.

In many ways, chloroplasts have a similar structure to mitochondria in that they have an inner and an outer membrane (compare Figures 05-10 and 05-15). However, they also have a third internal membrane system not present in mitochondria, known as the **thylakoid membrane**. Inside this membrane is a compartment known as the **thylakoid lumen**. Figure 05-15 shows the chloroplast of a flowering plant, which we can tell because of the organization of the thylakoid into **grana** (that look like stacks of coins). Other photosynthetic organisms (most notably the many different kinds of algae) will have a slightly different organization of their flattened thylakoid sacs.

In addition to being the site of photosynthesis, resulting in the production of large amounts of sugars, chloroplasts have other roles as well. Some of these functions will result in the production of material that must be stored in the stroma of the thylakoid for a time. In Figure 05-15 above, you can see this in the **starch granules** (storing glucose) and **plastoglobules** (storing lipids and proteins) that can be observed in the **stroma** of the organelle.

Chloroplast Function

Photosynthesis and the Production of Sugars

Of course, the most famous function of the chloroplast is the process known as photosynthesis. In other words, the chloroplast is the site where light energy is captured and used to transform CO2 into carbohydrates. Much like our approach when discussing the mitochondrion, here we are going to focus on understanding how the structure of the chloroplast facilitates its ability to do photosynthesis.

Chloroplasts must maintain a process that

- captures light energy, which is then used to…
- produce a proton gradient to drive the production of ATP and NADPH, and then…
- fix atmospheric carbon dioxide, reduce it, and turn it into carbohydrates that the cell can use.

These processes occur in different parts of the chloroplast. The **light-dependent reactions** are the ones where the energy of the sun is captured and then converted into chemical energy the chloroplast can use. The light reactions include the chloroplast electron transport chain and the ATP synthase. Just like in the mitochondrion, both are membrane bound. In the chloroplast, both are embedded in the thylakoid membrane and pump protons back and forth from the stroma, to the thylakoid lumen, and then back again. The chemical energy produced by the light-dependent reactions (i.e., ATP and another electron carrier known as NADPH) will be fed into the next part of the process, called the **carbon-fixation reactions** (also known as the light-independent reactions, or the Calvin cycle). Most of the enzymes involved in carbon fixation are soluble and pack the stroma. The product of carbon fixation is sugars, which then can be used in a number of ways in the plant cell (energy storage, plant cell wall synthesis, etc.).

Figure 05-16, below, helps illustrate where everything happens.

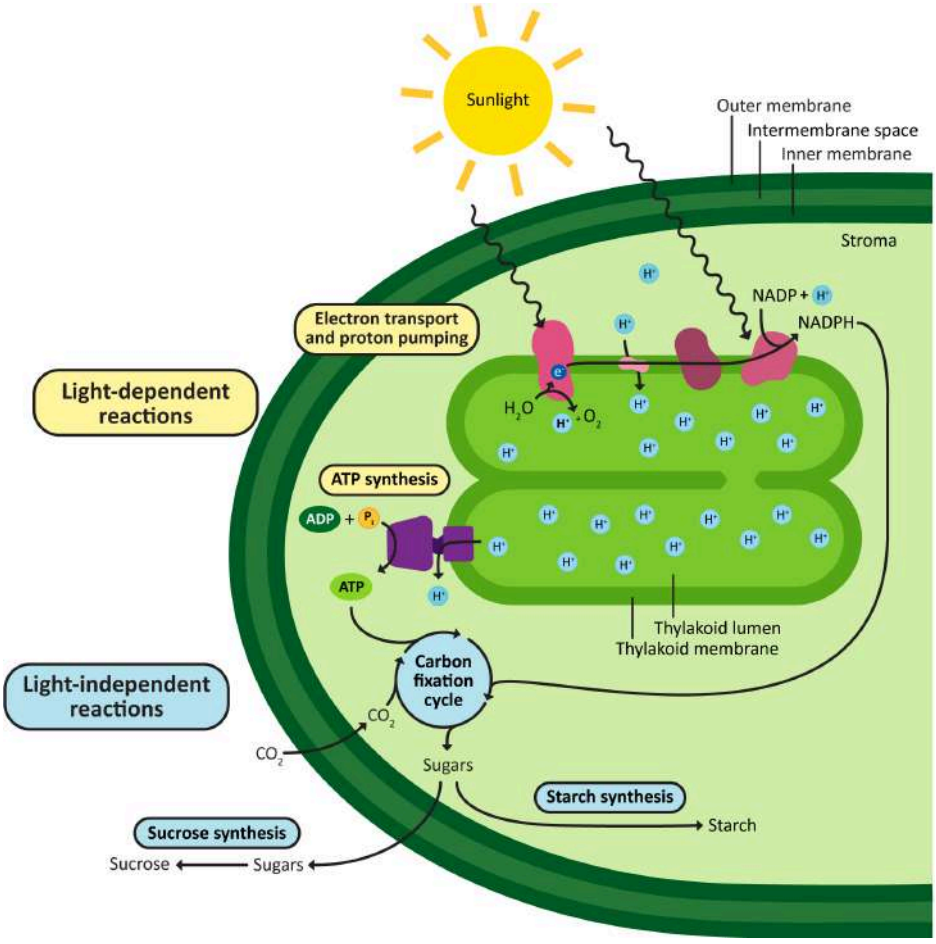


Figure 05-16: Photosynthesis in the chloroplast. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Once again, it's important to take time to really understand Figure 05-16. Consider carefully how this compares to the organization of mitochondrion and cellular respiration, as seen in Figure 05-12. Think again about where the areas of high and low pH would be, the direction of protein flow across the ETC, and the site of ATP synthesis.

A few other things to note:

- The goal of photosynthesis in the chloroplast is to produce and export sugars. ATP is made in an intermediate step but then used up by the synthesis of the sugars.
- While the chloroplast can synthesize ATP, *it cannot export it*. The ATP is entirely used up in the process of carbohydrate production.
- The plant's mitochondria must still be used to supply ATP to the rest of the plant cell.

Other Roles of the Chloroplast

Like the mitochondria, the chloroplast has additional important functions beyond photosynthesis. As autotrophs, plants must make *every* molecule that they need to survive, and the chloroplast plays a key role in the synthesis of many of these essential compounds, including the following:

- Almost all of the amino acids for the plant cell are synthesized by the chloroplast. The only exception is the sulfur-containing amino acids, cysteine and methionine, which are made elsewhere in the cell.
- The purines and pyrimidines used to produce nucleic acids are also synthesized here.
- It is involved (along with the ER) in the production of complex lipids, such as fatty acids and some sterols.
- It is involved in the production of stress response chemicals that are involved in triggering the plant immune response.

CHAPTER SUMMARY

In this chapter, we have taken a deep dive into the structure and function of the mitochondria and chloroplasts. First, we explored their evolutionary origins, which are described by the endosymbiont theory. Their evolutionary origins explain quite a bit of their structure and how similar it is to their closest living bacterial relatives. We also looked at how proteins that are translated in the cytosol are able to be imported into either the mitochondrion or the chloroplast (but not both).

We focused heavily on the similar purpose of these organelles, in that they both create chemical energy for the cell, but the products they create are different. Mitochondria focus on creating ATP, which can be exported and used in the rest of the cell, whereas chloroplasts do not export any of the ATP that they make. Instead, they use the ATP to create carbohydrates, which are then exported and used by the cell in a variety of ways.

Finally, we explored the structure of each of the organelles as well as the additional roles that these organelles play in the proper function of the cell.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required

to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

Topic 5.1: Evolutionary Origins and Protein Targeting to the Mitochondrion and Chloroplast

1. Shown below are a few major features of mitochondrial and chloroplast structure and function. Explain how the features of these elements are consistent with the notion that they originated from prokaryotes:
    a. cell-surface membranes
    b. internal membranes
    c. ribosomes
    d. genomes
    e. electron transport systems
    f. formation of ATP
    g. any other characteristics that you can think of
2. What features are shared by mitochondria, chloroplasts, and bacteria that are not shared by their eukaryotic hosts?
3. Compare and contrast the mechanisms of import into mitochondria and chloroplasts with import into the other organelles that were discussed in previous topics.

Topic 5.2: Function of Mitochondria and Chloroplasts—ATP Producers with Differing End Goals

1. What is the main energy-storing molecule produced by mitochondria? Chloroplasts?
2. Compare the nature of the electrochemical gradient involved in the generation of ATP in mitochondria with that in chloroplasts.
3. Compare and contrast mitochondria and chloroplasts with regard to structure and function.
4. List the main products of the mitochondrial electron-transfer reactions.
5. List the main products of the photosynthetic electron-transfer reactions. Why is light needed for these reactions?
6. The conversion of CO2 to carbohydrates does not directly require light energy. What necessary components does light energy indirectly provide for this conversion to occur?
7. If chloroplasts make ATP and synthesize sugars, then why do plants need mitochondria even when growing in the light?
8. Trace a molecule of oxygen from its entry into an organism, through the cell, and into the mitochondria for use in oxidative phosphorylation.
9. Trace energy from the sun to energy in ATP in the cytosolic compartment of a plant cell.

Topic 5.3: Structure-Function Relationships in Mitochondria and Chloroplasts

Make a diagram of a chloroplast and a mitochondrion and label the diagram(s) with the following terms and structures. Note that all structures are not expected to be in both organelles.

1. outer, inner, and thylakoid membranes and cristae

2. matrix, stroma, thylakoid lumen, and intermembrane space
3. relative pH of each compartment
4. ATP synthase and compartment where ATP synthesis occurs
5. mitochondrial electron transport chain
6. light-dependent reactions
7. carbon-fixation reactions
8. citric acid cycle enzymes
9. site of organelle DNA and ribosomes
10. pyruvate transferase
11. ATP exporter
12. porins
13. protein import machinery (i.e., TIM/TOM, TIC/TOC)

REFERENCES

Bölter, B., & Soll, J. (2016). Once upon a time—chloroplast protein import research from infancy to future challenges. *Molecular Plant*, *9*(6), 798–812. https://doi.org/10.1016/j.molp.2016.04.014

Wiedemann, N., & Pfanner, N. (2017). Mitochondrial machineries for protein import and assembly. *Annual Review of Biochemistry*, *86*(1), 685–714. https://doi.org/10.1146/annurev-biochem-060815-014352

CHAPTER 6.

THE CYTOSKELETON

INTRODUCTION

We often think of the many pathways of the cytoskeleton as "the highways of the cell," yet the reality is so much more complex and fascinating. They do not simply act as highways but highways that can disassemble and completely rebuild themselves when the need arises. Throughout the previous chapters of this book, we have made reference to the cytoskeleton and its role in the various cellular processes. This is because the cytoskeleton is involved in virtually every aspect of cellular function. It's impossible to ignore the central role the cytoskeleton plays in the cell. The cytoskeletal "highways" help provide structure and connection within the cell and between cells. The cytoskeleton influences the shape of the cell and aids in the relay of signals from one part of the cell to another. The cytoskeleton helps cells grow, move, retract, and even die. We can see the impact of the cytoskeleton in every chapter you've read thus far, and it will continue to play an important role in the remaining chapters.

In this chapter, we will explore the components of the cytoskeleton, the details of how it is built, and how it functions. We will highlight some specific examples of function that are considered more universal, but know that there are so many more examples that could also be included.

TOPIC 6.1: OVERVIEW OF THE CYTOSKELETON AND INTERMEDIATE FILAMENTS

Learning Goals

- Compare and contrast the structure and function of the three types of cytoskeletal elements: actin filaments, intermediate filaments, and microtubules.
- Identify the different types of cytoskeletal elements on different types of micrographs.
- Discuss different types of intermediate filaments, and correlate their structure with the strength of the assembled, functional cytoskeletal element.

Introduction

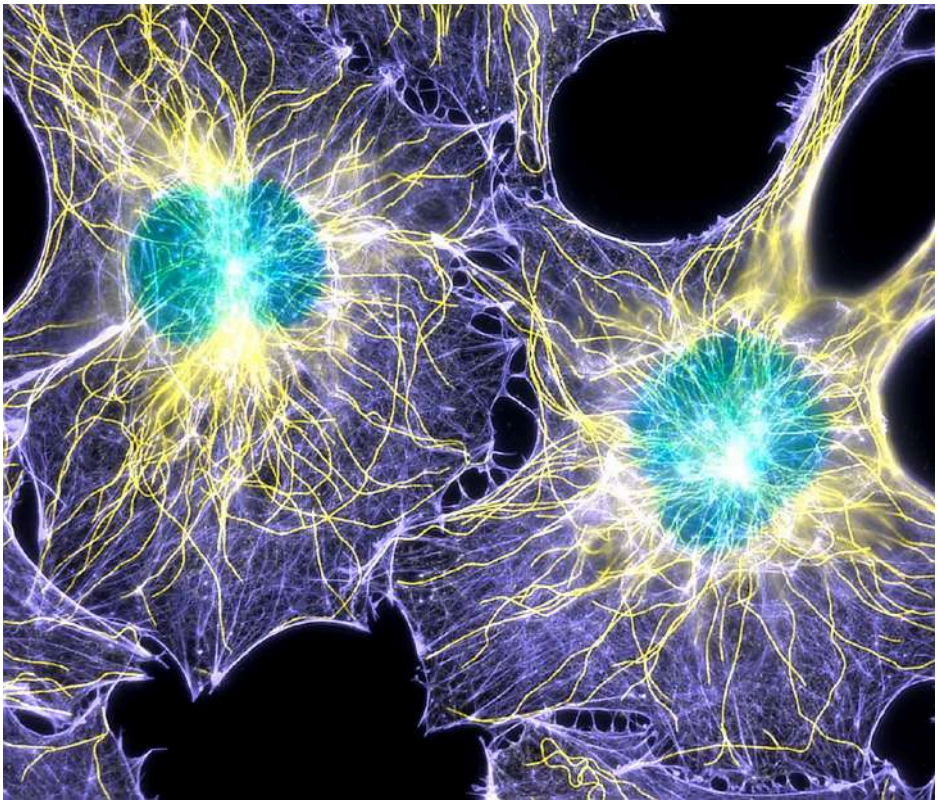


Figure 06-01: Two mouse fibroblast cells in interphase. Actin filaments (purple web), microtubules (yellow lines), and nuclei (green blobs in center) are labeled with fluorescent molecules. This image won first place in the Nikon 2003 Small World photo competition. Torsten Wittmann (2011), CIL:240. CIL. Dataset. https://doi.org/10.7295/W9CIL240. This image is in the public domain.

The cytoskeleton is a filamentous network that extends throughout the cell. There are three different types of cytoskeletal filament that exist. All known eukaryotic cells have both **actin filaments** and **microtubules** (Figure 06-01). The third type, **intermediate filaments** (Figure 06-02), is no less important in the cells that have them; however, they are really only found in the cells of vertebrates. Collectively, all three of these filaments have vital roles to play in virtually every function in the cell. Not only do they help give the cell its shape, but they also help resist mechanical stresses; act as an anchor for many proteins and organelles; provide tracks for transport of vesicles, organelles, and other cargo; and aid in the locomotion of many cell types. They also help with the division of both the cytoplasm and the DNA during mitosis.

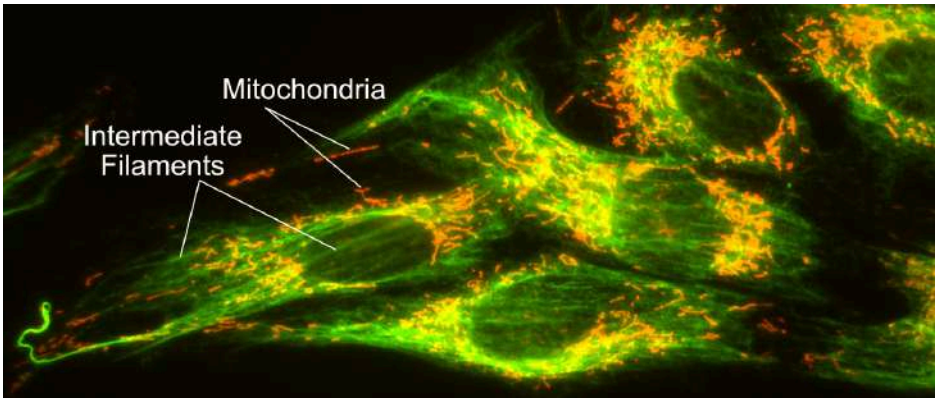


Figure 06-02: Mouse (Mus musculus) fibroblast cells, labeled to show the mitochondria (red squiggles) and intermediate filaments (green lines). Adapted from Linda Parysek and Trudy Aebig (2010), CIL:2, Mus musculus, fibroblast. CIL. Dataset. https://doi.org/10.7295/W9CIL2. This image is in the public domain.

Even cell types that were previously thought to lack a cytoskeleton have now been shown to have proteins that fulfill these roles. A very good example of this are bacteria, which were originally thought to have no cytoskeleton at all, but we now know have a cytoskeleton-like network that functions similarly. Another example is plant and insect cells, which do not have intermediate filament genes in their genomes, yet they both still have a nuclear lamina. To build it, they use genetically unrelated proteins, which perform the function of nuclear lamins (we call them *lamin-like proteins*). As such, it is safe to say that the cytoskeleton is (another) essential component of all living cells. Cells simply cannot survive without it.

The features of each of the types of cytoskeletal filaments that we list here will be explained in further detail throughout the rest of the chapter. For now, here is a quick summary, in order of filament size (see also Video 06-01):

1. **Microtubules (MT)—**described in Topic 6.2
    - The largest of the three types (~25 nm in diameter).
    - Made of repeating subunits called tubulin. Tubulin itself is a dimer, made of two smaller subunits, called alpha- and beta-tubulin.
    - Go through rounds of polymerization and depolymerization (**dynamic instability**).
    - Involved in a variety of cell functions.
        - In all cells, they are involved in mitosis, cargo trafficking, and maintenance of cell shape.
        - In animal cells, they radiate outward from the center of the cell and act as "highways" for vesicular transport.
        - In plants and algae, they play an important role in the secretion of cellulose, which happens at the plasma membrane. Thus, they are found in the **cell cortex** in plants (meaning that the microtubules are found directly underneath the plasma membrane).
        - In fungi and other protists, the arrangement of the cytoskeleton is more variable.
        - Both cilia (singular **cilium**) and flagella (singular **flagellum**) are specialized

structures that depend on microtubules for function.

2. **Intermediate filaments (IF)—**described later in this topic (6.1)
    - Their diameter is "in between" the other two types of cytoskeletal filaments (~10 nm).
    - They are a family of proteins classified into five major "types" and subtypes (e.g., keratins, vimentin, neurofilaments, and nuclear lamins). These types describe the subunit that they are made from.
        - Generally, only one or two subtypes are expressed in each cell types (i.e., nuclear lamins plus keratin, for example).
        - A single intermediate filament is always made from the same kind of subunit. They do not mix and match.
    - They do not undergo dynamic instability, but they are assembled and disassembled as needed.
    - They help with anchoring cells that are not motile, resist mechanical strain in tissues, and are vital to the maintenance of cell shape. Nuclear lamins are very important in the maintenance and organization of the nucleus.
3. **Actin filaments (AF)—**described in Topic 6.3
    - Also known as microfilaments.
    - The smallest of the three types (~7 nm in diameter).
    - Made of repeating subunits, also called actin.
        - Monomers of actin are sometimes referred to as *g-actin* to differentiate from the filament form, which is call *f-actin*.
    - Like microtubules, they also undergo dynamic instability.
    - In animal cells, they are heavily involved in cellular locomotion but also cellular adhesion and shape, cytokinesis, cargo trafficking, and muscle contraction. As a result, actin filaments tend to remain close to the plasma membrane, in the **cell cortex**.
    - In plant and algae cells, the cellular arrangement is different. They are found throughout the cytoplasm (not just underneath the plasma membrane). As such, they are involved more heavily in cargo trafficking and maintenance of cell shape (plant cells are not motile). They also help produce the phenomenon known as **cytoplasmic streaming**.
    - In other cells, once again, the arrangement is more variable.

One or more interactive elements has been excluded

— from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-1

Video 06-01: A general overview of intermediate filaments, microtubules, and actin filaments in animal cells.

Intermediate Filaments (IFs)

As mentioned, intermediate filaments get their name from the fact that they are "intermediate" in size between actin and microtubules. Their importance in the cell is often overshadowed by the focus on actin and microtubules when discussing the cytoskeleton and by the fact that both plants and insects do not appear to carry any genes in this family. This discovery initially implied they were "less important" than the "universally essential" filaments. However, we have since discovered that even though the genes for intermediate filaments are not always present, there are often other proteins that are able to functionally fill this role. For example, there is always a structure that resembles the nuclear lamina in all nuclei, and it functions similarly, suggesting that it plays a vital cellular role. In addition, in the cell types that contain additional intermediate filaments (like skin or neurons), they play a key role, helping to maintain shape and resist mechanical strain. Thus, intermediate filaments, or intermediate filament-like proteins, are an essential cytoskeletal element for all cells.

Intermediate filaments (**IFs**) are a genetic family of proteins rather than a single highly conserved type. Within one filament, only one type of protein subunit is present. Each intermediate filament type is tissue specific, which means that you can tell cell type based on what kind of IF it expresses. For example, neurons express neurofilaments, but skin cells do not.

Some of the most well-known types of intermediate filaments include the following:

Making Connections: IFs

Can you speculate how this might be useful in cancer diagnoses? Imagine that you found keratin, which is normally expressed in skin and hair, in an abdominal tumor. What type of cancer does the patient have?

- *Keratin*: Found in epithelial cells and their derivatives, hair, nails, and horn. They help hold skin cells to each other and to the underlying membrane and provide resistance against mechanical tension (also made famous by shampoo commercials promoting their effects on

keratin).

- *Vimentin*: Found in intracellular fibers of connective tissue cells, including muscle. They help hold tissues together and provide strength.
- *Neurofilaments*: Found in neurons. They help form and support the characteristic (and extreme!) shape of a nerve cell.
- *Lamins*: Components of the nuclear lamina. We discussed these previously in Chapter 3. They form a meshwork beneath the nuclear envelope and help maintain the integrity of the nuclear envelope.

Intermediate filaments are considered less dynamic compared to the other filaments (i.e., actin and microtubules). They do not undergo **dynamic instability**, which is a process of rapid switches between growth and shrinkage that is key to microtubule and actin function. However, they are still able to assemble and disassemble as the cell requires. A perfect example of this is the breakdown of the nucleus during mitosis, which is driven by the phosphorylation and breakdown of the nuclear lamina, which we saw in Chapter 3 and will discuss again later in this chapter.

Structure of Intermediate Filaments

Regardless of the specific subunit used to make it, all intermediate filaments are somewhat similar and follow the same pattern of assembly. Like all three of the cytoskeletal filaments, intermediate filaments are formed when smaller subunits, known as **monomers**, come together in a particular way to form a longer, filamentous **polymer**. So when we refer to the specific types of intermediate filament (i.e., keratin, vimentin, etc.), we are referring to the monomer being used as the building block.

All of the intermediate filament monomers are what are known as **filamentous proteins**, meaning that they are long and threadlike (Figure 06-03). They have a large central region that is a long alpha helix, which makes up the bulk of the filament. On either end are the "head" and "tail" region, which are specific to each of the different monomer types (keratin, vimentin, etc.). It is these head and tail regions that help the intermediate filament carry out its specific function. For example, the nuclear lamins will require sites that can bind to the chromatin, whereas keratin, which is not in the nucleus, does not require this feature.

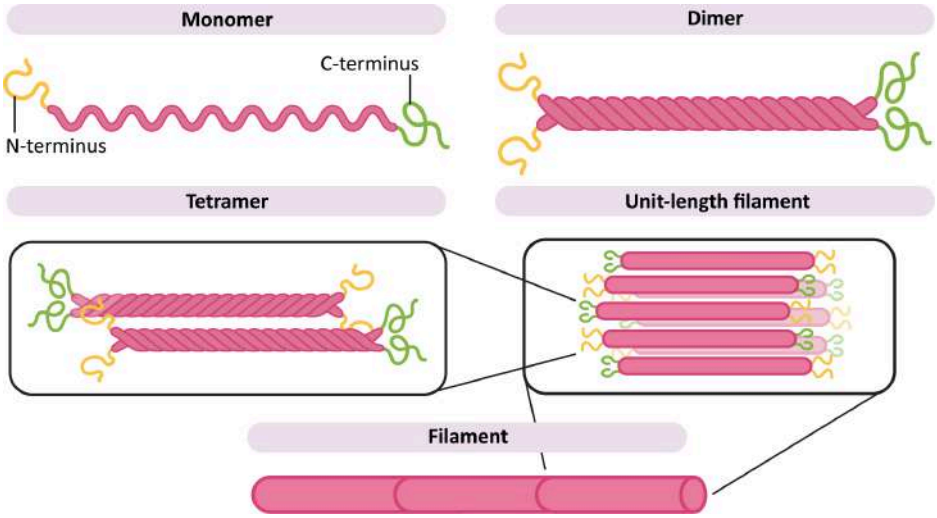


Figure 06-03: Intermediate filament assembly. Monomers coil together into dimers with the heads and tails on the same side. Then two dimers bind into an antiparallel (head-to-tail) tetramer. Groups of tetramers assemble into a ropelike structure. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

In order to build a filament (Figure 06-03), two monomers first come together to form a **dimer**. The dimers then assemble with each other in an **antiparallel** arrangement, meaning that they are oriented in the opposite direction from each other. This forms a **tetramer**. The tetramers then assemble to form the long filament. Since the ends of the monomers are staggered compared to each other in the filament, this forms a long, ropelike structure that doesn't stretch and is difficult to break.

Like many protein complexes, the intermediate filament subunits undergo **self-assembly** in order to spontaneously form the final filament. This means that energy is *not* required to facilitate this assembly process (this is different from actin and microtubule assembly).

Making Connections: SNAREs

SNAREs, which you encountered when learning about vesicle transport in Chapter 4, also use the coiled-coil strategy to come together.

Remember from Chapter 2 that the alpha helix is a secondary structure, commonly used in protein folding. Thus, the amino acid sequence of the IF subunit is such that the alpha helix takes shape on its own. The formation of the dimers and tetramers is also facilitated by the amino acid sequence. Chemical analysis of the alpha helix shows that they have a line of nonpolar amino acids on one side of the alpha helix. This facilitates the formation of a structure known as a **coiled coil** (Figure 06-04). This is a very common structural strategy, as it ensures precise alignment of subunits, and the nonpolar interactions are difficult to break in the aqueous environment of the cell.

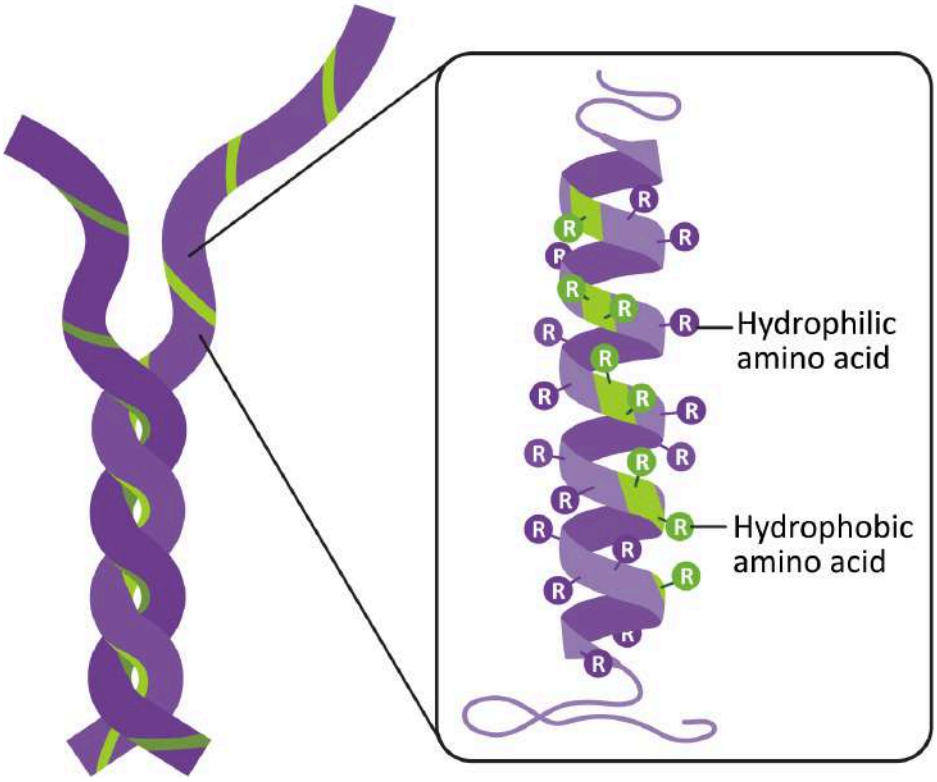


Figure 06-04: The arrangement of the polypeptides in a coiled coil. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Examples of Intermediate Filaments in Cells

Example 1

The Nuclear Lamina

Since all eukaryotic cells have a nucleus, the nuclear lamina is easily the most important example of intermediate filament function. We have already discussed this function back in Chapter 3, when we examined the structure of the nucleus in detail. The nuclear lamina is vital to the proper organization and function of the nucleus throughout the cell cycle. All Eukaryotes that have a nucleus will also have a nuclear lamina of some kind.

In interphase, the assembled nuclear lamins form a latticework just underneath the nuclear envelope (Figure 06-05). Not only does this help provide structure to the nuclear envelope, but it also helps with the organization of the chromatin. The chromosomes are attached to the nuclear envelope and the underlying lamina. This attachment helps maintain the regional organization of the chromatin, thus facilitating gene expression.

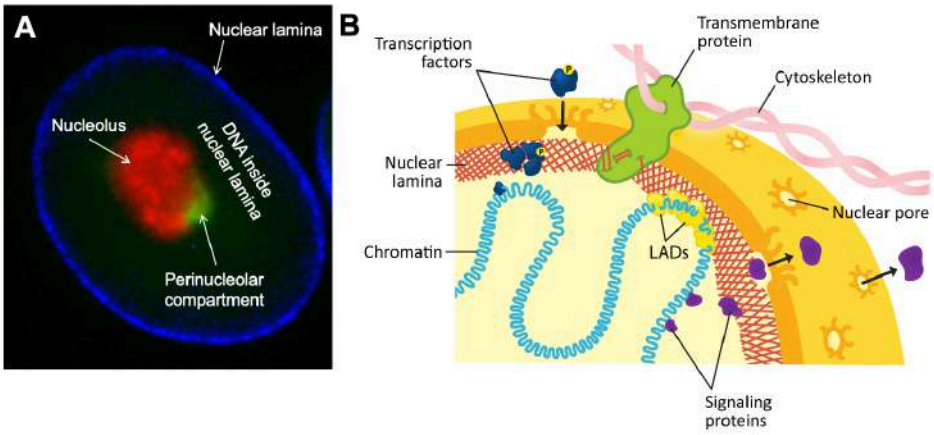


Figure 06-05: The nuclear lamina and its role in the nucleus. (A) Close-up, labeled fluorescent micrograph of the nucleus of a cell derived from the Henrietta Lacks (HeLa) cell line. The nuclear lamina is a blue circle (labeled) that lies directly underneath the (unlabeled, not visible) nuclear envelope. Adapted from Sui Huang, Thomas J. Deerinck, Mark H. Ellisman, and David L. Spector (2010), CIL:9233, Homo sapiens, cervical carcinoma of Ms. Henrietta Lacks (https://doi.org/10.7295/W9CIL9233) and shared under a CC BY-NC-SA 3.0 license. (B) Schematic of the interphase nucleus showing the location and function of the nuclear lamina. This image was created by Heather Ng-Cornish and is shared under a CC BY-NC-SA 4.0 license.

Not only is the nuclear lamina attached to the nuclear envelope and internal components of the nucleus, but it also maintains attachments to the cytoskeleton on the exterior of the nucleus (Figure 06-05). This attachment can help maintain the location of the nucleus within the cell (and allow it to be moved if necessary). A variety of proteins facilitate connections between the nuclear lamina, on the inside of the nucleus, and the cytoplasmic cytoskeleton. In some cases, specialized transmembrane proteins form complexes spanning the inner and outer membranes of the nucleus.

During mitosis, destabilization of the nuclear lamina (via phosphorylation of the lamins) is what drives the disassembly of the nuclear envelope (Figure 06-06). Once mitosis is over, the lamins are dephosphorylated, and the nuclear lamina reforms. Since it's attached to both the membrane and the chromosomes, the result is that everything ends up back where it's supposed to be.

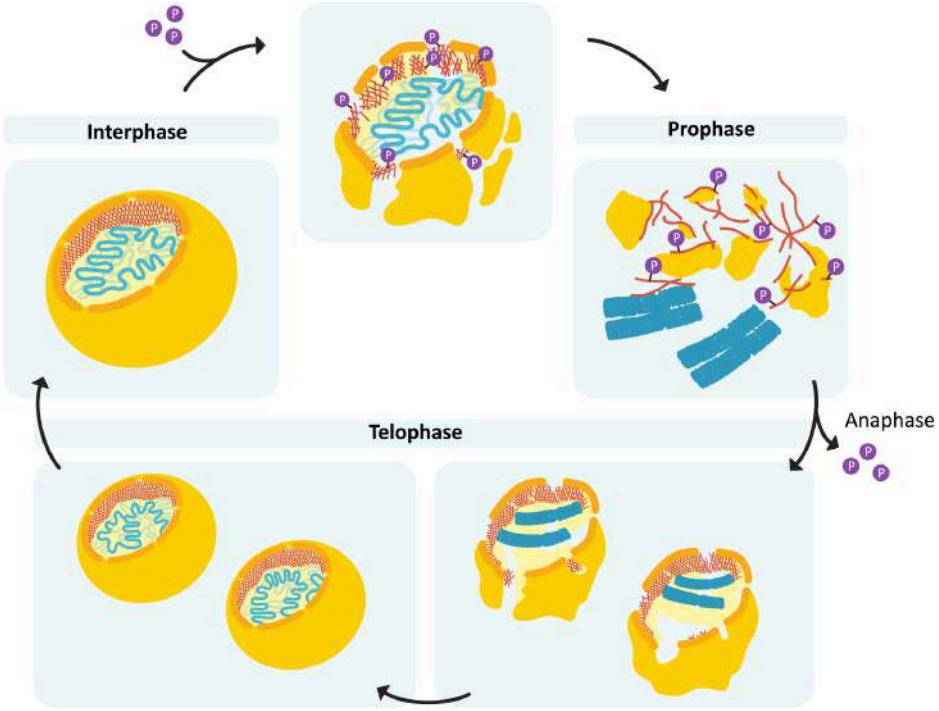


Figure 06-06: The breakdown and reformation of the nuclear lamina during mitosis. Lamins are phosphorylated at the beginning of mitosis, which triggers the nuclear lamins to depolymerize. This results in the breakdown of the nuclear envelope as well. This is reversed in anaphase when the lamins are dephosphorylated. The nuclear lamins polymerize and the nuclear envelope reforms. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

To further illustrate the importance of these nuclear lamins, here are a few examples of genetic diseases that result from mutations in lamin proteins. In one disease known as Hutchinson-Gilford progeria syndrome, children appear to age at a highly accelerated rate and often die in their early teens (Prokocimer et al., 2013). Emery-Dreifuss muscular dystrophy is caused by a few different genes, but one of them is a nuclear lamin, and others include proteins that are also involved with maintaining the shape of the nuclear envelope (Maggi et al., 2021).

Example 2

Neurofilaments and Cell Shape

As you likely know, neurons have a very distinct shape, characterized by the axon that projects out one side of the cell. Axons can be extremely long—the longest axon in the human body is the sciatic nerve, which runs from the base of your spinal cord to your big toe, so it's about 1 m long. However, larger animals likely have much longer axons than that. (The blue whale is thought to have a dorsal root ganglion that is 25–30 m long!!!)

Cells do not easily make this kind of extreme shape…it's a long way from the sphere that is considered to be the most energetically favorable shape for a cell! Thus, the cytoskeleton, including both microtubules and intermediate filaments, is required to maintain this structure. Neurons express a type of intermediate filament known as a *neurofilament*. Both neurofilaments and microtubules can be seen in the cross section of a nerve axon in Figure 06-07. If either the neurofilaments or the microtubules are disrupted, the axon will lose its integrity and retract, and the cell will die. Many kinds of dementia are the result of this kind of loss of shape due to impaired cytoskeleton function.

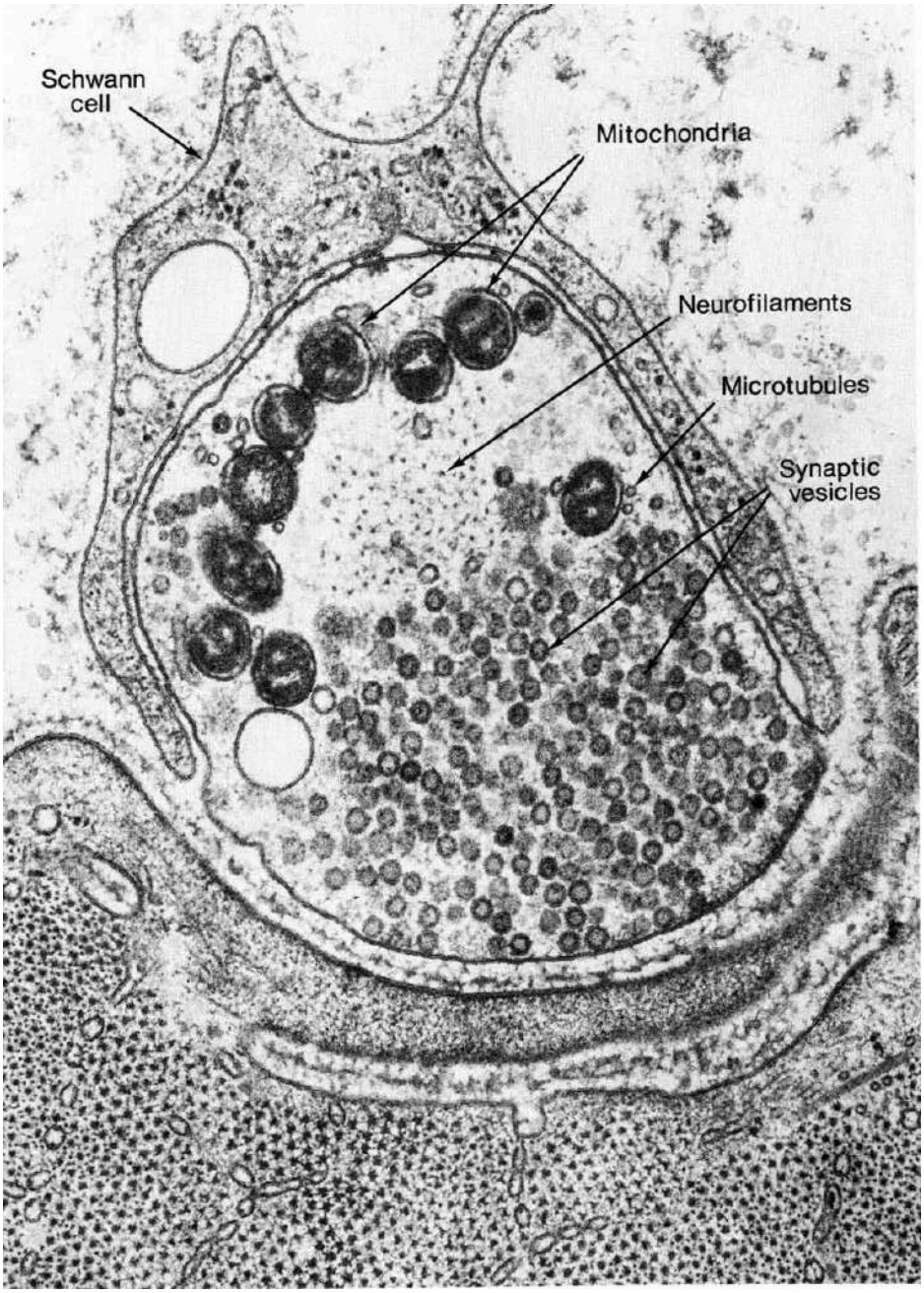


Figure 06-07: TEM cross section of a frog neuron at the edge of a neuromuscular junction. At the top, a Schwann cell can be seen (part of the myelin sheath of a neuron). At the bottom, the muscle cell is visible, with its highly ordered actin ready to initiate muscle contraction. Don W. Fawcett, John Heuser, and Tom Reese (2011), CIL:36010, Anura, peripheral neuron, muscle cell, myelinating Schwann cell. CIL. Dataset. https://doi.org/10.7295/W9CIL36010. Shared under a CC BY-NC-ND 3.0 license.

Example 3

Desmosomes and Hemidesmosomes

As was mentioned earlier, not all cells appear to require cytosolic intermediate filaments in high numbers, but in the cell types that have them, they are key to proper function. Our skin and the cells lining our gastrointestinal tract are perfect examples of this.

The role of our skin is not just to keep all of our internal bits on the inside but also to protect us

from the world around us. It forms a flexible yet impenetrable barrier to pathogens, like bacteria, viruses, and other toxins. Thus, even if we twist, push, or pull on our skin, we cannot penetrate it. Breaking the skin barrier requires a sharp object that can cut through the cells and expose the tissues underneath. (*We do not recommend trying this...*)

In order for our skin to resist everything that we put it through, the cells need to be much stronger than what a plasma membrane can provide on its own. Intermediate filaments, specifically keratin, are used to help our skin stay strong. The keratin filaments pass from one side of the cell to the other and bind to large protein plaques in the plasma membrane. These plaques, known as **desmosomes** (Figure 06-08), bind to similar plaques in adjacent cells, which are attached to that cell's keratin. The keratin + desmosome complex transmits mechanical stress along the ropes instead of through the membrane. Desmosomes act like the rivets in your jeans that hold the different pieces of fabric together. They also provide an anchor point for the keratin filaments. Not only do desmosomes attach cells to each other, but they also bind to the membrane layer underneath the skin (called the basal lamina). In this case we call them **hemidesmosomes**.

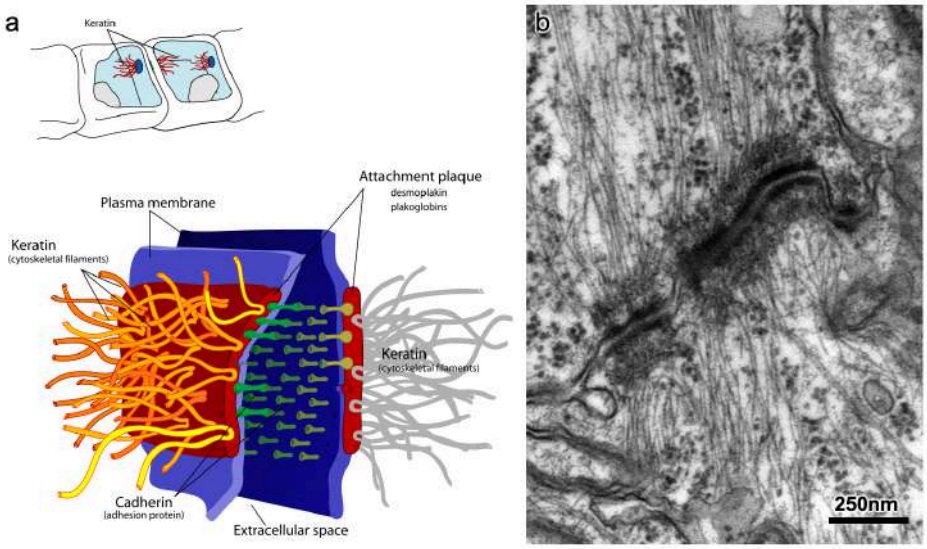


Figure 06-08: Desmosomes. (A) Schematic of the structure of desmosomes, showing the attachment plaque and the intermediate filaments (keratin) that connect desmosomes through the cytoplasm. (B) Electron micrograph showing two desmosomes in a cell membrane and the keratin filaments leading away on either side of the desmosome. Dr. Robin Young compiled the full figure from the following sources: Panel A is by Mariana Ruiz. Panel B is adapted from Wai Yan Lam and Marian Rice (2011), CIL:27225, Fundulus heteroclitus, epidermal cell. CIL. Dataset. https://doi.org/10.7295/W9CIL27225. Both images are in the public domain.

Once again, we can see the importance of intermediate filaments by examining the diseases that result from genetic mutations in keratin. Epidermolysis bullosa is a very painful disease in which the skin blisters and tears at the slightest touch. Even soft clothing and the act of eating can cause ruptures and blisters in the skin. The children born with this disease are often referred to as "butterfly children," as their skin is incredibly fragile (like the wings of a butterfly). It is caused by several different mutations affecting the strength of the attachment of the skin, including more than one mutation in keratin genes. Like the other diseases we've seen in this section, there is no cure. However, regeneration of skin through stem cell therapy and grafting has shown promise in recent years.

TOPIC 6.2: MICROTUBULES AND DYNAMIC INSTABILITY

Learning Goals

- Compare the mechanics of microtubule polymerization in vivo (in a cell) and in vitro (in a test tube).
- Describe the process of dynamic instability, focusing on how tubulin dimers (a/b) interact noncovalently to make microtubules with a "plus" and "minus" end.
- Interpret live-cell images of microtubules undergoing dynamic instability and explain the current model of how GTP binding to tubulin can bring about this dynamic behavior.
- Explain how motor proteins work and how their movement relates to the polarity of the microtubule.
- Provide examples of how accessory proteins can interact with microtubules or tubulin to influence their structure, which, in turn, will influence their function.

The second of the three types of cytoskeletal filament that we will discuss is **microtubules**. These filaments differ significantly from intermediate filaments in many fundamental ways. For example, their assembly and disassembly are considered more dynamic than that of intermediate filaments. Each microtubule filament is in constant flux, switching between phases of growth and shrinkage. This process of rapid growth and shrinkage is known as **dynamic instability**. It is fundamental to the function of both microtubules and actin filaments. As such, we will use microtubules as a case study to explore how dynamic instability works in both types of filaments. After that, we will discuss the function of microtubules prior to moving on to the final topic of this chapter, where we will discuss actin structure and function.

Microtubule Structure

Microtubules are protein polymers in the shape of hollow tubes (Figure 06-09). The tube is formed by linear associations of **tubulin** subunits, which we call **protofilaments**. There are 13 of these protofilaments arranged in a circle to form each microtubule.

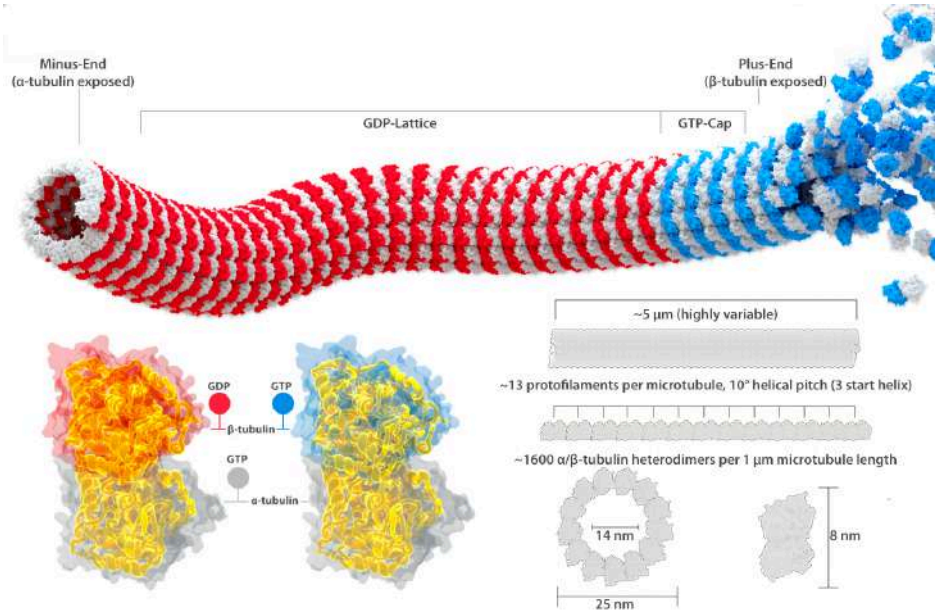


Figure 06-09: Structure of a microtubule. Alpha- and beta-tubulin arrange in a linear fashion to create a protofilament. Thirteen protofilaments form a sheet that wraps to create a hollow tube. Image by PSK615 is shared under a CC BY-SA 4.0 license.

The "subunit" that is added to a growing microtubule is actually a dimer, made up of one alpha- and one beta-tubulin (Figure 06-09). Alpha- and beta-tubulin are each about 55 kDa each, and together they are about 8 nm long. The positions of tubulin dimers in the microtubule in adjacent protofilaments are slightly out of phase so that they form a spiral with a relatively low pitch, as you can see in Figure 06-09. Each monomer in the tubulin dimer also carries a GTP molecule within it. The alpha-tubulin subunit grabs the GTP and keeps it in its core but does nothing else with it. On the other hand, the GTP in the beta-subunit can undergo a hydrolysis to convert from GTP to GDP. Thus, the state of the beta-subunit, and whether it is currently bound to GDP or GDP, will affect the chemical properties of the entire tubulin dimer. This shift in chemical composition affects how well the protein monomers assemble into a polymer as well as the stability of the assembled polymer.

Before we go further, we need to take a step back and address the name of these monomers. First they are *confusing*. The tubulin dimer is made of a couple of different tubulin monomers, and both the monomers and the dimers all get called "tubulin" at times. For the most part, we try to refer to the dimer subunits by their actual names (alpha- and beta-tubulin, or alpha- and beta-subunit) so that when we refer to tubulin without mentioning alpha- or beta- (or gamma-, as there are even more types of tubulin that you will learn about later), you will know that this is intended to refer to the *tubulin dimer*. We also very commonly call it the tubulin dimer, to be extra clear about what we're referring to, but the rest of the world does not necessarily use this term. Thus, it is important that you consider the context when "tubulin" is mentioned.

Since the tubulin dimers are all lined up in the same orientation, a different side is exposed at each end. At the end called the **plus end** in Figure 06-09, it is the beta-subunit that is exposed, whereas the alpha-subunit is exposed at the **minus end**. This is an important observation, as it tells us that there are going to be structural differences between the different ends. (This is unlike intermediate filaments, whose antiparallel subunit arrangement means that there is no obvious way to differentiate one end from the other structurally.) The structural difference between the ends of the microtubule is a key characteristic and influences microtubule function. This structural characteristic is also why

we say that microtubules have **polarity**. This simply means the ends are *different* from each other. Labeling them as the plus and minus ends is not related to electrochemical charge in any way. Instead, this nomenclature is related to their propensity for growth. The plus end grows more easily, and faster, than the minus end does. This is the direct result of the subunits that are exposed at each end.

The video shows microtubule assembly and disassembly first, and then moves onto actin assembly. When researchers examined how the assembly of the polymer worked, they found that much like polypeptides and nucleic acids, microtubules grow by addition of new subunits at the ends of the polymer only. The tubular structure of a microtubule is assembled in a manner that can be quite striking. Assembly is shown in Video 06-02.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-2

Video 06-02: Overview of assembly and disassembly of microtubules and actin filaments.

Microtubules and Dynamic instability

Re: Dynamic Instability

Dynamic instability can be challenging to understand! The entire cell is really just a collection of chemical reactions, which are the result of the properties of the molecules involved and the environment they're in. We know this, and yet we don't often think about it. As a result, biology can sometimes seem far removed from the chemistry that underlies it. In the case of dynamic instability, it helps if you work to remember your knowledge of chemistry and chemical reactions and consider the polymerization reaction in that light.

Microtubules are in a constant state of flux. They will grow, pause, and then shrink again in the span of a few seconds. They may shrink down to nothing or pause and then start growing again. Because of this constant state of flux, which shows that microtubules are dynamic (i.e., constantly changing from growth to shrinkage and vice versa), and because they are clearly very unstable (since they fall apart a lot), we say that microtubules undergo **dynamic instability**.

The discovery of microtubules dates back almost as far as the invention of the microscope, with observations of cilia and flagella in the late 1600s. However, it wasn't until the 1980s that scientists developed the theory explaining dynamic instability.

Dynamic Instability in Vitro (i.e., in a Test Tube)

The simplest way to understand microtubule polymerization is to remove tubulin from cells and see how it works on its own in a test tube. In essence, this is what Mitchison and Kirschner did in 1984 when they first described dynamic instability. They discovered a few key features, which were then built on by other scientists as time went on. Here, we outline the basic theory of dynamic instability. Later, we will see how dynamic instability is central to the function of both actin and microtubules.

Tubulin dimers have a certain affinity for each other, and if there is a high enough concentration of tubulin in the test tube, they will start to polymerize into a polymer spontaneously. The polymer forms using intermolecular forces exclusively (i.e., no covalent bonding! See the introduction if you need a refresher on what this means).

This follows a predictable pattern and, on a graph, looks very similar to other chemical reactions (Figure 06-10). In essence, once you add enough tubulin to a test tube, the reaction starts to take place. At first the reaction is slow (lag phase), followed by an exponential phase of fast polymer growth. Like all reactions in a test tube, eventually this will reach equilibrium, and the concentration of tubulin that is incorporated into microtubules will stabilize and remain constant. Even so, there will still be dynamic instability going on in the test tube, as we know that in chemistry, at equilibrium, both the forward reaction and the reverse reaction happen at the same rate.

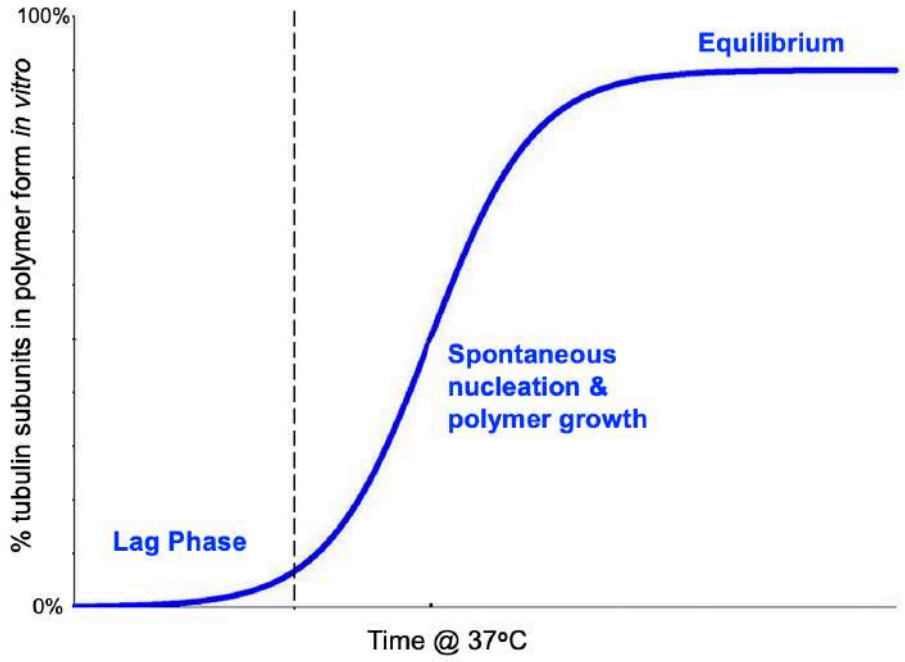


Figure 06-10: Graph showing typical curve for the formation of microtubules in vitro. Based on data from Roostalu and Surrey (2017). Image by Dr. Robin Young is licensed under CC BY-SA 4.0.

The **lag phase** is the slowest stage of the reaction in Figure 06-10. Initial nucleation is dependent on molecules running into each other in the right orientation, via random motion, in the aqueous solution of the test tube. Because of this, it's easy to imagine that the concentration will have a strong effect on how this reaction proceeds. If the concentration of tubulin is too low, then it will be difficult for the tubulin subunits to find each other in solution to start the process. In the experiment that Figure 06-10 is based on, the concentration of tubulin is quite high, almost 20 times the concentration in an actual cell, and yet the lag phase is still quite slow. The lag phase is the rate-limiting step of the overall reaction.

If the concentration of tubulin is too low in the test tube, the process will not even start. The entire reaction essentially gets trapped in lag phase, and polymerization will not occur. Because of this, we say that there is a **critical concentration (Cc)** of tubulin required for polymerization to proceed.

If we examine the relationship between rate of assembly and concentration more closely, there is more to learn. First of all, we must remember that each microtubule has *two* ends, the plus end and the minus end, which have different structural properties. At the start of this section, we said those structural differences caused the ends to be different in terms of rate and likelihood of polymerization. We can visualize this when comparing the rate of growth to the concentration of tubulin we add to the test tube, as shown in Figure 06-11. The graph teaches us two things:

1. Above the critical concentration, we see that the rate of assembly increases with increasing concentration of tubulin subunits in solution.
2. There is a difference in *both* the critical concentration *and* the rate of assembly at the plus and minus ends.

This means that the plus end *starts growing more easily* than the minus end and also *grows at a faster rate.* Thus, Figure 06-11 provides the evidence for what we said earlier about the plus end being "more prone to growth" than the minus end. Note that in a test tube, where the minus end is unbound, the

minus end is also able to grow, but it requires a higher concentration to get it started, and once it starts, it grows more slowly.

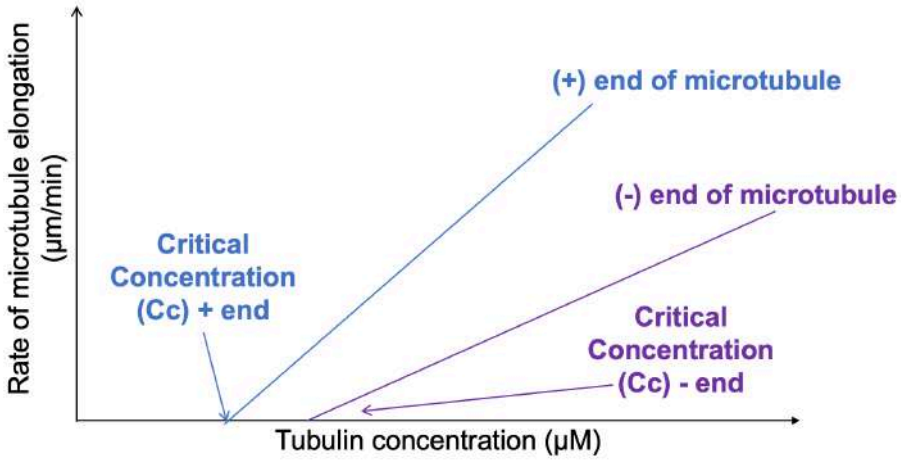


Figure 06-11: The rate of reaction is dependent on the concentration of the tubulin subunits available. The higher the concentration of tubulin subunits, the higher the rate of assembly. Notice also that the rate of assembly is always slower on the minus end of the polymer. On the plus end of microtubules, fewer subunits are needed (lower concentration) to begin filament formation. Graph is based on data from Walker et al. (1988). Image by Dr. Robin Young is licensed under CC BY-SA 4.0.

Interestingly, data collected on microtubule shrinkage rates show that unlike growth, the rate of shrinkage is *not* dependent on concentration. Shrinkage rate appears to be more or less constant at both ends at all concentrations. This is somewhat analogous to the fact that the rate at which a tower gets built is dependent on how many workers there are to do the work. However, destroying a tower has very little to do with the number of workers…once the primary supports are gone, the rest will come down very quickly.

Dynamic Instability in Vivo (i.e., inside Cells)

Now that we understand dynamic instability in the simpler environment of the test tube, we need to consider it in the context of the cell. The cell, of course, is a much more complex place, and the polymerization of microtubules will be heavily influenced by their environment. Local tubulin concentrations may go up and down as nearby microtubules grow and shrink. Additionally, accessory proteins will bind to both the tubulin dimers and the polymers to influence overall microtubule behavior based on the needs of the cell.

One important aspect of tubulin dynamics that we have not yet discussed is the difference between tubulin dimers that have GTP bound to their beta-subunit and those that have hydrolyzed their GTP into GDP. (Remember that the alpha-subunit can neither release nor hydrolyze its GTP, so we can ignore it here.) The hydrolyzation of GTP in tubulin causes a change in the conformation of the protein, which impacts its function. Here's what you need to know:

1. While both forms of tubulin are able to polymerize, the GTP-bound tubulin dimers are considered to be the *activated* form. GTP-bound tubulin has a higher affinity for the microtubule polymer, which results in a lower critical concentration, and it assembles at a faster rate compared to GDP-bound tubulin (which is often called the *inactive* form, even

though that's not entirely accurate).

2. GTP tends to be in excess concentration in the cytosol. As a result, GDP-bound tubulin that is not bound up in microtubules is immediately reactivated by replacing the GDP in the beta-subunit with GTP. So the concentration of free-floating, GDP-bound tubulin tends to be quite low in the cell.
    - As a result, in the cell, we really only need to consider GTP-bound tubulin during microtubule *assembly*, but *both* GDP- and GTP-bound tubulin play a role at other points in the life cycle of the microtubule.
3. As soon as a tubulin dimer binds and becomes part of a microtubule, the beta-subunit begins to slowly hydrolyze its GTP (Figure 06-12). The hydrolysis of GTP to form GDP reduces the affinity of the binding of the tubulin dimers to each other within the microtubule. This means that over time, the affinity of the tubulin dimers for each other (and, as such, the bonding energy holding the polymer together) will go down.
4. When hydrolysis of GTP occurs more slowly than the new monomers can be added at the plus end, the microtubule will grow at a maximum rate (i.e., it will be in *elongation phase*). As a result, there will be a buildup of GTP-bound beta-subunit at the plus end of the microtubule. This accumulation is known as a **GTP cap** at the growing end.
5. The lower affinity of GDP-bound tubulin dimers for the microtubule means that if growth slows down to the point that there are GDP-bound tubulin subunits exposed at an end of the polymer, the entire microtubule is more readily disassembled. The lower affinity of the GDP-bound monomers is due to a slight conformation change in the dimer, causing the protofilament to bend slightly (Figure 06-12 and Video 06-03).

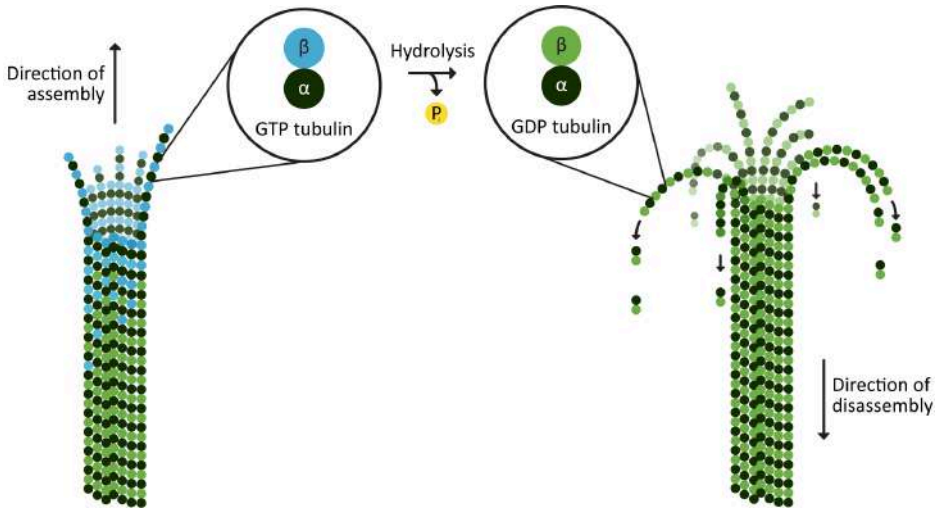


Figure 06-12: The cycle of assembly and disassembly of microtubules. GTP-bound tubulin dimers are added to a growing microtubule polymer. If the assembly is rapid, a section of GTP-bound tubulin (known as the GTP cap) forms at the growing end. If the end of the polymer contains GDP-bound tubulin, the conformation of the protofilament bends slightly, which favors disassembly. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-3

Video 06-03: Microtubule assembly with GTP cap and subsequent catastrophe event.

The loss of the GTP cap usually results in rapid shrinkage of the microtubule (Video 06-03). However, this may not be the end of it. If there is a high enough local concentration of activated tubulin, the microtubule can be rescued and then start to grow again (Figure 06-12). This means that if we watch microtubules over time, they are constantly changing length. This is the basis for the idea of dynamic instability, and it can be observed in fluorescence microscopy (Video 06-04).

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-4

Video 06-04: Assembly/disassembly of fluorescent microtubules in a live cell.

Microtubule Function

Microtubules have a wide variety of functions in the cell. Their most prominent ones include the following:

- acting as the "highways" on which vesicles and other cargo are transported,
- helping control and maintain specific cellular shapes,
- providing sites to anchor complexes and proteins in specific regions of a membrane,
- helping maintain organelle shape and location,
- coordinating the division of the DNA during mitosis, and
- forming the structural basis for eukaryotic cilia and flagella.

Each of these functions will require precise control of the microtubules so that they are always in the right place, at the right time, doing the right thing for the circumstances. However, if growth and shrinkage are dependent on conditions and protein chemistry, then how is the cell able to control this? The answer is that *the cell controls the conditions in order to promote growth, maintenance, or shrinkage as required*. In this section, we discuss the most prominent ways that this is done by the cell.

Microtubule Organizing Centers (MTOCs) Promote Growth

In many cells, especially animal cells, microtubules are observed growing from a specific site, at or near the center of the cell. This is known as a **microtubule organizing center (MTOC)** (Figure 06-13). In animal cells, it is also sometimes called the **centrosome**. MTOCs are sites where microtubule growth is specifically promoted through the use of a third type of tubulin, known as **gamma-tubulin**.

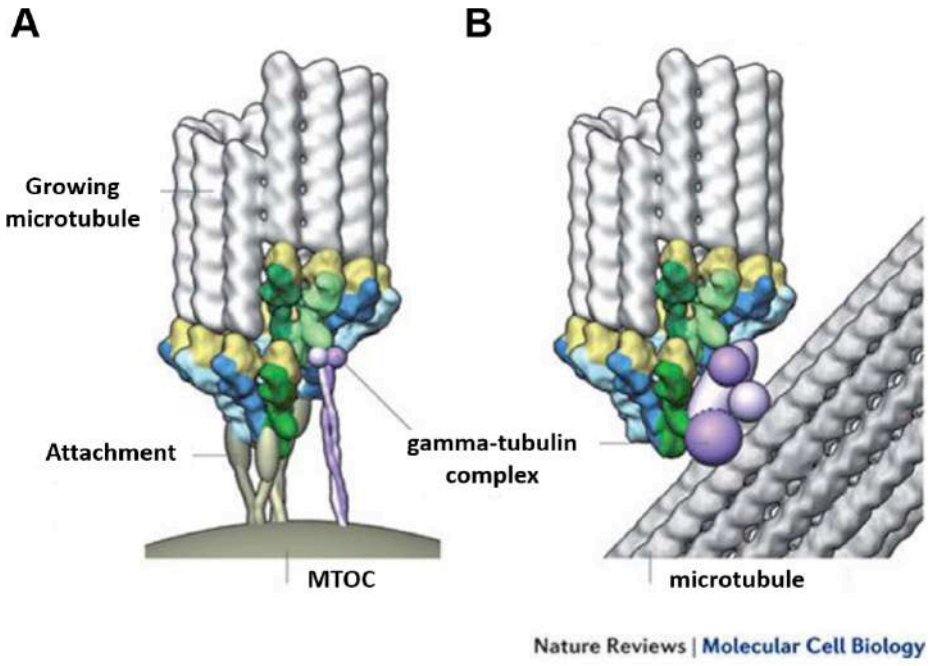


Figure 06-13: Gamma-tubulin serves as a nucleation point for microtubules. This helps promote microtubule growth at specific sites within the cell. (A) A gamma-tubulin assembly that is part of an MTOC, as is common in animal cells. (B) A gamma-tubulin assembly attached to the side of another microtubule, as is common in plant cells, since they do not have a centralized MTOC in interphase. This relabeled image is originally from Kollman et al. (2011) and is included under fair use as described in the CBPFUOER. Please do not redistribute without permission of the rights holder.

Gamma-tubulin helps promote microtubule growth by providing a premade template for the tubulin subunits to bind to. Having this template reduces the "lag phase" of the growth curve from Figure 06-10, which is the slowest stage of microtubule formation. An additional result is that the critical concentration required to promote assembly is lowered at these **nucleation sites**. As such, microtubules will grow more easily at these sites than they will elsewhere in the cell. In this way, the gamma-tubulin complexes help the cell control not only *when* but *where* microtubules grow.

A side effect of this method for promoting microtubule growth, which is evident in Figure 06-13, is that the minus end of the microtubule is bound to the gamma-tubulin ring. As a result, in most cells, the minus end is not free for growth or shrinkage, and it is only the plus end that is used for dynamic instability. However, there are important exceptions to this rule, which will be highlighted in Chapter 8.

Figure 06-13 also highlights that there are two ways that the MTOC can be arranged in cells: centralized (as we see in most animal cells, some fungi, and during mitosis in virtually all cells; Figure 06-13A) or decentralized (as we see in plant cells and some fungi in interphase; Figure 06-13B). Let's talk about each of these a little bit more.

Many cell types use centralized MTOCs in interphase, while others do not. In mammalian cells, which have a centralized MTOC throughout their life cycle, we call it the **centrosome**. The centrosome usually includes a pair of **centrioles** and their associated proteins (Figure 06-14A). Centrioles are also made of tubulin dimers, and their role is somewhat murky in a lot of cases.

During **mitosis**, *all* eukaryotic cells have two centralized MTOCs, often called the *spindle poles*, which coordinate the proper placement of microtubules for the separation of the chromosomes (Figure 06-14B).

Eukaryotic cells with cilia and/or flagella will have additional MTOC called basal bodies. These are a pair of centrioles found at the base of each **cilium** or **flagellum** that promote the growth of the microtubules within the structure (Figure 06-14C). Cilia and flagella are fascinating structures that are also more dynamic than we give them credit for. For example, they are often resorbed before mitosis and then regrown afterward.

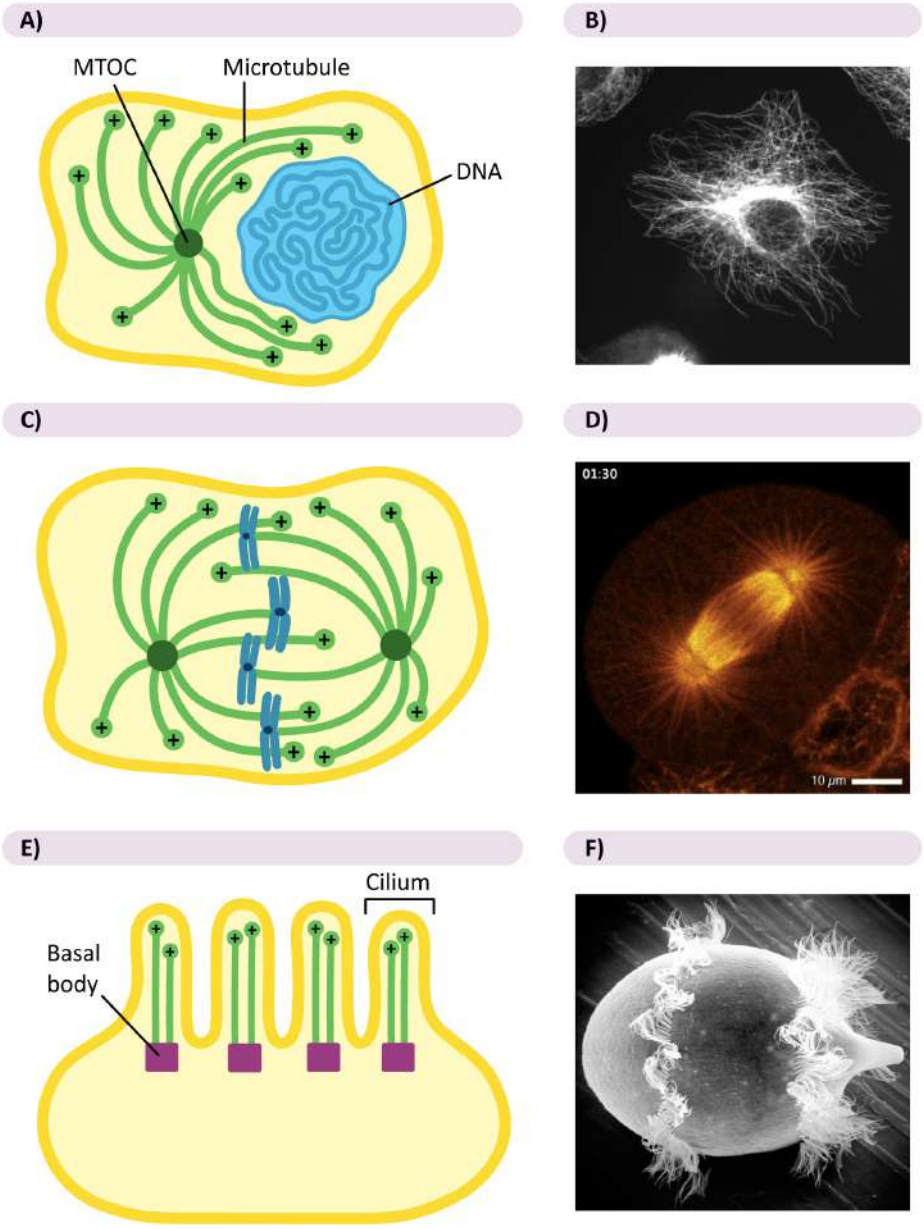


Figure 06-14: Examples of different kinds of microtubule organizing centers (MTOCs). Diagrammatic representations of MTOCs on the left with the corresponding example in an actual cell on the right. (A) An interphase animal cell is most commonly found with a centrosome near the center of the cell, with the microtubules radiating outward from there. (B) Fluorescent micrograph showing a microtubule stain in an unidentified animal cell. Micrograph by Jeffrey81 is shared under a CC BY-SA 3.0 license. (C) In mitosis, a mitotic spindle forms between two spindle poles, each of which is an MTOC. (D) Dividing cell in a sand dollar embryo (Dendraster excentricus), with fluorescently labeled microtubules. George Von Dassow (2011), CIL:15806, Dendraster excentricus. CIL. Dataset. https://doi.org/10.7295/W9CIL15806, shared under CC BY-NC-SA 3.0. (E) Cilia (pictured) and flagella grow from a basal body that is found directly below the plasma membrane and contains centrioles. These are also a form of MTOC. (F) SEM micrograph of an unidentified ciliated protist. The cilia can be seen in two rows down the sides. Gregory Antipa (San Francisco State University) (2011), CIL:39251, Didinium nasutum, eukaryotic cell, Ciliated Protist. CIL. Dataset. https://doi.org/10.7295/W9CIL39251. Diagrams created and image compiled by Heather Ng-Cornish, licensed under CC BY-NC-SA 4.0.

Decentralized MTOCs

- In plants and fungi, which do not have a centralized MTOC in interphase, the gamma-tubulin rings are generally carried along preexisting microtubules via motor proteins. With MTOCs present throughout the cell, microtubule growth is promoted evenly throughout the cell. The arrangement of microtubules in Video 06-04, above, is the product of decentralized MTOCs. It is a plant cell in interphase.
- During mitosis, cells that normally have decentralized microtubule growth will collect all of their gamma-tubulin into two MTOCs, which then act as spindle poles for the mitotic spindle so that cell division can take place.

Caps and Other Proteins That Control Microtubules

In order to harness the power of the microtubules and their capacity for dynamic instability, the cell uses a variety of proteins to control them. Some examples include the following:

- Nucleating proteins: Gamma-tubulin helps promote the growth of new microtubules.
- Stabilizing proteins: *XMAP215* helps stabilize preexisting microtubules and allows them to grow more efficiently.
- Severing proteins*: Katanin* is a protein that severs microtubules somewhere in the middle of the filament.
- Destabilizing proteins*: Kinesin 13* binds at the plus end and promotes disassembly.
- Monomer sequestration: *Stathmin* binds to unpolymerized tubulin dimers, which inhibits microtubule growth.
- MT organizing proteins: A variety of proteins, such as Tau and MAP2, are involved in building and stabilizing specific arrangements of microtubules known as **arrays**.
- Connecting proteins: *Plectin* binds to both microtubules and intermediate filaments, thus connecting the different cytoskeletal networks.
- **Protein caps**: Used to stabilize microtubules to stop them from growing and shrinking.
- **Motor proteins**: Used to help vesicles travel down the microtubule, among other specialized roles.

We aren't going to discuss all of these, as there simply isn't enough time. We will highlight only two: protein caps in this section and motor proteins in the next section.

Important note!

Capping proteins should not be confused

with the GTP cap. The GTP cap is an integral part of the microtubule consisting of tubulin dimers with GTP attached. Capping proteins are not made of alpha-beta tubulin subunits and are a separate entity from the microtubule.

Protein Caps

If a microtubule is required for a long period of time in the same place, it must be stabilized to reduce or eliminate dynamic instability. We've already seen an example of how this can be done at the minus end—the gamma-tubulin is acting as a "cap" that stops the microtubule from being able to depolymerize from that end. A protein cap can also be added to the microtubule at the plus end so that the microtubules cannot depolymerize from there either. Microtubules that are capped at both ends are generally quite stable and remain in place for as long as the cells need them.

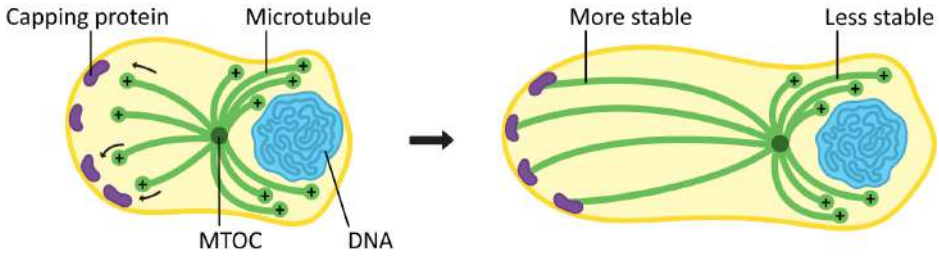


Figure 06-15: Microtubules help elongate and shape the cell as they develop. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Stabilization of microtubules can be an important part of cellular development. Microtubules bind to capping proteins at the cell surface to influence the final shape and structure of a cell, as seen in Figure 06-15.

In cells like axons, which have long projections that must be maintained, the stabilization of microtubules in the axon (in addition to the production and assembly of neurofilaments) is an important part of how the axons develop. Loss of the microtubules will have the same effect as loss of the neurofilaments in that the axon may start to retract. This will cause the cell to lose its connections to the rest of the neural network and die.

In addition to their role in stabilizing the axon, the microtubules in these long axons are required to transport vesicles and other cargo from the cell body to the plasma membrane at the axon terminal (Figure 06-16). Motor proteins are used for this, which we will describe next.

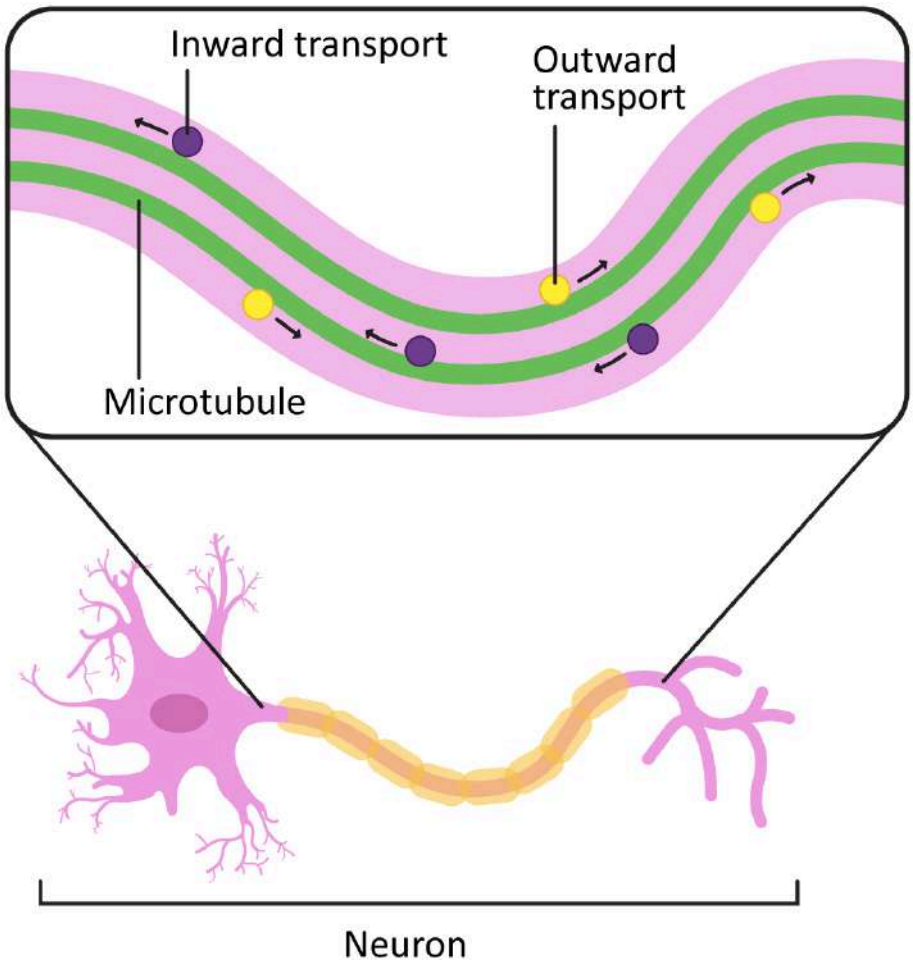


Figure 06-16: Schematic of a neuron, showing the arrangement of microtubules and the direction of vesicle movement within the axon. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Motor Proteins

The ability to move directionally along microtubules is a key requirement for many cellular functions, including cargo trafficking, organellar positioning, and mitosis. This is where the polarity of the microtubules becomes critical. Microtubule motor proteins have the ability to bind to the microtubule and move along it in a specific direction (using energy provided by ATP).

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-5

Video 06-05: Fluorescent kinesin walking along microtubules labeled in red. Yellow dots appear when kinesin attaches to the microtubule. You can track its movement over time, and eventually the fluorescence

goes away when the motor is released from the microtubule.

There are two classes of motor proteins that use microtubules to move (Figure 06-17):

1. **Kinesins**—most commonly move toward the plus-end of microtubules (see Video 06-05)
2. **Dyneins**—move toward the minus-end of microtubules

Since the microtubules is a **polar** molecule, due to the linear arrangement of the tubulin dimer, the motor protein can "walk" along the microtubule "highway," carrying its cargo toward the plus end (in the case of kinesins) or toward the minus end (in the case of dyneins).

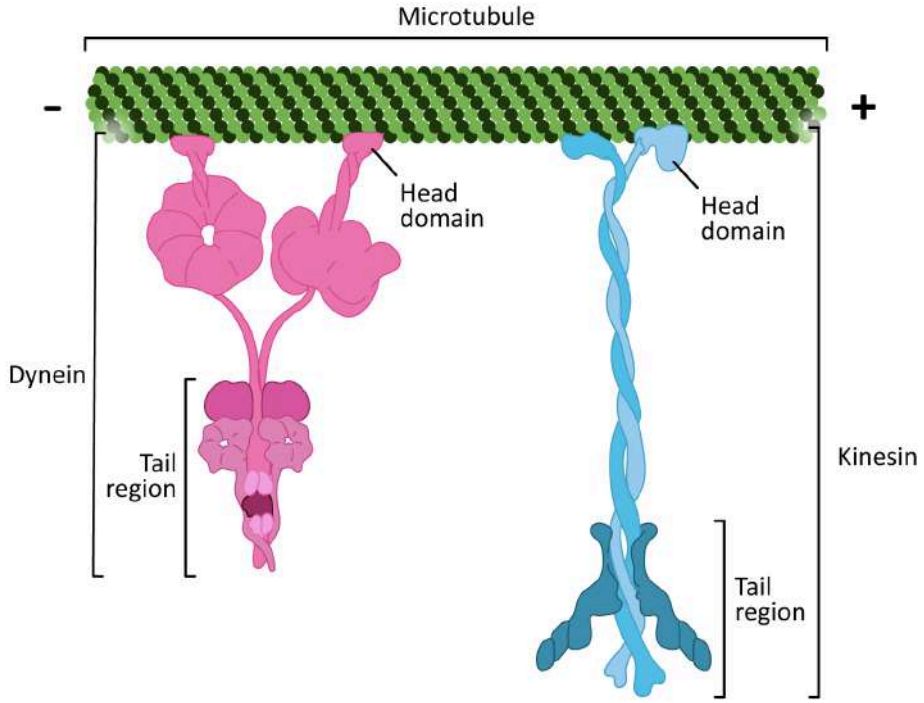


Figure 06-17: Structural representations of kinesin and dynein on a microtubule. The head domain binds to the microtubule reversibly (in an ATP-dependent manner). Through cycles of ATP hydrolysis, these motors are able to walk along the microtubule in a specific direction. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

There are some interesting similarities between the two classes of microtubule motor proteins (Figure 06-17). Kinesins and dyneins are both multisubunit complexes. They have globular head groups that attach to the microtubule. They also have a tail region that attaches to cargo. Both classes of motors are powered by ATP and have some kind of "stepping" cycle linked to ATP hydrolysis.

However, as you can see from Figure 06-17, their shape is very different. Their mechanism of action is also quite different. Kinesin "walks" along the microtubule, moving one motor head and then the other (Video 06-06). On the other hand, dynein appears to use more of a "strut"-like motion (Video 06-07).

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-6

Video 06-06: Kinesin motors "walk" along a microtubule.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-7

Video 06-07: Molecular model of dynein walking on a microtubule.

Motor proteins transport a number of different types of materials, including the following:

- vesicles
- proteins and/or protein complexes
- protein-mRNA complexes (also known as RNP particles)
- membrane-bound vesicles and organelles (including membranes of the ER and Golgi)
- other microtubules, as in flagella, cilia, and other motile microtubule arrays

Motor Proteins and Organelle Positioning/Maintenance

One question that is usually overlooked when studying the organelles of the cell is "How did they get to be shaped like that?" Membranes are not going to spontaneously form the flattened pancake of the Golgi apparatus or the interconnected tubular structure of the ER. So how does that structure happen?

Another question we rarely consider is how the different organelles manage to maintain their position in the cell. They are subject to gravity, like everything else on the planet, so why aren't the nucleus, mitochondria, and the rest of the organelles sitting on the bottom?

The answer to both of these questions is the same: molecular motors. The molecular motors have an important role to play in both the maintenance of structure and the location of the organelle within the cell.

Organelles have several motors attached to them. They likely have both dynein and kinesin, as well as the actin motor, myosin. These will work together to properly position all of the parts of the organelle in the 3D space of the cytosol. Let's consider two membrane-bound organelles:

1. The Golgi apparatus maintains a position near the nucleus in mammalian cells. This is also near the interphase centrosome. Dyneins have been shown to play an important role in the

maintenance of this position. When microtubules are experimentally depolymerized, the Golgi has been observed to fragment and float away from its original position.

2. On the other hand, the endoplasmic reticulum, which extends throughout the cytoplasm in most cells, has a different reaction when microtubules (MTs) are experimentally depolymerized. In this case, the ER has been shown to become more spherical and retract toward the nucleus. The ER tubules are moved outward by plus-end kinesin motors, walking toward the plus end of microtubules.

TOPIC 6.3: ACTIN FILAMENTS

Learning Goals

- Compare and contrast dynamic instability using actin as the monomer.
- Correlate *in vivo* microfilament polymerization and organization at the cell cortex with its function of deforming the plasma membrane and regulating cell shape.
- Relate how actin-binding proteins influence actin filament polymerization and organization to regulate filament function (e.g., cell motility).
- Interpret results from experiments using drugs that disrupt microfilaments and compare with equivalent experiments that disrupt microtubules.
- Explain the function of myosin motors and their role in contractile bundles as well as in other cellular contexts, such as cytoplasmic streaming or muscle contraction.

The final cytoskeletal element that must be discussed is **actin filaments** (also known as microfilaments, filamentous actin, f-actin, or just plain "actin"). In animal cells, these filaments are major components of the peripheral cytoskeleton, the **cell cortex**, and complex contractile systems such as muscle.

RuBisCo is thought to be the most abundant protein on the *entire planet*, so it beats actin in photosynthetic eukaryotes
In plants, they are found throughout the cell and are crucial for vesicle and organellar transport. They are also involved in a process called **cytoplasmic streaming**, which is when a current is created in the cytosol of the plant by the movement of vesicles and other organelles along actin filaments. The movement of the cellular components acts a bit like a whirlpool, such that proteins and other structures that are not attached to actin can also be carried along passively. Actin is thought to be the most abundant protein in a nonphotosynthetic eukaryote.
Actin filaments can form *either* stable or dynamic structures, depending on how they are cross-linked to other proteins. As a result, the array of **actin-binding proteins (ABPs)** that exist are considered extremely important to understanding actin function.

Actin Structure and Filament Formation

Like microtubules, actin filaments are a polymer of actin subunits (sometimes called actin, actin monomers, globular actin, or **g-actin**). The subunits line up end to end and polymerize in such a way that different sides of the subunit are exposed at either end (Figure 06-18). This means that the filament has a directionality to it, so actin filaments are also referred to as polar molecules (once again, this has nothing to do with charge or other chemical properties per se).

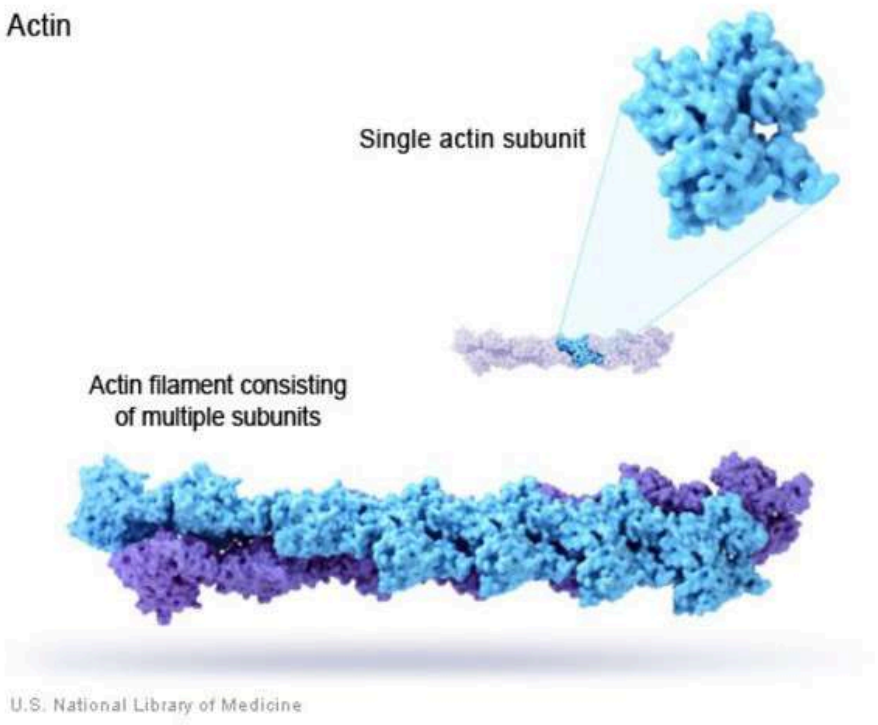


Figure 06-18: The molecular structure of actin filaments. Actin proteins polymerize into a filament with a twisted shape. Image source: MedlinePlus, National Library of Medicine. This image is in the public domain.

Like microtubules, actin also is capable of dynamic instability even though actin researchers do not commonly discuss actin in those terms. This is, in part, due to the large group of ABPs that are used to control actin monomers and filaments. As a result, the functions of actin filaments and microtubules are very different, as you will see throughout this topic.

Actin and Dynamic Instability

It's important to note here that *in vitro* (i.e., in a test tube) actin and tubulin behave very similarly. This means the following:

- They both undergo dynamic instability.
- The plus end has a higher affinity for monomers than the minus end, resulting in
    - a lower critical concentration and
    - an increased rate of assembly once polymerization starts.

Like tubulin, actin uses chemical energy to drive the process of polymerization (Figure 06-19). However, in this case, ATP is used instead of GTP. In a living cell (i.e., *in vivo*) there is a high concentration of both ATP and GTP in the cytosol. Thus, it is relatively straightforward to keep the actin monomers in their "active," ATP-bound state, which favors polymerization. Video 06-08 shows a molecular model of the polymerization of actin.

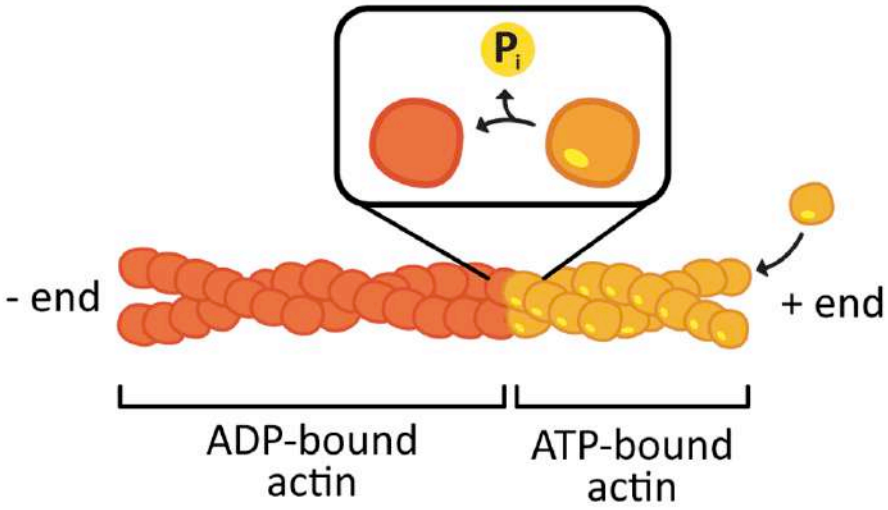


Figure 06-19: Actin also undergoes dynamic instability. ATP is used to drive polymerization. Since polymerization is faster at the plus end compared to the minus end of the filament, we can get the formation of an "ATP cap," which helps hold the filament together and promote continued growth. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-8

Video 06-08: Actin polymerization occurs by addition of monomers forming a helical structure.

Actin Function

Actin-Binding Proteins (ABPs)

The cell uses many of the same strategies to control actin filament formation, as we saw in the section on microtubules. Proteins known as ABPs are used extensively to control where and when actin filaments form in the cell. In fact, far more proteins exist to control actin than have been found to control microtubules. This is due, in part, to the number of highly specialized roles that actin plays in different cells. Of course, actin filaments have different functions in the cell compared

to microtubules, so the types of binding proteins and the structures formed by actin filaments are distinct from those of microtubules.

Similar to microtubules, actin formation is modulated by the following:

- *Controlling where an actin filament forms:* Decentralized nucleation sites provide locations to polymerize actin filaments easily. Unlike microtubules, there are no examples of a centralized organizing center for actin. In actin, there are two major types of actin nucleation sites:
    1. Branching networks—formed when **Arp2/3** binds to the side of a preexisting actin filament. We will see this when we discuss cellular locomotion.
    2. Parallel networks—formed when *formin* and *profilin* bind to actin.
- *Stabilizing actin filaments using protein caps:* Many of the structures formed by actin must be maintained for the long term. ABPs bind to the end of the filament and act as caps to prevent disassembly (Figure 06-20). Examples of these long-lasting, stable actin structures include the microvilli, found in the intestinal tract, and the actin filaments that form the basis for muscle contraction (along with the actin motor protein, myosin).

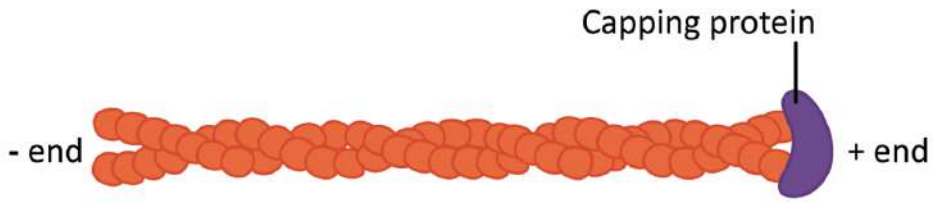


Figure 06-20: Capping of actin by ABPs helps stabilize actin filaments. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

By controlling where actin forms and when/where it gets stabilized, ABPs and actin can form several different types of actin networks in the cell (known as **arrays**; Figure 06-21):

- *Parallel bundles*—filaments are closely spaced and have the same polarity (directionality). For example, parallel bundles are found in filopodia or microvilli. The proteins formin and profilin build these arrays with additional proteins to cross-link.
- *Contractile bundles*—filaments are arranged antiparallel and are cross-linked by stabilizing proteins. Molecular motors drive the "contraction" of these arrays. Examples include stress fibers in cells and in the contractile ring that splits apart daughter cells during cytokinesis.
- *Cross-linked gel*—random orientation of fibers linked at the crossing by filamin. Actin gel is a major component of the cell cortex of most cell types and plays a big role in locomotion. Arp2/3 is used to build the filaments in these arrays, and additional proteins are used to cross-link and stabilize.

In association with these internal arrays, actin has also been observed to attach to a variety of plasma membrane proteins, which allows it to help shape the cell and to respond to the cellular environment. Actin is thought to play a key role in cell shape for all cells, but in animal cells, these arrays are the primary driver for most cell shapes.

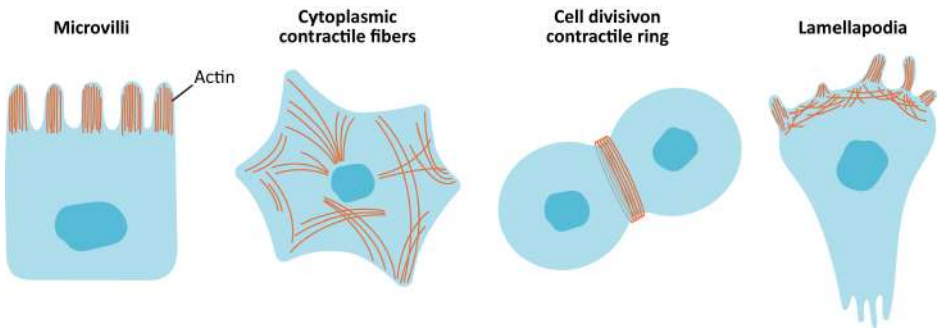


Figure 06-21: Example cell structures formed by actin arrays in animal cells. The actin filaments are shown in orange. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Myosin Motor Proteins

Just like microtubules, actin filaments also use motor proteins to move cargo along the actin filament. **Myosins** are a large and diverse class of motor proteins that use actin filaments as their substrate.

Myosin motors are involved in a number of cellular functions (Figure 06-22), including the following:

- vesicle transport (Myosin I is the most important myosin for this, but others also get involved.)
- organelle positioning
- cytokinesis
- cellular locomotion and shape
- phagocytosis
- muscle contraction (Myosin II: We know quite a lot more about this myosin than any of the others, as quite a lot of research is done on muscle function.)

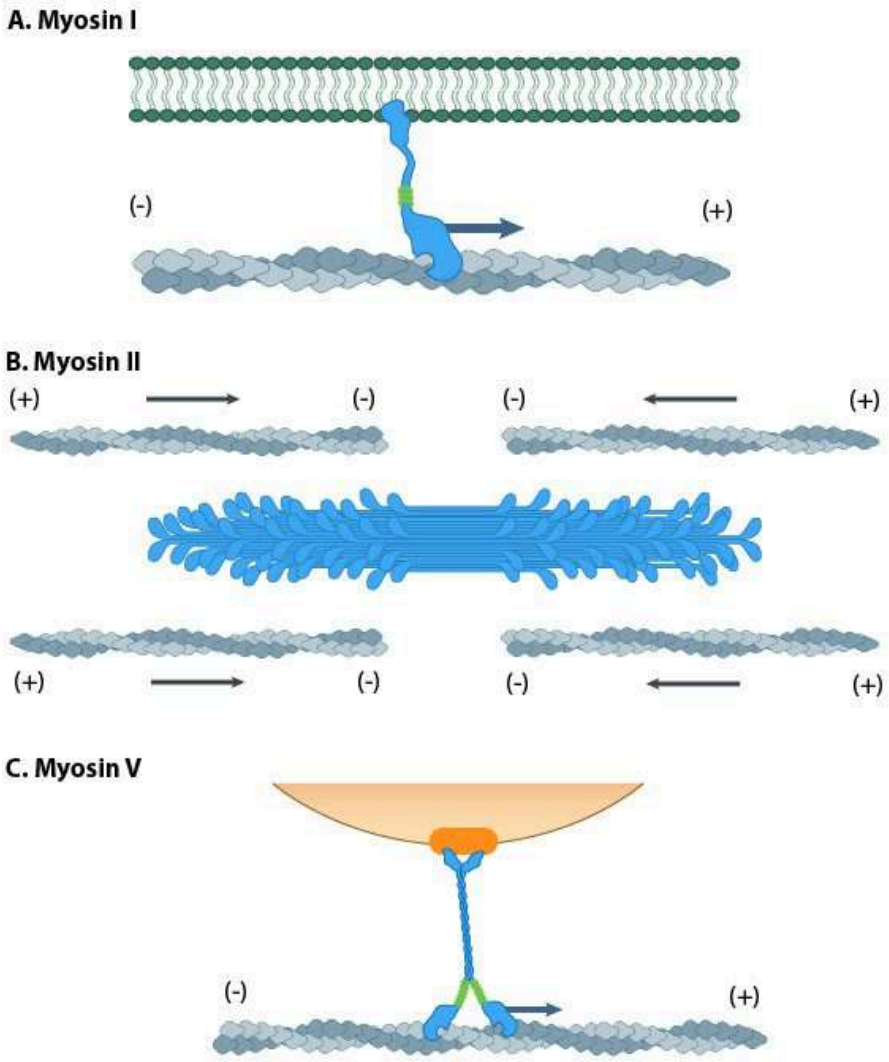


Figure 06-22: Examples of different functions of the myosin family. (A) Myosin I can bind to membrane lipids. (B) Bundles of Myosin II slide along the actin cytoskeleton network to drive actomyosin contractility. (C) Myosin V transports cargo by "walking" along actin filaments. Image appears in MBInfo and shared under a CC BY-NC 4.0 license.

The myosin family of proteins is very large. There are as many as 18 different types of myosin motors found in eukaryotes. Most of these myosins move toward the plus-end of actin filaments; however, there are some exceptions. One subtype, called myosin IV, has been shown to move toward the minus-end. Myosin motors are complexes of several polypeptide chains (Figure 06-22). Some subtypes have one motor head, and some have two. However, the motor head functions the same way regardless of how many there are.

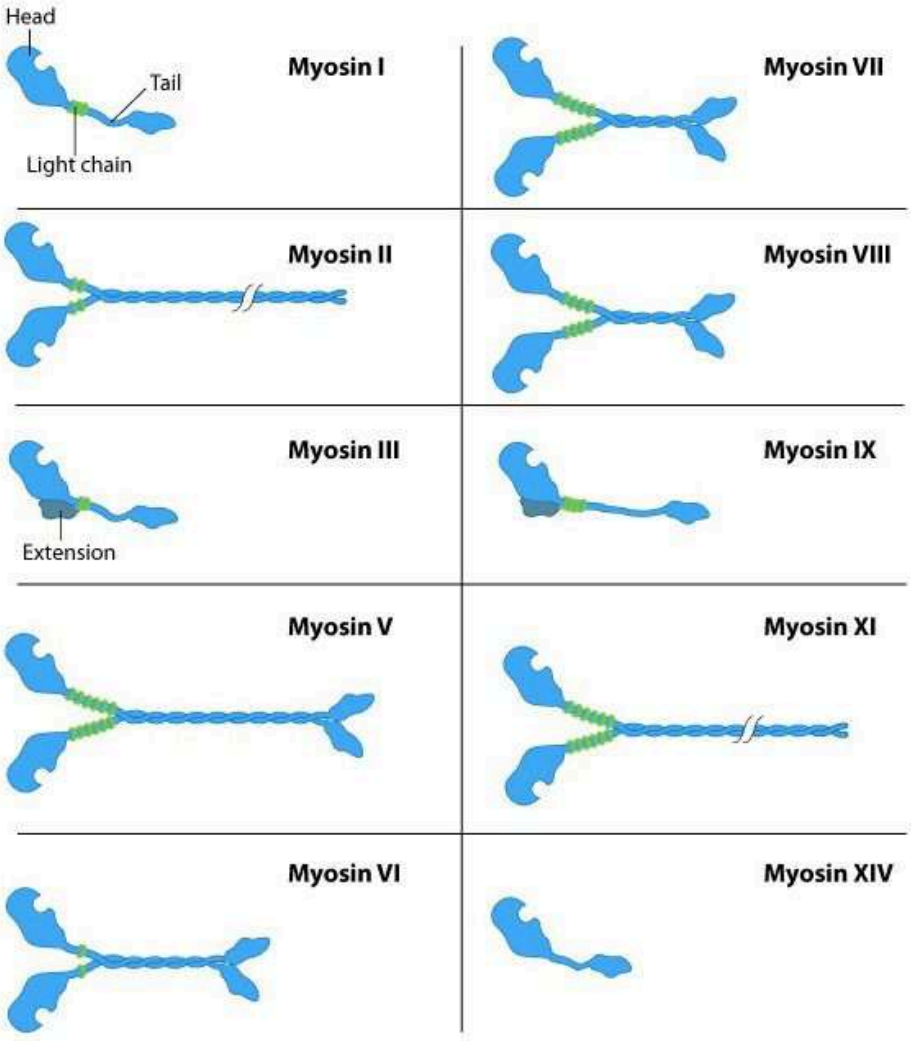


Figure 06-23: Some examples of the different subfamilies of myosin proteins. Image appears in MBInfo and shared under a CC BY-NC 4.0 license.

Sadly, we do not have time to further study muscle function here. Muscles are a highly specialized cell type and thus not representative of the majority of eukaryotic cells. Those who are interested in muscle function are encouraged to take upper-level animal biology / human physiology courses.

In a nutshell, they bind to the subunits of the actin filament, pull their motor head in, and then release with the help of ATP. Different myosins appear to move slightly differently, so it's more difficult to find a single video to represent them all. Muscle myosins (myosin II) work through a cycle that includes a "power stroke" that pulls the actin filament in toward the center of the muscle (Video 06-09). Myosin V, on the other hand, appears to use a very different mechanism that is similar to the movement of a Slinky toy (Video 06-10).

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-9

Video 06-09: Myosin binding to actin as part of muscle contraction.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-10

Video 06-10: Myson V interacting with an actin filament. Animation of model from Erin M. Craig and Heiner Linke, Mechanical model for myosin V, *Proceedings of the National Academy of Science, 106*(2009), 18261–66,https://doi.org/10.1073/pnas.0908192106.

Like the microtubule motors, myosin motors require ATP in order to function. The ATP hydrolysis helps prime the motor head to prepare for the next time it binds and pulls on the filament.

Interestingly, no ATP is required for binding to the actin filament, but every other step of the process requires either ATP or ADP. This is the opposite of the microtubule motors, which cannot bind to the microtubule (MT) unless they are bound to ATP.

As a result of this unique situation, myosin motors will freeze in place if the cell runs out of ATP, which happens at the point of death or shortly thereafter. Muscle contraction is dependent on successive rounds of binding and unbinding of myosin, so without ATP, the myosin binds and causes the muscles to become really rigid right after death. This is a well-known phenomenon called *rigor mortis*. It is commonly referenced in detective shows, especially those with a focus on forensic science.

Example: Cellular Locomotion

Locomotion in animal cells is a good example of how precise control of dynamic instability can be used to the cell's advantage. Cellular locomotion is a complex interplay of holding on to the substrate, building a network in one area of the cell, and deconstructing the network in another area (Figure 06-24). Every step of this process requires ABPs to make sure that it all goes according to plan.

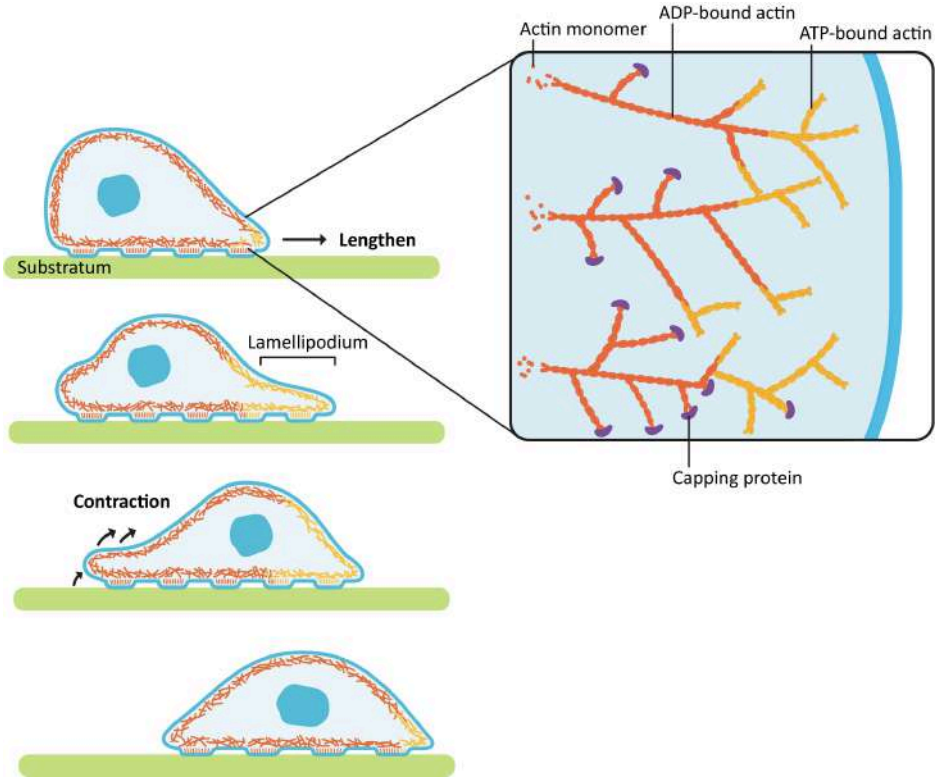


Figure 06-24: How cells crawl. An actin array is formed at the leading edge of the cell. This pushes the membrane forward. Subsequent depolymerization of actin on the opposite side allows the cell to contract in the direction of movement. Capping proteins help stabilize the new branches. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The act of forming the membrane protrusion at the **leading edge** requires actin polymerization directly underneath the plasma membrane. Nucleation of new actin filaments at the leading edge is mediated by Arp2/3 complexes. As you may recall, Arp2/3 binds to the side of a preexisting filament so that a branched array is formed (Figure 06-24). The growing actin network is what pushes the leading edge forward, creating membrane waves, or ruffles, called a **lamellipodium** (plural lamellipodia). At the same time, the polymerized actin at the back end of the leading edge (i.e., on the far side of the branched network from the membrane) is depolymerized so that free actin can be moved forward. This ensures that the cell continues to have access to free actin for polymerization at the leading edge.

As the leading edge is being pushed forward by the lamellipodia, the back of the cell also needs to be pulled forward, so the cell moves forward instead of just flattening out. This contraction is accomplished by a network of actin and myosin motors (sometimes referred to as the *actomyosin* network) that contract within the cell body, creating tension. As it does this, the proteins holding onto the substratum at the rear release so that the contraction can happen, and the cell moves forward.

Cellular locomotion is well illustrated by Video 06-11, filmed in the 1950s using a light microscope attached to a film camera. In it we see a white blood cell chasing, capturing, and engulfing a bacterium.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=153#oembed-11

Video 06-11: A very famous movie from the 1950s of a neutrophil (which is a type of white blood cell) chasing a bacterium. The original video was shot on 16 mm film in the 1950s by David Rogers. More information: https://embryology.med.unsw.edu.au/embryology/index.php/Movie_-_Neutrophil_chasing_bacteria.

Studying Cells: Drugs That Influence Actin and Microtubule Function

Before we leave our discussion of the cytoskeleton, we must once again explore how scientists study the function of the cytoskeleton. As we've seen before, a commonly used tactic in biology is to disrupt a biological system and observe the results. You may not be surprised to know that both actin and microtubules are essential cellular structures. So mutational analysis is not a very good option for experimental study.

However, since the cytoskeleton is so important, especially as a key player in cell division and muscle contraction, it has become the target of many natural toxins produced by a variety of plants, animals, and fungi in order to protect themselves from other organisms in their environment. As scientists, we are able to exploit these compounds to our advantage and use them to temporarily disrupt the cytoskeleton and then observe the impact on cellular function. Table 06-01 shows the most commonly used compounds and how they impact the cytoskeleton when applied.

Table 06-01: Chemical compounds that impact microtubules or actin filaments

CHEMICAL COMPOUNDS THAT IMPACT MICROTUBULES OR ACTIN FILAMENTS

Microtubule-specific

Compound	Action	Source
Paclitaxel	Binds to microtubule and prevents disassembly	Bark of Pacific yew (*Taxus brevifolia*)
Colchicine	Binds to unpolymerized tubulin and prevents assembly	Autumn crocus (*Colchicum autumnale*)
Vinblastine	Binds to unpolymerized tubulin and prevents assembly	Madagascar periwinkle (*Catharanthus roseus*)
Oryzalin	In plants only; binds to unpolymerized tubulin and prevents assembly	Synthetic herbicide in the dinitroaniline family

Actin-specific

Compound	Action	Source
Phalloidin	Binds to actin filament and prevents disassembly	Death cap mushroom (*Amanita phalloides*)
Latrunculin	Binds to unpolymerized actin and prevents assembly	Sea sponges in the *Latrunculia* and *Negombata* families
Cytochalasin B or D	Caps the plus end of actin filaments	Fungal plant pathogen of the *Helminthosporum* genus

What we hope you will notice in the table is how varied the effects of the drugs are. While some of them interact directly with the polymerized microtubule, others work by disrupting the chemical equilibrium through interactions with the monomer. This shows how essential dynamic instability is to the proper function of both actin and microtubules.

One final note: There is often confusion between the difference in "capping" and "stabilizing," especially with respect to the function of these compounds. A *capped filament* is also stabilized, as the polymer cannot fall apart. But it is possible to *stabilize without capping*. A perfect illustration of this is the difference between paclitaxel, which stabilizes microtubules without capping so that they may still grow (but not shrink), and something like cytochalasin, which acts as a protein cap at the plus end of the actin filament. Cytochalasin stops the microtubule from growing or shrinking, whereas paclitaxel only inhibits shrinking but not growth.

CHAPTER SUMMARY

By now, we hope you agree that the cytoskeleton is an intriguing cellular system that helps cells to take on very specific functions. If not for the cytoskeleton, cells would largely be shaped like amorphous blobs. In this chapter, we talked about the three major types of cytoskeletal elements, each with its unique mechanism of formation and regulation to harness these structures to meet cellular needs.

Intermediate filaments, so called because they are intermediate in size, are the only nonpolar cytoskeletal element. This means that the two ends of the filament have no identifiable differences between them. Because they are nonpolar filaments, motor proteins don't associate with them due to the lack of directional cues. These ropelike filaments provide tensile strength to help cells withstand mechanical strain.

Microtubules are the largest diameter cytoskeletal elements. These are hollow tubes formed from tubulin monomers, which themselves are made of an alpha and a beta-subunit. Assembly of microtubules is spontaneous, and these filaments will continuously assemble and disassemble based

on a number of different factors. This dynamic instability allows microtubules to form and reform quickly, which makes them quite easy to rearrange. Microtubule-associated proteins (MAPs) are used to control when and where microtubules assemble and when they are disassembled. Microtubules are responsible for a variety of cellular processes, including vesicle trafficking, chromosomal segregation in mitosis, organelle positioning, and specialized motility structures like cilia and flagella.

Finally, actin filaments are the smallest diameter filament. Like microtubules, they undergo dynamic instability and are influenced by the binding of cellular proteins to influence stability of the filament as well as where and when these can form within a cell. These filaments are also polar; have a wide variety of proteins that control them, known as actin-binding proteins (ABPs); and associate with motor proteins. Actin filaments are responsible for functions like cellular locomotion, muscle contraction, cytoplasmic streaming, cell shape, and intracellular transport.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

Topic 6.1: Introduction and Intermediate Filaments

1. Fill out the following table to compare and contrast the different components of the cytoskeleton. Be as specific and detailed as you can.

Table 06-02: Components of the cytoskeleton

Topic...	6.1 (Intermediate filaments)	6.2 (Microtubules)	6.3 (Actin filaments)
Monomer?			
Structural polarity (Y/N)			
Dynamic instability (Y/N)			
Energy source?			
Motor proteins?			
Function(s)?			

2. Describe the organization of intermediate filaments going from the level of single polypeptides to the microscopically visible filaments that are found in cells.
3. Explain how the structure of intermediate filaments contributes to their strength.
4. How do intermediate filaments contribute to the integrity of epithelial tissues?
5. Compare and contrast nuclear lamins with other known intermediate filaments, such as keratin or vimentin. Make a list that includes structure, function, and distribution of these different types.
6. Thought problem: There are no known motor proteins that use intermediate filaments as substrate. Why might that be?

7. Compare the organization of intermediate filaments and chromatin. What can you say concerning how one goes from molecules to microscopically visible structures?

Topic 6.2: Microtubules

1. Explain how tubulin dimers interact noncovalently to make microtubules with a "plus" and a "minus" end.
2. What is meant by polarity in microtubule structure? Explain what is meant by the "plus" and "minus" end of the microtubule. (Hint: It has nothing to do with charge or hydrophobicity.)
3. What are nucleation sites, why are they needed, and what is their role in microtubule formation?
4. What are microtubule organizing centers (MTOCs)? What do they do?
5. What does dynamic instability mean? How is GTP involved?
6. Explain how the different microtubule motor proteins work and how their movement relates to the polarity of the microtubule.
7. Describe how microtubules are arranged in animal cells and how their arrangement contributes to the structure and function of the endomembrane system. How might motor proteins contribute?

Topic 6.3: Actin Filaments

1. Compare and contrast the polymerization of actin with that of microtubules.
2. Compare and contrast the motor proteins associated with microtubules and actin microfilaments with regard to general properties, structure, and function.
3. Explain how animal cell motility is dependent on the dynamic instability of actin filaments.
4. Describe how associated proteins modify the function of both actin filaments and microtubules.
5. Drugs are often used to disrupt normal cytoskeletal function in experiments. Explain how this might be used to learn about actin and/or microtubule function.

REFERENCES

Kollman, J. M., Merdes, A., Mourey, L., & Agard, D. A. (2011). Microtubule nucleation by γ-tubulin complexes. *Nature Reviews Molecular Cell Biology*, *12*, 709–721. https://doi.org/10.1038/nrm3209

Maggi, L., Mavroidis, M., Psarras, S., Capetanaki, Y., & Lattanzi, G. (2021). Skeletal and cardiac muscle disorders caused by mutations in genes encoding intermediate filament proteins. *International Journal of Molecular Sciences*, *22*(8), article 8. https://doi.org/10.3390/ijms22084256

Mitchison, T., & Kirschner, M. (1984). Dynamic instability of microtubule growth. *Nature*, *312*(5991), 237–242. https://doi.org/10.1038/312237a0

Prokocimer, M., Barkan, R., & Gruenbaum, Y. (2013). Hutchinson–Gilford progeria syndrome through the lens of transcription. *Aging Cell*, *12*(4), 533–543. https://doi.org/10.1111/acel.12070

Roostalu, J., & Surrey, T. (2017). Microtubule nucleation: Beyond the template. *Nature Reviews Molecular Cell Biology*, *18*, 702–710. https://doi.org/10.1038/nrm.2017.75

Walker, R. A., O'Brien, E. T., Pryer, N. K., Soboeiro, M. F., Erickson, H. P., & Salmon, E. D. (1988). Dynamic instability of individual microtubules analyzed by video light microscopy: Rate constants and transition frequencies. *Journal of Cell Biology*, *107*, 1437–1448. https://rupress.org/jcb/article-pdf/107/4/1437/1461571/1437.pdf

CHAPTER 7.

CELL SIGNALING

INTRODUCTION

In order to survive, a cell must be able to understand its environment. This is true whether the cell is a single-celled organism or part of a larger, more complex multicellular organism. Cells communicate with their environment through a process called signaling. Cell signaling is how the cell collects information and then responds with an action at the correct time. Signaling is the initial event associated with many key cellular functions, from the correct timing of cell division, to the decision to migrate in a particular direction, and even to whether a cell needs to go through programmed cell death. So many of the cellular events we explore in biology are dependent on signaling to happen correctly. Not only that, but many of the concepts covered in upper-level biology courses are, at their hearts, studies of how the cell receives information and responds to it. For example, developmental biology, sensory perception, endocrinology, and even physiology will make much more sense if you have a foundational understanding of how signaling works.

Thus, our focus in this chapter is to dissect the process of signaling and take a look at the parts of a signaling cascade and some of the commonly recurring themes and patterns. That way, when you encounter any pathway in any context, you will have the ability to "read it" through and identify the patterns within it to correctly interpret the outcomes of that pathway.

TOPIC 7.1: GENERAL PRINCIPLES OF SIGNALING

Learning Goals

- Explain the three stages of a general signaling cascade.
- Provide examples of common protein families involved in each stage.
- List and explain the different types of intracellular responses possible for a given signal.
- Distinguish among relay, amplification, integration, and distribution steps in a signaling cascade.
- Correctly interpret schematic representations of signaling cascades.

While signaling is a beautiful coordination of cellular events, it also is a very complex process. It can be very confusing to keep track of what the cell is doing at any given time due to the sheer number of signals, the overlap in signaling pathways, and sometimes even the competing signaling events occurring simultaneously. Additionally, a wide range of cellular "behaviors" are mediated through

a small set of extracellular signals. Thus, *the way that a cell responds to a specific extracellular signal will depend on what genes are being expressed in the cell at that specific moment in time*. For example, acetylcholine is an extracellular signal that has different effects in different cell types. It is released at a neuromuscular junction by neurons to promote muscle contraction. In cardiac pacemaker cells, it signals that the heart rate should lower, so the pacemaker cells fire at a decreased rate. Finally, salivary glands also respond to acetylcholine by increasing the synthesis and secretion of saliva by the endomembrane system. Each of these scenarios is the result of different modes of release of the acetylcholine, different receptors receiving it on the cell surface, and also differential expression of the internal components of the acetylcholine pathway.

Our focus for this topic is to lay some groundwork for understanding signaling by looking at the general principles that underlie cell signaling. Then in the next topic, we'll look at some examples of actual mechanisms that are commonly used in cell signaling. These tools should allow you to interpret any signaling pathway in the future and begin to understand how they mediate a response.

The video below (Video 07-01) gives a quick overview of the mechanism of cell signaling and the terminology you will need to understand this process.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=156#oembed-1

Video 07-01: Introduction to cell signaling. This introduces much of the common terminology you will see in this chapter and some of the patterns as well.

Types of Cell Signaling

Before we start dissecting the various components of a signaling cascade, we must first understand the different forms of signaling that can occur. There are five different types of signaling that are common in cells: **endocrine**, **neuronal**, **paracrine**, **autocrine**, and **juxtacrine** (Figures 07-01 and 07-02, as well as Video 07-01). A multicellular organism (such as ourselves) is likely to use most or all of these different forms in different tissues/scenarios. Comparatively, single-celled organisms are more limited in their options, since they do not specialize, nor do they require communication as frequently between the different cells.

The different types of signaling modes differ in both the distance the signal travels and the speed with which it travels from the site of signal release to the site of **receptor** binding. They will also differ in the **affinity** that the receptor will have for the signaling molecule (also known as a **ligand**). The affinity of a ligand for its receptor refers to how "well" the ligand binds to the receptor. A receptor that is regularly flooded with a high concentration of ligands will not need to be as sensitive as one that must find a single molecule in the vast sea of the extracellular environment. Receptors that are more sensitive are said to be high affinity compared to those that are less sensitive to the presence of

their ligand. It is also true that high-affinity receptors bind more tightly to their ligands and are less likely to uncouple after binding than lower affinity receptors.

Long-Distance Signaling: Endocrine (Slow) and Neuronal (Fast) Signaling

Signaling over long distances can be challenging, as there may be hundreds, if not thousands, of cells between the source of the signal and the cell that must receive it. Both plants and animals are able to transport signals over long distances; however, animals are the true experts at long-distance signaling, with both a fast and a slow mechanism to do so. Plants, on the other hand, mostly focus on short-range signaling. They have very few options for long-distance signaling, none of which are considered to be terribly fast. Endocrine and neuronal signaling are illustrated in Figure 07-01.

Endocrine signaling uses **hormones**, which are sent throughout the organism, as the initial signaling molecule. In animals, this is done via the bloodstream. Using the bloodstream for transport allows the signal to be produced at a single site in the organism (e.g., the adrenal glands), but it can have an effect in many places that are far away:

- Endocrine signaling is considered to be *slooow*, as the signal must be produced and secreted into the bloodstream, moved throughout the body, and then picked up by another cell that is likely very far from the site of ligand release. It can take minutes for the signal to be received, which is quite a long time in the world of cells and signaling.
- On the other hand, since the hormone will become very dilute as it moves through the body and the receptor must be able to find and bind the ligand even in these low-concentration conditions, the receptors for endocrine signals are quite sensitive. In some cases, a single molecule from the bloodstream can be detected. A receptor that can be activated by a small, dilute dose of signal is said to have *high affinity* for its ligand.
- Examples of endocrine signaling ligands include most hormones (i.e., insulin, adrenaline, estrogen, growth hormones, etc.).

Neuronal signaling is the type of signaling used by the nervous system. In a nutshell, an electrochemical signal is sent through our neurons, across large distances, to elicit a response:

- This kind of signaling is quite *fast*, which is good, considering that neuronal signaling is what makes you move your hand when you touch a hot surface by accident. The signal is passed along and received in a matter of milliseconds.
- Quite a bit of this process happens within the nerve cell. Nerves are very long cells—as an example, your sciatic nerve starts at the base of your spine and ends in your foot and is about 1 m long! Keeping the signal inside the cell for as long as possible makes it much easier for the signal to move quickly.
- At some point the signal will need to exit the neuron and be passed to the muscle or other tissue that is expected to respond. Multiple chemical signals are used to help with this part. Collectively, we call these signals **neurotransmitters**. You likely have heard of many of them: dopamine, epinephrine (also known as adrenaline), and acetylcholine, for example. Some recreational drugs act by affecting the ability of neurons to send and/or receive neurotransmitter signals.

- Neurotransmitters help the signal move from the neuron to the target cell (like a muscle cell). The gap between the nerve cell and the muscle cell is really small, so the signal doesn't have to diffuse very far. To increase the chances that the signal will be received as quickly as possible, the nerve cell floods the entire area with the neurotransmitter. As such, these receptors are not as sensitive as those used for endocrine signaling…they don't really have to be. Thus, we consider the receptors in the synapse to be *low affinity*.

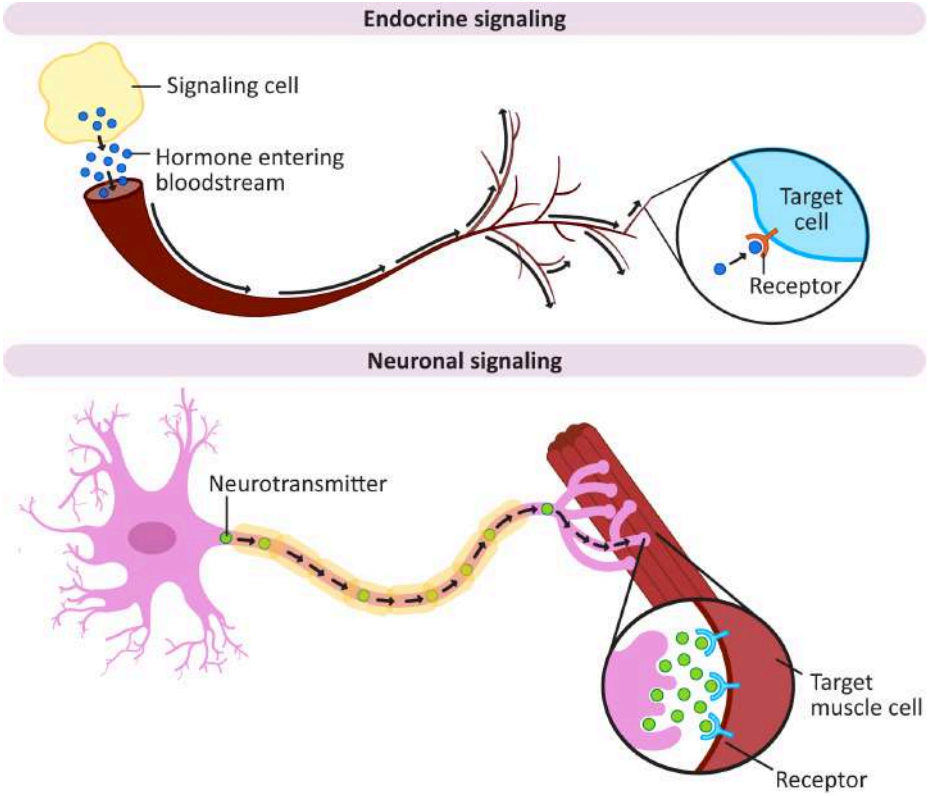


Figure 07-01: Long-distance signaling—endocrine and neuronal signaling schematics. In endocrine signaling, the signal molecule is transported long distances. In the image example, the blood transports the signal molecule away from the signaling cell to the target cell, where it binds to the receptor on the target cell. Neuronal signaling is a long-range signaling process where an electrical signal is transported down the axon, which can be extremely long. The electrical signal is converted into a chemical signal when neurotransmitters are released into the synaptic junction, where they can bind to the target cells (a muscle cell in this example). This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Medium- to Short-Distance Signaling: Paracrine (Diffusion Based), Juxtacrine (Contact Dependent), and Autocrine (Self-)Signaling

In many ways, signaling at closer range is much easier than long-distance signaling, as the signal can be released into the extracellular space and simply left to diffuse. This is one of the most common types of signaling in development. Plants, algae, and fungi don't really have any long-distance options that are as efficient as the bloodstream or neurons, so for them, almost all signaling is local and diffusion based. In addition, there are a number of examples of cells that release their own signaling molecules and then detect them with receptors. Here we look at the three types of short-range signaling in more detail. They can all be observed in Figure 07-02, below.

Paracrine signaling is considered to be a local signaling mechanism. The ligand is released into the extracellular space, and it diffuses through the extracellular matrix to be picked up by nearby receptors.

- As a result, cells that are farther from the source are expected to be exposed to a lower dose of the ligand. This kind of gradient-dependent ligand is heavily used throughout development, as different doses of the signal can be detected, and responded to, differently.
- Compared to endocrine and neuronal signaling, described above, the affinity for the ligand is likely to be more moderate. However, it will also be somewhat more variable than in other forms of signaling. Dose-dependent responses often use receptors with different affinities for the ligand so that they cannot be activated at concentrations that are lower than their sensitivity. Again, this is very common during embryonic development across kingdoms.
- Examples of paracrine signaling include the synaptic signaling we discussed earlier, where the neurotransmitter is released into the synapse and received by the receptors on the other cell in the synapse. Also, many of the growth hormones used in development will have dose-dependent responses, thus allowing multiple responses from the same growth hormone.

Autocrine signaling is when the signal is both released and received by the same cell. It is considered a type of paracrine signaling, since the signal must diffuse to the receptor through the extracellular environment.

- This form of signaling is often used by the immune system in order to help it ramp up the immune response when activated. It is also something that happens in cells that have become cancerous; it helps the cancerous cells break free of the normal regulatory controls so that they can grow and divide without restriction.
- Sometimes a signal is released by a cell and is received by both itself *and* neighboring cells. Like much in biology, and in signaling, these categories are not mutually exclusive, and a single ligand can be used in a number of ways to illicit cellular responses.

Finally, **juxtacrine** signaling isn't very common overall compared to the other kinds of signaling. It's also known as *contact-dependent signaling*, which gives you an idea of what is involved in this form of signaling. In this case, the cell that receives the signal comes into direct contact with the one that is sending out the signal. Most commonly, this would mean direct cell-to-cell contact, but it may also be an interaction between a cell-surface receptor and a glycoprotein of the extracellular matrix.

- Just because this form of signaling is not common, we don't want you to get the impression that it's not important when used. Some of the most important survival signals are juxtacrine. Integrins are receptors that bind directly to the extracellular matrix and help the cell know that it's not (yet) time to undergo programmed cell death (also known as **apoptosis**). Complete loss of integrin binding, without undergoing programmed cell death, is also considered a hallmark of cancer and the start of the process known as *metastasis*, which is when cancerous cells migrate and form new tumors in different parts of the body.
- Another example of cell-to-extracellular matrix signaling is when leucocytes use this form of binding to "roll" through blood vessels (Video 07-02) prior to leaving the bloodstream at the

site of an infection.

The most famous example of cell-to-cell juxtacrine signaling is the *Notch-Delta* signaling cascade, which is essential to embryonic development in animals. The receptor (Notch) is a plasma membrane protein on one cell, and the ligand (Delta) is on the other. When they bind, the cytosolic side of the Notch receptor is cut off and becomes a transcription factor that then enters the nucleus. As a result, the transcription factor activates gene expression, and the "fate" of this cell is permanently and irreversibly changed. Notch-Delta signaling is extremely important in development, when cells are gaining their identities.

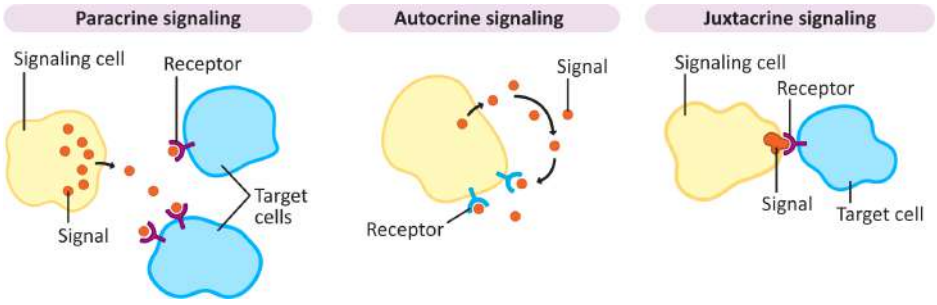


Figure 07-02: Short/medium-range signaling—paracrine, autocrine, and juxtacrine signaling. Paracrine signals are released into the extracellular matrix and allowed to diffuse to neighboring cells. Autocrine signaling happens when a cell releases a ligand that is received by its own receptor. Juxtacrine signaling requires physical contact between cells, as the ligand is attached to the signaling cell and not secreted as in other examples. Sometimes this is called contact-dependent signaling. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=156#oembed-2

Video 07-02: Immune cells use juxtacrine signaling to find sites of injury. White blood cells travel through the blood vessels by a mechanism known as "rolling" and then slip between the cells of the blood vessel wall at the point where they are needed to fight infection. This video describes the cellular mechanism for rolling and escaping from the blood vessel.

The Stages of Cell Signaling

At the most basic level, signaling happens in the following stages:

1. *A signal is sent.* There are a number of ways this could happen. A cell could release a molecular cue, or the environmental conditions themselves could provide the molecule that is being

detected. We'll do a more in-depth exploration of signals and ligands shortly.

2. *This signal is received by the cell.* This will require a receptor to recognize the signal and respond to it. Most often the receptor is at the cell surface, since that's where signals would arrive first, but it doesn't have to be. Nitric oxide (NO) is an example of a signaling molecule that can diffuse across cell membranes, so its receptor is in the interior of the cell.
3. *The signal is "interpreted" by the cell that receives it.* This could involve a number of things, like splitting the signal so that multiple responses are possible; it also likely involves transferring the signal across the plasma membrane by activating specific internal responses. This interpretation step is called *transduction*.
4. Finally, *the cell responds to the signal.* A cell could have multiple responses to a signal. Some responses are fast, while others are slow. If the signal is used to start a process, like cell division, apoptosis, or changes in cell identity, then there could be multiple changes in the cell as it prepares for the new behavior.

We'll look at each of these parts in turn as we continue to explore this topic.

1. Signaling Molecules Are Varied

Signaling events start with a signal that originates from the cell exterior. These signals could be something directly from the environment or something that originated from another cell (which may be part of the same multicellular organism or not). There are very few rules for what constitutes a signaling molecule other than the fact that there is a receptor capable of receiving the signal and responding to it. Some examples of signals include the following:

- **Odorants***:* Each of the chemicals you taste and smell has a receptor that can bind to it and sends a slightly different signal to your brain so that you can tell the difference between them.
    - Animals that are considered to have a better sense of smell will have many (thousands!) more receptors available than those that are not considered to have a good sense of smell. The increased number and variation in their receptors mean that their sense of smell is more nuanced comparatively. Among the vertebrates, mammals and fish are thought to have relatively good senses of smell, whereas birds are thought to have much worse senses of smell compared to other vertebrates.
    - In the mammals, the grizzly bear is thought to have a truly spectacular sense of smell—possibly one of the best in the world. It is said to be 7 times stronger than that of a basset hound (the superstar sniffer of the dog world). In turn, basset hounds are thought to have a sense of smell that might be as much as 100 million times more sensitive than our human noses. Interestingly, but maybe unsurprisingly, whales have a terrible sense of smell compared to their fellow mammals. However, they have an excellent sense of taste.
- *Biological macromolecules:* There are examples of signaling molecules from all four of the biomolecules (lipid, protein, carbohydrate, nucleic acid). Insulin, which helps you regulate blood sugar, is a short peptide. Glucose itself acts as a signal for the release of insulin.
- **Neurotransmitters***:* They're usually small chemicals that are synthesized in the cell for

eventual release. We've seen examples of these already (acetylcholine, dopamine, epinephrine, serotonin, etc.).

- *Photons:* There are many receptors that are able to respond to light. The photoreceptors in our eyes are a great example, but there are further examples across the various kingdoms of life. Photons are often used to help organisms regulate circadian rhythms, among other things, so that they know when it is time to sleep and when to be awake. Plants use light sensing extensively to decide when and how to grow and when to flower to maximize the change that their pollinators will be available.
- *Mechanical force and other nonchemical signals:* Your skin can respond to a variety of nontraditional signals, like physical pressure and temperature. Your ears respond to sound waves. Plants can also sense a number of the physical properties of their environment, such as temperature and gravity. They can even respond to injury (usually as a result of herbivory) and manage the wound while also ramping up their chemical defenses.

Most of the signals we're going to explore in this chapter will be of the type that are released by one cell and then travel to the receptor in one of the ways described earlier. As such, they will most likely be small chemicals (like neurotransmitters) or biomolecules (like short peptides or carbohydrates).

2. Receptors Are Used to Receive Extracellular Signals

A cell's first point of contact with the outside world is when one of its receptors binds to a ligand and a signaling pathway is activated. As mentioned above, since the ligand is coming from the exterior of the cell, it is quite common for the receptor to be on the cell surface, with the ligand-binding domain facing outward. However, if the ligand is able to diffuse across membranes (*Aside: What chemical properties should the ligand possess to make this possible?*), then the receptor could be somewhere inside the cell. We won't be discussing too many examples of internal receptors, but the following are a few examples of signals that can cross the plasma membrane:

- **Nitric oxide (NO)** is detected by the nervous system and results in a relaxation of smooth muscle cells. NO is a very small, nonpolar molecule, so it slips easily through the plasma membrane without help. The receptors for NO are inside the cell.
- Estrogen, testosterone, and other sterol-based hormones are often received inside the cell by a family of receptors known as the **nuclear receptors**. Nuclear receptors are receptors that either live full time in the nucleus or move from the cytosol to the nucleus upon activation. They bind directly to the DNA and act as a transcription regulator when in their active state.
- Auxin is a key plant hormone that appears to be involved in…well…everything! It has a strong influence on growth and development. Since plants cannot get up and move from where they were planted, they respond to their environment, in part, by growing toward or away from a specific stimulus. Auxin helps with this kind of differential growth. It is also a small, uncharged molecule that is able to diffuse right into the nucleus. The auxin receptor is also a transcription factor, even though it's not genetically related to the nuclear receptors described above.
- As you know, light is also a signal that is received and processed by many organisms and can penetrate the cell without help. While the light receptors in the rod and cone cells of the

human eye are officially on the cell membrane, they are within a specialized structure where the plasma membrane has folded in on itself many times, so they're no longer on the surface per se. In plants there are light receptors that are found it the cytosol, in addition to the photosynthetic apparatus, which is inside the chloroplast (as you know from Chapter 5).

A common theme for these internal receptors is that the activated pathway is often very short. In the case of auxin and the nuclear receptors, it is the receptor itself that enters the nucleus and causes the change in gene expression. From the perspective of exploring the principles of signaling, that doesn't give us very much to work with. So in this chapter, we mostly focus on plasma membrane–bound receptors, as those tend to activate longer, more complex pathways; thus, they require more explanation to understand. A ligand that is received by a receptor on the plasma membrane will usually not enter the cell. Instead, the signal is **transduced** across the plasma membrane and then activates a series of events inside the cell that allow the cell to respond to the signal. All of these events together are known as the **signaling cascade**, and understanding how these cascades fit together is the purpose of the rest of this chapter.

Signaling cascades are often categorized by the type of receptor used to initiate them. These fall into several categories based on their structural features. Later in this chapter, we will explore some of the most common types of receptors and the responses that they activate.

3. Once the Signal Is Received, It Must Be "Interpreted" by the Cell

The binding of the ligand to the extracellular portion of the receptor causes a conformation change in the receptor that carries through the membrane and into part of the receptor that is inside the cell. This is how the signal "passes through" the membrane, and it's the first step in a process called **signal transduction**. Transduction is when a signal is converted from one form to another. In this case, the extracellular signal molecule binds and initiates the conformation change in the receptor, which activates it and allows it to generate a response inside the cell.

Once the signal is perceived inside the cell, there are several things that could happen (Figure 07-03). However, not all possibilities will occur in every cell. Different cells express different genes, which will code for different signaling cascade components. Thus, a single ligand can have different effects in different cell types. This is helpful in multicellular organisms, as it means that an organismal-level response can be coordinated using a single signal molecule. An excellent example would be our response to adrenaline. When our body releases adrenaline, there are a wide variety of physiological effects (e.g., increased heart rate and breathing, perspiration) that are all designed to prepare us for the "fight or flight" response to danger. Each of these physiological responses begins when a receptor on different cell types binds to adrenaline, and a signaling cascade is triggered.

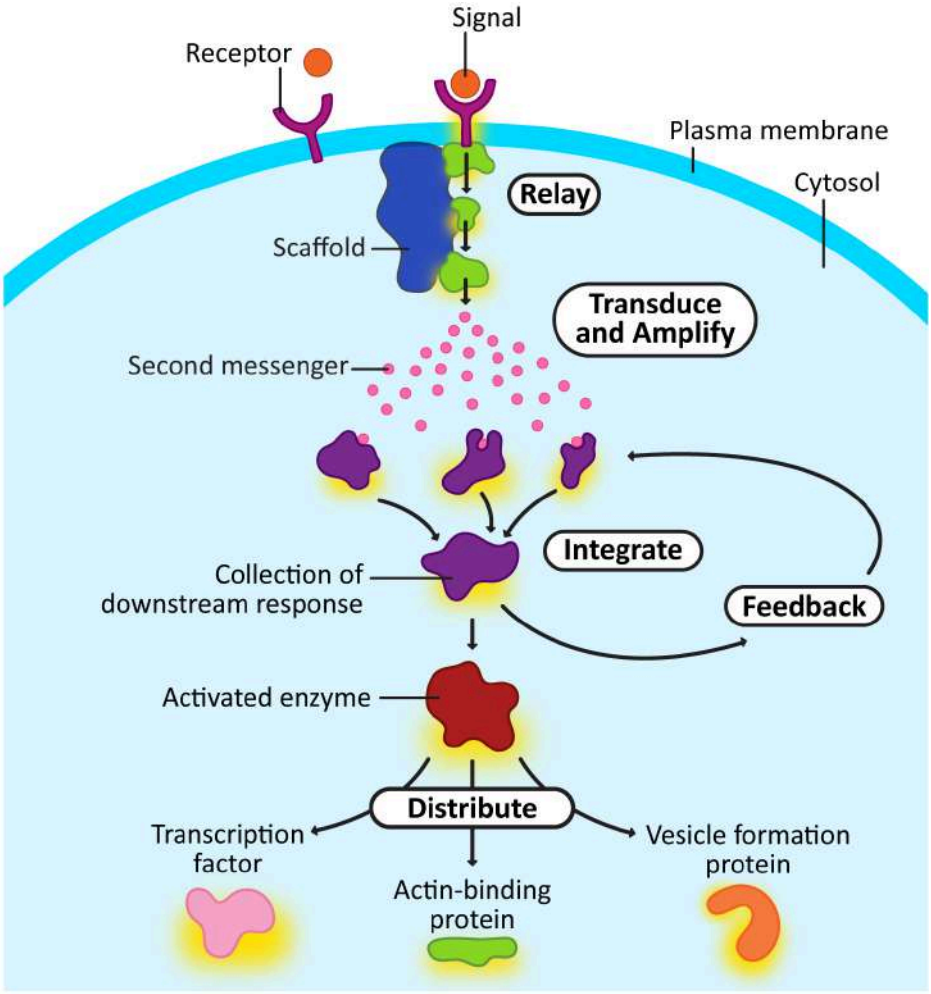


Figure 07-03: A generic signaling cascade showing some possible intracellular steps and responses to an extracellular signal. This is not based on a specific "real" cascade. It contains elements that may or may not be found together, depending on the pathway. It is designed to illustrate all of the possible options in one figure rather than be representative of any single pathway. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Here's a quick explanation of each of the intracellular signaling steps from Figure 07-03. Later, we'll look at some of the more famous examples of proteins that perform these roles. Remember that not all of these are present in every signaling cascade, and sometimes an individual element could be performing more than one function.

- *Relay:* Relay is an event where the signal is passed from one intracellular signaling molecule to the next. This is not dissimilar from passing a baton in a relay race. In Figure 07-03, this is shown right after the receptor, but in reality, it can happen anywhere in the signaling cascade. Often, the "signal" that is being passed along is the addition or removal of a phosphate group on a target protein (known as **phosphorylation** or **dephosphorylation**, respectively). You'll also notice a larger protein that connects three intracellular relay proteins together. This protein is called a **scaffold protein**. They often aid in relay signaling events. Their job is to hold several proteins together so that activation can be more efficient.
- *Transduce and amplify:* Often, a signal needs to be transformed so that it can be used in other ways. Phosphorylation is a common modification in signaling cascades but only works on

other proteins. Small "messenger" molecules (often referred to as **second messengers**) are used to amplify the signal and quickly spread it throughout the cell. These molecules are either produced by proteins that have been activated as part of the response or released into the cytosol from another compartment (like the ER). Some of the most famous second messengers include the following:

    - *Calcium* ions flood the cytosol after they have been released from the ER.
    - *Cyclic AMP* (cAMP) is produced quickly in the cytosol from adenosine monophosphate (AMP).
    - PIP_2, a specific phospholipid molecule, can be split into a membrane-bound second messenger called *DAG* (diacylglycerol) and a soluble second messenger called IP_3 (inositol triphosphate).

- *Integrate:* In this process, multiple receptors feed into a single "downstream" response. In some cases, both signals will be required for the response to occur, while in other cases, either signal can cause the response in question, regardless of whether the other is present.
- *Feedback:* As the signal moves through the cascade of responses, later events can influence earlier components of the pathway. We call this a **feedback loop**. Feedback loops can be either *positive* (i.e., later events help further activate earlier steps) or *negative* (i.e., later events stop earlier events from continuing).
- *Distribute:* At some point, it's probable that an enzyme will be activated whose function is to activate many different proteins. The proteins that are being acted upon will have a wide variety of functions (e.g., transcription factors, actin- or microtubule-binding proteins, vesicle-formation machinery). This is how the signal is distributed so that cellular function changes.

4. Responses to the Signal Can Also Be "Fast" or "Slow" and Are Usually Both

The cellular responses at the end of the signaling cascade usually fall into two categories: fast responses and slow responses. The difference between these is very simple. *Cellular responses that require protein synthesis will always be slower than responses that rely on preexisting proteins, which can be simply altered.*

Slow Responses Come from Changes in Gene Expression

Often, a signaling event will result in an overall change in cellular identity. Changing the identity of a cell will require a number of changes in gene expression: some proteins will need to be synthesized, while others will likely need to be removed from the cell. For example, most of the signaling events involved in development will cause changes in gene expression, as cell differentiation is a key step in embryonic development. As a secondary example, the immune system also uses alterations in gene expression so that antibodies can be produced in response to a pathogen or vaccine. In order to change gene expression, signaling cascades cause the activation or deactivation of transcription factors or proteins that affect chromatin structure (i.e., histone or chromatin-modifying enzymes). As we discussed in Chapter 3, proteins that influence histones and/or chromatin will impact the accessibility of the DNA, thus impacting the expression of genes in that area. The end result will be changes in the protein composition in the affected cell.

Alternatively, extracellular signals can stimulate signaling cascades that result in shorter-term changes in cellular function by altering preexisting proteins. A very good example of this, which we will explore in detail in Chapter 8, is mitosis. When mitosis is activated, the cell is vulnerable, so this process must be completed quickly. Most of the proteins involved in rearranging the cytosol for mitosis already exist before mitosis starts. They are present in the cytosol, waiting for the signal to initiate cell division. Once the signal is perceived, these preexisting proteins are modified as a result of the signaling cascade. This causes a rapid shift in cellular physiology, and the cells progress through mitosis as efficiently as possible.

Deactivation of a Signaling Cascade

So far in this chapter, we have discussed how signaling cascades are activated. In some cases, this is the end of the story, as the signal never needs to be turned off. An example of this would be the Notch-Delta signaling pathway, which changes the cell permanently. It doesn't really need to be "turned off" in the traditional sense. Another great example would be apoptosis. The signaling cascades of apoptosis lead directly to the death of the cell, so once apoptosis is activated, there is no cellular mechanism to deactivate it.

On the other hand, a great many signaling events need to stop eventually so that the cell can go back to "normal." Mitosis is a good example, as the signaling that started mitosis must be stopped in order for mitosis to end. Another good example is sensory perception. Every single one of our senses is controlled through signaling. If we can't stop receiving and responding to the signal, then we can't go back to normal so that the signal can be received again. Imagine if our pain receptors never turned off once they were turned on!

Figure 07-04 illustrates the most common ways that signaling cascades are stopped. Quite often, cells deactivate signals by inhibiting, or removing, the receptor. In some cases, it is an internal component that is deactivated.

Some of these changes are permanent. An example of permanent deactivation occurs in **receptor downregulation**, where the receptor is brought into the cell via endocytosis and then degraded in the lysosome. However, many other deactivation steps are reversible. **Receptor sequestration** pulls the receptor off the cell surface so that it can't bind to a signal for a period of time but the receptor isn't destroyed. In both **receptor inactivation** and **signaling protein inactivation**, one of the proteins that forms part of the signaling cascade is inactivated in some reversible way (most likely by phosphorylation and/or binding to an inhibitory protein).

Deactivation of signaling cascades is usually controlled via negative feedback loops. This means that the activation of the signaling cascade produces products that help shut down the signaling cascade as well.

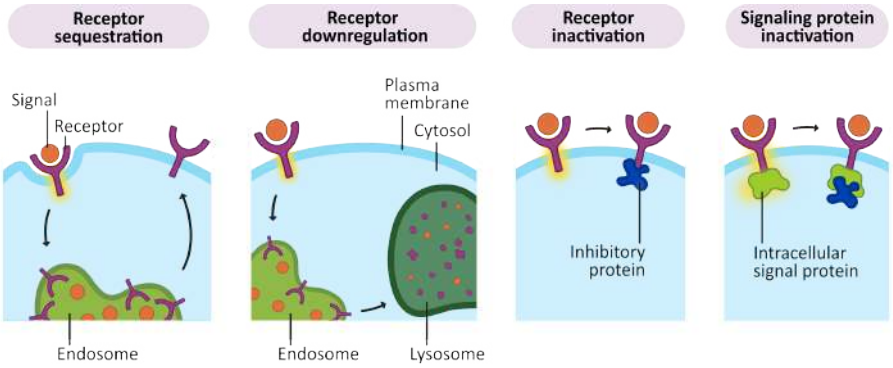


Figure 07-04: Common mechanisms to deactivate signaling cascades. In receptor sequestration, the receptor is brought from the membrane into an internal compartment (usually an endosome) where it can't interact with the signal. Downregulation of a receptor means that it ends up in the lysosome and is thus destroyed. Alternatively, the receptor can be inactivated, usually through the binding of an inhibitory protein, so that the signal can no longer be relayed. Finally, it may not be the receptor itself that is inhibited but a downstream protein, which has more or less the same effect of stopping the signaling cascade. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Studying Cells: Reading and Interpreting Diagrams of Signaling Cascades

Before we move on to the next topic, we should take a moment to discuss the way that scientists write out signaling cascades and how to interpret them.

As you likely have seen in the figures here or elsewhere in your course notes or the internet, signaling cascades are often written out with the proteins lined up and connecting symbols between them. Two important symbols are used in between the proteins to identify how they impact each other:

- a regular arrow (, which is used when protein 1 *activates* protein 2, and
- a blunt arrow (, which is used when protein 1 *inhibits* the function of protein 2.

We often discuss the steps of a signaling cascade as being **upstream** or **downstream** of each other. So for example, the signal binding to the receptor is *upstream* of the cellular response. Conversely, the response will be *downstream* of ligand binding. These terms and symbols are important for interpreting the diagrams used to represent signaling pathways in textbooks and papers. Video 07-03 provides further explanation.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=156#oembed-3

Video 07-03: Example of a signal transduction pathway, with an explanation of how to read the different symbols that are used to

represent activation and inhibition as well as some ways mutations can impact them.

TOPIC 7.2: EXAMPLES OF COMMON SIGNALING MECHANISMS

Learning Goals

- Compare and contrast ion-gated channels, G-protein-coupled receptors (GPCRs), and enzyme-coupled receptors, including their mechanism of activation, response, and deactivation.
- Name and describe the possible downstream pathways used after GPCR and enzyme-couple receptor activation. Identify which ones are more commonly used with each type of receptor.
- Explain how researchers can experimentally investigate the role of posttranslational modifications in the proper functioning of a cell signaling cascade.

All right! Now that we have a better understanding of some of the general principles of signaling, the next step is to look at examples of the proteins that do this work and also how they do it. As you may have already noticed, signaling can be quite confusing. There are a lot of components that interact, so it can be difficult to keep it all straight. Our suggestion is to take it slow and to stop and think about each step and how it might influence the others. *Memorizing will not be very helpful here*, as there are too many things to keep track of, so it's better to take a step back and try to understand what's going on. We're not going to dive too deep, since this is still only a brief introduction to the beautiful complexity of signaling.

We'll start by discussing the most common cell-surface receptors (Figure 07-05 and Video 07-04). Along with each receptor type, we'll also examine some of the more common "downstream" responses to each receptor type. Remember that there is quite a lot of mixing and matching when it comes to the proteins involved in signaling cascades.

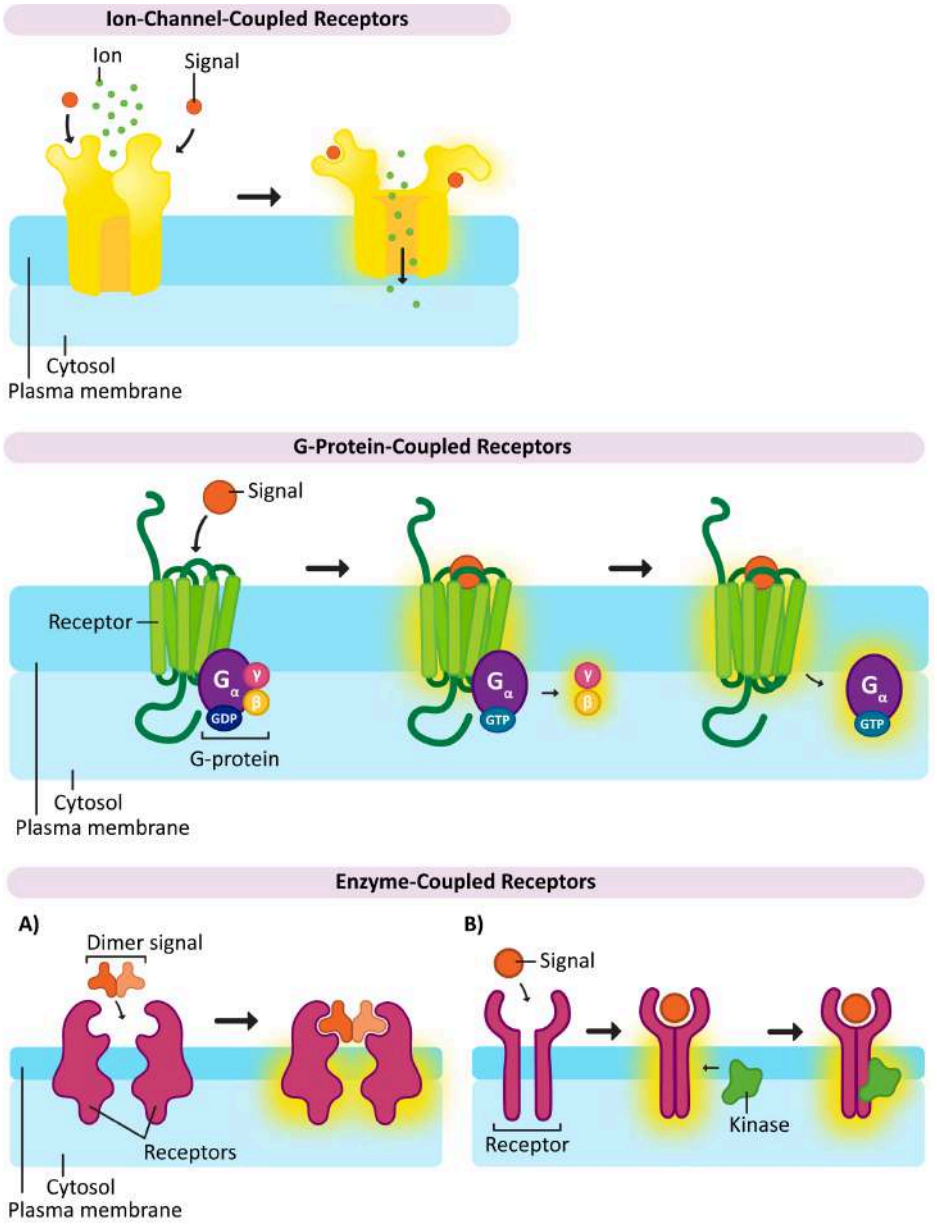


Figure 07-05: The three most common types of cellular receptors. When bound to their signal, ion-coupled receptors (A) open to allow an influx of ions into the cell. G-protein-coupled receptors (B) activate G-proteins when activated by their ligand. Finally, enzyme-coupled receptors (C) activate downstream proteins enzymatically once activated by binding to a signal. Most often, this takes the form of phosphorylation. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=156#oembed-4

Video 07-04: Overview of the major types of receptors.

Ion-Channel-Coupled Receptors

As the name says, **ion-channel-coupled receptors** are receptors that, when activated, open (or close) an ion channel (Figure 07-06). This is very common in the nervous system, where action potentials are produced through ion channels and perpetuated in part by the release of neurotransmitters. In the case of the nervous system, the purpose of these ion-channel-coupled receptors is to convert a chemical signal (like a ligand) into an electrical one. Unfortunately, this aspect of neuronal function is well beyond the scope of this textbook. As a result, this function of ion-channel-coupled receptors is not going to be discussed much. Another name for an ion-channel-coupled receptor is a *ligand-gated ion channel*.

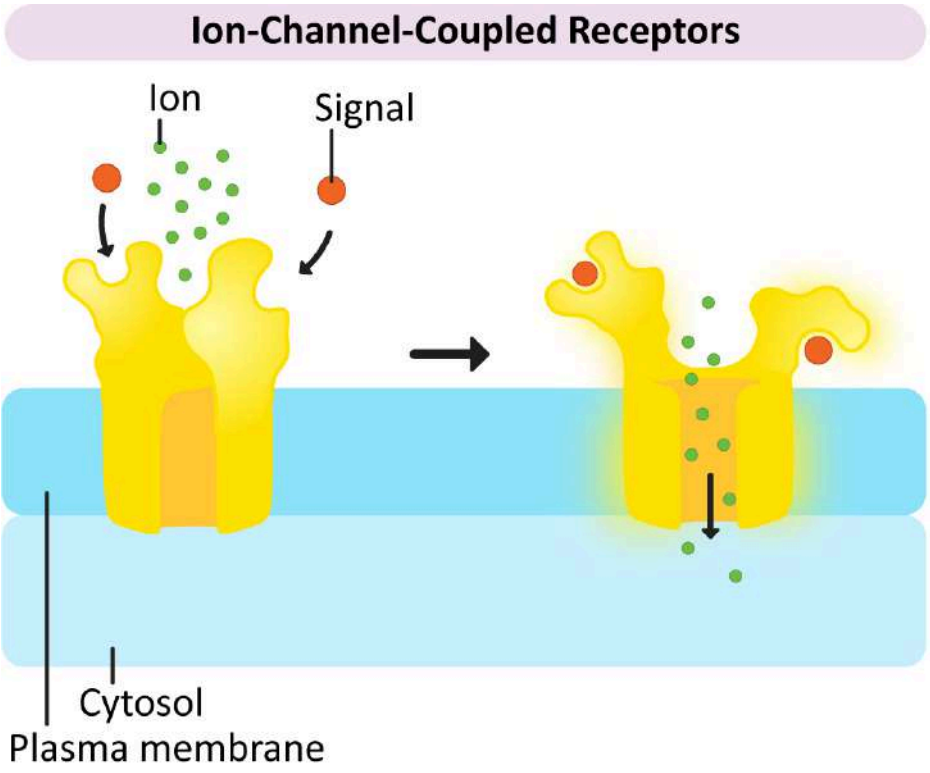


Figure 07-06: Ion-channel-coupled receptors bind to their ligand. This induces a confirmational change causing the opening of the channel to allow an influx of ions across the membrane. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

There is one important example of a ligand-gated ion channel that we will be discussing a bit later, and that's the one used to release calcium into the cytosol from the ER. Remember that calcium is an important **second messenger** in cell signaling and is often used to amplify and spread the response throughout the cell. The ER-localized calcium channels that are used in this process open to a specific signal as well. In this case, another second messenger is used, **IP_3**, which is released when the phospholipid, **PIP_2**, is cleaved as part of a signaling response. We will see how this works a bit later in this topic.

G-Protein-Coupled Receptors (GPCRs)

The next group of receptors we'll look at is **G-protein-coupled receptors**, or GPCRs for short. This

type of receptor is by far the most common in animals. They are heavily involved in animal sensory perception, especially our sense of smell. Humans have over 700 GPCRs in total, most of which are dedicated to smell. On the other hand, a mouse, which has a much sharper sense of smell than we do, has over 1,000 different GPCRs involved in its sense of smell alone!! Even better sniffers, like the basset hound and grizzly bear we talked about earlier, are thought to have an even higher number of olfactory GPCRs.

GPCRs are a very large family of proteins that includes many subfamilies with specific functional differences. We won't get into all of that detail here (it fits better with an upper-level signaling textbook), but we can identify some important characteristics of GPCRs that are common across the family. For example, the GPCR itself is almost always a single polypeptide chain with seven transmembrane domains that cross the membrane as alpha helices. This is considered to be a characteristic structural feature of these receptors. GPCRs have a large pocket for ligand binding, which allows for a lot of variation. Thus, the ligands that can bind to GPCRs are also quite variable. This is why they are such a good candidate to manage our sense of smell—they most commonly bind to small organic molecules rather than peptide hormones. Some of the light receptors on our retinas are GPCRs as well.

Once again, the name of this type of receptor tells you something about it. The receptor itself works in conjunction with another set of proteins called **G-proteins**. The "G" in G-protein stands for GTP, as GTP is required to activate the G-proteins and allows a response to be generated. G-proteins also come in a variety of flavors to match up with the GPCRs that they work with.

Despite variation in the G-protein family, they also have many similarities:

- Each G-protein complex contains three subunits: the alpha (α), beta (β), and gamma (γ) subunits.
- The three subunits can all bind together in a trimeric complex, or G-alpha can separate from the other two. G-beta and G-gamma almost always stay together as a unit.
- The alpha-subunit is able to bind to GDP or GTP. Both the alpha- and the gamma-subunits have covalently linked lipid tails, which allow them to remain associated with the membrane.
- The beta unit does not have a lipid tail, but since it remains attached to the gamma-subunit, it also remains attached to the cell membrane.

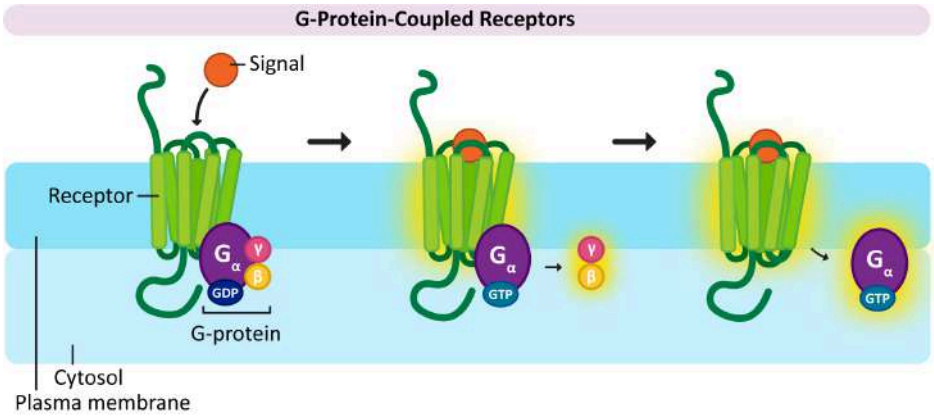


Figure 07-07: The activation of G-proteins. When a signal molecule binds, this activates the G-protein-coupled receptor. The activated receptor swaps out GDP for GTP on Gα. When bound to GTP, the Gα separates from Gβγ. These two molecules stimulate further downstream signaling. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

In a nutshell, here's how GPCR activation works (Figure 07-07):

1. When the receptor binds to its ligand, it undergoes a conformation change and binds to the α/β/γ complex. (At this point, the three subunits are bound together as a single complex attached to the GPCR on the cytosolic face of the plasma membrane.) This causes the GDP to be released from the α-subunit and get replaced by GTP.
2. The binding of the GTP molecule causes a conformation change in the G-protein complex, and the β/γ subunits are released from Gα.
3. The separated G-protein subunits (Gα and Gβγ) are now considered active and activate their appropriate downstream targets. So by this mechanism, we effectively split the signal as it is being relayed so that we can create more change in the cell.
   It's worthwhile to note that as the G-proteins are activated and move away from the receptor, the lipid tails keep them associated with the interior surface of the plasma membrane. As we learned back in Chapter 2, membranes are fluid, so the G-proteins will be able to diffuse laterally through the membrane until they come into contact with proteins that they can bind to and activate.
4. Eventually the G-proteins will deactivate themselves. The Gα subunit will slowly hydrolyze the GTP so that it is converted back into GDP (and a phosphate molecule, which is released). This inactivates the Gα subunit, which causes all three G-proteins to return to their inactive state.
   Because the activation of the G-proteins is dependent on the presence of GTP in the Gα subunit, this provides a built-in system for shutting down the response. The GTP is short lived, so the response lasts only as long as the GTP does. However, there may be many cycles of activation and deactivation of G-proteins during a single response.

Common Downstream Effects of GPCRs—Second Messengers: Cyclic AMP (cAMP)

Now that we've discussed how the GPCR is activated and deactivated, we turn our discussion to an important downstream signaling molecule commonly used in GPCR signaling cascades: the

production/release of **second messenger** molecules. We'll specifically discuss the two most common second messengers that are activated by G-proteins: **cyclic AMP** and **IP_3/DAG**.

Cyclic AMP, or cAMP, as it is often abbreviated, is a modified version of adenosine monophosphate (AMP). cAMP is made from ATP, a molecule you should know well that is extremely abundant in the cell. Figure 07-08 shows the structural changes required to interconvert ATP, cAMP, and AMP.

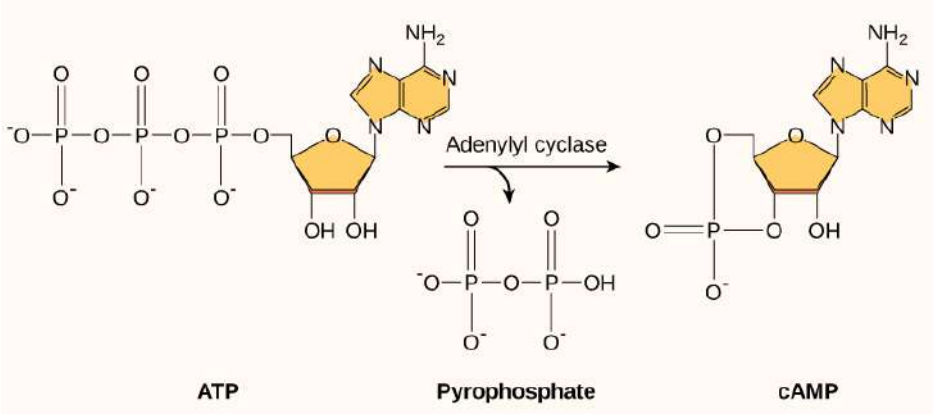


Figure 07-08: The formation of cyclic AMP from ATP. Image from OpenStax Biology "Chapter 9—Propagation of a signal—Figure 9.12" is shared under a CC BY 4.0 license.

Cyclic AMP is produced when a plasma membrane–bound enzyme, **adenylyl cyclase**, is activated by the Gα subunit of the activated G-protein (Figure 07-09). Cyclic AMP is then capable of moving throughout the cell and interacting with any enzymes that have binding sites for it. cAMP is a short-lived molecule—its half-life in the cell is less than 1.5 minutes. An enzyme known as **phosphodiesterase** breaks it down so that the cellular response is kept short. Figure 07-09 and Video 07-05 (below) review the activation of GPCRs and discuss the production of cAMP.

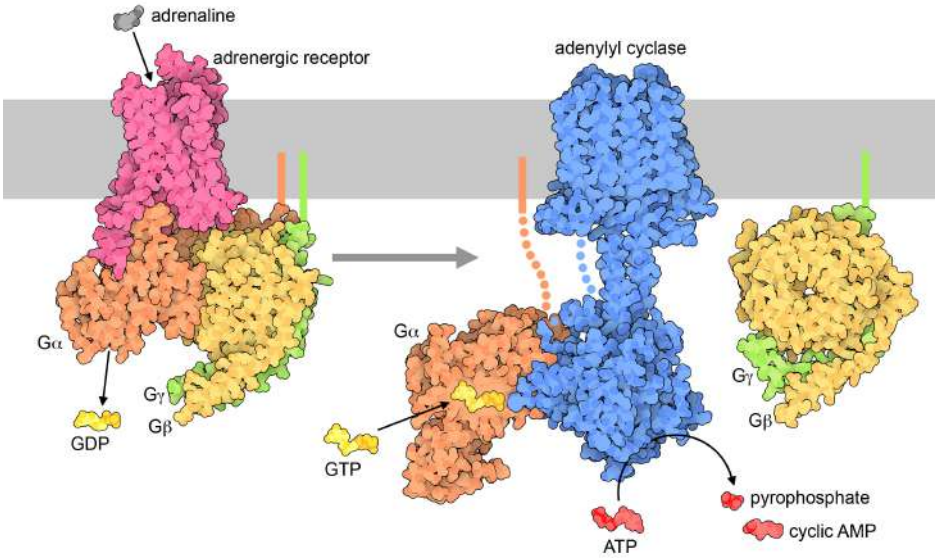


Figure 07-09: Activation of adenylyl cyclase by G-proteins. Note that the lipid tails of the G-proteins are indicated by the colored lines that enter the gray area that depicts the membrane. Image by Goodsell, D. S. Molecule of the month: G-Proteins is shared under a CC BY 4.0 license. http://doi.org/10.2210/rcsb_pdb/mom_2004_10.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=156#oembed-5

Video 07-05: Walk-through explaining the activation of GPCRs and the production of cAMP.

One of the most important enzymes activated by cAMP is a cytosolic protein called **protein kinase A (PKA)**. When activated, PKA can have a number of different effects. Figure 07-10 shows two examples, one of which is a "fast" response, illustrated by the breakdown of glycogen in the liver. The other is a "slow" response—PKA can also enter the nucleus and phosphorylate transcription factors, which will activate these transcription factors, resulting in gene expression. Remember that different cells/tissues will have different proteins present in their cytosol that can be phosphorylated by PKA, which, in turn, will result in different effects.

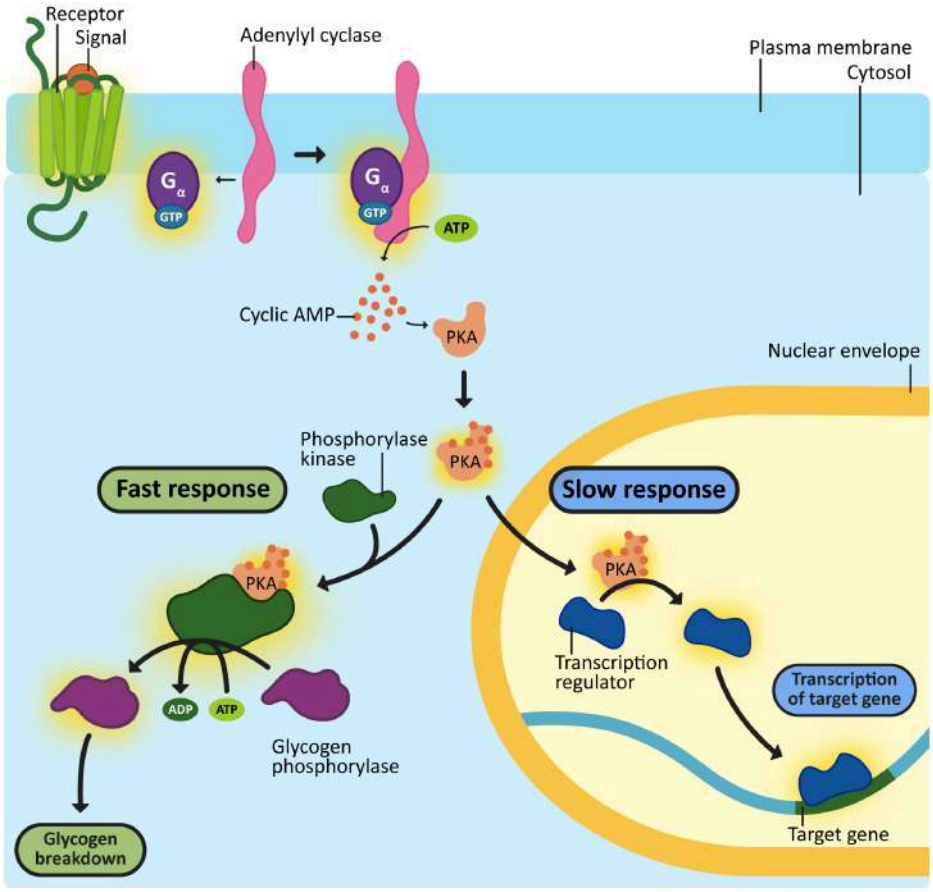


Figure 07-10: Downstream effects of PKA activation by cAMP. Activated GPCR activates Gα by swapping out GDP for GTP. This, in turn, activates adenylyl cyclase to produce cyclic AMP (cAMP) to activate protein kinase A (PKA). PKA, depending on the cell type, can elicit different responses. In the fast response, PKA activates proteins that alter other proteins. In the above example, glycogen phosphorylase is activated in response to PKA and results in the breakdown of glycogen (a glucose storage molecule made by animals). In the slow response, PKA activates a transcription factor that alters the transcription of a target gene. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Common Downstream Effects of GPCRs—Second Messengers: DAG, IP_3, and Calcium

DAG and **IP_3** are another set of common second messengers. They are always produced together, as they are two parts of the same initial phospholipid (called PIP_2, which is a generic term for a *phosphatidyl inositol phosphate with two phosphates added to the inositol sugar*). PIP_2 gets split in two by an enzyme known as **phospholipase C (PLC)** Figure 07-11) to produce the two second messengers. PLC is activated by a Gβγ protein complex. While the breadth of lipid function in cells doesn't often get discussed at the introductory level, it is worth noting that there are a number of membrane lipids, especially phospholipids, that have a key role to play in cell signaling and membrane identity.

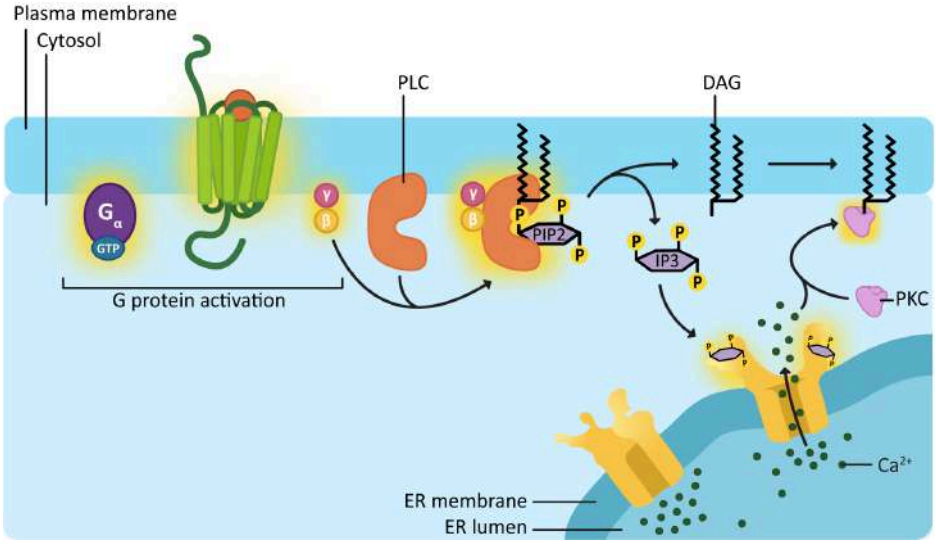


Figure 07-11: Production of second messengers IP_3 and DAG by GPCR activation resulting in the release of calcium ions from the ER lumen. In this image, the lipid tails can be observed as zigzags. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

As illustrated in Figure 07-11, phospholipase C clips the soluble "head" portion of the phospholipid from the glycerol backbone and lipid tails. The head portion is a specific type of six-carbon sugar, with three phosphate groups on it, called inositol 1,4,5-triphosphate, or IP_3 for short. The hydrophobic "tail" portion that is left in the membrane is now a molecule known as **diacylglycerol**, or DAG. Both of these molecules will be used to activate further components of the signaling cascade. DAG will diffuse laterally through the membrane and act as a docking site for a protein known as protein kinase C (PKC). IP_3 is now soluble, so it can diffuse through the cytosol and bind to and activate its downstream targets.

One important protein that IP_3 can bind to is a ligand-gated ion channel in the ER. This ligand-gated channel is responsible for releasing calcium ions, another important second messenger in the cell, into the cytosol. Much like cAMP, calcium can flood the cytosol very quickly, producing a rapid change in the membrane potential of the plasma membrane, among other things. Calcium waves that rapidly spread through the cell are a key event in many cellular processes, such as muscle contraction and oocyte fertilization.

The released calcium can also interact with many different targets in the cytosol. In Figure 07-11, we see that PKC requires calcium to bind in order for it to become active. So PKC activation requires *both* the docking site, produced by DAG, and the calcium released by IP_3 before it can be activated. Activated PKC will then be able to phosphorylate a number of specific targets to generate change in the cell.

Eventually, the calcium is pumped back into the ER using *different* ion channels than those that were opened during calcium release. The new channels are opened by the presence of calcium in the cytosol, which makes this an excellent example of a **negative feedback loop**, where the signaling event activates the process that will also deactivate the signal.

Incidentally, calcium is by far the most commonly used second messenger in plants. Also of interest is that GPCRs are not nearly as common in plants as they are in animals. The most commonly used receptors in plants are receptors coupled to **serine/threonine kinases**. Ser/Thr kinases are enzymes, so these are a type of enzyme-coupled receptor. This leads us very nicely to the last type of receptor we need to discuss.

Enzyme-Coupled Receptors

This final set of receptors that we need to discuss is much more structurally variable compared to GPCRs, but they all have one very important thing in common: when the ligand binds to the receptor, an enzyme known as a **kinase** is activated. Kinases add phosphate groups to other proteins. We've seen many examples of this kind of enzyme throughout this textbook (specific examples include PKA and PKC from the previous section). There are also enzymes that remove phosphate groups, which we call **phosphatases**.

Phosphate groups can only be added to proteins at specific sites—not all R groups on an amino acid are capable of having phosphate groups added. Kinases fall into two major categories based on the amino acid R groups they modify:

- *Tyrosine kinases:* These are the largest class of kinases that associate with receptors and thus are better studied overall within the context of signaling. In this case, phosphates are added to tyrosine residues on the target protein. Tyr kinases are often also receptors, or directly associated with receptors, so these are the specific type of kinase we discuss most often in this section.
- *Serine/threonine kinases:* Serine and threonine amino acids have very similar sidechains in that they are short hydrocarbon chains with a reactive hydroxyl group (-OH). Thus, the kinases that work on them are structurally similar as well. As such, we call this group serine/threonine (Ser/Thr) kinases. These kinases are more common in plants and in soluble kinase proteins not associated with receptors. PKA and PKC, from the previous section, are Ser/Thr kinases.

The enzyme-coupled receptors themselves will either

- have their own *intrinsic* phosphorylation capacity, meaning that they can act on themselves or their dimerized neighbor, which is activated upon ligand binding, or
- have no intrinsic enzymatic activity but will be closely associated with a kinase that is activated upon ligand binding.

We won't really differentiate between the types (integrated kinase or associated kinase) here, as there really isn't a whole lot of difference between them from a functional standpoint. Additionally, for the most part, the examples we'll see will be **receptor tyrosine kinases** (RTKs for short), where the kinase acts on tyrosine and is an intrinsic function of the activated receptor. Remember that there are also receptor-associated Ser/Thr kinases. Ligands for these kinds of receptors are often small peptides, such as growth hormones and insulin.

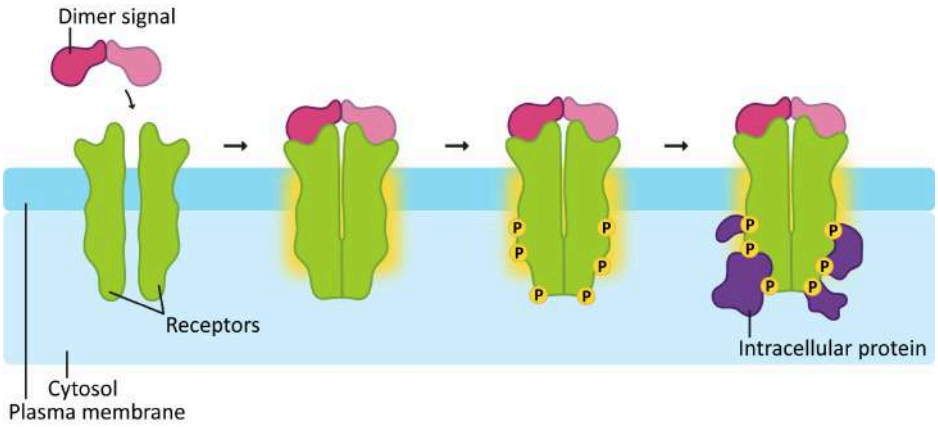


Figure 07-12: Activation of receptor tyrosine kinases. The ligand is a dimer (shown in pink) and binds to the receptor on the extracellular side, causing two receptors to dimerize, cross phosphorylate, and activate. Phosphorylation creates docking sites, allowing for recruitment of other enzymes and downstream signaling targets. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

A few important points about RTKs are highlighted by Figure 07-12:

- The first is that RTKs usually **dimerize** as they bind to the signal molecule. So the inactive receptor is solo, and when the ligand binds, the receptor forms a dimer with either a second version of itself (i.e., a homodimer) or a closely related, but not identical, version (i.e., a heterodimer).
- Once they dimerize and the ligand is bound, the enzymatic activity is turned on. As a general rule, each protein in the dimer will phosphorylate the other one.
- The added phosphate groups on the receptor most commonly form the basis for **docking sites** so that a bunch of other proteins can bind and be phosphorylated. These docked proteins will help relay the signal to the next step. At this point the receptor is acting as a **scaffold protein** in addition to its role in phosphorylating target proteins.

Much like GPCRs, there are both membrane-bound and soluble proteins that will be activated by enzyme-coupled receptors. Second messengers may or may not be involved with these receptors, though we are not going to explore any examples of second messenger activation by RTKs here (we already did that). We're going to look at two new examples of common downstream effects: Ras and mitogen-activated protein (MAP) kinase signaling as well as another example of lipids that have roles in signaling.

Downstream Effects of Enzyme-Coupled Receptors—Ras and MAP Kinase Cascades

Just like GPCRs, an activated receptor will need to relay the signal and amplify it somehow so that it can be spread around the cell and induce changes in the cell. We've already seen a couple of ways this can happen with GPCRs. In the case of RTKs, the first commonly activated molecule after receptor activation is a protein called **Ras**.

Ras is a **molecular switch** that is somewhat related to G-proteins. Molecular switches like these are very common in the cell. In fact, there is a very large family of these proteins (so big that we call it a *superfamily*) that is named after Ras. The members of the Ras superfamily of proteins are all small

monomers (i.e., they are **monomeric**) and are all GTPases. This means that they convert GTP to GDP. We've seen many examples of GTPases in other chapters of this textbook, including the following:

- Ran, which helps proteins to enter/exit the nucleus (Chapter 3)
- Rabs, which help with vesicle docking (Chapter 4)
- GTPases, which are part of the vesicle coat (such as Arf and Sar) and help the coat fall off after the vesicles form (Chapter 4)
- Tubulin, a very distant relative of Ras that works in a very similar way (Chapter 6)

The reason that we call these proteins molecular switches is that they have an "on" and "off" conformation, which is determined by the presence of GDP (off) or GTP (on). This is extremely similar to the G-proteins we saw earlier when discussing GPCRs. Like the G-proteins, Ras also has a small lipid tail covalently linked to it, which is exposed when the protein is activated. Thus, when the receptor activates Ras, it will act very similarly to a G-protein in that it will associate with the plasma membrane, diffuse laterally, and activate other proteins at the cell surface, as shown in Figure 07-13.

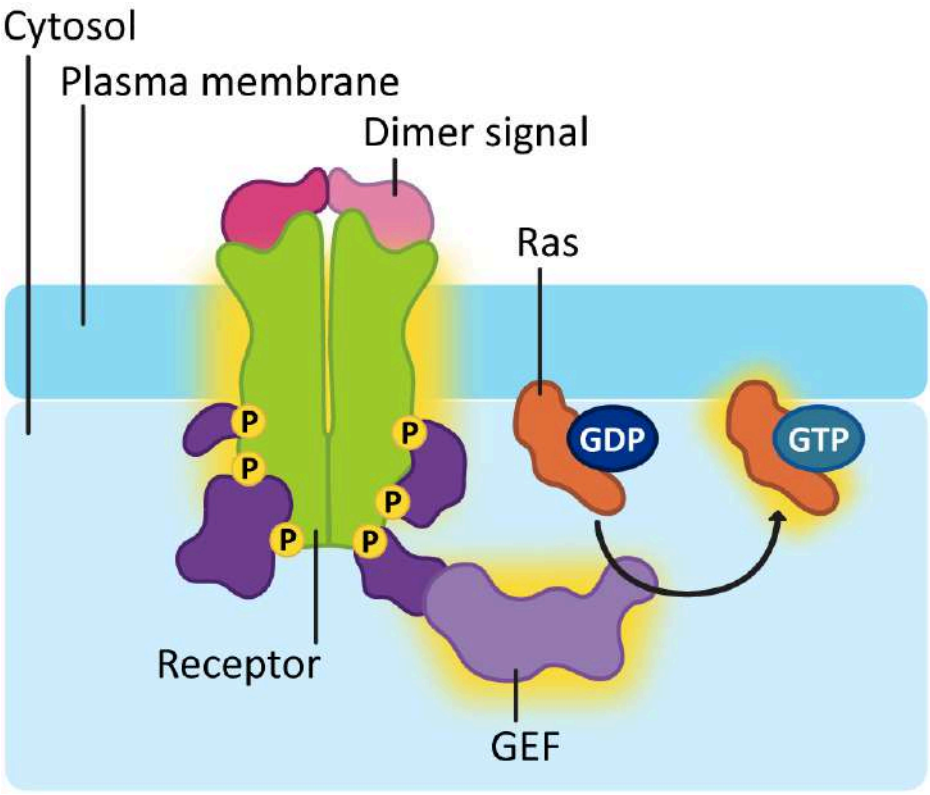


Figure 07-13: Activation of Ras by a receptor tyrosine kinase (RTK). Receptor activates a guanine-exchange factor (GEF), which swaps GDP for GTP on Ras thereby activating it. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

In order for Ras to be activated, a second enzyme known as a **guanine-exchange factor**, or GEF, is used to swap out the GDP for GTP. As you can see in Figure 07-13, the activation of the receptor allows the GEF to dock, which, in turn, makes sure it's in the right spot to activate Ras.

Ras activation is also short lived, like G-proteins. The GTP that is bound to the activated Ras will eventually be hydrolyzed with the help of another protein known as **GTPase-activating protein**, or

GAP. One of the more common downstream effects of Ras is that it activates a cascade of kinases that phosphorylate each other. You can think of it as very similar to a domino effect, where one domino knocks down the next, which knocks down the next, and so on. In this case, however, serine/threonine kinases are being activated in sequence, as each one phosphorylates the next. We call this a **MAP kinase cascade** (Figure 07-14).

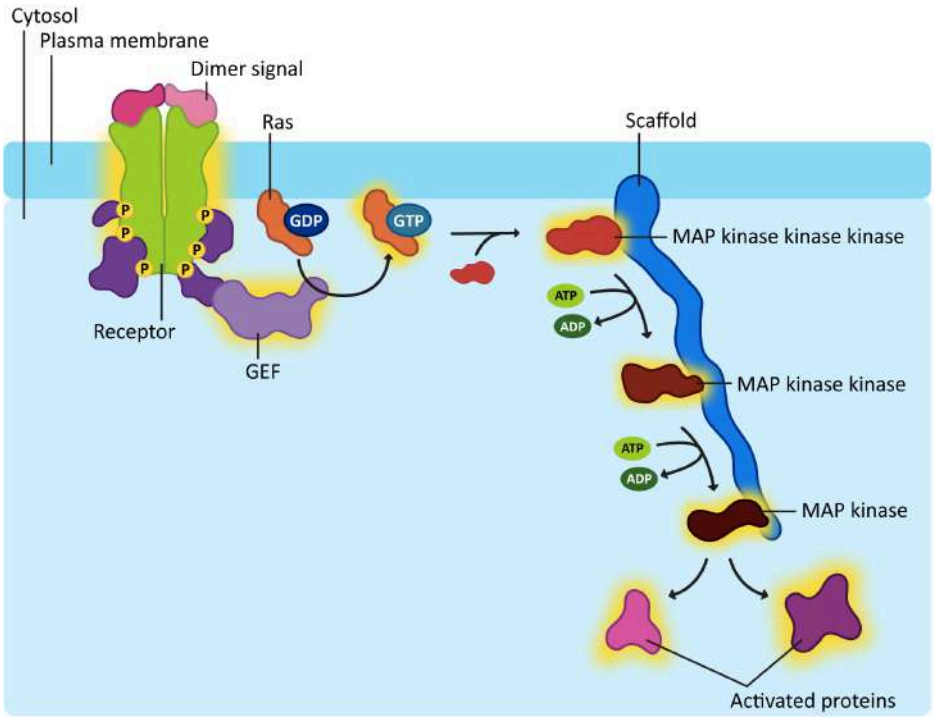


Figure 07-14: A generic MAP kinase cascade activated by Ras. MAP KKKs activate MAP KKs, which activate MAP kinases. The end MAP kinase is able to activate other downstream proteins through phosphorylation. Scaffold proteins bind to specific enzymes to allow specificity as well as temporal and spatial regulation to MAP kinase cascades. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

To be quite frank, MAP kinase cascades are very confusing. There are a lot of proteins involved, and the names are not really very creative, which makes things worse. Each of the kinases that gets activated in the cascade has the capacity to go on and activate many other proteins and enzymes, effectively amplifying the signal inside the cell. Usually, scaffold proteins help hold together the correct MAPK proteins. With so many of them in the cell and the fact that they are structurally similar, it would be easy to activate the wrong ones without taking specific measures to avoid that.

The downstream targets that get activated by a MAP kinase cascade will be involved in many different aspects of cellular function and are likely to have both fast and slow responses, much like all of the rest of the signaling we've seen so far. To show you how these cascades can be used, here is an example of a MAP kinase cascade simulated by the growth factor EGF (Figure 07-15). We are lucky to also have a video representation of this cascade to help us visualize what's going on (Video 07-06).

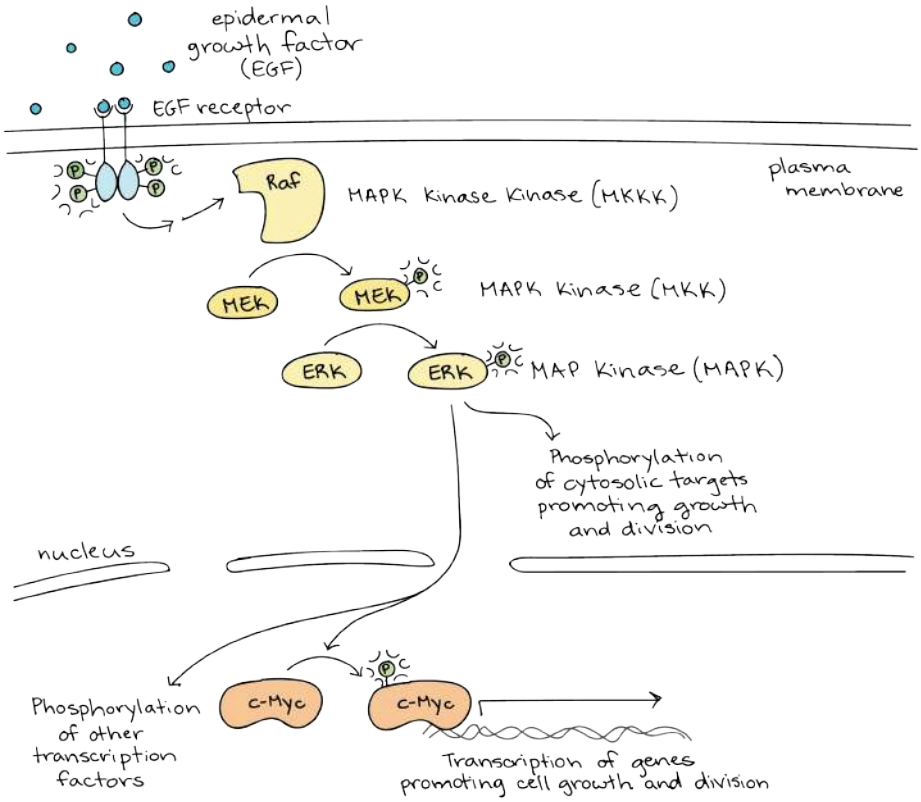


Figure 07-15: The EGF response pathway, showing the role of a MAP kinase cascade. From Khan Academy and shared under a CC BY-NC-SA 3.0 license. Note: All Khan Academy content is available for free at www.khanacademy.org.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=156#oembed-6

Video 07-06: The MAPK pathway. We start this video at 3:06 minutes in, as this is where the discussion of the MAPK cascade begins (using the EGF pathway as an example). It's worth watching this video more than once, as it is somewhat confusing due to the number of proteins involved.

Downstream Effects of Enzyme-Coupled Receptors—PI 3-Kinase

Another common mechanism that gets activated in an RTK pathway is an enzyme called **PI-3-kinase**. Phosphatidylinositol (PI) is a type of phospholipid that we've seen before. These phospholipids have a 6-carbon sugar (called **inositol**) as the head group. As we learned earlier in this topic, IP_3 and DAG are also produced from a phosphatidylinositol known as PIP_2. The reason these particular phospholipids are so useful in signaling is that the inositol ring is exposed at the surface of the membrane and has

three different sites exposed where phosphates can be added to it. As a result, you can have PIs with one, two, or three phosphates added in many different combinations. In addition, it is easy to switch the locations of the phosphate groups around. The job of PI-3-kinase is to add a phosphate group at the third carbon in the inositol ring (Figure 07-16).

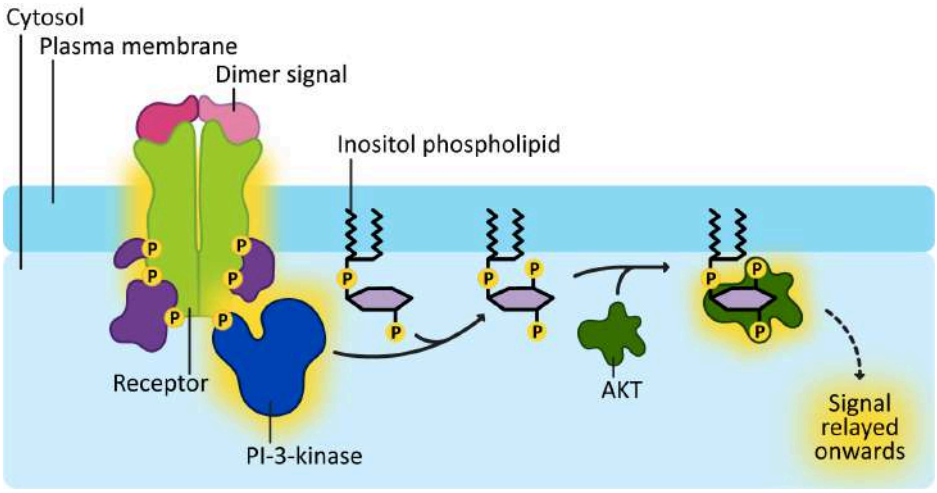


Figure 07-16: PI-3-kinase activation by RTKs. This causes a phosphate to be added on the 3-carbon of the inositol ring for phosphatidylinositol lipids. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The addition of the phosphate at the third carbon results in the production of specific docking sites for a variety of proteins. For example, Figure 07-16 shows a protein called AKT docking on the modified PI molecule. AKT is a signaling protein specifically involved in cell survival and suppression of programmed cell death. PI-3-kinase gets involved in all kinds of cellular functions, from the formation of vesicle budding sites, to growth and proliferation, to cell motility.

There Is Extensive Crosstalk between Pathways

The final point that is important to discuss, before we move on to experimental techniques, is a reiteration of how complex signaling really is. These pathways and mechanisms that we discuss do not function in isolation. Each and every cell on the planet is in a constant state of sending, receiving, and responding to signals from its environment. In fact, a cell that stops receiving signals from the environment for whatever reason will immediately die, as there will be nothing to stop programmed cell death (a.k.a. **apoptosis**) from occurring.

Very few signaling pathways are linear. They are more like an intricate and interconnected web, where each part can be influenced by many other parts. One signaling pathway could make another pathway easier or harder to turn on/off. To end this section, we offer an *extremely simplified* representation of a tiny subset of the signaling that can occur in a cell (Figure 07-17). We encourage you to explore and try to recognize the different parts of the signaling that we have explored in this chapter. Also look for the visual cues of activation and inhibition that we discussed at the end of Topic 7.1, which is indicated in the figure by either a pointed or blunt arrow.

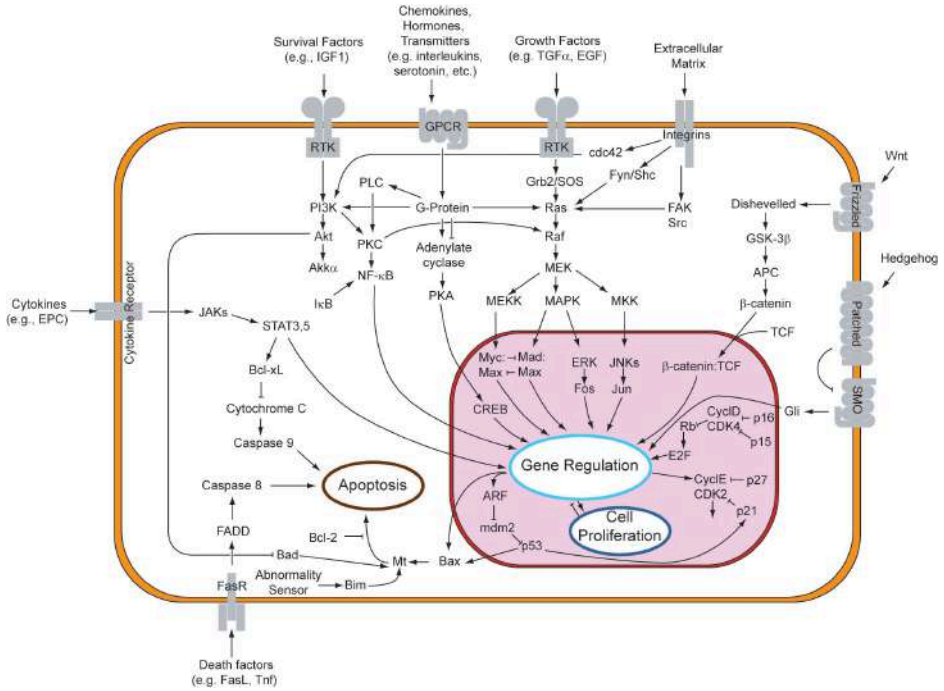


Figure 07-17: Some common signaling pathways in mammals. BohHog2. This image is in the public domain.

Once you've done that, if you'd like to see a more realistic version of the complexity of signaling, the Cell Signaling Technologies website has an excellent set of signaling pathways that you can explore. (*Hint: We recommend starting with insulin signaling or the regulation of actin dynamics, as they have most/all of the components we've explored, but there are lots of options.*) It's excellent practice to look at pathways you've never seen before and try to figure out what you can recognize. Once you start being able to recognize the various parts and proteins of a signaling pathway, it becomes a little bit less scary.

Studying Cells: Investigating the Effects of Phosphorylation in Signaling

You may remember from earlier in the chapter that phosphorylation and dephosphorylation of proteins are very common posttranslational modifications that are used in signaling to regulate protein activity. You may also remember that phosphate groups are added to amino acids by kinases using ATP as the source of both the activation energy and the phosphate group. Most commonly, kinases will add phosphate groups on one of three different amino acid side chains: serines, threonines, and tyrosines (Figure 07-18). In signaling, phosphorylation of the amino acid residues in preexisting proteins allows for quick responses to external cues, which can be easily reversed by later removing the phosphate.

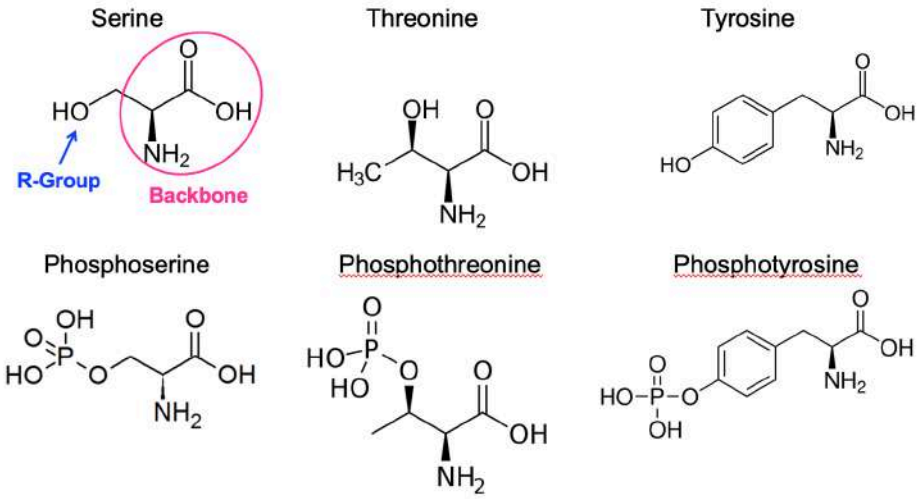


Figure 07-18: Phosphates are added by kinases to one of three different amino acid R groups: serine, threonine, and tyrosine. All the amino acids in the figure are aligned the same way, with the backbone atoms on the right and the R group on the left, as labeled for serine. The top row is the original amino acid, without the phosphate. The bottom row is the phosphorylated version. Note that in each case, the hydroxyl group on the R group reacts to allow the phosphate to be added. Image created by Dr. Robin Young and is shared under a CC BY-SA 4.0 license.

Signaling researchers spend quite a lot of time trying to understand the function of each phosphorylation site on a protein. Dissecting the contribution of an added phosphate to a protein's overall activity is technically challenging. Bioinformatic analysis has given us the ability to know how many serines, threonines, and tyrosines are present in the protein sequence. However, it cannot tell us which of them are modified by kinases and what the effect of that modification would be on the function of the protein. That's where experimentation comes in. We will explore two different approaches that are currently use by scientists: **phosphomimetics** and **genetic code expansion (GCE)**.

Phosphomimetics Attempt to Copy Specific Conformations of Phosphorylated Proteins

In this approach, cell and molecular biologists genetically engineer proteins that have been modified in a particular way using a technique called site-directed mutagenesis. They can use site-directed mutagenesis to specifically replace an amino acid that would normally get phosphorylated with a different amino acid intended to "mimic" the phosphorylated state. Most commonly, negatively charged amino acids are used to replace the phosphorylated amino acid, such as aspartic acid (Asp) or glutamic acid (Glu). The negatively charged Asp and Glu residues are somewhat similar in size and charge to the phosphorylated amino acid, as can be seen in the example in Figure 07-19. The altered proteins have the added advantage that they cannot be dephosphorylated like a phosphorylated amino acid could be. As such, they remain in the "phosphorylated" conformation, which gives us a chance to study the effects of the "phosphorylated" protein in the cell.

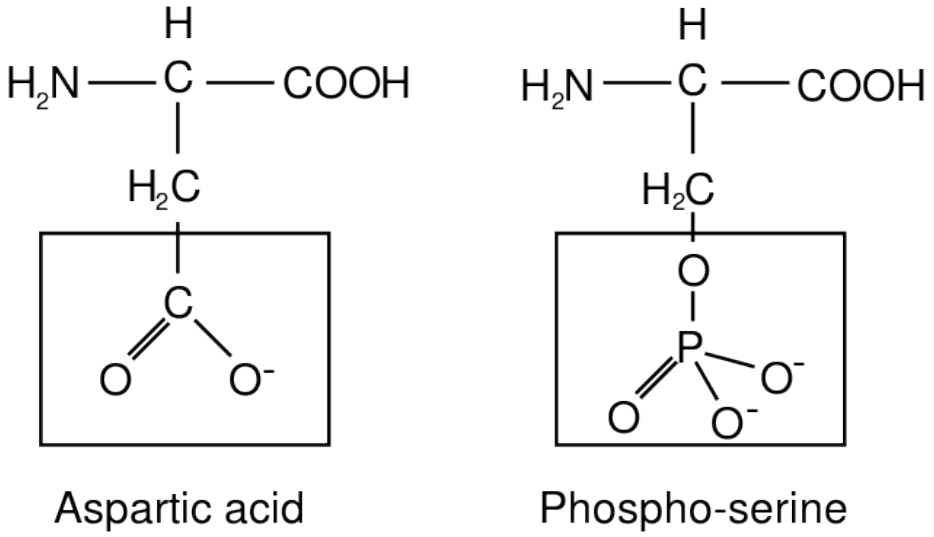


Figure 07-19: Aspartic acid mimics the structure and charge of a phosphoserine residue. They are similar in size and charge to each other, though not identical. Often, the charge allows them to "mimic" the phosphorylated version of the protein in the cell, and so give us an idea of how phosphorylation at a specific site impacts the function of our protein of interest. Image by Clarkdn and is shared under a CC BY-SA 3.0 license.

We can also mimic the dephosphorylated form of the protein by genetically engineering an alanine substitution at the site where a serine or threonine would have been originally. Since alanine cannot be phosphorylated, it blocks phosphorylation at that site and "mimics" the unphosphorylated version of the protein.

While these approaches have proven valuable in probing the biological function of phosphorylation of specific residues in some proteins, these approaches also have limitations. As you know from Chapter 2, the primary amino acid sequence determines folding and structure. Replacing one amino acid with another will never be a perfect replication of either the phosphorylated or unphosphorylated form of the original amino acid. The chemical properties of aspartic and glutamic acid are significantly different from serines, threonines, and tyrosines in a variety of ways, including electronegativity, volume, polarity, and hydrogen bonding. All of these chemical properties contribute to how a protein will fold and thus how it will function. Further, the nonphosphorylatable alanine amino acid has a much smaller R group than any of the amino acids it's meant to replace, and since it's nonpolar, it lacks their capacity to form hydrogen bonds. Thus, a different approach was devised.

Genetic Code Expansion (GCE) Allows Us to Code for Phosphorylated Versions of Amino Acids

To get around the problems that can arise with phosphomimetics, researchers created a new way to build phosphorylated versions of proteins without requiring the use of amino acids that "mimic" the original function. GCE is a technique that allows us to treat the phosphorylated version of the amino acid as its own entity within the genetic code so that we can insert it directly into the protein sequence. GCE is still a relatively new technique, but it is proving to be very powerful as we seek to learn more about posttranslational modifications, such as phosphorylation, and their role in protein function. Oregon State University is a world leader in this technique through their research center, the GCE4All Center.

The essence of this technique depends on how cells read and interpret the genetic code. The idea is that we "hijack" and modify how the cells interpret specific codons of the genetic code so that

they now code for new, unique amino acids. Of course, the genetic code has 64 codons that code for 20 amino acids, and it is interpreted virtually identically across the entire tree of life. However, there are exceptions, most commonly in bacteria, where some codons are interpreted differently and code for different, unique amino acids. In these organisms, we have a different, *expanded* set of amino acids with which to create proteins. In GCE, we genetically engineer this exact same scenario to include specific new amino acids of our own choosing. For example, we could hijack a codon to code for phosphoserine instead of regular serine so that a phosphorylated version of the protein can be translated directly.

While the premise for GCE is relatively straightforward, hijacking the codons and making them code for a different amino acid requires careful manipulation of several components of the protein translation machinery. Remember that a key component of proper protein translation is the correct addition of the amino acid to the tRNA (for a refresher on translation, see this link from the Khan Academy). So if we want to change which amino acid a particular codon codes for, we must find a way to change which amino acid is attached to that codon. Adding the amino acid to a specific tRNA is done by a family of proteins known as **amino acyl tRNA synthetases**. Changing the genetic code begins with genetically modifying these synthetases so that the amino acid that is added to the tRNA is altered. Figure 07-20 provides a visual representation of this. In the figure, one of the stop codons (UAG) has been repurposed to code for phosphoserine (pSer in the figure). Once you have created this machinery, you can use site-directed mutagenesis to add your newly repurposed codon in any site in the genetic sequence where you want the new amino acid to be added, and the translation machinery will do the job for you. It's worth noting that this system is mostly done in test tubes (i.e., *in vitro*), as changing the codon language used to create every protein in a cell would likely be detrimental and kill the cell.

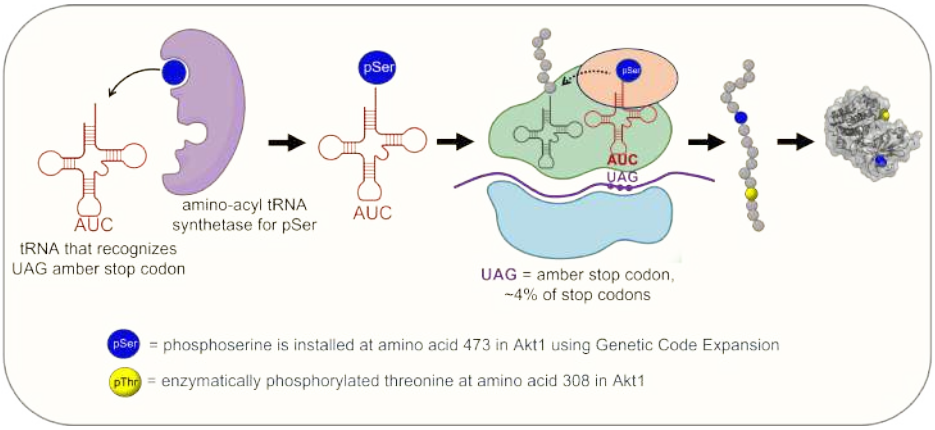


Figure 07-20: Schematic of genetic code expansion through the repurposing of a stop codon (UAG) so that it now codes for phosphoserine (pSer). To achieve this, an amino-acyl tRNA synthetase was genetically modified so that it could add a pSer to the UAG tRNA. Once this system is complete, genes can be modified to include UAG at a site where a pSer is desired, and it will be translated that way. This schematic was made by Dr. Richard Cooley and Dr. Kari van Zee in the GCE4All Research Center at Oregon State University and is shared under a CC BY-SA 4.0 license.

The advantage of this technique, compared to using phosphomimetics, is that you can add an actual phosphoserine rather than using a closely approximated amino acid, such as aspartic acid. Researchers do not know which kinases are involved in phosphorylation of a protein at a particular site. Using

this method, they can create a phosphorylated version for use *in vitro* and then examine the protein's activity by looking at its effect on known downstream targets.

GCE can be and has been used to address a wide variety of research questions that go far beyond the phosphorylation research described here. Many research areas have benefited from GCE, such as cell signaling, disease regulation, protein localization, protein-protein interactions, therapeutic development, and novel biomaterials and sensors. Its strength is in the fact that the novel amino acids that can be introduced into the genetic code are almost endless as long as the translation machinery can still work with them. They have been used to study topics such as posttranslation modification of proteins, linking two proteins together, adding on a stable fluorescent dye for microscopy in live cells, and other forms of labeling proteins to learn more about their structure. There are some examples of research that have been supported by Oregon State's GCE4All Center, which gives you an idea of the versatility of this technique.

CHAPTER SUMMARY

In this chapter, we explored how cells communicate and respond to their environment through cell signaling. In the first topic, we explained the fundamental anatomy of a signaling cascade, from the properties of a signal molecule, to the receptors to which it binds, to the intracellular signaling molecules, and finally to the types of responses that can be elicited from any given signal. Then we delved into specific examples of signaling systems and common pathways you might encounter in past and future coursework. Finally, we explored how we can use a couple of different genetic techniques to study how posttranslational modifications affect signaling cascades to understand the regulation of these complex systems. Hopefully, after reading this chapter, you will agree that this fundamental cellular process is as beautiful as it is complex.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

Topic 7.1: Introduction and Cell Signaling Overview

1. We described five major types of signaling used by multicellular organisms. What are they? What are the benefits of each one? How do they impact the affinity of the receptors that are used in this form of signaling?
2. Explain the difference in the chemical properties of signals that use internal receptors compared to those that are restricted to using cell-surface receptors.
3. Describe the types of responses that can be elicited in cells when a signal is received.
4. Compare the pros and cons of fast versus slow response rates for a signaling pathway.
5. List the pros and cons of long-range (endocrine) signaling as opposed to short-range (contact or paracrine) signaling.

6. Explain/define the following terms and know the function and significance of each:

    a. cell-surface receptors
    b. signal transduction
    c. second messenger
    d. cell response
    e. endocrine signaling
    f. paracrine signaling
    g. contact-dependent signaling
    h. synaptic signaling

Topic 7.2: Examples of Common Signaling Mechanisms

1. Describe the activation of an ion-gated channel.
2. Explain why it can be said that ion-gated channels convert chemical energy into electrical energy.
3. Name some examples of signal molecules that bind to receptor tyrosine kinases.
4. Describe activation of a G-protein-coupled receptor (GPCR) pathway from the binding of a signal at a GPCR to the activation of downstream signaling molecules.
5. Name and describe common potential downstream signaling molecules/events in a GPCR pathway.
6. Explain/define the following terms, know which ones are the same, and know the function of each:

    a. GPCR
    b. heterotrimeric G-protein
    c. hydrolyzation (hydrolysis)
    d. IP_3
    e. DAG
    f. adenylyl cyclase
    g. CAMP (cyclic AMP)
    h. phosphomimetic
    i. phosphorylation
    j. posttranslational modification
    k. genetic code expansion (GCE)
    l. receptor tyrosine kinase
    m. autophosphorylation
    n. Ras, structure and features of Ras
    o. MAP kinases

p. scaffold proteins

7. Name some examples of signal molecules that bind to receptor tyrosine kinases.
8. Explain the events following the binding of a signal to a receptor tyrosine kinase and how these events bring about change in cells.
9. Name and describe common potential downstream signaling molecules/events in a receptor tyrosine kinase pathway.
10. Describe a mechanism that allows the specificity of signaling pathways despite redundancy of signaling molecules.
11. Describe how phosphorylation is an important posttranslational modification in signaling pathways.
12. Explain how phosphorylation affects proteins chemically.
13. List some pros and cons of using phosphomimetics or GCE to study the role of phosphorylated proteins in signaling pathways.

CHAPTER 8.

THE CELL CYCLE AND MITOSIS

INTRODUCTION

For most of this textbook, we have looked at each structure and process separately so that we can better understand them. We have focused on how the underlying chemistry of structures is essential to understanding the function of each organelle and each individual process. Here we are finally able to integrate all of the different bits of the cell and see how it all comes together as a whole and drives complex cellular processes. More specifically, we will be exploring the process of cell growth and division. We assume that this is not your first exposure to the concepts of cell growth and division (also known as **mitosis** and **cytokinesis**), and we approach the topic from that lens. If you need a refresher on the basics, we encourage you to return to the introduction and explore the review material.

Cell growth and division are moments within a larger process known as the **cell cycle**. The cell cycle impacts every aspect of cellular function. Every organelle gets involved in some way or another at every step of the way. Growth requires protein and membrane synthesis, energy production, transport of materials, and more. Selecting the correct moment to undergo cell division requires careful coordination of both the internal and external environment of the cell. Cells that divide at the wrong moment are unlikely to survive the process, so it must be carefully timed and controlled. Mitosis itself requires a complete and total disruption of cellular function as well as a complete but temporary rearrangement of the cell's contents. This, too, requires precise coordination of the organelles. In this, the final chapter of this textbook, we will explore how the cell coordinates all of the various processes we've learned about so far in order to grow and then, at just the right moment, divide into two daughter cells.

TOPIC 8.1: REGULATING THE CELL CYCLE: CHECKPOINT CONTROL

Learning Goals

- Review the stages of the cell cycle, including the checkpoints, and identify the key features of each stage.
- Describe how specific protein modifications (e.g., phosphorylation and ubiquitination) result in activation/deactivation of cyclin-CDK complexes to regulate cell cycle checkpoints.
- Explain how the activation of the cyclin-CDK complexes results in the start of the next phase of the cell cycle. Use examples from both M- and S-cyclin-CDK complexes to explain this.

- Detail how fluorescence-activated cell sorting (FACS) can be used to identify the stage of the cell cycle for a population of cells.

Introduction to the Cell Cycle and Checkpoints

A discussion of the **cell cycle** and mitosis is a very good way to end this book, as it is a wonderful example of how the concepts we've covered in this book are interconnected. The progression of the cell from interphase to cell division is precisely regulated, and it involves every other cellular component in some way.

The cell cycle is defined as the events that enable cells to proceed from one cell division event to the next. Cell division itself consists of the overlapping processes of **mitosis** (nuclear division) and **cytokinesis** (division of the cytoplasm).

The cell cycle is divided up into four separate phases based on the primary event that is taking place in that stage:

- **G_1 (gap or growth 1) phase**: This is the "gap" between the end of cytokinesis and the start of DNA synthesis. A lot of the work of this phase involves cell growth so that it can support itself and also have the resources it needs for the next phase.
- **S (synthesis) phase**: This phase is defined by the initiation and termination of DNA synthesis.
- **G_2 (gap or growth 2) phase**: This second "gap" phase lasts from the end of DNA synthesis to the onset of mitosis. The cell continues to grow but also prepares for what's to come in the next phase.
- **M (mitosis) phase**: This is the phase in which cell division occurs.

Figure 08-01 shows an overview of the stages of the cell cycle. Collectively, we consider G_1, S, and G_2 to be **interphase** (i.e., the phases "in between" M phase).

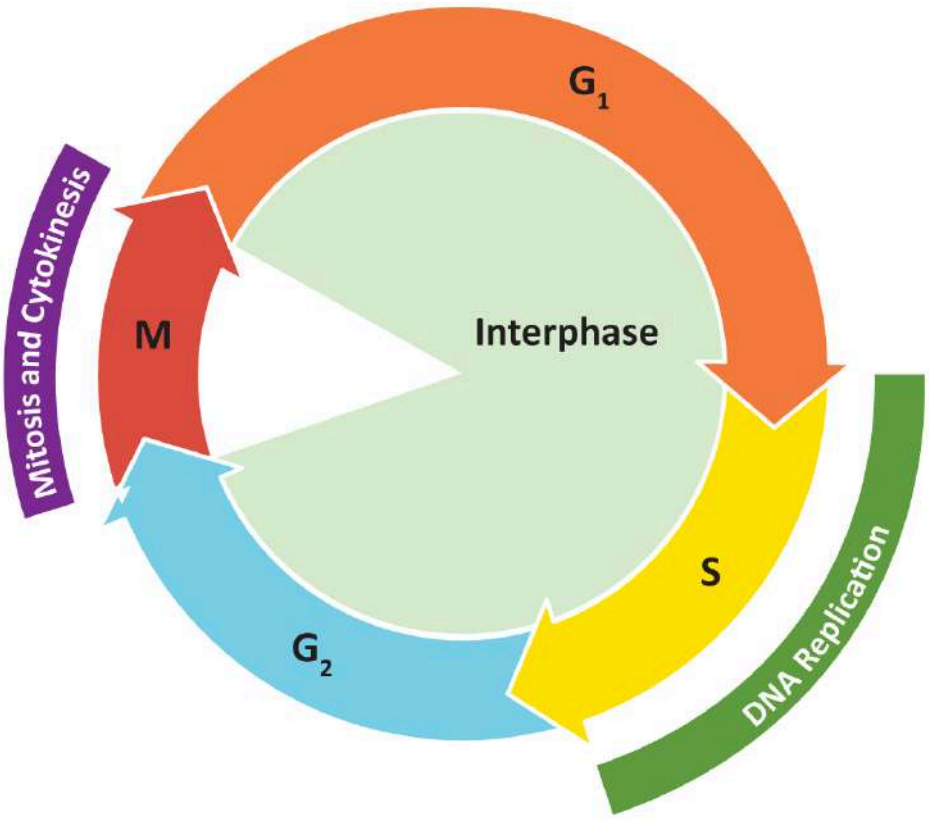


Figure 08-01: The four stages of the cell cycle. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

It is important to keep in mind that while cells do need a mechanism to control growth, that is not always their primary concern. There must also be provision for cells to "step away" from the cell cycle for a time and into other developmental pathways (mating, meiosis, differentiation). Some cell types will hit a point in their development where division isn't a good option anymore for one reason or another, so they withdraw from the continuous cycle of growth and replication. Whether the cell leaves the cell cycle on a temporary or permanent basis is also highly controlled, and we will discuss how the cell makes these kinds of decisions later in this topic. (*Hint: Signaling proteins are involved.*)

Signaling and regulation are at the heart of the proper progression of the cell cycle. Using these tools, the cell ensures that

- discrete events, such as DNA synthesis and mitosis, do not occur before the cell is ready, and they occur in the right order
- the cell pauses DNA replication or mitosis when errors are identified and repairs are attempted
- only cells that *should* divide are allowed to do so
    - Some cells must undergo **terminal differentiation**, stall temporarily or permanently, or even go through programmed cell death (i.e., apoptosis). These are all cell cycle decisions.
- growth is coordinated with cell division so that the size of cells is maintained over the many cellular generations
    - Some cells are designed to get bigger or smaller over several cell cycles. Some even

divide asymmetrically so that one large and one small daughter cell is produced. All of these will require precise cell cycle control.

One of the most important controls placed on the progression of the cell cycle is a series of **checkpoints** in which the cell is required to meet certain criteria before it is allowed to proceed. In this way, these checkpoints act as a form of quality control. The control of the checkpoints and the mechanism of when/how cell division takes place is a perfect example of how the cell uses signaling to understand its environment and to effect internal change. Cells do not pass through these checkpoints randomly. They are constantly receiving cues from the exterior (such as growth factors, for example) and from the interior that help them decide exactly when and how to divide.

Each of these checkpoints is controlled by one or more "gatekeeper proteins" that respond to cellular conditions and will only allow the cell to move forward into the next phase of the cell cycle if conditions are "right." As a result, the checkpoints help ensure that specific criteria are met before the cell cycle is allowed to continue. There are a few different checkpoints, but the ones we are going to focus on are the following: the **G_1/S checkpoint**, which allows the cell to pass into S phase, and the **G_2/M checkpoint**, which controls when the cell enters mitosis. We will also look at the checkpoint that is in the middle of mitosis, which ensures that the mitotic spindle is set up correctly prior to chromosome separation. We call this the **metaphase checkpoint**, but we have seen many other names used as well (spindle assembly checkpoint, M phase checkpoint, M/G_1 checkpoint, etc.).

Cells can only move through the checkpoint and into the next stage of the cell cycle when they have met the required conditions. For example, if biotin, a vitamin, is missing from the growth medium, yeast cells will not pass the G_1/S checkpoint even if all other conditions are perfect. Low nutrient levels reduce the growth rate, which, if severe enough, can make it so a newly divided cell will not survive. Thus, having the ability to confirm that everything is in place *before* dividing is key to survival for the cell.

G_1/S Checkpoint

The first checkpoint a new cell will encounter is the **G_1/S checkpoint** (also sometimes known as the "restriction" point or "Start"). As its name implies, this checkpoint marks the transition from G_1 to S phase. Since S phase involves the replication of DNA, it is important that the cell and the environment are both ready before replication starts. Not only is this an energy- and nutrient-intensive process, but replication is when any preexisting errors in DNA become permanent mutations, so the DNA must be in good shape before the cell starts this process. Once again, making an error with the timing of replication could result in the death of the cell, so the checkpoint plays a key role here.

Cell signaling is important for ensuring that conditions are ideal for cell division. The cell responds to internal and external cues in order to "decide" when to divide. Some of the conditions that must be met include the following:

- Proper nutrients (carbon source, energy source, inorganic phosphate, nitrogen, vitamins, etc.) must be present at specific concentrations.
- Sister chromatid separation (from the previous mitosis) must be complete.
- There must be no detectable DNA damage.
- The cell must have reached a critical threshold size.

Additionally, external factors must also be appropriate. For example, in yeast, if the appropriate mating factor is present in the environment, the cells cannot proceed to S phase and are switched instead into an alternative pathway (called the sexual pathway). Similarly, in mammalian cells, appropriate growth factors must be present to allow cells to pass this checkpoint. If not, the cell remains in G_1.

In some cases, cells are stalled for extended periods of time…maybe indefinitely. We say that these cells have *removed* themselves from the cell cycle and that they are in **G_0 phase**. This could be due to a long-term deprivation of nutrients or other resources required for cell division. More commonly, however, this is a normal part of the development of certain cell types. For example,

- **Stem cells** for specific tissues will enter G_0 for short periods of time until replacement cells are required. This allows tissues to grow to a certain size and then stop growing and maintain a relatively stable size and distribution. Your blood cell system (called the *hematopoietic system*) does this. New blood cells are grown only when specific cell types are needed or cells are lost through injury.
- Some cell types undergo **terminal differentiation**, which means that when these cells reach maturity, they do not need to go through mitosis anymore. Good examples of this are *muscle cells* (multiple muscle cells fuse together at maturity to produce multinucleate muscle fibers), *neurons* (with their extremely long axons), and *osteocytes* (bone cells, which are intricately embedded in the calcified matrix of the bone).

G_2/M Checkpoint

The second is the **G_2/M checkpoint**, which stops the cell from entering mitosis before its ready. It is also sometimes called "CD" or "Commitment to Division," as the cell cannot stop the process of mitosis once it has passed this point.

As its name suggests, this checkpoint controls the transition from G_2 to M phase. Some conditions that must be met here are the following:

- DNA replication must be complete and accurate. No DNA damage can be detected (through a robust biochemical surveillance system) or this checkpoint cannot be passed. This is the most important factor for passing the G_2/M checkpoint.
- The cell must also have reached a certain minimum size so that it is big enough that, when split in two, the two daughter cells will also be large enough to survive.

It is considered exceedingly rare that cells would stall and enter G_0 from G_2. One would think that there is not really much point in doing all of the work of replication unless there is an intention for the cell to complete mitosis. However, since this is biology, there are examples of chromosomes called polytene chromosomes that can have thousands of sister chromatids, instead of simply two. These cells undergo repeated rounds of replication without moving forward to mitosis. It is thought to help increase the number of copies of genes in the cell, which can significantly impact gene expression. While this is thought to be relatively common in some cell types, such as the salivary glands of certain insects (like flies), and examples can be found throughout eukaryotes, it is still considered to be a relatively rare occurrence overall.

Metaphase Checkpoint

This checkpoint marks the halfway point of mitosis, but it's also the point right before the actual division of the genetic material, so it makes sense that this would be a point that the cell would verify before proceeding. To separate the sister chromatids of the chromosome, the **mitotic spindle**, composed of microtubules and associated motor proteins, must be assembled, and the chromosomes must be properly attached. If the spindle is not assembled correctly, an entire chromosome could get destroyed or mislocalized. Considering how important the genetic material is to the cell, you can imagine how bad this kind of "mitotic misfire" would be; both daughter cells would likely die, if they could even complete mitosis. Interestingly, you can see this checkpoint if you watch closely as cells divide under the microscope (Video 08-01). The chromosomes line up at the **metaphase plate** and then wait at the checkpoint for a little while until suddenly the sister chromatids split and move to opposite ends of the spindle.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=158#oembed-1

Video 08-01: Several examples of mitosis observed with live fluorescence imaging. In all cases, DNA is labeled in red and often looks like small worms in the center. Microtubules are labeled in green and look like long filaments. These videos are sped up and do not show mitosis in real time.

Cell Cycle Checkpoint Control: Cyclins and CDKs

Like all things in cell biology, there was a time when the details of the cell cycle were not known to scientists. Mitosis was observed extremely early (as early as the 1600s), as even the most rudimentary light microscopes could be used to observe it. The fact that cells arose from other cells was identified in the 1800s, but it wasn't until the 1950s that the rest of the cell cycle was suggested. By the late 1960s, scientists believed that "something" in the cytoplasm was controlling the cell cycle, but they had no proof. Dr. Yoshio Masui, a Japanese Canadian researcher at the University of Toronto, ran some experiments while working as a postdoctoral fellow at Yale University that provided the evidence needed. In these experiments, Dr. Masui extracted cytosol from frog's eggs that were in mitosis and then injected it into a cell that was stalled at the end of interphase. (*Aside: This is a normal part of frog oocyte maturation.*) As a control, they compared this with oocytes that were injected with cytosol from another oocyte also in interphase. They found that when the cytosol from a mitotic oocyte was injected into the interphase oocyte, a mitotic spindle would begin to form. This did not happen in the control, in which interphase cytosol was injected into an interphase oocyte.

The results of these experiments provided evidence of the following:

1. The active agent that promotes the next stage of the cell cycle is in the cytosol.

2. Control of DNA synthesis and mitosis is *positive*—that is, the active agent *promotes* mitosis in the recipient cell.
3. Cells can be advanced into the next stage before they planned it by adding the appropriate factors to their cytosol.

After this initial discovery, it was replicated using cytosol from mitotic cells from many different species and by taking cytosol from one species and using it to induce mitosis in another species. This meant that the "factor" in the cytosol was universal and helped promote the "maturation" of the cell. As a result, they called this molecule the *maturation promoting factor (MPF)*. It was called a factor, and not a protein, because at this point, no one knew what it was. About 20 more years of research were required to figure out that the "factor" in question was actually a protein.

MPF = Activated Cyclin-CDK Complex

We now know that MPF is actually a set of proteins that work together to control the checkpoints and thus the entire cell cycle. The active agent is a protein called **cyclin-dependent kinase (CDK)**. As you remember from Chapter 7, a **kinase** is an enzyme that specifically adds phosphate groups to other proteins. This particular kinase is only active in the presence of a second protein known as **cyclin**. Cyclin binds to CDK (Figure 08-02), and acts as a regulatory unit; CDK can *only* perform its function as a kinase when cyclin is bound. If cyclin is removed from CDK, then CDK is inactivated. This is also the source of its name…cyclin-*dependent* kinase. These two proteins combine to produce a single enzyme called the **cyclin-CDK complex**. The cyclin-CDK system is somewhat unique due to the use of cyclins to regulate CDK activity. The cell controls the *activity* of CDKs by controlling the *synthesis* and *destruction* of cyclins.

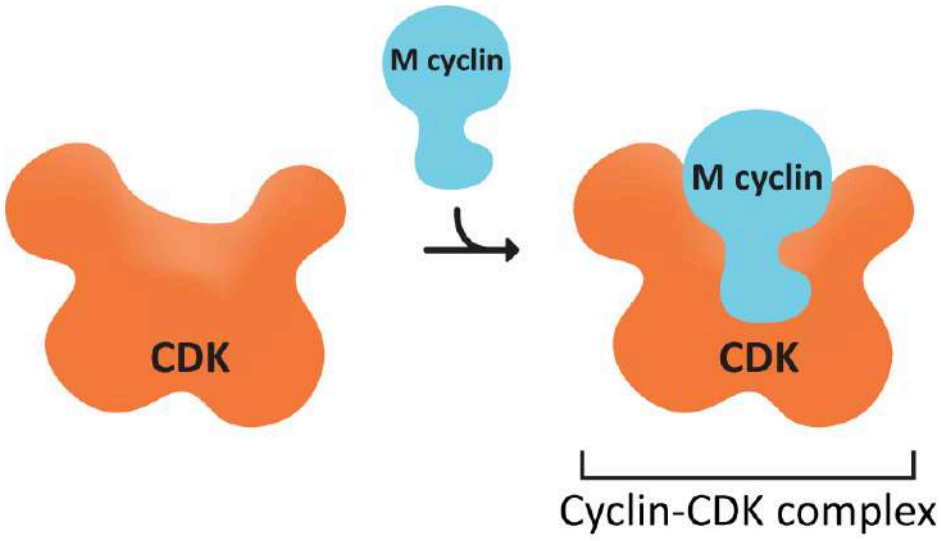


Figure 08-02: Cyclin and cyclin-dependent kinase (CDK) work together to control the cell cycle checkpoints. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

CDK concentrations in the cell remain constant throughout the cell cycle, but they are not always active. On the other hand, cyclin concentrations show a cyclical pattern. They increase as the cell moves through the cell cycle, which coincides with increasing enzymatic activity of CDK, peaking at

the appropriate point in the cell cycle (usually a checkpoint). After the relevant checkpoint is passed, cyclin concentrations crash down to almost nothing. Once again, this change in cyclin concentration coincides with the end of the CDK enzymatic activity. It is this "cycling" of the cyclin concentrations that regulates CDK activity, allowing the cell to progress through checkpoints of the cell cycle.

Since the initial studies last century, we have learned that there are actually several classes of cyclins and CDKs, and each of these classes is responsible for controlling a specific part of the cell cycle. The four major classes of cyclins are listed in Table 08-01.
<ta

Table 08-01: The different classes of mitotic cyclins

MITOTIC CYCLIN CLASSES	
Cyclin class	*Function*
G_1 cyclins	Unique cyclins that are thought to help the cell respond to external signals to leave G_0 and initiate cell division. We won't discuss these any further in this book.
G_1/S cyclins	Cyclins that control the G_1/S checkpoint and control the transition from G_1 to S phase.
S cyclins	Cyclins that activate at the start of S phase (by the G_1/S cyclins) and directly induce replication of DNA. The concentration of these cyclins remains high right through to M phase.
M cyclins	Cyclins that control the G_2/M checkpoint. They remain active in the first half of mitosis until their destruction is signaled by the anaphase-promoting complex (APC).

Increasing cyclin concentrations occur directly prior to a cell passing a cell cycle checkpoint. This is because the cyclin-CDK activity must hit a certain level before the checkpoint can be passed, and cyclins are required components of CDK activity. In some cases, such as the M cyclin shown in Figure 08-03, the activity of the CDK decreases rapidly not long after the checkpoint has been passed, but in others, like the G_1-cyclin shown here, the activity of the CDK is activated at one checkpoint and then stays high throughout the rest of the cycle.

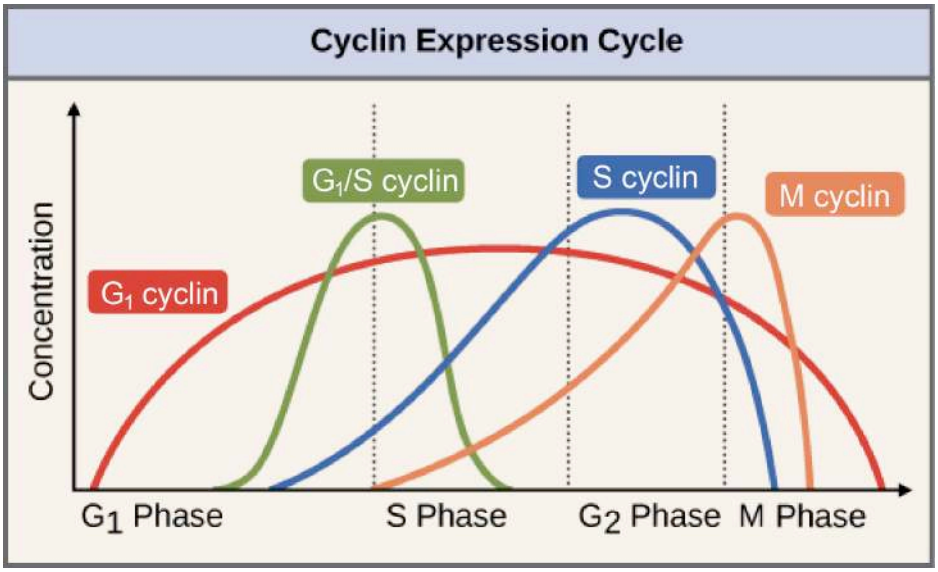


Figure 08-03: The concentration curves of the different cyclins. The three checkpoints are indicated by the vertical dashed lines. "Control of the cell cycle: Figure 2" by OpenStax College, Biology, shared under a CC BY 3.0 license, was modified by WikiMaMa.

Phosphorylation Controls Cyclin-CDK Enzymatic Activity

The activity of the CDKs must be very tightly controlled. The consequences of a cell moving to the next stage of the cell cycle before its ready could be disastrous. As such, cyclin-CDKs are at the heart of a complex signaling pathway involving potentially hundreds of enzymes that are fighting with each other to either activate or deactivate the cyclin-CDK complex. Figure 08-04 shows an extremely simplified version of a regulatory pathway that includes two different cyclin-CDKs (highlighted by yellow arrows). In particular, take note of how the pathways labeled *cell growth* (labeled as cell proliferation in the figure) and *programmed cell death* (also known as **apoptosis**) are connected to each other. This gives a strong hint about how important it is to get this right. If it goes wrong, the cell has the option to initiate apoptosis and die if necessary.

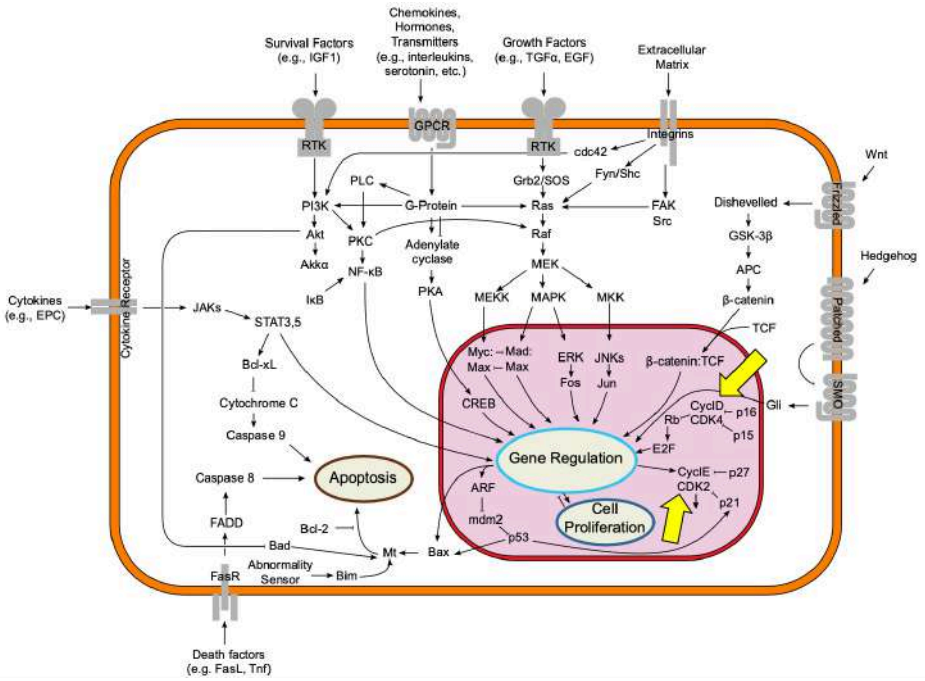


Figure 08-04: CDK2 (and its cyclin, CyclE) and CDK4 (+CyclD) within the signaling cascade that controls growth proliferation and/or apoptosis. "Signal Transduction pathways" by BohHog2 modified by Dr. Robin Young. This image is in the public domain.

Because CDK activation must be tightly controlled, there are multiple layers of regulation. Note that while the binding of cyclin to CDK is *necessary* for CDK activity, it is not *sufficient* for activation of the CDK on its own. (Remember our discussion of necessary and sufficient from Chapter 3.) This means that other factors are also needed to activate the CDK complex. In this case, the CDK-cyclin complex itself must also be phosphorylated (by other kinases). Interestingly, phosphorylation of CDK can also be used to keep the CDK inactive as well, depending on the location of the phosphate addition on the protein.

The activation of CDKs follows a set pattern, which requires both the presence of cyclin *and* proper phosphorylation of the CDK. We'll use the activation of the M-CDK/cyclin complex as an example, which is illustrated by Figure 08-05 and Video 08-02. In a nutshell, the binding of cyclin initiates the activation process. The signaling proteins upstream activate two kinases in particular: Wee1 and CAK. Both phosphorylate the CDK, but for different purposes. The phosphate group that Wee1 adds is considered to be an *inhibitory phosphate*, which limits the activity of the CDK. On the other hand, CAK (which stands for cyclin-activating kinase) adds a phosphate that is necessary for the CDK to become active. Once everything else is ready and it's time for the checkpoint to be passed, a third protein called cdc25 comes in and removes the inhibitory phosphate (thus cdc25 is a phosphatase), and the CDK becomes fully active.

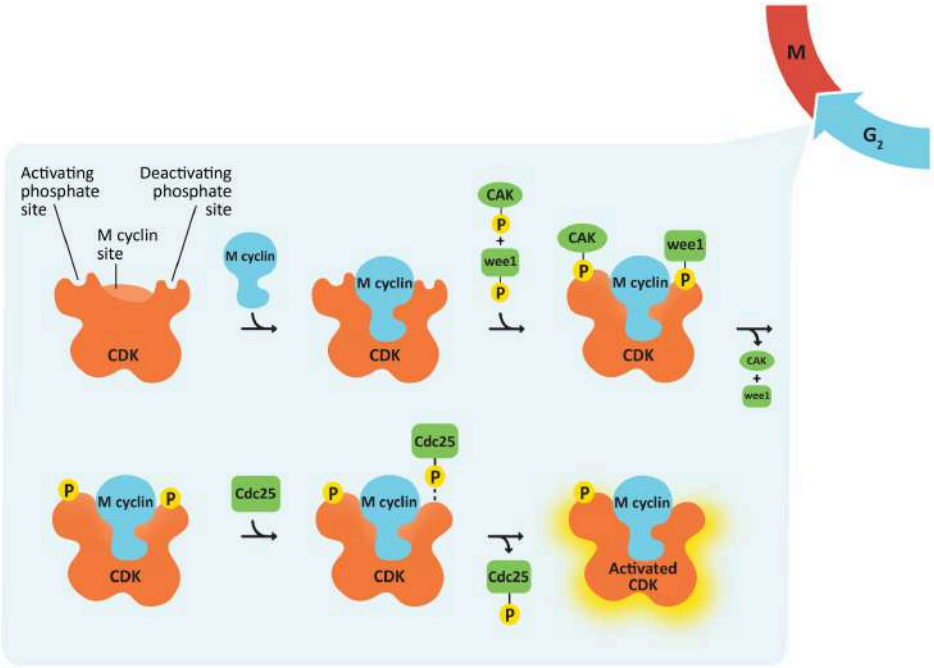


Figure 08-05: The activation of the cyclin-CDK complex by both the addition of phosphate at an activating site and the loss of a phosphate at an inhibitory site. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

You can think about this a little bit like a race car at the starting line of a race. By the time you get to the starting line, the car is all gassed up (cyclin) and running (activating phosphate). The drivers might even be revving the engines a bit, but with the break on (inhibitory phosphate) so that the car can't go anywhere until the signal is received that the race has started. By being ready to go before the race starts, they can leave the starting line as quickly as possible once the signal to start is received.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=158#oembed-2

Video 08-02: The molecular mechanisms required to initiate M phase.

In yeast, Wee1 and Cdc25 are key regulators of M-CDK. The *antagonistic* (i.e., opposing) relationship of Wee1 and Cdc25 was discovered using genetic experiments in which the "dosage" of the genes was experimentally manipulated. This simply means that the amount of protein present in the cell was increased or decreased depending on the mutation. Increasing the gene dosage (i.e., a **gain-of-function mutation**) increases the concentration of enzyme in the cell, whereas decreasing gene dosage (i.e., a **loss-of-function mutation**) leads to a decrease in enzyme concentration. The effects of these mutations on cell division are summarized in Table 08-02.

Table 08-02: The effects of mutations in *WEE1* and *CDC25* on yeast cell growth and division

THE EFFECTS OF MUTATIONS IN AND ON YEAST CELL GROWTH AND DIVISION		
Gene	*Gain-of-Function Mutation*	*Loss-of-Function Mutation*
***WEE1*—inhibitory kinase**	Cells divide later at a larger-than-normal size	Cells divide early at a smaller-than-normal size (*i.e., they're "wee"!*)
***CDC25*—activating phosphatase**	Cells divide early at a smaller-than-normal size	Cells divide later at a larger-than-normal size

*i.e., too much protein produced or produced all of the time without regulation
**i.e., no protein produced

Specific Cyclins and Their Role in the Cell Cycle

As mentioned earlier in this topic, we are focusing on three specific checkpoints in the cell cycle. Of those, two are controlled directly by cyclin-CDKs: the G_1/S checkpoint and the G_2/M checkpoint. Each of these checkpoints is controlled by its own cyclin-CDK complexes. Cyclins and CDKs involved in the start of mitosis are called M cyclins. The progression into S phase is more complicated, and multiple cyclin/CDK complexes are needed (see Figure 08-03). Interestingly, the metaphase checkpoint is indirectly controlled by cyclin/CDK complexes. The M cyclin-CDK complex also activates the process by which the metaphase checkpoint is set up and then passed. We'll look at how the different checkpoints work here.

Control of the G_1/S Checkpoint and S Phase Progression

The transition from G_1 to S phase is controlled by multiple cyclin-CDK complexes, including a G_1 cyclin-, an S-cyclin-, and a G_1/S cyclin-CDK complex (Figure 08-03). The G_1/S cyclin-CDK complex will be de-activated once the checkpoint is passed, but the others will continue to function. Once the G_1/S checkpoint is passed, DNA replication will begin.

One of the most important criteria for the passing of the G_1/S checkpoint is that there can be no detectable DNA damage. The protein that manages checking for DNA damage is a transcription regulator called p53. p53 is sometimes called "the guardian of the genome" because of the vital role it plays in ensuring that DNA remains intact and undamaged. It also functions in a way that is somewhat counterintuitive—in its *inactive* state, p53 is constantly translated and then immediately degraded (Figure 08-06). When DNA damage is detected, p53 gets phosphorylated, which stops it from being degraded. It then binds to the promoter sequence for a CDK inhibitor called p21, thereby activating its transcription and subsequent translation. p21 blocks the activity of the G_1/S cyclin-CDK complex, which will stop the cell from passing the G_1/S checkpoint.

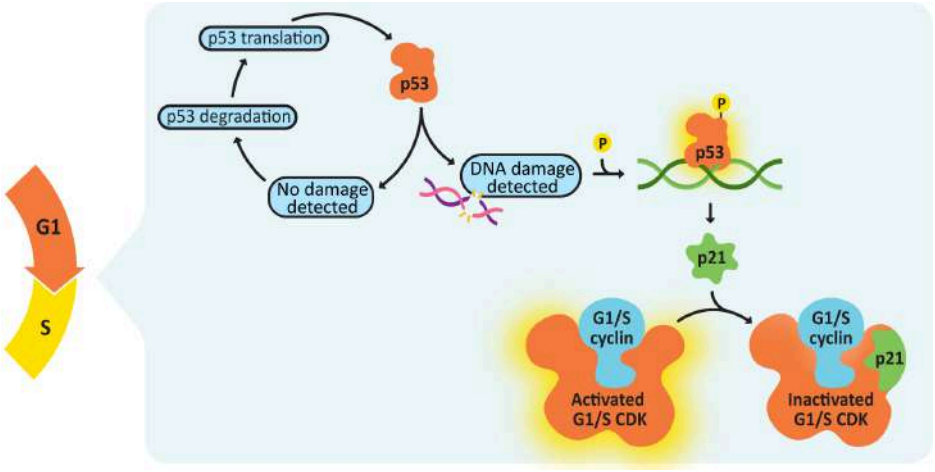


Figure 08-06: The protein p53 inhibits entry into S phase when DNA damage is detected. Presence of p53 upregulates the production of p21, an inhibitor of the G_1/S cyclin-CDK complex. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

It is also interesting to note that p53 is mutated in as many as *50% of all cancers*, which is a clear indication of just how important this protein is in the proper function of the cell cycle. If you lose p53 function in a cell, you lose your ability to delay the cell cycle so that there's time to repair DNA damage. At that point, damage is less likely to get fixed before replication, and mutations accumulate at a much faster rate, with each new round of replication.

Once the G_1/S checkpoint has been passed successfully, the active cyclin-CDK complexes help initiate S phase via phosphorylation of the replication machinery, which, in turn, helps it assemble on the DNA. *DNA helicase*—responsible for DNA unwinding—is also activated via phosphorylation, allowing replication to begin.

While there is much focus on how S phase is initiated, it is equally important to consider how it is terminated. DNA replication machinery must be deactivated at the end of replication as well. This ensures that the entire genome is replicated once, and only once, during S phase. Again, the cyclin-CDKs control this by phosphorylating key enzymes, which will lead to the eventual shutdown of replication.

Prior to the start of S phase, each chromosome in the genome was made of a single helix of DNA that is in a complex with histones to form a chromatin fiber. (Review Chapter 3 if needed.) It has its own **telomeres** at each end of the chromosome and a **centromere** near the middle of the DNA. During DNA replication, the DNA double helix is duplicated, and the chromosome now includes two **sister chromatids**. This requires that a whole new set of histones is synthesized and imported into the nucleus to form the new chromatin fiber. Also, the two sister chromatids must be held together until the point at which they separate, during mitosis. This is done using a protein we've seen before (in Chapter 3) known as **cohesin** (Figure 08-07). Cohesin attaches the two chromatids together along their entire length at specific sites on the DNA called **cohesin attachment regions (or CARs)**.

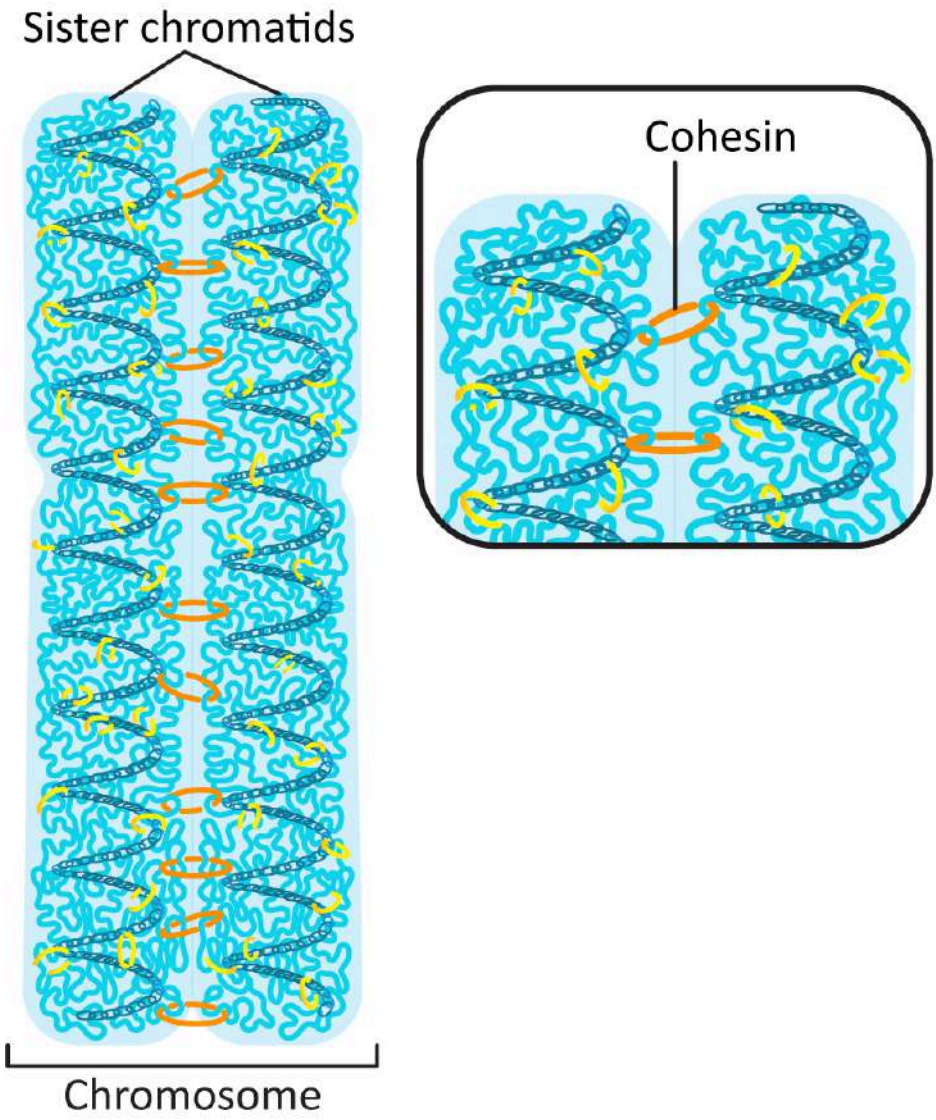


Figure 08-07: Cohesins holding the two sister chromatids together by binding the DNA (light blue) from two adjoining sister chromatids. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One interesting final point: S-CDK has been shown to remain active right until the start of mitosis (Figure 08-03) despite the fact that S phase is over long before that. The reasons for this are not entirely clear, but there is some evidence that S-cyclin-CDK helps with the activation of M-CDKs. This provides further continuity in the cell cycle and shows that the different cyclins are influenced by each other.

M-CDK Controls the G_2/M Checkpoint and Reentry into G_1

The transition from G_2 phase to M phase is complex. It requires an almost complete rearrangement of the cytoplasm, including shutting down all transcription and translation, preparing all of the organelles for separation (which often means deconstructing them), building a second microtubule organizing center (MTOC) for the mitotic spindle, and completely rearranging the cytoskeleton to allow for division. M-CDK-cyclin becomes enzymatically active at the end of the G_2 phase and is the primary control for this transition. M cyclin concentrations peak in metaphase, before crashing, which deactivates the M-CDK and marks the beginning of the transition back to G_1.

As a kinase, the role of M-CDK is to phosphorylate other proteins, which, in turn, will activate or deactivate them depending on the protein. Some of the targets of the activated M cyclin-CDK complex include the following:

- **Histone H1—**Phosphorylating H1 leads to changes in chromatin configuration and, in conjunction with other proteins, leads to the tighter packing of chromatin required for mitosis.
- **Condensins—**These are a class of DNA-binding proteins that bind to chromatin to help with higher-order chromosome condensation. They work to loop up the chromatin fiber into the tightly coiled mitotic chromosome.
- **Nuclear lamins—**Phosphorylated lamins have a lower affinity for each other, and as such, the nuclear lamina falls apart. Disassembly of nuclear lamina results in the breakup of the nuclear envelope. (We explored this in detail in Chapter 3.)
- *Structural proteins of the* **nucleolus—**Since DNA from multiple chromosomes is used to form the nucleolus, it must be taken apart prior to mitosis. Phosphorylation of the structural proteins results in dispersion of the nucleolar proteins and disintegration of the nucleolus.
- *A variety of protein kinases that regulate the cytoskeleton—*In order for mitosis to happen, the cytoskeleton needs to be completely taken apart and rebuilt. A number of proteins are involved in this, including microtubule-associated proteins (MAPs) and even some actin-binding proteins (ABPs), and must be activated directly or indirectly by the M-CDK.
- *cdc25, the M-CDK-activating phosphatase—*This creates a positive feedback loop, resulting in further activation of M-CDK-cyclin. As a result, M-CDK activity rises increasingly rapidly as more M-CDK becomes active (Figure 08-08).
- **Anaphase-promoting complex (APC)—**This protein is key to passing the metaphase checkpoint. At the beginning of anaphase, APC degrades the cohesin proteins that bind sister chromatids together, releasing the daughter chromosomes. Interestingly, APC also activates enzymes that tag M cyclin for degradation. This ensures that M-CDK will be properly deactivated when its task is done. The cell cannot complete mitosis and return the cytoplasm to its interphase state unless M-CDK is inactive.

We will be breaking down the events of mitosis and how CDKs and other proteins drive that process in the next topic of this chapter.

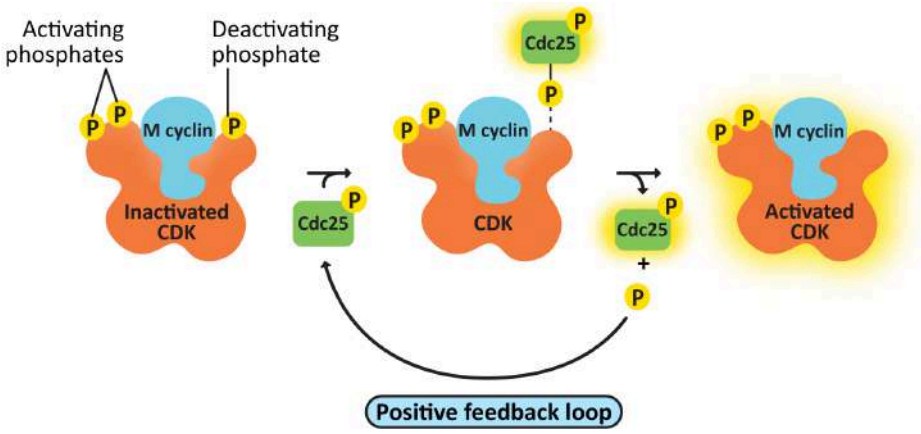


Figure 08-08: Positive feedback loop of Cdc25. Active M-CDK helps activate Cdc25, and this activation helps activate more M-CDK. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The Metaphase Checkpoint Ensures Proper Mitotic Spindle Formation

Unlike the other two checkpoints, this checkpoint is *not directly* controlled by a cyclin-CDK complex. However, that doesn't mean that the CDKs and cyclins have no role to play. This checkpoint is in the middle of M phase and ensures that the mitotic spindle is formed properly before mitosis is allowed to proceed. The M cyclins and their associated CDKs are involved in this checkpoint, as they activate the proteins required to create the mitotic spindle. Passing the metaphase checkpoint not only allows mitosis to proceed but also sets in motion a series of events that will eventually shut down M phase, which we will talk about in the next section.

Proper spindle assembly is essential to the metaphase checkpoint, so it stands to reason that the proteins involved in regulating this checkpoint would be interacting directly with the spindle itself. At some point during G_2 phase, a large protein complex known as a **kinetochore** assembles at the centromere of each sister chromatid in the chromosome. This complex will help capture microtubules during the assembly of the mitotic spindle so that the sister chromatids can be separated. Several proteins are also assembled at the kinetochore, the most important of which is one called anaphase-promoting complex, or APC, and they will remain there until microtubules are properly attached to the kinetochores of *both* sister chromatids. Once that happens, the checkpoint proteins assembled at the kinetochore, including APC, are released and activated. Once that happens, the cohesins holding the sister chromatids together are degraded and anaphase begins.

This is not the end of APC's role in mitosis, however.

Deactivation of the Cyclin-CDK Complex

Just like in S phase, the cell must eventually end M phase, put everything back where it was, and allow the new daughter cells to reenter G_1. To do this, the M cyclin-CDK must be deactivated so that it stops phosphorylating its target proteins and the cell can exit M phase. APC is directly involved in the degradation of M cyclins, which deactivates the M-CDK.

Incidentally, that means that M cyclin-CDK not only creates a **positive feedback loop** to activate itself, through cdc25 (Figure 08-08), but it also creates a **negative feedback loop**, using APC, which will result in its eventual deactivation. This is a fascinating case study on the complexity of cell signaling and how it regulates cellular function.

Like most signaling events, the cyclin-CDK complex must be deactivated once its job is complete. This is done by a *combination* of

1. shutting down the transcription and translation of new cyclin and
2. degrading the cyclin proteins that already exist.

Negative feedback loops are often built into signaling pathways, and CDKs are no different. The mechanism is illustrated in Figure 08-09. APC is phosphorylated by M-CDK so that it is active and controls the metaphase checkpoint. Unlike many of the regulatory proteins we've seen, APC is *not* a kinase or a phosphatase. APC is a protein that transfers a small protein tag called **ubiquitin** to other proteins. (Review Chapter 4 if needed.) APC tags the cyclin with ubiquitin, which results in it being sent to the proteasome for degradation. Once the cyclin is degraded, the CDK is no longer active, and all of the phosphorylation targets of the CDK can then be dephosphorylated and returned to their original state. The result is that the concentration of cyclin in the cell drops rapidly, and the CDK is also deactivated, and the cell can start the process of putting the cytosolic contents back together again in the new daughter cells, complete cytokinesis, and return to G_1. For its part, APC is deactivated in G_1. It is one of the phosphorylation targets of the activated G_1/S-cyclin-CDK complex.

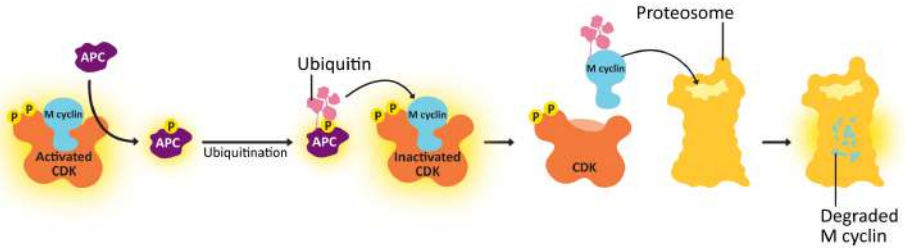


Figure 08-09: Ubiquitination and degradation of cyclin, which deactivates CDK. The anaphase-promoting complex (APC) is activated by M-CDK. Activated APC adds ubiquitin to target proteins including m-cyclin. Ubiquitination targets proteins for degradation in the proteasome. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Studying Cells: Experimental Techniques to Identify the Cell Cycle

Initially, we knew two things about the cell cycle: there was mitosis, and there was the time in between mitosis (i.e., **interphase**). We could observe this in the most rudimentary light microscopes roughly 200 years before we knew that DNA was the molecule that stored genetic information. As we now know, there are four major stages to the life cycle of a cell, and only one of them is mitosis. When researchers study the cell cycle, it is important that they can differentiate among the four stages. DNA content is a useful cue, since the amount of DNA changes in predictable ways throughout the cycle. In addition, sometimes researchers will need to work with a population of cells that are all in the same stage of development. In this section, we look at two ways to identify, and possibly synchronize, a population of cells based on cell cycle.

We will explore two different techniques and discuss each in turn:

1. Chemically synchronizing cells in a population: This technique is the starting point for many experiments to ensure all cells are starting at the same stage in the cell cycle.

2. **Fluorescence-activated cell sorting (FACS)**: This is a subtype of a more commonly known procedure called flow cytometry, where cells are monitored for certain properties and grouped based on these properties.

Chemically Synchronizing Cells in a Population

When studying cells as they progress through the cell cycle, it's common to want to take measurements based on a particular parameter. We can watch cells go through mitosis using a simple light microscope, but some of the other phases are more difficult to detect. As such, we may want to work with a population of cells that we know are all at the same phase. However, that generally does not happen naturally in a population of cells. Just like your classmates in elementary and high school hit their growth spurts at slightly different times even though you were all roughly the same age, cells will not naturally align their cell cycles.

In order to synchronize a population of cells, the scientist must apply some kind of block that stalls the cell cycle at a known step in the cell cycle. The two most common blocks are based on biological concepts we've seen before in this book.

- In Chapter 4, we discovered the existence of **temperature-sensitive mutations**, and how they could be used to study the essential proteins of the secretory pathway. Most cell cycle proteins are equally essential, as a cell that cannot progress through its cell cycle cannot divide. So temperature-sensitive mutants are useful in this context as well.
- In Chapter 6, we also saw that chemical inhibitors could be used to disrupt cytoskeletal function temporarily. This, too, is a strategy that can be applied to the cell cycle. There are a few known chemical inhibitors that block further progression in the cell cycle. DNA replication is often targeted by these compounds, and the cell stalls at the start of S phase as a result of the chemical inhibition.

When these blocks are applied, the cells will continue to progress through the cell cycle until they hit the blockage, and then they can go no farther. You can think of this like construction on a major bridge or roadway. The cars are able to move freely elsewhere as they work toward their destination, but once they arrive at the construction, they stop and stay there until the construction block is lifted so that the stopped traffic can again begin to move forward.

Fluorescence-Activated Cell Sorting (FACS)

This final technique is, in some ways, explained by its name. When we do **FACS**, we fluorescently label cells and then use that fluorescence to differentiate between cells that are in different states (Figure 08-10). The machine (called a **flow cytometer**) that measures the fluorescence is also able to separate a mixed population of cells based on this fluorescence (i.e., the cells are "sorted" into different groups based on the measured fluorescence).

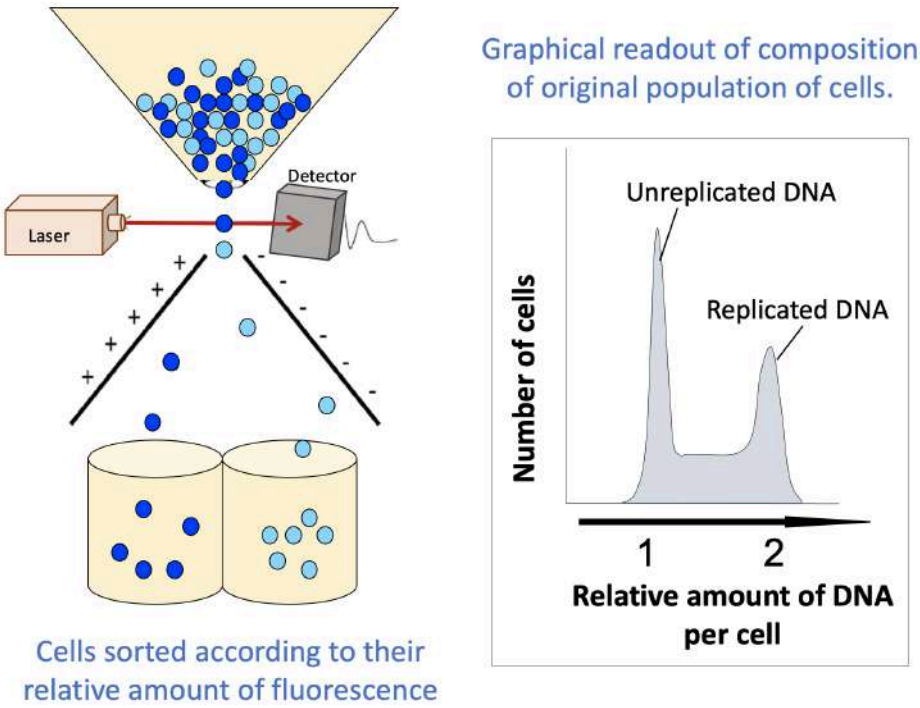


Figure 08-10: A schematic representation of fluorescence-activated cell sorting (FACS). The cells (blue circles) are first labeled with fluorescent dye that binds quantitatively to DNA. The cells are sent through the machine, and a laser detects the fluorescence and redirects the cells into different test tubes based on the fluorescence intensity. The machine also produces a readout of the cells that were measured and their fluorescence intensity. This image was created by Dr. Robin Young and is shared under a CC BY-SA 4.0 license.

When using FACS to learn about the stage of the cell cycle that each cell in a population is in, a live fluorescent DNA stain like DAPI is commonly used. Since the amount of DNA changes as the cell cycle progresses, this is a very simple way to separate the cells in a sample. The result of this technique is twofold:

1. You now have synchronized populations of cells without using chemical inhibitors or temperature-sensitive mutations. These can be used for further experiments.
    - Cells in G_1 will have one "full set" of their DNA, while cells in G_2 and M phase will have duplicated their DNA (i.e., two "full sets").
    - Cells in S phase will have more than one set but less than two sets, as replication has started in S phase but is not yet complete.
    - Additional visual separation may be required to differentiate between G_2 and M phase, but this can be done relatively easily, as a cell undergoing mitosis can be identified using a standard light microscope.
2. The flow cytometer also produces a graphical readout (Figure 08-10) that summarizes how many cells were found with each "amount" of fluorescent material (DNA in this case). These readouts can be used to learn important information about your population of cells.

Interpreting FACS Readouts

A simple FACS readout is shown in Figure 08-10. It is very much like a histogram, in which the y-

axis represents the number of cells counted at each point on the x-axis (often listed as a percentage of the total population), and the x-axis measures the amount of DNA. (Note: This is usually done *indirectly*, by measuring the fluorescence intensity.) The amount of fluorescence intensity correlates to the amount of DNA in a given cell.

Since the amount of DNA in a cell correlates with the phase of the cell cycle, this readout can help us determine how many of the cells in our sample are in each stage of the cell cycle at the moment at which they were measured. We can compare cells at different time points, which might tell us whether our cells are able to progress through the cell cycle or whether they have stalled.

The following are some things to remember about FACS and the readout in Figure 08-10:

- The readout indicates "replicated" and "unreplicated" DNA, not whether it's in G_1, S, G_2, or M phase. *This is an important limitation of this technique*. There are *two* phases in which the DNA has been replicated and as such are lumped together in the second peak (G_2 and M phase). FACS cannot differentiate between these, since the DNA content is the same for both. Depending on the question you're trying to answer, this might be key to correctly interpreting your results. To separate G_2 from M phase in your sample, you would need to do something additional, like look at the cells in a microscope, for example.
- The part *in between* the peaks is more important than it might seem. In this area, we have more than one "full set" of DNA but less than two. There's only one stage of the cell cycle where the amount of DNA goes from one set to two: S phase. Cells that are in the process of undergoing S phase will be found between the peaks that represent G_1 cells and G_2/M cells.

Like many of the techniques we've seen so far (FRAP, SDS-PAGE, etc.), FACS is a way to *quantify* what's going on in your samples. It's not actually the experiment itself. The experiment would be done before you put your samples through the FACS machine. If, for example, you wanted to explore the impact of a newly discovered mutation, you could compare the FACS readout from a wild-type (i.e., unmutated) and a mutated sample of cells and see how similar or different their FACS readouts are.

Changes in the FACS readout, even really strange and unexpected changes that don't make a lot of sense at first glance, can tell us very important things about the samples and treatments we choose to study. Like all of the experiments we've covered in this book, understanding what the technique can, and cannot, reveal is essential to being able to accurately interpret the data you receive from your experiment.

TOPIC 8.2: MITOSIS AND CELL DIVISION

Learning Goals

- Relate the progression of mitosis to the activation and deactivation of proteins by the mitotic CDK-cyclin complex.
- Explain the role of the cytoskeleton (and associated motor proteins) during mitosis and cytokinesis, including how dynamic instability contributes to the formation and function of the

mitotic spindle.

- Describe how signaling events are used to end mitosis, starting at anaphase and followed by the transition of the cell back into G_1.

The stages of mitosis are likely something you have been learning about since grade school, so it might seem odd at first that we are going to revisit them here. However, the reality of what actually has to happen, at the cellular level, in order for mitosis to occur is incredibly complex. As you may have guessed by now, nothing in the cell just happens by chance. It's all driven by proteins acting on other molecules in the cell. For a cell to go through mitosis, here's just a short list of cellular events that occur:

- The entire transcription and translation machinery needs to be shut down so that the DNA can be replicated, condensed, accurately divided, and then set up again in the new cell.
- The microtubule network in the cell needs to be dismantled and completely rearranged so that it can help with DNA separation.
- The actin network also needs to be rearranged to help with **cytokinesis** (division of the cytosol). This means that all of the regular work of the cytoskeleton during interphase (transport, organelle positioning, etc.) is also disrupted.
- Every organelle and structure in the cell must be duplicated and/or distributed into the two daughter cells so that both of the new cells have everything they need to survive. This often requires a complete dismantling of larger organelles (like the endomembrane system or the nucleus).
- A whole series of signaling events needs to be started up to initiate all of this change, and then, at the end, it all needs to be shut down again so that everything can transition back to the interphase and regular cellular function can be rebooted.

This incomplete list of events highlights the complexity of the process and also how vulnerable the cell is while this takes place. As long as the cell is focused on division, it cannot respond to changes in the environment or defend itself, metabolize new foods, or deal with damage. The cell needs to be efficient and get this work done quickly so that it reduces the time that it is vulnerable. The work also needs to be done very accurately, as any mistake in the process of mitosis is very likely going to kill one or both daughter cells.

Mitosis is a highly regulated dance that includes every part of the cell. This final topic of the book works to highlight some of the key elements of that dance and expand your understanding of the beautiful complexity involved in M phase.

The Hypercondensation of Chromosomes for Mitosis

Waaaay back in Chapter 3, we discussed the structure of the genome in detail, including how the DNA that forms each of the chromosomes is organized and compacted with the help of histone proteins. In this chapter, we will look at how the cell transforms the organization of the DNA in the interphase nucleus into the hypercompacted mitotic chromosomes that are required for cell division.

Just like we did in Chapter 3, we want to remind you that this is an area of active research that is extremely difficult to study, so we don't have all of the answers for how this process works. We will do our best to tell you what science currently believes to be true and also what we still have yet to discover.

There are two parts to chromosome management during mitosis: The first is the actual condensation of the chromosomes at the start of mitosis. The second is the maintenance of the structure of the mitotic chromosome throughout the stages of mitosis, including the separation of the sister chromatids at anaphase. Remember that chromosomes are, in essence, extremely long strands of DNA that could easily become a tangled mess in the dynamic environment of the cell during mitosis. Based on what we currently know, it appears to use at least some (but not all) of the same proteins that are used to maintain interphase packing of DNA. We'll focus primarily on the formation of the mitotic chromosomes and highlight what we know about the maintenance of these structures as we discuss the proteins involved.

Condensation of the chromosome for mitosis is thought to require only five additional proteins beyond the histones used to form the chromatin fiber. They are **condensin** I and II, **cohesins**, a **kinesin** known as Kif4A, and an enzyme called DNA topoisomerase II alpha. The roles of Kif4A and DNA topoisomerase II alpha are not entirely clear, even though it is clear that without them the mitotic chromosome is unable to form. From the histones, a key player in chromosome condensation for mitosis appears to be **histone H1**, also known as the linker histone. We'll start our discussion with the H1 histone, as it connects the most directly with what we learned about interphase chromatin in Chapter 3.

Phosphorylation of Histone H1 Promotes Higher-Order Chromatin Packing

While the structure of histone H1 is slightly different from the core histones, it still has some key features that are common to all histones. For example, it has two tail regions, which are sites where posttranslational chemical modifications (phosphorylation, acetylation, methylation) commonly occur. Histone H1 is a key phosphorylation target for M-CDKs. Phosphorylation is thought to have a few different effects on the structure of the chromatin, which helps it begin the process of disassembling the interphase organization so that the mitotic chromosome structure can take shape. Our discussion of H1 in Chapter 3 showed us that H1 is often used to help pack up the chromatin more tightly, which reduces access for gene expression, so the involvement of histone H1 in preparing the DNA for mitosis makes quite a bit of sense.

Condensin I and Condensin II Work Together to Pack the DNA into a Tight Column

The major work of packing the DNA up into the tight column that forms the mitotic chromosomes falls to two proteins that are part of the **condensin** family. Condensin I and II each have their own role to play in the process (Figure 08-11 and Video 08-03).

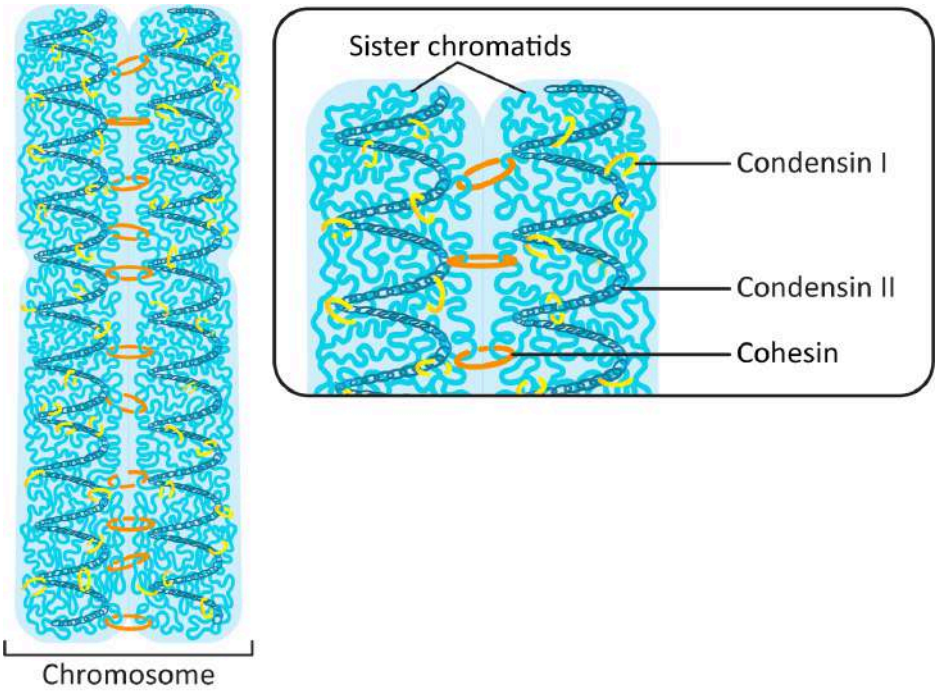


Figure 08-11: Condensin I (yellow) and II (dark blue) help compact the DNA (light-blue squiggles), while cohesin (orange) holds the sister chromatids together. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=158#oembed-3

Video 08-03: Condensins loop DNA into specific regions to develop a spiralized cylindrical DNA structure that is characteristic of the mitotic chromosome.

Cohesins Hold the Sister Chromatids Together

Interestingly, while **cohesins** have an important role in organizing interphase chromatin (they help form topologically associated domains, or TADs), they are *removed* from much of the DNA as the genome prepares for mitosis. Condensins pick up the slack and help loop up the DNA during the extreme condensation required for mitosis. However, cohesins still have an important function at this stage. As we saw earlier, cohesins are used to hold the sister chromatids together after replication, right until they are degraded during anaphase so that chromatid separation can take place (Figure 08-11).

The end result of all of this packing is an extremely condensed chromosome that is as much as 20,000 times more compacted than the original DNA strand. When observed in electron microscopy

(Figure 08-12), the looped fibers of DNA can be observed as well as the sister chromatids and the centromere. This level of packing is so extreme that no transcription is possible.

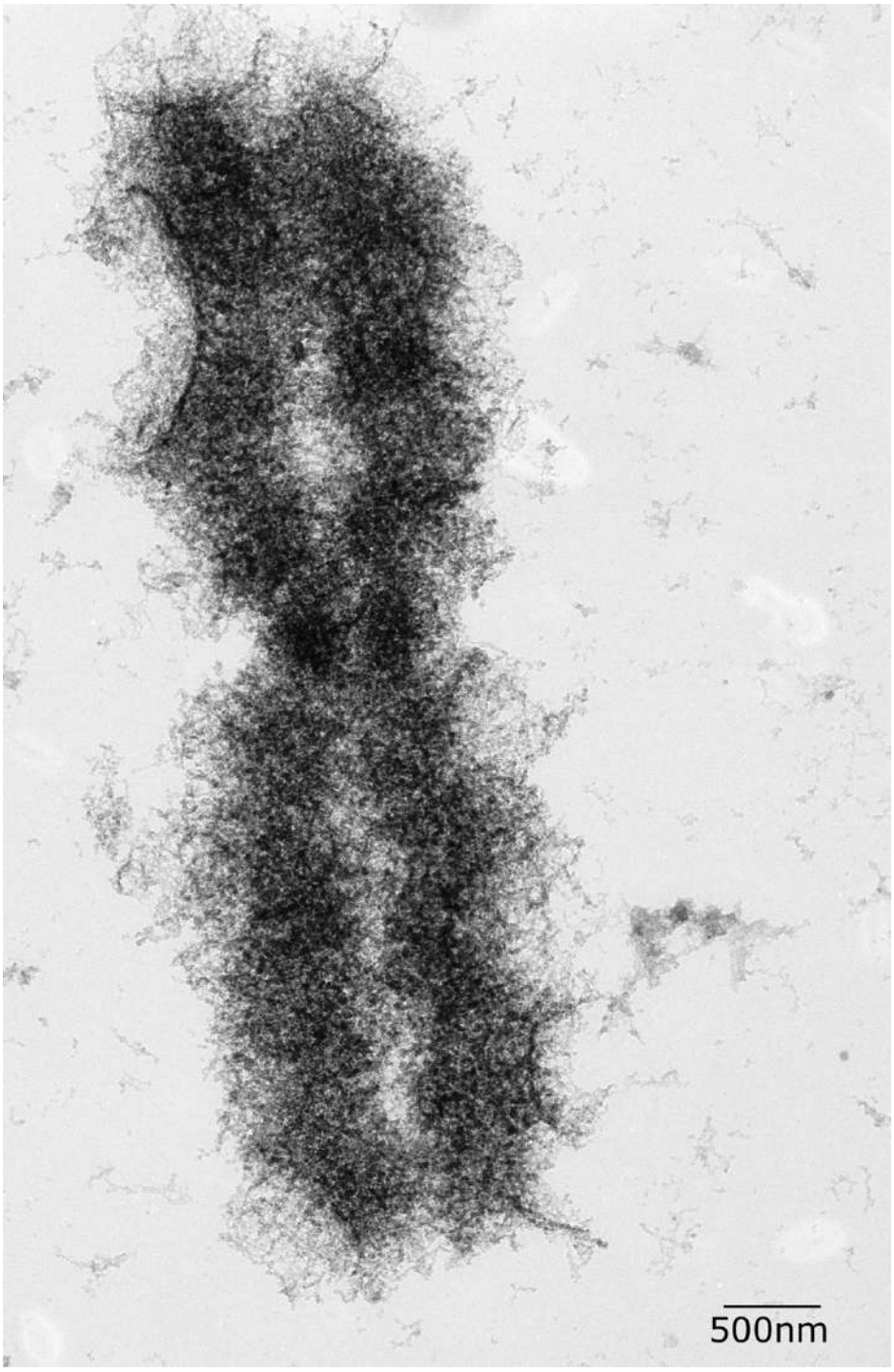


Figure 08-12: A mitotic chromosome observed in high-voltage TEM. The scale bar was added by Dr. Lauren Dalton. Original image from Hans Ris (2012), CIL:40682, https://doi.org/doi:10.7295/W9CIL40682. Chromosome extracted from Mus musculus fibroblast in the cell image library. This image is in the public domain.

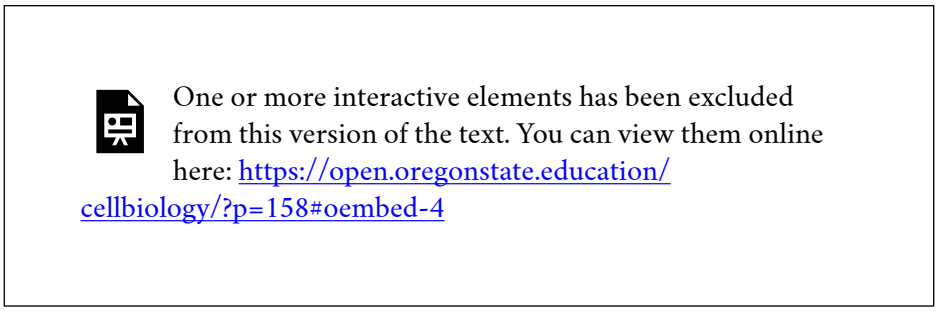
M Phase: Mitosis and Cytokinesis

M phase consists of two separate events:

- **Mitosis**: This is the process of division of the nucleus and its contents (i.e., the DNA!).
- **Cytokinesis**: This is where the rest of the cytoplasm gets divided into two new cells.

Both events are managed and controlled by a variety of proteins, which are activated and deactivated by the CDKs. It's also important to note that the cytoskeleton plays a key role in both events.

To start, we're going to take a step back and look at the process of mitosis through a microscope (Video 08-04).



Video 08-04: Live-action microscopy showing nuclear division of a cell. The DNA is shown in red, while the microtubules are shown in green.

It is likely that you have already learned the names of the stages of mitosis, as they are taught in elementary and/or high school. However, they are summarized in Figure 08-13.

We are going to look at each of these stages in turn. Our primary focus will be the rearrangements of the cytoskeleton that control the cellular events and the key signaling events that are responsible for triggering different parts of the process.

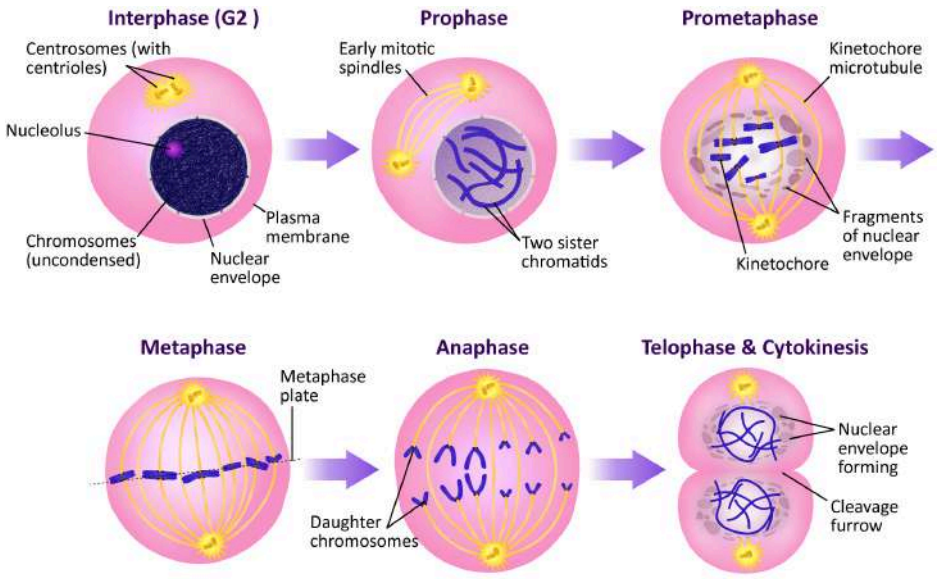


Figure 08-13: The stages of mitosis. "A diagram of mitosis stages" by Ali Zifan is shared under a CC BY-SA 4.0 license and has been rearranged to fit the page.

Mitosis Stage 1: Prophase (and Prometaphase)

Prophase is the first stage of mitosis. At this point, the cell has passed the G_2/M checkpoint, and its M-CDK is fully activated, so it begins to prepare itself in earnest for mitosis. Each of these events is the result of M-CDK phosphorylating one or many of the targets mentioned earlier in this chapter.

- Chromosomes condense as H1 and nucleolin are phosphorylated.
- The nuclear lamins depolymerize, resulting in the nuclear envelope breaking apart.
- The mitotic spindle will begin to form.

The first step in mitotic spindle formation is the production of two centralized microtubule organizing centers (MTOCs) in the cell, which will become the opposite poles of the spindle. The process of spindle formation actually begins well before the start of M phase. As early as the transition to G_2 phase, some of the components of the mammalian MTOC (also called the centrosome) will duplicate. At the onset of prophase, the centrosome splits in two, and the two MTOC will begin to migrate to either side of the nucleus. Interestingly, not very much is currently known about how plant cells transition from a decentralized gamma-tubulin distribution to the highly organized mitotic spindle.

There is more than one model for how the mitotic spindle is formed. Here we are going to focus on one called the "search and capture" model (Figure 08-14). In this model, the formation of the mitotic spindle relies very heavily on the dynamic instability of the microtubules. As microtubules grow outward from the spindle pole, they "search" the cytosol for things to grab onto (chromosomes, the plasma membrane, or microtubules from the other pole of the spindle). As you will remember from Chapter 6, microtubules are highly unstable unless capped or stabilized in some way. The formation of the mitotic spindle uses this characteristic to its advantage—a microtubule that doesn't "capture" anything in the cytosol will depolymerize. As such, microtubules will grow and shrink at random until they encounter the proper components to form the mitotic spindle.

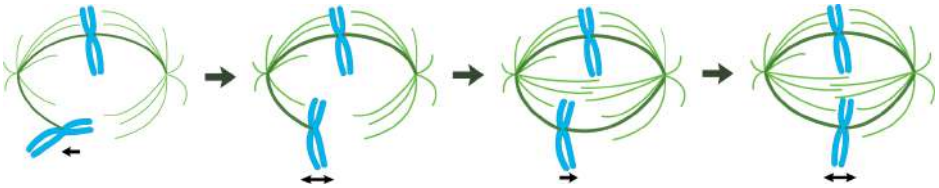


Figure 08-14: In the search and capture model, microtubules drive the formation of the mitotic spindle. All microtubules that compose the mitotic spindle originate with their minus end at the MTOC, which stabilizes that end. The other end is subject to dynamic instability and grows and shrinks until it becomes stabilized (see numbered text below). When a microtubule attaches to the kinetochore of the chromosome, this stabilizes the interaction. Spindles from both poles must attach to kinetochores on opposite sides of the chromosome. Heather Ng-Cornish adapted this image from "Spindle assembly models" by Lordjuppiter, shared under a CC BY-SA 3.0 license.

As the components of the spindle are stabilized and come together, three types of microtubule structures begin to form (Figure 08-15):

1. If microtubules from opposite poles happen to interact with each other, they will stabilize each other (with the help of motor proteins), thus forming a bridge between the two spindle poles. These are aptly named **interpolar microtubules**, as they span the two poles.
2. Conversely, if they happen to interact with a chromosome's kinetochore during their random growth and shrinkage, they will attach and become a **kinetochore microtubule**.
3. **Astral microtubules**, the third type of microtubule, are anchored at the plasma membrane (also with the help of motor proteins).

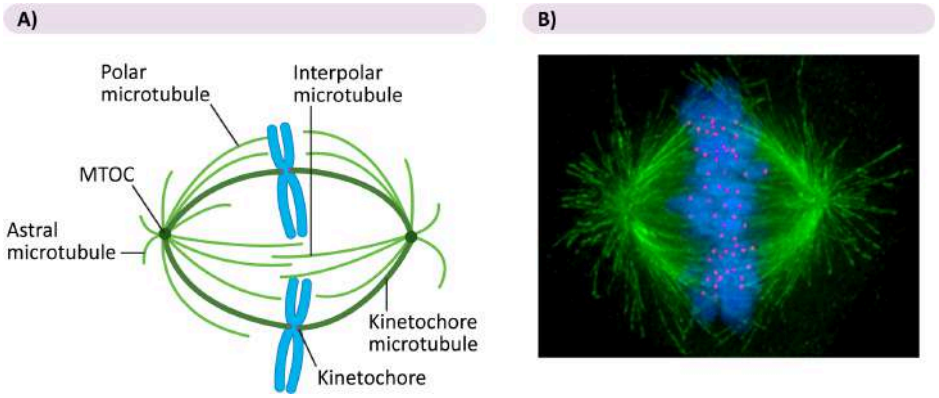


Figure 08-15: Labeled diagram (A) and microscopy (B) of mitotic spindle formation, including interpolar microtubules, astral microtubules attached to plasma membrane (not shown), and kinetochore microtubules attached to chromosomes. Heather Ng-Cornish adapted this image from (A) "Spindle Apparatus" by Lordjuppiter, shared under a CC BY-SA 3.0 license, and (B) "Kinetochore" by Afunguy, which is in the Public Domain.

Though there is much that is still unknown about the formation of the mitotic spindle, scientists do know that motor proteins are absolutely crucial to this process. Several motor proteins have been identified as having a role in the formation of the mitotic spindle, including the following:

- **Kinesins** are found at the kinetochore and help the chromosomes stay attached to the kinetochore microtubules as they disassemble in anaphase.
- **Dyneins** hold the astral microtubules and also help pull the spindle poles apart in anaphase

- Interpolar microtubules have a special double-headed kinesin, which walks along each of the interpolar microtubules in a set. They also help to push the spindle poles apart in anaphase B.

There is a distinction sometimes made between the beginning and end of prophase. We often say that once the nuclear envelope breaks down, we have entered **prometaphase**. The breakdown of the nuclear envelope is an important event, as the mitotic spindle cannot be completed until that happens and the condensed chromosomes are free in the cytosol. It is only then that the "search and capture" part of spindle formation can begin in earnest.

By the end of prometaphase, each of the chromosomes is held at the center of the mitotic spindle by microtubules attached to the **kinetochore**. The kinetochore is a specialized region of the chromosome associated with the centromere. As you may recall, the centromere is a DNA sequence that is the site of assembly of the kinetochore structure. The kinetochore is a plaque or button-shaped structure composed of proteins. There is one kinetochore associated with each of the chromatids, and the microtubules get inserted into them. They face in opposite directions on the two sides of the chromosome.

Mitosis Stage 2: Metaphase

After the nuclear envelope breaks down and the chromosomes are attached to the mitotic spindle by the microtubules, we transition to **metaphase**. Notably in this phase, the chromosomes align on the **metaphase plate**—an imaginary plane equidistant from the two spindle poles. To accomplish the chromosomal alignment, chromosomes are pulled simultaneously toward both spindle poles by the kinetochore microtubules. Even tension on the chromosome from *both* sides of the spindle is required for everything to stabilize and stay in place. The proper tension is only achieved when the chromosome is properly aligned at the metaphase plate.

It is known that even though *both* ends of the microtubules are being held in the mitotic spindle, the state of kinetochore microtubules is still very dynamic—tubulin units are continually being added and removed at *both* ends. This is intriguing and unusual, considering that each microtubule is held at the minus end by the MTOC (presumably with the help of **gamma-tubulin**), and the plus end is held by the kinetochore. While it is not entirely clear how the minus end can both be held by the MTOC and allow for dynamic instability, the story of how this works at the kinetochore is taking shape. Video 08-05 summarizes how we think it works. Something worth noting is the release of the metaphase checkpoint signal, which is the signal to activate the anaphase-promoting complex (APC) and initiate anaphase. This event leads us to the next stage of mitosis.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://open.oregonstate.education/cellbiology/?p=158#oembed-5

Video 08-05: Molecular representation of the metaphase checkpoint. This video emphasizes the dynamic nature of this checkpoint.

Mitosis Stage 3: Anaphase

Sister chromatids do not remain associated with each other by accident—they are held together by **cohesins**. At anaphase, the cohesins are dissolved, which allows the sister chromatids to separate. The abrupt destruction of the cohesin linkage between sister chromatids is brought about by activation of APC by M-CDK. Remember that APC's function is to tag proteins with ubiquitin, marking them for degradation via the proteasome (Figure 08-16). APC first acts on a regulatory protein called *securin*, which then releases an enzyme called *separase*. Separase cleaves the cohesins, allowing the sister chromatids to separate.

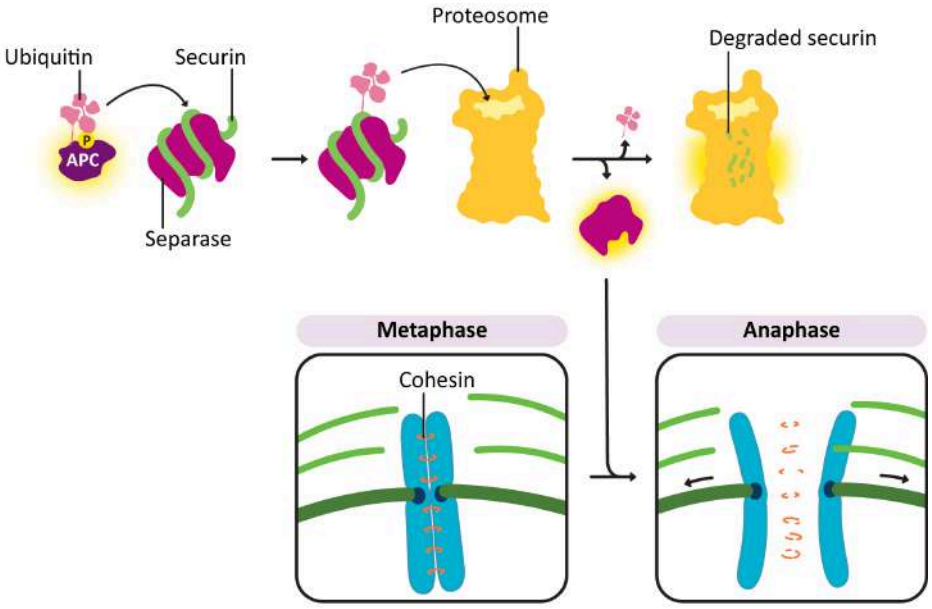


Figure 08-16: The anaphase-promoting complex (APC) adds ubiquitin to a protein called securin. This marks the protein for degradation. Once securin is removed, separase is active and cleaves cohesins that hold the sister chromatids together. This results in the separation of the sister chromatids during anaphase. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

The movement of kinetochores toward the spindle poles in anaphase takes place through the action of motor proteins. The motor proteins that helped with spindle assembly now work to separate the sister chromatids from each other. In addition, the kinetochore microtubules are depolymerizing rapidly from their plus end. These two forces work together to separate the sister chromatids and move them to the opposite poles.

Anaphase movement of chromosomes is separable into two phases with different mechanisms: **anaphase A** and **anaphase B** (Figure 08-17):

- **Anaphase A**. During the initial part of anaphase, kinetochore microtubules shorten and move the chromosomes toward the opposite poles. This requires a combination of depolymerization of kinetochore microtubules and the action of motor proteins at the kinetochore that help keep the chromosome connected to the shrinking microtubule.
- **Anaphase B**. During anaphase B, the length of the kinetochore microtubules (MTs) remains more or less constant. Now it is the other MTs of the mitotic spindle that begin to do the work. Again, there is much that is still unknown about this process, but three separate forces

have been shown to play a role in anaphase B:

- Tubulin subunits are added to the plus ends of the interpolar microtubules, making them longer.
- Motor proteins move the overlapping interphase microtubules from the two poles apart to make the spindle longer. This requires the help of the kinesins that are holding the paired interpolar microtubules together.
- Dyneins, attached to the astral microtubules and the cell cortex, walk toward the spindle poles on either side. This results in a shortening of the distance between the poles of the spindle and the plasma membrane. It also helps pull the spindle poles farther apart from each other.

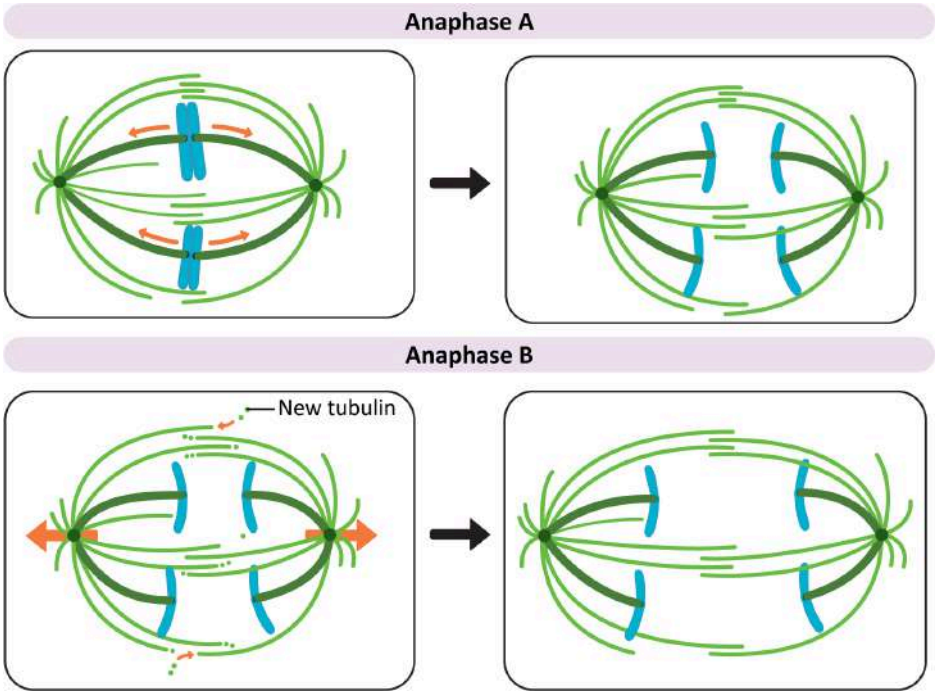


Figure 08-17: Anaphase A and B. In anaphase A, the chromosomes are pulled apart by the depolymerization of the kinetochore microtubules from the plus end. In anaphase B, the interpolar microtubules elongate. The motor proteins that anchor these microtubules together walk toward the plus end. This puts pressure on the spindle pole and pushes the poles farther apart. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

Mitosis Stage 4: Telophase

By **telophase**, the division of the nuclear contents is now complete; thus the cell needs to shut down mitosis and prepare for entry into G_1. Several processes occur:

- The nuclear lamina must reassemble so that the nuclear envelope can be rebuilt around the chromosome.
- The mitotic spindle disintegrates, and the cytoskeleton reassembles into its interphase conformation.
- The chromosomes must decondense and reform their interphase arrangement, which

includes the reassembly of the nucleolus.

If you examine the above list, you should notice that this is the reverse of what happened in preparation for mitosis and prophase. Since preparation for mitosis requires the activation of M-CDK, it makes sense that all of these changes to prepare for G_1 are a result of the deactivation of the same M-CDK. APC shuts down M-CDK by tagging M cyclin with ubiquitin. This results in the destruction of M cyclin by the proteasome (Figure 08-09). Once the cyclin is gone, the CDK can no longer function and is turned off. Because of this, the balance between active kinases and phosphatases is drastically shifted in favor of the phosphatases, which remove the phosphates that were previously added to M-CDK and its target molecules. This allows the proteins to return to their pre–M phase state, which allows the cell cycle to reenter G_1.

As a direct result of the M-CDK shutdown, the **nuclear lamina** gets dephosphorylated and begins to reform around the chromosomes (Figure 08-18). The chromosomes retain their connection to the nuclear lamina throughout the cell cycle. As such, the chromosomes themselves provide a template for nuclear envelope reformation, ensuring that the nucleus reforms without excluding any of the chromosomes. APC continues to be active throughout the end of M phase and into G_1, when it gets deactivated by the G_1/S-cyclin-CDK complex.

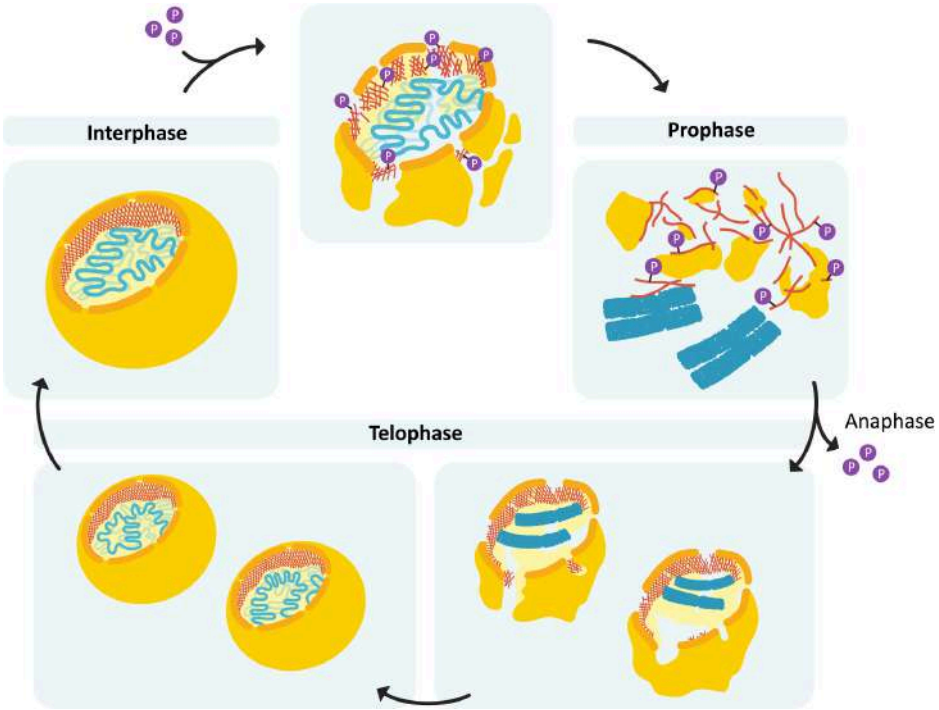


Figure 08-18: The cycle of breakdown and reformation of the nuclear envelope during mitosis. In prophase, the lamins (red) are phosphorylated. This results in depolymerization. In Anaphase, the lamins are dephosphorylated. This allows them to reform the meshwork (red) just below the surface of the membrane. This image was created by Heather Ng-Cornish and is shared under a CC BY-SA 4.0 license.

M Phase Final Stage: Cytokinesis

Once the division of the nuclear contents is complete, the rest of the cytoplasm also needs to be divided between daughter cells. This occurs during **cytokinesis**, which is also the final stage of M

phase (Figure 08-19). This process happens quite differently in plants, animals, and yeast and other fungi. However, in all cases, actin is the cytoskeletal fiber that plays the most important role.

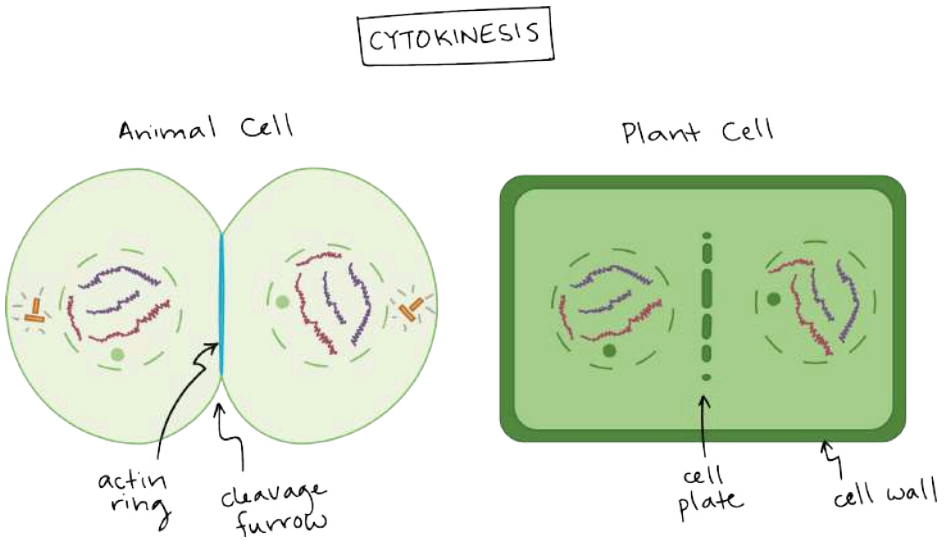


Figure 18-19: Cytokinesis in animal cells versus plant cells. Animals use an actin contractile ring to separate the cytoplasm. In contrast, plants primarily build a new cell wall to divide the cytoplasm of the two cells. "Cytokinesis" by Khan Academy is shared under a CC BY-NC-SA 4.0 license.

Cytokinesis in Mammalian Cells

In mammalian cells, actin arranges itself in an antiparallel array right underneath the plasma membrane. Together with myosin motors, they help form a **contractile ring** that pulls the plasma membrane inward and eventually splits the cell in two. Like many other parts of M phase, there is still much that is uncertain about the process of cytokinesis. However, the contractile ring has been found to always form around the midzone of the mitotic spindle, where the original metaphase plate was found. Thus, it is believed that the mitotic spindle itself plays a role in determining where the contractile ring will form.

As the myosin walks along the actin filaments of the ring, the ring contracts around what's left of the mitotic spindle (namely, the interpolar MTs). The organelles have been either broken down into vesicles (ER and Golgi) or duplicated (mitochondria) to ensure that there will be more or less equal quantities of them in each of the daughter cells. Thus, dividing the cytoplasm also divides the organelles. At the same time, vesicles fuse rapidly with the plasma membrane, which helps ensure that there is enough plasma membrane to fit around each daughter cell.

Cytokinesis in Plant Cells

In plants, the cell not only must divide its cytoplasm, but a new cell wall must be produced between the two daughter cells. As a result, the Golgi apparatus, a key synthesizer of cell wall components, remains functional throughout. The many small mobile plant Golgi do *not* vesiculate (i.e., break into small pieces) like they do in animal cells but instead divide by fission prior to mitosis. Electron microscopy shows that they congregate on either side of the site of the new cell wall. The interpolar microtubules also help guide this process through the formation of a structure known as the **phragmoplast** in the center of the cell. The Golgi secrete cell wall compounds very rapidly to the center of the phragmoplast to build a **cell plate**. Actin has an important role to play in forming the

phragmoplast and directing the vesicles to the cell plate, though the exact mechanism is still unclear. The cell plate grows from the center of the cell toward the outer walls and then fuses with the side walls of the dividing cell. At that point, the new daughter cells have been fully formed with their full complement of organelles and DNA.

Mitosis and Cytokinesis in Other Eukaryotes

The information in this chapter covers the most commonly discussed forms of cytokinesis—namely, plants and animals. It's important to remember that much like every other process in the cell, Eukaryotes are a hugely diverse set of organisms, which means that mitosis and cytokinesis can vary quite a bit from what is described in this chapter. In addition, specific tissues within organisms will vary. While we don't have time to go into all of the different ways that mitosis happens, we did want to end this chapter with a few interesting examples of how different it can be:

- Plant and animal cells undergo what's called *open mitosis*, which means that the nuclear envelope breaks down completely in order to form the mitotic spindle. There are also two other forms:
    - *Partially closed mitosis* occurs when the nuclear envelope remains intact and the spindle is formed both inside and outside of the nuclear envelope. Budding yeast (*Saccharomyces cerevisiae*) uses a partially closed format, which is especially interesting considering how often they are used as a model organism for all Eukaryotes.
    - A third form is *closed mitosis*, which happens when the nuclear envelope not only remains intact, but the nuclear pores are plugged to close the nucleus off completely from the rest of the cell. This is more common in protists, fungi, and other single-celled organisms.
- Many cell types don't divide evenly at mitosis. Lots of plant and animal cells undergo **asymmetric division**, which results in one daughter cell that is larger than the other. Additionally, lots of cells "grow" their daughter cell off of one side and move the new daughter nuclei into the bud once the bud is large enough. *Saccharomyces cerevisiae* is called "budding yeast" for this very reason. There are many additional cells that do this, especially single-celled algae, fungi, and other protists.

CHAPTER SUMMARY

As you can tell now that you are at the end of this chapter and the end of the book, the cell cycle is truly a great process to integrate all the knowledge that you have gained thus far. In addition, we hope that you appreciate the additional layers of complexity required to regulate one of the most important cellular processes.

The cell cycle is an intricate dance that requires the cooperation of most of the cellular systems. We've emphasized a few in this chapter. But we challenge you to think of even more connections to the other units in the book.

- Signaling: The cell cycle requires integrated information from the extracellular and intracellular environments through signaling. The balance of activating forces versus inactivating forces helps ensure that cells divide only when the time is right.

- Cytoskeleton: The cytoskeleton is highly involved, especially in M phase. It's clear that dynamic instability, a hallmark of the cytoskeletal system, is vital for the formation of the mitotic spindle but also the correct alignment of the chromosomes at the metaphase plate.
- DNA packing: The molecular packers and organization help unpack all the DNA during synthesis and do the opposite in M phase, completely packaging the DNA.
- Other organelles: Of course, these organelles need to be properly divided between daughter cells, but that does not end their contribution to this process. Prior to shutdown for mitosis, they need to make sure that everything the cell will need is available. This will include a great deal of energy, in the form of ATP and GTP, membrane, vesicles, and even polysaccharides if a wall needs to be made between the two daughter cells.

With that, we've come to the end of the book. We congratulate you for getting to the end! No matter your starting place, we hope that you have at least a little more appreciation for the complex beauty of the cell.

Review Questions

Note on usage of these questions: Some of these questions are designed to help you tease out important information within the text. Others are there to help you go beyond the text and begin to practice important skills that are required to be a successful cell biologist. We recommend using them as part of your study routine. We have found them to be especially useful as talking points to work through in group study sessions.

Topic 8.1: Regulating the Cell Cycle: Checkpoint Control

1. Make a list of the four phases of the cell cycle, and describe in detail what has to happen in each phase before it is complete. Be as specific as you can.
2. For each of the stages, make a list of experimental techniques that could be used to identify cells that are in each of the four stages. Explain the rationale for your choice in each case.
3. Explain the concept of checkpoint control of the cell cycle and give examples of what types of conditions should be met before S and M phases proceed. Explain why those conditions need to be met before the cell moves to the next phase.
4. Describe the sequence of events leading to CDK activation and deactivation.
5. Define negative and positive feedback. How are they thought to play a role in cell cycle control?
6. Discuss the role of CDK-cyclin complexes in regulation of the cell cycle. How do they contribute to the cell "knowing" the conditions are right for the next stage of the cell cycle?
7. Compare and contrast the roles of CDKs that control entry into S phase versus entry into M phase. Make a list of the different examples of proteins that they are thought to interact with and their effects.
8. Explain the role of phosphorylation in regulation of the cell cycle. Explain how phosphorylation can affect protein folding and how that might result in changes in protein

activity.

Topic 8.2: Mitosis and Cell Division

1. Describe the role of dynamic instability in the formation and maintenance of the mitotic spindle.
2. Illustrate with labeled diagrams the relationship between chromatids and chromosomes at each of the stages of the cell cycle.
3. Explain how each stage of mitosis is ultimately controlled by the activation and deactivation of CDKs.
4. Compare and contrast the following stages:
    a. prophase and prometaphase
    b. prophase and telophase
    c. anaphase A and B
5. What is cytokinesis, and how does it differ in plants and animals?

GLOSSARY

0-9

2D gel electrophoresis (also 2-DE)
"A form of gel electrophoresis commonly used to analyze proteins. Mixtures of proteins are separated by two properties in two dimensions on 2D gels." *Source:* Two-dimensional gel electrophoresis. (2023, March 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Two-dimensional_gel_electrophoresis

3′ untranslated region (3′ UTR)
The region of DNA that is downstream (3′) of the transcription termination site.

5′ untranslated region (5′ UTR)
The region of DNA that is upstream (5′) of the transcription start site.

A

A/B compartments
Distinct chromosomal regions. The A compartment tends to be located closer to the center of the nucleus and contains more genes. The B compartment contains relatively more constitutive heterochromatin and is more likely to be at the nuclear periphery.

acetylation
A chemical modification that can be added to a variety of biomolecules to change their function. In this process, an acetyl group is added to the molecule at a specific location.

actin filaments
Polymers of actin subunits that are part of the cell's cytoskeletal system.

actin-binding proteins (ABP)
Proteins that bind to actin (monomers or filaments) in vivo to modulate their function.

activator proteins
A DNA binding protein that has positive control over gene expression, often causing an increase of transcription of a particular gene.

adaptor protein
Proteins that play an important role in vesicle formation, where they help to gather scaffold proteins that are needed to form the clathrin coat of the vesicle.

adenylyl cyclase
An enzyme responsible for making cAMP (cyclic AMP) from ATP. Also called adenylate cyclase.

affinity
In the context of proteins, refers to the favorability of a binding interaction. For instance, the protein when phosphorylated has greater affinity for its substrate. This means that the protein is likely to bind more tightly when the protein is phosphorylated.

agarose gel electrophoresis
A method of gel electrophoresis used in biochemistry, molecular biology, genetics, and clinical chemistry to separate a mixed population of macromolecules such as DNA or proteins in a matrix of agarose, one of the two main components of agar. *Source:* Gel electrophoresis. (2023, July 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Gel_electrophoresis

aggregation
When referring to proteins, it means that they cluster together into clumps. Oftentimes these clumps are no longer functional.

allele
"A variation of the same sequence of nucleotides at the same place on a long DNA molecule." *Source:* Allele. (2023, April 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Allele

alpha helix (α-helix)
"A common motif in the secondary structure of proteins and is a right hand helix conformation in which every backbone N–H group hydrogen bonds to the backbone C=O group of the amino acid located four residues earlier along the protein sequence." *Source:* Alpha helix. (2023, July 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Alpha_helix

alternative splicing
"An alternative splicing process during gene expression that allows a single gene to code for multiple proteins. In this process, particular exons of a gene may be included within or excluded from the final, processed messenger RNA (mRNA) produced from that gene. This means the exons are joined in different combinations, leading to different (alternative) mRNA strands." *Source:* Alternative splicing. (2023, June 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Alternative_splicing

amino acyl tRNA synthetases
Responsible for putting the amino acid on the tRNA.

amphipathic
A chemical compound or molecule that possesses both a hydrophilic region and a hydrophobic region. Phospholipids that make up the majority of the structure of phospholipid bilayers are an example of an amphipathic molecule.

anaphase
The stage of mitosis when the chromosomes are separated.

anaphase A
In this part of Anaphase, the kinetochore microtubules shorten to facilitate chromosomes moving closer to their respective spindle poles.

anaphase B
In this part of Anaphase, the interpolar microtubules lengthen. This pushes the spindle poles further apart.

anaphase-promoting complex (APC)
Causes the degradation of the cohesion proteins that bind sister chromatids together, which releases the daughter chromosomes so that they can be pulled toward their respective spindle pole.

anterograde traffic
Involves movement of coated vesicles in the "forward" direction. Depending on the trafficking pathway, this can be different destinations.

antiparallel

Refers to two molecules that bind in a head-to-tail orientation.

apoptosis

A regulated process of cell death.

Arp2/3

An actin-binding protein that binds to actin filaments and promotes a new branched actin filament.

arrays (cytoskeletal)

Networks of either actin or microtubules.

astral microtubules

Microtubules that start at the spindle pole but extend toward the plasma membrane to anchor the spindle pole.

asymmetric division

When cell division does not equally split a cell in half. This occurs in a variety of cell types.

ATP synthase

A large protein complex that uses the flow of protons to synthesize ATP from ADP and inorganic phosphate. It dimerizes and is found in the cristae membrane of mitochondria.

autocrine

In this type of signaling, the signal is released and received by the same cell.

autotroph

An organism that creates its own complex biomolecules (carbohydrates, proteins, lipids) from simple substrates (carbon dioxide and water).

B

beta barrel

"A beta sheet composed of tandem repeats that twists and coils to form a closed toroidal structure in which the first strand is bonded to the last strand (hydrogen bond). Beta-strands in many beta barrels are arranged in an antiparallel fashion. Beta barrel structures are named for resemblance to the barrels used to contain liquids." *Source:* Beta barrel. (2022, May 5). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Beta_barrel

beta sheet

"A common motif of the regular protein secondary structure. Beta sheets consist of beta strands (β-strands) connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet. A β-strand is a stretch of polypeptide chain typically 3 to 10 amino acids long with backbone in an extended conformation." *Source:* Beta sheet. (2022, September 9). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Beta_sheet

bilayer

Generally refers to a lipid bilayer that has two lipids stacked tail to tail with the polar head groups facing the aqueous environment.

binary fission

A type of asexual reproduction where an organism divides into two of the same.

bioinformatics

"An interdisciplinary field of science that develops methods and software tools for understanding biological data, especially when the data sets are large and complex." *Source:* Bioinformatics. (2023, July 7). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Bioinformatics

biological membrane

A selectively permeable barrier that is designed to separate the cell from the external environment, allows communication of activities between cells, and functions to form intracellular compartments. It is composed of lipids, proteins, and carbohydrates.

brightfield light microscopy

"The simplest of all the optical microscopy illumination techniques. Sample illumination is transmitted (i.e., illuminated from below and observed from above) white light, and contrast in the sample is caused by attenuation of the transmitted light in dense areas of the sample." *Source:* Bright-field microscopy. (2023, May 8). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Bright-field_microscopy

C

carbon-fixation reactions

The biochemical reactions that occur in the chloroplast stroma, where carbon dioxide is used to make three carbon sugar molecules.

cargo protein

Proteins that are carried within vesicles to a particular location. These can be proteins that are secreted but also other resident proteins that need to travel to a particular destination in the endomembrane system.

cargo receptor (protein)

Responsible for binding to specific forms of cargo like proteins. Cargo receptors help the cargo get loaded into the correct vesicle for transport.

cell cortex (also known as the actin cortex, cortical cytoskeleton, or actomyosin cortex)

"A specialized layer of cytoplasmic proteins on the inner face of the cell membrane." *Source:* Cell cortex. (2023, February 12). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cell_cortex

cell cycle

"The series of events that take place in a cell that causes it to divide into two daughter cells." *Source:* Cell cycle. (2023, June 7). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cell_cycle

cell fractionation

"The process used to separate cellular components while preserving individual functions of each component." *Source:* Cell fractionation. (2023, February 22). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cell_fractionation

cell plate

The site at which a new cell wall forms when a plant cell undergoes cytokinesis.

cell wall

"A structural layer surrounding some types of cells, just outside the cell membrane. It can be tough, flexible, and sometimes rigid. It provides the cell with both structural support and protection, and also acts as a filtering mechanism." *Source:* Cell wall. (2023, March 30). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cell_wall

centriole

A cylindrical organelle at the center of a centrosome in animal cells.

centromere

A structural region of a chromosome where the two sister chromatids are connected. Additionally, the kinetochore and mitotic spindle fibers attach to the sister chromatids during mitosis and meiosis.

centrosome

A microtubule organizing center (MTOC) in animal cells. It contains a pair of perpendicular centrioles.

chaperone protein

"Proteins that assist the conformational folding or unfolding of large proteins or macromolecular protein complexes." *Source:* Chaperone (protein). (2023, April 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Chaperone_(protein)

checkpoints (cell cycle)
Molecular mechanisms that ensure that proper cellular conditions are met prior to allowing a cell to enter the next phase of the cycle.

chemiosmotic coupling
Refers to the cellular mechanism that harnesses the potential energy stored in a chemical gradient to power a biochemical reaction.

ChIP-seq
"A method used to analyze protein interactions with DNA. ChIP-seq combines chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing to identify the binding sites of DNA-associated proteins. It can be used to map global binding sites precisely for any protein of interest." *Source:* ChIP sequencing. (2023, May 9). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/ChIP_sequencing

cholesterol
A molecule in the sterol family that plays a role in the fluidity of membranes.

chromatid
Represents one-half of a replicated chromosome. Two sister chromatids together form one chromosome.

chromatin
In the simplest terms, the combination of DNA and protein. When DNA is packaged using proteins, this combination structure is referred to as chromatin.

chromatin fiber (30 nm fiber)
A structure of chromatin that contains the nucleosome core as well as histone H1. This is considered the most abundant form of chromatin in the interphase nucleus.

chromatin immunoprecipitation (ChIP)
"A type of immunoprecipitation experimental technique used to investigate the interaction between proteins and DNA in the cell. It aims to determine whether specific proteins are associated with specific genomic regions, such as transcription factors on promoters or other DNA binding sites." *Source:* Chromatin Immunoprecipitation. (n.d.). In *Wikipedia, The Free Encyclopedia.* https://en.wikipedia.org/wiki/Chromatin_immunoprecipitation

chromatin remodeling
"The dynamic modification of chromatin architecture to allow access of condensed genomic DNA to the regulatory transcription machinery proteins, and thereby control gene expression." *Source:* Chromatin remodeling. (2023, May 24). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Chromatin_remodeling

chromatin remodeling complexes
Protein complexes responsible for making changes to chromatin structure. This can include nucleosome sliding, nucleosome eviction, and nucleosome swapping.

chromosome
Long strands of DNA that contain the genetic material of the given organism. Chromosomes usually contain histone and nonhistone proteins that add to the packaging of the DNA strand.

chromosome territory
A region in the nucleus preferentially occupied by a chromosome.

cilium
A specialized structure in Eukaryotic cells containing microtubules and a specialized MTOC called a basal body.

cisterna
"A flattened membrane vesicle found in the endoplasmic reticulum and Golgi apparatus. Cisternae are an integral part of the packaging and modification processes of proteins occurring in the Golgi." *Source:* Cisterna. (2023, January 3). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cisterna

cisternal maturation model
A model to describe the movement of proteins through the Golgi apparatus. In this model, each cisterna matures into the next. This means that the cis cisternae become the medial, then the trans cisternae, which eventually turn into vesicles in the *trans* Golgi network (TGN) for further trafficking.

citric acid cycle
A series of biochemical reactions that produces ATP and the electron donors NADH and FADH2 from pyruvate. Also known as the tricarboxylic acid (TCA) cycle.

clathrin (coated vesicle)
Transport vesicles that contain a coat with clathrin triskelions. These coated vesicles participate in a variety of pathways, notably endocytosis and the lysosomal targeting pathway.

coat protein
Help support the formation of a budding transport vesicle. Coat proteins bind to adaptor proteins and cargo receptors to ensure that the proper cargo is recruited into these structures.

coding strand
"The DNA strand whose base sequence is identical to the base sequence of the RNA transcript produced (although with thymine replaced by uracil)." *Source:* Coding strand. (2021, December 8). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Coding_strand

cofactor
Refers to a protein to aid in regulation of gene expression.

cohesin
A protein complex that helps keep sister chromatids attached after DNA replication and also helps organize the interphase genome.

cohesin attachment regions (or CARs)
Specific sites on DNA where cohesins bind to attach sister chromatids together.

coiled coil
A form of protein structure that is produced when two alpha helices wind around each other to produce a super coiled structure.

collagen
"The main structural protein in the extracellular matrix found in the body's various connective tissues." *Source:* Collagen. (2023, July 3). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Collagen

condensin
A family of proteins with a role in DNA condensation needed for mitosis and meiosis.

confocal laser scanning microscopy
Uses a laser to excite fluorescent molecules inside cells, similar to epifluorescence. However, the light shines through a pinhole to block out-of-focus light. This allows for "optical" sectioning (dividing the cells) along an axis and can result in a sharper image.

consensus sequence
"The calculated sequence of most frequent residues, either nucleotide or amino acid, found at each position in a sequence alignment. It represents the results of multiple sequence

alignments in which related sequences are compared to each other and similar sequence motifs are calculated." *Source:* Consensus sequence. (2023, May 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Consensus_sequence

constitutive heterochromatin
Regions of chromatin that are packaged and inaccessible for transcription. These cannot be unpackaged. Most often these regions are important for chromosome structure and do not contain genes.

constitutive secretion
The process of continuous secretion of molecules from the cell or organelle that is not dependent on external factors. This is also considered to be the default secretory pathway.

contractile ring
A special formation of actin filaments and myosin motors that will continually rachet smaller. This structure is needed for cytokinesis in animal cells.

COPI (vesicle coat)
Transport vesicles that contain a coat with COPI units. These coated vesicles participate in a variety of pathways, notably Golgi to ER (retrograde pathway).

COPII (vesicle coat)
Transport vesicles that contain a coat with COPII units. These coated vesicles participate in a variety of pathways, notably ER to Golgi (anterograde pathway).

core histones
H2A, H2B, H3, and H4 are the histone proteins that make up the core of the nucleosome. Two copies of each are needed to create the octamer.

crista
"A fold in the inner membrane of a mitochondrion. The name is from the Latin for crest or plume, and it gives the inner membrane its characteristic wrinkled shape, providing a large amount of surface area for chemical reactions to occur on." *Source:* Crista. (2022, November 5). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Crista

critical concentration (Cc)
Can be thought of in three ways: the monomer concentration needed to initiate polymer formation, the monomer concentration at reaction equilibrium, and the concentration at which a polymer goes from shrinking to growth.

cryofixation
"A technique for fixation or stabilization of biological materials as the first step in specimen preparation for electron microscopy and cryo-electron microscopy." *Source:* Cryofixation. (2023, May 27). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cryofixation

cyclic AMP (cAMP or 3′,5′-cyclic adenosine monophosphate)
"A second messenger, or cellular signal occurring within cells, that is important in many biological processes. cAMP is a derivative of adenosine triphosphate (ATP) and used for intracellular signal transduction in many different organisms, conveying the cAMP-dependent pathway." *Source:* Cyclic adenosine monophosphate. (2023, April 30). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cyclic_adenosine_monophosphate

cyclin
"A family of proteins that controls the progression of a cell through the cell cycle by activating cyclin-dependent kinase (CDK) enzymes or group of enzymes required for synthesis of cell cycle." *Source:* Cyclin. (2023, January 29). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cyclin

cyclin-CDK complex
A complex of a cyclin protein with a cyclin-dependent kinase (CDK). When this complex is activated, it promotes progression to the next stage of the cell cycle.

cyclin-dependent kinase (CDK)
A protein kinase that, when activated, phosphorylates target proteins needed for progression to the next stage of the cell cycle. Activation requires binding with a cyclin protein as well as phosphorylation of particular locations within the protein.

cytokinesis
The process of cytoplasmic division.

cytoplasmic streaming
"The flow of the cytoplasm inside the cell, driven by forces from the cytoskeleton." *Source:* Cytoplasmic streaming. (2023, July 7). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cytoplasmic_streaming

D

darkfield microscopy
"Describes microscopy methods, in both light and electron microscopy, which exclude the unscattered beam from the image. As a result, the field around the specimen (i.e., where there is no specimen to scatter the beam) is generally dark." *Source:* Dark-field microscopy. (2022, September 7). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Dark-field_microscopy

Delta G (ΔG)
The change in Gibb's free energy. When this term is negative, it suggests that a reaction will be spontaneous and occur without the input of energy from the environment.

Delta H (ΔH)
The change in enthalpy between two states. In the case of proteins, usually a bound and unbound state. Enthalpy encompasses the amount of bound energy in the system. When applied to proteins, this can relate to the amount of energy contained within all the bonds of the molecule. A positive ΔH indicates there is more bond energy in the second state compared to the first state, usually as a result of more bonds or stronger bonds.

Delta S (ΔS)
The change in entropy of the system. Entropy is the measure of motional freedom within the system. Different molecules will have more and less freedom. A higher ΔS indicates that the resulting state will have more motional freedom than the original in a chemical reaction.

dephosphorylation
The process of removing a phosphate chemical group from a protein.

desmosome
"A cell structure specialized for cell-to-cell adhesion." *Source:* Desmosome. (2023, May 27). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Desmosome

diacylglycerol (DAG)
"A glyceride consisting of two fatty acid chains covalently bonded to a glycerol molecule through ester linkages." *Source:* Diglyceride. (2022, November 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Diglyceride

dichroic mirror
A special filter that allows some wavelengths of light to pass through but reflects others. These are used in microscopy to shine certain wavelengths of light on the sample but not others.

dimer

"The word dimer has roots meaning 'two parts,' di- + -mer. A protein dimer is a type of protein quaternary structure." *Source:* Protein dimer. (2023, May 15). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Protein_dimer

dimerize
The process of two biomolecules binding.

diploid
Refers to organisms that contain two copies of each chromosome.

direct staining
When colored or fluorescent dyes are added to a biological sample to help differentiate cellular structures under the microscope.

disulfide bridge
The covalent linkage of thiol groups between two thiol residues. Commonly, these are found on the R groups of cysteine amino acids.

docking sites
Protein locations where other biomolecules are able to bind.

domain map
A schematic representation of the protein sequence with highlighted areas indicating protein domains (regions of the protein with a particular function).

downstream
Refers to the region of the DNA that is 3′ of the area referenced. For example, in a sentence, you could say, "The transcription stop site is located downstream of the transcription start site."

dynamic instability
"Refers to the coexistence of assembly and disassembly at the ends of a microtubule. The microtubule can dynamically switch between growing and shrinking phases in this region." *Source:* Microtubule. (2023, July 15). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Microtubule

dynamin
Part of a large family of enzymes known as GTPases that play a crucial role in endocytosis, cytokinesis, scission of newly formed vesicles, and division of organelles.

dynein
A molecular motor that uses microtubules. All known dynein motors move toward the minus end. Plants do not have dyneins and instead use minus-end directed kinesin to fulfill this role.

E

ectotherm
"An animal in which internal physiological sources of heat are of relatively small or of quite negligible importance in controlling body temperature." *Source:* Ectotherm. (2023, June 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Ectotherm

electrochemical gradient
"A gradient of electrochemical potential, usually for an ion that can move across a membrane. The gradient consists of two parts: 1) The chemical gradient, or difference in solute concentration across a membrane. 2) The electrical gradient, or difference in charge across a membrane." *Source:* Electrochemical gradient. (2023, July 18). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Electrochemical_gradient

electron carrier (donor)
"A chemical entity that donates electrons to another compound. It is a reducing agent that, by virtue of its donating electrons, is itself oxidized in the process." *Source:* Electron donor. (2022, August 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Electron_donor

electron tomography
The chemical gradient or difference in solute concentration across a membrane.

electron transport chain (ETC)
"A series of protein complexes and other molecules that transfer electrons from electron donors to electron acceptors via redox reactions (both reduction and oxidation occurring simultaneously) and couples this electron transfer with the transfer of protons (H+ ions) across a membrane." *Source:* Electron transport chain. (2023, May 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Electron_transport_chain

emitted light microscopy (a.k.a. fluorescence light microscopy)
A microscopy technique where samples are illuminated with light of a particular wavelength. The fluorescent molecules in the sample become excited and emit light of a longer wavelength. The microscope catches this light to display the location of the fluorescent molecules bound to a cellular structure of interest. There are two main types of fluorescent light microscopy: epifluorescence and confocal microscopy.

endocrine
A long-range signaling pathway where a cell secretes a signaling molecule that is picked up by target cells not in the local vicinity.

endocytic pathway
The pathway by which materials are brought into the cell. Once materials are in the cell, they transit to the endosome and then to the lysosome.

endocytosis
The process in which outside materials and cargo are brought into the cell after passing through the membrane and forming a vesicle once inside. Endocytosis includes pinocytosis, phagocytosis, and receptor-mediated endocytosis.

endoplasmic reticulum (ER)
"The transportation system of the eukaryotic cell… [that] has many other important functions such as protein folding. It is a type of organelle made up of two subunits—rough endoplasmic reticulum (RER), and smooth endoplasmic reticulum (SER). The endoplasmic reticulum is found in most eukaryotic cells and forms an interconnected network of flattened, membrane-enclosed sacs known as cisternae (in the RER), and tubular structures in the SER." *Source:* Endoplasmic reticulum. (2023, June 12). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Endoplasmic_reticulum

endosome
Intracellular sorting centers that are the first location molecules are delivered to once they are endocytosed. There are two classifications of endosomes, early and late, which are marked by different cellular markers. Some vesicles from the TGN can be routed to the endosome as well. Endosomes sort materials that need to go to the lysosome from those that need to go back to the Golgi or cell membrane.

endosymbiont theory
"The leading evolutionary theory of the origin of eukaryotic cells from prokaryotic organisms. The theory holds that mitochondria, plastids such as chloroplasts, and possibly other organelles of eukaryotic cells are descended from formerly free-living prokaryotes (more closely related to the Bacteria than to the Archaea) taken one inside the other in endosymbiosis." *Source:* Symbiogenesis. (2023, May 28).

In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Symbiogenesis

endotherm

"An organism that maintains its body at a metabolically favorable temperature, largely by the use of heat released by its internal bodily functions instead of relying almost purely on ambient heat." *Source:* Endotherm. (2023, April 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Endotherm

enhancer region

"A short (50–1500 bp) region of DNA that can be bound by proteins (activators) to increase the likelihood that transcription of a particular gene will occur." *Source:* Enhancer (genetics). (2023, July 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Enhancer_(genetics)

epifluorescence

The most general ubiquitous form of fluorescence light microscopy. Light at the excitation wavelength is applied to the sample. Then the detector will pick up the emission wavelength of light emitted from the sample.

epigenetics

"The study of stable changes in cell function (known as marks) that do not involve alterations in the DNA sequence." *Source:* Epigenetics. (2023, July 5). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Epigenetics

ER exit site

The site on the endoplasmic reticulum (ER) where vesicles accumulate and bud off bound for the Golgi.

ER insertion sequence

A sequence of hydrophobic amino acids (approx. 8–10) that serve as a binding site for SRP. This triggers its cotranslational insertion into the ER. There are several variations of this sequence depending on the location in the protein's sequence, including N-terminal START sequence, internal START sequence, and internal STOP sequence.

ER lumen

The interior fluid of the endoplasmic reticulum (ER).

ER resident protein

"Refers to proteins that remain in the endoplasmic reticulum, or ER, after folding has finished and are not exported to other organelles." *Source:* ER resident proteins. (n.d.). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Endoplasmic_reticulum_resident_protein

ER retention signal (KDEL)

A targeting sequence defined by the amino acids KDEL (lysine, aspartate, glutamate, and leucine). If these proteins travel to the Golgi, this sequence helps them to be recognized by a receptor to be returned to the ER.

euchromatin

Loosely packed chromatin with regions rich in genes that can be transcribed in order to be later translated into proteins.

exon

A coding region within a DNA sequence that is spliced together after the introns have been removed from an mRNA sequence allowing for the formation of a mature mRNA segment.

exon skipping

"A form of RNA splicing used to cause cells to `skip' over faulty or misaligned sections (exons) of genetic code, leading to a truncated but still functional protein despite the genetic mutation." *Source:* Exon skipping. (2023, January 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Exon_skipping

exportin

The nuclear export receptor. They bind to proteins with a nuclear export receptor to aid in transport of proteins through the nuclear pore out of the nucleus into the cytosol. They are closely related to importins (which help protein enter the nucleus) and together form a larger family of proteins known as the karyopherins.

extracellular matrix (ECM)

"A network consisting of extracellular macromolecules and minerals, such as collagen, enzymes, glycoproteins and hydroxyapatite that provide structural and biochemical support to surrounding cells." *Source:* Extracellular matrix. (2023, June 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Extracellular_matrix

F

facultative heterochromatin

Regions of chromatin that are densely packaged and not accessible for transcription. However, these regions contain genes and can be opened for transcription when needed by the cell.

feedback loop

A regulatory mechanism whereby the product of a process will influence (upregulate or downregulate) itself.

filamentous protein

"Made up of elongated or fibrous polypeptide chains which form filamentous and sheet-like structures." *Source:* Fibrous protein. (2022, September 7). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Fibrous_protein

flagellum

"A hairlike appendage that protrudes from certain plant and animal sperm cells, and from a wide range of microorganisms to provide motility." *Source:* Flagellum. (2023, July 13). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Flagellum

flippase

"Transmembrane lipid transporter proteins located in the membrane that belong to ABC transporter or P4-type ATPase families. They are responsible for aiding the movement of phospholipid molecules between the two leaflets that compose a cell's membrane (transverse diffusion, also known as a 'flip-flop' transition)." *Source:* Flippase. (n.d.). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Flippase

flow cytometer

A machine that analyzes properties (such as fluorescence) as it passes through.

fluorescence recovery after photobleaching (FRAP)

A method to determine the mobility of molecules within a membrane. A region of the cell is photobleached (the fluorescent molecules are destroyed). The same area is monitored over time to determine if fluorescence returns to the bleached area.

fluorescence-activated cell sorting (FACS)

"In this process, a sample containing cells or particles is suspended in a fluid and injected into the flow cytometer instrument. The sample is focused to ideally flow one cell at a time through a laser beam, where the light scattered is characteristic to the cells and their components. Cells are often labeled with fluorescent markers so light is absorbed and then emitted in a band of wavelengths. Tens of thousands of cells can be quickly examined and the data gathered are processed by a computer." *Source:* Flow cytometry. (2023, June 27). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Flow_cytometry

fluorescent light microscopy (a.k.a. emitted light microscopy)

A microscopy technique where samples are illuminated with light of a particular wavelength. The fluorescent molecules in the sample become excited and emit light of a longer wavelength. The microscope catches this light to display the location of the fluorescent molecules bound to a cellular structure of interest. There are two main types of fluorescent light microscopy: epifluorescence and confocal microscopy.

G

g-actin
The monomeric actin subunit that is not attached to any other actin units.

G-protein-coupled receptors (GPCRs)
"A large group of evolutionarily related proteins that are cell surface receptors that detect molecules outside the cell and activate cellular responses. They are coupled with G-proteins." *Source:* G-protein-coupled receptor. (2023, April 16). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/G_protein-coupled_receptor

G-proteins (guanine nucleotide-binding proteins)
"A family of proteins that act as molecular switches inside cells and are involved in transmitting signals from a variety of stimuli outside a cell to its interior. Their activity is regulated by factors that control their ability to bind to and hydrolyze guanosine triphosphate (GTP) to guanosine diphosphate (GDP)." *Source:* G-protein. (2023, July 21). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/G_protein

G0 phase
A life stage that exists outside of the cell cycle. Cells can enter this phase due to poor environmental conditions, terminal differentiation, or other cellular cues.

G1 (Gap or Growth 1) phase
"The first of four phases of the cell cycle that takes place in eukaryotic cell division. In this part of interphase, the cell synthesizes mRNA and proteins in preparation for subsequent steps leading to mitosis. G1 phase ends when the cell moves into the S phase of interphase. Around 30 to 40 percent of cell cycle time is spent in the G1 phase." *Source:* G1 phase. (2022, December 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/G1_phase

G1/S checkpoint
A cell cycle checkpoint that regulates the progression of a cell into the S (synthesis) stage of the cell cycle.

G2 (Gap or Growth 2) phase
"The third subphase of interphase in the cell cycle directly preceding mitosis. It follows the successful completion of S phase, during which the cell's DNA is replicated. G2 phase ends with the onset of prophase, the first phase of mitosis in which the cell's chromatin condenses into chromosomes." *Source:* G2 phase. (2022, December 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/G2_phase

G2/M checkpoint
A cell cycle checkpoint that regulates the progression of a cell into the M (mitosis) stage of the cell cycle.

gain-of-function experiment
An experiment where a protein is given a new function through introducing a mutation and the result on the system is measured. Gain-of-function experiments are used to explore what the absolute minimum requirements are for a particular function (i.e., what is sufficient).

gain-of-function mutation
A form of mutation that causes a gene to produce altered products that have gained new molecular functions compared to the unaltered gene products.

gamma-tubulin
A specialized tubulin network that acts as a nucleating site for microtubule growth.

gel electrophoresis
"A method for separation and analysis of biomacromolecules (DNA, RNA, proteins, etc.) and their fragments, based on their size and charge." *Source:* Gel electrophoresis. (2023, July 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Gel_electrophoresis

general transcription factors (GTFs)
"A class of protein transcription factors that bind to specific sites (promoter) on DNA to activate transcription of genetic information from DNA to messenger RNA." *Source:* General transcription factor. (2023, April 11). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/General_transcription_factor

genetic code expansion (GCE)
A biochemical technique to encode designer amino acids into genetically engineered proteins.

glycocalyx
"A glycoprotein and glycolipid covering that surrounds the cell membranes of bacteria, epithelial cells, and other cells." *Source:* Glycocalyx. (2023, July 9). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Glycocalyx

glycolysis
"The metabolic pathway that converts glucose (C6H12O6) into pyruvate." *Source:* Glycolysis. (2023, July 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Glycolysis

glycosylation
"The reaction in which a carbohydrate (or 'glycan'), i.e., a glycosyl donor, is attached to a hydroxyl or other functional group of another molecule." *Source:* Glycosylation. (2023, June 21). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Glycosylation

glycosyltransferase
An enzyme that adds a carbohydrate group ('glycan') to a biological molecule.

Golgi apparatus
A compartment at the heart of the endomembrane system responsible for sorting and processing proteins in that system.

grana
"Chloroplast thylakoids frequently form stacks of disks referred to as grana (singular: granum)." *Source:* Thylakoid. (2023, May 24). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Thylakoid

green fluorescent protein (GFP)
A naturally occurring protein found in jellyfish that glows green when excited with blue wavelength light. It is commonly fused genetically to proteins of interest to visualize the interior of cells.

GTP cap
A buildup of GTP bound beta subunit accumulates at the end of a microtubule, protecting it from depolymerization (catastrophe).

GTPase
Enzyme that functions as a molecular switch. GTPases bind to GTP and GDP. Depending on their bound state, they will have different functions. This is a large class of proteins with a wide variety of functions.

GTPase-activating protein, or GAP
A family of proteins that helps a G protein accelerate its GTPase activity (converting GTP into GDP).

guanine-exchange factor, or GEF

A family of proteins that will swap a GDP for GTP on a G protein.

H

haploid

Refers to organisms that contain one copy of each chromosome.

hemidesmosomes

"Very small stud-like structures found in keratinocytes of the epidermis of skin that attach to the extracellular matrix." *Source:* Hemidesmosome. (2023, April 21). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Hemidesmosome

heterochromatin

Tightly packed chromatin that is not transcriptionally active.

heterotroph

"An organism that cannot produce its own food, instead taking nutrition from other sources of organic carbon, mainly plant or animal matter." *Source:* Heterotroph. (2023, February 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Heterotroph

histone

Proteins that bind to DNA to package into nucleosomes. They are highly basic, resulting in a positive charge. This allows them to bind to the negative backbone of DNA nonselectively.

histone exchange

A chromatin remodeling process where an enzyme removes one histone core and replaces it with a histone core with slight variations in composition.

histone-modifying enzymes

"Enzymes involved in the modification of histone substrates after protein translation and affect cellular processes including gene expression." *Source:* Histone-modifying enzymes. (2023, May 9). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Histone-modifying_enzymes

hormone

"A class of signaling molecules in multicellular organisms that are sent to distant organs by complex biological processes to regulate physiology and behavior." *Source:* Hormone. (2023, July 24). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Hormone

hyaluronic acid (HA)

A carbohydrate polymer often found in the extracellular matrix and often serves a lubrication function.

hydropathy index

An index (reported on the y axis) that is a measure of the average hydrophobicity of a given amino acid residue (hydropathy score) and the average hydropathy score of a range of amino acids on either side combined into one value.

hydropathy plot

A graph used to predict the number of position of transmembrane domains in an amino acid sequence. Each amino acid's hydropathy score is plotted with the N terminus of the protein starting at position 0 on the x axis. A peak on this plot above the threshold indicates a potential transmembrane domain.

hydropathy score

A set value indicating the hydrophobicity of a given amino acid.

hydrophilic

Latin for "water loving." Refers to chemical groups that readily associate with polar molecules like water.

hydrophobic

Latin for "water fearing." Refers to chemical groups that do not readily associate with polar molecules like water. Instead, these are generally nonpolar in nature and group with other nonpolar molecules.

hydrophobic effect

"The observed tendency of nonpolar substances to aggregate in an aqueous solution and exclude water molecules." *Source:* Hydrophobic effect. (2023, June 13). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Hydrophobic_effect

I

immunolabeling

"A biochemical process that enables the detection and localization of an antigen to a particular site within a cell, tissue, or organ. Antigens are organic molecules, usually proteins, capable of binding to an antibody." *Source:* Immunolabeling. (2023, February 9). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Immunolabeling

immunoprecipitation

A technique that uses an antibody bound to an insoluble bead to selectively isolate a protein out of a solution. This will purify and concentrate a protein of interest and anything that is bound to that protein.

importin

"A type of karyopherin that transports protein molecules from the cell's cytoplasm to the nucleus. It does so by binding to specific recognition sequences, called nuclear localization sequences (NLS)." *Source:* Importin. (2022, November 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Importin

in silico

Refers to experiments conducted in a computer; computer simulations.

inositol

A ringed sugar molecule that is commonly used in signaling pathways.

integral membrane protein

A type of protein that is embedded into biological membranes. Transmembrane proteins are an example of integral membrane proteins that span the entire biological membrane and are used to transport material from one side to the other.

intermediate filaments

A family of proteins that polymerize to form ropelike structures. These provide mechanical strength as part of the cytoskeletal system.

intermolecular forces (IMF)

"The force that mediates interaction between molecules, including the electromagnetic forces of attraction or repulsion which act between atoms and other types of neighboring particles, e.g., atoms or ions. Intermolecular forces are weak relative to intramolecular forces—the forces which hold a molecule together." *Source:* Intermolecular force. (2023, June 22). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Intermolecular_force

internal START transfer sequence

A start signal sequence that is somewhere internal on the protein as opposed to at the N-terminus of the protein. The START sequence is always either the first ER insertion sequence or one that directly follows a START (if there are multiple ER insertion sequences). Unlike the N-terminal START sequence, this is not cleaved off after the transfer of the protein has been completed.

interphase

All phases of the cell cycle that are not mitosis or meiosis. This stage contains G1, S, and G2 phases.

interpolar microtubules
The microtubules emanating from a spindle pole that attach in the center (between the two poles).

intron
A region of noncoding DNA that is usually located between two exons and is spliced out of the mRNA segment after transcription.

ion-channel-coupled receptors
When a signal binds to this type of receptor, an ion channel is opened as a result.

IP3 (inositol triphosphate)
"An inositol phosphate signaling molecule. It is made by hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP2), a phospholipid that is located in the plasma membrane, by phospholipase C (PLC). Together with diacylglycerol (DAG), IP3 is a second messenger molecule used in signal transduction in biological cells." *Source:* Inositol trisphosphate. (2023, May 1). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Inositol_trisphosphate

J

juxtacrine
In this contact dependent signaling pathway, a cell attaches to another cell to initiate the signaling pathway.

K

karyopherin
"Proteins involved in transporting molecules between the cytoplasm and the nucleus of a eukaryotic cell." *Source:* Karyopherin. (2023, February 22). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Karyopherin

KDEL
A targeting sequence defined by the amino acids KDEL (lysine, aspartate, glutamate, and leucine). If these proteins travel to the Golgi, this sequence helps them to be recognized by a receptor to be returned to the ER.

KDEL receptor
A protein that binds to the KDEL amino acid sequences in the Golgi to help load them into vesicles back to the ER.

kinase
An enzyme that adds a phosphate chemical group to another enzyme. These can be added on serines, threonines, or tyrosine.

kinesin
A molecular motor that uses microtubules. Most kinesin motors move toward the plus end, but not all.

kinetochore
"A disc-shaped protein structure associated with duplicated chromatids in eukaryotic cells where the spindle fibers attach during cell division to pull sister chromatids apart." *Source:* Kinetochore. (2023, March 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Kinetochore

kinetochore microtubule
A microtubule emanating from the spindle pole that attaches to the kinetochore of a chromosome during mitosis.

L

(to column two ->)

lag phase
A beginning stage of a chemical reaction where the reaction proceeds slowly.

lamellipodium
"A cytoskeletal protein actin projection on the leading edge of the cell." *Source:* Lamellipodium. (2022, November 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Lamellipodium

lamin (nuclear)
"Fibrous proteins in type V intermediate filaments, providing structural function and transcriptional regulation in the cell nucleus. Nuclear lamins interact with inner nuclear membrane proteins to form the nuclear lamina on the interior of the nuclear envelope." *Source:* Lamin. (2023, January 21). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Lamin

lamin-associated domain (LAD)
Areas on the nuclear lamina that link the nuclear lamina to specific regions of chromosomes.

lariat loop
The intronic DNA between two exons that is 'looped' out during the splicing process by the spliceosome. It is so called because the shape mimics a lasso loop used by farm hands when wrangling animals.

leading edge
The portion of the cell that is at the "front" with respect to the direction of movement.

leaflet
A single layer of a lipid bilayer.

ligand (signaling)
A molecule that binds to a receptor to initiate a signaling cascade.

light-dependent reactions
The biochemical reactions that occur in the plant chloroplast that depend on light. These include the process of photophosphorylation.

linker DNA
The DNA between two nucleosomes.

lipid raft
A distinct membrane region that contains specialized lipids and proteins. Often these regions are a bit thicker than surrounding membrane areas and provide specialized functions like cell signaling.

liposomes
"A small artificial vesicle, spherical in shape, having at least one lipid bilayer." *Source:* Intermolecular force. (2023, June 22).
In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Intermolecular_force

loss-of-function experiment
An experiment designed to test how the functional inhibition of a protein/gene impacts the system. Loss-of-function experiments help us determine what aspects of the system are required (i.e., *necessary*).

loss-of-function mutation
A form of mutation that causes a gene to produce proteins that are lacking molecular functions present in the nonmutated protein.

lumen
The area inside a membrane-bound organelle. For instance, the ER lumen, Golgi lumen, and lysosomal lumen.

lysosomal pathway
Refers to the vesicle trafficking pathway originating at the TGN and ending at the lysosome for lysosomal resident proteins. Proteins are recognized by a receptor at the TGN and transported to the endosome where the protein is disassociated from the receptor. The protein then travels to the lysosome. The receptor is transported back to the TGN for further rounds of trafficking.

lysosome

A roughly spherical membrane-bound organelle that contains digestive enzymes. This compartment breaks down materials delivered from the endocytic pathway.

M

M (mitosis) phase
The stage of the cell cycle where mitosis and cytokinesis occur.

magnification
Refers to enlarging the apparent size of an image. This is akin to scaling or zooming in on an image and does not change its inherent resolution.

mannose-6-phosphate (M6P)
A mannose sugar phosphorylated on the sixth carbon with a phosphate.

mannose-6-phosphate receptor (M6PR)
The protein receptor that recognizes M6P on proteins.

MAP kinase cascade
A three-part signaling cascade where a MAPKKK phosphorylates a MAPKK, which then phosphorylates a MAPK. These are part of a larger signaling pathway.

matrix
The area inside the mitochondria where the TCA cycle, mitochondrial DNA, and mitochondrial ribosomes are located.

messenger RNA (mRNA)
"A single-stranded molecule of RNA that corresponds to the genetic sequence of a gene, and is read by a ribosome in the process of synthesizing a protein." *Source:* Messenger RNA. (2023, June 7). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Messenger_RNA

metaphase
The stage of mitosis where chromosomes are lined up in a line on the metaphase plate.

metaphase checkpoint
Ensures that chromosomes are properly lined up prior to starting the next phase, anaphase.

metaphase plate
An invisible axis where chromosomes line up during metaphase.

methylation
"Denotes the addition of a methyl group on a substrate, or the substitution of an atom (or group) by a methyl group." *Source:* Methylation. (2023, July 4). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Methylation

microtubule organizing center (MTOC)
"A structure found in eukaryotic cells from which microtubules emerge." *Source:* Microtubule organizing center. (2023, March 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Microtubule_organizing_center

microtubules (MT)
"Polymers of tubulin that form part of the cytoskeleton and provide structure and shape to eukaryotic cells." *Source:* Microtubule. (2023, July 15). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Microtubule

minus end
The portion of the microtubule filament that grows slower and has a higher critical concentration. This is the end with alpha tubulin exposed.

mitochondrial processing peptidase (MPP)
"An enzyme complex found in mitochondria which cleaves signal sequences from mitochondrial proteins." *Source:* Mitochondrial processing peptidase. (2023, January 29). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Mitochondrial_processing_peptidase

mitochondrion
A membrane-bound organelle in a cell known to produce and secrete a chemical currency called ATP.

mitosis
A cellular process of nuclear division.

mitotic chromosome
Refers to the structure of a chromosome during the cellular process of mitosis. In this stage, the DNA is fully condensed, and the DNA appears as the characteristic X's.

mitotic spindle
A microtubule arrangement established during mitosis.

molecular switch (protein)
"Proteins can switch between active and inactive states, thus acting as molecular switches in response to another signal. For example, phosphorylation of proteins can be used to activate or inactivate proteins." *Source:* Molecular switch. (2023, June 10). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Molecular_switch

monolayer-associated protein
A protein that is embedded in a biological membrane but only interacts with one layer of the bilayer.

monomer
Composed of one unit. With a protein, this refers to having a single polypeptide chain.

monomeric
Existing as a single unit. With a protein, this refers to having a single polypeptide chain.

motor proteins
Proteins that bind to cytoskeletal filaments and use energy to move along them.

myosin
A molecular motor that uses actin filaments. Most myosin motors move toward the plus end, but not all.

N

N-linked glycosylation
"The attachment of an oligosaccharide, a carbohydrate consisting of several sugar molecules, sometimes also referred to as glycan, to a nitrogen atom (the amide nitrogen of an asparagine [Asn] residue of a protein), in a process called N-glycosylation." *Source:* N-linked glycosylation. (2023, July 9). In *Wikipedia, The Free Encyclopedia.* https://en.wikipedia.org/wiki/N-linked_glycosylation

N-terminal START transfer sequence
A sequence of hydrophobic amino acids located at the N-terminus of a protein. This sequence will be cleaved after the protein has been imported into the ER.

native gel electrophoresis
"Native gels are run in non-denaturing conditions so that the analyte's natural structure is maintained. This allows the physical size of the folded or assembled complex to affect the mobility, allowing for analysis of all four levels of the biomolecular structure." *Source:* Gel electrophoresis. (2023, July 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Gel_electrophoresis

negative feedback loop
A regulatory mechanism whereby the product of a process will negatively influence (downregulate) itself.

negative staining
"An established method, often used in diagnostic microscopy, for contrasting a thin specimen with an optically opaque fluid. In this technique, the background is stained, leaving the actual

specimen untouched, and thus visible." *Source:* Negative stain. (2021, December 4). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Negative_stain

neuronal
Referring to neurons, an electrically excitable cell that is associated with the brain.

neurotransmitters
"A signaling molecule secreted by a neuron to affect another cell across a synapse." *Source:* Neurotransmitter. (2023, April 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Neurotransmitter

nitric oxide (NO)
"In mammals, including humans…a signaling molecule in many physiological and pathological processes." *Source:* Nitric oxide. (2023, July 11). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Nitric_oxide

nonhistone chromatin-associated protein
A protein whose function aids in maintenance of DNA packing but is not in the histone family.

nonpolar
Refers to molecules that have *do not* have one end of the molecule that is different from the other end. Most often, this refers to electron distribution. Nonpolar molecules share electrons evenly and never carry a dipole or charge. *Nonpolar* can also refer to the symmetrical shape of molecules, such as the tetrameric form of intermediate filaments.

nuclear bodies
Membraneless structures found in the nucleus. A common example is the nucleolus.

nuclear envelope
The lipid bilayer that surrounds the nucleus. It is a double bilayer with a perinuclear space in between. The outer membrane is continuous with the ER, and nuclear pores go through both layers of the envelope.

nuclear export receptor
A protein that binds to a nuclear export signal to help proteins exit the nucleus.

nuclear export signal (NES)
A short targeting sequence found in proteins that allows them to bind to the nuclear export receptor to be transported out of the nucleus into the cytoplasm.

nuclear import receptor (NIR)
A protein that binds to a nuclear localization signal (NLS) to help proteins enter the nucleus through the nuclear pore.

nuclear lamina
Fibrous mesh formed by intermediate filaments and membrane-associated proteins found just inside the nuclear envelope of cells. It functions as a structural support and helps maintain chromosome organization among other functions.

nuclear lamins
A type of intermediate filament found in animal cells. They polymerize to form the nuclear lamina.

nuclear localization sequence (NLS)
A sequence of amino acids (KKKRK) that marks the protein to be sent to the nucleus of the cell via nuclear transport. Without an NLS tag, the protein will remain in the cytosol and won't be capable of entering the nucleus.

nuclear pore
A complex of proteins that spans the double membrane of the nuclear envelope, which surrounds a cell's nucleus. The pore serves as a method of transport for molecules to be able to enter or exit the nucleus through either a passive or active process.

nuclear receptors (signaling)
"A class of proteins responsible for sensing steroids, thyroid hormones, vitamins, and certain other molecules. These intracellular receptors work with other proteins to regulate the expression of specific genes thereby controlling the development, homeostasis, and metabolism of the organism." *Source:* Nuclear receptor. (2023, May 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Nuclear_receptor

nucleation sites
A preformed structure that allows molecules in a reaction to be more favorably arranged. In regards to the cytoskeleton, nucleation sites include small preformed pieces of cytoskeleton. These make it easier for other cytoskeletal monomers to attach and thus speeds up the rate of reaction.

nucleolus
A membraneless organelle within the nucleus where ribosome biogenesis takes place.

nucleolus organizing regions (NORs)
"Chromosomal regions crucial for the formation of the nucleolus." *Source:* Nucleolus organizer region. (2022, October 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Nucleolus_organizer_region

nucleoplasm
The fluid and molecules found within the nucleus.

nucleosome
"The basic structural unit of DNA packaging in eukaryotes. The structure of a nucleosome consists of a segment of DNA wound around eight histone proteins and resembles thread wrapped around a spool." *Source:* Nucleosome. (2023, June 18). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Nucleosome

nucleosome eviction
A chromatin remodeling process where an enzyme removes a core set of histone from a region of DNA. This opens up this region so that it is able to be accessed by other proteins.

nucleosome sliding
A chromatin remodeling process where a chromatin remodeling enzyme moves a nucleosome thereby exposing the DNA that was originally around the histone core. This opens up this region so that it is able to be accessed by other proteins.

O

O-linked glycosylation
"The attachment of a sugar molecule to the oxygen atom of serine (Ser) or threonine (Thr) residues in a protein." *Source:* O-linked glycosylation. (2021, October 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/O-linked_glycosylation

octamer
Refers to the eight histone proteins that form the core of the nucleosome. It contains two copies of each of the four core histone proteins (H2A, H2B, H3, and H4). DNA is wrapped around this structure to create a nucleosome.

odorants
The chemicals that make up smell molecules. These are detected by odorant receptors in the nose. Subsequent signaling and activation of neurons to the brain help us sense these molecules.

oxidative phosphorylation
"The metabolic pathway in which cells use enzymes to oxidize nutrients, thereby releasing chemical energy in order to produce adenosine triphosphate (ATP)." *Source:* Oxidative phosphorylation. (2023, July 17). In *Wikipedia, The Free*

Encyclopedia. https://en.wikipedia.org/wiki/Oxidative_phosphorylation

P

paracrine
A short-range signaling pathway where a cell secretes a signaling molecule that is picked up by target cells in the local vicinity.

PCR (polymerase chain reaction)
"A method widely used to rapidly make millions to billions of copies (complete or partial) of a specific DNA sample, allowing scientists to take a very small sample of DNA and amplify it (or a part of it) to a large enough amount to study in detail." *Source:* Polymerase chain reaction. (2023, June 27). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Polymerase_chain_reaction

perinuclear space
The region between the two membranes of the nuclear envelope.

peripheral membrane protein
Membrane proteins that do not embed in the biological membrane. Instead, they bind to other integral proteins or more rarely interact with the lipids themselves.

phagocytosis
The process of bringing in a large particle (like a bacterium or a cancerous cell) for digestion.

phagosome
"A vesicle formed around a particle engulfed by a phagocyte via phagocytosis." *Source:* Phagosome. (2023, February 11). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phagosome

phase-contrast microscopy (PCM)
"An optical microscopy technique that converts phase shifts in light passing through a transparent specimen to brightness changes in the image." *Source:* Phase-contrast microscopy. (2023, July 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phase-contrast_microscopy

phosphatases
Proteins that remove a phosphate from a target protein.

phosphodiesterase
An enzyme that cleaves a phosphodiester bond. These occur in between nucleic acids or within a ringed nucleotide such as cyclic AMP.

phospholipase C (PLC)
"Phospholipase C's role in signal transduction is its cleavage of phosphatidylinositol 4,5-bisphosphate (PIP2) into diacyl glycerol (DAG) and inositol 1,4,5-trisphosphate (IP3), which serve as second messengers." *Source:* Phospholipase C. (2023, June 24). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phospholipase_C

phospholipid
"A class of lipids whose molecule has a hydrophilic 'head' containing a phosphate group and two hydrophobic 'tails' derived from fatty acids, joined by an alcohol residue (usually a glycerol molecule)." *Source:* Phospholipid. (2023, May 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phospholipid

phospholipid bilayer
Composed of two sheets of phospholipids. The lipids are orientated so that the polar heads are facing outward and the nonpolar fatty acid tails are facing inward.

phosphomimetics
An amino acid substitution in a engineered protein to mimic the rough charge and shape of a phosphorylated protein at that location.

phosphorylation
The attachment of a phosphate group to a biological molecule. When applied to a protein, this can often change the conformation of the protein, resulting in a functional change.

photobleach
Refers to fading. Often this is in reference to a fluorescence molecule that is no longer able to be seen.

photophosphorylation
The biochemical process of turning light energy into ATP.

phragmoplast
"A plant cell specific structure that forms during late cytokinesis. It serves as a scaffold for cell plate assembly and subsequent formation of a new cell wall separating the two daughter cells." *Source:* Phragmoplast. (2023, May 12). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phragmoplast

PI-3-kinase
"A family of related intracellular signal transducer enzymes capable of phosphorylating the 3 position hydroxyl group of the inositol ring of phosphatidylinositol (PtdIns)." *Source:* Phosphoinositide 3-kinase. (2022, October 27). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phosphoinositide_3-kinase

pinocytosis
An endocytic process that occurs continually and nonselectively. The membrane invaginates, bringing in small molecules from the external environment.

PIP (phosphatidylinositol)
A family of lipids containing a phosphate head, two fatty acid tails, and an inositol head group. These compose a small but functionally relevant component of the cytosolic side of a cell's membranes.

PIP2
"Phosphatidylinositol 4,5-bisphosphate or PtdIns(4,5)P2, also known simply as PIP2 or PI(4,5)P2, is a minor phospholipid component of cell membranes. PtdIns(4,5)P2 is enriched at the plasma membrane where it is a substrate for a number of important signaling proteins." *Source:* Phosphatidylinositol 4,5-bisphosphate. (2023, January 26). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Phosphatidylinositol_4,5-bisphosphate

pKa
A measure of a molecule's readiness to donate a proton in solution. When an amino acid has a higher pKa, it is more likely to be a basic amino acid, preferring to keep its proton. Conversely, an amino acid with a low pKa will be willing to donate its proton and is called acidic.

plasma membrane
A biological membrane that separates the cell from the surrounding environment. It is composed of a phospholipid bilayer, proteins, and carbohydrate components.

plastid
A membrane-bound organelle found in plants. Commonly plastids will differentiate into chloroplasts, but other examples include leucoplasts and chromoplasts.

plastoglobules
"Spherical bubbles of lipids and proteins about 45–60 nanometers across. They are surrounded by a lipid monolayer. Plastoglobuli are found in all chloroplasts." *Source:* Chloroplast. (2023, July 14). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Chloroplast

plus end
The portion of the microtubule filament that grows faster and has a lower critical concentration. This is the end with beta tubulin exposed.

polar
Refers to molecules that have one end of the molecule that is different from the other end. Most often, this refers to an uneven electron distribution, where extreme uneven distribution of electrons results in ionized molecules with a charge. *Polar* can also refer to shape differences on different ends of a molecule, as is the case with polar cytoskeletal elements.

polarity
Refers to molecules that have a distinct endedness. If referring to electron distribution, molecules with unequal electron sharing are said to be polar. Extreme versions are ions. With regards to the cytoskeleton, the protein face that is exposed is distinct and refers to the polarity in these elements.

polarized light microscopy
A subset of transmitted light microscopy where polarized light is used to enhance the contrast of the image. DIC (differential interface contrast) microscopy is one common example of this technique.

polyadenylation
"The addition of a poly(A) tail to an RNA transcript, typically a messenger RNA (mRNA). The poly(A) tail consists of multiple adenosine monophosphates; in other words, it is a stretch of RNA that has only adenine bases. In eukaryotes, polyadenylation is part of the process that produces mature mRNA for translation." *Source:* Polyadenylation. (2023, March 21). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Polyadenylation

polymer
A strand of repeating units called monomers.

porin
A protein in the outer membrane of chloroplasts and mitochondria that allows small molecules to pass through via diffusion.

positive feedback loop
A regulatory mechanism whereby the products of a process will positively influence (upregulate) themselves.

presequence
The mitochondria and chloroplast targeting signal. It is an N-terminal sequence that forms a helical structure needed for correct receptor recognition and transport to the proper organelle.

primary endosymbiosis
"Involves the engulfment of a cell by another free living organism." *Source:* Symbiogenesis. (2023, May 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Symbiogenesis

primary producers
Autotrophic organisms that provide the base of the food chain.

primary sequence (structure)
Refers to the sequence of the protein. Amino acids are joined into a polymer. The unique sequence of each polypeptide chain dictates how it will fold.

prometaphase
The intermediary mitotic phase between prophase and metaphase. Distinct in this phase is nuclear lamina breakdown and further attachment of kinetochore microtubules.

promoter
The sequence of DNA just prior to the transcription start site. General transcription factors bind here to recruit RNA polymerase to initiate the start of transcription.

prophase
The first phase of mitosis. The DNA begins to condense and the mitotic spindle begins to form.

protein aggregates
Proteins that form clusters with other proteins. The aggregation tends to be composed of misfolded proteins that join to other misfolded proteins.

protein backbone
The repeating atomic structure formed when amino acids are joined through a peptide bond. These form N-C-C repeats referring to the central atoms in the amide, R-group, carboxylic acid groups that form each amino acid.

protein cap
Proteins that bind to microtubule polymers. This prevents them from growing or shrinking.

protein coding region
The region of a gene that directly translates into protein.

protein domain
"A region of a protein's polypeptide chain that is self-stabilizing and that folds independently from the rest." *Source:* Protein domain. (2023, June 4). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Protein_domain

protein kinase A (PKA)
"A family of serine-threonine kinase whose activity is dependent on cellular levels of cyclic AMP (cAMP)." *Source:* Protein kinase A. (2023, July 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Protein_kinase_A

protein processing
The process of modifying proteins after they are translated.

proteoglycans
"Proteins that are heavily glycosylated. The basic proteoglycan unit consists of a 'core protein' with one or more covalently attached glycosaminoglycan (GAG) chain(s)." *Source:* Proteoglycan. (2023, February 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Proteoglycan

proteolysis
"The breakdown of proteins into smaller polypeptides or amino acids." *Source:* Proteolysis. (2023, May 10). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Proteolysis

proteosome
"A protein coplex that is responsible for degrading unneeded or damaged proteins through proteolytic chemical reactions that break peptide bonds." *Source:* Proteasome. (2023, January 29). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Proteasome

protofilament
A single polymerized strand of microtubules. Thirteen such strands associate into a hollow tube-like structure.

proton pump
A protein that moved protons from one side of the membrane to the other, creating a gradient.

proton-motive force
A description of the potential energy captures in a proton gradient. When the protons are allowed to flow down their gradient through ATP synthase, this allows for ATP to be made much like a hydroelectric dam makes electricity from the flow of water through turbines.

Q

quaternary structure

Refers to the interactions between two independent polypeptide chains. Not all proteins have quaternary structure. Some will only have quaternary structure transiently.

R

R group (or side chain)

A variable chemical structure off the central carbon of an amino acid. The different chemical features of the R group define each amino acid and give it its overall characteristics (size, shape, charge etc.).

Rab (protein)

A small protein that is a member of the GTPase family and is important in the processes of vesicle trafficking.

Ran cycle

A process by which the Ran protein is shuttled in and out of the nucleus to facilitate the import and export of molecules through the nuclear import channel.

Ras

"A family of related proteins that are expressed in all animal cell lineages and organs. All Ras protein family members belong to a class of protein called small GTPase, and are involved in transmitting signals within cells." *Source:* RasGTPase. (2023, March 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Ras_GTPase

receptor (signaling)

A protein that binds a signal molecule to transmit a cascade intracellularly.

receptor downregulation

A regulatory process where a receptor is brought into the cell via endocytosis, and then degraded in the lysosome.

receptor inactivation

A regulatory process where a biological molecule directly inhibits the cell receptor's function.

receptor sequestration

A regulatory process where a receptor is endocytosed such that it isn't available to bind to signal molecules on the cell's surface.

receptor-mediated endocytosis

A process in which cargo from outside the cell enters, creating inward budding vesicles from the plasma membrane. This process is controlled by receptors located on the cell surface, and cargo is only allowed to enter after a ligand has bonded to a corresponding receptor.

receptor-tyrosine kinases (RTKs)

"The high-affinity cell surface receptors for many polypeptide growth factors, cytokines, and hormones." *Source:* Receptor tyrosine kinase. (2022, February 19). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Receptor_tyrosine_kinase

recovery curve

The graph formed by a FRAP experiment. It measures the fluorescence intensity of a region over time to assess how much fluorescence "recovers" after the initial photobleaching step.

regulated secretion

A form of secretion that is regulated by external factors like signals that allow certain secretory vesicles to fuse with the plasma membrane.

regulatory DNA sequences

Refers to particular segments of the regulatory region that binds proteins. These proteins modify the level of gene expression. For example, there can be an enhancer protein that, when bound to the regulatory sequence, would upregulate the amount a particular gene that is expressed.

regulatory region

The region outside of the gene region of DNA that binds proteins to modify the level of gene expression for a particular gene.

repressor protein

A protein that inhibits the transcription of a particular gene when it binds to a regulatory DNA sequence.

resolution

"Quantifies how close lines can be to each other and still be visibly resolved. Resolution units can be tied to physical sizes (e.g., lines per mm, lines per inch), to the overall size of a picture (lines per picture height, also known simply as lines, TV lines, or TVL), or to angular subtense." *Source:* Image resolution. (2023, June 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Image_resolution

retrograde traffic

Vesicle tracking that moves opposite of the default movement of traffic. In the secretory pathway, an example would be going from the Golgi back to the ER.

ribonucleoprotein (RNP)

"A complex of ribonucleic acid and RNA-binding protein. These complexes play an integral part in a number of important biological functions that include transcription, translation and regulating gene expression and regulating the metabolism of RNA." *Source:* Nucleoprotein. (2022, July 26). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Nucleoprotein

ribosomal RNA (rRNA)

"A ribozyme which carries out protein synthesis in ribosomes. Ribosomal RNA is transcribed from ribosomal DNA (rDNA) and then bound to ribosomal proteins to form small and large ribosome subunits." *Source:* Ribosomal RNA. (2023, July 3). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Ribosomal_RNA

ribosomes

"Macromolecular machines, found within all cells, that perform biological protein synthesis (mRNA translation). Ribosomes link amino acids together in the order specified by the codons of messenger RNA (mRNA) molecules to form polypeptide chains. Ribosomes consist of two major components: the small and large ribosomal subunits." *Source:* Ribosome. (2023, June 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Ribosome

ribozyme

"RNA molecules that have the ability to catalyze specific biochemical reactions, including RNA splicing in gene expression, similar to the action of protein enzymes." *Source:* Ribozyme. (2023, February 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Ribozyme

RNA capping (5′ cap)

"A specially altered nucleotide on the 5′ end of some primary transcripts such as precursor messenger RNA. This process, known as mRNA capping, is highly regulated and vital in the creation of stable and mature messenger RNA able to undergo translation during protein synthesis." *Source:* Five-prime cap. (2023, January 31). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Five-prime_cap

rosette

A special protein structure where cellulose is made in plant cells destined for use in the cell wall.

rough ER (rER)

A section of the ER where the membrane is studded with ribosomes and is responsible for the process of protein production and protein folding.

S

S (synthesis) phase
A discrete stage in the cell cycle where the whole genome is duplicated.

saturated lipids
A lipid that has no double bonds within the hydrocarbon chain.

scaffold protein
Proteins that "interact and/or bind with multiple members of a signaling pathway, tethering them into complexes. In such pathways, they regulate signal transduction and help localize pathway components (organized in complexes) to specific areas of the cell." *Source:* Scaffold protein. (2021, September 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Scaffold_protein

scanning electron microscopy (SEM)
"A type of electron microscope that produces images of a sample by scanning the surface with a focused beam of electrons." *Source:* Scanning electron microscope. (2023, July 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Scanning_electron_microscope

scission
A term used to describe the process of vesicles detaching from the donor compartment.

scramblase
A protein that will move lipids from one monolayer (leaflet) to the other in a lipid bilayer. They are not selective in direction and do not require an energy input.

SDS-PAGE (SDS–polyacrylamid gel electroporesis)
A method to separate molecules (primarily proteins but sometimes DNA) by size. For protein separation, SDS (a detergent molecule) helps denature (unfold) the protein and gives it a consistent size-to-charge ratio. Then the proteins are loaded into a gel matrix and induced with an electric charge. The negatively charged proteins move toward the positive terminal. Larger molecules take longer to move through the gel and stay at the top, while smaller molecules can move more easily through the gel and reside at the bottom.

second messengers
Biological molecules released as part of a signaling cascade as a way to amplify the signal.

secondary endosymbiosis
"Occurs when the product of primary endosymbiosis is itself engulfed and retained by another free living eukaryote." *Source:* Symbiogenesis. (2023, May 28). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Symbiogenesis

secondary structure
Refers to local repeated backbone-backbone interactions of the polypeptide chain. There are two common forms of secondary structure: alpha helices and beta sheets.

secretion
Refers to proteins leaving the cell through the secretory pathway.

secretory pathway
Refers to the protein trafficking pathway originating at the ER and ending at the plasma membrane, where soluble proteins are released into the extracellular space and integral membrane proteins are embedded in the plasma membrane. Proteins are first imported into the ER and moved to the Golgi for further processing before being sent to the plasma membrane / cell exterior. This is the default pathway for proteins targeted to the ER and does not require additional targeting sequences beyond ER insertion sequences.

selectively permeable
A feature of biological membranes that allows only certain molecules and ions to pass through to the inside of the cell. Selective permeability is useful in controlling the composition of the internal cellular environment.

self-assembly
A cellular structure that assembles without input of energy.

serine/threonine kinases
Kinases (enzymes that add a phosphate group) on serine or threonine amino acids of the target protein.

signal peptidase
An enzyme that is responsible for cleaving the N-terminal start sequences from the rest of the protein once trafficked and inserted into the ER.

signal recognition particle (SRP)
"An abundant, cytosolic, universally conserved ribonucleoprotein (protein-RNA complex) that recognizes and targets specific proteins to the endoplasmic reticulum in eukaryotes." *Source:* Signal recognition particle. (2022, October 18). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Signal_recognition_particle

signal transduction
"The process by which a chemical or physical signal is transmitted through a cell as a series of molecular events." *Source:* Signal transduction. (2023, June 20). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Signal_transduction

signaling cascade
"A series of chemical reactions that occur within a biological cell when initiated by a stimulus." *Source:* Biochemical cascade. (2023, June 16). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Biochemical_cascade

signaling protein inactivation
When a biological molecule that is part of a signaling cascade is prevented from participating in its role.

sister chromatid
Identical copies of a chromatid that are attached at the centromere after genome duplication in S phase. They will remain attached until anaphase, when they are pulled apart toward opposite spindle poles. Each sister chromatid represents one half of a replicated chromosome.

small nuclear ribonucleoproteins (snRNPs; pronounced "snurps")
"RNA-protein complexes that combine with unmodified pre-mRNA and various other proteins to form a spliceosome, a large RNA-protein molecular complex upon which splicing of pre-mRNA occurs." *Source:* SnRNP. (2021, September 20). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/SnRNP

smooth ER (sER)
A region of the ER that does not contain any ribosomes. It is involved in many metabolic processes, including phospholipid synthesis, needed for membrane formation.

SNARE
Part of a class of proteins needed for vesicle fusion. When a vesicle nears a target membrane, if there are appropriately matched SNARE proteins, they will bind. This binding causes the proteins to wind up, physically bringing the membranes in close proximity such that the lipid bilayers can fuse.

soluble protein

A protein that is not embedded in a membrane and can diffuse freely in an aqueous environment.

spliceosome

"A spliceosome is a large ribonucleoprotein (RNP) complex found primarily within the nucleus of eukaryotic cells. The spliceosome is assembled from small nuclear RNAs (snRNA) and numerous proteins." *Source:* Spliceosome. (2023, July 20). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Spliceosome

splicing (introns)

Process in which introns are removed from a pre-mRNA and exons are joined back together. This is required as part of the process to form a mature mRNA.

starch granules

"Storage areas where starch molecules are stored at night." *Source:* Starch. (2023, July 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Starch

stem cells

"Undifferentiated or partially differentiated cells that can differentiate into various types of cells and proliferate indefinitely to produce more of the same stem cell." *Source:* Stem cell. (2023, July 17). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Stem_cell

STOP transfer sequence

An ER insertion sequence that is located after a START sequence. It signifies to the translocation channel to stop threading the translated protein into the ER lumen. This sequence eventually is released into the membrane and exists as a transmembrane domain for the protein.

stroma

The area inside the chloroplast inner membrane that surrounds the thylakoids.

stromal processing peptidase (SPP)

A chloroplast enzyme that cleaves off the transit peptide of proteins in the stroma of chloroplasts.

structural maintenance of chromosome (SMC) complex

"Represent a large family of ATPases that participate in many aspects of higher-order chromosome organization and dynamics." *Source:* SMC protein. (2023, May 23). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/SMC_protein

substrate-level phosphorylation

"A metabolism reaction that results in the production of ATP or GTP supported by the energy released from another high-energy bond that leads to phosphorylation of ADP or GDP to ATP or GTP." *Source:* Substrate-level phosphorylation. (2023, March 13). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Substrate-level_phosphorylation

superresolution microscopy

"A series of techniques in optical microscopy that allow such images to have resolutions higher than those imposed by the diffraction limit, which is due to the diffraction of light." *Source:* Super-resolution microscopy. (2023, May 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Super-resolution_microscopy

T

target-SNARE (t-SNARE)

A type of SNARE that is located in the membranes of target membranes and will pair with specific v-SNAREs found on vesicle membranes.

targeting sequence

A short peptide sequence located in a protein that dictates to which organelle it is transported. Proteins without any such sequences remain in the cytosol.

telomeres

A region at the end of linear chromosomes with repetitive sequences. Specialized proteins will bind to this area. In addition, these repetitive sequences contain no genes and protect the genetic loss when the chromosomes are repeatedly replicated.

telophase

The last stage of mitosis, characterized by the DNA decondensing and the nuclear envelope reforming.

temperature-sensitive mutation

Mutants of a gene where the result is a lack of activity when exposed to its restrictive temperatures. Often proper function is restored when brought back to the optimal (permissive) temperature.

template strand

The strand of DNA that is read by RNA polymerase to make mRNA during transcription. It is read 3′ to 5′.

terminal differentiation

"During terminal differentiation, a precursor cell formerly capable of cell division permanently leaves the cell cycle, dismantles the cell cycle machinery and often expresses a range of genes characteristic of the cell's final function (e.g. myosin and actin for a muscle cell)." *Source:* Cellular differentiation. (2023, April 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Cellular_differentiation

tertiary structure

All the structure of the protein that is not primary or secondary or quaternary. This defines the 3D shape and encompasses interactions between side chains, side chains–backbone, and backbone-backbone (that are not local or repeated).

tether (protein)

Tether proteins in the target membrane help with vesicle docking. Tether proteins bind to the Rab proteins to bring the vesicle closer to the target membrane such that SNARE proteins are then able to bind and promote vesicle fusion.

tetramer

A group of four proteins.

thylakoid lumen

"The thylakoid lumen is a continuous aqueous phase enclosed by the thylakoid membrane. It plays an important role for photophosphorylation during photosynthesis." *Source:* Thylakoid. (2023, May 24). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Thylakoid

thylakoid membrane

"The thylakoid membrane is the site of the light-dependent reactions of photosynthesis with the photosynthetic pigments embedded directly in the membrane." *Source:* Thylakoid. (2023, May 24).
In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Thylakoid

topologically associated domains (TADs)

"A topologically associating domain (TAD) is a self-interacting genomic region, meaning that DNA sequences within a TAD physically interact with each other more frequently than with sequences outside the TAD." *Source:* Topologically associating domain. (2023, June 10). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Topologically_associating_domain

trans Golgi network (TGN)

The end location of the Golgi where proteins are sorted into vesicles bound for their next destination. Primary post-Golgi locations include the plasma membrane and the endosome.

transcribed region
A region of DNA that is transcribed into RNA. There are three types of transcripts: mRNA, rRNA, and tRNA.

transcription factors
Proteins involved in regulating the transcription of genes.

transcription start site (+1 site)
The location of a gene where the DNA begins to be transcribed into RNA.

transcription stop site
A specific sequence that RNA polymerase recognizes and that causes transcription of DNA to stop. Normally the transcription stop site is located and the end of a gene.

transcytosis
"A type of transcellular transport in which various macromolecules are transported across the interior of a cell. Macromolecules are captured in vesicles on one side of the cell, drawn across the cell, and ejected on the other side." *Source:* Transcytosis. (2022, September 25). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Transcytosis

transduced
Refers to the process of signal transduction. When a signal is transduced, it is changed from its original form along the pathways. For example, when a signal binds to a receptor, this induces a conformational change in the receptor to activate it. The signal can be said to be transduced from a physical molecule to an activated receptor.

transfer RNA (tRNA)
A special type of RNA that carries an amino acid. During the process of translation, the tRNA binds to a correct codon and the amino acid is transferred to the growing peptide.

transit peptide
An N-terminal region of amino acids that serves as the chloroplast targeting sequence.

transitional ER (tER)
A region of ER where the rER and sER meet. Often it is the exit site for vesicles leaving the ER.

translation start site
The three-letter start codon signals the ribosome to begin translation of the messenger RNA transcript.

translation stop site (stop codon)
Termination marks the end of translation. This happens when a three-letter stop codon in the messenger RNA is reached and placed in the A site of the ribosome.

translocation channel (ER)
Channel necessary for inserted proteins into the ER. When an ER insertion sequence emerges from a protein, SRP binds and brings it to the SRP receptor, which interacts with the translocation channel. Anything after this sequence is threaded through the channel in the ER lumen unless it runs into a STOP translocation sequence.

transmembrane protein
A protein that has a region that extends through the full membrane. As a result, the protein will have exposure to both sides of the membrane.

transmission electron microscopy (TEM)
"Transmission electron microscopy (TEM) is a microscopy technique in which a beam of electrons is transmitted through a specimen to form an image. The specimen is most often an ultrathin section less than 100 nm thick or a suspension on a grid. An image is formed from the interaction of the electrons with the sample as the beam is transmitted through the specimen." *Source:* Transmission electron microscopy. (2023, June 2). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Transmission_electron_microscopy

transport proteins
Refers to proteins in a membrane that facilitate the movement of molecules across a biological membrane.

transporter inner chloroplast membrane (TIC)
Part of the TIC/TOC complex, TIC (transporter inner chloroplast membrane) lives in the inner chloroplast membrane to help facilitate chloroplast proteins' transfer from the cytoplasm into the chloroplast stroma.

transporter inner mitochondrial membrane (TIM)
Part of the TIM/TOM complex, TIM (transporter inner mitochondrial membrane) lives in the inner mitochondrial membrane to help facilitate mitochondrial proteins' transfer from the cytoplasm into the mitochondrial matrix.

transporter outer chloroplast membrane (TOC)
Part of the TIC/TOC complex, TOC (transporter outer chloroplast membrane) lives in the outer chloroplast membrane to help facilitate chloroplast proteins' transfer from the cytoplasm into the chloroplast stroma.

transporter outer mitochondrial membrane (TOM)
Part of the TIM/TOM complex, TOM (transporter outer mitochondrial membrane) lives in the outer mitochondrial membrane to help facilitate mitochondrial proteins' transfer from the cytoplasm into the mitochondrial matrix.

triskelion
The name for clathrin monomers that form a cage-like structure around a forming transport vesicle.

tubulin
A family of proteins. Notably, alpha and beta tubulin dimers will polymerize into microtubules.

U

ubiquitin
"A small (8.6 kDa) regulatory protein found in most tissues of eukaryotic organisms." *Source:* Ubiquitin. (2023, July 18). In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Ubiquitin

unfolded protein response
A cellular response to unfolded proteins whereby they are targeted for destruction.

unsaturated lipid
A lipid that has at least one double bonded carbon in its chain.

upstream
The region 5′ of the indicated location in a piece of DNA.

V

vacuole
An organelle present in plants and fungi that is responsible for the degradation of materials originating in the endocytic pathway. Vacuoles also play a storage role that is not present in the animal lysosome, which is the functional equivalent in animal cells.

vesicle
A small membrane-bound structure containing cargo proteins destined for a particular cellular location in the endomembrane system.

vesicle budding
The process of a vesicle forming from the donor membrane in a cell or organelle.

vesicle docking
Involves the use of tether proteins to get closer to the target compartment prior to fusion.

vesicle fusion
The process of merging a vesicle membrane with the membrane of the target compartment, thereby releasing the contents into the lumen of the target (or the extracellular

space in the case of the cell membrane). The process is mediated by SNARE proteins.

vesicle transport

The movement of proteins and other biological molecules through the cells in membrane compartments called vesicles.

vesicle-SNARE (v-SNARE)

A type of SNARE located in the membrane of transport vesicles.

vesicular transport model

A model describing the function of the Golgi. In this model, the cisternae stay stable in a Golgi. Proteins traveling through this system will bud from each cisterna and fuse with the next subcompartment using vesicles.

Δ

ΔG = ΔH – TΔS

The Gibbs free energy equation. When the ΔG is negative, this indicates a spontaneous reaction and does not need extra energy input to occur.

ABOUT THE AUTHORS AND ILLUSTRATOR

Dr. Lauren Dalton is a Senior Instructor I / Academic Advisor in the Department of Biochemistry and Biophysics at Oregon State University in Corvallis, Oregon, where she teaches introductory and advanced cell biology as well as a course on scientific writing. Her PhD work at the University of British Columbia–Vancouver leveraged the "awesome power of yeast genetics" to understand the molecular mechanisms of vesicle trafficking in the endomembrane system. From her first time as a teaching assistant for Cell Biology, it was clear that teaching was her passion. Since then, she has devoted her career to teaching students with modern pedagogy and is an advocate for inclusive teaching. With two young kids at home, she has been known to bring in a toy or two to illustrate scientific concepts!

Dr. Robin Young is an Assistant Professor of Teaching at the University of British Columbia's Okanagan campus in Kelowna, British Columbia, Canada. She is an award-winning microscopist with a research background in plant an animal cell biology, focusing on the morphology and function of the endomembrane system. She has been teaching cell biology in some capacity since 2003 and has been curating precursor forms of this textbook since 2010. She's an advocate of active classrooms, accessibility in higher education, and showcasing the differences in cellular function across the various Eukaryotic kingdoms. Her favorite organelle is the plant Golgi.

Heather Ng-Cornish is a graphic designer from St. John's, Newfoundland, currently based in Victoria, British Columbia. She completed her BSc (Hons) with co-op distinction in environmental chemistry at the University of British Columbia in Kelowna, British Columbia. During her degree, Heather completed her honors thesis using analytical chemistry to study smoke taint in Okanagan wine grapes due to wildfires. Throughout her degree she gravitated toward visual representations of scientific processes to solidify her understanding. Now Heather uses her skills and education to aid scientists in communicating complex concepts to all audiences. Combining sciences and art is a dream come true! For more information, see heatherngcornish.com.