THE GREAT COURSES®

Topic
Science & Mathematics

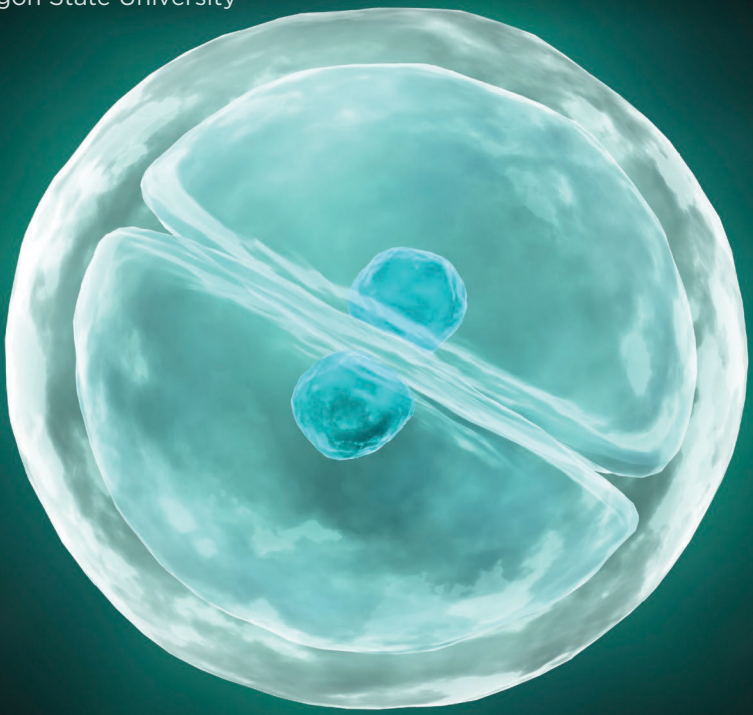
Subtopic
Biochemistry

Biochemistry and Molecular Biology

How Life Works

Course Guidebook

Professor Kevin Ahern
Oregon State University

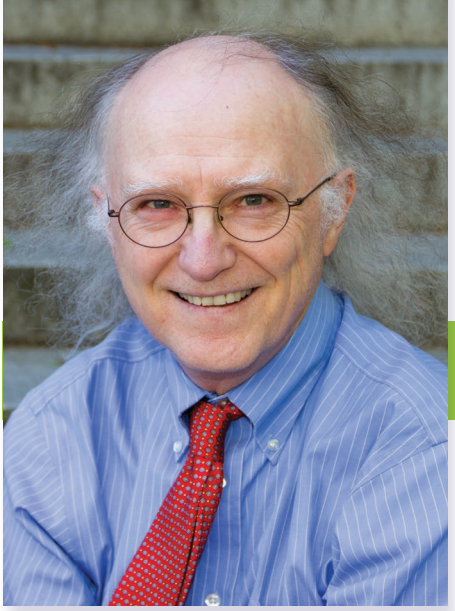


Published by
THE GREAT COURSES

Corporate Headquarters

4840 Westfields Boulevard | Suite 500 | Chantilly, Virginia | 20151-2299

[PHONE] 1.800.832.2412 | [FAX] 703.378.3819 | [WEB] www.thegreatcourses.com



Kevin Ahern, PhD

Professor of Biochemistry and Biophysics

Oregon State University

Kevin Ahern is a Professor of Biochemistry and Biophysics at Oregon State University (OSU). He received BS and MS degrees from Oklahoma State University and a PhD in Biochemistry and Biophysics from OSU. Trained as a molecular biologist, Professor Ahern's primary area of scholarly activity is in biochemistry instruction. He has served on the OSU faculty in Biochemistry/Biophysics since the mid-1990s, teaching undergraduate and graduate courses and serving as head advisor of the department. A passionate supporter of undergraduate research, Professor Ahern served for 14 years as director of OSU's Howard Hughes Medical Institute Summer Undergraduate Research Program and for 3 years as OSU's first Director for Undergraduate Research. He is also the principal investigator and has served since 2014 as director of OSU's National Science Foundation–funded STEM Leaders Program, which uses research experiences as tools for retention of underrepresented minority students in STEM disciplines.

Professor Ahern coauthored 3 popular biochemistry textbooks: *Biochemistry* (3rd edition with Christopher K. Mathews and Kensal E. van Holde), *Biochemistry Free and Easy*, and *Biochemistry Free for All*. The last 2 books were cowritten with Indira Rajagopal and are open educational resources that have been downloaded more than ¼ of a million times.

Paralleling his academic career, Professor Ahern had an extensive career in scientific publishing, serving as a contributing editor to *Science* magazine, *Genetic Engineering & Biotechnology News*, and *BioTechniques*. He has written and published more than 700 articles, including 39 in *Science* magazine and 42 in *Biochemistry and Molecular Biology Education*. Professor Ahern's approach to classroom teaching is novel. He writes and performs music, verses, and limericks to help students learn complicated material. He has more than 100 metabolic melodies he sings to, and with, students in his classes.

Professor Ahern's efforts have been recognized with numerous local, regional, and national awards. He was a 2-time national finalist for Baylor University's prestigious Robert Foster Cherry Award for Great Teaching and was OSU's nominee for the US Professor of the Year Award in 2009. Professor Ahern received OSU's highest teaching recognition, the Elizabeth P. Ritchie Distinguished Professor Award, and is an Eminent Professor of OSU's Honors College. He was named an OSU top professor by students a record 14 times and was inducted into OSU Libraries' Open Access Hall of Fame in 2013. In 2019, Professor Ahern was the recipient of a Fulbright US Scholar Award that provided funding for him to teach biochemistry at the University of Malta.

As an academic advisor, Professor Ahern won every university award given in that field. His mentoring of students was also recognized regionally with the Oregon Health & Science University Foundation's Medical Research Foundation Mentor Award and nationally with the National Academic Advising Association's Outstanding Advising Award. ◆

Table of Contents

INTRODUCTION

GUIDES

SUPPLEMENTARY MATERIAL

Course Scope

Biochemistry is a fascinating subject that explains how all life on Earth functions. Unfortunately, the details of biochemistry often cause students to get lost, and consequently, they don't get to see the big picture. Without connections to real life, the details simply become things to memorize for tests—and are soon forgotten.

These lectures were designed to help everyone learn and enjoy biochemistry using explanations of real-world biochemistry problems as tools to understand the subject—while still introducing college-level biochemistry and not dumbing it down. And without any need for memorization, you won't spend your time memorizing facts, giving much more time to listen to explanations and think through the concepts.

Biochemistry combines biology—the science of life—and chemistry, the science of molecules. Biochemistry, then, is the molecular basis of life. Every single thing that makes people alive is due to biological molecules and their interactions.

The first 24 lectures of this course focus on biological molecules and metabolism, the thousands of different biochemical reactions that are occurring in human cells. You'll start by learning some basics about water, the most abundant molecule in the body, in lecture 2. You'll then move on to amino acids, proteins, and enzymes in lectures 3 through 8. Along the way, you'll discover how the flexibility of proteins enables everything from speeding up reactions to meeting the body's needs for oxygen.

Lectures 9 and 10 introduce other major biomolecules: lipids and carbohydrates. You'll learn what metabolic energy is in lecture 11 and how much of it you go through. Then, in lectures 12 and 13, you'll discover how

cells break down dietary and stored carbs and fats to generate the energy your body needs. Lecture 14 will bring all of the cell's energy considerations together under one umbrella, and you'll learn how photosynthesis—the capture of solar energy by plants—is related to metabolic energy generation in humans.

Lectures 15 through 19 complete the tour of metabolism by examining how human cells build carbs, fats, amino acids, and cholesterol. In lectures 20 through 24, you apply your biochemical knowledge to fascinating topics that take you well beyond individual cells. These lectures discuss healthy eating and fad diets, the language of cellular communication, the incredible speed of neurotransmission, and the 5 manifestations of neurotransmission that give rise to your senses.

The last 12 lectures of the course delve into biological information—specifically, the specifications for building an organism and how that information is stored in molecules to be used, maintained, and transmitted to future generations. This is often called molecular biology. Whereas biochemistry focuses on biomolecular reactions, molecular biology covers the underlying cellular architecture that gives rise to the proteins that enable all of the biochemical reactions.

DNA's role in this process is dissected in lectures 25 through 29. Prepare to be surprised at how you have DNA that is different from both of your parents' and how little DNA would be needed to store all of the world's computer information. Lecture 30 illuminates the role of another information molecule, RNA, in making proteins and how it may have played additional roles in early life on Earth. The actual process of making proteins, and the many ways cells control when and how much of each protein to make, is the subject of lectures 31 and 32.

Beginning with lecture 33, you'll apply knowledge you've acquired to understand human genetic diseases. You'll also learn about some clever and promising approaches to fixing some of them. Lecture 34 covers the molecular basis of cancer and its treatment. The last 2 lectures (35 and

36) focus on exciting, modern applications of biochemistry knowledge—especially on medicine that is personalized and that works at the molecular level to cure disease.

You can skip around a bit, if you like, but you'll probably benefit more from following the topics sequentially in each of the 2 modules—biochemistry and molecular biology—as the lectures were created to build on lessons learned from preceding lectures. The biochemistry and molecular biology modules have also been structured to be independent so that watching the first 24 lectures isn't essential to understanding the last 12. And if you want to go deeper at any point in your biochemical journey, there is an extensive reading list in the Bibliography and 2 biochemistry textbooks you can download for free at www.davincipress.com/freeforall.html. ◆

Acknowledgments

The effort that went into putting these lectures together was enormous. I can take credit for what is said in front of the camera, but much more happened behind the scenes. Major credit for the scripts goes to Indira Rajagopal and Jay Tate. Video production, editing, and graphics were the product of a talented team led by Trish Golden, Kristen Westphal, and Trisa Barnhill. Finally, I'd like to acknowledge Christopher Mathews for giving me the opportunity to bring my teaching ideas to life, George Pearson for being an incredible mentor and friend, and Neal Gladstone for inspiring me to write lyrics and verse. ◆

01

BIOCHEMISTRY IS THE SCIENCE OF US

Biochemistry helps design life-saving treatments, improve the food supply, and answer questions in areas spanning from archaeology to criminal justice. In fact, there are few areas of science where biochemistry hasn't had some impact.

Biochemistry: A Young Science

Biochemistry is new, as the sciences go. Separately, the subjects of chemistry and biology have been around since the time of the ancient Greeks and Egyptians. But the idea of combining them is much more recent. About 200 years ago, there was no evidence that a science of biochemistry was even possible. Even 100 years ago, the word *biochemistry* was barely coming into use, and the field had scarcely begun.

Yet biochemistry has already unraveled chemical pathways and processes in cells that make us who we are. Personalized medicine is becoming possible thanks to disciplines biochemistry has spawned, including genomics and metabolomics.

Biochemistry is truly the science of us—a single grand story that is much bigger than all of humanity.

The earliest roots of biochemistry date from almost 2 centuries ago, when an experiment by Friedrich Wöhler accidentally created a compound now known as urea. It turned out to be identical to a crystal that appeared when he dried urine. This demonstration made it clear, for the first time, that ordinary chemistry had to be possible inside of cells.

As early investigators began to take apart the cell, they increasingly began to discover both the magic of biochemistry and its roots in traditional chemistry. Two discoveries at roughly the same time in the 1860s were inheritance and the existence of DNA. But the significance of these remained unrealized for decades.

The famous physicist Erwin Schrödinger wrote a seminal book for lay readers entitled *What Is Life?* based on lectures he had been giving in Dublin. This book set the theoretical framework for what we now take for granted—that everything we associate with life has its roots in molecules. Watson and Crick in 1953 pointed to Schrödinger's book as inspiration for their search to find the structure of DNA.

Everyone knew that the chemistry of reproduction had to be distinctive, but it took a long time to realize that what enabled cells to reproduce was their ability to store information, read that information, and reproduce that information. No chemistry anyone saw before had ever done that.

Classical biochemistry had been concerned with enzymes and other molecular reactions found inside and among cells. Now a new field called molecular biology came into being, with the master instructions that cells need to carry out their activities. The 2 fields became intertwined, and the 2 names for this combined subject are basically interchangeable.

The periodic table has more than 100 elements, and with only a few exceptions, biochemistry is not concerned with those individual atoms.

Molecules and Minerals

Biochemistry is mostly about molecules. The molecules of biochemistry are overwhelmingly built using primarily just 6 bonding elements: carbon, oxygen, nitrogen, and hydrogen, with supporting help from sulfur and phosphorus.

What matters for joining these atoms together is electrons. Chemical bonds always involve electrons—the smallest parts of their atoms by far. Electrons are negatively charged and located outside the nucleus of the atom. The number and kind of bonds an atom can make is due to the number of electrons it can share, release, or steal. Of those, the bonds that biochemistry cares about most involve electron sharing. It is these covalent bonds that stick us together.

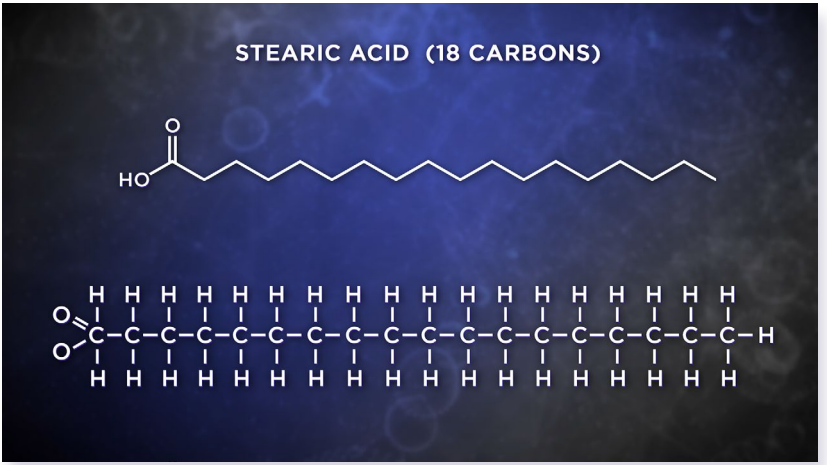
Cells are mostly water, which is just hydrogen covalently bonded with oxygen—hydrogen dioxide. The oxygen can make 2 bonds; each hydrogen can only make one bond. More importantly, they share electrons with each other. The other 4 elements that are of primary importance for making bonds in biochemistry also like to share.

The element whose atoms are arguably the most important for life is carbon. Carbon's importance is directly traced to its electrons. It has 4 electrons involved in reactions, and they all participate in sharing with other atoms. Sometimes the sharing is equal, while other times it's unequal, but carbon never gives up its electrons entirely or takes those of another atom.

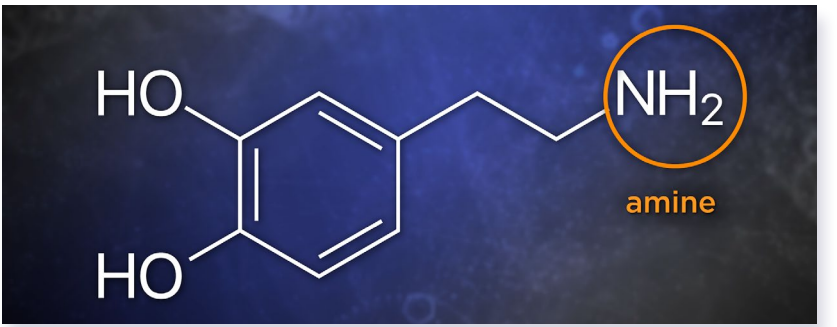
Carbon's ability to make 4 bonds also makes it central to the construction of large and complicated molecules. A carbohydrate molecule, for example, is a bunch of carbons that have been hydrated with a bunch of water. Biochemistry is all about large carbon-centered molecules and their relationships to the water of the cell.

Conventions for the Depiction of Molecules

In the simplified diagram shown above right, the carbon is assumed to be at the meeting point of angled lines depicting the bonds. Hydrogens attached to the carbons are also omitted entirely. These simplifications help a lot when showing bigger carbon compounds, including the fatty acid depicted below.



Besides carbon, hydrogen, and oxygen, the fourth most abundant element in our bodies—and the most abundant in our atmosphere—is nitrogen. Nitrogen can make 3 bonds with its electrons. When nitrogen bonds to hydrogen, as it commonly does, it forms an amine, which is related to ammonia. Amines also give their name to amino acids, which are the building blocks of proteins—the stars of biochemistry.



A fifth element that is important in many proteins is sulfur, which can make 2 bonds and is found in 2 of the amino acids.

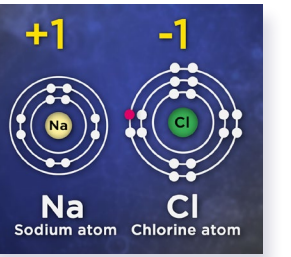
The sixth element that's important for building biomolecules is phosphorus. In cells, phosphorus is pretty much only seen when bound to 4 oxygen atoms to create a unit known as a phosphate. Phosphates are important for storing energy and for being on-off switches for proteins. Phosphates are also part of the backbone of the structure of DNA.

There are some other players in biochemistry. Many of these are atoms and molecules that are referred to as minerals. Eight minerals are most abundant in biochemistry: sodium, potassium, magnesium, calcium, iron, zinc, copper, and chloride. Minerals fall into a different category from the previous 6 elements because minerals do not usually share electrons when they form bonds. Chloride steals electrons; each of the other minerals is a supplier of electrons.

In the watery environment of the body, such molecules are broken up and become charged ions. Our bodies are not made of ions; instead, we make use of ions.

Stealing electrons involves gaining a negative electric charge, while supplying electrons means a loss of negative charge by the atom participating. Because atoms start out with a charge of zero, we can tell the number of electrons an atom has gained or lost by its charge.

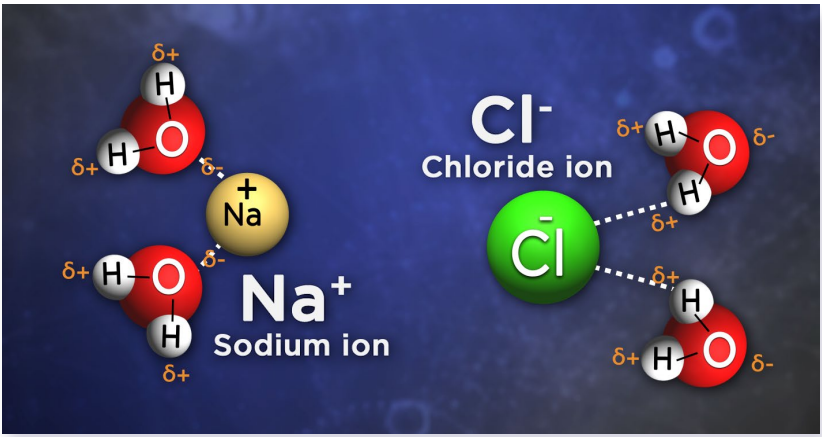
Sodium is a supplier of one electron, so it typically has a charge of +1, and chlorine steals electrons, so it has a charge of –1. Put them close together and they make NaCl, also known as table salt. But in the watery environment of our bodies, the ionic bonds holding the salt are broken, and we use the ions.



Iron and copper are 2 ions with some additional abilities. We find them in cells with 2 possible charges each. For iron, it can exist as Fe^{++} or Fe^{+++}, and copper can exist as Cu^{+} or Cu^{++}. The ability of iron and copper to flip between 2 states differing by one electron turns out to be very important for handling energy in cells.

Copper plays 2 important roles in the body. One involves energy generation: It helps in respiration by transferring electrons to oxygen to make water, and this is the reason you breathe. The second role of copper is in helping to protect cells from random damage actually caused by the oxygen you breathe.

The process of making an ion is called ionization. Ions are very important to cells. When minerals dissolve in water, they form ions and come apart. Sodium chloride splits into sodium ions and chloride ions when you dissolve it in water. Because water is 70% of the weight of cells, interactions with water in the cell are crucially important.



Ions are not building blocks of cells, but they perform critical tasks in and around the cells of our bodies. Individual atom ions, such as sodium and potassium, are tiny compared to proteins and other giant cell molecules with thousands of atoms. The tiny size of these ions allows cells to regulate their movement using nanoscopic protein channels specific for each one and allows them to move very rapidly, while their electrical charge gives them important roles in the electrical circuits of our nerve cells.

Proteins and Enzymes

Proteins are the workhorses of the cell. Proteins are polymers, which are long strings of molecules joined end to end. But there's more than one way to build a polymer.

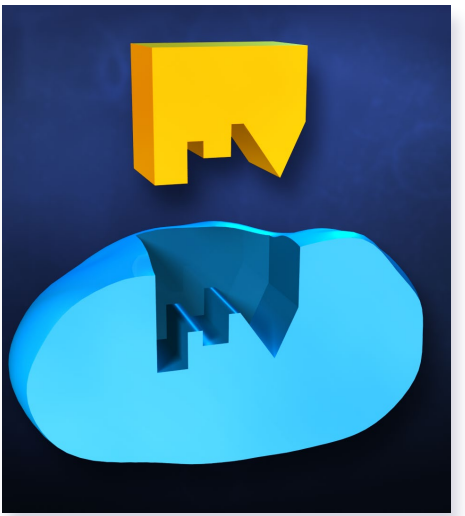
Nonliving polymers, such as plastics, tend to repeat a single building block over and over. Such repetition provides uniformity: A sheet of plastic will appear the same no matter what part of it you examine.

By contrast, protein polymers have much more individuality. Instead of repeating a single building block monotonously, proteins have at their disposal, and often use, 20 different building blocks—all amino acids. Each of the thousands of proteins found in cells has a unique order of building blocks joined to each other, and that sequence of building blocks determines the shape of the protein and what it does.

Many proteins are quite large, often with more than 1000 amino acids in a single protein molecule. That's why they're called macromolecules.

Some proteins speed up chemical reactions occurring inside of cells. A huge family of proteins that is responsible for speeding up all sorts of reactions is called enzymes. Nonprotein molecules that speed up reactions are called catalysts.

What makes enzymes almost magical is their incredible speed. Cells rely on enormously sped-up reactions. In the absence of enzymes, some reactions can take millions of years to occur, so the speed of enzymes is crucial.



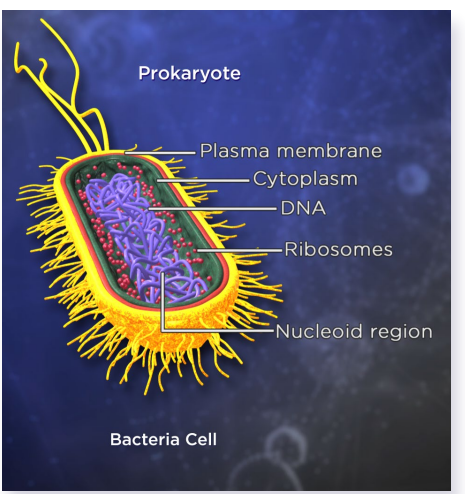
Enzymes are also very particular about their partners. This is called high specificity. The specificity of an enzyme is rooted in the very specific arrangement of the building blocks in its polymer. Change the order of amino acids and you change the shape of the enzyme.

The shape of molecules is one of the most important considerations in biochemistry. A basic principle is that structure determines function. Enzymes have highly specific shapes, and those characteristic shapes only allow molecules with corresponding shapes to fit into them. Fitting into enzymes in a very specific way is critical because that's where enzymes catalyze reactions. If a molecule can't fit into an enzyme, its reaction will merely grind along at the speed of nonlife.

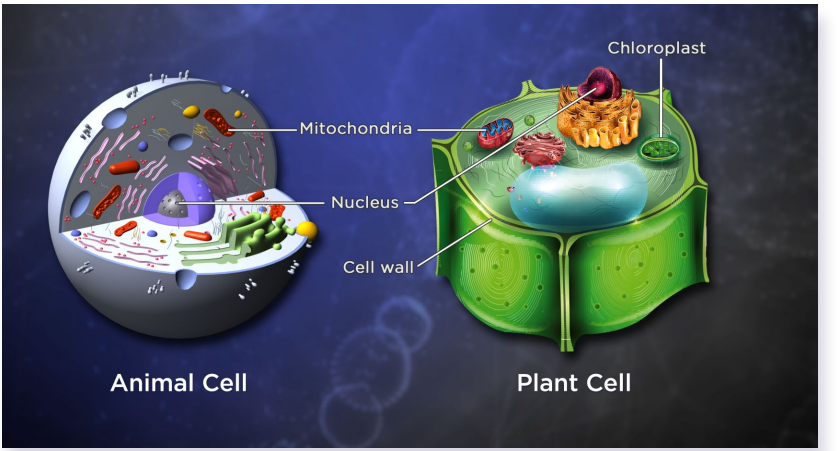
The vast majority of biomolecules found in a cell get made in reactions catalyzed by enzymes in that cell. Because each enzyme catalyzes reactions quite specifically, on only one or a handful of different molecules, any given cell must make thousands of enzymes to catalyze the thousands of different types of molecules made in that one cell.

Cells as Houses for Reactions

Cells are the fundamental unit of life. Overall, you could think of cells as houses for molecular reactions. Like houses, cells have surrounding walls in the form of a membrane. Some cellular houses are more like tents, with a simple design, small size, and only a single room for the reactions to live in. Bacteria cells have single rooms. They are also called prokaryotes. With no internal walls, bacterial houses keep all their possessions in the same room.



Most other cells, such as the ones in our body, are more like houses subdivided into rooms, each of which has a function. Rooms in cells are called organelles. There are several, including a nucleus, where DNA is stored, and mitochondria, where energy is made.



Cells with the organizational scheme of a house are called eukaryotes. Most eukaryotes are multicellular—like cities containing many houses. However, a few eukaryotes, such as yeast and amoebas, happily maintain the rural life of a little house on the prairie.

One of the remarkable things about the biochemistry of cells is this: All living things are remarkably similar. The molecules they contain, the reactions their molecules undergo, and the ways they store information inside themselves are similar. Whether it's the breakdown and synthesis of fat, the storing of chemical energy, or the synthesis of proteins, all cells take the same basic approach.

Drug makers work hard to exploit minor differences between human and bacteria cells when designing antibacterial drugs for infections.

This similarity provides a great simplification for learning biochemistry. When you learn the biochemistry for one organism, you've already learned a great deal about every other organism. These deep similarities make it possible for us to use bacteria, yeast, and animals as biological factories to make human proteins, such as insulin, for medicinal use.

Don't focus on the terminology used in these lectures. This isn't a course about names, equations, memorizing, or any of the standard things people stereotypically associate with science courses. This is a one-of-a-kind course where you can enjoy biochemistry while you learn it.

READINGS

Gray, *Molecules*.
———, *Reactions*.
Jonsson, et al., "Essential Chemistry for Biochemists."
Pross, *What Is Life?*

QUESTIONS

1 The 6 elements that are used to construct biomolecules are not the 6 most abundant elements on Earth. Discuss why abundance is not an important consideration for biomolecules. What properties do these elements have that allow them to better perform the roles of making molecules that are important for life?

2 The great scientist Louis Pasteur mistakenly believed that the process of fermentation was a biological process, not one that was possible using ordinary chemistry. He was wrong, of course. One type of protein bridged the biological/chemistry gap and allowed reactions to occur both inside of cells and, when removed, outside of them. What are these proteins called, and how do they enable what seemed impossible to Pasteur?

CLICK HERE TO SEE THE ANSWERS.

02 WHY WATER IS ESSENTIAL FOR LIFE

Water is uniquely valuable because it has some special and peculiar properties that make it different from all other liquids in ways that are crucial to the processes that support life.

Water's Electronegativity

The key to the behavior of water is in a property of atoms called electronegativity, which is a measure of the affinity of an atom's nucleus for its outermost electrons. You can think of it as how strongly the nucleus of the atom holds on to those electrons. Atoms whose nuclei hold on tightly to outer electrons have high electronegativity values.

> Life on Earth almost certainly began in water, and liquid water seems to be an irreplaceable requirement for all life. About 2/3 of Earth's surface is covered with water, and about 2/3 the weight of a human being is water.

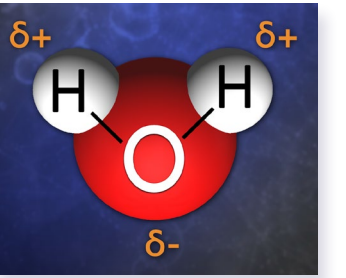
Electronegativities of elements vary considerably. Fluorine and oxygen have the highest electronegativities. As you move from right to left along the Periodic Table, the electronegativity values tend to decrease: Elements on the left side of the table, such as sodium and potassium, have very low electronegativity values. The exception is hydrogen, but its single electron is also a special case.

A molecule of water is made up of an oxygen atom and 2 hydrogen atoms. Oxygen has a very high electronegativity value compared to hydrogen, which means that the nucleus of the oxygen atom "likes" electrons more than the nucleus of each hydrogen atom does. So, although oxygen can make a bond to "share" electrons with hydrogen, the oxygen does not share equally.

You could think of oxygen as an electron bully, keeping shared electrons in each bond closer to itself and farther away from hydrogens. Although this setup is still a covalent bond, in that the electrons are shared, the oxygen holds the bond electrons closer, so it acquires a partial negative charge, written δ^-. And because hydrogen is farther away from the bond electron, the hydrogen gets a partial positive charge, written δ^+. It is this nonuniform charge distribution that gives water its very interesting properties.

Water's Polarity

When you look closely at the structure of a water molecule, the shape is a relaxed V. The bottom of the V, where the oxygen is, will be slightly negative, and the tops of the V, where the hydrogens are, will be slightly positive. A molecule with an uneven distribution of charge like this is described as being polar.



Each water molecule is in close proximity to several other water molecules, each with the same partial charge distributions. Opposite charges attract, so partial positive charges on hydrogens of one water molecule are attracted to the partial negative charge on each oxygen of neighboring water molecules. Interactions arising from these partial charges are called hydrogen bonds.

Although hydrogen bonds are possible between atoms of many molecules, they mostly involve oxygen, nitrogen, sulfur, and hydrogen.

Two of the most important structures in our cells rely on hydrogen bonds for stability:

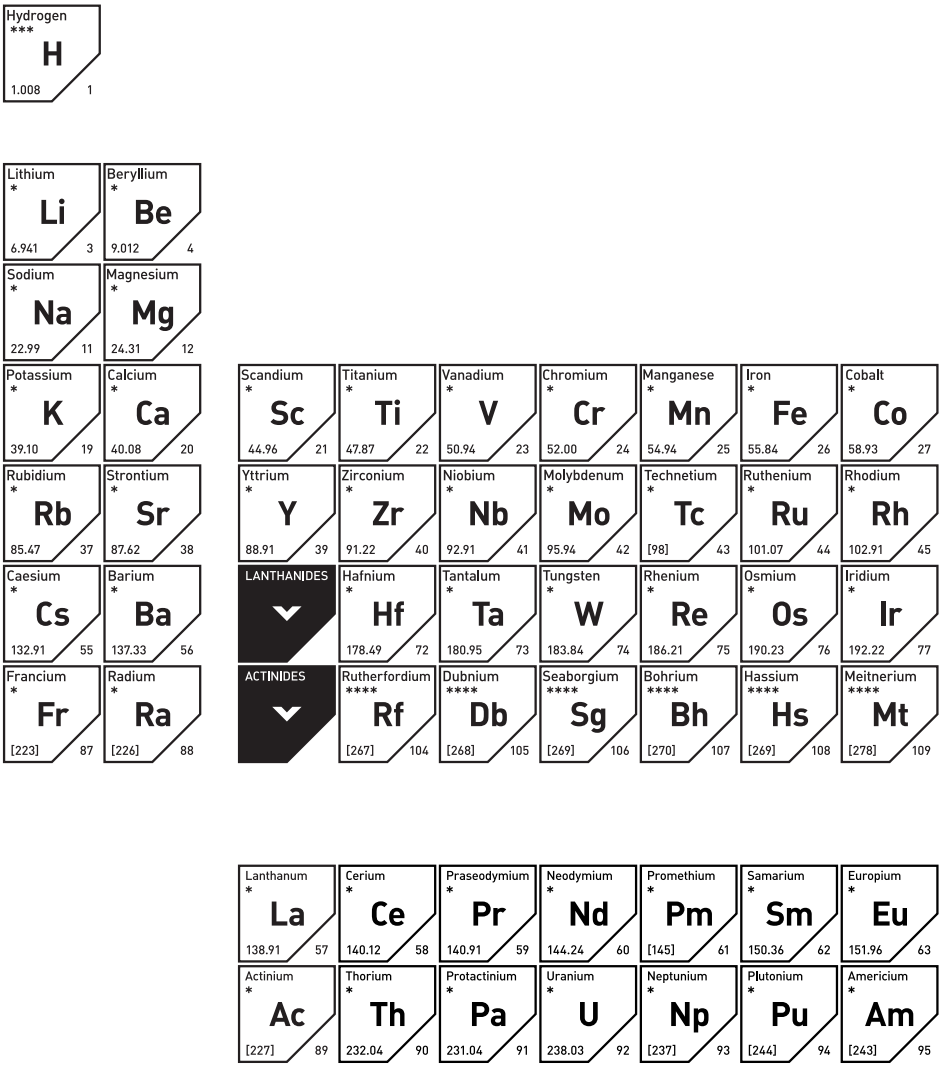
- ◇ **The complex folded structure of proteins is stabilized by hydrogen bonds.**
- ◇ **Double strands of DNA are held together by hydrogen bonds.**

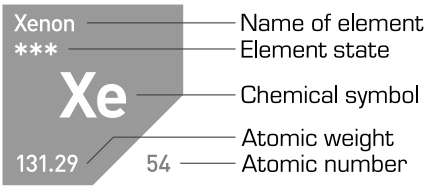
Hydrogen bonds are why water remains liquid at room temperature; these bonds are how water molecules attract and hold on to each other. Although such bonds are not permanent, each water molecule in liquid water is constantly making and breaking links with its neighbors in a sort of group dance, swapping partners but staying connected. Breaking all these bonds at the same time to turn water into steam takes much more heat than would be needed for molecules not held together by such bonds.

Unlike most liquids, as water freezes, it expands, thus reducing its density. This is why ice floats in liquid water. And the fact that water is a liquid across the wide range of temperatures considered to be normal is unusual, too.

Conversely, when temperatures drop and thermal motion is minimal, H_2O molecules take a tiny step back and each water molecule hydrogen-bonds to 4 neighboring H_2Os. This creates the highly ordered structure in an ice crystal. And because frozen water is in this crystal form, ice is less dense than water.

The Periodic Table of Elements





* Solid
** Liquid
*** Gas
**** Unknown

						Helium *** He 4.003 2
		Boron * B 10.81 5	Carbon * C 12.01 6	Nitrogen *** N 14.01 7	Oxygen *** O 16.00 8	Fluorine *** F 19.00 9	Neon *** Ne 20.18 10
		Aluminium * Al 26.98 13	Silicon * Si 28.09 14	Phosphorus * P 30.97 15	Sulfur * S 32.07 16	Chlorine *** Cl 35.45 17	Argon *** Ar 39.95 18
Nickel * Ni 58.69 28	Copper * Cu 63.55 29	Zinc * Zn 65.39 30	Gallium * Ga 69.72 31	Germanium * Ge 72.63 32	Arsenic * As 74.92 33	Selenium * Se 78.96 34	Bromine ** Br 79.90 35	Krypton *** Kr 83.80 36
Palladium * Pd 106.42 46	Silver * Ag 107.87 47	Cadmium * Cd 112.41 48	Indium * In 114.82 49	Tin * Sn 118.71 50	Antimony * Sb 121.76 51	Tellurium * Te 127.60 52	Iodine * I 126.90 53	Xenon *** Xe 131.29 54
Platinum * Pt 195.08 78	Gold * Au 196.97 79	Mercury ** Hg 200.59 80	Thallium * Tl 204.38 81	Lead * Pb 207.2 82	Bismuth * Bi 208.98 83	Polonium * [209] 84	Astatine * [210] 85	Radon *** [222] 86
Darmstadtium **** Ds [281] 110	Roentgenium **** Rg [281] 111	Copernicium **** Cn [285] 112	Ununtrium **** Uut [286] 113	Flerovium **** Fl [289] 114	Ununpentium **** Uup [289] 115	Livermorium **** Lv [293] 116	Ununseptium **** Uus [294] 117	Ununoctium **** Uuo [294] 118

Gadolinium * Gd 157.25 64	Terbium * Tb 158.93 65	Dysprosium * Dy 162.50 66	Holmium * Ho 164.93 67	Erbium * Er 167.26 68	Thulium * Tm 168.93 69	Ytterbium * Yb 173.04 70	Lutetium * Lu 174.97 71
Curium * Cm [247] 96	Berkelium * Bk [247] 97	Californium * Cf [251] 98	Einsteinium * Es [252] 99	Fermium * Fm [257] 100	Mendelevium * Md [258] 101	Nobelium * No [259] 102	Lawrencium * Lr [262] 103

Hydrogen bonds can also form within and between molecules. The same electronegativity difference that gives rise to hydrogen-bonding in water also allows amino acids to interact within proteins.

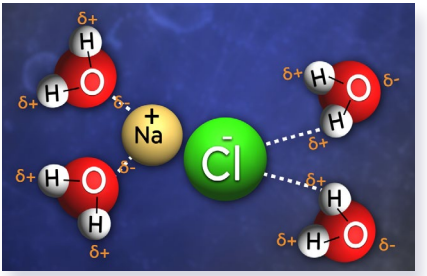
Whenever large electronegativity differences exist between hydrogen and an atom, hydrogen bonds are possible. For example, hydrogen has much weaker electronegativity than nitrogen and forms hydrogen bonds when linked together. Carbon, on the other hand, is much more similar to hydrogen, so it does not form hydrogen bonds to any significant extent.

Every amino acid in every protein contains nitrogen-hydrogen bonds, and carbon-oxygen double bonds and hydrogen bonds can form between the hydrogens of the nitrogen-hydrogen bonds and the oxygens of the carbon-oxygen double bonds. Similarly, every nucleotide in every DNA or RNA molecule contains these bonds.

And because most of the body is water, the hydrogen bonds within these biomolecules are constantly being invited to dance with water. In other words, because water is held together by hydrogen bonds, water is an effective medium and solvent for the molecules important for life, both inside cells and in the fluids that surround them.

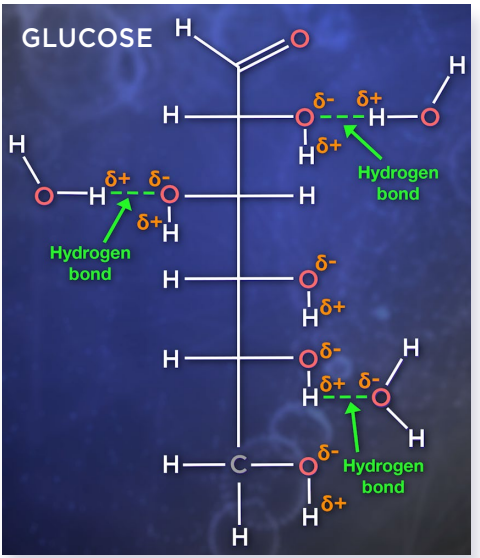
A solvent is a liquid that dissolves something. Ocean water is full of dissolved salts. Water uses its partial charges to dissolve compounds.

Table salt, or sodium chloride, can be ionized in water. Sodium's very low affinity and chlorine's very high affinity for electrons means that sodium essentially gives up an electron to chlorine when they bond. Sodium then has a net charge that's fully positive and is a sodium ion.



On the other side of the molecule, chlorine gains sodium's electron and is fully negatively charged, forming a chloride ion. Sodium's positive charge is strongly attracted to chloride's negative charge, creating an ionic bond.

By contrast, many organic compounds, such as sugar, dissolve in water because the partial charges of their own hydrogen bonds can mingle directly with the hydrogen bonds of water. There is no need to ionize. Water's partial charges can dance with the partial charges of glucose, a typical sugar, to pull individual sugar molecules into solution.

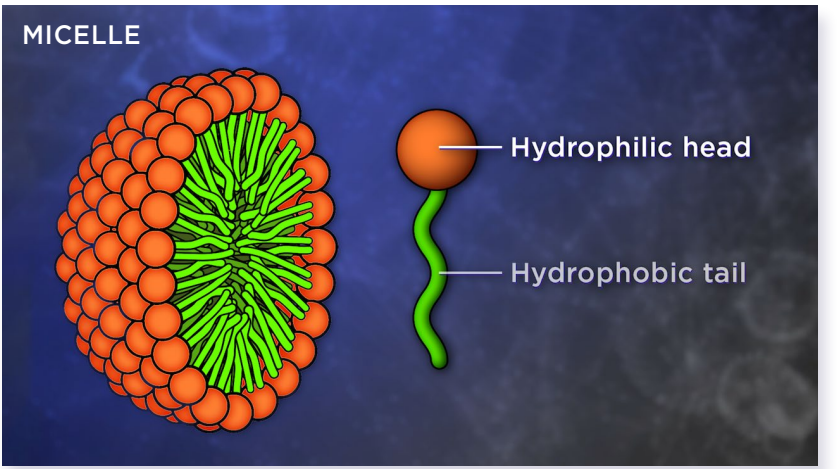


But other molecules, such as oil and fat, do not dissolve or mix in water. This is because they have long chains of carbon-hydrogen bonds that do not ionize or form hydrogen bonds. The electronegativities of hydrogen and carbon are similar enough that electrons in these bonds are pretty equally shared. The charge distribution is effectively uniform, and the bond is effectively nonpolar. So, fats can neither ionize nor readily form hydrogen bonds; therefore, water can't pull fat molecules apart from each other.

Molecules that do not dissolve in water are hydrophobic, meaning "water-fearing." Molecules that do dissolve in water are hydrophilic, or "water-loving." Molecules that are ambivalent about water are amphiphilic, which means they have an intermediate level of solubility in water compared to hydrophilic and hydrophobic compounds.

Like a clique of middle schoolers huddling with each other and rejecting those who are different, the hydrophobic regions of amphiphilic molecules arrange themselves to be near each other and away from water. Meanwhile, their hydrophilic ends happily associate with water molecules. All this combines to form a spherical bubble-like structure called a micelle.

This chemical property of nonpolar parts of molecules to associate with each other to the exclusion of water is called the hydrophobic effect, and it is important for forming cellular membranes.



Soap is an example of an amphiphilic compound. It is partly hydrophilic and partly hydrophobic. The hydrophilic end interacts with water readily; the hydrophobic end avoids interacting with water. Soap's micellar structure is critical to how it removes grease when you wash your hands or do the dishes.

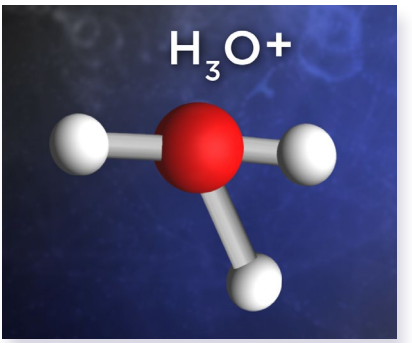
Water's Ability to Ionize

Another property of water that is important in biological chemistry is water's own ability to ionize, or break into ions. This happens only to a very small number of water molecules, but it does occur and is important.

Only about one water molecule in more than 500 million will ionize—meaning that this is less likely to happen than a single Powerball ticket winning the big prize.

When water ionizes, it breaks into a positively charged hydrogen ion (H^+), called a proton, and a negatively charged hydroxide ion (OH^-). The process results in the formation of ions because oxygen's higher electronegativity means it takes the unequally shared electron entirely for itself when the split occurs, leaving the positively charged proton with no electron. Under these conditions, water has gone from being H_2O, with no net charge, to H^+ and OH^-, each of which has a charge.

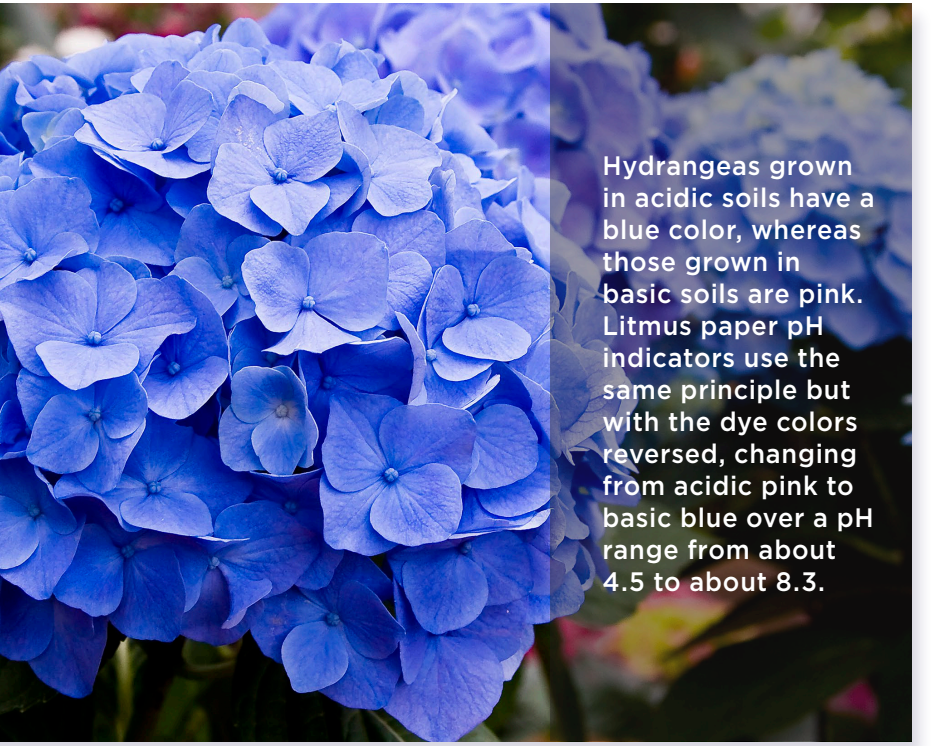
A fundamental principle in chemistry is that any substance that can donate a proton is called an acid, while one that can accept a proton is called a base. In addition to its other peculiarities, water acts as both an acid and a base: When it ionizes, it releases a proton, acting as an acid, but it simultaneously releases a hydroxide ion, which can accept a proton to act as a base.



Water is not a very strong acid/base, because only a miniscule amount of it ionizes; most of it remains intact as water molecules. The proton joins to a water molecule, making H_3O^+. In pure water, the concentration of H_3O^+ and hydroxide ions is equal, with each being 1×10^{-7} molar (M), so the solution remains neutral overall.

The fact that the concentration of protons (hydrogen ions) in pure water is 1×10^{-7} M is the basis for the pH system to measure the acidity of a solution, where the *H* in *pH* stands for the hydrogen proton. pH is defined as the negative logarithm of the proton concentration. For pure water, the concentration is 10^{-7} M.

The negative logarithm of 10^{-7} is 7, so a neutral solution, which is neither acidic nor basic, is said to have a pH of 7. If the hydrogen ion concentration is increased to 10^{-6} M by the addition of an acid, the pH goes down to 6; each change of one pH unit reflects a tenfold increase in proton concentration.



Hydrangeas grown in acidic soils have a blue color, whereas those grown in basic soils are pink. Litmus paper pH indicators use the same principle but with the dye colors reversed, changing from acidic pink to basic blue over a pH range from about 4.5 to about 8.3.

All cellular processes involve water, so the pH of cellular and extracellular fluids plays an enormous role in biological chemistry. The pH of the cytoplasm of cells has to remain within a narrow pH range for cellular reactions to occur appropriately.

Likewise, the pH of our bloodstream is confined to a very narrow range of values, almost always between 7.35 and 7.45. If the blood pH drops below 6.8 or rises above 7.8, bonds holding together critical proteins can get destroyed, and death results. Fortunately, there are buffering mechanisms in our bodies to keep pH within that narrow range and avoid catastrophic consequences.

Weak acids act as buffering agents to keep pH constant. Different weak acids have different pHs at which ½ of their carboxyl groups (COOH) are protonated, and this specific pH for a given acid is called its pKa. When the pH around a weak acid is perturbed, there's a range of pH values, on either side of the pKa, within which there is a stubborn reluctance to change, called buffering. Buffering, or resistance to change in pH, is enormously important in living organisms, which have to maintain the pH of cells and extracellular fluids within very narrow ranges.

There is no evidence that diets—such as the alkaline diet, which shuns meats, cheeses, and grains in favor of fruits and vegetables—can do anything to the pH of your blood or cells.

In fact, if a fad diet did change blood pH, it almost certainly would be detrimental, not beneficial. Your body's buffering systems already keep your pH where it needs to be.

READING

Ball, *Water*.

QUESTIONS

1 Water has 2 important considerations for giving it the properties it has. One is the unequal sharing of electrons. The other is a structural feature. What is the feature, and why might it be important for water's properties?

2 Ionic substances like sodium chloride readily dissolve in water. Ionization is not required, though, to dissolve substances in water. Some nonionizing biomolecules like sugars dissolve readily in water, whereas other nonionizing biomolecules like fats and cholesterol do not. What property do water-soluble nonionizing biomolecules have that water-insoluble biomolecules lack?

3 The Henderson-Hasselbalch equation shows the relationship between the pH, the pKa, and the ratio of salt to acid. Given that the log of a number that is greater than 1 is a positive number, the log of 1 is 0, and the log of a number that is less than 1 is a negative number, predict whether a solution will have more salt than acid, less salt than acid, or equal amounts of salt and acid under the following conditions:

a $pH < pKa$

b $pH = pKa$

c $pH > pKa$

CLICK HERE TO SEE THE ANSWERS.

03

AMINO ACIDS: 20 BUILDING BLOCKS OF LIFE

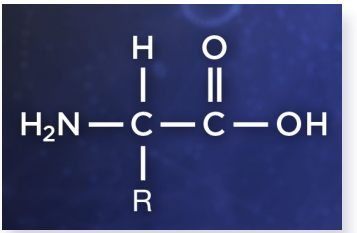
Besides water, proteins are the most abundant molecules in all known forms of life. Proteins are the most diverse class of biological molecules, making up everything from enzymes and hormones to antibodies.

About 70% of your body weight is water, and about 17% is protein.

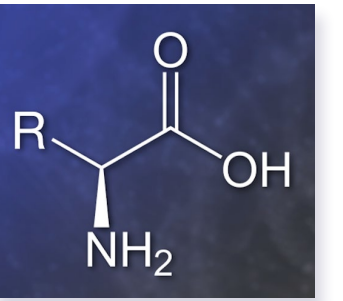
R Groups and Chiral Compounds

There are 20 kinds of amino acids that are strung together in proteins. Because amino acids are the building blocks of all proteins, all cells need amino acids.

Amino acids found in proteins all have the same simple core structure: a central carbon, an amino group with NH_2, an acid group with COOH, and a lone hydrogen opposite the amino group.



=



The central carbon is the alpha carbon. Attached to it is an alpha amino group, an alpha carboxyl group, a hydrogen, and a fourth group that varies in structure between amino acids called an R group.

The R group is the only part of an amino acid's structure that varies from one to the other; the other parts of the structure are common to all of them.

R groups are aliphatic when they contain only carbons and hydrogens, which are so similar in electronegativity that they are nonpolar—meaning they are hydrophobic, or can't make hydrogen bonds with water and therefore avoid it.

Other R groups contain other atoms and can ionize or make hydrogen bonds, so these are hydrophilic—they like water.

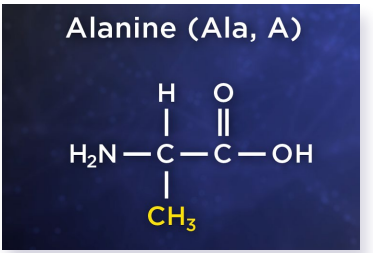
About ½ of the amino acids are nonpolar and hydrophobic. These are the inward-turning introverts of a folded protein molecule; when proteins fold, these amino acids try to hide inside the folded molecule. The other ½ are the polar, water-loving extroverts.

More than 300 amino acids exist in nature, but only 20 are coded for in our DNA.

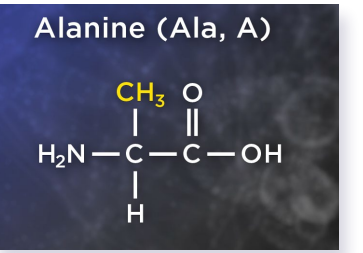
Amino acids are the building blocks of proteins. The bonds holding proteins together are known as peptide bonds, and they form by joining the carboxyl group (COOH) of one amino acid to the amino group (NH_2) of the next.

All amino acids except glycine can exist in 2 mirror image forms. These differ in the arrangement of the 4 groups around the alpha carbon. It's like right and left hands that cannot be superimposed on each other. The 2 forms are called stereoisomers: the L form and the D form. Glycine is different from all the other amino acids in having an H across from another H, so there's only one form of glycine.

L FORM



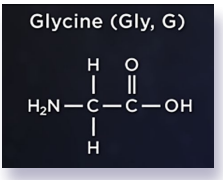
D FORM

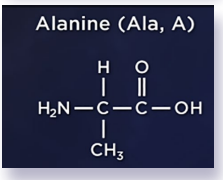


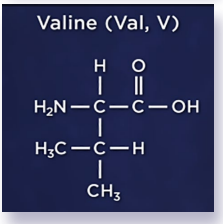
Molecules with handedness are chiral. All chiral compounds, such as amino acids when made apart from cells, have a 50/50 mixture of the D and L forms. However, amino acids made in cells for use in protein synthesis are almost completely in the L form. This is because the enzymes synthesizing these amino acids have specific 3-D structures that arrange reactants such that only one form can be made.

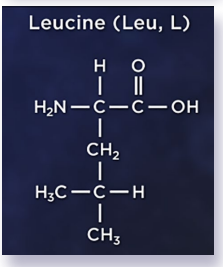
We can tell if a mixture of molecules was made by an enzyme or by nonbiological chemistry simply by analyzing its chirality. If nonenzyme chemistry produced the molecules, they'll have equal amounts of D and L forms because that's what ordinary chemical reactions in test tubes yield. Molecules made by enzymes, though, will have one form or the other.

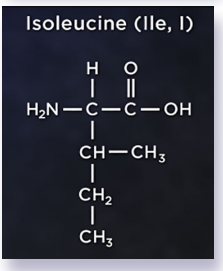
Nonpolar, Hydrophobic Amino Acids











Hydrophobic amino acids, which have R groups that mostly contain carbons and hydrogens, include glycine, alanine, valine, leucine, and isoleucine. The degree of hydrophobicity increases steadily from glycine to isoleucine as the R groups increase in size and complexity.

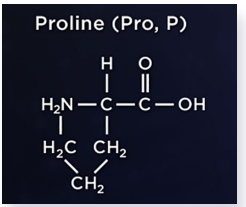
Glycine is the simplest amino acid, with a lone hydrogen as its R group. Despite its simplicity, glycine has important functions in proteins—especially in collagen, the most abundant protein in the human body. Glycine also acts as a signaling molecule for transmitting nerve signals in the spinal cord, brainstem, and retina.

The second simplest amino acid is alanine, which has a CH_3 methyl group for its R group. It is one of the amino acids that is sometimes seen in the D form in cells. One place D-alanine is found is in the cell wall of bacteria, where it is linked to other amino acids to form supports for the cell wall. The integrity of the cell wall is important for all cells; mess with it and a cell dies.

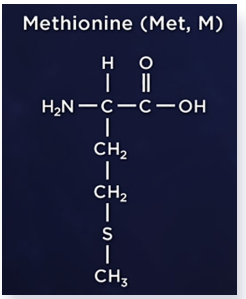
The amino acids valine, leucine, and isoleucine are sometimes referred to as the branched-chain amino acids (BCAAs). These amino acids have R groups with longer chains of carbon and hydrogens, making them more hydrophobic than glycine or alanine because hydrophobic properties are somewhat additive in nature—longer means more. Like the other amino acids, they are important for making proteins.

Unfortunately, there is a widespread belief that BCAAs are especially effective in helping build muscle, which leads some people to load up on BCAA supplements. But that could actually be doing harm, because BCAAs are present at higher levels in the blood of people with insulin resistance associated with type 2 diabetes.

Two other amino acids, proline and methionine, sometimes get grouped with the hydrophobic amino acids because they share characteristics with them.

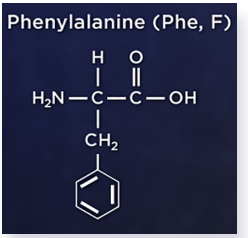


Proline has a formula that is almost like valine, but instead of making a V, its R group loops around from the alpha carbon to bond to the alpha amino group. The resulting ring structure is unique among the amino acids because it's the only one that connects directly to the alpha amino group.

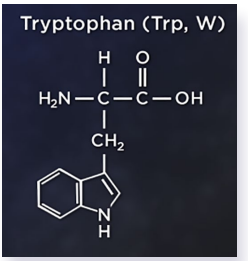


Methionine's chemical oddity is that it is one of 2 amino acids that contains sulfur. Sulfur's electronegativity is similar enough to carbon that there is fairly even sharing of electrons, meaning that sulfur bonded to 2 carbons in methionine is nonpolar and the amino acid is hydrophobic. Methionine is noteworthy because the cellular machinery for making proteins uses it as the first amino acid for making almost every protein on Earth.

A subgroup of the hydrophobic amino acids is the aromatic amino acids, named for the large and quite stable aromatic ring structures in their side chains.



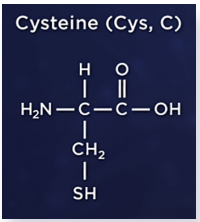
Phenylalanine, which is needed in the diet, is a phenyl ring of 6 carbons attached to an alanine. The R group is hydrophobic. If an OH hydroxyl is attached to the far side of the phenyl ring of phenylalanine, then it's tyrosine, which is just a hydroxyl-containing version of phenylalanine. It exhibits some properties of hydrophilic amino acids, but the ring makes it also hydrophobic.

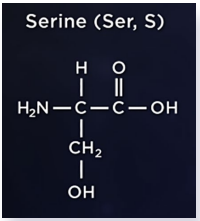


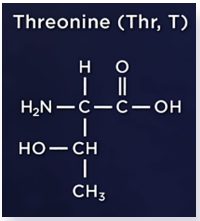
The largest hydrophobic amino acid is tryptophan, which has an R group of 9 carbons and 1 nitrogen in a structure known as an indole ring. Tryptophan is a starting point for the synthesis of many important biomolecules, including the vitamin niacin and the plant hormones known as auxins. Tryptophan is also a component of the neurotransmitter serotonin, which converts into the sleep-related hormone melatonin.

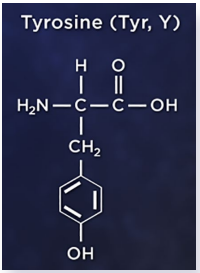
It's an urban legend that tryptophan in Thanksgiving turkey is why we get sleepy after that annual binge. If tryptophan were the cause, you'd be just as likely to fall asleep after a hamburger—beef has as much or more tryptophan than turkey.

Polar, Hydrophilic Amino Acids









The first polar amino acid is cysteine, which has sulfur, but unlike methionine, cysteine's sulfur is in the form of an SH group, known as sulfhydryl. The SH group of cysteine is reactive, because the H^+ proton can readily be lost near physiological pH of 7.4, and loss of a proton leaves the side chain with an S that is negatively charged. The electrons in the sulfurs are lost in the reaction that creates the disulfide bond between them. It is not uncommon for a protein to have multiple cysteines.

Similar to the SH group of cysteine are the OH hydroxyl groups in serine, threonine, and tyrosine. Unlike the SH sulfhydryl group on cysteines, the hydrogen on the OH of these amino acids doesn't come off readily at physiological pH. Unlike the OH groups found in COOH carboxyl groups, the OH groups of these R groups normally will not form ions. Instead, the OH stays together, remains uncharged, and happily hydrogen-bonds with water.

The simplest of these water-friendly amino acids is serine, which has only a single carbon in its R chain, just like alanine, plus OH at the end. Serine plays a distinctive role as a signaling molecule in the brain thanks to D-serine, a rare D-amino acid in our bodies. Meanwhile, ordinary L-serine is important for the synthesis of numerous cellular molecules, including proteins, membrane components, and nucleotides for making DNA and RNA.

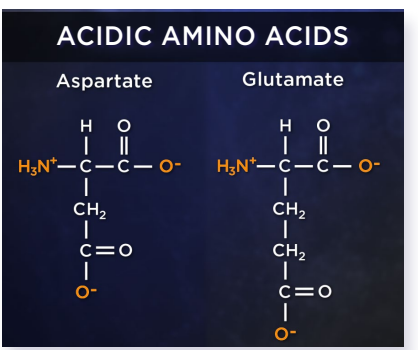
Threonine is similar to serine in many respects, and both function in the active sites of many enzymes.

The largest hydroxyl-containing amino acid, tyrosine, is the starting point for the synthesis of many biologically important molecules, including the melanin pigments that color our hair, skin, and eyes; thyroid hormones; adrenaline; and the nerve-signaling molecule dopamine, made in the reward circuits of our brains.

All 3 of the hydroxyl amino acids in proteins can have their hydroxyl groups modified. Enzymes called kinases catalyze replacement of the hydrogen on hydroxyl groups with a phosphate, a modification that is hugely important for protein regulation.

Phosphorylation can act as a switch, where toggling between the phosphorylated and unphosphorylated state determines whether a protein is active or not. The most common targets for phosphorylation are the hydroxyls of serine and threonine. Tyrosine phosphorylation also occurs in cells for signal relaying.

At physiological pH, 2 of the amino acids are notable for being acidic, meaning their R groups lose a proton, while 3 of the amino acids are basic, meaning their R groups gain a proton. All of these amino acids have charged R groups at the pH found in cells. The acidic amino acids contain carboxyls in their R groups—aspartic acid and glutamic acid. Ionized forms of these are called aspartate and glutamate.



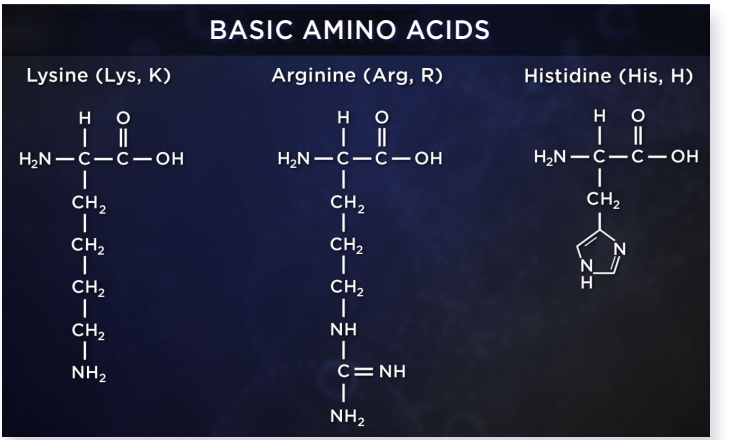
Glutamate famously affects what we taste. Our tongues have receptors that are sensitive to free glutamate, which is responsible for the savory, or umami, taste of some cheeses and soy sauce. The sodium salt of glutamate, known as monosodium glutamate (MSG), is often used for enhancing this flavor.

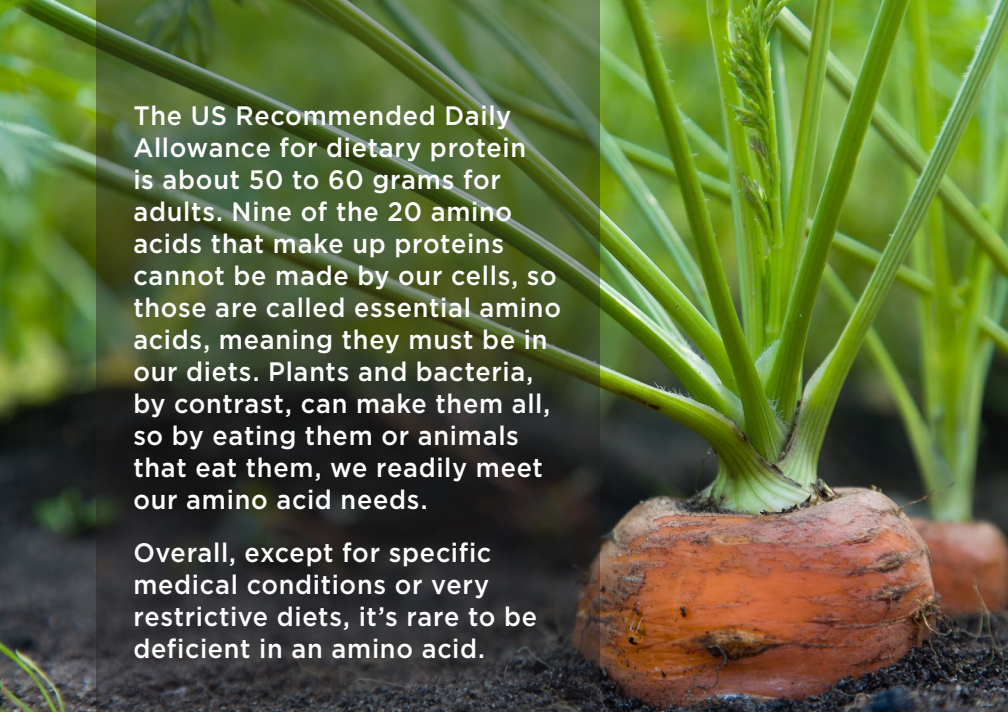
The pKa values of the R-group carboxyls of acidic amino acids are around 4, which means their protons are gone at the physiological pH of 7.4, leaving a negatively charged carboxyl group. The R groups of these 2 amino acids are identical, except the aspartate chain is smaller, while the glutamate side chain has one extra CH_2 group.

Glutamic acid got its name from being discovered in gluten, the wheat protein. Aspartic acid was first obtained from a related amino acid called asparagine, which was in asparagus.

Both aspartate and glutamate help cells use protein as an energy source when supplies of sugar run low or when a person goes on a low-carbohydrate diet. In cells, aspartate and glutamate are important in managing ammonia (NH_3), a toxic by-product of metabolism.

The 3 basic amino acids—lysine, histidine, and arginine—are called basic because their R groups accept protons at physiological pH, giving them positively charged R groups.





Lysine's R group is a long, flexible chain with a single positively charged nitrogen in the amine at the end. Similar, but slightly longer, is arginine, with 3 nitrogens near the end of the side chain. Arginine plays an important role in the urea cycle as the source of urea. Arginine is also a source of a cellular signaling molecule called nitric oxide, which stimulates the relaxing of blood vessel walls and improves blood flow.

Histidine is important in many enzymes and in the blood proteins myoglobin and hemoglobin. It is also the most chemically interesting of the basic-charged amino acids. Histidine's ring contains 2 nitrogen atoms that share a proton when the pH is sufficiently low, giving the group as a whole its positive charge. Histidine's ring-shaped R group is the basic amino acid with the lowest pKa of 6, meaning at physiological pH of 7.4, the proton is gone. Losing a proton means histidine's ring is mostly uncharged. But reductions in pH can drastically increase the percentage of histidine rings with the proton on.

READINGS

Gutiérrez-Preciado, Romero, and Peimbert, "An Evolutionary Perspective on Amino Acids."

QUESTIONS

1 Several chemical and physical properties of R groups of amino acids have effects on proteins that contain them. Describe these considerations and how they might affect protein structure and function. Include examples for each.

2 Phosphates are acid molecules, and they react primarily with the hydroxyls of serine, threonine, and tyrosine in cells. Based on the types of bonds described in the previous lecture, what kind of bond does phosphate make with these amino acids?

3 Histidine is an amino acid that some people might be tempted to try to reduce intake of with a certain health concern, but it would not be a good idea. What is the health concern, and why might it not be a good idea to try to reduce intake of it?

CLICK HERE TO SEE THE ANSWERS.

04 FROM PEPTIDE BONDS TO PROTEIN STRUCTURE

Amino acids are linked together by peptide bonds. When molecules have many peptide bonds, they are called polypeptides. Chains of amino acids take on unique structures that allow them to carry out a wide variety of specific functions.

The Amino Acid Alphabet

The English language has 26 letters in its alphabet, and we can construct every word in the language using different combinations of the letters. Likewise, the 20 amino acids can be used to create many different protein "words." When we write words, we know that which letters we use and the order of the letters matters. Likewise, which amino acids are used to build a protein and what order they are in distinguishes one polypeptide from another.

A polypeptide will have at least a "sentence" worth of "letters," and even a chain of 100 amino acids is, by cellular standards, quite small! That's the protein equivalent of a tweet worth of letters.

Protein letters differ from letters in a sentence, though. Letters in proteins are physically linked to each other by the formation of peptide bonds between the alpha carboxyl group of one amino acid and the alpha amino group of the next. These very strong covalent bonds leave a free amino group on the first amino acid and a free carboxyl group on the second one. The carboxyl group is free to make another peptide bond. Polypeptides grow in this way—by the addition of new amino acids onto the free carboxyl ends.

This scheme results in a chain where the amino acid on one end has a free amino group, while the amino acid on the other end has a free carboxyl group. The end with the free amino group is called the amino-terminus, or N-terminus. The other end is called the carboxyl-terminus, or C-terminus.

Amino Acid Alphabet

NAME	ABBR.	CODE
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

To make it easier to keep track of amino acids in a protein chain, we assign single letters to correspond to each amino acid. It's an amino acid alphabet. About ½ of the letters are obvious, because they are the first letter of the amino acids' names. But 4 amino acids start with A, 3 start with G, 2 start with L, 2 start with P, and 3 start with T, so some adjustments are made.

Because each amino acid has slightly different chemistry than all of the others, changing the order and composition of the amino acids changes not only the "words" they make, but also what the polypeptide chain does and how it works.

The Primary Structure of a Protein

Proteins are made up of chains of amino acids that vary in length, the kinds, and the order of their constituent amino acids. This sequence of amino acids is referred to as the primary structure of a protein, and that determines its properties.

Two different proteins will never have the same sequence of amino acids. But every molecule of a given protein will be identical.

But just like a sequence of words alone does not make a novel, a sequence of protein letters does not make a functional protein. Like the elements of a plot must work together properly for a novel to be enjoyable, interactions between amino acids in a protein must occur for a protein to function.

These interactions give rise to a protein's characteristic structure. This structure is so important to function that when a protein loses structure, its function is lost, too.

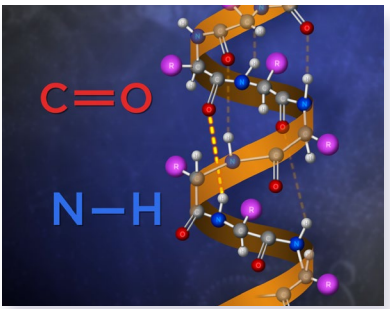
Amino acids have different side chains, or R groups, and the chemistry of the R groups affects the way any 2 side chains interact with each other. A positively charged R group is attracted to a negatively charged one, creating a force holding 2 groups close together. Conversely, any 2 groups with the same charge are pushed apart by repulsive effects.

Charges, though, are not the only kinds of interactions that affect the structure of proteins. The hydrophobic R groups all try to hide away from water. In a protein with hydrophobic side chains, the side chains associate with each other and exclude water. This is a driving force for the protein to assume its functional shape.

The other amino acids are hydrophilic and are more likely to seek and interact with water because their R groups can form hydrogen bonds.

R groups aren't the only parts of a protein where interactions occur; even the double-bonded oxygens and amine groups on the peptide backbone get into the act.

This is why the order and location of the amino acids in a protein are so important: They determine what interactions can occur and ultimately the shape of the protein.



Secondary Structure

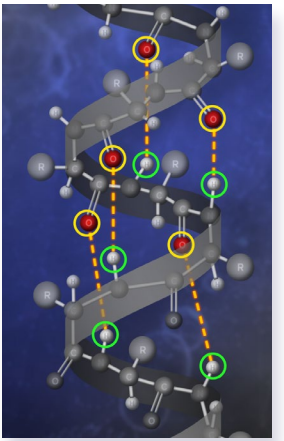
The process of establishing protein structure is called protein folding. This process is complicated and occurs at multiple levels, depending on how far apart in the primary sequence amino acids are that are interacting.

Interactions between amino acids that are close in primary sequence give rise to secondary structure and occur mostly as hydrogen bonds between atoms in the peptide backbone. Specifically, these involve the carbonyl group (C=O) of one amino acid and a nearby NH group of another.

Such hydrogen bonds can, depending on the sequence of the polypeptide in a given region, give rise to 2 main types of characteristic helical structures, plus 2 lesser variations.

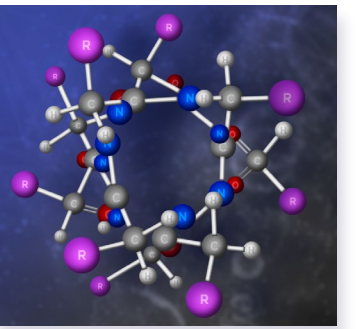
The alpha helix structure is one such secondary structure. It was originally predicted by Linus Pauling and Robert Corey, for which Pauling was awarded the Nobel Prize in Chemistry in 1954.

There are 2 image formats that are commonly used to depict proteins in biochemistry.

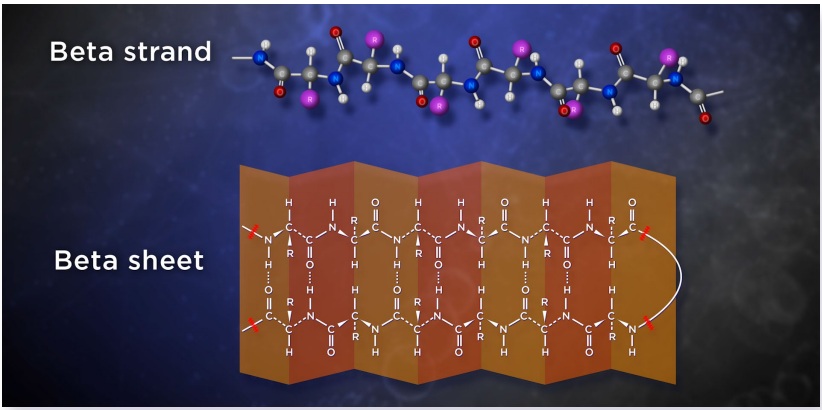


A ribbon diagram is a simple cartoon depiction of the structure of a protein that guides your eyes to depict the secondary structures within it. To help you see the forces that help stabilize an alpha helix, the individual atoms of the polypeptide's backbone are superimposed onto a ribbon. The stabilizing forces are shown as dotted lines, each of which is a hydrogen bond between a hydrogen and an oxygen. The hydrogens are located on NH groups in the polypeptide backbone, and the oxygens are on the carbons of a carbonyl group on the backbone of a different amino acid.

Pauling's alpha helix is a 3-D coil. If you look at it from the top instead of looking at it from the side, you can see that the R groups are all oriented to point away from the central core of the helix. R groups can have a large influence on the stability of a secondary structure. Their size and flexibility influence the type of secondary structure a given sequence assumes.



A second structure Pauling predicted is a flattened helix—a 2-D pleated structure called a beta strand, which can readily form hydrogen bonds with other beta strands to give rise to higher-order structures called beta sheets.



Alpha helices and beta sheets are the 2 most common secondary structures, but there are also variants. The alpha helix has 18 amino acids per 5 turns of the helix, which works out to 3.4 amino acids per turn. But there are also variants with slightly different numbers of amino acids per turn that only occur over very short amino acid stretches.

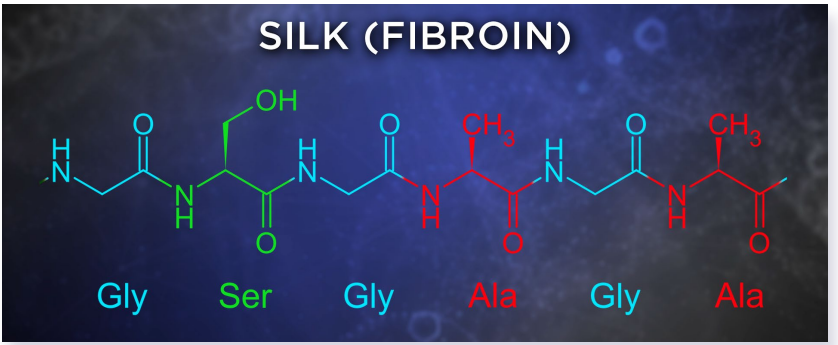
Another consideration is the way a helix coils. There are 2 possibilities, referred to as right- and left-handedness. The easiest way to determine the handedness of a helix is by looking at the direction of the coil moving away from the viewer. Looking end-on, a right-handed helix will have coils extending away from the viewer in a clockwise direction; a left-handed helix will coil counterclockwise. In proteins, right-handed helices are much more common.

The different R groups of the amino acids in a stretch of polypeptide play important roles in determining secondary structures. By examining the known structures of thousands of proteins and tallying how often each amino acid is found in the different forms of secondary structure, a likelihood of each amino acid being in any of the structures can be determined.

The data shows only tendencies, as there are no absolutes for finding any amino acid in any particular secondary structure. Indeed, some amino acids, such as the fibrous proteins, go counter to the values shown in the table in their secondary structures.

			ALPHA HELIX	BETA SHEET	REVERSE TURN	ALPHA – BETA
Q	Gln	Glutamine	1.59	0.52	1.01	1.07
A	Ala	Alanine	1.41	0.72	0.82	0.69
L	Leu	Leucine	1.34	1.22	0.57	0.12
M	Met	Methionine	1.30	1.14	0.52	0.16
E	Glu	Glutamic Acid	1.27	0.98	0.84	0.29
K	Lys	Lysine	1.23	0.69	1.07	0.54
R	Arg	Arginine	1.21	0.84	0.90	0.37
F	Phe	Phenylalanine	1.16	1.33	0.59	–0.17
I	Ile	Isoleucine	1.09	1.67	0.47	–0.58
H	His	Histidine	1.05	0.80	0.81	0.25
W	Trp	Tryptophan	1.02	1.35	0.65	–0.33
D	Asp	Aspartic Acid	0.99	0.39	1.24	0.60
V	Val	Valine	0.90	1.87	0.41	–0.97
N	Asn	Asparagine	0.76	0.48	1.34	0.28
T	Thr	Threonine	0.76	1.17	0.96	–0.41
Y	Tyr	Tyrosine	0.74	1.45	0.76	–0.71
C	Cys	Cysteine	0.66	1.40	0.54	–0.74
S	Ser	Serine	0.57	0.96	1.22	–0.39
G	Gly	Glycine	0.43	0.58	1.77	–0.15
P	Pro	Proline	0.34	0.31	1.32	0.03

Fibrous proteins are comprised mostly of primary and secondary structures, which sometimes combine into bundles, making long, strong fibers. Some fibrous proteins contain only a subset of the 20 amino acids that is organized in a regularly repeating pattern.



Silk is a fibrous protein called fibroin that uses the 3 smallest amino acids to make beta sheets over and over. The structure of silk fibers gives them very high tensile strength, and woven silk is one of the most durable fabrics.

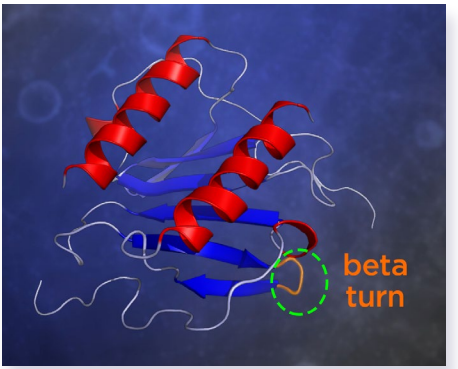
Fibrous proteins are grouped into several large superfamilies. Perhaps the most important fibrous protein for multicellular organisms is the superfamily known as collagen. Collagen, which functions to hold cells together, is in the extracellular space of connective tissue and makes up 25% to 30% of the protein in human bodies.

Tertiary Structure

The regular, repeating secondary structures of collagen and other fibrous proteins are very important. But most proteins are not fibrous, and in these proteins, secondary structures do not extend for long distances, as they do in fibrous proteins.

Instead, for nonfibrous proteins, secondary structures are local patterns connected by bends and folds between the local secondary structures to create the tertiary structure.

The most distinguishing structural features of nonfibrous proteins are turns, of which the most common is the beta turn, which essentially reverses the direction of a polypeptide chain when drawn in 2 dimensions. Such a turn is stabilized by a hydrogen bond. Interruptions like turns lead to a more complex 3-dimensional structure than a simple alpha helix or beta sheet.



Another nonrepeating feature of polypeptide chains is the random coil, which is a short, random-looking stretch of a polypeptide. Like turns, random coils interrupt repeating secondary structures and allow the polypeptide chain to fold on itself.

Bends and random coils provide enormous diversity for the shapes and structures of folded proteins, all of which started out as an amino acid chain. It is this diversity of form that allows proteins to perform their specific functions.

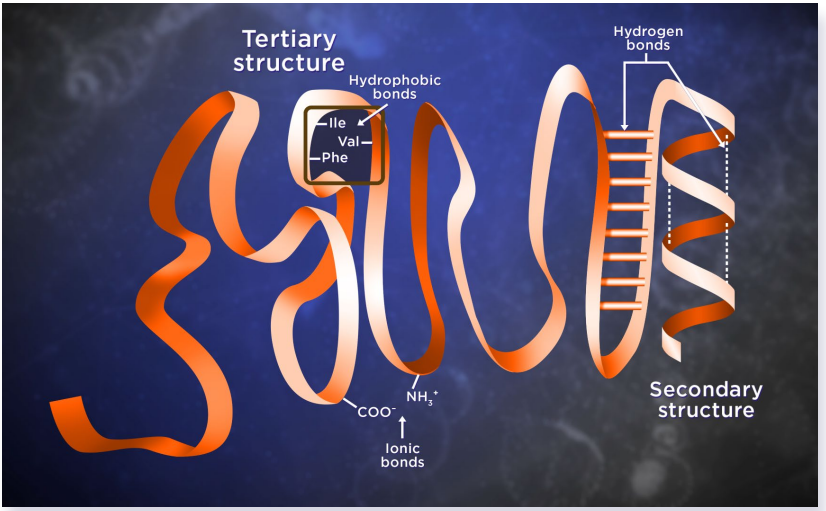
Proteins that fold into tertiary structures are called globular proteins, and they are the most abundant proteins in cells.

You can see thousands of known protein structures in the online **Protein Data Bank**.

For a tertiary structure to exist, there must be forces to stabilize it. In contrast to secondary structure, in which the stabilizing forces are hydrogen bonds, tertiary structures are stabilized by several different forces. Most of these are not covalent bonds; rather, they're weaker interactions between amino acids in a protein sequence.

Like secondary structure, hydrogen bonds form and help hold the structure together. Almost as strong as covalent bonds, ionic bonds between a positively charged R group of one amino acid and the negatively charged R group of another amino acid are stabilizing forces.

Most distinctive, hydrophobic interactions are major considerations in tertiary structure of most globular proteins. Hydrophobic side chains of molecules tend to associate with each other and exclude water. As a result, in an aqueous cellular environment, polypeptides tend to fold so that hydrophobic amino acids are buried in the interior of a globular protein. This allows their water-fearing R groups to cluster together and avoid water. The tendency of hydrophobic amino acids to behave this way is called the hydrophobic effect, and it plays a very important role in protein structure.



Conversely, the amino acids with hydrophilic R groups tend to orient on the outer surface of a protein, where they can interact with water.

Many proteins also contain metal ions. A single metal ion can interact with the R groups of multiple amino acids by coordination. The most common metals found in proteins are iron, zinc, and copper.

A significant stabilizer of tertiary structure is the covalent disulfide bond. Cysteine has the only side chain that commonly forms these strong covalent bonds.

> **A protein's function depends on its structure. If a protein loses its 3-D structure and undergoes denaturation, its structure is changed and its function is lost. That's what happens when we cook things; proteins get denatured because the heat overcomes the forces that stabilize 3-D structure.**

Quaternary Structure

Proteins that are made up of more than one polypeptide have quaternary structure, which is defined by interactions among individual folded polypeptides. Quaternary structure can help regulate the activity of proteins.

Hemoglobin, for example, has 4 folded polypeptides that associate with each other to form the functional protein. Note that these are 4 distinct polypeptide chains that interact with each other but are otherwise not physically connected.

READINGS

Al-Karadaghi, "Protein 3D Structure," https://proteinstructures.com/Structure/protein-structure1.html.

Goodsell, *Our Molecular Nature*.

SWISS-MODEL, "Part 1: Introduction to Protein Structure," https://swissmodel.expasy.org/course.

QUESTIONS

1 Hydrogen bonds hold together the 2 strands of DNA molecules as well as alpha helices and beta sheets in proteins. It takes boiling temperatures to separate the 2 strands of DNA, but most alpha helices and beta sheets can have their hydrogen bonds broken with much lower temperatures. Speculate on why there is a difference between hydrogen bonds in DNA and in proteins. (Hint: The answer to this is very simple.)

2 Fibrous proteins often use a small subset of the 20 amino acids, whereas globular proteins almost always use all 20 of the amino acids. Speculate why this might be.

3 The interior of cells' membranes is very nonpolar, unlike the aqueous environment of the rest of the cell. Most proteins found in water fold so as to keep amino acids with hydrophobic side chains in the interior of the protein. Predict the arrangement of amino acids that is found primarily in membranes. Note that not all parts of membrane proteins are solely in the membrane; some project into the cytoplasm or exterior of the cell.

CLICK HERE TO SEE THE ANSWERS.

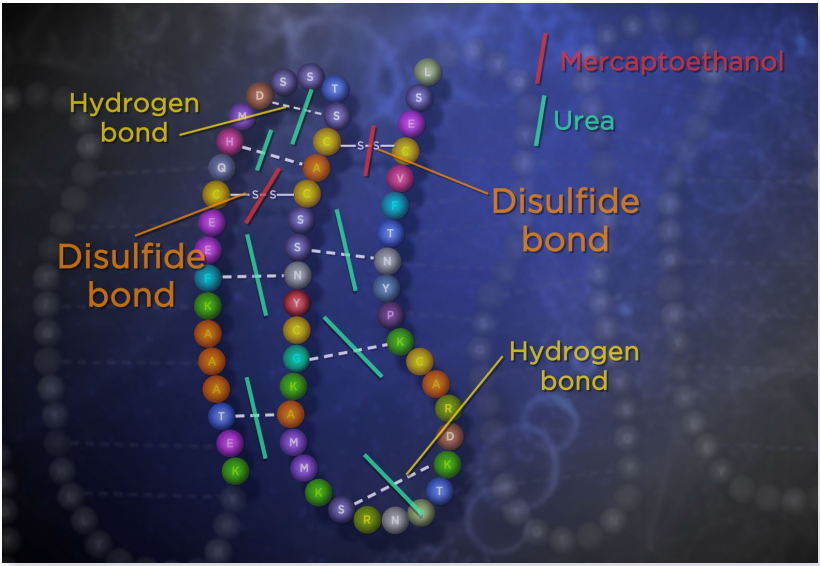
05 PROTEIN FOLDING, MISFOLDING, AND DISORDER

Proteins take on an enormous variety of shapes, stabilized by several types of bonds. How each protein gets its shape is the realm of protein folding.

Protein Folding

In the 1950s, the first experiments began by taking a protein out of the cell, unfolding it, and then seeing if it could refold in a test tube, independent of any cellular factors. The protein that Christian Anfinsen serendipitously chose was the enzyme ribonuclease (RNase) A, which is relatively small—made up of only about 100 amino acids—and extraordinarily stable. While most enzymes are exquisitely sensitive to changes in temperature and pH, ribonuclease A is not easily denatured, thanks to disulfide bonds connecting cysteine molecules in 4 different locations.

Proteins fold mind-boggingly quickly, often on the order of milliseconds.



To completely unfold RNase, the disulfide bonds must be broken using a chemical agent such as mercaptoethanol, and the hydrogen bonds can be broken using something like urea. Using both together completely unfolds RNase, which is then no longer capable of enzymatic activity.

Once the enzyme is unfolded, it can at least partly refold outside the cell, as Anfinsen showed in work that won the 1972 Nobel Prize in Chemistry. This is called renaturation because the protein has been returned to its native, or natural, state.

But the renaturation was not perfect: Sometimes cysteines misaligned and formed the wrong disulfide bond. Nevertheless, the fact that at least some renaturation occurred properly meant that the folding was dictated by the sequence and not by some cellular factor.

Experimentally altering amino acid sequences backs this up: Changing amino acid sequence changes 3-D structure.

Still, Anfinsen's choice of RNase was fortunate: Not many proteins are as close to self-sufficient about folding themselves as ribonuclease A. Many proteins have little or no ability to refold after you denature them.

Some proteins begin folding as they're being made, meaning that the beginning of a polypeptide chain can start folding before the rest of the sequence is attached. This simplifies the number of available options. For example, if side chains of amino acids in the first few amino acids have already folded when interacting with each other, they are less likely to be available to interact with amino acids added later, thus reducing the likelihood of interactions producing incorrect folds.

By contrast, a protein that is being denatured doesn't have such restrictions. First, many other random motions of amino acids in the polypeptide become possible, enabling interactions between regions that normally wouldn't associate. The result is an improperly folded protein that's inactive. Second, amino acids of one molecule of the polypeptide can also gain opportunities to interact with those of another polypeptide, forming a protein aggregate.

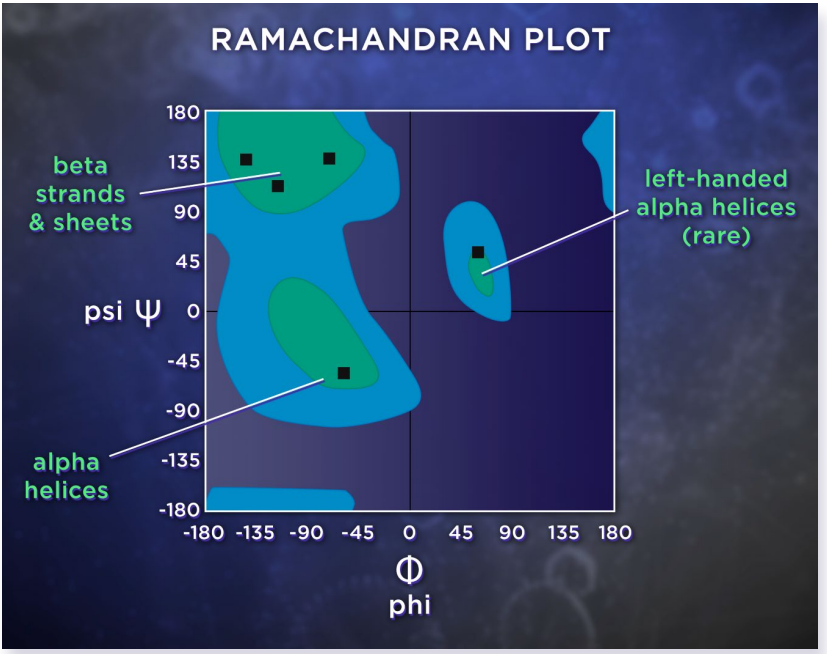
And, at least in principle, there are many ways to fold. A calculation by Cyrus Levinthal in the 1960s showed that the number of ways a small polypeptide of 100 amino acids could fold would be in the same ballpark as the number of atoms in the universe.

Fortunately, there are a few simplifications resulting from structure. The double bond holding the oxygen atom to the carbon in the peptide bond is electron-rich and can electronically rearrange into an alternative structure that turns the adjacent carbon-nitrogen bond into a double bond.

Both of these structures exist interchangeably through a phenomenon called resonance. Any resonant double bond is fundamentally different from a single bond because it is locked in place and cannot rotate, whereas a single bond can.

Because of this, it can be assumed that the peptide bond is fixed and cannot rotate. The 2 single bonds on the atoms on either side of it, though, remain completely free to rotate. This limits the polypeptide to conformations resulting from rotation around those 2 single bonds: the nitrogen–alpha-carbon bond (called phi) and the carbon–alpha-carbon bond (called psi).

Most theoretically possible phi and psi angles give rise to structures that are unstable, because such angles would result in putting 2 atoms too close together. The phi and psi angles that give stable structures correspond mostly to rotational angles found in well-known secondary structures.

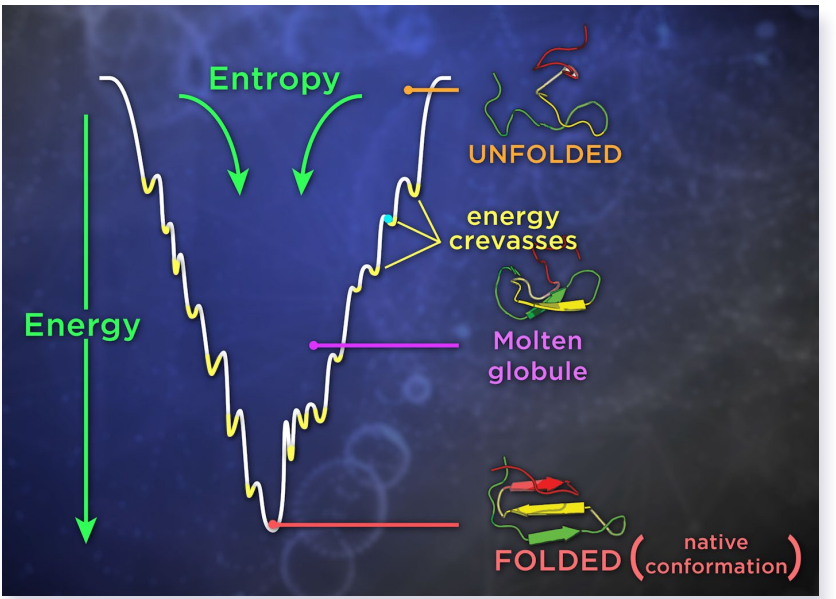


A Ramachandran plot shows all theoretically possible rotational angles for phi and psi.

Another consideration for folding is in terms of energy. Experimental evidence shows that protein folding is thermodynamically favorable, whether by releasing heat or increasing entropy, or both.

Formations of bonds/forces that are thermodynamically favorable include several known to stabilize protein structure: hydrogen bonds, ionic bonds, hydrophobic effects, and disulfide bonds.

Researchers think that folding proceeds energetically by an overall energy funnel mechanism, not unlike that of a skier moving down a long hill with numerous miniature hills along the way. The unfolded protein is at the top of the hill, in the highest energy state. As the protein makes its way down the energy hill, it accumulates favorable interactions that lower its energy progressively.



Just as different skiers could take slightly different paths down to the bottom of the hill, not every molecule of a protein must take an identical path to the bottom of the energy funnel. The path is not always smooth. Along the trail to the folded state, a protein might fall into a local energy crevasse and would have an energy barrier to overcome as a result.

But overall, the native conformation is the lowest energy state that the protein can achieve, and as folding proceeds, interactions that favor the native state accumulate. Each favorable interaction lowers the energy state while also narrowing the range of remaining options for the folding of the polypeptide, in effect increasing the odds of the protein achieving its native form. But that isn't the whole story.

Simulations can approximate the folding of a few small-size proteins, but we are still far from a general algorithm capable of using primary structure to predict how secondary structures are arranged into a folded 3-D conformation.

Helper Proteins

Even small proteins (with 100 amino acids or fewer) taken out of the cellular environment can have delays or small errors when they refold. And some larger proteins, once unfolded outside the cell, fail to regain their native conformation at all. Clearly, there is something different about the cellular environment that allows them to fold correctly.

But conditions within the cell are in some ways less favorable for the proper folding of proteins, because the inside of the cell is very crowded. This crowding tends to bump partly folded proteins up against each other, increasing the likelihood of misfolding or aggregating into mistaken clumps.

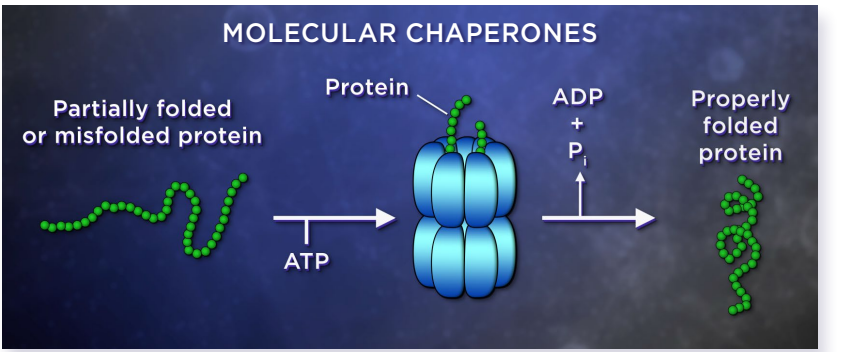
This is where helper proteins—which work in part by preventing inappropriate interactions during folding—come in.

To assist in correct protein folding in challenging conditions, all cells have molecular chaperones, which are any proteins that help other proteins achieve their functionally active conformation but do not remain associated with the proteins once folding has taken place.

There are several classes of chaperones that work together to ensure that proteins are properly folded. Some help in the actual folding process, while others work with misfolded proteins to allow them to refold correctly. Other chaperones tag hopelessly misfolded proteins for destruction.

Chaperone proteins were originally identified in cells transiently exposed to high but nonlethal temperatures. Heat can have the effect of breaking weak interactions like hydrogen bonds and could result in partly unfolding cellular proteins. The response of cells to heat is to make large amounts of proteins called heat shock proteins (HSPs), which are molecular chaperones that help refold the partially heat-denatured proteins in these cells.

In fact, under any stressful conditions with an increase in misfolded or unfolded proteins, cells increase their production of molecular chaperones.



But all cells make chaperones, even when they are not under such stress. That's because when polypeptides are being made, their hydrophobic regions are at risk of associating with hydrophobic components of other partly folded protein molecules in the crowded environment of the cell. This could lead to formation of toxic misfolded aggregates.

Chaperones prevent this from happening. But because all proteins eventually need to be replaced, there are also chaperones that identify and target proteins for degradation.

Protein Folding Gone Wrong

Some of the most mysterious diseases result from protein folding gone wrong. These are the prion diseases, also known as transmissible spongiform encephalopathies (TSEs), which affect humans and other animals. They are a group of degenerative disorders that affect the brain. Mad cow disease, or bovine spongiform encephalopathy (BSE), is one of the best-known prion diseases.

Prion diseases are caused by a misfolded protein that somehow causes its normal counterparts to fold in abnormal ways and remake themselves in the same misfolded image. All known prion diseases in mammals affect and destroy brain tissue. All are progressive, meaning that they worsen and spread. There is no known treatment, and all are fatal. Prions are thought to be the cause of several human diseases, including Creutzfeldt–Jakob disease, fatal familial insomnia, and kuru.

Like prion diseases, Alzheimer's and Parkinson's diseases are also associated with misfolded proteins that aggregate.

We don't understand exactly how prions are transmitted from one individual to another; prion diseases sometimes appear without a known cause, sometimes are inherited, and sometimes are acquired through exposure to infected material.

Drug research began around 2003 to develop pharmacological chaperones, known as pharmacoperones. Some studies in cell cultures and animal models suggest that this approach could be useful in treating or even preventing a variety of protein misfolding diseases.

Flexible Proteins

Intrinsically disordered proteins (IDPs) have no fixed 3-D conformation in at least part of the protein. Surprisingly, this is not a harmful aberration. Numerous proteins—perhaps 40% of all proteins—may have at least one region that is intrinsically disordered.

But if structure determines function, then why would so many proteins naturally have portions with no fixed shape? Their lack of fixed structure makes possible a different kind of function: flexibility.

Proteins are not rigid structures; instead, they are flexible. And IDPs take this flexibility to the extreme. Some have no fixed shape anywhere throughout their structure and are called fully disordered. Others are partly folded into a fixed shape but also have regions without a fixed shape. Such proteins are described as partly disordered.

A protein with a fixed 3-D conformation is limited in the molecules it can interact with. By contrast, an IDP could bind to a wider variety of partners, adopting a different structure complementary to each partner. Indeed, intrinsically disordered regions are found in proteins that are known to interact with multiple molecular partners, with functions that include signaling and regulation.

During the 21st century, IDPs have come to be recognized as a class of proteins in their own right. Analysis of the sequence of proteins known to be disordered reveals that they are unusually rich in hydrophilic amino acids, such as cysteine.

The association of hydrophobic amino acids inside a protein to avoid water—meaning primarily the branched-chain amino acids, such as leucine, phenylalanine, and tryptophan—is one of the forces favoring protein folding. Conversely, a lack of those hydrophobic amino acids in IDPs may help lead to and determine their disordered structure, because there will be much less of the hydrophobic effects that help drive the folding process.

IDPs have vastly expanded our ideas about protein structure and folding. For some proteins, function results from proper folding into a regular 3-D structure. But for other proteins, function may arise from exactly the opposite—the inability of a protein, or a part thereof, to form a precise structure. This may be what allows for flexible interactions with a much wider variety of other biomolecules.

READINGS

Crow, "Go with the Fold," https://www.chemistryworld.com/features/go-with-the-fold/3008748.article.

Foldit: Solve Puzzles for Science, https://fold.it/portal/info/science.

QUESTIONS

1 To completely denature RNase, you must break hydrogen bonds with heat and/or urea and disulfide bonds with mercaptoethanol. If you remove the heat/urea and the mercaptoethanol, some of the original enzymatic activity returns. Speculate on why this is the case. Interestingly, if you remove the heat/urea but leave a trace of mercaptoethanol in the mix, you get more activity than when there is no mercaptoethanol. Why?

2 In chaperonins, unfolded proteins enter the chamber, and then the interior of the chaperonin changes from having exposed hydrophobic side chains to having hydrophilic ones exposed. Why is this important for favoring folding, and what feature of a folded protein is it favoring in the folding process?

3 Intrinsically disordered proteins have relatively few hydrophobic amino acids. Speculate on how this could contribute to their lack of a stable or fixed structure. What kinds of interactions would you predict they would have on binding to another protein?

CLICK HERE TO SEE THE ANSWERS.

06 HEMOGLOBIN FUNCTION FOLLOWS STRUCTURE

One of the most remarkable proteins in virtually all higher animals is hemoglobin: the protein in red blood cells that carries oxygen from the lungs to tissues. Its lesser-known function is to haul away carbon dioxide from those same tissues. In the presence of oxygen in the lungs, hemoglobin changes shape to bind as much oxygen as it can. Elsewhere in the body, hemoglobin changes shape again to fine-tune delivery of oxygen: less to tissues with low oxygen needs and more to tissues with high oxygen needs.

Carrying Oxygen

Our cells need oxygen to efficiently obtain energy from food. While some microbes survive in anaerobic conditions—that is, without oxygen—their use of foodstuffs is inefficient compared to ours. This is why animals have oxygen-carrying proteins to provide a steady supply of oxygen. Almost every vertebrate has hemoglobin as the oxygen carrier in its blood.

Although hemoglobin is the best-known oxygen-binding protein in the body, a related protein called myoglobin is found in heart and skeletal muscle, where it functions as an oxygen reservoir. Myoglobin is great at holding onto oxygen, but only when the oxygen concentration gets very low does myoglobin release its oxygen.

The different roles these 2 oxygen-binding proteins play is reflected in their structures, as was discovered in 1958, when hemoglobin and myoglobin became the first proteins to have their 3-D structures determined.

Both hemoglobin and myoglobin are globular proteins, with bends and turns bringing amino acids that are not necessarily close in the primary structure of the protein into close proximity in the tertiary structure.

Myoglobin is a single polypeptide containing 154 amino acids, which fold into 8 helices connected in between by proline-rich loops. The result is a distinctive globin protein structure.

Hemoglobin has a similar secondary and tertiary structure but adds a crucial quaternary structure. This is because hemoglobin contains 4 polypeptide subunits, making it a tetramer, while myoglobin is a monomer.

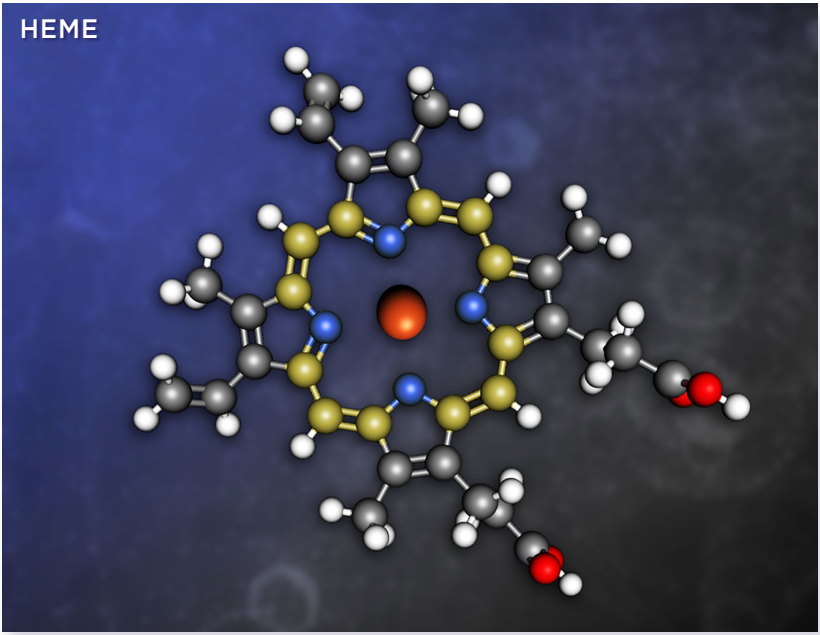
Remarkably, the number of amino acids in the alpha and beta subunits remains the same for virtually all animals, while the precise sequence of amino acids differs increasingly with evolutionary distance.

Humans and chimpanzees have identical hemoglobin sequences, but humans and horses have alpha or beta subunits that differ by a few dozen amino acids each.

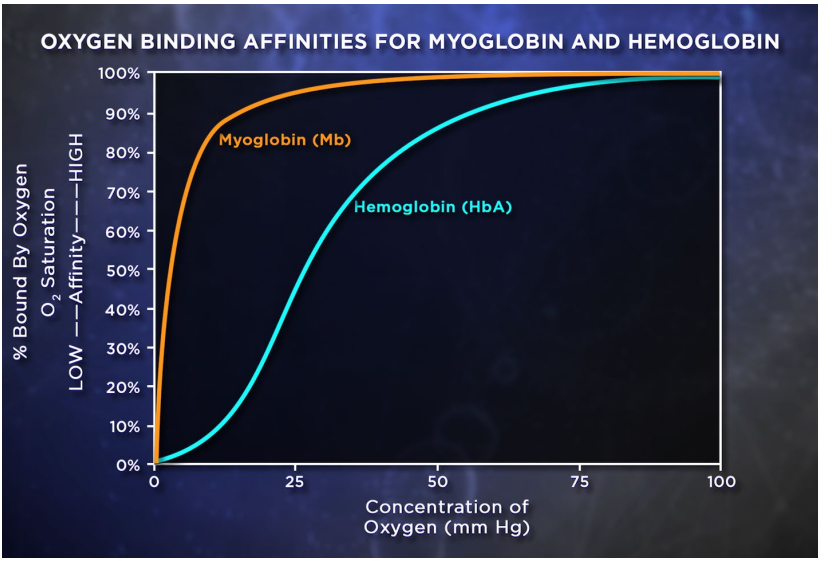
Hemoglobin has 2 identical alpha subunits, each with 141 amino acids arranged in 7 helical regions. There are also 2 identical beta subunits, each with 146 amino acids divided into 8 helical regions. Hemoglobin is like a quartet of 4 closely coupled myoglobins.

Proteins performing specialized tasks often need help from nonprotein components called cofactors. The cofactor in hemoglobin is called heme, which binds oxygen, while the polypeptide chains modulate how it binds oxygen.

The distinctive ring structure of heme is found in molecules classified as porphyrins. In hemoglobin, each globin subunit is covalently linked to the ring structure of heme.



Hemoglobin goes racing through the lungs in the red blood cells. A single oxygen molecule binds there and tugs an iron atom a tiny fraction of a nanometer, and this results in the entire hemoglobin getting loaded with oxygen before it exits. Were it not for that miniscule change in iron's position, resulting in cooperativity—in which the initial binding facilitates subsequent binding—there would be almost no vertebrates, including humans.



A binding curve moved to the left, like that of myoglobin, indicates a greater oxygen binding affinity, whereas a curve shifted to the right indicates a reduced oxygen binding affinity. And an S-shaped curve, like hemoglobin has, means the affinity for oxygen is changing with increasing oxygen.

On average, a red blood cell does a complete circuit through your body in only about 20 seconds.

Carrying Protons and Carbon Dioxide

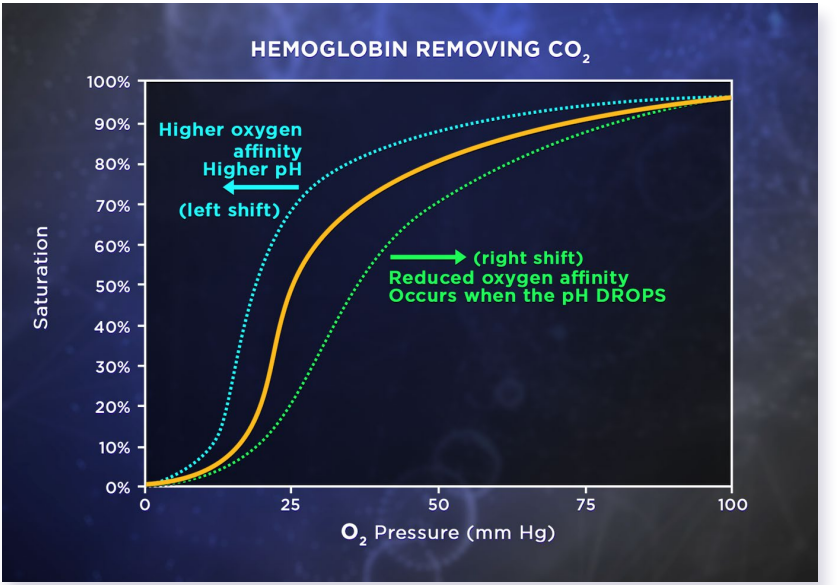
Besides carrying oxygen, hemoglobin has other functions. It also picks up 2 sets of molecular baggage in the tissues and dutifully hauls them back to the lungs.

- Carbon dioxide is one type of baggage, a direct product of various oxidative processes in cells that create energy. Small quantities of carbon dioxide can be tolerated by cells, but at higher concentrations, it is toxic. Hemoglobin binds carbon dioxide and takes it back to the lungs to be exhaled.

- Protons also need to be removed, especially from actively metabolizing tissues, such as working muscle. Protons come from acids, which are by-products of cellular metabolism.

Both protons and carbon dioxide are great indicators of cellular metabolism: The higher the rate of metabolism, the more these compounds are produced.

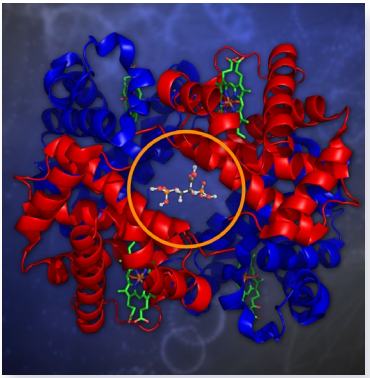
Hemoglobin binds protons and carbon dioxide. Later, it delivers them to the lungs, where they get released. Oxygen binds to the iron in heme, whereas protons and carbon dioxide instead are carried on amino acid side chains of the globin subunits—most commonly histidine. The binding of carbon dioxide and protons to hemoglobin also facilitates the release of oxygen from hemoglobin to tissues.



When tissues are rapidly metabolizing, the protons and carbon dioxide they release bind to hemoglobin. This lowers hemoglobin's affinity for oxygen, so it releases more oxygen to the tissues that need it. When tissues are not metabolizing rapidly, there are fewer protons and carbon dioxide, so hemoglobin has a higher oxygen affinity and releases less oxygen where it is not needed as much.

Hemoglobin's release of oxygen is increased by another molecule produced abundantly by rapidly metabolizing tissues. Besides oxygen, protons, and carbon dioxide, hemoglobin can bind a small compound known as 2,3-bisphosphoglycerate (2,3-BPG).

Hemoglobin quickly binds oxygen in the lungs and releases it in tissues, and it "reads" the need for oxygen by the amount of carbon dioxide and protons it binds.



On hemoglobin, 2,3-BPG only binds to one very special site. Hemoglobin has a donut-like shape with a hole in the middle, and this is where 2,3-BPG binds.

The combined effects of protons, carbon dioxide, and 2,3-BPG reduce the affinity of hemoglobin for oxygen. Thus, increasing concentrations of protons, carbon dioxide, and 2,3-BPG all favor the release of oxygen.

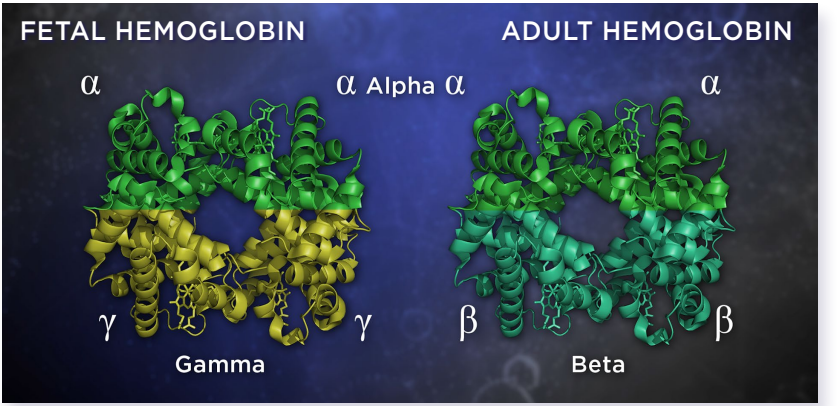
When 2,3-BPG binds to hemoglobin, though, it is important that it be released before hemoglobin returns to the lungs. Otherwise, the hemoglobin will remain in the state of low affinity for oxygen and will not bind oxygen readily when it passes through the lungs. 2,3-BPG is not bound to hemoglobin covalently, so normally this little molecule readily detaches from hemoglobin during its journey back to the lungs, leaving hemoglobin free to flip to the high-affinity state when it gets to the lungs.

Adult versus Fetal Hemoglobin

Hemoglobin's ability to pick up and release oxygen is carefully regulated. A related system is at work during pregnancy, when hemoglobin has to serve the mother and the developing fetus.

Adult hemoglobin has 2 alpha subunits and 2 beta subunits. A growing fetus also needs oxygen, but in the womb, its only source is the mother's bloodstream. Because blood cells and hemoglobin don't cross the placental barrier, it is important that the fetus be able to take oxygen from the mother's hemoglobin. Fortunately, the fetus has a slightly different form of hemoglobin that enables this.

In fetal hemoglobin, the beta subunits of adult hemoglobin are replaced by 2 similar units called gamma. The same alpha units are made as in adult hemoglobin. The overall structure of fetal hemoglobin is called alpha-2 gamma-2.



Fetal hemoglobin has a higher affinity for oxygen than maternal hemoglobin. Fetal hemoglobin behaves much like adult hemoglobin with respect to cooperativity as well as to the binding of protons and carbon dioxide.

But one difference is that fetal hemoglobin doesn't bind 2,3-BPG very well. And that's a good thing. Despite looking much like the beta globins of adult hemoglobin, fetal hemoglobin's binding region for 2,3-BPG lacks a positively charged histidine, which is replaced by a neutral serine.

This and other amino acid substitutions reduce the overall positive charge of the gamma subunit, making it less likely that the negatively charged 2,3-BPG will bind to fetal hemoglobin. Because the fetal hemoglobin binds 2,3-BPG less frequently than maternal hemoglobin, the fetus can take oxygen away from the mother.

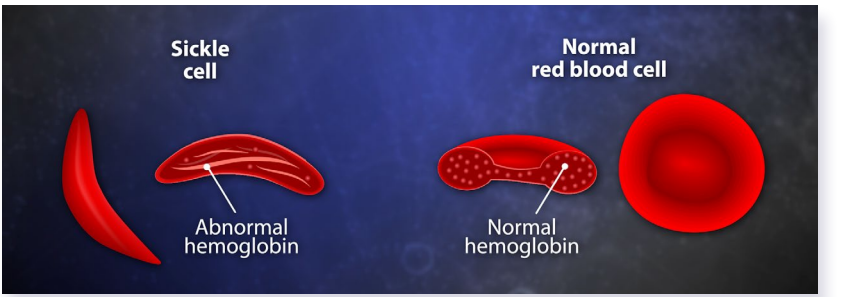
The hemoglobin in a fetus differs from the mother's hemoglobin, thus allowing the growing fetus to obtain sufficient oxygen from the maternal bloodstream.

Soon after a baby is born, its body stops making fetal hemoglobin as beta globin production increases. The heme released in the switch is broken down into the compound bilirubin, which gets disposed of by the liver. If a baby's liver is not able to process the bilirubin released as the fetal hemoglobin is broken down, it accumulates in the blood, leading to neonatal jaundice and a sickly, yellowish appearance. Premature babies, whose livers are insufficiently developed, are most likely to have this problem.

Sickle Cell Anemia

Although the fetus shifts hemoglobin production toward adult hemoglobin, a small amount of fetal hemoglobin continues to be made into adulthood. This affords an opportunity for helping individuals with sickle cell disease, or sickle cell anemia.

The word *anemia* refers to any deficiency of hemoglobin or red blood cells. The anemia in sickle cell disease is the result of a genetic condition in which red blood cells assume the shape of a sickle in capillaries.



Sickle cell anemia was the first disease ever linked to a specific mutation, and it is a fascinating instance of how a very small change in a protein has profound effects on an organism.

The sickling of red blood cells results from a mutation in the beta globin gene that causes aggregation of the hemoglobin molecules under low-oxygen conditions, such as during exercise. The result is the formation of long polymers of mutant hemoglobin that distort the shape of the red blood cell.

Once cells take on the sickle shape, they get stuck in capillaries and block blood flow, depriving tissues of needed oxygen. The result is pain and cellular damage at a minimum. Chronic oxygen deprivation can lead to vision loss, damage to vital organs, and even death. Sickled cells are recognized as damaged and removed from the blood supply by the spleen, causing the reduced red blood cell counts associated with anemia.

One approach to treat such a disease is to try to increase the synthesis of fetal hemoglobin in adults. A drug called hydroxyurea has been shown (albeit with some side effects) to induce synthesis of fetal hemoglobin. This has 2 targeted effects: It increases the amount of oxygen-carrying hemoglobin so that tissues can be better supplied with oxygen, and the increased numbers of gamma subunits interfere with the polymerization of the alpha-2 beta-2 hemoglobin to reduce cell sickling.

So, a change in a single amino acid in one subunit of hemoglobin can cause a life-threatening disease. But this is only true for a person who inherits 2 copies of the mutant beta globin gene: one from the father and one from the mother. Having only a single copy does not result in the disease. In fact, having one copy of the hemoglobin mutation seems to protect against malaria!

READINGS

Storz, *Hemoglobin*.
Thom, et al., "Hemoglobin Variants."

QUESTIONS

1. Hemoglobin differs from myoglobin in having quaternary structure and cooperativity. In general, proteins lacking quaternary structure do not exhibit traits like cooperativity. Given what you know about hemoglobin, describe a protein lacking quaternary structure that might exhibit cooperativity and how it would work.

2. Myoglobin has a greater affinity for oxygen than hemoglobin under most oxygen concentrations and was described as being able to take oxygen away from hemoglobin in the lecture. Myglobin's high affinity for oxygen is also important for its role in muscles. What do you think that role is?

3. Fetal hemoglobin has a higher affinity for oxygen than adult hemoglobin because fetal hemoglobin mostly remains in the R state. With most of its time spent in the R state, fetal hemoglobin is great for taking oxygen away from the mother's hemoglobin but is not so good at letting it go, because it doesn't as readily flip into the T state. Explain why this is usually not a problem for the fetus. (Hint: It would be a problem for the mother.)

CLICK HERE TO SEE THE ANSWERS.

07 ENZYMES' AMAZING SPEED AND SPECIFICITY

At any given time, cells of all kinds carry out thousands of metabolic reactions, breaking down some molecules and building others. And all of this has to happen fast enough to meet cellular needs. Because ordinary chemical reactions would happen far too slowly to support life, enzymes act as catalysts to accelerate chemical reactions to extraordinary extents.

Enzymes as Catalysts

Enzymes are called catalytic for 2 reasons: They speed up the rate at which biochemical reactions happen in cells, and each enzyme molecule can be used over and over without being changed.

Enzymes are proteins. They contain amino acids, whose unique side chains play important roles in the mechanism of catalysis. But what makes enzymes truly amazing catalysts is their efficiency. It's not uncommon for enzymes to increase reaction rates by a factor of a trillion or more, as happens with adenosine deaminase, which affects the synthesis of nucleotides needed for making DNA and RNA.

Not all catalysts are found in cells. A nonbiological example is the catalytic converter in cars, which converts pollutants like carbon monoxide and nitrogen oxides into less harmful substances.

Another remarkable thing about enzyme catalysis is the fact that it happens at the temperature and pressure conditions of living cells. Consider the conditions needed in factories to synthetically make something as simple as ammonia (NH_3). A widely used industrial process to make it combines nitrogen from the air with hydrogen in the presence of a metal catalyst—but they form ammonia only if heated to 450°C (840°F) at 200 times atmospheric pressure.

In contrast, nitrogen-fixing bacteria make ammonia from nitrogen in the air much more efficiently and without the need for high temperature and pressure because they use enzymes as catalysts.

The molecule being acted on by an enzyme is called its substrate. In the reaction for making ammonia, nitrogen and hydrogen are the substrates. Substrates are also called reactants because they are the molecules undergoing the reaction. The molecule produced by the reaction is called the product. The example had only one product, ammonia, but some reactions may have multiple products.

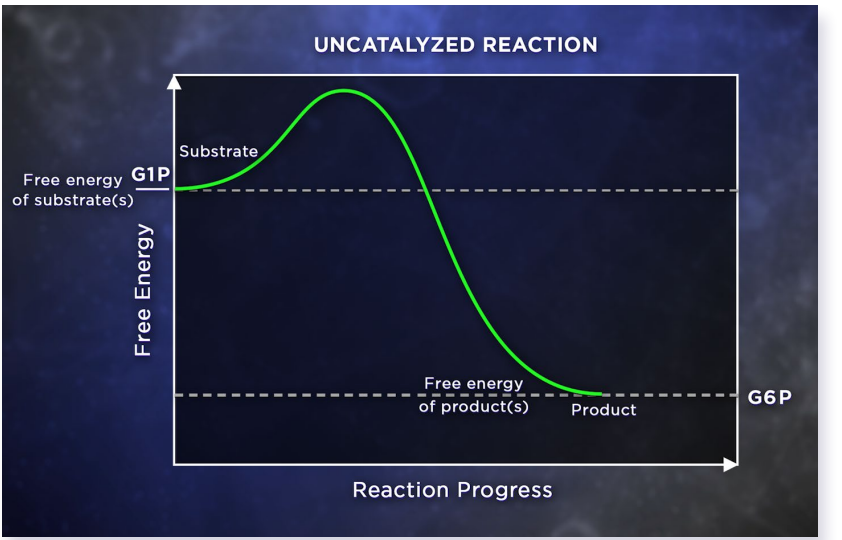
Most enzymes have names that end with *-ase*. The few exceptions include some of the first-discovered enzymes for digesting protein, such as pepsin, trypsin, and chymotrypsin. The rest of an enzyme's name tells you something about what the enzyme does.

Sometimes, the name tells you which substrate the enzyme breaks down. A lipase breaks down a fat; a protease breaks down a protein. Other enzyme names indicate an output or product: ATP synthase is the enzyme that catalyzes synthesis of adenosine triphosphate (ATP). Many other enzymes have names that describe a type of reaction, using words like *synthetase*, *oxidase*, *reductase*, *transferase*, or *isomerase*. In general terms, any reaction involving enzymes can be described as substrate *S* being converted to product *P*.

Enzyme Energy

To look at what's happening in catalysis, we use Gibbs free energy (*G*), a measure of the amount of energy available to do useful work in a process. In a biochemical reaction, the work being done is the making and/or breaking of chemical bonds.

The energy changes for a reaction can be shown in a simple graph. On the *y*-axis is the Gibbs free energy of the substrate and product of the reaction, and the *x*-axis follows the progression of the reaction. On the far left of the graph is the free energy of the substrate, G1P; on the far right is the free energy of the product, G6P. The line tracks the progress of the reaction from an energy perspective, starting with the uncatalyzed reaction.



Notice that the free energy of the final product is lower than the free energy of the substrate. The difference between these 2 energies is the change in Gibbs free energy during the reaction. The net change in free energy, or ΔG, is negative in the conversion of *S* to *P*, which means that energy is released by the reaction when *S* is transformed to *P*.

The Greek letter Δ is commonly used to denote change.

Also note that at first, the line increases from the energy level of the substrate, indicating that some energy input is needed before the reaction proceeds downhill. The difference between the energy of the substrate and the peak of the line is the activation energy for the reaction.

But if the final product is at a lower energy level than the initial substrate, why would you need to add any energy for the reaction to happen?

To convert one or more substrate molecules into a product, some bonds must be broken and new ones must be made. For example, the substrate molecule or molecules might have to be forced or bent into a form that will allow existing bonds to break or form, just as you might need to bend a stick to weaken it at the spot you want it to break. This contorted form of the reactants is called the transition state, and to reach it takes energy, just as you need to put in effort to bend a stick.

A reaction involved in making a crucial component of DNA takes about 78 million years. The enzyme OMP decarboxylase speeds up the reaction so that it's completed in less than 20 milliseconds—an increase in speed of 140 thousand trillion times.

That's what the activation energy is—the energy needed to get molecules to that transition state. The input of the activation energy is why the transition state is the highest energy level on the graph.

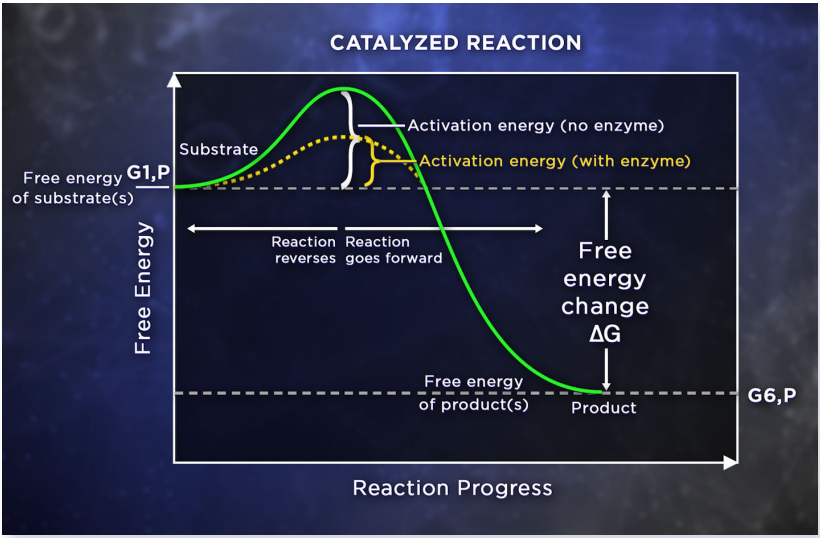
The transition state is a short-lived and unstable state, just as a bent stick does not stay that way for long—it breaks or returns to its original state. Once the substrate molecules are in the transition state, conditions are favorable for the original bonds to break and for new ones to be made. But where does the activation energy to get to the transition state come from?

Under normal circumstances, molecules are in constant motion, driven by thermal energy. They bounce around and bump into each other. A particularly energetic jostle can send some of them into a state of higher energy, so they are now at the highest point of the energy curve, the transition state, from which they can then go downhill easily.

This happens in the absence of catalysts, but rarely. Things can be sped up a bit by raising the temperature; the input of heat energy makes molecules move faster and bounce around more. But to make a significant change in the rate of the reaction, cells need an assist from a catalyst because the range of temperatures they can survive at is pretty narrow.

In a graph of the same reaction catalyzed by phosphoglucomutase, dotted lines show where the energies differ from the uncatalyzed reaction. Notice that the activation energy in the presence of the enzyme is significantly lower, but neither the beginning Gibbs free energy nor the ending Gibbs free energy is altered. Thus, the energy released in going from substrates to product is the same regardless of whether or not an enzyme was involved.

The only thing that has changed is the activation energy. In other words, the catalyst has reduced the activation energy it takes to get the substrate into the transition state. When the activation energy is lower, many more substrate molecules reach the transition state at a given temperature, so the conversion of substrate to product is correspondingly faster.



Enzyme Specificity

For a reaction to happen, reactants must be brought into a transition state that allows bonds to be broken so new ones can be made. This requires activation energy. In the presence of an enzyme, the amount of activation energy needed is reduced compared to the uncatalyzed reaction. How does the enzyme make it easier to reach the transition state?

Enzymes have an active site, a pocket into which the substrate fits and where the chemical reaction occurs. The pocket is created by the 3-D folding of the polypeptide. Amino acid side chains protrude into the cavity of the active site, giving it its very specific shape, with chemical characteristics that derive from their side groups.

Each will be positioned at specific places in the active site, ensuring that only substrates with the appropriate surface characteristics will be able to bind.

The lock-and-key model is called the Fischer model after Emil Fischer, who postulated it.

The high specificity of substrate binding sites by enzymes gave rise to a simple, but not entirely accurate, analogy that a substrate fits an enzyme like a key fits a lock. In this lock-and-key model, the enzyme is envisioned as an exact complement to the shape of the substrate it acts on.

But for most enzymes, an induced-fit model is a better description of what happens. In it, binding of the substrate at the active site changes the enzyme-substrate combination transiently, making the 2 fit closer together. This involves slight alteration of the shapes of the enzyme and the substrate, like inserting your foot into a shoe that is not the shape of your foot. The footwear stretches, and your toes might squish a bit, resulting in a snug fit.

An enzyme's flexibility lends itself to several catalytic mechanisms. One mechanism of enzymes is to bind the substrate in a way that provides a shortcut to the transition state. In other words, the best fit of the enzyme for the substrate favors the transition state, where bonds can be broken and new ones can be made.

Another mechanism, if there is more than one substrate, is as a juxtaposer that brings molecules into close proximity. Outside the enzyme, the molecules depend on random molecular collisions to bump into each other. There are many unproductive collisions, while useful interactions are rare. But inside the active site, the reactants are closer and optimally positioned to react with each other.

Enzymes commonly use other molecules to help in catalysis. Metal ions, such as iron, zinc, and copper, serve as cofactors that aid reactions. In other cases, bigger molecules called coenzymes help move electrons, atoms, or entire functional groups. Some enzymes directly participate in the reaction they are catalyzing, usually by the enzyme's amino acid side chains forming transient bonds with the substrate.

Enzymatic reactions do not occur in isolation. In cells, the vast majority of them are interdependent, with the product of one reaction serving as the substrate for the next, like a factory's assembly line. Production of biomolecules depends on chains of reactions, each fed by the one before it and each catalyzed by a specific enzyme.

Enzyme Velocity

Enzyme velocities are measured in product made per time. If substrate is in short supply, the enzyme will spend time idling, and maximum velocity will not be reached. However, if the enzyme is saturated with substrate—meaning that as soon as a product molecule is released, a new substrate is bound—then the reaction will reach a maximum rate, or velocity.

Maximum velocity is set by the number of enzyme molecules performing catalysis. If there are twice as many enzyme molecules, the maximum velocity will be twice as large, so long as there is plenty of substrate for conversion into product.

Maximum velocity gives information about the activity of an enzyme, but it is not useful for comparing different enzymes. To do this, biochemists use a measure called the catalytic constant (k_{cat}), also called the turnover number of the enzyme.

The k_{cat} is useful because it takes into consideration the concentration of the enzyme as well as the maximum velocity. The k_{cat} is the maximum velocity divided by the concentration of enzyme used, which essentially gives the number of product molecules made per enzyme molecule per second. This is used to gauge productivity.

The extremely rapid catalytic action of enzymes makes it harder to fine-tune production to changing demand. The synthesis of biomolecules by enzymes with high k_{cat} values could result in more product than desired being made before control mechanisms decrease production. Consequently, most enzymes have not evolved to have k_{cat} values that are as high as possible.

There is a maximum rate that enzymes can achieve. So-called perfect enzymes perform at the theoretical optima of enzyme activity. Perfect enzymes also exemplify the concept of affinity—the snugness of fit of an enzyme for its substrate. Enzymes with a high affinity for substrate will have significant amounts of it bound, even when substrate concentrations are low.

READINGS

Kornberg, *For the Love of Enzymes*.
Neitzel, "Enzyme Catalysis."

QUESTIONS

1 The induced-fit model of catalysis was described in the lecture as favoring a reaction in 2 ways: serving as a shortcut to the transition state and as a juxtaposer for reactions with 2 substrates. For reactions with one substrate, the flexibility of the enzyme may be important for breaking it into 2 pieces by means of another model described in the lecture. What is the model, and how might it play a role in the splitting of one molecule into 2?

2 The alkoxide ion of serine proteases is extraordinarily reactive. Why do you think it only acts on the peptide bond of the bound substrate and not on the millions of other molecules floating around in the cell?

3 Some enzymes create intermediates in their catalytic mechanism and can release them before the reaction is complete. At least one perfect enzyme called triose phosphate isomerase is in this category, and its perfectness is needed because of its ability to release an intermediate. What property might such an intermediate have, and how does perfectness play a role?

CLICK HERE TO SEE THE ANSWERS.

08 ENZYME REGULATION IN CELLS

There are thousands of distinct enzymes in cells, catalyzing billions of reactions. Unleashing such tremendous power, without mechanisms to control it, could create chaos. How does the cell control this enzyme activity?

How Cells Regulate Enzymes

Cells regulate enzyme activity, exerting precise control over synthesis and breakdown of biomolecules, for every molecule they make. And they must do so continuously to maintain stable internal conditions, called homeostasis.

One reason that biochemists care so much about the control of enzyme activity is that many medical conditions result from enzyme activity that's excessive or insufficient. By figuring out ways to inhibit or increase these reactions, we can treat the associated diseases.

The Goldilocks principle applies to enzyme activity: You don't want too much or too little; it needs to be just right. If enzymes in your thyroid gland make too much hormone, you become irritable and sleepless; if they make not enough hormone, you become tired.

These have serious consequences for your health. Enzymes are like automobiles, with accelerator and brake pedals to adjust speed so that it's appropriate to existing conditions.

Not all of the thousands of reactions that enzymes catalyze are needed at any given moment. Cells must sense levels of molecules and adjust actions on the fly. For example, suppose your muscle cells need glucose for energy to help you run. Something needs to turn on the enzyme that breaks down stored glycogen to release glucose. The glycogen breakdown enzyme must be active long enough to provide the necessary amount of glucose, and then stop—or you'd deplete all your stored reserves, which you need for emergencies.

Other situations call for maintaining a steady supply of molecules used by cells on a regular basis. If levels of a particular nutrient are low within a cell, its synthesis must be triggered. But once levels of the molecule are restored, its production must be halted. This is like the thermostat regulating temperature in your house. Regulating production of biomolecules depends on controlling the activity of the enzyme(s) involved. There must be something that senses the conditions and then a way to switch an enzyme on or off.

Cells use assembly lines to build molecules they need. In other words, there is an ordered sequence of reactions, called a pathway, in which the product of reaction 1 is the substrate for reaction 2, and so on. Each reaction in a pathway is catalyzed by one enzyme, and the product of its activity is sent on to the next enzyme, until the final product is obtained.

Chaining together multiple reactions into a single pathway gives cells a simple way to regulate the pathway. Because each enzyme depends on substrate from the reaction before it, turning off one enzyme stops the activity of all those that come after it in the pathway. As a result, each enzyme in a pathway does not need to be micromanaged by the cell. Usually, an enzyme early in a pathway is controlled by a mechanism called feedback inhibition.

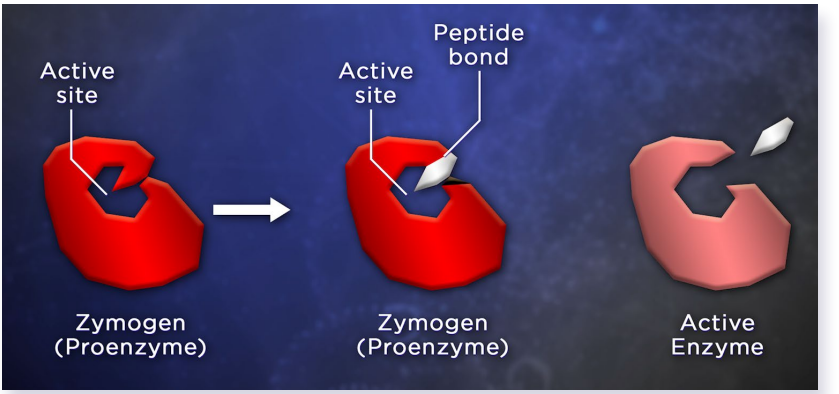
Some of the major ways by which the activity of enzymes is controlled are synthesis, allostery, physical inhibition, and covalent modification.

Allosteric Regulation

A simple, yet much more nuanced, mechanism for control is allosteric regulation, which works like a dimmer on a light to adjust brightness. In this form of regulation, the binding of a regulatory molecule in one part of the enzyme can adjust the enzyme's behavior. The binding of these regulatory molecules, called effectors, is reversible, allowing them to exert or remove control as needed.

Allosteric regulation is dependent on the concentrations of allosteric regulator molecules. High concentrations are necessary for occupying the allosteric site, and occupation of the allosteric site is what modulates the enzyme's synthesis of product.

Allosteric controls are great if what you want to do is control the "volume" of the reaction up or down. What if you prefer an actual on-off switch? One kind of on-off switch acts like a sealed-for-your-protection package. Until you break the seal, you can't get at the goodies. The "seal" in this on-off switch is a peptide bond that blocks access to the active site.



The term *zymogen* comes from *zyme*, meaning "enzyme," and *gen*, meaning "to give rise to."

Enzymes that start with the active site being covered up like this are known as zymogens or proenzymes. They are one way to keep enzymes inactive until needed. The inactive form of an enzyme like lipase or trypsin will have a name like *pro*lipase or trypsin*ogen* to tell you that it's not the active form.

It's only after one or more peptide bonds in the covered enzyme are broken that substrates get access to the active site. So, really, this is an *off-on* switch, because the zymogen is always off until activated by bond cleavage.

For example, chymotrypsin acts as a protease in intestines but will attack the pancreas that makes it if activated too soon. It is made in a form called chymotrypsinogen, which is inactive and is only referred to as chymotrypsin after the active site is exposed by peptide bond cleavage.

So, the breakage of peptide bonds is important in the body to quickly activate proenzymes, or zymogens, that would not be safe if they were always in their active form. But because breaking peptide bonds to activate the enzymes is irreversible, you might wonder how such enzymes can be turned off.

That brings us to another control mechanism for enzymes: using other proteins to cover the active site. Proteins called protease inhibitors, for example, inhibit the activity of protease enzymes by binding to their active sites tightly and preventing further catalysis. Though this might sound like the cover that was removed by the protease cleavage, inhibitors are separate peptides and physically project into, or disrupt action of, the active site.

The crudest but most effective way to control enzymes is managing whether or not the enzyme is made by a cell at all. An example of an enzyme regulated in this way is chymotrypsin, which is made in the pancreas. Other kinds of cells in the body don't make this enzyme, saving them the work of having to regulate its activity.

Covalent Modification

Another common mechanism for enzyme regulation also acts like an on-off switch, but in this case, the switch is reversible—without needing an inhibitor protein to cover the active site. How it works depends on the covalent addition or removal of what is usually a phosphate to the enzyme.

These enzymes exist in 2 states: with a phosphate or without. Switching between these 2 states determines whether it is active or inactive, though which is which depends on the specific enzyme. In other words, some enzymes are "on" when phosphate is linked to them, such as glycogen phosphorylase; while other enzymes are "on" when phosphate is removed, such as glycogen synthase.

Enzymes controlled by this mechanism get phosphate groups added to amino acid side chains—usually the polar but uncharged serines, threonines, or tyrosines. Special phosphate-adding enzymes called kinases covalently link a phosphate group to the appropriate side chain in a process called phosphorylation.

Phosphate groups are negatively charged, so adding them on to the uncharged side chain changes it from neutral to negative. This affects forces stabilizing protein structure.

This could cause a small movement and make a subtle change to the enzyme's conformation. Because function is related to structure, the new structure might be active while the previous one was inactive, or vice versa. Think of light switches; even just a small movement can change them from off to on.

However, unlike allosteric regulation, where the binding of the regulatory molecule is noncovalent and readily reversible on its own, phosphorylation can only be reversed by removing the

Enzymes can toggle between their active and inactive states, using phosphorylation and dephosphorylation.

phosphate group using another enzyme called a phosphatase. So, the reversible switch needs 2 enzymes: one enzyme for phosphorylation and another to remove the phosphate.

If the enzyme is turned on when it is phosphorylated, it gets turned off when the phosphate is removed. On the other hand, if it was active in the absence of the phosphate group, addition of the phosphate will turn it off.

A typical mammalian cell contains thousands of kinds of phosphorylated enzymes. Switching those enzymes back and forth are several hundred kinases and phosphatases, which are often responding to neural and hormonal signals.

Physical Inhibition

Many familiar drugs in use today work by changing the activity of a specific enzyme. The majority of enzyme regulators used as drugs work by reducing the activity of the target enzyme. Enzyme inhibition by drugs involves molecules binding to enzymes reversibly or irreversibly.

The irreversible ones are sometimes called suicide inhibitors. Irreversible inhibitors usually resemble the enzyme's normal substrate, but once bound, the inhibitor forms a covalent bond to the active site, blocking it permanently.

A common suicide inhibitor is penicillin, which acts by binding to a bacterial enzyme responsible for making bacterial cell walls. Unable to complete synthesis of functional cell walls, the bacteria die.

Unlike penicillin, most enzyme inhibitors used in medicine are reversible because many might act on our own enzymes, and we don't want to permanently disable them!

Reversible inhibition comes in several forms, on a spectrum from competitive, to partially competitive, to noncompetitive and uncompetitive.

For medical purposes, the most common of these is competitive inhibition, in which the inhibiting molecule resembles the normal substrate of the enzyme and competes with it for binding at the active site. When the inhibitor binds to the enzyme, the normal substrate cannot bind and the enzyme is inactive.

This is similar to what happens with suicide inhibitors, with one crucial difference: The competitive inhibitor does not make covalent links to the enzyme. With no connection between the enzyme and the inhibitor, their interaction is temporary and readily reversible.

The extent of inhibition depends on the relative concentrations of the inhibitor and substrate because the inhibitor molecules and substrate molecules are aiming for the same active site. If there are more substrate molecules than inhibitor molecules, the chance that the active site will bind substrate will be higher. Conversely, if there are large numbers of inhibitor molecules and few substrate molecules, the odds of the inhibitor occupying the active site are higher.

Thus, in competitive inhibition, addition of more substrate results in decreasing the extent of inhibition. If the substrate concentration is high enough, the inhibition can be completely overcome.

The approach to reversible enzyme inhibition that is typically attempted in drug design is competitive inhibition. Because competitive inhibitors bind the active site, such drugs can be designed to resemble the substrate of the target enzyme. Mimicking the substrate simply involves some clever chemistry.

By contrast, reversible inhibition that is noncompetitive or uncompetitive is not commonly used medicinally. The inhibitors used in these mechanisms do not resemble the substrate and do not bind to the active site, so it's harder to figure out what such an inhibitor should look like.

READING

Robinson, "Enzymes."

QUESTIONS

1 Cells are the ultimate control freaks, and their regulation of enzymes is one manifestation of that. Cells efficiently regulate pathways by controlling one or 2 enzymes of a pathway and then controlling the entire pathway. How is that possible? Why doesn't it make sense for the cell to control the activity of all the enzymes of a metabolic pathway? (Hint: Some say that cells actually *do* regulate all enzymes in a pathway.)

2 A compound called PALA resembles ATCase's normal substrate and is a potent inhibitor of the enzyme. Tiny amounts of PALA, however, actually activate the enzyme. Speculate on a possible mechanism by which this might occur.

3 A person has a mutation that inactivates his or her alpha-1 antitrypsin. Speculate on the person's health outcomes.

CLICK HERE TO SEE THE ANSWERS.

09 FATTY ACIDS, FATS, AND OTHER LIPIDS

Lipids—a varied group that includes fats, oils, waxes, steroids, hormones, and even some vitamins—are essential for life. Many lipids lack nitrogen and are built with the same 3 elements as carbohydrates; others do include nitrogen, meaning that they build using the same 4 basic elements as proteins. But what unites all lipids and defines them as a distinct group is that they are molecules with at least one part that is hydrophobic and that overall are not hydrophilic.

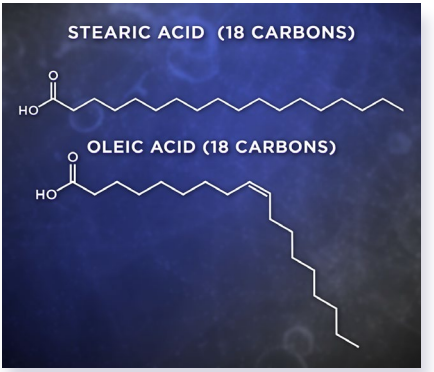
Fatty Acids

Lipids do not play in the watery neighborhoods that other biomolecules do. Some lipids, such as fats and cholesterol, are fully phobic about water. Others, such as membrane lipids, are ambivalent—a part of them fears water, but another is drawn to it.

Unlike proteins, carbohydrates, and nucleic acids, lipids are not polymers. Their molecules are also smaller. Most fats, for example, have molecular weights of less than 1000 grams per mole, whereas even a small protein has a molecular weight of 64,000 grams per mole.

Fatty acids are the smallest, simplest lipids. They consist of a single chain of 2 to 28 carbons and are amphiphilic: One end of a fatty acid, called the head, has a hydrophilic carboxyl group, while the remainder of the molecule, called its tail, is hydrophobic, being made up of only nonpolar carbon/hydrogen bonds.

Almost 2/3 of the dry weight of the human brain is made up of lipids.



Fatty acids come in 2 forms: saturated and unsaturated. Saturated fatty acids, such as stearic acid, have long, straight hydrocarbon tails with no double bonds between carbons.

By contrast, unsaturated fatty acids, such as oleic acid, have one or more carbon double bonds in their tails.

Fatty acids with just one double bond, such as oleic acid, are called monounsaturated fatty acids. Those with more than one double bond, such as linoleic acid, are called polyunsaturated.

For oleic acid, which is the leading component of olive oil, the double bond between the carbons has the 2 hydrogen atoms on the same side of the bond. This sort of double bond where hydrogens are on the same side is called a cis double bond. Cis bonds create kinks in fatty acids.

But if the 2 hydrogens are on opposite sides of the double bond, it is called a trans double bond, and the fatty acid is described as a trans fatty acid. Trans fatty acids are found in small amounts in nature, but most of them are created by the industrial production of food. The process called partial hydrogenation uses hydrogen to convert some of the double bonds in fatty acids to single

bonds. This increases the saturation, which raises the melting point. This was used originally to make vegetable shortening and margarine as substitutes for more heavily saturated fats, such as butter or lard.

Unfortunately, trans fatty acids, whose unsaturated bonds are in the trans configuration, are by-products of partial hydrogenation. These unsaturated bonds lack a kink, and it leads to their association with cardiovascular problems.

But that connection is the reason why the US FDA required labeling of trans fats beginning in 2006 and announced a ban on hydrogenated oils in 2018 for use in packaged foods and restaurants. The FDA estimates that banning hydrogenated oils will prevent 20,000 heart attacks and 7000 additional deaths from heart disease each year in the US.

Triacylglycerols

Our bodies can make most, but not all, of the fatty acids we need. Two classes of polyunsaturated fatty acids we cannot make are omega-3 and omega-6 fatty acids, and this makes them essential to get from the food we eat.

When the doctor checks your cholesterol, the lipid profile includes a number for your triglycerides, which is the amount of fat present in your blood.

Whether cis or trans, fatty acids are rich energy sources. In the body, they are typically bound to a sugar alcohol called glycerol, which has 3 carbons and 3 OH groups. A more general name for any fatty acid is acyl, so 3 fatty acids linked to a glycerol makes a triacylglycerol. An older technical term still in widespread use—triglyceride—means the same thing.

Triacylglycerols are the densest energy storage form in our body. When fatty acids are attached to glycerol, their polar head groups get hidden in the bonds with the glycerol. As a result, triacylglycerols, unlike fatty acids, are uniformly fully hydrophobic, making them insoluble in water.

Think of how oil mixed with vinegar in a salad dressing eventually separates. Left to their own chemistry, fats in your body would do the same thing and separate in the blood. To prevent this, fats are stored in specialized cells called adipocytes that contain a fat/oil droplet surrounded by cytoplasm. Fat cells grow in size, to a point, as obesity increases.

In the body, fats must move from the adipocytes they are stored in to tissues that need their energy. This movement occurs in the bloodstream in special micromolecular care packages called lipoprotein complexes. These complexes, which have an outer shell of proteins, contain an inner shell of fats and other lipids. These complexes are why your blood results come back with information about your high-density and low-density lipoproteins (HDLs and LDLs)—the concentration of bundled-up care packages.

Triacylglycerols make up solid cooking fats, such as butter, as well as liquid oils, such as olive oil. The differences between these compounds arise only from the kinds of fatty acids that are bundled to the triacylglycerols.

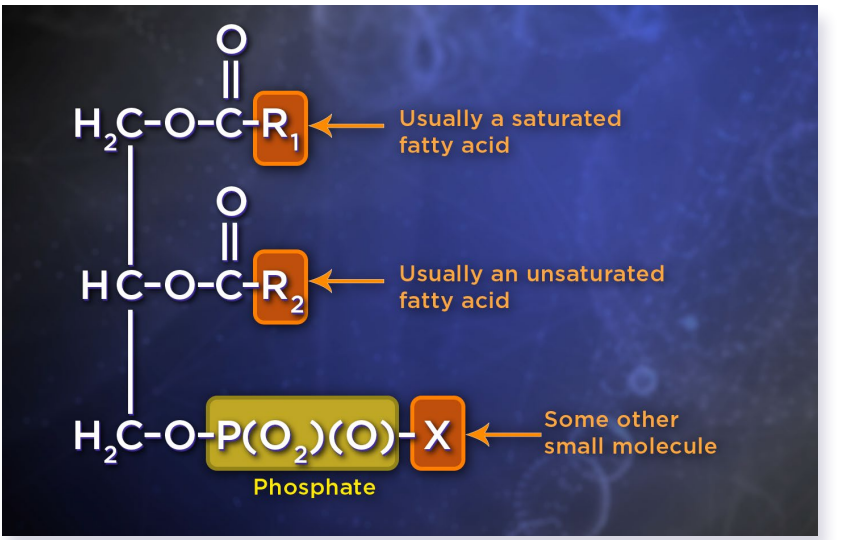
If the bundles of fatty acids are mostly saturated fatty acids, with their straight hydrocarbon tails, the fats can pack closely together. The tight packing of saturated fatty acids is why butter is solid at room temperature. Saturated fats are generally more abundant in fats of animal origin, such as butter, lard, and beef tallow—all of which are solid.

If, instead, the fatty acids bundled together are mostly the kinky, unsaturated ones, the kinks in the tails keep them from being able to line up neatly. Because of this, triacylglycerols with unsaturated fatty acids can't be packed as easily. Most plant-derived oils have a high proportion of the loosely packed unsaturated fatty acids that make them liquid. Olive oil and safflower oils are common examples.

Unlike other plant oils, coconut oil—which is plant-derived yet normally solid at room temperature—is overwhelmingly composed of saturated fatty acids, with only about 8% mono- or polyunsaturated. In fact, all common cooking oils and fats are a mixture of saturated, monounsaturated, and polyunsaturated oils; what's different are the proportions of each.

Glycerophospholipids

A related group of molecules have only 2 fatty acids attached to a glycerol, and in the place of the third one, they have small molecules linked to a phosphate at that position. These are called glycerophospholipids, or phosphoglycerides. They have a family resemblance to the triacylglycerols but differ by featuring a phosphate in place of one of the 3 fatty acids.



This small difference makes glycerophospholipids different from triacylglycerols in a crucial way. The phosphate groups, which are negatively charged, allow phospholipids to be hydrophilic at the phosphate end. They're still hydrophobic on the long hydrocarbon tails of the fatty acids, but overall they are amphiphilic—no longer fully hydrophobic.

Simple amphiphilic molecules like fatty acids arrange themselves in water so that the hydrophilic portions face outward toward water, while the hydrophobic portions face inward and associate with each other to the exclusion of water. Individual fatty acids are long, skinny molecules that form spherical structures called micelles.

However, glycerophospholipids are not skinny and cannot readily form micelles. Instead, these compounds arrange themselves in water into lipid bilayers. Interestingly, lipid bilayers form spontaneously. They don't require anything but a solution of water and phospholipids and result simply from the tendency of the hydrophilic heads of glycerophospholipids to face water while their hydrophobic tails hide from water. This property of phospholipids creates the most important cellular structure: the cell membrane, which regulates traffic in and out of cells.

Not all lipids are made from fatty acids. The most familiar of these lipids is cholesterol, which is a vital component of cellular membranes, especially brain membranes.

Sphingolipids

Another important group of amphiphilic membrane lipids is the sphingolipids. These molecules get made by joining the amino acid serine to a 16-carbon fatty acid known as palmitic acid followed by several modifications. The amine of serine provides an attachment point for a second acyl group, so sphingolipids, like glycerophospholipids, contain 2 fatty acids.

Though some sphingolipids, such as sphingomyelin, also contain a phosphate like the glycerophospholipids, most of them do not and instead have one or more sugar residues linked in place of the phosphate. Either way, most sphingolipids, like all glycerophospholipids, have an amphiphilic nature.

Fatty acids are also the starting point for a variety of other molecules that affect us in significant ways. Two essential fatty acids that our cells cannot make are linoleic acid and alpha linolenic acid, both of which are polyunsaturated 18-carbon molecules and each of which is a starting point for synthesis of other lipids.

Linoleic acid can be used by cells to make a 20-carbon polyunsaturated fatty acid called arachidonic acid. The compounds derived from arachidonic acid are a diverse group of important molecules called eicosanoids, all of which have 20 carbons, like their parent compound, arachidonic acid.

Eicosanoids include 2 important groups of molecules—one that is involved in pleasure and another that is involved in pain. Endocannabinoids play roles in mediating pleasure/reward pathways and regulation of the appetite, among other things. Arachidonic acid also gives rise to molecules that mediate pain and inflammation, a group called the prostanoids.

Fatty acids also play important roles specific to plants. Linolenic acid, one of 2 fatty acids we have to obtain in our diet, is made in plants, such as walnuts, flaxseed, soybeans, and several vegetable oils.

Fat-soluble vitamins—A, D, E, and K—are all hydrophobic lipids. Each fat-soluble vitamin has a function that is completely unrelated to the other fat-soluble vitamins. With the exception of vitamin D, they ultimately come from plants. Excessive doses are to be avoided. With vitamin A, in particular, it can be toxic.

READINGS

Kohli, et al., "Designer Lipids for Drug Delivery."
Pond, *The Fats of Life*.

QUESTIONS

1 Fatty acids are a component of fats but have a distinct chemical difference apart from fats that would allow them to travel fairly freely in the aqueous environment of the bloodstream. They don't, however, and are transported with serum albumin. What about the chemistry of fatty acids makes it so that it might not be good to have a lot of them traveling freely in the bloodstream?

2 Images drawn of glycerophospholipids in lipid bilayers tend to show them with fatty acids all nicely lined up, but that is not completely accurate. In fact, lipid bilayers can be somewhat leaky, and cholesterol helps give them integrity. What is inaccurate in depictions of lipid bilayers that may help account for their leakiness?

3 Eskimos can have difficulty making vitamin D in the winter due to their lack of exposure to sunlight. Other ethnic groups have problems making vitamin D even though they may get exposed to sunlight. Speculate on variables that might come into play.

CLICK HERE TO SEE THE ANSWERS.

10 SUGARS: GLUCOSE AND THE CARBOHYDRATES

Sugars provide us with instant energy, feed our brains, direct proteins to their destinations, and communicate the identity of our cells. But they also make some people ill when they eat ice cream and can poison our cells when present in large quantities.

What Are Sugars?

Biochemically, sugars are saccharides, a group of molecules with a particular chemical makeup—a simple carbohydrate. The general chemical formula for sugars is CH_2O, with one carbon atom for every H_2O water molecule.

In simple sugars, or monosaccharides, the number of carbons in the molecule equals the number of H_2Os. The common monosaccharides we eat—glucose, fructose, mannose, and galactose—each have 6 carbons and 6 H_2Os, giving them each the formula $C_6H_{12}O_6$. They only differ in how they're organized.

Combine any 2 monosaccharide sugars and you get a disaccharide. Glucose combined with fructose is ordinary table sugar, sucrose. Of all the sugars traveling in and around your cells, fructose is the second most abundant sugar, after glucose, and the 2 can be readily interconverted.

The USDA estimates that the average American eats about 160 pounds of sugar each year.

Simple sugars are building blocks for complex carbohydrates. If a few simple sugars, between 3 and 10, get linked together, they are called oligosaccharides. Longer polymers, typically with more than 10 simple sugars, are called polysaccharides.

In animals, the sugar polymer is a very large polysaccharide called glycogen. Plants make a mix of polysaccharides we refer to as starch. Plants don't make glycogen, but they do use sugars to make other big carbs beside starch. Very similar to starch, but with a crucial difference, is a cousin polymer called cellulose, and there are also pectins, which are polymers of modified sugars. All these molecules differ in the ways the sugars are organized in the polymer.

Why Do Sugars Appeal to Us?

We have sweet receptors on our tongues, which are linked to pleasure centers in our brains. We also have sweet receptors in our guts. Our bodies reward us for eating sugar and enthusiastically encourage us to eat more. But why?

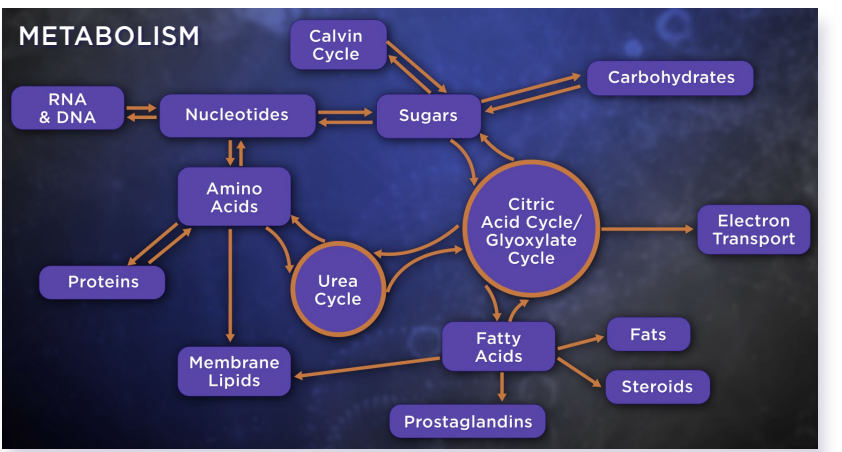
Sugars are full of energy, and eating high-energy foods is a simple recipe for survival.

Sugars are not the most energy-dense molecules in the body. That distinction goes to fat, which, gram for gram, stores more than twice as much energy as sugar. But fats are water-insoluble, so they must be packaged in complexes to travel in the bloodstream. Then, on arrival at their cellular destinations, fats have to be separated from their carrier molecules before they can be used. This all takes time, so fats are a poor choice for quick energy.

When we need immediate energy, sugars are the go-to source. Not only are sugars a source of ready energy, but their water solubility means they can easily travel in the bloodstream.

Virtually every cell in every organism uses sugar in a central metabolic pathway—a part of metabolism that's so essential we can't live without it. This is why sugar is bundled together to make glycogen in our liver and muscle cells.

But at any given time, we only have about a 24-hour supply of sugar present in our bodies, stored as glycogen. If you did not eat for a day or so, you'd use up all your glycogen. And that can be problematic, especially because our brains strongly prefer glucose to fuel activities. If blood glucose levels fall too low, the resulting hypoglycemia can be dangerous—even fatal.



How Do Cells Make Glucose?

Because glucose is so important, when supplies get low, even temporarily, our bodies start making it from other molecules. Cells in the liver and kidney do the work of making glucose to help prevent this from happening when glucose levels drop. Glucose synthesis is called gluconeogenesis, which translates to "new synthesis of glucose."

Our bodies cannot make glucose from the fatty acids in fat. But our cells can break down proteins and use the amino acids released to make glucose. This is also why people on high-protein, low-carbohydrate diets still get enough glucose.

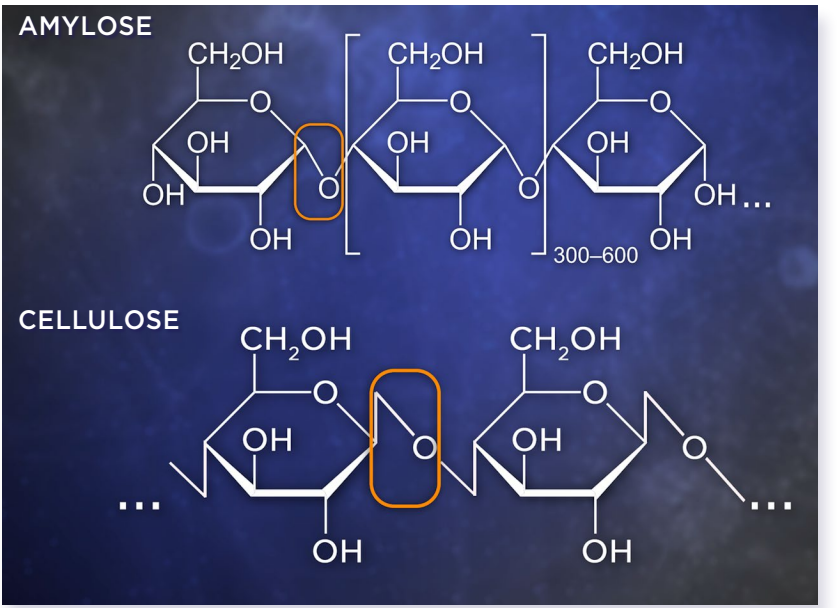
Stored chemical energy is extracted from sugar molecules by the process of oxidation. Energy can also be extracted from sugars in the absence of oxygen in a process called fermentation. Bacteria and yeasts carry out fermentation—which isn't the most efficient way to extract energy from sugar, but you do get alcohol out of it!

The term *oxidation* comes from the oxygen that is used in the process.

The story is slightly different for organisms that make their own food rather than eating it. Plants use light energy to make sugars from carbon dioxide and water. In plants, glucose is stored in large molecules of starch. The stored starch is broken down by plant cells to release the glucose needed for their activities, just as our bodies do when we eat plant starches.

Starch is made up of 2 related polymers called amylose and amylopectin. Amylose is a long linear molecule, whereas amylopectin is branched, similar to glycogen. Starch mostly consists of branched amylopectin, with smaller amounts of long amylose molecules. Exact proportions vary among the starchy foods.

Plants also use glucose to make a distinctive polymer that animal cells completely lack: cellulose, which provides structural integrity to plant cell walls. Cellulose is almost identical to the long molecules of amylose but with a subtle structural difference.



The chemical bond linking the glucose units of cellulose differs from amylose only in the configuration of carbon 1 of each glucose. Enzymes are very specific about the structure of molecules they will act on, and that minor difference means that the enzyme for digesting amylose cannot act on cellulose. As a result, we can digest and obtain glucose from amylose in starch, but not from cellulose, which passes undigested through our systems.

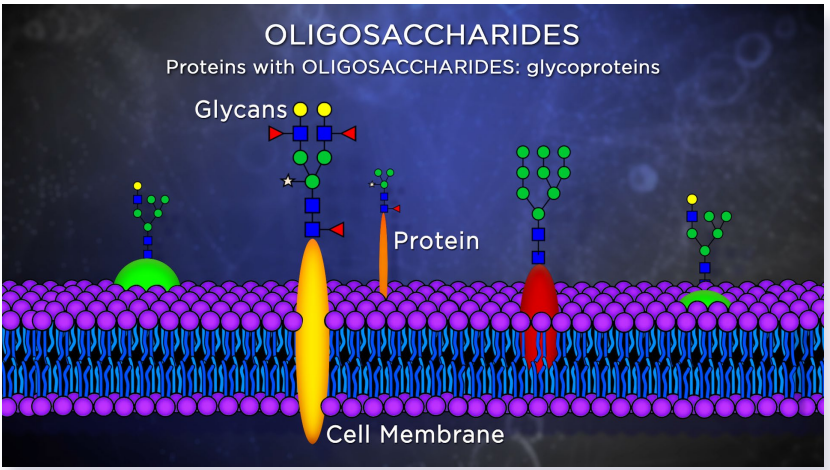
The large amount of glucose energy stored in cellulose can also be used to produce biofuels. A great deal of plant waste is not eaten by humans or most livestock but is full of cellulose. If that's broken down into glucose, the

glucose in it can be used to make ethanol by bacterial fermentation. Most of the gasoline sold in the US contains 10% ethanol, and some of that ethanol is made from cellulose.

But sugars are used for more than energy. For example, ribose contributes to the structure of RNA, and a related sugar called deoxyribose contributes to the structure of DNA.

Glucose can also be converted into other forms. For example, glucose can combine with an amine group to become a glucosamine molecule, which joins with an acetyl group to become N-acetylglucosamine. Long polymers of N-acetylglucosamine contribute to cartilage and produce chitin, which plays important roles in many organisms.

In addition to providing energy and structure, another fascinating use that cells have for sugars is establishing cellular ID. Proteins in animal cell membranes have small groups of sugars, called glycans, attached covalently to them in such a way that the sugars stick out on the outside of the cell. These are oligosaccharides that are often branched, so they look like little trees growing out of the membrane proteins.

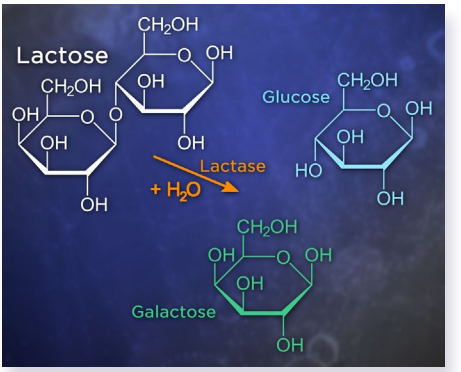


The kinds of oligosaccharide trees attached to the membrane proteins serve as a sort of cellular bar code, identifying what type of cell it is. Proteins with oligosaccharides attached to them are called glycoproteins, and each cell membrane has thousands of them.

The membrane proteins involved in cellular ID are examples of glycoproteins, but there are also glycoproteins that are involved in cellular ID—and that are not even bound to membranes, such as the mucins in mucus.

One example of glycoproteins that are involved in cellular ID are the blood group antigens or markers, which help determine blood types. Whichever kind of blood group antigens are found on your cell membranes determines whether you are A, B, or AB. If your blood group is O, your cells don't have any of the markers.

Problems Caused by Sugar



Common sugars can cause problems. For example, consider lactose intolerance. Lactose is a disaccharide sugar found in milk that needs to be digested by the enzyme lactase into glucose and galactose to be used for energy.

Human babies make lots of lactase to help them digest the lactose in milk. Early humans only drank milk while nursing as babies, and the production of lactase was turned off as children moved into adolescence. A few individuals in early human populations may have had mutations that allowed them to continue making lactase all through their lives, a condition called lactase persistence.

And once people began domesticating animals that produced milk, between 8000 and 10,000 years ago, the lactase-persistence mutations spread rapidly in those areas of the world where animals were raised for milk. Today, many of the descendants of those early dairy farmers are able to digest milk throughout their lives. In parts of the world where milk-producing animals were not commonly raised, such as much of East Asia, most individuals to this day can tolerate and break down only a small amount of lactose as adults.

This condition has led to the creation of lactose-free milk, which is typically made using the lactase enzyme, which leaves behind the glucose and galactose sugars instead of lactose.

A more serious issue with sugar arises when levels of glucose in the blood are not properly controlled, as happens in diabetes mellitus, commonly referred to simply as diabetes.

After a meal, the nutrients in the food are broken down by the digestive system into simpler molecules to be absorbed and transported by the bloodstream to various tissues of the body. Complex carbohydrates in our diet will eventually be broken down into simple sugars. The more complex the carbohydrates are, the longer it takes for digestive enzymes to act on them, and therefore the slower the sugars in them get released into the bloodstream.

We can gauge this rate by the glycemic index, a measure of the rate at which a food or drink increases the concentration of blood glucose. The scale is set with glucose itself having a score of 100. Everything else is assigned a number that reflects how fast it increases blood glucose compared to pure glucose.

Glycemic index is only one measure of sugar's impact on the body. The quantity of sugar in food, for example, is measured by the glycemic load. Foods with a very low glycemic load, like carrots or apples, often include fiber, which means insoluble cellulose and/or water-soluble pectin. In general, foods with low glycemic loads are recommended to avoid spikes in blood glucose.

High blood glucose, as found in diabetes, can damage blood vessels and cause blindness, organ failure, cognitive impairment, and cardiovascular disease. Preventing excessively high levels of glucose in the bloodstream is where the hormone insulin comes in.

The reason why ancient Egyptians could use honey as an antibiotic for wounds is because sugar at high levels is a cellular poison.

Insulin is a small protein made in specialized cells in the pancreas in response to increases in the blood glucose concentration. Hormones are molecules that are made in one part of the body and travel in the blood to another part of the body to exert their effects. It's insulin's job to bring glucose levels down into safe territory to prevent the damage that sugar can do to cells.

People with type 1 diabetes make little or no insulin. This may arise from infection or autoimmune attack on the insulin-making cells of the pancreas. The only way to deal with it is to carefully manage/monitor blood sugar levels and inject insulin regularly to help keep them normal.

In type 2 diabetes, the pancreas makes insulin, but cells don't respond to it as they normally should. This is called insulin resistance, and it can cause a roller-coaster effect with blood sugar. It appears that what causes insulin resistance in the first place is a diet high in sugar and fat, which results in the storage of fat in the liver, which in some way triggers cells to become less responsive to insulin.

Do artificial sweeteners—such as aspartame, sucralose, or saccharin—eliminate the problems associated with sugar and diabetes?

Unfortunately, artificial sweeteners still lead to high levels of blood sugar, and they even appear to have side consequences that are worse than sugar, such as disruption of the communities of bacteria that live in the gut.

READINGS

Genetic Science Learning Center, "Spotlight on Sugar," https://learn.genetics.utah.edu/content/metabolism/sugar/.

QUESTIONS

1 The lecture described how various ethnic groups tend to have lactose intolerance. Based on the lecture, speculate on how widespread lactose intolerance is among nonhuman mammals.

2 The release of insulin into the bloodstream stimulates cells to take up glucose, and as a result, blood glucose levels fall. A fad diet called grazing proposes that people should eat many small meals throughout the day. Speculate on what the grazing diet might do to insulin and hunger levels.

3 It has been proposed that the activation of sweet taste receptors is a signal to the body to begin to release digestive juices and insulin in preparation for a meal. Speculate on the effects arising from the stimulation of those receptors in people who consume diet drinks.

CLICK HERE TO SEE THE ANSWERS.

11 ATP AND ENERGY TRANSFORMATIONS IN CELLS

But how does life safely use and manage all the energy that's made in biochemical reactions? It's an incredible story of energy storage and transformation that centers on a molecule known as adenosine triphosphate (ATP).

ATP and Oxidation

Amazingly, we make and break down our weight in ATP each day! To understand how it's possible to do this, consider the sugar in a marshmallow. There are basically 2 ways the sugar in that marshmallow can be converted to energy.

- If you accidentally set fire to a marshmallow you were trying to toast, the burning sugar of the marshmallow makes impressive flames. That's oxidation, and heat is released in the process.

- When you eat an uncooked marshmallow, sugar again converts to energy, and it's still oxidation, but you do not burst into flames or start breathing fire.

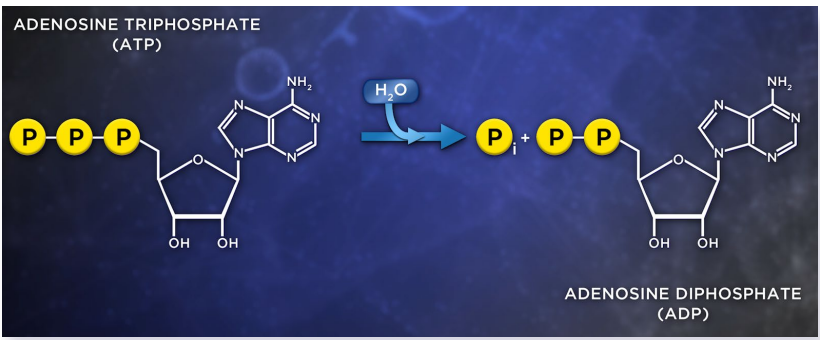
What's the difference?

What cells do during their oxidation of glucose happens in many small steps instead of one big one. The energy released in each step is correspondingly lower, so no fire results.

What's remarkable is how much of the small amounts of energy released in glucose oxidation in cells is recaptured as biochemical energy. Essentially, one form of chemical energy, glucose, is converted to another form of chemical energy, ATP, with relatively little loss as heat.

In cells, the complete oxidation of glucose requires more than 30 individual chemical reactions, and the ultimate product of these reactions is another high-energy molecule: ATP.

Going to all that trouble is worthwhile because ATP works like a battery, storing energy to be used instantaneously by cells. The fully charged battery of ATP can instantaneously discharge its energy and become adenosine diphosphate (ADP). The spent battery of ADP can be recharged to ATP. ATP is a triphosphate, while ADP is a diphosphate, so all that's going on is the removal and addition of a phosphate. This can happen again and again, using energy from food.



Our cells get energy by oxidizing food and capturing the released energy in the bonds of ATP. Why do we need so much ATP?

Cells must build and maintain complex molecules and structures. These activities require energy, whether stringing together amino acids to make proteins or carrying out the many reactions that build cholesterol for membranes.

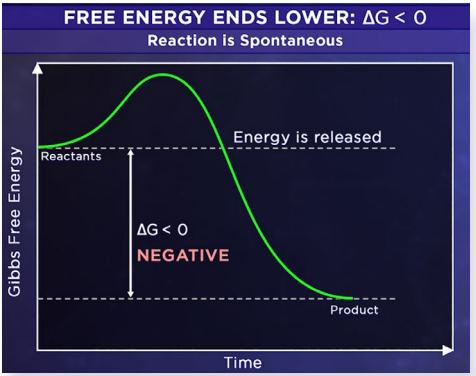
Moreover, all cells have 2 kinds of activities going on simultaneously: the breakdown of food to obtain necessary building blocks and energy, called catabolism, and synthetic pathways that use the building blocks and energy from catabolism to create the molecules the cell needs, called anabolism. Some reactions are used for catabolism in one direction and anabolism in the other.

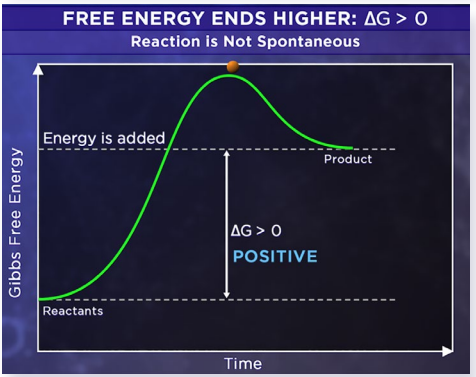
Using Thermodynamics

Recall that Gibbs free energy (G) is a measure of the amount of energy available to do useful work in a process. In biochemical reactions, the work being done is making or breaking chemical bonds. As biomolecules react with each other, there is a change in the available Gibbs free energy. The size of that free energy change is ΔG: If energy inputs are required for a reaction, then ΔG is positive, and if ΔG is negative, then the reaction gives away excess energy.

Imagine a reaction where the free energy is lower for the final product than for the starting molecule. This is described as moving downhill energetically.

In everyday terms, envision that a ball at the top of a hill will roll down to the bottom. The higher the hill is, the steeper the slope is, and the greater the likelihood of balls rolling down to the bottom would be.





Because the product is at a lower energy level than the starting molecule, energy is given off in the reaction. Mathematically, the Gibbs free energy change, or ΔG, is negative.

Conversely, to change a molecule that has low free energy to a molecule that has high free energy requires input of energy—just as pushing an object from the bottom of a hill to the top requires added energy. Because the object is gaining energy in this process, the change in Gibbs free energy is positive, meaning the value of ΔG is positive.

But how do we tell for a given reaction what its ΔG is? Consider a simple reaction:

$$A \Leftrightarrow B$$

When you think of a reaction going from *A* to *B*, you might imagine that it only goes one way. But it turns out that cellular reactions are never that simple: Instead, it's far more common for a reaction to go mostly one way, but also somewhat in the other direction, at the same time. Because of this, it's rare that a reaction goes until all the *A* is converted to *B*.

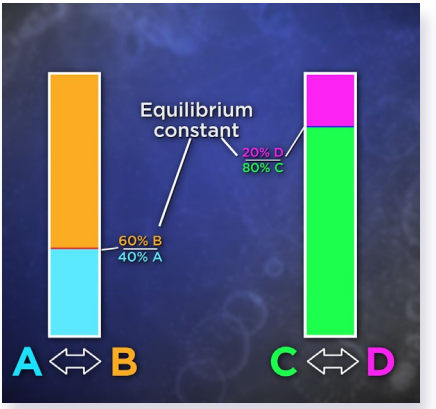
Instead, there will always, to a small extent, be conversion of *B* back to *A*. The outcome of the reaction is a matter of how much of *A* is converted to *B*, and vice versa.

Another consideration is the relative concentrations of *A* and *B* you start with. Suppose you start with only *A*. The reaction converting *A* to *B* would proceed, making more and more molecules of *B*.

What if you started with only *B* and no *A*? Because the reaction is reversible, *B* would start going to *A*. In neither case, though, would the reactions get completely converted into *A* or into *B*. They will stop at a point somewhere in between. In fact, when you measure the ratio of *B* to *A* at the point where they stop, the value of that ratio will be the same no matter what amount of *A* or *B* you started with.

The point at which the ratio of *B* to *A* stops changing is the equilibrium, and the ratio of *B* to *A* at that point is the equilibrium constant. There is one such equilibrium constant for every reaction, and the equilibrium is almost never when the 2 concentrations are equal.

The equilibrium for one reaction might be at 60% *B* and 40% *A*. Another reaction, *C* ⇔ *D*, might be at equilibrium at 20% *D* and 80% *C*.



Each reaction has its own equilibrium constant. It doesn't matter what the starting concentrations of the molecules in the reaction are. At equilibrium, the ratio of *B* to *A* or *D* to *C* will always have the same value of the equilibrium constant that is specific to that particular reaction.

Suppose that for the $A \Leftrightarrow B$ reaction, the equilibrium constant is 70% B and 30% A. What will happen if you start with equal amounts of B and A (5%50)? In that case, the reaction will go forward until it reaches 70% B and 30% A because reactions always move toward equilibrium.

If you start with 90% B and 10% A, the reaction will go backward until it reaches 7%30. So, the equilibrium constant and the amounts of A and B you start with determine the direction the reaction goes.

ΔG offers an easy way to predict whether a reaction will go forward or backward or is already in equilibrium.

When a reaction is at equilibrium, its ΔG will be equal to zero. When ΔG is negative, energy is released, and the reaction moves forward. If ΔG is positive, then energy would have to be added for the forward reaction, so it moves backward.

Whether a cellular reaction goes forward or backward is governed by thermodynamic laws that cannot be violated. Cells perform the remarkable task of making all the ATP our bodies need within statutes defined by the Gibbs free energy equation.

Fortunately, altering concentrations of reactants and products is all that's needed to move a reaction in a desired direction, and cells are very good at changing their concentrations.

Determining ΔG

ΔG tells us a lot about reactions. How do we determine the value of ΔG? This turns out to be related to the equilibrium constant.

For temperature, scientists use the Kelvin temperature scale, which puts 0 K as absolute zero, so 25°C is 298 K.

Suppose you start with a 1-liter solution of *A* and *B*, 50% of each, at standard conditions for a reaction: a temperature of 25° Celsius (298 kelvin) and an atmospheric pressure of 1 atmosphere.

You know that every reaction is going to move toward equilibrium. You simply need to determine if there is energy released by the reaction ($-\Delta G$, going downhill) or absorbed by the reaction (ΔG, going uphill).

First, you'll measure the value of the energy change in the reaction. Eventually, the reaction will reach equilibrium and the energy will not change. At equilibrium, you'll measure the concentration of *B* and *A*, because their ratio will give you the equilibrium constant.

So, there are 2 conditions, and they correspond to 2 energies: There's a starting condition not at equilibrium, giving off or absorbing energy, and there's an ending condition at equilibrium, with no energy change.

In essence, the starting condition must be neutralized by the forces making the ending condition. In other words, the starting condition gives off energy ($-\Delta G$) or absorbs energy (ΔG). The forces giving rise to equilibrium must be countering that negative or positive energy by creating an opposite energy of an equal amount with the opposite charge.

In the beginning, the countering energy is equal to zero at standard conditions. That absence of countering energy is what allows the reaction to move toward equilibrium.

ΔG at start = initial energy + 0 (countering energy at standard conditions)

Because the reaction is not at equilibrium, it moves toward it. Then, it reaches equilibrium, and when that happens,

ΔG at equilibrium = 0 = initial energy + countering energy at equilibrium.

Thus, the countering energy at equilibrium must be opposite in charge and equal in magnitude to the initial energy to bring ΔG to zero.

The initial energy was the energy released at the beginning under standard conditions. Biochemists call the initial energy release the change in standard Gibbs free energy, or $\Delta G^{\circ\prime}$, where $^\circ$ indicates that the measurement is at pH 7, as would exist in a cell, and the $'$ is used to distinguish it from ΔG.

Each reaction has its own distinctive standard Gibbs free energy change, and it remains a constant for that reaction.

The countering energy at equilibrium is the negative, or opposite in sign, of the initial energy.

The following equation captures all of this. The key idea is that ΔG changes depending on the concentrations of product B and reactant A.

$$\Delta G = \Delta G^{\circ\prime} + RT \ln \{[B]/[A]\}$$

And because there can be more than one product created by more than one reactant, $\{[B]/[A]\}$ can be simplified to [products]/[reactants].

In words, this means the following:

- ΔG (the change in Gibbs free energy) is the actual change in free energy for a reaction, which gives the direction of the reaction. A negative value means the reaction goes forward; a positive value means more energy would be needed, so the reaction on its own must go backward. Reactions at equilibrium have $\Delta G = 0$ and have no further energy change.

- The G under standard conditions is always the same for a given reaction, so it's a constant in the calculation of actual ΔG.

- The only variables in the equation that cells can change are the concentrations of reactants and products.

- R is simply a constant.

- T is the temperature in Kelvin.

- The term $\ln\{[B]/[A]\}$ is positive when $B > A$, 0 when $B = A$, and negative when $B < A$.

So, the ratio of the product B to the reactant A is the main term influencing the whole reaction. Thus, the countering energy is positive if $B > A$ and negative if $B < A$.

And as the concentration of B increases, the $\ln\{[B]/[A]\}$ term becomes more positive. And the more positive it is, the more positive the overall ΔG becomes. But if more and more energy is required, then the reaction will tend to move backward. Conversely, increasing reactant A will eventually make B/A smaller than 1, which means $\ln\{[B]/[A]\}$ will become negative—because it's giving away energy—and the reaction will be favored to move forward.

Cells exploit these principles all the time, altering the concentration of reactants and products to favor a reaction going one way or the other. If a cell manipulates the concentration of either products or reactants, the ΔG will change—from negative to positive, or vice versa—simply by altering the concentrations of products and reactants.

Changes in ΔG are the secret to how cells control the direction of reactions. It is, in fact, the only control cells can exert over the direction of a reaction.

READING

Jonsson, et al., "Essential Chemistry for Biochemists."

QUESTIONS

1 Le Chatelier's principle states that when the equilibrium of a system is upset, it will adjust to reestablish the equilibrium. Imagine you have a reaction $A \Leftrightarrow B$ at equilibrium and you disturb the equilibrium by doubling the amount of B. Describe how the ΔG equation illustrates Le Chatelier's principle.

2 Many athletes take creatine supplements because they believe it will help them improve their athletic performance. Imagine someone taking a large quantity of creatine just before a sprinting race. Using the creatine equation below, predict the effect of the creatine on ATP production as the race proceeds, compared to someone who did not take creatine.

$$\text{creatine} + \text{ATP} \Leftrightarrow \text{creatine phosphate} + \text{ADP}$$

CLICK HERE TO SEE THE ANSWERS.

12 BREAKING DOWN SUGARS AND FATTY ACIDS

A metabolic pathway is a series of biochemical reactions in which the product of one serves as the substrate for the next. Each reaction is catalyzed by an enzyme, and each molecule is called an intermediate.

Metabolic pathways are essentially identical for humans and manta rays—or any other organism. Consequently, if you learn a metabolic pathway for one organism, you've already learned a pathway that's in almost every other organism.

Sugar Glycolysis

Cells act on glucose in 10 steps in glycolysis. Glucose contains 6 carbons and gets broken down to 2 identical molecules of pyruvate, each with 3 carbons. Other sugars can be oxidized here as well. In the process, glucose is oxidized. This involves a release of energy, which is partly stored in the phosphate bonds of ATP.

The glycolysis pathway starts in the cytoplasm of cells.

1 A phosphate from ATP is transferred onto carbon 6 of glucose to create a molecule known as glucose 6-phosphate (G6P). This requires ATP instead of producing it. To make ATP, we have to use some ATP. Enzymes that put phosphates onto molecules are kinases, and this one is called hexokinase. Adding a phosphate onto glucose makes it negatively charged and therefore makes it not able to easily exit cells on its own.

2 There is a simple rearrangement of G6P. In other words, nothing is added or taken away; the atoms in the G6P are just rearranged to make fructose 6-phosphate (F6P).

3 F6P is the substrate for reaction 3, a crucial control point for the pathway. A second ATP is invested to add a second phosphate onto F6P, this time at carbon 1, making fructose 1,6-bisphosphate (F1,6-BP). The enzyme catalyzing the reaction is called phosphofructokinase (PFK). This enzyme is capable of controlling the entire pathway.

4 Catalyzed by the enzyme called aldolase, this step of glycolysis splits F1,6-BP into 2 pieces of 3 carbons each. These 2 pieces are not identical but instead are isomers, meaning that one of them can be rearranged into the other.

5 The piece called dihydroxyacetone phosphate (DHAP) gets converted to its isomer, glyceraldehyde 3-phosphate (G3P).

6 Glyceraldehyde 3-phosphates are converted to 1,3-bisphosphoglyerates. This involves the addition of a phosphate. For this to happen, the G3P aldehyde is converted to an acid, which is an oxidation. And in that process, the molecule being oxidized loses 2 electrons. It gives those electrons to NAD^+, which accepts them plus a proton and makes NADH. In the same

Starting at step 6, glycolysis becomes a process of energy generation, where ATP is made.

reaction, the acid group created by the oxidation reaction gets a phosphate linked to it, yielding 1,3-bisphosphoglycerate. This time, the phosphate is not from ATP, but rather a free phosphate added onto the acid. The energy for doing this comes from the oxidation of this same step; this is why ATP was not needed to add the phosphate.

7 The formation of 1,3-bisphosphoglycerate kicks off the last reactions of glycolysis, where ATP begins to get made! First, a phosphate gets transferred away from each of the 1,3-bisphosphoglycerates to ADPs, resulting in 2 ATPs! And 2 molecules called 3-phosphoglycerate are left behind. The energy invested in steps 1 and 3 has been recovered.

8 The molecules of 3-phosphoglycerate undergo rearrangement.

9 The molecules of 3-phosphoglycerate form a molecule named phosphoenolpyruvate (PEP), which is the substrate for the final step in glycolysis: the big bang.

10 The big bang reaction, which is catalyzed by pyruvate kinase, gets its nickname from a very large energy release that results. In the reaction, 2 PEPs are converted to 2 pyruvates, and 2 ATPs are produced. There is enough energy released in the reaction to almost produce 2 more ATPs, but not quite. Energy that is released during reactions that is not captured or used is lost as heat.

Glucose has been broken down into 2 pyruvates, but what's been achieved? The products of glycolysis include 2 ATPs (after subtracting the 2 input ATPs), 2 NADHs, and 2 molecules of pyruvate. The payoff is double the investment.

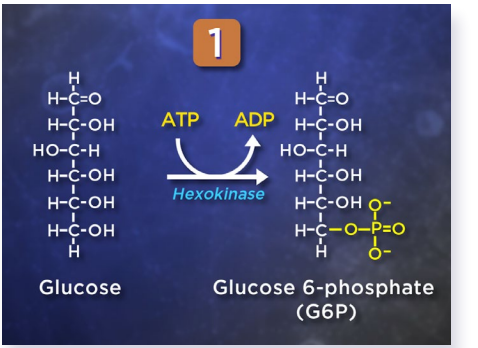
Moreover, the NADH made in glycolysis can be used to produce more ATPs, and the 2 molecules of pyruvate can be further oxidized in the citric acid cycle. In total, there is enough energy in one glucose molecule to give 32 to 38 ATPs. But glycolysis mostly just sets the stage for the main act of energy extraction.


Glucose
1
Hexokinase
ATP
ADP
Glucose 6-phosphate
2
Phosphoglucose isomerase
Fructose 6-phosphate
3
Phosphofructokinase
ATP
ADP
Fructose 1,6-bisphosphate
4
Aldolase
Glyceraldehyde 3-phosphate
5
(2 molecules)
Glyceraldehyde-3-phosphate dehydrogenase
2 NAD+ + 2PI
2 NADH + 2H+
6
1,3-Bisphosphoglycerate
(2 molecules)
7
Phosphoglycerate kinase
2 ADP
2 ATP
3-Phosphoglycerate
(2 molecules)
8
Phosphoglycerate mutase
2-Phosphoglycerate
(2 molecules)
9
Enolase
2H2O
Phosphoenolpyruvate
(2 molecules)
10
Pyruvate kinase
2 ADP
2 ATP
Pyruvate
(2 molecules)

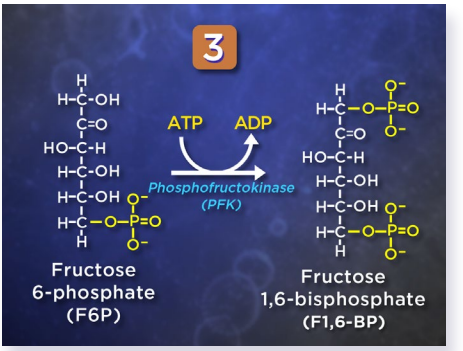
Meanwhile, the 10 steps of glycolysis do more than get the breakdown of glucose started. Every pathway is part of a network, and intermediates of glycolysis are used in other metabolic pathways. Specifically, glycolysis creates building blocks that can be tapped for making DNA, other sugars, amino acids, and more. Having multiple steps also provides several points at which the flow of intermediates through glycolysis can be adjusted.

Metabolic pathways must be regulated.

- Step 1 of the pathway, catalyzed by the enzyme hexokinase, makes G6P from glucose. However, the enzyme is inhibited by high levels of the product, G6P. This feedback keeps cells from piling up large amounts of phosphorylated glucose.

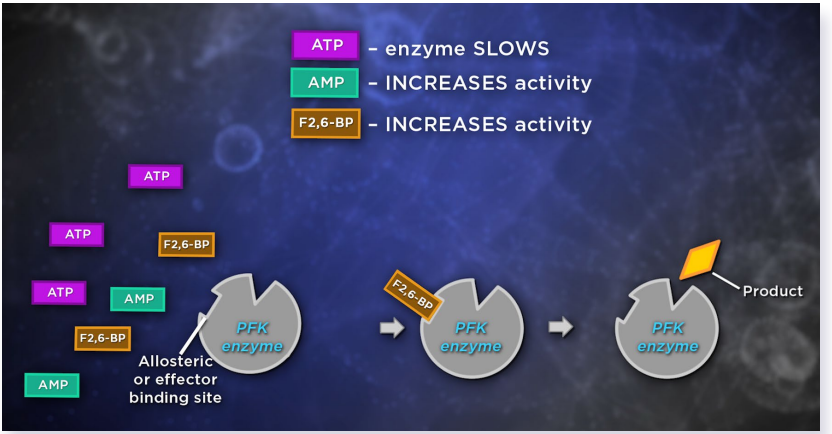


- Step 3 of glycolysis—in which the enzyme phosphofructokinase (PFK) adds a phosphate onto F6P, making F1,6-BP—is allosterically regulated by 3 different molecules: ATP, a low-energy cousin called adenosine monophosphate (AMP), and a double-phosphate molecule known as fructose 2,6-bisphosphate (F2,6-BP). Allosteric regulating molecules bind to an enzyme and alter its activity. When any of these molecules binds to PFK, the resulting change in the enzyme's conformation dials the activity of the entire pathway up or down.



An abundance of ATP indicates that the cell has plenty of energy, so when it binds to PFK, the enzyme's activity slows down. Large amounts of AMP, on the other hand, signal low energy and a need for glycolysis to make ATP. The binding of AMP to PFK increases PFK activity.

- F2,6-BP is made in response to insulin, which stimulates cells to take up glucose following a meal. When this happens, cells must run glycolysis to break it down. So, F2,6-BP increases PFK enzyme activity and speeds up glycolysis.



Regulation affects the aldolase reaction that splits glucose into DHAP and G3P, but not through control of its enzyme. This reaction would move backward, not forward, under standard conditions. But regulation of other enzymes in the pathway helps ensure that conditions are not standard when glycolysis needs to run. It does this by ensuring that the substrate F1,6-BP is in high concentration and that the products' concentrations are kept low.

Remember that cells must adjust molecules' concentrations to control reactions.

It's easy to keep the levels of the substrate F1,6-BP high as long as there is glucose flowing through glycolysis, and that happens when PFK is activated.

For products, cells decrease their concentrations by quickly converting them into something else. This is not a simple 1-step process, but it is a clever one that relies on the connectedness of reactions in a pathway. It involves the substrate for the reaction, F1,6-BP, and the final enzyme of the pathway, pyruvate kinase.

In addition to glucose, glycolysis is also important for the metabolism of other sugars, such as fructose, mannose, and galactose. Each of these sugars is readily converted into an early intermediate of the glycolysis pathway and can get made into pyruvate.

Just as there is more than one way to enter the glycolysis pathway, there are also alternatives about what happens at the end. In particular, when there is no oxygen, there are alternative exits from glycolysis.

Lactic acid was once thought to be the culprit in the muscle soreness that develops after heavy exercise. But studies have shown that soreness results primarily from microscopic damage to muscle fibers.

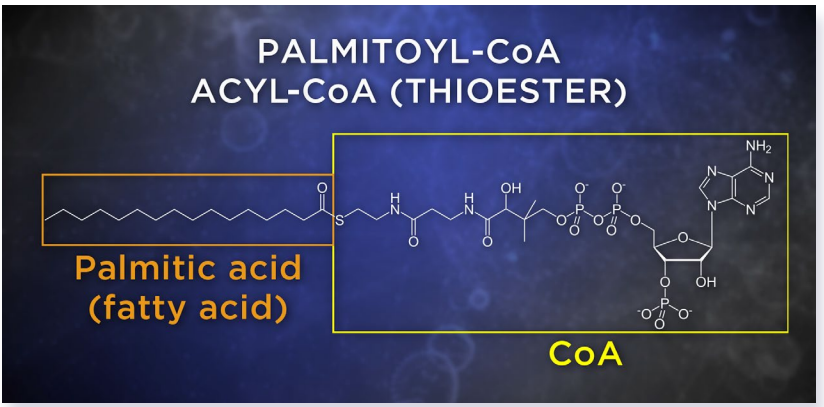
One such exit is lactic fermentation. It produces lactic acid, which is used to make foods like yogurt or sauerkraut. It is also the path used by our muscle cells during prolonged, vigorous exercise, when the oxygen supply to muscles is unable to keep up with the demand.

A second fermentation pathway that other yeasts and bacteria run in the absence of oxygen produces alcohol from pyruvate. This is the exit from glycolysis that wine and beer makers depend on. In this 2-step process, pyruvate is first converted to acetaldehyde, and carbon dioxide is released in the process. The acetaldehyde is then turned into alcohol using electrons from NADH, thus regenerating NAD^+.

Fatty Acid Oxidation

Sugars are not the only sources of energy in our bodies. In fact, they are not even the primary energy storage molecules. Those would be fats, also known as triglycerides or triacylglycerols. A triacylglycerol is a molecule that binds 3 fatty acids (3 acyls) to a glycerol. To get energy out of a fat, the acyls in it must first be cleaved from the glycerol they are attached to in a reaction catalyzed by lipases. These acyl fatty acids are a primary energy source for many tissues, including heart muscle.

The fatty acid acyls enter cells with the help of membrane transport proteins. All entering acyls first get linked to a common cellular "handle" known as coenzyme A, creating an acyl-CoA thioester. It is these acyl-CoA molecules that undergo the subsequent steps in fatty acid oxidation.



In eukaryotes, 2 cellular organelles are the controlled sites for fatty acid oxidation: the peroxisomes and the mitochondria. In plants, fatty acid oxidation is carried out entirely in the peroxisomes, which are organelles specialized for the oxidation of substances. In animal cells, including our own, peroxisomes merely start the job if the fatty acid has 22 carbons or more, whereas fatty acid chains that are shorter go straight to the mitochondria.

Peroxisomes in our cells cannot completely oxidize those long fatty acids; they merely trim the very long chain of fatty acids to a more manageable size and then export the smaller pieces to be dealt with by the mitochondria.

Most fatty acids of animal cells get oxidized in mitochondria.

- Upon arrival at mitochondria, acyl-CoAs bearing fatty acid chains of up to about 20 carbons swap out their CoA for a molecule called carnitine to create acylcarnitine. This switching of partners is temporary; acylcarnitine, but not acyl-CoA, has a transporter that will let it into the mitochondria. Once inside, the carnitine is swapped back out for CoA to recreate acyl-CoA, which is what enzymes involved in oxidizing fatty acids look for.

Deficiency of carnitine has health consequences, particularly for tissues like heart muscle that are heavily dependent on fatty acids for energy.

- Both in the peroxisomes and the mitochondria, fatty acid oxidation happens in cycles of 4 reactions, with each cycle yielding a fatty acid chain that is shortened by 2 carbons. This continues until ultimately only acetyl-CoAs are left.

- The oxidation of fatty acids is also known as beta oxidation because all reactions involve the fatty acids' beta carbon (2 carbons in from the carboxyl group). The first enzyme involved in beta oxidation is acyl-CoA dehydrogenase, and it catalyzes the removal of 2 protons and 2 electrons from the bond between carbons 2 and 3 of the fatty acid. These are donated to the electron carrier FAD, creating $FADH_2$.

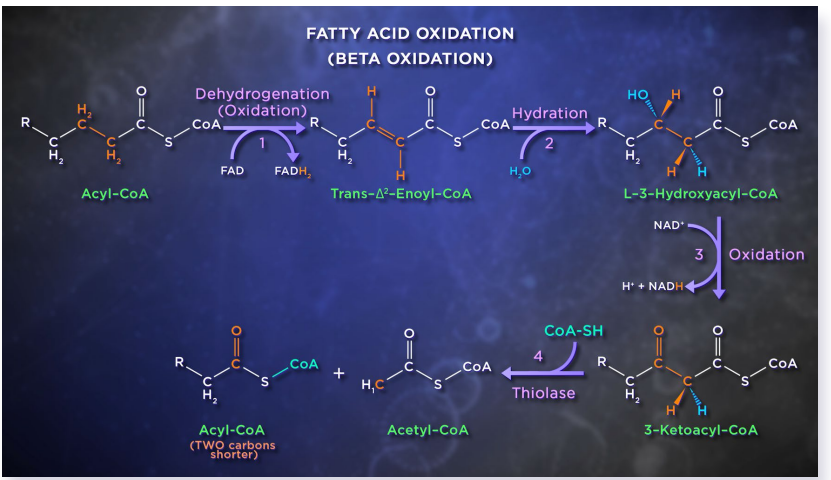
There are 3 remaining reactions in fatty acid oxidation:

- Water is added first, with the hydroxyl getting added to carbon 3. This removes the double bond.

- The next reaction oxidizes the hydroxide that was just added and generates an NADH.

- A 2-carbon chunk known as acetyl-CoA splits off from the fatty acid chain. The enzyme catalyzing this step is thiolase. The remaining acyl group from which the acetyl-CoA was removed gets attached to another coenzyme A, creating an acyl-CoA that is 2 carbons shorter. This becomes the substrate for another round of oxidation. Each time the pathway is repeated, it chops off an acetyl-CoA, so the fatty acid gets 2 carbons shorter with each round, until the final 4-carbon-CoA molecule is split into 2 acetyl-CoAs.

At this point, fatty acid oxidation is done, at least for fatty acids that had an even number of carbons.



Oxidation of fatty acids with an odd number of carbons, though, yields a 3-carbon-CoA molecule at the end. This molecule is converted into a 4-carbon molecule known as succinyl-CoA, which gets oxidized in the citric acid cycle. The enzyme that creates succinyl-CoA needs vitamin B_{12} as a coenzyme to help with catalysis.

Based on its role in energy production, vitamin B_{12} is often added to energy drinks targeted at young people. There is no credible evidence, though, that the added vitamin B_{12} plays any role in increasing energy. Excess vitamin B_{12} is simply excreted.

READINGS

OpenStax, "Carbohydrate Metabolism," https://opentextbc.ca/anatomyandphysiology/chapter/24-2-carbohydrate-metabolism/.

Rogers, *The Science of Booze*.

QUESTIONS

1 The molecule 2,3-BPG, which can bind to hemoglobin and favor the release of oxygen, is a by-product of glycolysis, although it is not directly made in the pathway. When muscle cells are working heavily, they run low on oxygen, and when this happens, they rely on fermentation, which requires much more glucose and glycolysis than when oxygen is abundant. Describe how this complementary system of 2,3-BPG production and oxygen release serves the body's needs.

2 NAD^+ is an important substrate for the glyceraldehyde 3-phosphate dehydrogenase reaction. It is produced readily when oxygen is abundant, but when oxygen is low, cells use fermentation as an alternate means of making NAD^+. A by-product of fermentation in muscle is lactic acid. Describe how its production will affect oxygen release based on the Bohr effect, which is that protons and carbon dioxide affect hemoglobin on binding by favoring the release of oxygen.

3 The metabolism of galactose normally proceeds through glycolysis after galactose is converted to glucose. People with deficiencies in the enzymes necessary to convert galactose to glucose can readily form cataracts as a result of a galactose-related crystal that can form in the eyes. Describe a dietary modification that might be recommended to someone with this deficiency.

CLICK HERE TO SEE THE ANSWERS.

13 METABOLISM MEETS AT THE CITRIC ACID CYCLE

Sugars, fats, and amino acids all converge on a single oxygen-dependent metabolic pathway known as the citric acid cycle, or the Krebs cycle. Even more than glycolysis, the citric acid cycle is a central metabolic pathway where everything comes together. And this one pathway is used by every cell, from the cells in bacteria to amoebas to you.

The Role of Energy Extraction

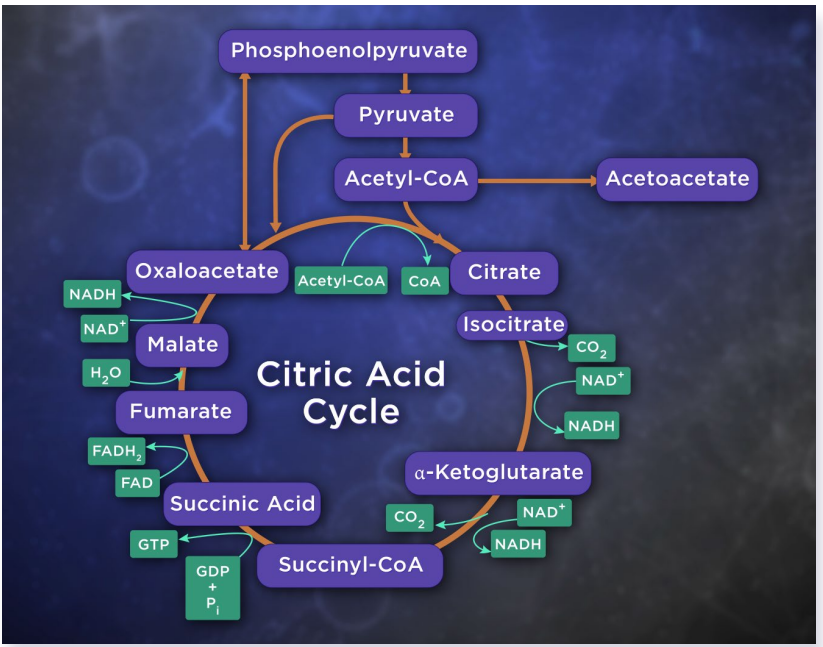
Glycolysis generates some ATP, but its main role is to deliver molecules that can be further oxidized to extract more energy. This is where the citric acid cycle comes in, providing a second stage for the efficient production of energy from food.

Like glycolysis, the citric acid cycle extracts energy and provides intermediates for other pathways. As a circle, the citric acid cycle returns to its starting point. And molecules from other pathways can enter it at multiple points and leave as needed. Just as in glycolysis, the oxidation of the intermediates in the citric acid cycle is coupled with energy capture, both in phosphate bonds and in molecules of the activated carriers, NADH and $FADH_2$.

The threshold of the citric acid cycle is arrived at with pyruvate or acetyl-CoA. Most commonly, the glycolysis pathway that oxidized sugars to yield pyruvate has been traveled, or acetyl-CoA has been chewed off in several successive oxidations of fatty acids.

Acetyl-CoAs can enter the citric acid cycle directly. But pyruvate made in glycolysis needs an extra step to get converted to acetyl-CoA, and then it, too, can enter. To start the process of entering the pathway, pyruvate needs to move from the cytosol (the cytoplasm), where glycolysis produced it, to an innermost section of mitochondria called the (mitochondrial) matrix.

To get converted to acetyl-CoA, pyruvate is first transported into mitochondria, where an enzyme complex called pyruvate dehydrogenase acts on it. Remember, the conversion of pyruvate to acetyl-CoA is necessary for the entry of pyruvate into the cycle.



The citric acid cycle has historically also been called the Krebs cycle, named for its discoverer, Sir Hans Krebs, who won the Nobel Prize in Physiology or Medicine in 1953 for his work. It's also called the tricarboxylic acid cycle because the citric acid molecule has 3 carboxyl groups.

The reason pyruvate dehydrogenase is called a complex is because it is not just one enzyme, but several enzymes that catalyze the multiple reactions necessary for converting pyruvate to acetyl-CoA. The enzyme complex uses 5 different cofactor molecules that act as helpers for the reactions. These cofactors include lipoic acid, coenzyme A, thiamine pyrophosphate, FAD, and NAD^+. The last 4 of these molecules are either vitamins or are derived from them.

The main events that the enzymes and their cofactors bring about are as follows:

- A molecule of carbon dioxide is released from pyruvate. This is a decarboxylation.

- Electrons also need to get transferred to electron carriers. This is an oxidation—loss of electrons—and is a 2-step process. First, the electrons are removed, and then they are donated to NAD^+, along with a proton, to form NADH.

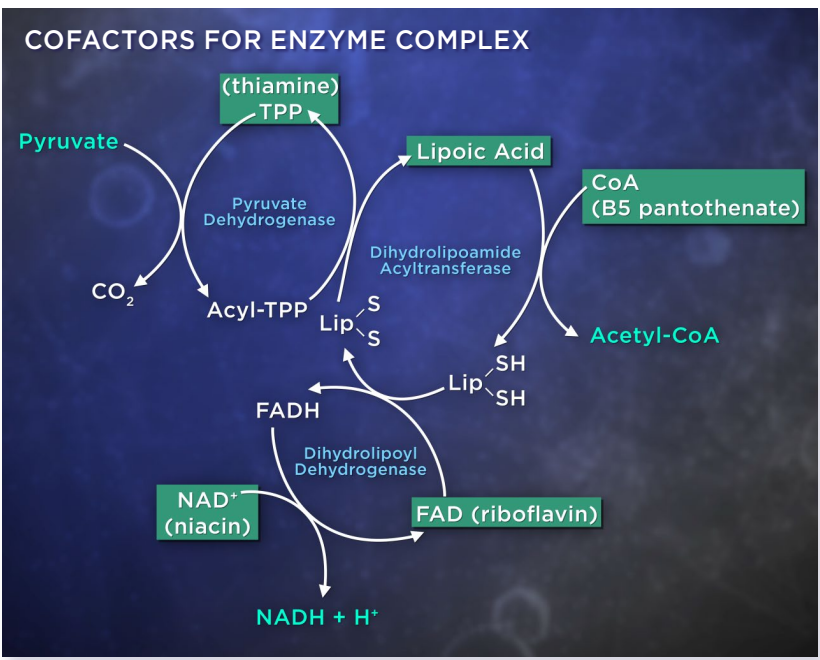
The gain of electrons is a reduction. Oxidation/reduction reactions are kind of like accounting. If you transfer money in the bank from one account to another, one account gets debited, and the other gets credited. For every loss, there is a gain, and vice versa.

What happened to the positively charged NAD^+ to make it become the neutrally charged NADH? In cellular oxidations, electrons move in pairs. Thus, NAD^+ gains 2 negative charges and one positive charge (H^+) to go from

NAD^+ to the neutrally charged NADH. Almost all cellular oxidations occur this way: Molecules lose electrons in oxidations, and electron carriers receive the electrons in reductions.

As a result of all of this, the 3-carbon molecule of pyruvate is made into the 2-carbon acetyl group in acetyl-CoA, which can then enter the citric acid cycle.

In addition to sources from sugar oxidation and fatty acid oxidation, acetyl-CoA can also be produced by breaking down amino acids and ketone bodies, which can join the citric acid cycle, just like sugars and fatty acids. And this is why these molecules can serve as energy sources when glucose supplies are low.



Reactions of the Citric Acid Cycle

By whichever means it got there, acetyl-CoA enters the citric acid cycle, which is comprised of 8 reactions. Though the pathway is a circle, the reactions are numbered starting from the entry point for acetyl-CoA.

1 Acetyl-CoA transfers its 2-carbon acetyl group to a 4-carbon molecule of oxaloacetate, which comes from the "end" of the cycle. The 4-carbon oxaloacetate and the 2-carbon acetyl group combine to produce a 6-carbon molecule called citrate, which is the name for citric acid when it's ionized. This reaction to create citrate is very favorable because breaking the bond between the acetyl group and coenzyme A releases a lot of energy. This turns out to be important when the circle is completed to make oxaloacetate.

Without the citric acid cycle, cells struggle to harvest sufficient energy from food and may die.

2 Citrate is rearranged to form its isomer, isocitrate. The enzyme catalyzing the reaction, aconitase, is the site of action of a poisonous compound known as fluoroacetate, which cells readily convert into fluorocitrate. That's a big problem because fluorocitrate is a potent inhibitor of the aconitase enzyme. Inhibiting this enzyme blocks the entire citric acid cycle.

3 The isocitrate loses a carbon dioxide and a pair of electrons in a process called oxidative decarboxylation. The loss of a carbon as carbon dioxide converts the 6-carbon isocitrate into a 5-carbon molecule, alpha-ketoglutarate. Meanwhile, the electrons from isocitrate get transferred to NAD^+, making NADH. Remember, the oxidation of isocitrate (the loss of electrons) causes NAD^+ to be reduced to NADH.

4 The same thing happens to alpha-ketoglutarate as another carbon is lost as carbon dioxide. Again, electrons are given away, and they combine with NAD^+ to form an NADH. The product of the reaction, a 4-carbon compound, links to a CoA, forming succinyl-CoA. The enzyme

here, alpha-ketoglutarate dehydrogenase, is closely related to pyruvate dehydrogenase; it uses the same 5 coenzymes and has a similar reaction mechanism.

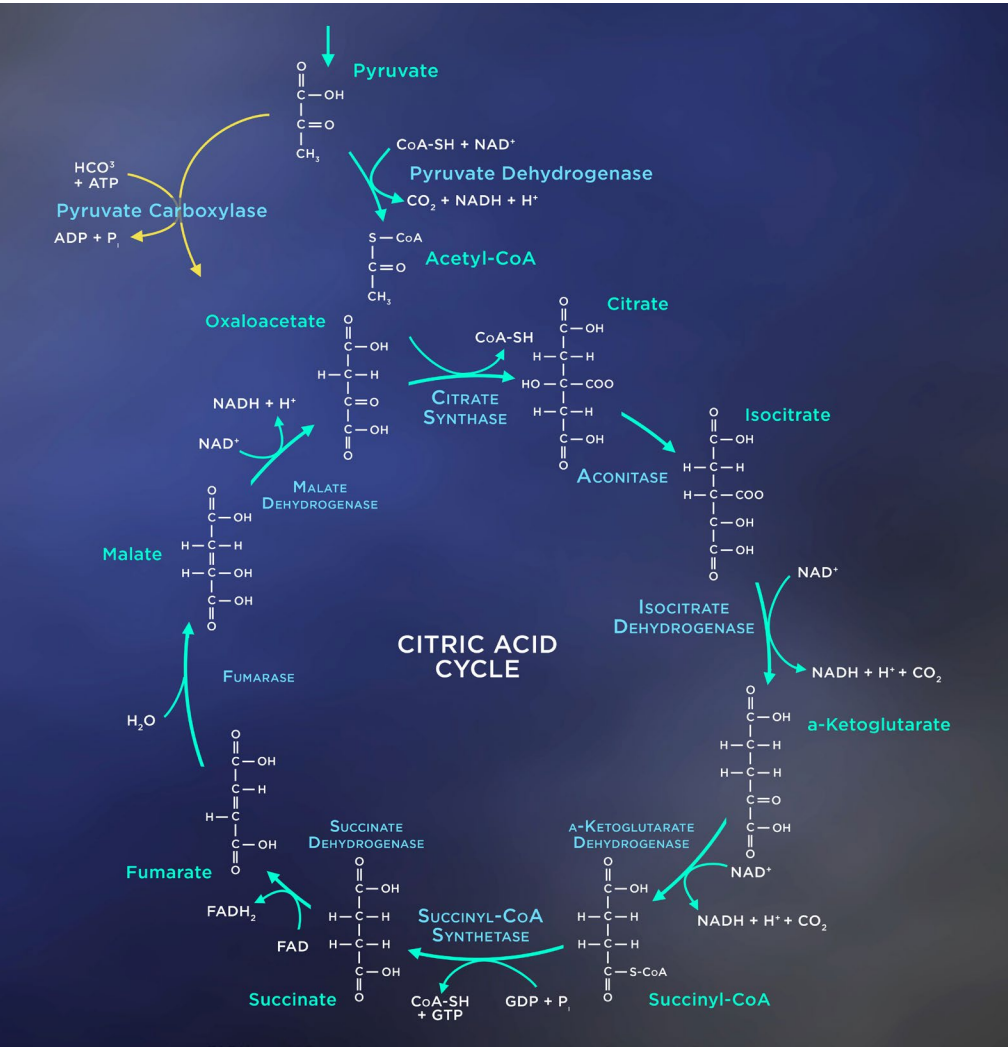
5 The coenzyme A is removed from succinyl-CoA to produce succinate. The removal of the CoA from succinyl-CoA is also a very energetically favorable reaction, just like when the CoA was released from acetyl-CoA in the first reaction. That's because the bond between succinate and coenzyme A is a high-energy one. The energy released is used to convert guanosine diphosphate (GDP) to guanosine triphosphate (GTP). The energy in GTP is equivalent to that in ATP.

6 The remaining reactions in the cycle are focused on recreating the oxaloacetate to complete the circle. In the first of these reactions, succinate is oxidized to form fumarate by the enzyme succinate dehydrogenase, which gives up electrons and protons, which are accepted by FAD to make $FADH_2$.

7 Water is added across the double bond of fumarate to form the molecule malate.

8 Malate is oxidized to oxaloacetate and, just as in many oxidation reactions, NADH is produced. The reaction, catalyzed by malate dehydrogenase, is notable for going backward under standard conditions. Fortunately, though, the cell is not under standard conditions. The product gets removed, and the reaction gets pulled forward by the citrate synthase reaction, which follows in the next round of step 1. This reaction is energetically favored and removes the product, oxaloacetate, as it reacts with acetyl-CoA to form citrate. With oxaloacetate being quickly removed to make citrate, the production of oxaloacetate can proceed.

Not only is citrate a source of the raw material needed for making fatty acids, but it also turns on the enzyme that stimulates the process.


Pyruvate
CoA-SH + NAD+
Pyruvate Dehydrogenase
CO2 + NADH + H+
HCO3 + ATP
Pyruvate Carboxylase
ADP + P
Acetyl-CoA
Oxaloacetate
Citrate
CoA-SH
Citrate Synthase
NADH + H+
NAD+
Malate Dehydrogenase
Aconitase
Isocitrate
Malate
NAD+
Isocitrate Dehydrogenase
NADH + H+ + CO2
CITRIC ACID CYCLE
Fumarase
H2O
a-Ketoglutarate
Succinate Dehydrogenase
a-Ketoglutarate Dehydrogenase
Fumarate
NAD+
FADH2
FAD
Succinyl-CoA Synthetase
NADH + H+ + CO2
Succinate
CoA-SH + GTP
GDP + P
Succinyl-CoA

With that, we've traveled around the cycle and returned to where oxaloacetate reacts with acetyl-CoA to make citrate. There are 2 turns of the cycle per glucose molecule being oxidized. That's because each glucose gives rise to 2 pyruvates, which in turn gives rise to 2 acetyl-CoAs.

The amount of energy that has been extracted through the oxidative reactions of the citric acid cycle for the 2 acetyl-CoAs is 6 NADHs, 2 $FADH_2$s, and 2 molecules of GTP. And if pyruvate was the source for the acetyl-CoAs, one NADH for each pyruvate was received from glucose for a total of 8 NADHs, 2 $FADH_2$s, and 2 molecules of GTP.

Connections to Other Metabolic Pathways

The citric acid cycle links to many pathways, feeding intermediates into some and being fed by others. In its role as a hub for many pathways, portions of the cycle can operate independently, and some molecules can enter and exit the cycle without going all the way around.

- Step 1 of the citric acid cycle combines acetyl-CoA with oxaloacetate to make citrate in the matrix of mitochondria. Lots of citrate is made when the cycle is fed abundantly with acetyl-CoA. When more citrate is made than is needed for immediate energy production, the citric acid cycle gets overloaded. When this occurs, citrate accumulates and gets transported into the cytoplasm, where the citrate gets split back into oxaloacetate and acetyl-CoA. In the cytoplasm, the acetyl-CoA is used to construct fatty acids, while oxaloacetate gets recycled back into mitochondria. When fatty acids are synthesized, they get bundled together as triacylglycerols, or fat, to be stored. Acetyl-CoA is also the raw material for making cholesterol. Citrate that is not split up in the cytoplasm serves as an allosteric effector, binding and activating the most important enzyme in fatty acid synthesis: acetyl-CoA carboxylase.

- The cycle intermediate alpha-ketoglutarate has several roles besides being an intermediate in the citric acid cycle. It can grab excess nitrogen and easily be made into the 5-carbon amino acid glutamic acid. From glutamate, it's just one step to glutamine. This is just one of at least 10 instances where an intermediate of the citric acid cycle is linked to amino acid synthesis. Many of these reactions can reverse, too. Glutamic acid and glutamine can be converted to alpha-ketoglutarate; thus, glutamic acid and glutamine enter the citric acid cycle via alpha-ketoglutarate. Alpha-ketoglutarate plays a crucial role in nitrogen metabolism.

- The next multifunctional citric acid cycle intermediate is succinyl-CoA at step 4. Succinyl-CoA allows 3 amino acids to enter the cycle, and it's also one of the building blocks used by cells in the synthesis of heme, the flat ring structure found in hemoglobin that is the site of oxygen binding.

- Fumarate and oxaloacetate also have numerous connections to other pathway molecules. Fumarate is made in both the urea cycle and in the metabolism of purine bases of DNA/RNA. It is also an entry point for 2 more amino acids: phenylalanine and tyrosine.

- Oxaloacetate is also connected to multiple pathways. It is an intermediate in glucose synthesis. It can also easily mop up an excess amino group to become the amino acid aspartate, or it can mop up an excess 2 amino groups to become asparagine. In the reverse reactions, the breakdown of those amino acids can give rise to oxaloacetate, providing entry points for 2 more amino acids.

Metabolic connections allow control of the comings and goings of metabolic intermediates. But having intermediates that interact with other metabolic pathways is not unique to the citric acid cycle. Glycolysis has that, as do other pathways. However, none are connected to so many pathways as the citric acid cycle is, with its direct connections to amino acid metabolism, the urea cycle, sugar metabolism, fatty acid metabolism, nucleotide metabolism, and the synthesis of heme.

Rapidly proliferating cells, like cancer cells, take in lots of glucose and glutamine.

Glucose is used for making ATP energy, NADPH for fatty acid/cholesterol synthesis, and ribose—which together supply vital molecules to rapidly dividing cells.

Glutamine turns into alpha-ketoglutarate to feed the citric acid cycle.

Fermentation conditions raise levels of NADH and slow or stop the forward movement of the citric acid cycle. So, reactions run in reverse, starting with alpha-ketoglutarate and resulting in the release of acetyl-CoA into the cytoplasm, where it can make fatty acids and cholesterol.

The million-dollar question is whether altering what the citric acid cycle does can inhibit the growth of cancer cells. Ongoing investigations are focused on key steps of the cycle.

READINGS

Brown, *The Energy of Life*.
Lane, *Power, Sex, Suicide*.

QUESTIONS

1 Cells need oxaloacetate to make glucose. Imagine a cell that was deficient in the enzyme aconitase and had a good supply of all of the amino acids from the bloodstream. Would it be able to make glucose abundantly? If it would, explain how. If it would not, explain the limitation.

2 Succinyl-CoA is a precursor of heme. If succinyl-CoA is taken away to make heme, though, the citric acid cycle is broken and could not continue. Cells, nevertheless, manage to keep it going. Explain how this might be possible.

3 The observation that cancer cells might run the citric acid cycle partly backward requires several things. Describe one reaction from the pentose phosphate pathway (PPP) that you might inhibit as a means of treating a cancer operating this way.

CLICK HERE TO SEE THE ANSWERS.

14 ENERGY HARVESTING IN ANIMALS AND PLANTS

The innermost, liquid-filled region of mitochondria, the matrix, is where the reactions of the citric acid cycle and fatty acid oxidation pathways occur, leading to the final phase in energy harvesting.

The Structure of Mitochondria

The structure of mitochondria is key to energy generation. Mitochondria are bounded by 2 membranes, the innermost of which is mostly impermeable to ionic substances, such as protons. This inner membrane is the site of the final stages of energy capture.

Electrons lost in the oxidative processes of glycolysis and the citric acid cycle are taken up by electron carriers. NADH and $FADH_2$ are charged batteries from having accepted electrons and a proton or 2 from oxidation reactions in the metabolic pathways you've learned about: the citric acid cycle, glycolysis, and fatty acid oxidation.

In the reactions examined next, high-energy electrons are stripped from these molecules in mitochondria and are passed through a series of increasingly electron-loving protein complexes called the electron transport chain (ETC). A single cell has thousands of mitochondria, and there are thousands of ETCs embedded in the inner membrane of each of the thousands of mitochondria. That means millions of ETCs per cell.

Two important things are happening in the inner membrane:

1. Energy is released that is used to pump protons out of the mitochondrial matrix and to create a proton gradient across the inner mitochondrial membrane. This part of the process is known as electron transport.

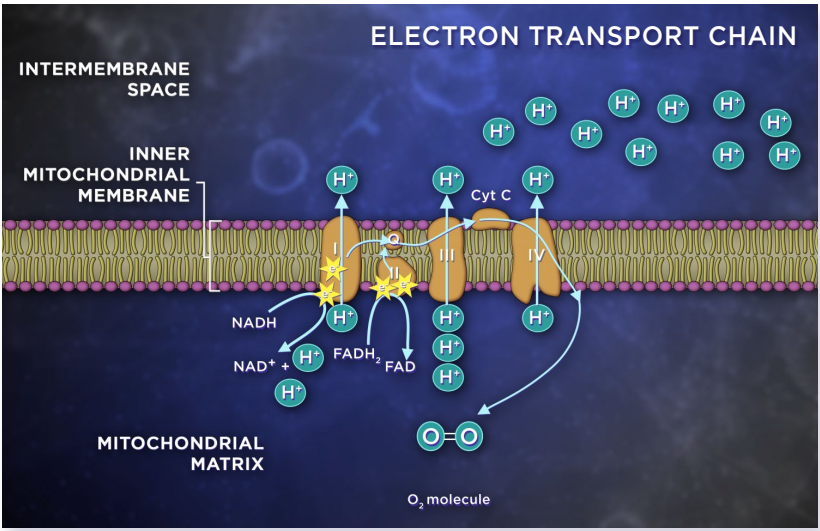
2. Protons reenter the mitochondrial matrix by passing through a protein machine called ATP synthase, which uses this energy to make ATP. This part is known as oxidative phosphorylation.

These 2 components to energy extraction are completely interdependent, or tightly coupled. The mechanism is an ancient one, used by organisms as different as bacteria and humans as well as by plants. For organisms like bacteria, which don't have mitochondria, the ETC is embedded in their cell membranes.

The Electron Transport Chain

Electrons move through the ETC in 4 steps in a process managed by 2 small shuttle systems and 4 large protein bundles called complex I through IV. Through these carriers, electrons start a journey that ultimately brings them to their final destination: molecular oxygen (O_2).

Oxygen's role as the final electron acceptor in the process of energy extraction is why it is necessary for most life.



One way that electrons first arrive at the ETC is delivery by NADH, which donates its pair of electrons to the first of 4 membrane protein complexes in the ETC called complex I. Movement of electrons from NADH through complex I releases energy.

As electrons traverse the complex, the released energy is used to pump protons out of the matrix and into the intermembrane space between the inner and outer mitochondrial membranes. This creates a difference in proton concentration across the inner membrane, with more protons on the outside and fewer in the matrix. Because protons are charged, there is a charge difference on the 2 sides of the inner membrane as well. The combined effect of differences in charge and chemical concentration of protons across the membrane is referred to as an electrochemical gradient.

After traveling through complex I, electrons move to a small molecule called coenzyme Q (CoQ). Where complex I was a humongous multiprotein complex, CoQ is a tiny shuttle that acts as a courier, ferrying electrons from complex I to their next destination in the ETC.

Complex II is an alternate entry point for electrons. It accepts electrons from $FADH_2$. The enzyme succinate dehydrogenase, which produces $FADH_2$ from the oxidation of succinate, is actually complex II. Embedded in the mitochondrial membrane, succinate dehydrogenase adds electrons and protons to FAD to make $FADH_2$, which then transfers electrons to CoQ.

Thus, electrons from NADH and $FADH_2$ are transferred via complexes I and II to CoQ. Mitochondrial inner membranes contain thousands of molecules of CoQ, so complexes I and II are not handing off their electrons to the same CoQ. The CoQ molecules ferry the electrons to complex III.

A fascinating electron carrier, cytochrome c is a tiny protein that is genetically conserved across the evolutionary spectrum, from bacteria to human beings.

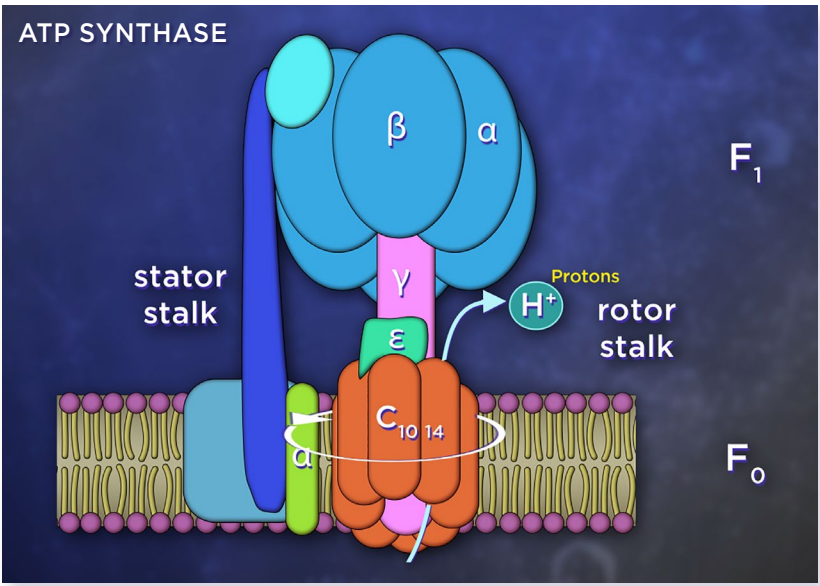
When electrons arrive at complex III, more proton pumping into the intermembrane space occurs, adding to the proton gradient. Electrons from complex III are handed off to another courier molecule known as cytochrome c, whose small size and water solubility provide excellent mobility, allowing it to shuttle electrons between complexes III and IV.

Cytochrome c transports electrons to the next complex, known as complex IV or cytochrome c oxidase. Complex IV hands the electrons over to their final acceptor, molecular oxygen. Four electrons eventually reach the oxygen molecule, and when combined with 4 protons, 2 molecules of water are created. And just as complexes I and III did, complex IV pumps protons from the mitochondrial matrix into the intermembrane space.

Oxidative Phosphorylation

Compared to electron transport, oxidative phosphorylation is simple. All of the magic occurs within a remarkable protein machine called ATP synthase, also known as complex V. ATP synthase is made up of 2 connected multiprotein assemblies, which together look like a mushroom. The base of the stalk is rooted in the mitochondrial inner membrane, while the cap, or head, projects into the mitochondrial matrix. The stalk in the lipid bilayer is called F0, while the cap reaching into the mitochondrial matrix is called F1.

The power for making ATP in ATP synthase is the high concentration of protons dammed outside the inner membrane in the intermembrane space. The ATP synthase provides a means for the protons to traverse the inner membrane through an opening in the F0 base. As the protons flow back into the matrix through the ATP synthase, they turn the stalk. This in turn causes conformational changes in the F1 head, the part of the enzyme that actually makes the ATP.



Starting with glycolysis, moving through the citric acid cycle, and ending up with oxidative phosphorylation, the balance sheet for ATP looks like this for each molecule of glucose:

- Glycolysis (glucose to pyruvate): 2 ATP (net) and 2 NADH
- Conversion of 2 pyruvate to 2 acetyl-CoA: 2 NADH
- 2 citric acid cycles: 2 GTP (= 2 ATP), 6 NADH, and 2 $FADH_2$

The total so far is 4 ATP, 10 NADH, and 2 $FADH_2$.

Electrons coming from NADH pass through 3 proton-pumping complexes: I, III, and IV. Electrons entering from $FADH_2$ only pass through 2 complexes that pump protons, III and IV, because complex II doesn't pump any protons. As a result, electrons from NADH cause more protons to be pumped across the membrane than electrons from $FADH_2$.

Back-of-the-envelope calculations indicate that between 2 and 3 ATPs per NADH and 1.5 to 2 ATPs per $FADH_2$ can be expected. This results in between 27 and 38 ATPs per molecule of glucose.

Compared to the paltry 2 ATPs from glycolysis, the subsequent oxidations and the ETC yield handsome dividends.

Oxygen: Key to Both Processes

In both the ETC and oxidative phosphorylation processes, oxidation creates activated electron carriers to pass on to the ETC. As electrons move through the ETC on their way to making water by combining with molecular oxygen, they generate a proton gradient, and protons flowing back into the matrix through ATP synthase power the production of ATP.

Oxygen is key to both processes. Without oxygen, electrons in the ETC have nowhere to go. This stops the ETC. When the ETC stops, no proton pumping occurs. And when proton pumping stops, there is no proton gradient, no spinning the rotor of ATP synthase, and no ATP synthesis.

Compounds that block electron transport, such as cyanide, are potent cellular poisons. Tiny quantities of cyanide can kill a person in a matter of a few minutes.

Oxygen supports life by accepting electrons at the end of the ETC as part of the process of making the large amounts of ATP we need to stay alive. But as much as we need it, oxygen has a dark side. If all goes smoothly, 4 electrons from the ETC are accepted by each O_2 molecule and water is made. But sometimes when electrons are passed to oxygen, it doesn't get reduced all the way to water, instead forming a free radical.

A free radical is simply an atom or molecule that has one or more unpaired electrons. Oxygen can form several different kinds of free radicals, depending on the number of electrons it has accepted. Molecular oxygen with one added electron is a superoxide; if it has 2 unpaired electrons, it forms a peroxide.

The presence of unpaired electrons makes a molecule very chemically reactive. And high chemical reactivity can cause damage to vital cellular components, such as DNA, proteins, and membrane lipids.

The Light Reactions of Photosynthesis

Early in the history of life, the atmosphere had very little oxygen in it. It was only when a group of tiny organisms called cyanobacteria began to perform photosynthesis almost 3 billion years ago that oxygen levels gradually began to rise.

The availability of abundant oxygen was dangerous because of the damage that oxygen radicals could do. The emergence of cellular antioxidant defense mechanisms helped cells minimize the oxidative damage while allowing them to benefit from the extra energy that can be harvested by aerobic metabolism.

The history of life on Earth would have unfolded very differently had it not been for the production of oxygen by those cyanobacteria. Today, our atmosphere is about 21% oxygen, a level maintained by vast numbers of photosynthetic organisms. As these organisms photosynthesize, they pump out oxygen, keeping our atmosphere supplied with the air we need to breathe.

Photosynthesis also provides the food molecules made by plants using nothing but carbon dioxide, water, and light energy.

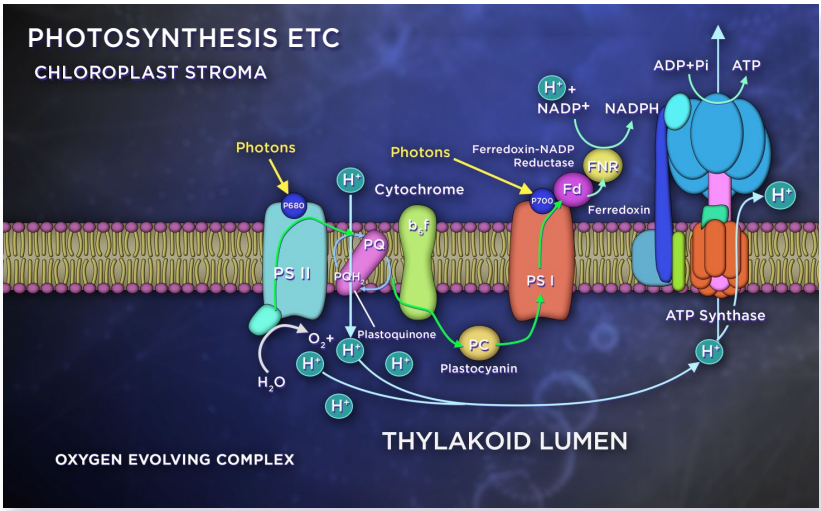
Photosynthesis has 2 phases: The first captures solar energy and uses it to make ATP, while the second uses the ATP to synthesize sugars and other food molecules. The first phase, called the light reactions because of their dependence on sunlight, makes ATP (and the activated carrier NADPH).

The process plants use to make ATP in the light reactions, called photophosphorylation, is remarkably similar to how animals extract energy via oxidative phosphorylation.

Photosynthesis employs an ETC embedded in a membrane, creates a proton gradient across the membrane, and uses ATP synthase to make ATP.

But plants are not us, and photophosphorylation differs in 3 things:

1. the cellular compartment where it occurs,
2. the source of the high-energy electrons passing through the ETC, and
3. the destination of the electrons.



In plant cells, photosynthesis is carried out in organelles called chloroplasts, which, similar to mitochondria, have 2 membranes that surround a fluid-filled inner space, the stroma. But this is not where plant cells have their ETC and ATP synthases. Instead, chloroplasts also have an internal membrane system called thylakoids where the light-absorbing pigment chlorophyll is found. The chlorophyll molecules, together with other light-absorbing pigments and associated proteins, form microscopic solar panels called photosystems. Within the thylakoid membranes is a space called the thylakoid space or lumen.

Chlorophyll is what gives plants their green color.

Light energy absorbed by chlorophyll and other pigments is the energy source for electrons in the chloroplasts' ETC. This energy is funneled to a pair of chlorophyll molecules called the special pair. There, the energy excites, or boosts, an electron from chlorophyll to a higher energy state. This high-energy electron then passes to the electron transport system embedded in the thylakoid membrane.

Once the electron from chlorophyll is passed on, the chlorophyll is left in the oxidized state and the lost electron must be replaced. The replacement electron comes from a system called the oxygen-evolving complex, which strips electrons from water to replace the ones lost from chlorophyll. When water loses electrons, 4 protons are released together with molecular oxygen. And this is how photosynthesis provides the world with oxygen.

Plants oxidize water in photosynthesis. But this oxidation does not generate energy. Instead it requires energy—from the Sun.

Meanwhile, the original electron passed to the ETC moves through a series of complexes that are similar to the ones in mitochondria. As the electrons move from one complex to the next, the energy released is used to pump protons into the thylakoid lumen, creating a proton gradient. The electrons are eventually accepted by $NADP^+$, forming NADPH. The protons flow back across the thylakoid membrane into the stroma through ATP synthase, making ATP.

Electron transport in the mitochondrion started with NADH donating electrons that were ultimately accepted by oxygen to make water. In the chloroplast, electrons start ultimately from water, creating O_2, and end up reducing $NADP^+$ to make NADPH. In this respect, the light reactions of photosynthesis look much like the reversal of the mitochondrial process, and it is made possible thanks to energy from a photon of light. The ATP and NADPH made in the light reactions is used in the synthesis of carbohydrates by plants in the phase known as the dark cycle, or Calvin cycle.

READINGS

Lane, *Oxygen*.

———, *The Vital Question*.

Royal Society of Chemistry, "Photosynthesis," https://www.rsc.org/Education/Teachers/Resources/cfb/Photosynthesis.htm.

QUESTIONS

1 Electrons can have alternate ways of entering the electron transport system. Given that, compare the relative poisonousness of an inhibitor of action of complex I versus complex III or complex IV.

2 Electron movement in the light cycle of photosynthesis is an uphill process compared to electron transport in mitochondria, which is referred to as a downhill process. Explain.

3 Imagine you made the photosynthetic fish described in the lecture, but you didn't give it a carbon source. Predict the outcome.

CLICK HERE TO SEE THE ANSWERS.

15 HOW ANIMALS MAKE CARBS AND FATS

One of the most important uses of ATP is for building new cellular components. Anabolic pathways lead to the synthesis of some of the most important groups of molecules: sugars, complex carbohydrates, fatty acids, and other lipids. These pathways need energy, obtained from catabolic pathways, which extract and store energy.

Gluconeogenesis

Glucose is produced through a pathway known as gluconeogenesis, which means "new synthesis of glucose." The body needs a continual supply of glucose, which is stored as the polymer glycogen, which is broken down to provide glucose as needed. But we only have about a 24-hour supply of glycogen in our bodies, so keeping glucose levels stable often requires gluconeogenesis.

Like other pathways, gluconeogenesis can be defined to begin at various points, but let's begin with pyruvate, the end product of glycolysis.

Starting with pyruvate, many of the reactions of the pathway are the reverse of those of glycolysis. Specifically, 11 reactions are required to synthesize glucose, compared to 10 reactions in glycolysis. Seven gluconeogenesis reactions are simple reversals of glycolysis reactions, while 4 are different.

At the end of glycolysis, pyruvate is made from phosphoenolpyruvate (PEP) in the big bang reaction. It is very energetically favorable in the glycolysis direction. For going the opposite direction in gluconeogenesis, cells use a 2-step detour to get pyruvate back to PEP.

1. In mitochondria, pyruvate combines with carbon dioxide to make oxaloacetate. This reaction is catalyzed by pyruvate carboxylase, and it requires energy from ATP. Oxaloacetate produced by this reaction is exported out to the cytoplasm.

2. Oxaloacetate is converted into PEP by an enzyme abbreviated PEPCK. The reaction requires a molecule of guanosine triphosphate (GTP).

With these 2 steps and expenditure of an ATP and a GTP, pyruvate gets made into PEP. But pyruvate and PEP each has only 3 carbons, so it will take 2 pyruvates to make glucose. So, the overall energy cost to get to 2 PEPs is 2 ATP and 2 GTP.

The next few reactions are reversals of glycolysis reactions, including the reconversion of 2 molecules of 3-phosphoglycerate to 1,3-bisphosphoglycerate (1,3-BPG). This requires one ATP per molecule and brings the energy input to 4 ATP and 2 GTP. Reduction of the two 1,3-BPG to glyceraldehyle 3-phosphate requires 2 NADH molecules.

Between meals and during exercise, blood glucose levels drop, but the need for glucose does not, particularly in the brain. Our brains require about 60% of the total glucose we use at rest, and even temporary drops in blood glucose can lead to impaired cognitive function.

We continue moving up the metabolic ladder, and only when we get to fructose 1,6-bisphosphate (F1,6-BP) do we encounter another non-glycolysis enzyme. It's part of a metabolic sleight of hand the cell is doing here. If the reaction to convert F1,6-BP to fructose 6-phosphate (F6P) had been the reversal of the phosphofructokinase (PFK) reaction, then ATP would have been synthesized by the transfer of the phosphate from F1,6-BP to ADP. But making ATP from ADP takes energy.

Instead, F1,6-BP is converted to F6P by a different enzyme, fructose 1,6-bisphosphatase (F1,6-BPase), which simply snips a phosphate off of F1,6-BP to leave F6P and a phosphate. This is instead of transferring the phosphate to ADP, and it makes the reaction much more favorable.

Next, F6P is converted into glucose 6-phosphate (G6P) in another glycolysis reaction reversal. Finally, there's another sleight of hand at the last step, involving glucose 6-phosphatase (G6Pase) in place of hexokinase to produce glucose. As with F1,6-BPase, clipping a phosphate off instead of remaking ATP makes the reaction more energetically favorable.

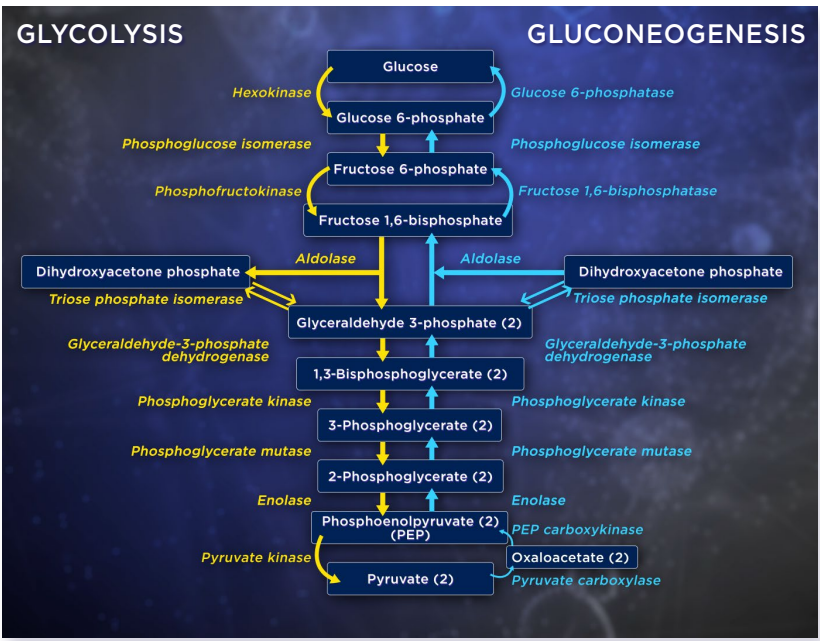
So, glucose is made from pyruvate in gluconeogenesis at the cost of 4 ATP, 2 GTP, and 2 NADH. The breakdown pathway of glycolysis to produce pyruvate, on the other hand, yields 2 ATP and 2 NADH. So, it takes 4 more triphosphates to make each molecule of glucose from pyruvate than can be obtained from glucose breakdown. This is the price you pay to have a constant level of glucose.

Other molecules can convert into pyruvate to provide additional gluconeogenesis sources. The most common feeder molecules are alanine and lactate, which can each be converted to pyruvate in a single step.

Another point of entry for gluconeogenesis is oxaloacetate, which can be obtained from the citric acid cycle, fed by amino acids like glutamate or glutamine. Aspartate and asparagine can also readily be made into oxaloacetate to help fuel gluconeogenesis.

Fat is yet another source of an intermediate for gluconeogenesis, but only a tiny portion of it. Hydrolysis of fat separates off glycerol, which gets converted to dihydroxyacetone phosphate, which is part of the pathway. However, the fatty acids and their breakdown products cannot be used to make glucose in complex animals like humans. So, glucose can be made by the liver and kidneys to maintain blood glucose levels.

Gluconeogenesis from pyruvate uses 4 unique reactions and 7 reversals of glycolysis reactions. And many molecules besides pyruvate can feed gluconeogenesis, including lactate, oxaloacetate, amino acids, and glycerol from fats.



A key idea for understanding the building of molecules is reversing the breakdown of those same molecules. But there are differences:

- ◇ Building up requires more energy than the same pathway yields in breaking down.
- ◇ Sometimes we use many of the same enzymes but take a few detours to reduce the amount of energy needed.
- ◇ Other times we use similar reactions in reverse but with different enzymes in a different cellular location.
- ◇ We straddle a catabolic/anabolic divide on the pathway for making sugars used in DNA and RNA with no fixed starting or ending point.

Glycogenesis

Glycogen is made when glucose supplies are abundant. It is stored in skeletal muscle and the liver. Making glycogen only requires 3 reactions:

1. G6P from glycolysis is converted to glucose 1-phosphate (G1P).

2. G1P reacts with the nucleotide uridine triphosphate (UTP) to make the high-energy intermediate known as uridine diphosphate glucose (UDPG). Like ATP, UTP has 3 high-energy phosphate bonds, 2 of which are broken in the process of making UDPG, which is the direct starting material for building glycogen.

3. Glucose of UDPG is added to an existing glycogen molecule by glycogen synthase. If the cell needs to make glycogen from scratch, another enzyme, glycogenin, is used. But both enzymes are catalyzing the same kind of reaction: Each UDPG reacts with the end of the growing glycogen chain, adding a glucose and releasing the uridine diphosphate (UDP).

Glycogen is a highly branched molecule, not just a linear chain of glucoses. So, in the last step, glycogen branching enzyme grabs a part of the chain and moves it to the point where it attaches to another part of the glycogen chain as a branch.

Why not continue making one long string, as opposed to a branched molecule? When glycogen is broken down for energy, its chains are cleaved, starting at the ends. The more branched glycogen is, the more ends there are and thus the more glucose can be quickly released.

The Pentose Phosphate Pathway

Glucose is only one of the sugars our bodies need. Sugars with 5 carbons, called pentoses, are needed for RNA and DNA. The pathway for the synthesis of sugars and related compounds is known as the pentose phosphate pathway (PPP).

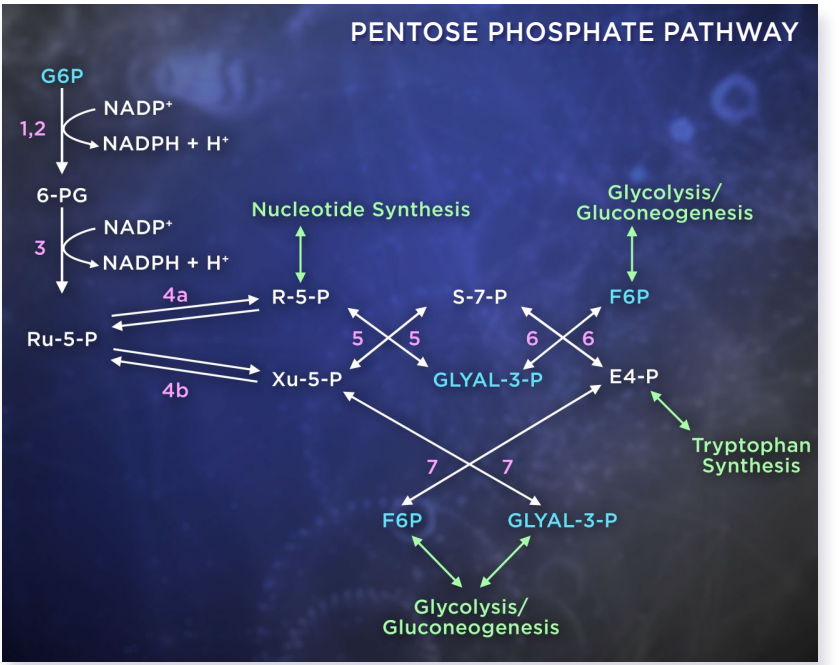
The PPP has at least 2 important possible ouputs:

1. Hexose sugars like glucose can be converted to pentose sugars like ribose.
2. NADPH can be formed.

Both of these molecules may be needed by rapidly growing cells.

The PPP allows cells to use whatever sugar phosphate molecules they have to make other compounds for cells.

The PPP is a very unusual metabolic pathway that has no real beginning or end. Its role is serving as a cellular interchange for materials. The citric acid cycle is also an interchange with no beginning or end, but the PPP, unlike the citric acid cycle, does not remake its starting material.



As a result, cells can easily use one part of the PPP and do nothing with another part of it without consequence or implication for the rest of the pathway.

The reactions in the PPP, like almost all reactions, are bidirectional.

NADPH is a simple output from the PPP, but it does have a required starting material. And G6P is required to make it. This can come from glycolysis or gluconeogenesis.

Another output of the PPP is ribose 5-phosphate. It can also use G6P as a starting point, but other intermediates from glycolysis in the PPP—such as F6P and glyceraldehyde 3-phosphate—would also work. The PPP could just as easily have a starting point of ribulose 5-phosphate from the diet.

Fats are the most efficient cellular molecules for storing energy, providing more than twice as many calories per gram as proteins or carbohydrates do.

The Synthesis of Fatty Acids

Fatty acids are the simplest of lipids, with a carboxyl group of COOH at one end and a long nonpolar hydrocarbon tail with a CH_3 methyl group at what is called the omega end. Fatty acids are the acyl groups joined to glycerol to make fats, also known as triacylglycerols.

Though lipids have many sizes, shapes, and functions, a large number, including fatty acids, are made from the same starter molecule: acetyl-CoA. This is convenient because there are several pathways that provide our cells with acetyl-CoA, including the breakdown of sugars, amino acids, and fats.

Though the pathway for building fatty acids is chemically similar to a reversal of the breakdown pathway, synthesis has several distinct aspects:

- A 3-carbon intermediate is used to add 2 carbons to the growing chain.
- NADPH is the electron source for the reduction reactions.
- Reactions are not reversals of fatty acid oxidation.
- It all takes place in the cytoplasm, versus the mitochondria for oxidation.

Although it is not directly used to add 2 carbons as the chain grows, acetyl-CoA plays 2 roles in the process: It provides the 2-carbon "seed" needed to start the process, and it is used for making the 3-carbon malonyl intermediate.

Fatty acids are made by the addition of 2-carbon units to provide a basic 16-carbon molecule. This molecule can be extended, if needed, to make longer fatty acids.

To get acetyl-CoA into the cytoplasm for these purposes, it must be moved from mitochondria. This occurs when mitochondria have an excess of acetyl-CoA or citrate. No protein, however, will transport acetyl-CoA to the cytoplasm, so it is moved as citrate, which is already being made in the mitochondria in the citric acid cycle: Citrate synthase joins oxaloacetate to acetyl-CoA, splitting out CoA in the process.

So, citrate turns out to be a branch point for 2 pathways: one being catabolic, in the citric acid cycle; and the other being anabolic, in the pathway toward synthesis of fatty acids, by providing a shuttle to move 2 carbons as acetyl-CoA. Which path citrate takes is a function of whether the cell needs energy from the citric acid cycle or has excess to make fatty acids.

Upon citrate's arrival in the cytoplasm, it is split back to oxaloacetate and acetyl-CoA, thanks to the addition of a coenzyme A to the acetyl group in the cytoplasm. Oxaloacetate returns to the mitochondrion and can transport additional acetyl groups by remaking citrate as needed.

Because not all reactions go to completion, there will be some free citrate floating around in the cytoplasm. That turns out to be quite useful, as citrate is an allosteric activator of the first enzyme in fatty acid synthesis, acetyl-CoA carboxylase, which catalyzes the addition of a carbon dioxide to acetyl-CoA to make a 3-carbon molecule called malonyl-CoA.

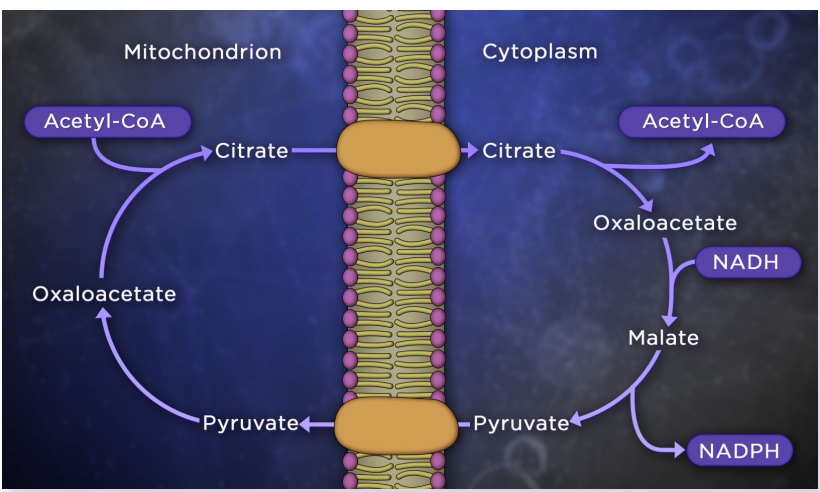
Once malonyl-CoA is made, the rest of the reactions to make the fatty acid are performed on a single multifunctional complex called fatty acid synthase that brings to bear 7 distinct enzyme actions and swaps in a replacement for CoA.

The swap takes place because fatty acid synthesis uses acyl carrier protein (ACP) as a handle for the growing acyl chain during the process of building up to a 16-carbon fatty acid. So, an enzymatic activity in the fatty acid synthase complex catalyzes replacement of CoA on acetyl-CoA for ACP. This acetyl-ACP is the seed used to start the process.

The newly made malonyl-CoA undergoes the same reaction. Once there are acetyl-ACP and malonyl-ACP, a fatty acid can begin to be assembled.

The next reaction joins malonyl-ACP with the acetyl-ACP seed, splitting out a carbon dioxide from the malonyl-ACP to make a 4-carbon molecule linked to ACP. Starting with a 2-carbon acetyl group and a 3-carbon malonyl group, this reaction ends up adding 2 carbons, with the loss of CO_2.

Compare this with beta oxidation, where 2 carbon units were cleaved off at each step. In fatty acid synthesis, it's the reverse: Two carbon units are added to the growing fatty acid chain at each step. This process is chemically very similar to the reversal of fatty acid oxidation, even if the enzymes and location are different.



READING

Widmaier, *Why Geese Don't Get Obese (and We Do)*.

QUESTIONS

1 Describe a diet in which a person would have plenty of energy but would be making mostly ketone bodies.

2 Futile cycles occur in cells when catabolic and anabolic pathways for a given molecule are occurring simultaneously. Eukaryotic cells do not have to worry too much about a futile cycle between making fatty acids and oxidizing them. Why not?

CLICK HERE TO SEE THE ANSWERS.

16 CHOLESTEROL, MEMBRANES, LIPOPROTEINS

A precursor to other molecules, cholesterol spends most of its time in cellular membranes and moves around in lipoproteins. Because a good deal of our cholesterol is synthesized in the body and does not come from diet, understanding how cholesterol synthesis occurs is important for developing strategies to reduce the body's cholesterol level.

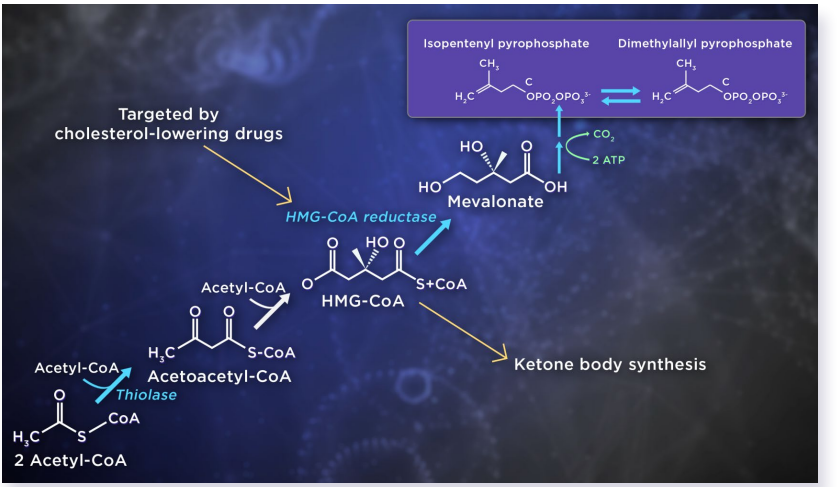
Key Steps in Cholesterol Synthesis

The first step in cholesterol synthesis starts with acetyl-CoA molecules—the same molecules that are used for fatty acid synthesis and are obtained from the breakdown of numerous molecules. The pathway to cholesterol from acetyl-CoA is long, involving 28 steps, but many cause only very minor changes. Most of the action in cholesterol synthesis takes place in the first 9 reactions.

Cholesterol's synthesis pathway partly overlaps with that of ketone body synthesis through the joining of 3 acetyl-CoA molecules to make a 6-carbon molecule known as HMG-CoA. But at this point, their paths diverge. A reaction catalyzed by an enzyme called HMG-CoA reductase converts HMG-CoA into mevalonate.

> **Because HMG-CoA reductase is central to cholesterol synthesis, it is the target of anticholesterol statin drugs, also known as reductase inhibitors, which mimic the structure of HMG-CoA and bind to the enzyme in its place, thus acting as competitive inhibitors.**

HMG-CoA reductase regulates the entire cholesterol synthesis pathway. The enzyme is controlled by several mechanisms, the most important of which occurs when cholesterol binds to it and turns it off. This is classic feedback inhibition, where the end product of a pathway binds to an enzyme to shut it off. When this happens, the entire pathway is halted; there's no need to waste resources and energy when cholesterol is abundant.



On the path to making cholesterol, mevalonate gets decarboxylated to convert into 2 slightly different 5-carbon molecules called isoprenes whose shortened names are DPP and IPP.

The next several steps involve joining these isoprenes together. Connect an IPP to a DPP and you get a 10-carbon molecule. Adding another IPP makes a 15-carbon molecule called farnesyl pyrophosphate, which provides alternative pathways to make coenzyme Q or heme.

However, on the pathway to cholesterol, what happens next is a joining of 2 farnesyl pyrophosphates to make a 30-carbon molecule known as squalene, which can be rearranged into a cousin of cholesterol called lanosterol. This 30-carbon molecule goes through numerous minor alterations (19 in total) before finally arriving at cholesterol.

Cholesterol is a precursor for the synthesis of several molecules. The first is a class known as the bile acids, which allow your body to dissolve fats in your digestive system. Bile acids are detergent-like lipids made from cholesterol by the addition of polar groups, creating a much more amphiphilic molecule than cholesterol is. The digestive system is a polar aqueous environment, and the detergent actions of bile acids solubilize nonpolar molecules like fats, cholesterol, and fat-soluble vitamins.

Gallstones—which up to 25 million people have—are mostly cholesterol. If the ratio of cholesterol to bile acids gets too high, crystals of cholesterol form, with painful results.

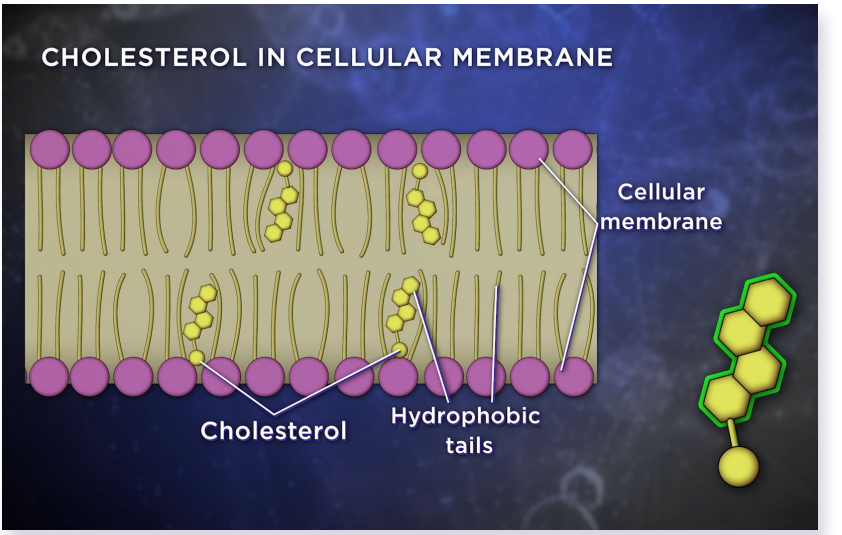
Another group of molecules made from cholesterol is the steroid hormones, which do everything from balancing ions and controlling inflammation to acting as sex hormones. The starting reactions for all steroid hormones are additions of hydroxyl groups to cholesterol.

Cellular Membranes

Probably the most important role for cholesterol is in its presence in cellular membranes to help manage their fluidity. The rigid, planar structure of cholesterol fits nicely in the gaps between the hydrophobic tails of the fatty acids, making membranes stiffer at moderate temperatures.

It also helps to significantly reduce the movement of small water-soluble molecules into and out of cells, thus maintaining membrane integrity and ensuring orderly movement of materials through protein transporters, rather than randomly through the membrane. At very low temperatures, cholesterol can insert itself between the lipid tails and keep them from packing together too closely.

The lipid bilayer, thanks to cholesterol, is remarkable in protecting the cell. Though some molecules—such as water, carbon dioxide, oxygen, and carbon monoxide—are allowed passage, most other molecules are blocked and only get in through specific gates, or channels, in the lipid membrane.



So, cholesterol helps cell membranes exclude nonspecific passage of most molecules across the bilayer. That means that the bilayer is a barrier to the movement of most molecules. On the other hand, specific passage of molecules in and out of the cell is managed by gatekeeper proteins known as membrane proteins. Together, membranes and their protein gatekeepers control the movement of materials in and out of cells. This control of transport determines everything from the vitality and fate of cells to the beating of our hearts.

Membranes not only create the outer boundaries of cells, but they also define inner compartments within cells, such as nuclei and mitochondria in eukaryotic cells. The same basic rules for making cell membranes hold for making organelle membranes: The foundational molecules are glycerophospholipids and sphingolipids. There are some differences, though, in composition.

Mitochondrial membranes are rich in a lipid not found elsewhere. It is known as cardiolipin. Damage to this lipid is part of the cellular signal for suicide that arises when mitochondrial function is impaired. In this process, a cellular peroxidase enzyme catalyzes an oxidation reaction on cardiolipin, which leads to membrane rearrangement and release of the protein cytochrome c. This is what activates the suicide mechanism.

Membranes must be fluidlike and semipermeable to function for a living cell. Cells and their membranes are like balloons filled with water: They are extremely flexible and can very easily change shape, but the membrane must be strong enough to avoid bursting and spilling its contents.

Cells biochemically alter their own membranes to keep them fluid at the temperature of the organism. Membranes containing fatty acids transition from a more fluid state to a more solid state as temperature drops. For human cells, at a relatively constant body temperature, this isn't a big deal, but it is a big deal for organisms like fish in icy waters.

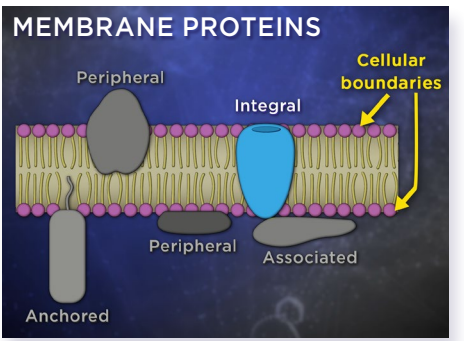
The biochemical adaptation involves changing the mix of fatty acids present in the glycerophospholipids and sphingolipids. Unsaturated fatty acids have kinks in their hydrocarbon tails that keep them from packing too closely. Cells increase the unsaturated fatty acid content of membrane lipids to keep them fluid at lower temperatures.

Membranes also contain a fair amount of cholesterol, occupying 14% of the dry weight of brain tissue.

Membranes are important cellular barriers, but barriers can be a bane as well as a boon. Cells need nutrients, such as glucose, which can't cross the membranes as they are. So, cells must have mechanisms for transporting needed molecules in and other molecules out. This is why cells have a vast array of membrane proteins, each of which is typically very specific for the molecules or ions they move.

The importance of membrane proteins is also shown by their relative abundance. On average, about 50% to 60% of a cellular membrane's weight is made up of membrane proteins. The other ½ are lipids, though because proteins are much heavier than lipids, this means there are far more lipid molecules.

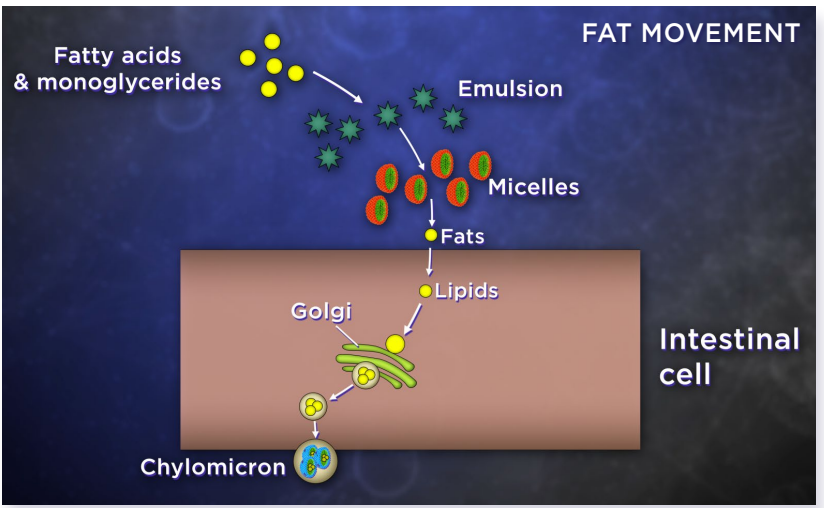
Proteins in the lipid bilayer have numerous functions, but an important function is serving as gatekeepers that control material crossing the cellular boundaries provided by the lipid bilayer. These proteins are called integral membrane proteins. Each of them also provides an opening or binding site that is specific for the molecule(s) being transported.



Lipoprotein Complexes

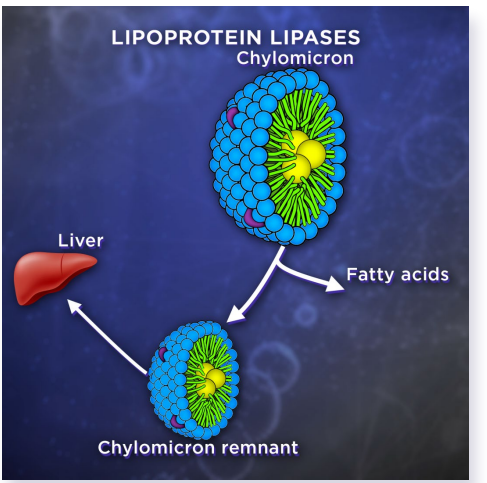
Membranes need cholesterol, and cells need other lipids as well. In order for cholesterol to do its magic, or for fat to provide energy to hungry cells, these water-hating molecules must first overcome their hydrophobic urges and travel through the body of an organism that is mostly water. Doing this takes extra steps, and this is one of the reasons that fat and fatty acids are not used as rapid sources of energy. They are insoluble in water, so cholesterol and other lipids must be made water-soluble so that they can travel through their watery surroundings.

Just as you send a letter through the mail with an enclosing envelope, cells package up lipids in the body. That process with dietary lipids begins in the digestive system when they encounter the aqueous environment of the stomach. Just as you use soap on your hands to make grease soluble in water, the digestive system's detergent-like bile acids prepare dietary fats, cholesterol, and other lipids for transport across the intestinal wall. The process, called emulsification, creates small droplets called micelles that make the lipids water-soluble and therefore capable of being delivered to the intestinal cells.



Once inside the intestinal cells, fats—together with other lipids—get packaged into protein/phospholipid bundles called lipoprotein complexes. The lipoprotein complexes made in intestinal cells are ultra-low-density lipoproteins (ULDLs) known as chylomicrons. These complexes are organized so as to hide the hydrophobic triglycerides in the interior of the structure surrounded by phospholipids. The outside includes lipoproteins and phospholipids, both of which are water-soluble. So, lipids are shielded from their aqueous surroundings by the water-loving proteins and phospholipids in the complex.

In the body, dietary lipids begin their journey in chylomicrons by entering the lymphatic system, which empties into the bloodstream. There, enzymes known as lipoprotein lipases, which are released by cells, begin to digest the fats in the chylomicrons. As these fats break down to fatty acids and glycerol, other lipids—though not cholesterol—may be released from the chylomicron to be absorbed by cells.



The chylomicron shrinks considerably as it loses these components, and its reduced size allows it to move freely through the capillaries. This shrunken structure travels the bloodstream to the liver, where it is absorbed and its contents are stashed away.

The liver is the major modulator of metabolites the body needs. Blood glucose concentrations are managed in the liver, and the liver monitors and addresses the body's lipid needs.

The liver must also release packaged-up lipids to meet the body's needs. When the liver senses that the body needs fat and cholesterol, it begins creating some lipoprotein bundles of its own. These complexes are known as very-low-density lipoproteins (VLDLs), which contain cholesterol and fats that travel in the bloodstream, much like chylomicrons do.

After VLDLs leave the liver and enter the bloodstream, lipases digest the fats, but not the cholesterol, in them. As a result, VLDLs shrink, becoming slightly higher-density lipoproteins called intermediate-density lipoproteins (IDLs) and, eventually, low-density lipoproteins (LDLs)—so-called bad cholesterol.

LDLs contain the highest concentration of cholesterol of all lipoprotein complexes. The problem isn't the cholesterol itself, but mostly the unsaturated fatty acids that LDLs contain. The things you can get from the unsaturated fats that are supposed to be good for you are susceptible to oxidation, and this is the first step in the development of an atherosclerotic plaque. The higher the levels of LDLs are, the greater the likelihood of damaging reactions with the unsaturated fatty acids is. Oxidized LDLs are swallowed by immune cells called macrophages, triggering a process that ultimately results in a physical block called an atherosclerotic plaque in the blood vessel.

Atherosclerotic plaques can cause heart attacks by blocking blood flow through arteries that supply the heart. Plaques can also rupture and trigger clotting. In contrast to the blockage of coronary arteries, clots are free to move anywhere in the body. If they lodge in a critical supplier of blood to the brain, a stroke can result.

If LDLs are so dangerous, why does the body make them?

LDLs are a means of delivering cholesterol directly into cells, which have proteins called LDL receptors embedded in their plasma membranes. These receptors stick out of the cell membrane and latch onto LDLs from the bloodstream.

Like the other lipoprotein complexes, high-density lipoproteins (HDLs)—so-called good cholesterol—are involved in the movement of lipids, but unlike the others, levels of HDLs are inversely correlated with atherosclerotic events, meaning that the more HDLs you have, the less likely it becomes for you to get atherosclerosis.

Given all the good effects of HDLs, can we do anything to increase their levels?

Exercise, weight loss, consumption of omega-3 fatty acids, increased intake of fiber, reduced carbohydrate intake, and removal of trans fats from the diet are all known to increase HDLs.

HDLs remove cholesterol from macrophages during the formation of atherosclerotic plaques, thus reducing the incidence of plaques. HDLs also take extra cholesterol from other cells back to the liver or to adrenal glands, testes, or ovaries, where it is used to make steroid hormones. This reduces the concentration of free cholesterol in the body. HDLs also inhibit oxidation, inflammation, blood coagulation, and blood platelet aggregation, all of which are factors in coronary heart disease.

READINGS

American Society for Biochemistry and Molecular Biology, "A Spring-Loaded Sensor for Cholesterol in Cells," https://www.eurekalert.org/pub_releases/2017-12/asfb-ass120717.php.

Yang, et al., "Cellular Cholesterol Homeostasis and Alzheimer's Disease."

QUESTIONS

1 Some side effects of digoxin include faintness, confusion, and vision changes. Based on what you know about the protein it affects, speculate on why these might not be unexpected.

2 Enzymes that break down LDL receptors on the cell surface are important for the control of how much cholesterol cells take in. Predict the effect on blood LDL/cholesterol levels of inhibiting the action of these enzymes that break down the receptors.

CLICK HERE TO SEE THE ANSWERS.

17 METABOLIC CONTROL DURING EXERCISE AND REST

With metabolic control, there is no central processor in a cell that directs activities through any sort of master plan. Instead, cellular responses to varying conditions are built into the enzymes that each run their own small piece of the show. Many cellular mechanisms operate through simple if-then responses: If a certain condition is met, then a particular response automatically follows.

Regulation of Metabolic Pathways

Ways that enzymes can be controlled include allosteric regulation, where the binding of a molecule modifies an enzyme's activity; covalent modification, where a phosphate group gets chemically linked to specific side chains in an enzyme protein to turn it on or off; and control of whether or not a cell even makes an enzyme in the first place.

Enzyme activity can be tuned to increase or decrease in response to a variety of conditions. This property is used to great advantage in controlling metabolic pathways.

Regulation of key enzymes in certain pathways ensures that a given cell doesn't waste energy breaking down and building the same molecules simultaneously. And because neither the synthesis nor the breakdown process is 100% efficient, it will take more energy to build the molecules than is obtained by breaking them down.

In breaking down glucose in glycolysis, for example, the net products are 2 pyruvates, 2 NADH, and 2 ATP. To make glucose in gluconeogenesis requires 2 pyruvates, 2 NADH, 4 ATP, and 2 GTP.

Put glycolysis and gluconeogenesis together and there is a net loss of 4 triphosphates each time both cycles are run. This is an example of what biochemists call a futile cycle. It's futile because metabolically it's pointless; in fact, there is an energy loss accompanying it. Naturally, it is in a cell's interest to prevent this from happening.

In glycolysis, the most important regulatory enzyme for breaking down glucose is phosphofructokinase (PFK) at step 3. Recall that gluconeogenesis doesn't try to reverse the PFK reaction; instead, it employs a different enzyme called fructose 1,6-bisphosphatase (F1,6-BPase).

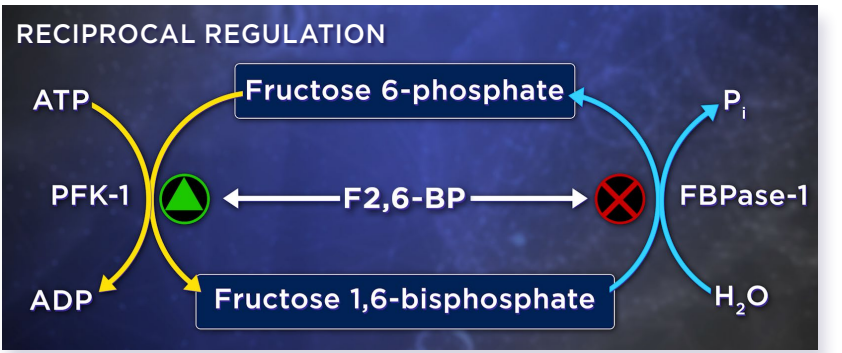
These 2 key enzymes control the flow of intermediates through glycolysis and gluconeogenesis, respectively. To ensure that the 2 pathways don't run simultaneously—which cannot happen—the cell turns one of these enzymes off any time the other is on. Both enzymes are allosterically regulated by the same molecules, but in opposite ways.

A more important regulatory switch with the same results is fructose 2,6-bisphosphate (F2,6-BP). This is another molecule made in cells that activates PFK and glycolysis to eat up all of the glucose while it inhibits glucose synthesis from F1,6-BPase and gluconeogenesis.

By contrast, a regulator where both effects are reversed is citrate, which is formed in the citric acid cycle. Citrate accumulates when cells have abundant energy. When lots of citrate is present, glucose breakdown will taper off to

slow down ATP formation. On the other hand, high levels of citrate activate fructose 1,6-bisphosphatase (F1,6-BPase), which will cause more glucose to be made, thus using ATP to store energy.

Regulation is simple when a single molecule has opposite effects on regulatory enzymes of corresponding catabolic and anabolic pathways. Such regulation is called reciprocal regulation. This is convenient and especially important when corresponding catabolic and anabolic processes can occur in the same cellular location—in this case, both sets of reactions are happening in the cytoplasm.



Reciprocal regulation also plays a role in the building and breaking down of glycogen, which is made when cells have more glucose than they can immediately use and is broken down to release glucose when energy is needed.

The regulated enzyme for glycogen breakdown is glycogen phosphorylase, and the corresponding enzyme for glycogen synthesis is glycogen synthase. Both enzymes are regulated by phosphorylation/dephosphorylation. Once again, the switch that turns one enzyme on turns the other one off. Phosphorylation makes glycogen phosphorylase more active and glycogen synthase less active. The results reverse when the phosphates are removed.

The stimuli for phosphorylation/dephosphorylation of these enzymes are hormones that mount a cellular response to bodily needs. Epinephrine, also known as adrenaline, is released in times of stress or excitement. It stimulates a cascade of phosphorylation, which activates glycogen phosphorylase. This favors the release of glucose from glycogen to fuel muscles and prepare them for action.

The hormone insulin, on the other hand, is produced when blood levels of glucose increase. Insulin stimulates dephosphorylation, activating glycogen synthase and favoring storage of glucose as glycogen.

In addition to allosteric effectors (in glucose metabolism) and phosphorylation (in glycogen metabolism), a third important mechanism for regulation of metabolic pathways is the availability of substrates—specifically electron carriers—that are needed to run the reactions.

The electron carriers NAD^+ and FAD accept electrons released in cellular oxidation reactions to become NADH and $FADH_2$, which are the fully loaded electron carriers. The relative availability of oxidized electron carriers like NAD^+ compared to reduced carriers like NADH is a factor in metabolic control.

Other players include the proton gradient made by those electron carriers, but also ADP and oxygen.

The Role of Electron Transport

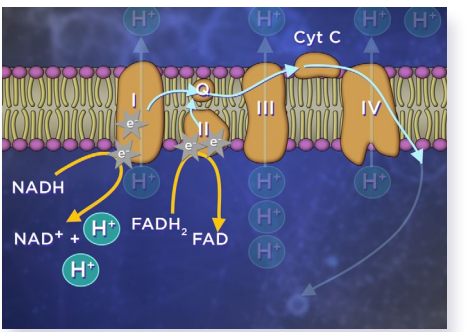
The movement of electrons through the electron transport chain pumps protons out of the mitochondrial matrix, creating a proton gradient. The terminal acceptor of electrons during electron transport is molecular oxygen. The proton gradient generated by electron transport is used by ATP synthase to create ATP from ADP and phosphate. All of these components are interdependent on each other. If one is limiting, all are affected.

Imagine a sedentary person named Sam who is sitting, eating pizza, drinking beer, and watching TV. With plenty of food available and no exercise, ATP levels in Sam's cells are high and ADP levels are low, having been mostly used up to make ATP. When this occurs, the ATP synthase slows to a crawl.

ATP synthase is the cellular enzyme through which protons reenter the mitochondrial matrix in the process of oxidative phosphorylation. If the ATP synthase slows down, the proton gradient grows, so long as the electron transport system is running.

But this cannot go on indefinitely. After a point, further buildup of the proton gradient is no longer possible. It's like a rechargeable battery nearing full capacity. The proton gradient depends on the movement of electrons through the electron transport complexes, and vice versa. If the proton gradient gets fully charged, electrons can't move in electron transport. When electrons stop moving, oxygen use drops. That's why couch potatoes don't breathe heavily when they're sitting and eating.

Sam's sluggish electron transport system from lack of exercise means that NADH and $FADH_2$ have no place to hand off their electrons. The change of NADH back to NAD^+ and $FADH_2$ to FAD slows down. As NADH and $FADH_2$ accumulate, NAD^+ and FAD concentrations fall. And when this happens, oxidative pathways relying on NAD^+ and FAD by necessity decrease, too.



This means that the citric acid cycle and fatty acid oxidation slow or stop. Lower citric acid cycle activity and decreased fatty acid oxidation explain why you're not going to burn off fat if you don't exercise.

Perhaps Sam decides to start running to lose weight. As Sam runs, the muscles in the legs use ATP to fuel their movement, leading to a drop in ATP levels. As a result, Sam's ATP synthase begins to spin a little faster, and more protons begin to enter the mitochondrial matrix. The proton gradient decreases. With a smaller proton gradient, electron transport speeds up, and cells need more oxygen to give the electrons to. Sam begins to breathe more deeply to supply the needed oxygen via the bloodstream.

As more and more electrons flow, more NADH and $FADH_2$ get converted to NAD^+ and FAD. Increased levels of these carriers stimulate the citric acid cycle and fatty acid oxidation. And as more fatty acids get used up, Sam begins to burn off the extra pounds.

Sam's experiences show the codependence of factors in electron transport, oxidative phosphorylation, and oxidative metabolism that occurs in glycolysis, fatty acid oxidation, and the citric acid cycle. If cells had unlimited quantities of NAD^+, FAD, and ADP, then stopping electron transport or oxidative phosphorylation would have no upstream effect on the citric acid cycle and fatty acid oxidation. But electron carriers and ADP are limiting resources; they must be recycled by electron transport and oxidative phosphorylation.

Oxidative metabolism, electron transport, and oxidative phosphorylation all respond to changes in each other. This is what metabolic control is all about.

Fermentation

When people exercise vigorously, their bloodstream has difficulty supplying all the oxygen the muscles could optimally use to generate ATP. And when cells have insufficient oxygen, electron transport slows and NAD^+ levels fall—and that could be a problem.

Fortunately, there's a backup: fermentation. In fermentation, oxygen is not needed, and cells convert pyruvate to lactic acid to regenerate NAD^+.

When Sam is running really hard, lactic fermentation serves as an alternate source of NAD^+ to keep glycolysis going. Cells need glycolysis, at a minimum, because it can provide ATP even when the citric acid cycle and fatty acid oxidation are stopped.

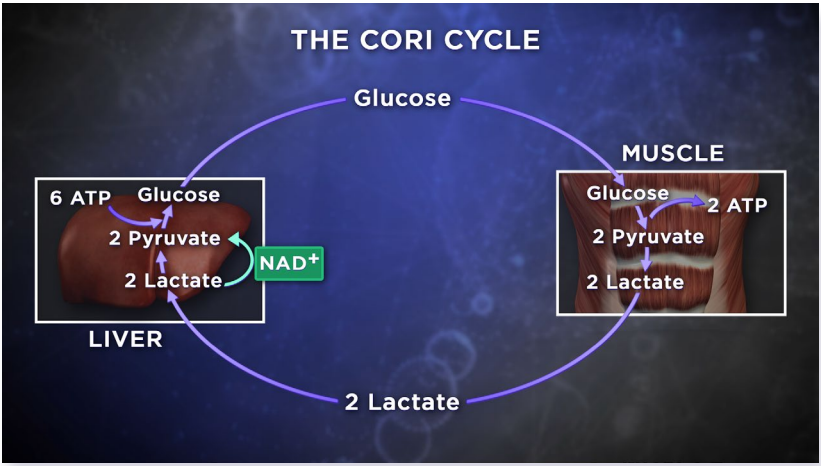
But it takes almost 20 times as much glucose to make ATP in fermentation as it does in the presence of oxygen. So, Sam's muscles need a lot of glucose to keep going. Fortunately, as levels of glucose in the blood fall, Sam's liver releases glucose into the bloodstream, using both gluconeogenesis and glycogen breakdown as quick and easy sources of glucose.

Meanwhile, in the muscle cells, fermentation is churning out the lactic acid—which is a metabolic dead end. It is made from pyruvate, but the only thing lactic acid can be metabolized to is pyruvate once again. Reoxidation of lactic acid back to pyruvate requires NAD^+, but that is in short supply in the muscle cells.

So, rather than accumulate lactic acid that's useless and even unhealthy if enough is on hand to lower pH, the muscle cells instead dump it into the bloodstream. And the lactic acid travels to Sam's liver. Unlike the muscle cells, the liver is not exercising and has no shortage of oxygen from its neighbors, the lungs. As a result, the liver has electron transport going on, so it has plenty of NAD^+. The liver cells use this NAD^+ to convert lactate back to pyruvate.

Lactate is not a dead end for the liver, from which it makes pyruvate, which is then used to build glucose through gluconeogenesis. So, the liver cranks out glucose and dumps it into the blood for muscle cells. Thus, the circle is closed. Muscle cells export lactate to liver cells, which convert lactate to glucose, which muscle cells take up and use to make more lactate. This cycle of metabolites traveling between the liver and muscle cells is called the Cori cycle.

The Cori cycle was named for Carl and Gerty Cori, who worked out the connections between liver glycogen, hormones, and blood glucose to win a Nobel Prize in 1947.



Let's compare the energy used when the Cori cycle is running and when it is not. When people exercise vigorously, they breathe very heavily as their bodies try to provide enough oxygen to their cells to keep electron transport going. There are 2 possibilities:

1 In the aerobic scenario, breathing provides enough oxygen so that electron transport can work efficiently.

2 In the partially anaerobic scenario, the individual is exercising even more rapidly than oxygen can be delivered, so the Cori cycle is going full force.

In both situations, Sam is doing aerobic exercise: The exerciser is breathing hard, with an elevated heart rate. But muscles in the oxygenated scenario deliver 36 ATP from each glucose while muscles running fermentation get 2 each, so muscles in fermentation need 18 times as much glucose to get the same energy as muscles rich in oxygen.

READING

Chandel, *Navigating Metabolism*.

QUESTIONS

1 Futile cycles are generally not useful to cells because they waste energy. They are useful, however, for heat generation. Compare a futile cycle involving glycolysis and gluconeogenesis versus fatty acid synthesis and oxidation with the speed of operation and the amount of heat generated using information from previous lectures.

2 PFK is an unusual enzyme in that it uses one molecule, ATP, as both an allosteric inhibitor and a substrate. ATCase, by contrast, uses aspartate as both an allosteric activator and a substrate. This makes sense, because increasing the amount of aspartate increases enzyme activity both allosterically and by increasing the level of substrate binding. How is it that increasing amounts of ATP inhibit PFK and don't result in increased substrate binding?

3 The person named Sam described in the lecture had 2 situations: nonexercise and exercise. Describe the effects these 2 situations would have on brown fat cells in Sam's body.

CLICK HERE TO SEE THE ANSWERS.

18

HOW PLANTS MAKE CARBS AND OTHER METABOLITES

Plants hold the secret to capturing solar energy and using it to build biological molecules out of air and water. In photosynthesis, plants convert light energy into chemical energy in ATP, producing NADPH, which gains electrons for driving anabolic reduction reactions. That first phase of photosynthesis is also called the light reactions, and it supplies plants with 2 important molecules—NADPH and ATP—which they need for the second phase, where sugar is synthesized.

An estimated 450 gigatons of carbon are in plant biomass, built by plants, beginning with air and water. The biomass of the roughly 8 billion humans is a relatively puny 100 million tons. And we're here at all only because of food made possible by a crucial set of reactions that happen in the chloroplasts of plants.

The Calvin Cycle

The Calvin cycle constitutes the second phase of photosynthesis, where the light energy is used to build the molecules that make up the plants, support their growth, and feed the rest of the world.

In biochemical terms, carbon dioxide is being chemically reduced to a carbohydrate. When a molecule is reduced, it accepts electrons. A carbohydrate is a more complex molecule than carbon dioxide, so the Calvin cycle involves a gain of electrons by CO_2 and an increase in complexity—meaning that a source of electrons and a source of energy are needed. Another name for what's going on is carbon fixation.

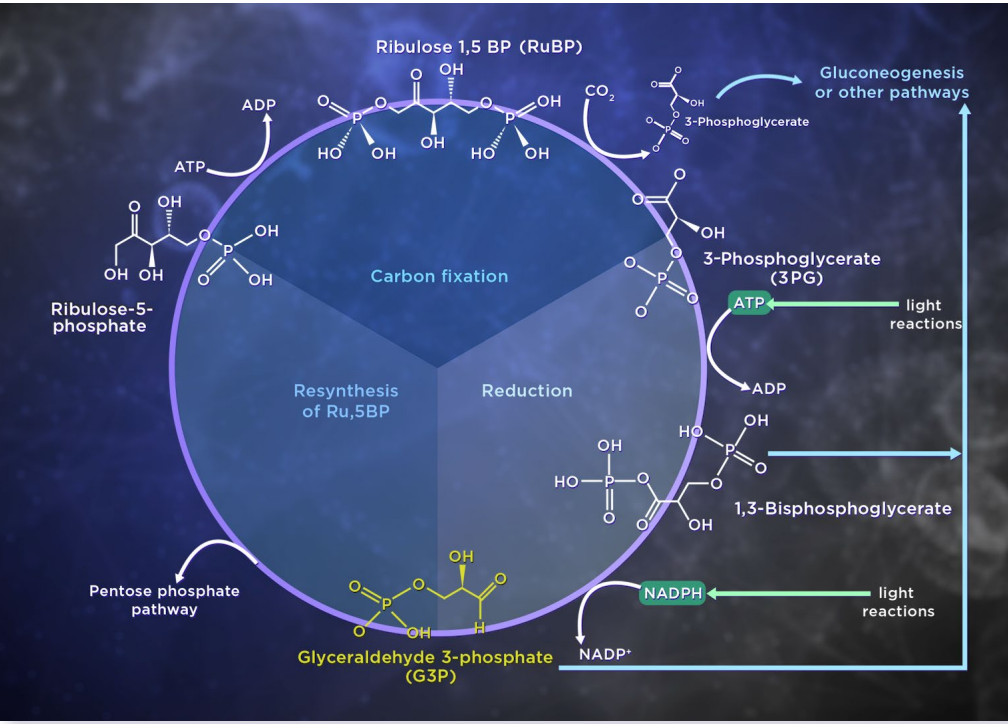
The RuBisCO enzyme is the most abundant protein in the world. And as the catalyst for the first reaction in making food from carbon dioxide, RuBisCO is an enzyme that the world depends on.

The first of the reactions of the Calvin cycle is the addition of CO_2 to a 5-carbon phosphorylated molecule called ribulose 1,5-bisphosphate (RuBP). This reaction is catalyzed by the enzyme RuBP carboxylase/oxygenase (RuBisCO).

The product of the RuBisCO reaction is a 6-carbon compound called 3-keto-2-carboxyarabinitol 1,5-bisphosphate. That molecule turns out to be extremely unstable. When it breaks down, it splits into 2 identical molecules that each contain 3 carbons known as 3-phosphoglycerate (3PG), which all living organisms make in glycolysis and gluconeogenesis.

At this point, the first of 3 stages in the Calvin cycle is complete.

In the next stage, 3PG is reduced to make the molecule glyceraldehyde 3-phosphate in reactions that use ATP and NADPH from the light reactions. NADPH donates the electrons needed for the reduction, and ATP provides energy.



Note that the reactions taking 3PG to G3P are identical to steps in gluconeogenesis. This makes some sense, as both pathways synthesize sugar. But in the Calvin cycle, the reactions do not go on to make glucose—not directly, anyway.

Getting to glucose is a bit more complicated because so far only one CO_2 has been taken up. Only after running the cycle 6 times are there enough carbons captured to make a 6-carbon sugar like glucose.

To continue running the cycle, some of the G3P molecules are recycled to make more RuBP. Regeneration of RuBP is the third stage in the Calvin cycle. And RuBP from the third stage is also the molecule that reacts with CO_2 in the first reaction of the cycle.

This third stage has several possible inputs and multiple outputs, but let's focus only on the regeneration of RuBP. The amount depends on the availability of ATP from the light reactions.

So, it takes 6 turns of the Calvin cycle to get enough carbon to make a glucose. Each complete turn requires 3 ATP and 2 NADPH. So, the 6 turns needed for making a single glucose molecule require 18 ATP and 12 NADPH. And these are made in the light reactions.

The G3P made in the Calvin cycle has 2 fates: Some of it is siphoned off from the cycle to go into making sugar, and some of it goes back into making RuBP to continue CO_2 capture.

And that is how photosynthesis makes food out of water and CO_2. Don't forget that water enters the picture in the light reactions: Water provides electrons for the electron transport chain in chloroplasts to make NADPH and, indirectly, ATP. Splitting water also releases oxygen.

Photorespiration

Plants don't make food or produce oxygen for our benefit; they need glucose for their own cellular reactions. Extra glucose is stored as starch in tubers like potatoes and in the seeds of grain plants like wheat. Plants also use glucose to make cellulose, which is a polymer of glucose that makes up the plant cell wall.

The production of oxygen during photosynthesis is useful to us, but it has also had some unintended consequences. Early photosynthetic organisms functioned in an atmosphere that contained much less oxygen than we now have. But oxygen levels rose as photosynthesis kept pumping it out. This made it possible for a lot of new oxygen-dependent organisms to thrive. But that created a problem for photosynthesis.

It turns out that RuBisCO, the enzyme that captures CO_2 from the air, is not very good at distinguishing oxygen from CO_2. About ¼ of the time, RuBisCO accidentally binds oxygen instead of carbon dioxide. Consequently, no CO_2 is added to the 5-carbon RuBP. So, instead of getting a 6-carbon intermediate that splits into 2 molecules of 3PG, one molecule of 3PG and the 2-carbon molecule 2-phosphoglycolate result.

This reduces the efficiency of carbon dioxide capture. Instead of 2 3PGs, you get one. Even worse, 2-phosphoglycolate is toxic to the cells, so it must be dealt with. Plants manage this through a series of reactions called photorespiration that converts the 2-phosphoglycolate into 3PG.

But this recycling costs plants more than 12 ATP per molecule of 2-phosphoglycolate, and in the process, some of the carbon is lost as CO_2. Moreover, getting oxygen confused with carbon dioxide becomes even more likely when temperatures rise. So, this problem will get worse as the world warms up.

Plants are in a bind regarding photorespiration. It greatly reduces the efficiency with which the Calvin cycle runs, but they need it to detoxify 2-phosphoglycolate. Some plants have come up with interesting fixes for this problem.

Most plants directly capture CO_2 from the air, combining it with RuBP to produce 3PG. Because this inefficient approach to carbon capture directly creates a 3-carbon molecule, these plants are called C_3 plants.

Some flowering plants have evolved a different carbon fixation approach to reduce the oxygen problems. The leaves of these plants have a specialized anatomy that separates the capture of carbon dioxide from the reactions involving RuBisCO.

One set of cells, called the mesophyll, closer to the leaf surface, is in contact with atmospheric CO_2. But unlike C_3 plants that use RuBisCO to take carbon dioxide directly from the atmosphere, these plants use a different enzyme, phosphoenolpyruvate (PEP) carboxylase, which combines the CO_2 with phosphoenolpyruvate to form oxaloacetate.

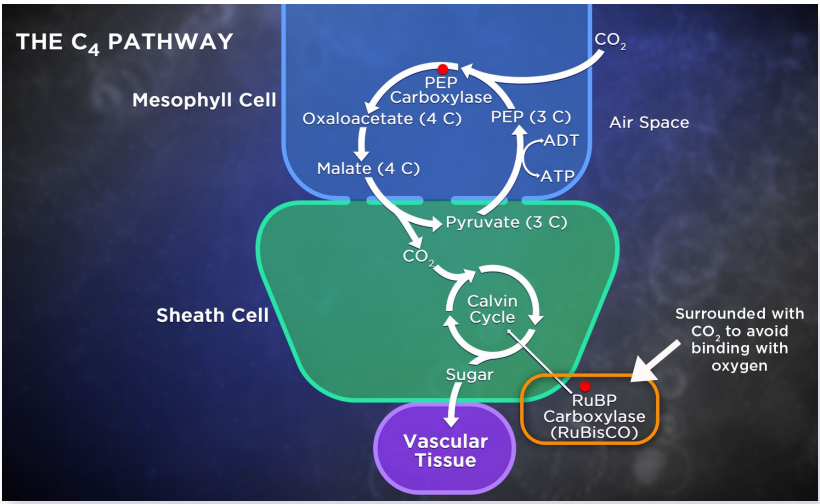
Because this alternative approach to carbon capture yields the 4-carbon oxaloacetate, these plants are called C_4 plants. The C_4 pathway is used by plants like prairie grasses and sugarcane, which grow in hot places where photorespiration could quickly dry them out.

By not using RuBisCO to capture CO_2 initially, C_4 plants avoid RuBisCO's mess from binding to unwanted oxygens. Instead, C_4 plants keep the RuBisCO tightly packed in another set of cells called the bundle sheath with less access to oxygen.

Meanwhile, the oxaloacetate in the mesophyll is converted to malate, another 4-carbon molecule, which gets transported to the bundle sheath and then is broken down to release the carbon dioxide right where RuBisCO is. And by bringing the carbon dioxide to the RuBisCO rather than letting it help itself from the atmosphere, where the oxygen concentration is high, C_4 plants reduce photorespiration greatly.

So, C_4 plants solve the photorespiration problem by keeping enzymes for carbon dioxide capture in different parts of the leaves from where RuBisCO is located.

Another solution to the photorespiration problem evolved in plants native to very hot, dry places. Their adaptation is called crassulacean acid metabolism (CAM), named for the Crassulaceae family, in which this was discovered. CAM plants—such as orchids, cacti, and pineapple—use the C_4 strategy of reducing the need for photorespiration and also save water.



How do these plants do both? Leaves on their surfaces have pores through which gases and water vapor diffuse. These pores, called stomata, can be opened and closed by the plant. CAM plants, which are adapted to desert conditions, open stomata only at night, when it is cooler and humidity is higher. This reduces water loss, but it also limits when carbon dioxide is available to the leaf cells.

At night, CO_2 diffuses into the leaf and is captured by PEP carboxylase to make oxaloacetate, just like C_4 plants. This is then stored in saclike structures called vacuoles in the cells until the next morning. When the sun comes up, CAM plants close their stomata, limiting the entry of gases, including oxygen. Then, they break down the oxaloacetate in the presence of RuBisCO, just like C_4 plants, to release CO_2.

Because they release all their CO_2 during the day while their stomata are closed and oxygen is not being admitted, these plants, too, are able to surround RuBisCO with carbon dioxide and avoid binding to oxygen.

The C_4 and CAM plants evolved adaptations that compensate for RuBisCO's inefficiency. But unfortunately, most plants are C_3 plants, and in these plants, photorespiration reduces crop yields significantly—by more than 1/3 for soybeans, for example.

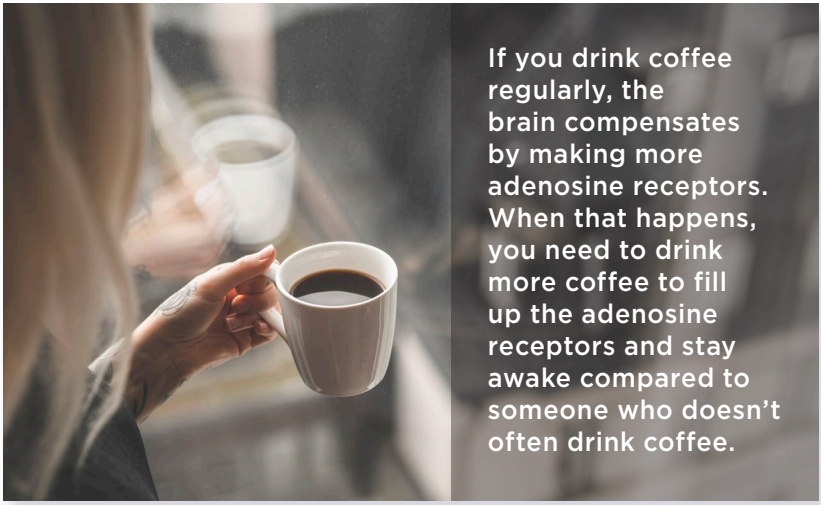
Researchers wondered whether photorespiration could be made less of a drain on the system. And in 2019, by introducing the genes for just 2 enzymes into chloroplasts, researchers bypassed photorespiration in C_3 plants.

Secondary Metabolites

In addition to the 2 primary things plants do for us—food and oxygen—there are 3 families of secondary metabolites that plants also make. By definition, secondary metabolites are not essential for plant growth and reproduction. But they often serve to protect plants against pests. It just so happens that many of these molecules are also valued by humans.

The first family is the alkaloids. Their most notable member, caffeine, is biochemically a methylxanthine, which is a purine compound similar to the bases adenine and guanine in DNA and RNA. And indeed, caffeine is derived from the purine nucleotides.

Caffeine, which is structurally similar to adenosine, binds to adenosine receptors in the brain. Adenosine normally accumulates during our waking hours and binds to those brain receptors. Adenosine receptors filling up with adenosine is a signal to the body to slow down and rest. This is why you get sleepy after a long day. When you sleep, adenosine levels decrease, emptying receptors and eventually leading to wakefulness. Caffeine binds to those adenosine receptors, but it doesn't make you sleepy. And because adenosine can't bind when caffeine is occupying its receptors, you stay alert.



Another group of secondary metabolites is the phenolic compounds, all defined by a phenol ring structure. Examples include the main flavor of vanilla and salicylic acid, the starting molecule for aspirin. Phenol is a benzene ring with a hydroxyl group attached. Plants make a bewildering variety of phenolic compounds, including polyphenols, which can have multiple phenol rings. But they all start from the amino acid phenylalanine.

A group of secondary metabolites that gives many herbs and spices their characteristic aromas is the terpenes and terpenoids. Terpenes are hydrocarbon compounds often manufactured using resin from evergreen trees. There are a few different biochemical pathways to terpenes, including one beginning with mevalonate, the same molecule that animals use to make cholesterol. Terpenoids are modified terpenes that add other groups, often containing oxygen. Also included in the terpene group are the carotenoids, pigments that make carrots orange and corn and daffodils yellow.

Plants make many terpenoids, including those that make up latex rubber and amber.

READINGS

Morton, *Eating the Sun*.
Sumner, *The Natural History of Medicinal Plants*.
von Cammaerer, et al., "C_4 Photosynthesis."

QUESTIONS

1 Assuming that 3 ATPs can be made from the energy in each NADPH, calculate the efficiency of energy capture for the oxidation of glucose by cells. Remember that glucose oxidation in the presence of oxygen yields approximately 36 ATPs, and in the absence of oxygen, 2 ATPs are produced.

2 Predict the difference in oxygen generation of C_4 plants versus C_3 plants.

3 Given that cells have limiting amounts of ADP/ATP, is it more efficient for the light cycle to operate only in the light and the dark only in the dark? Explain your answer.

CLICK HERE TO SEE THE ANSWERS.

19 RECYCLING NITROGEN: AMINO ACIDS, NUCLEOTIDES

Earth's atmosphere is overwhelmingly made of nitrogen. But our cells are unable to directly take up the free nitrogen from the atmosphere and use it to build amino acids or the nitrogenous bases of DNA. Fortunately, some bacteria can pull nitrogen from the atmosphere, and there are a few plants that can partner with those bacteria to help capture nitrogen. All the rest of life depends indirectly on those few microbes partnered with a select group of plants. The result is a major bottleneck in the movement of nitrogen into living systems.

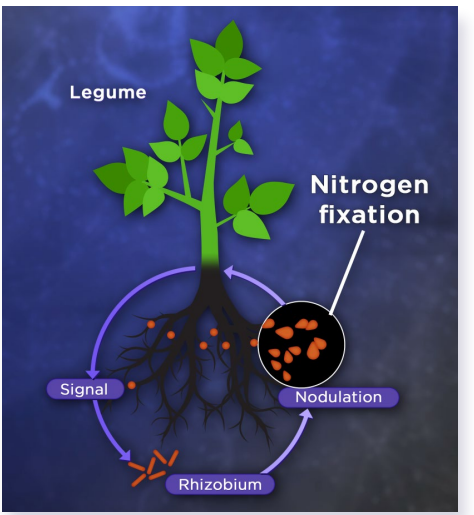
Nitrogen Fixation

To be absorbed and used by animals or plants, nitrogen in the air must be converted to forms of nitrogen that plants can use—either ammonia (NH_3) or nitrate (NO_3). Converting nitrogen to ammonia is no easy task. The industrial process for making ammonia from nitrogen requires 400°C to 500°C and more than 200 atmospheres of pressure. Fortunately, cells capture nitrogen at normal temperatures and pressure, thanks to the remarkable enzyme nitrogenase.

Some of the bacteria that grab, or fix, nitrogen are free-living bacteria in the soil or in water. Other nitrogen-fixing bacteria form close associations with the roots of certain plants. Most of the nitrogen gets fixed by bacteria called rhizobia, which colonize the roots of leguminous plants, such as peas and beans.

Nitrogen makes possible our amino acids, which contain amino groups of NH_2, and our DNA and RNA, which need 2 to 5 nitrogens for every base.

How the rhizobia and legumes establish their partnership is a bit like dating. First, the plant sends signals, in the form of special flavonoid attractant molecules it secretes, to lure the bacteria to its roots. The bacteria respond by attaching themselves to the root hairs of the plant and sending a signal of their own, called a nodulation factor.

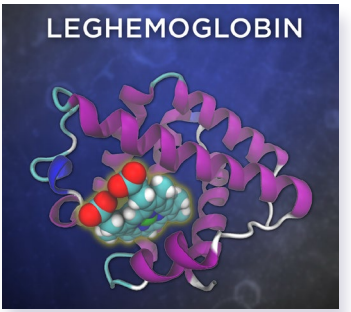


Different species of bacteria produce slightly different nodulation factors. And the specific bacterial nodulation factor determines whether the plant will allow them to colonize the roots of the plant. Once this compatibility is established, bacteria move into root cells, which divide to form little swellings on the roots called nodules. It is inside these nodules that nitrogen fixation occurs.

The first thing to remember about converting nitrogen from N_2 to NH_3 is that it gains electrons and gets reduced. The electron source for reducing nitrogen comes from a small iron-containing protein called ferredoxin.

The enzyme that catalyzes this reduction is called the nitrogenase complex. It is made up of 2 components. The first funnels electrons, one at a time, from an iron-sulfur protein (ferredoxin) to the nitrogenase. The second component catalyzes the reduction of nitrogen to ammonia by the nitrogenase. This results in each nitrogen atom forming bonds with 3 hydrogen atoms. In aqueous environments, NH_3 forms an NH_4^+ ammonium ion.

That seems simple enough, but here's why we don't fix nitrogen ourselves: Nitrogenases are inactivated by oxygen and need protection from it, so the root nodules produce a protein that removes oxygen by binding it with high affinity.



Recall that hemoglobin and myoglobin are proteins that bind oxygen. Interestingly, plants called legumes make a related protein called leghemoglobin, which resembles myoglobin. Leghemoglobin is a joint construction project between bacteria, which make the heme, and plant cells, which make the globin protein.

Leghemoglobin grabs and binds oxygen in the root nodules. Its affinity for oxygen is very high, about 10 times greater than that of hemoglobin. This sharply lowers the free oxygen concentration within the nodule, allowing nitrogenase to do its work to make ammonia (NH_3) or the ammonium ion (NH_4^+) in water. Plants can use these forms of nitrogen.

The ammonia made by nitrogen fixers can also be converted to other useful chemical forms, such as nitrite and nitrate, by other bacterial species. But plants are limited by the amount of nitrogen that is fixed by these microbes. If there isn't enough available nitrogen, plants don't grow optimally. To counter this, farmers used fertilizers, such as manure and fish, for thousands of years to increase crop yields.

Amino Acids

Plants and bacteria synthesize all 20 of the amino acids for making proteins, but animals, including humans, make only some of them, with the rest coming from diet.

Amino acids have a characteristic structure: The alpha carbon is attached to an amino group, a carboxyl group, and a hydrogen, plus an R group, or side chain, that is different for each amino acid. To make amino acids, cells start with molecules produced in metabolic pathways. Starter molecules include alpha-ketoglutarate, oxaloacetate, and pyruvate.

There is no single unifying pathway for synthesis of all of the amino acids. There are, however, a few commonalities called families, and that is how their synthesis is organized.

- The glutamate family starts from alpha-ketoglutarate of the citric acid cycle and goes to glutamate, which in turn is also progenitor to arginine and proline.

- The aspartate family, also from the citric acid cycle, starts from oxaloacetate and goes to aspartate and 5 other members: asparagine, isoleucine, lysine, methionine, and threonine.

- The serine family comes from 3-phosphoglycerate of glycolysis.

- The aromatic family begins with phosphoenolpyruvate (PEP) at the big bang step of glucose metabolism.

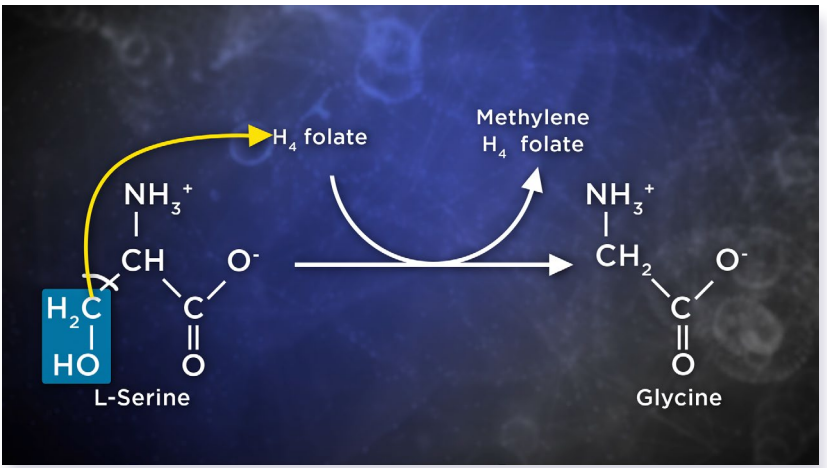
- The pyruvate family starts with pyruvate to make alanine, valine, and leucine.

- The histidine family traces its roots to ribose 5-phosphate from the pentose phosphate pathway.

The reactions leading to amino acids from their starting materials are as varied as the amino acids themselves, with some requiring one or 2 steps and others needing 10 or more. Some, such as lysine, can be made on more than one pathway. But there are 2 classes of reactions to focus on here.

1 Transaminations involve the transfer of an amino group from one molecule to another. The amine group transfers commonly involve glutamate as a donor or acceptor molecule. These reactions are catalyzed by enzymes called aminotransferases, which use a coenzyme called pyridoxal phosphate, which is a derivative of vitamin B_6.

2 One-carbon transfers involve the movement of a single carbon from one molecule to another. Transfers depend on several forms of folate/vitamin B_9. You see this type of transfer when the amino acid glycine is made from serine. In this reaction, one carbon from serine is moved to a folate (tetrahydrofolate), and glycine is what remains.



Together, a series of transaminations, one-carbon transfers, and other reactions allow us to build about ½ of our amino acids. The rest we get from food.

Our bodies take a use-it-or-lose-it approach to protein. Muscles are made of protein and can even be broken down during emergency, but muscle proteins are created to support the purpose of movement, not storage. We break down and excrete any dietary proteins that cannot be used right away, whether for new protein synthesis or oxidized for energy.

And there's a lot to get rid of. First, for any well-fed person, the intake of protein in the diet will generally exceed bodily needs. Second, proteins in the body are also in a state of constant change, with many enzymes turning over in hours or minutes.

Cells have 2 built-in ways to recycle protein.

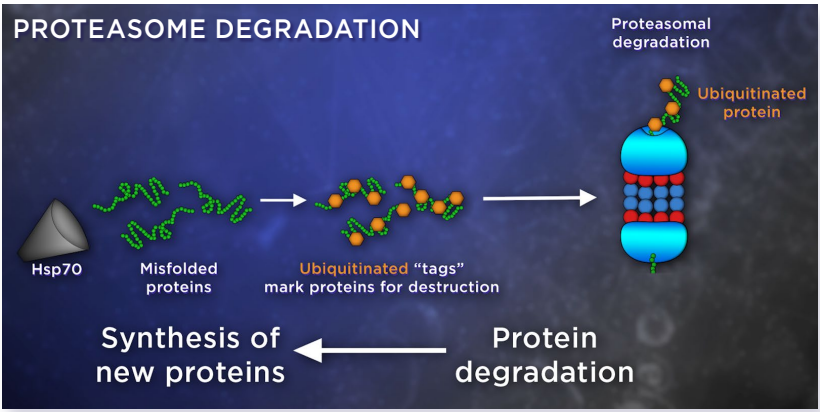
1. A lysosome is full of enzymes that are capable of degrading not only proteins to individual amino acids, but also all other sorts of biomolecules.

2. By contrast, the only job of proteasomes is to break down proteins.

Some of the amino acids released by protein degradation are reused in the synthesis of new proteins.

The breakdown of most amino acids is a bit like the reverse of their synthesis.

Transamination adds an amino group onto pyruvate from glycolysis to make alanine, and the molecules oxaloacetate and alpha-ketoglutarate from the citric acid cycle undergo similar transaminations. Those same intermediate molecules are recovered during the breakdown of most amino acids and can enter into the citric acid cycle, or other pathways in the cell, to be used for energy production.



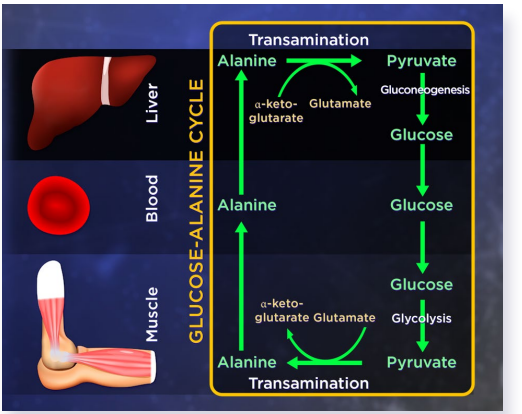
But when amino groups are removed from amino acids, where do they go?

The body must balance amines in the body because too many are toxic. Consequently, pathways for excretion are important. Transamination reactions just shuffle amino groups around. They are not lost—yet.

In brains, glutamate functions not only as a component of proteins, but also as a neurotransmitter, a molecule involved in signaling between nerve cells. In this role, glutamate gets broken down after use, and the ammonia released is highly toxic. Fortunately, there's a lot of glutamate in the brain, so the ammonia released gets added onto other glutamate molecules to make glutamine. And it is in the form of glutamine that nitrogen moves out of the brain to the liver via the bloodstream.

A different set of deamination reactions can occur in muscles, particularly during strenuous exercise or starvation. In muscle cells, amino groups get handed off to pyruvate, making alanine. The alanine is sent through the bloodstream to the liver, where it also gives up its amino group to make glutamate, and pyruvate, which the liver can use to make glucose to send back to the muscles.

This is not a cycle inside individual cells but a larger-scale cycle operating between muscles and liver, much like the Cori cycle with lactate and glucose. This one is known as the glucose-alanine cycle, and it keeps muscles supplied with energy in times of need. The glutamate, meanwhile, joins the pool of glutamate in the liver that will be broken down.



In our bodies, the liver plays a major role in recycling amino acids. It converts the ammonia released in their breakdown into urea, which is excreted mostly in urine but also in trace amounts in sweat.

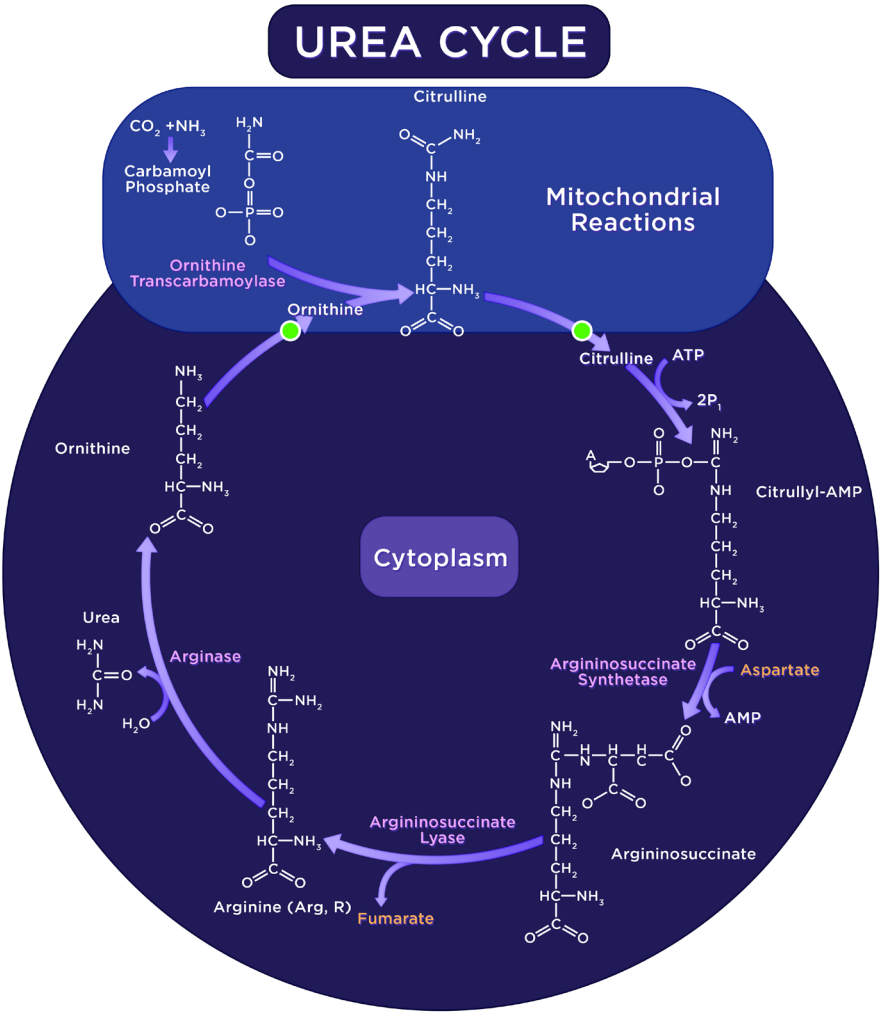
After the amino groups from amino acids are collected in the liver into a few amino acids, primarily as glutamate and aspartate, a major player is glutamate. It is broken down to release ammonia, which gets mopped up by the metabolic pathway known as the urea cycle.

With each turn of the cycle, one molecule of ammonia and one amino group from aspartate are combined with carbon from carbon dioxide to ultimately form one molecule of urea, which is very water soluble and provides an efficient means of lowering the body's load of nitrogen.

The urea cycle encompasses 6 reactions. Intermediates include 2 amino acids found in proteins, aspartate and arginine, and 2 other amino acids not used to make proteins, ornithine and citrulline. Urea is formed by splitting arginine, and the cycle is circular, like the citric acid cycle.

The urea cycle was worked out by Sir Hans Krebs of the citric acid cycle, also called the Krebs cycle.

One urea cycle by-product is fumarate, which is part of the citric acid cycle. Fumarate can reenter the citric acid cycle and get oxidized to oxaloacetate, which can then be transaminated to form another aspartate that feeds back into the urea cycle. So, fumarate and aspartate link the citric acid cycle to the urea cycle.



Nucleotides

In addition to its other roles, glutamine is also the nitrogen control point for creation of nucleotides, which are the building blocks of nucleic acids, another critical area of nitrogen metabolism.

Nucleotides have 3 components: a sugar, one or more phosphates, and a nitrogenous base.

All nucleotides are initially synthesized as ribonucleoside monophosphates—that is, there's a base, a ribose sugar, and one phosphate. If the nucleotides being made are for DNA, then the ribose sugar gets converted to deoxyribose. Additional phosphates are added to make the triphosphate nucleotides needed for building DNA and RNA.

There are 2 categories of bases: purine bases—adenine or guanine—and pyrimidine bases—cytosine, thymine, or uracil.

The source of the sugar phosphate for all nucleotides is a molecule called phosphoribosyl pyrophosphate (PRPP). But nucleotide assembly differs depending on whether it contains a purine or a pyrimidine. Purine nucleotides are assembled atop a PRPP molecule, while the pyrimidine bases are assembled completely before they are added to PRPP.

High levels of the purine adenosine triphosphate (ATP) favor synthesis of pyrimidines, while high levels of the pyrimidine cytidine triphosphate (CTP) stop pyrimidine synthesis. This is how purines and pyrimidines are kept in balance.

So far, the focus has been on the synthesis of ribonucleotides, the building blocks of RNA. To make DNA nucleotides, the ribose sugar in the ribonucleotides must be converted to deoxyribose. The other difference between RNA and DNA is that DNA replaces uracil with thymine-containing nucleotides.

When it comes to the breakdown of nucleotides, pyrimidines are simpler. They go straight to ammonia, which is removed as urea. Purine catabolism produces a less soluble molecule called uric acid that is normally excreted by the kidneys. If its levels in the blood are high, uric acid can form crystals that may accumulate near joints, causing a painful condition called gout. Uric acid can be excreted as is or converted into urea for excretion.

READING

Bröer and Bröer, "Amino Acid Homeostasis and Signaling in Mammalian Cells and Organisms."

QUESTIONS

1 Three of the most important molecules for cells are poisonous or form toxic compounds. Name them and describe the nature of their toxicity.

2 RuBisCO, which assimilates oxygen from the atmosphere, is the most abundant protein on Earth. Nitrogenase, which assimilates nitrogen, isn't even close. Why not?

3 Transamination is critical for the movement of amine groups, but it is not involved in the elimination of nitrogen from the body of higher animals. Explain.

CLICK HERE TO SEE THE ANSWERS.

20 EATING, ANTIOXIDANTS, AND THE MICROBIOME

The wear and tear of living means our cells are constantly in need of repair. The oxygen we breathe and the energy we extract from food require oxidation reactions that occur naturally in our cells, but these same reactions cause damage to DNA, proteins, and lipids. Our cells work to minimize the harm, but the accumulated damage plays a role in aging and in the development of chronic and degenerative diseases.

Free Radicals

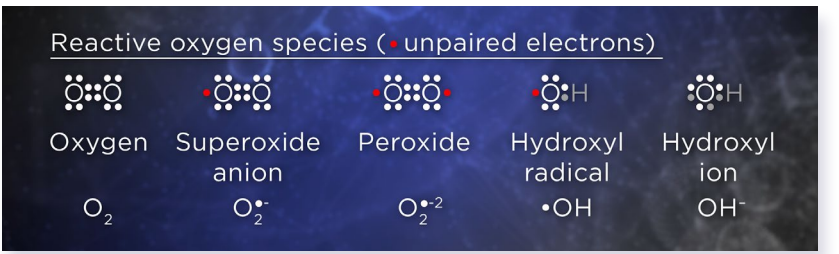
We begin to age the moment we are born. And this is primarily because of events in our cells that occur in the course of normal metabolism.

A cellular cause of aging is the production of free radicals: unstable atoms or molecules with one or more unpaired electrons. Electrons like to be in pairs, so atoms with unpaired electrons are highly reactive, meaning that these free radicals react with other molecules to steal their electrons.

Chain reactions of radical formation can occur for any free radical. When free radicals form in the cell, their chain reactions can disrupt membranes, DNA may be mutated, and the structure of proteins and lipids may be altered.

This is the inevitable by-product of living as our cells carry out oxidative reactions to obtain energy from food. The electron transport chain is essential for the powerhouse production of ATP in the mitochondria. But it is also a prime source of free radicals involving oxygen.

Normally, the electron transport chain carries electrons that combine with oxygen and protons, producing water. We cannot live without this process. But electrons sometimes escape from the electron transport chain and attach to other oxygen molecules, forming free radicals known as reactive oxygen species (ROS).



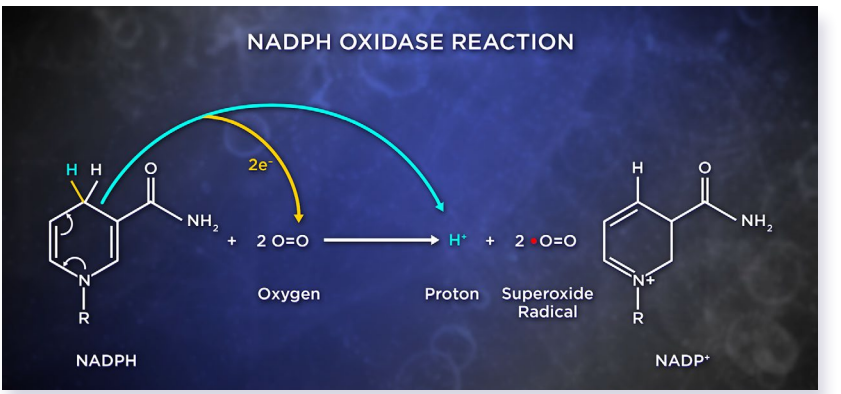
Ordinary molecular oxygen (O_2) has 12 electrons in nicely balanced pairs. But if an extra unpaired electron attaches as well, a free radical results called superoxide. Molecular oxygen can also have 2 extra unpaired electrons, one on each atom, making a free radical called peroxide.

A third kind of free radical starts with a single oxygen bonded to a single hydrogen. Most free OH groups in cells have a negative charge (OH^-). They can accept a proton, the charges balance, and water forms. But if hydroxyl loses an electron, although the charge becomes zero, that also leaves the molecule with an unpaired electron—a free radical. Like superoxide and peroxide, hydroxyl radicals are extremely reactive.

Besides mitochondria, free radicals can also be made during normal enzymatic reactions and when cells are exposed to x-rays, air pollutants, cigarette smoke, and industrial chemicals.

Another potentially damaging molecule made during normal cellular metabolism is hydrogen peroxide (H_2O_2). Unlike plain peroxide, hydrogen peroxide isn't a free radical, because all of its electrons are paired, but it very easily can form free radicals, including the hydroxyl radical. Consequently, H_2O_2 does kill bacterial cells, which is why it is in medicine cabinets—to kill bacteria. H_2O_2 does not harm our cells, because we have a protective enzyme that bacteria lack.

Though it may seem odd, cells sometimes turn molecular oxygen into superoxide on purpose. Superoxide is made in cells by an enzyme known as NADPH oxidase. Such an enzyme exists because macrophages and neutrophils in the immune system use oxygen radicals as weapons to kill bacteria.



So, there's no escaping reactive oxygen molecules—which steal electrons from, and oxidize, molecules they bump into. Over time, cellular components get oxidized by reactive oxygen reactions. Fortunately, cells have mechanisms to limit the harm done by oxygen free radicals. Enzymes that dismantle free radicals are one line of defense.

One such enzyme, known as superoxide dismutase, is so important that it has evolved to have the highest catalytic efficiency of any known enzyme. That speaks to the importance of protection from damage by superoxides, because if the enzyme doesn't disable the superoxide first, damage to an important cellular molecule may occur. In fact, speed is essential for every protective mechanism against radicals; it's a race between the damaging radical and the protecting material. When the race is lost, damage occurs.

In addition to enzymes, other defenses against free radicals are provided by antioxidant molecules, which prevent oxidation of other substances by themselves becoming oxidized. They take one for the team to spare other molecules. Antioxidants are small molecules themselves, not enzymes, and they work by bumping into radicals and giving them electrons.

While enzymes do help prevent oxidation of cellular components and are important in antioxidant defense, the term *antioxidant* specifically refers not to enzymes, but to molecules that are easily oxidized.

The most important ones are vitamins C and E. Vitamin C, also known as ascorbic acid, is water soluble. When completely oxidized, it loses 2 protons and 2 electrons to form dehydroascorbate. An intermediate, partially oxidized form of vitamin C results from loss of one proton and one electron. That's what's produced when vitamin C bumps into a radical.

Vitamin C also helps reduce oxidation from reactive oxygens in the cytoplasm's aqueous environment. But for parts of the cell, such as membranes, that do not contain water, the fat-soluble vitamin E can provide protection against oxidative damage if it bumps into a radical before the radical bumps into something else. Oxidation of fatty acids in lipid membranes can create a chain reaction, so stopping it is important.

In cells, the multiple layers of antioxidant protection work together. But sometimes there are more free radicals than protective enzymes and antioxidants can mop up. When these defenses are overwhelmed, the cells are said to be under oxidative stress, and damage can occur.

Oxidative Stress

Oxidative stress may be a factor in aging and is believed to play a role in the development of a whole slew of diseases. For example, emphysema arises from oxidative damage to a critical protective protein in the lungs called alpha-1 antitrypsin. Reactive oxygens have been implicated in atherosclerosis, stroke, arthritis, hypertension, cancer, vision loss, and neurological conditions.

Reducing oxidative stress, then, seems to be a no-brainer for good health. And that means reducing things that increase oxidative damage and maintaining optimal levels of antioxidants in our bodies. But this is not a simple matter of popping pills to supply antioxidants.

There are 2 immediate reasons, supported by many studies, why increasing intake of antioxidants cannot solve all oxidation-related problems. First, the antioxidant balance in cells is complex. And because of the interdependence of various antioxidants, increasing the intake of some haphazardly might be useless at best and harmful at worst.

Large-scale studies done to test whether antioxidant supplements improved health outcomes show that many effects obtained are minimal. A second point against antioxidant supplements is that a small amount of oxidative stress may actually induce our cells to increase their defenses. Instead of flooding cells with supplements, it might be better to find ways to stimulate the cell's intrinsic antioxidant defenses.

A related approach is to study people who exhibit the signs of good health and observe what they eat. In numerous studies, people who enjoy better health have been found to eat the most fruits and vegetables and whole grains, together with nuts, oils, a serving or 2 of dairy products, and small amounts of plant-based or lean animal protein. They also limit red meat, saturated fats, and refined carbohydrates. What are the biochemical reasons for this?

Many nutritionists advise people to load plates with colorful vegetables—because the compounds that make vegetables so attractive are also good for you.

A typical Western diet is high in refined carbohydrates like white flour, simple sugars, animal protein, and saturated fat. If you compare that diet with one with large amounts of vegetables and fruits, more like a so-called Mediterranean diet, the first noticeable difference is energy density: There's a lot more energy in meat, fat, and refined carbohydrates than there is in vegetables and fruits.

We obtain this energy by the oxidation of food molecules that eventually culminates in the electron transport chains in mitochondria. The more food there is to burn for energy, the more the electron transport chain runs. And the more it runs, the greater the chances are of generating reactive oxygens.

Highly caloric meals are likely to increase the number of free radicals. Because too many calorie-dense foods generate high levels of reactive species that can cause damage, one way to reduce oxidative damage may be to eat foods that are less energy dense. By substituting vegetables and fruit and whole grains, all of which are high in fiber but lower in energy content, it is possible to feel full without generating as many free radicals.

The other advantage of eating lots of fruits and vegetables is that plant foods are naturally high in antioxidants. Many of the secondary metabolites made by plants are antioxidants.

Even if you hate vegetables, you probably love one or more plant-based drinks, such as orange or pomegranate juice or almond or soy milk. Even coffee, tea, and wine contain complex mixtures of bioactive compounds, many with antioxidant, anti-inflammatory, and even anticancer effects.

Plants produce molecules that reduce oxidative damage and are healthy in other ways, too. Nuts and oils contain vitamin E and unsaturated fats that make it easier to absorb fat-soluble vitamins and are associated with better cardiovascular health.

But if we know what these compounds do, why not just take a pill that has them all? Eating the foods seems to provide benefits that pills don't. It may be that the particular combinations or proportions of the molecules found in plant foods may be important for their activity, or there may be additional compounds in the whole foods we don't know about.

The Gut Microbiome

An additional reason has emerged for why eating plant-derived foods is better than taking supplements: The trillions of bacteria that live in our guts affect our health. The gut microbiome, as it is referred to, is turning out to be an enormous factor in whether we are fat or thin, in good health or sick, and even whether we feel upbeat or depressed.

The roughly 2 to 6 pounds of microbes in the colon help protect us against pathogens, train our immune cells to know what to target, and contribute to metabolism through the compounds they make and release into our bodies. Our biochemistry is interwoven with that of our microbial guests.

It turns out that each of us, when healthy, has more than 1000 species of bacteria that make up our microbiome. A diverse mix of species predicts better health than having fewer kinds. And the specific kinds of bacteria and their proportions may have a bearing on, for example, whether you tend to gain weight or stay skinny.

Your colon bacteria also produce a variety of neurotransmitters—chemicals that are used for signaling in your nervous system and brain. More serotonin is made by your microbial guests than is made in the brain.

What we eat feeds us, but it also must feed our gut microbes. It turns out that what we feed them determines which species flourish.

So, healthy eating is not directed just at our cells, but also at our microbes. And we must eat plant foods to get the nutrients needed for health-promoting effects.

The more easily that foods are digested and absorbed in our small intestines, the less of it passes on to our microbes in the colon. If much of what we eat is absorbed by us, there are 2 ill effects: We gain weight and our guests starve.

What's more, the undigested plant material that enters the colon contains some compounds that bacteria convert into useful products for us. If we took those same plant compounds as purified supplements, they could be directly absorbed in the small intestine and never get processed by microbes in the gut. In addition, when plant compounds don't make it to the colon, the composition of the microbiome changes, and not necessarily in a positive way.

So, feeding our microbiome is best done by eating foods that are less digestible. This is one of the reasons that juicing and making smoothies, even if the drinks contain kale, is not as good as eating whole fruits or salads. The pulverized food releases nutrients for easy absorption, but less of it ends up feeding your bacterial friends.

Overcooking can potentially cause similar issues, but there are very good reasons to cook food: to kill disease-causing organisms, inactivate harmful compounds, and enhance taste.

How food is cooked also makes a difference. Grilling meats, for example, can create a number of cancer-causing chemicals. Cooking can also cause the formation of compounds called advanced glycation end products (AGEs).

Glycation is the haphazard bonding of sugar to a fat or protein without guidance from an enzyme. Many foods contain some glycation end products to start with, but these increase when cooked, especially when cooked with dry heat.

At low levels, glycation harmlessly contributes to many of the yummy flavors made possible by cooking. But at high levels, AGEs are associated with increased oxidative stress and inflammation as well as possible formation of carcinogens.

Using cooking methods that include moisture, such as braising or poaching, reduces the formation of glycation products. Marinating meat in acidic solutions like lemon juice or vinegar before cooking also limits AGEs.

Our metabolic pathways have daily rhythms, and limiting the time during which we eat to the most active times of day is conducive to the proper regulation of feeding.

Periodic fasting stresses our bodies gently and stimulates all manner of repair processes. This is also what happens to us during exercise, when transient stress increases restorative mechanisms.

READINGS

Gluckman and Hanson, *Mismatch*.
Yong, *I Contain Multitudes*.

QUESTIONS

1 When superoxide is converted to molecular oxygen by superoxide dismutase, describe what is getting oxidized and what is getting reduced.

2 Given what you've learned about free radicals, what kinds of reactions can hydroxyl radicals start that hydroxide ions cannot? Why can hydroxyl groups not participate in these reactions?

CLICK HERE TO SEE THE ANSWERS.

21 HORMONES, STRESS, AND CELL DIVISION

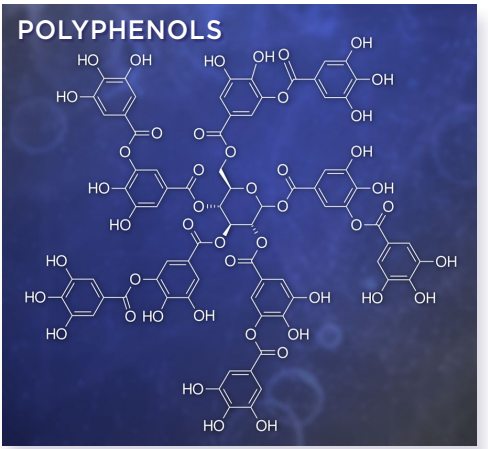
Inside our bodies, communication systems—invisible to us yet constantly active—relay information about conditions internal and external to our cells. Cellular communication depends on specific molecular interactions, where the message and the receiver are biomolecules. Binding of a message by the cellular receptor molecule triggers information transmission inward to the cell and changes its activities.

Cellular Signals

Cellular communications always involve sensing and signaling. It has long been known that even bacteria sense and move toward food sources and away from toxic substances. Even more interesting, bacteria talk to each other, kind of like teenagers at a mall, to find out how many "friends" are nearby.

Plants sense and communicate in response to light, gravity, temperature, and water. Inside a plant, messages carry information for specific cells. Some plants recognize family members—that is, genetically related individuals. Such plants share resources with relatives by limiting the amount of soil nutrients they take up. If they sense a stranger, though, they compete by putting out extra roots to take up more nutrients. Molecules exuded from roots may help with distinguishing strangers from family.

Plants also respond to insect attacks by releasing volatile organic molecule signals into the air around them to alert other plants to predators. The signals trigger increased production of chemicals like polyphenols that affect the taste of the plant or proteinase inhibitors that inhibit digestion of the plant by the predator.



Not surprisingly, the most complex communication systems are found in multicellular animals. Many of the signaling molecules come from the endocrine system and are generally known as hormones.

Hormones can be divided into 2 broad groups:

1. molecules that act within minutes by binding to receptors on the surface of cell membranes, and

2. steroid hormones that act more slowly and bind to receptors inside of cells.

In animals, hormones originate in a few specialized organs that make up the endocrine system. These glands release hormones only inside ("endo") the bloodstream. Major endocrine glands include the hypothalamus, pineal gland, pituitary gland, thyroid gland, parathyroid, thymus, adrenals, pancreas, and ovaries/testes.

How signals get to their target cells is varied.

- Signaling in animals that uses the circulatory system is called endocrine signaling.
- Signaling that only involves neighboring cells is called paracrine signaling.
- A signal molecule that acts only on the outer surface of the same cell where a signal was created is called autocrine.
- An intracellular signal molecule that stays and signals inside the membrane of the same cell is called intracrine.
- There are also exocrine glands that release outside the body, such as sweat glands that send signals to communicate with other organisms.

Cells receive a signal starting when the signal molecule binds to a receptor protein, often located on a cell's membrane. But for steroid hormones, and also thyroid hormones, the receptor is located inside the cell, so such signals must cross the membrane and enter the cell to meet up with their receptors.

For intracellular receptors involved in steroid signaling, a typical outcome is that cells adjust the amount and/or timing of the synthesis of specific proteins. For signaling occurring through membrane receptors, the receptor protein conveys the message through a chain of other cellular molecules into the interior of the cell.

The Stress-Response System

In some types of membrane receptor signaling, enzyme activities are altered. Affected enzymes include protein kinases, which phosphorylate target proteins, and phosphatases, which remove phosphates. Adding and removing phosphates is an easy way to control whether cellular proteins are active or not.

A good example of such signaling is the fight-or-flight response that evolved to prepare an organism in a dangerous situation for fighting with or fleeing from predators or competitors. That same stress-response system kicks in when we get cut off in traffic by an aggressive driver, interview for a job, or even make a presentation before the boss.

Any response to stress takes energy, as well as a toll. The energy involves mobilizing stores of food to provide additional glucose. It begins in a section of the brain called the amygdala, which is involved in memory, decision making, and emotions such as fear, aggression, and anxiousness.

When the brain perceives danger, it sends messages to the adrenal glands above the kidneys, stimulating them to secrete hormones: epinephrine (also known as adrenaline), norepinephrine, and cortisol.

Adrenaline is familiar to most people and is often credited with giving people extraordinary strength or endurance under stress. At the molecular level, there are quite a few steps in the pathway from adrenaline binding at the cell surface to the final cellular changes, but there are 3 important themes:

1. The signal arriving at a cell must be bound by a receptor for any action to follow.

2. The receptor passes the message to a chain of molecules until finally a kinase or phosphatase is activated.

3. The relevant enzyme either phosphorylates or dephosphorylates target proteins and changes their activity.

As a result of adrenaline signaling, a series of reactions is set in motion, culminating in providing plenty of glucose to the body, both by breakdown of glycogen and new synthesis of glucose in the liver. The resulting flood of glucose in the bloodstream helps fuel the stress response.

The enzymes involved are activated by the addition of a phosphate, so the off switch involves the removal of the phosphate that was added. An enzyme known as protein phosphatase removes phosphates from the enzymes when there is an abundance of glucose.

In addition to the short-term stress response for energy production by adrenaline, a longer-term response to stress comes with the production of cortisol, a steroid hormone that is a member of the class known as glucocorticoids.

Cortisol's effects overlap somewhat with those of adrenaline, but in addition, cortisol keeps the system on high alert over extended periods of time. Normally, once adrenaline has done its thing, the system returns to the "no danger" state. But if the threat is perceived to be ongoing, cortisol keeps the stress pedal depressed.

Like other steroid hormones, cortisol exerts its effects by binding to a steroid hormone receptor inside of cells. When a steroid binds, the steroid hormone receptor acts to affect the synthesis of specific proteins. This route of communication is relatively slow and may require hours to fully activate.

Among the proteins whose production is increased by cortisol are the enzymes used in gluconeogenesis. To obtain even more glucose, cortisol stimulates fat and muscle breakdown to provide fatty acids and amino acids that could be funneled into glucose synthesis. In addition, cortisol can suppress activity of the immune system in an attempt to divert all resources to dealing with immediate threats. These actions of cortisol are useful in continuing to provide energy to an organism that needs to get its brain and body geared for action in the face of continued danger.

The downside is that low levels of chronic stress, such as worrying about being able to pay your bills, also trigger the cortisol response. In addition to elevating your existing blood glucose levels, even when you aren't literally going to fight or flee, cortisol also stimulates your appetite so that you load up on calories even more.

Stress eating and chronically elevated blood glucose are not recipes for good health. Over time, increased blood sugar and chronically elevated appetite, coupled with loss of lean muscle mass, can lead to obesity and insulin resistance.

Paradoxically, there's a partial upside. Cortisol's suppression of the immune system can be exploited in treating diseases related to inflammation and overactivity of the immune response, such as rheumatoid arthritis and allergies as well as eczema. The long-term use of glucocorticoids, though, does have potential negative effects, such as loss of bone density, weight gain, cataracts, and glaucoma. Research into drugs that could deliver the benefits without the side effects is ongoing.

Another signaling molecule that is involved in the brain's response to stress is norepinephrine, which also functions to maintain alertness, readiness for action, and normal levels of respiration. Norepinephrine is also produced at low levels normally and acts similarly to epinephrine.

Cell Division

There is another category of signaling in which external signals can induce changes in cells that tell them to divide. This pathway also relies on receptor binding, relaying of the message through various go-betweens, and finally activating kinases that bring about changes in the cell.

A classic example of such a signaling system involves a hormone called epidermal growth factor (EGF), one of the functions of which is to stimulate cells to divide. EGF is a polypeptide that binds to a cellular receptor known as the epidermal growth factor receptor (EGFR), which belongs in a class of receptors called receptor tyrosine kinases.

There is an entire family of proteins, each slightly different from the other, that can serve as receptors for EGF and related peptide signals. All of the receptors are transmembrane proteins that are anchored in the cell membrane. One end of the receptor peptide sticks out of the cell, while the other end extends into the cytoplasm.

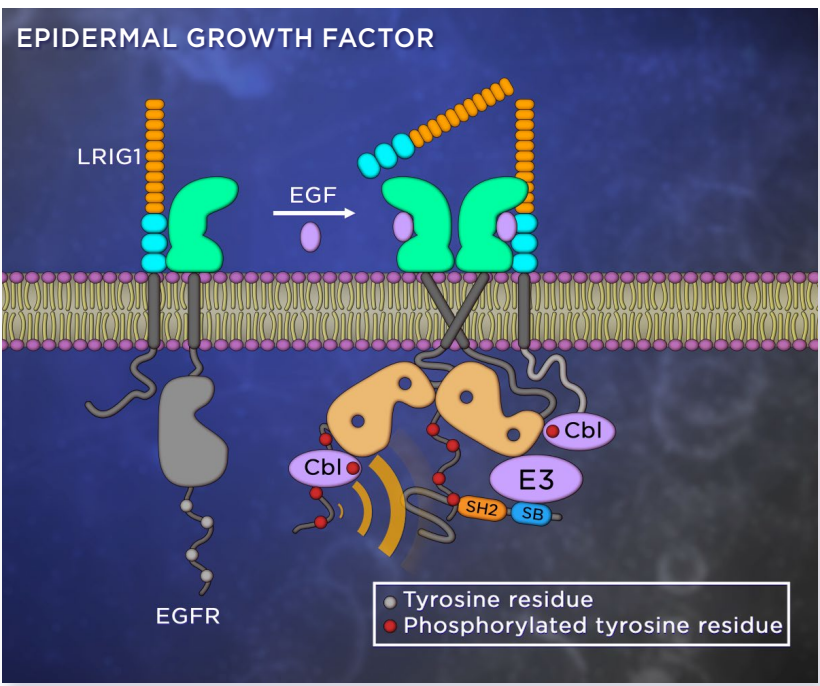
Receptors for EGF work in pairs, with 2 of them needed to pass on the message when EGF binds to them. Before they encounter EGF, the receptors exist as monomers. When EGF shows up, 2 monomers move closer to each other, drawn together by their shared affinity for EGF. When they get close

The name *receptor tyrosine kinases* derives from the fact that they have a latent kinase activity. Kinases add phosphates, and the EGF receptors specifically add phosphates onto tyrosines in proteins.

enough, the kinase activity in tails of the receptor monomers gets stimulated, and the receptor tail phosphorylates its partner on tyrosines in its amino acid sequence.

The phosphorylated tyrosines on the EGF receptors act as beacons, attracting other proteins in the signaling pathway located in the cytoplasm. Several different proteins, upon interacting with the phosphotyrosines, then activate other enzymes that pass on EGF's message.

One enzyme pathway that is very important in stimulating cells to divide is known as the MAP kinase pathway. It involves numerous intracellular protein messengers, including ras, which is inactive when carrying guanosine diphosphate (GDP), but when it binds guanosine triphosphate (GTP), it is activated and stimulates the rest of the pathway.



The ras signaling pathway involves a kinase cascade, and the last kinase in the cascade is mitogen-activated protein (MAP) kinase, where a mitogen is a *mitos*is *gen*erator, or something that induces cell division. MAP kinase brings about the changes that lead to cell division by phosphorylating a slew of target proteins.

The interesting aspect of EGF signaling is how the pathway gets turned off. This is important because uncontrolled cell division is cancer. There is more than one off switch, but an important one is found at the ras step. Ras slowly hydrolyzes GTP to GDP, thus inactivating itself over time. Mutations that destroy ras's ability to hydrolyze GTP leave it perpetually in the "on" state. And when that happens, cells continue to receive the message to divide, even when they shouldn't.

READINGS

Hancock, *Cell Signaling*.
Niehoff, *The Language of Life*.

QUESTIONS

1. Some say that blocking the action of molecules used by pathogenic bacteria for quorum sensing is the best way to block an infection, but others say that an argument can be made for spreading these molecules broadly. Discuss this latter strategy.

2. Exocrine signaling is used as a signaling mechanism by many animals. Describe what it might be communicating.

3. Imagine that you gave cells an inhibitor to stop conversion of GTP to GDP. Predict the physiological effects it would have based on the information in this lecture.

CLICK HERE TO SEE THE ANSWERS.

22 NEUROTRANSMITTERS, THE BRAIN, AND ADDICTION

Our bodies are wired with incredible nerve circuitry, just like houses are wired to provide electricity to lights and all the other conveniences of modern life. And like the light switches and other controllers for the electrical circuits in our houses, nerve cells are activated by electrical signals from tiny neurotransmitters.

The Nervous System

The nervous system can be divided into 2 main parts:

1. The central nervous system (CNS) is made up of the brain and the spinal cord.
2. The peripheral nervous system (PNS) consists of all the nerves that branch off from the spinal cord and extend to all parts of the body.

Fast-acting hormones like adrenaline take minutes to act, while slower-acting steroid hormones like cortisol may require hours. By contrast, neurotransmitters send signals that work in fractions of a second.

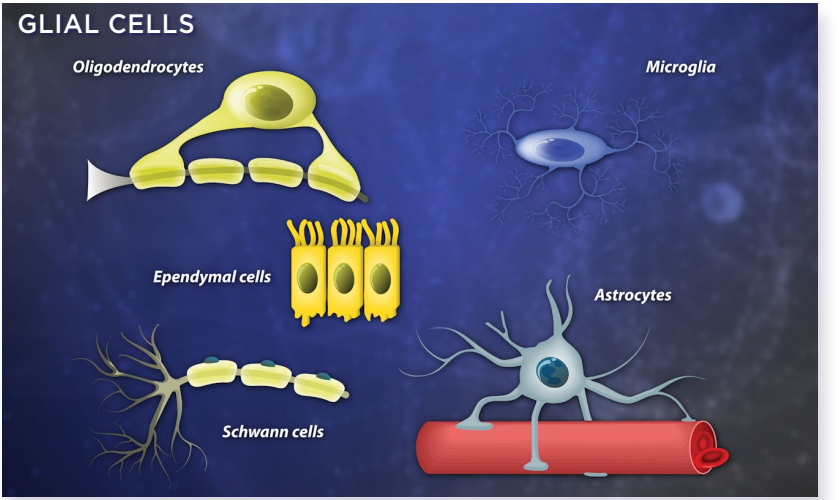
The peripheral system controls both our voluntary movements and the involuntary ones, such as breathing and the beating of your heart. The central system guides conscious and other actions. The CNS and PNS are interconnected parts of a functional whole.

The brain and spinal cord are connected, so you could think of the spinal cord as being a sort of tail that hangs off the base of the brain, and runs down your back, protected by the spinal column made up of bony vertebrae. At regular intervals along the backbone, nerves emerge from the spinal cord. There are 2 kinds of PNS nerves connecting the spinal cord to the rest of the body. One kind, the afferent nerves, brings information in from the various parts of the body to the spinal cord. The other, called efferent nerves, carries signals from the brain and spinal cord to the rest of the body.

Your brain, spinal cord, and nerves are made up of 2 major kinds of cells.

1. The neurons do the work of receiving and transmitting information. Everything you have ever felt, thought, learned, or sensed—every movement you've made and every experience you've had—is a result of the activity of your neurons.

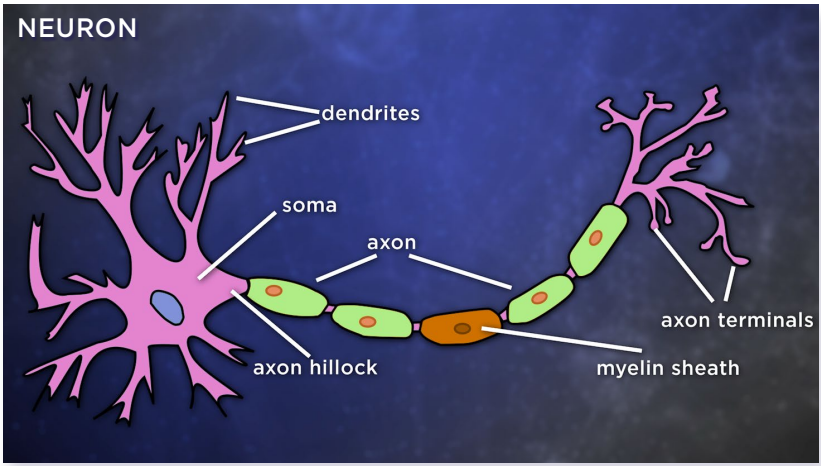
2. The glia—from the Greek word for "glue"—hold neurons together and play important supporting roles in your neurons. Among other functions, glial cells regulate the levels of chemical signals available near neurons, dispose of dead neurons, and even take responsibility for immune defense of the CNS.



The main body of a neuron is called the soma, and it contains the cell's nucleus and most of its other organelles. Bringing input into the cell body are branching structures called dendrites, which are the neuron's receivers—the part of the neuron that receives signals from other neurons. Dendrites on a neuron can get signals from many other neurons simultaneously. It's the combination of those signals that tells the neuron whether it should fire and send a signal onward.

The axon running from the spinal cord to your big toe is about 3 feet long—and it's a single cell!

Stretching away from the cell body is a single axon that carries the cell's output. It's like a long cable that runs from the neuron to the cell it sends information to. The axon starts from a special point called the axon hillock and carries the outgoing signal.



Axons can branch, so a single neuron can send messages to multiple cells. At the tip of each axon branch is a bulbous structure called the axon terminal or presynaptic terminal, which is the part of the neuron that will send signals to the dendrites of another neuron.

The axons of some neurons are coated in an insulating layer, the myelin sheath, made by glial cells. The myelin sheath allows the neuron's electrical signals to be sent speedily and efficiently.

To send messages, neurons use electrical signals called action potentials that move down the axons from the soma. These action potentials send signals from the neuron to its own axon terminals.

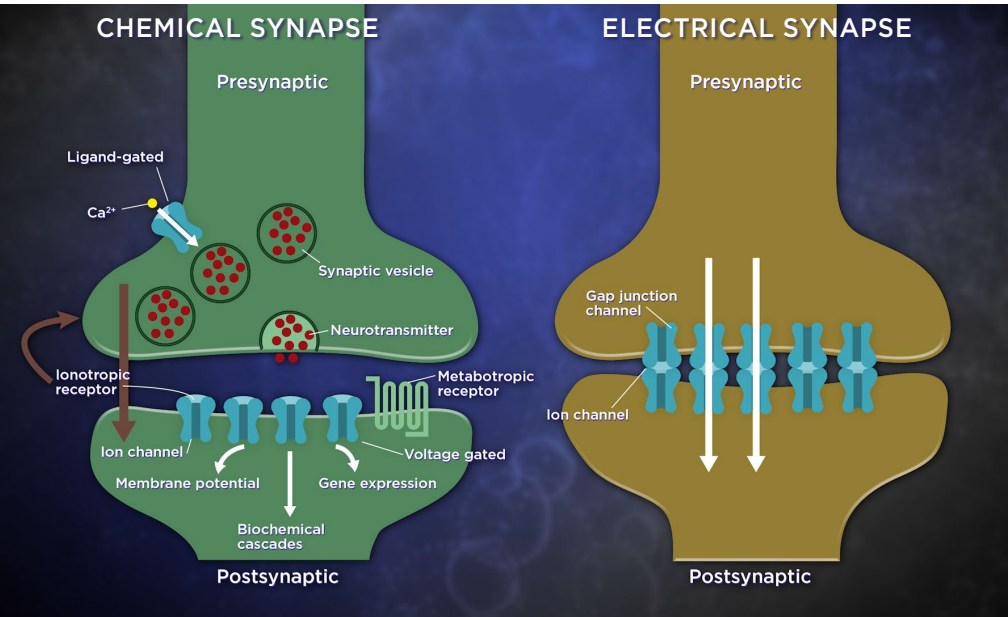
At the axon terminal, the signal must then be handed off to the next cell. This could be another neuron, or perhaps a muscle cell. Sometimes, the receiving cell is physically connected to the first neuron through gap junctions, which actually involve direct connections between cells with no "gap" between them. The electrical signal of a nerve cell is directly passed from the originating neuron through the junction to the cell receiving the message. The sending and receiving cells effectively function as a single unit.

But in many signals passing from neurons, there is a physical gap that the electrical signal cannot cross between the axon terminal of the neuron and the target cell. This gap without a connecting junction is called a synaptic cleft. At this point, the axon terminal will release chemical signals called neurotransmitters that diffuse across the synaptic cleft and bind to receptors on the next cell to pass on the message. The cell on the receiving end of the synaptic cleft is called the postsynaptic cell.

Both the electrical and biochemical signaling of neurons depend on the actions of special channels in the cell membrane. These are tiny pores that control the movement of ions in and out of the cell. Because their function is to regulate ion traffic, these pores are called ion channels, and they open and close as needed.

Ion channels are extremely picky. Each one is specialized for a particular ion.

Opening an ion channel is like opening the gate in a dam. Ions rush through the channel at a rate of millions of ions per second. Whether they rush in or out depends on which side of the membrane has the higher concentration of the ion. Just as water always flows downhill, ions move down concentration gradients.



To regulate the flow of ions across membranes, channels have gates. Membrane channels are made up of proteins, so the gates open or close as proteins making up the channel change in conformation. Ion channels involved in neurotransmission have gates that are opened by either an electrical signal or the binding of a molecule.

Ion channels cause major changes in the concentration of ions within cells when the channel opens. This changes the relative amount of an ion on the 2 sides of a membrane. When the number of ions and electrical charge outside of a membrane do not equal the number of ions and charge inside, the membrane is said to be polarized. This charge difference across the membrane is called the membrane potential.

Neurotransmitters

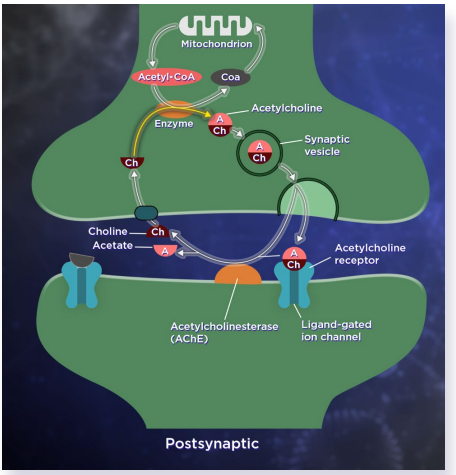
Neurotransmitters are small molecules made by neurons to cross the gap of the synaptic cleft. In the absence of a signal, neurotransmitters are stored in small membrane sacs called synaptic vesicles near the axon terminals. The arrival of the action potential at an axon terminal is a signal to vesicles containing the neurotransmitters to fuse with the cell membrane and release their contents into the synaptic cleft.

The major neurotransmitter used at neuromuscular junctions is acetylcholine. When acetylcholine is released, it diffuses across the synaptic cleft to the muscle cell, whose membrane has ion channels. Acetylcholine molecules bind to these channels and cause them to open, causing depolarization of the muscle cell membrane. This triggers a series of changes within the muscle cells that stimulates them to contract, causing movement.

After acetylcholine opens ion channel gates, it is broken down quickly by an enzyme of the muscle cells called acetylcholine esterase. This action is important for 2 reasons.

1 It prevents prolonged muscle contraction that would cause the muscle to spasm.

2 It releases the remnants of the neurotransmitter—in this case, choline—into the synaptic cleft for recycling. The choline is taken back up by the neuron and reused in making more acetylcholine.



So, muscle movement is dependent on the delivery of neurotransmitters from the axon of a neuron nearby. Anything that blocks this neuromuscular signaling pathway blocks muscular contraction. And that is what the neurotoxins in the venom of some snakes do. Muscle paralysis following snakebite is common.

Some snake toxins block the acetylcholine receptors on muscle cells and prevent transmission of the nerve signal. Unless the victims of snakebite receive prompt treatment to counter the effects of such neurotoxins, they will die when the toxins immobilize the muscles involved in breathing.

In addition to communicating with cells, such as muscle cells, using neurotransmitters, neurons can also communicate with each other. The axon of the neuron that is sending the message releases neurotransmitter molecules. Recipient neurons have receptors that bind these signals and initiate changes within them.

In the brain, large networks of neurons communicate with one another and with the spinal cord and nerves to control the activities of the body. Neurotransmitters are central to all of these activities. These signaling molecules work to promote or prevent the production of an action potential in the cells they affect. In the case of neuromuscular signaling, acetylcholine acts to stimulate the production of an action potential in the muscle cell.

Neurotransmitters that enhance production of an action potential in its target cell are called excitatory. If the binding of a neurotransmitter to its receptor stops creation of an action potential in the target cell, it is called inhibitory.

A given neurotransmitter may function as excitatory or inhibitory, depending on what kind of receptors it binds. Acetylcholine is excitatory when binding to skeletal muscle receptors but is inhibitory when binding receptors on heart muscle cells, where it slows down heartbeat.

Serotonin and Dopamine

Some neurotransmitters affect the way we see the world. Serotonin is popularly associated with happiness and well-being. The vast majority of serotonin in the body is produced and acts in the gastrointestinal system, where it is necessary for normal gut function. Pathogenic amoebas producing their own serotonin can cause diarrhea in humans.

But neurons also produce serotonin in the central nervous system, where it positively affects mood, appetite, and sleep. Serotonin is also involved in learning and cognitive functions and is a target of action by pharmacological antidepressants.

Another important neurotransmitter in the central nervous system is dopamine, which is commonly referred to as the feel-good chemical due to a pathway called the mesolimbic dopamine system—the most important

reward pathway in the brain. Dopamine's normal function is to make you associate good feelings with things that are necessary for survival and reproduction, such as eating and sex, so that you keep doing them.

The 2013 update to the Diagnostic and Statistical Manual of Mental Disorders added gambling disorder to the same category as substance abuse. Smartphones and video games may also activate the same pathways.

When you do something good for your survival or propagation of your genes, neurons in your brain's mesolimbic dopamine pathway release a small dose of dopamine. But these are also pathways that get hijacked in drug addiction and in behaviors like compulsive gambling.

Drugs can enhance activities of reward circuits by inducing neurons in the brain to release large amounts of dopamine. Amphetamines work this way. Other drugs, such as cocaine, act by inhibiting the removal of dopamine from the synaptic cleft. Dopamine release may also be stimulated by nicotine and by opioid drugs.

Each mechanism increases the level of dopamine much more than the small natural increases of eating a good meal, for example. This originally led scientists to believe that dopamine increases were the highs that drug addicts sought.

But researchers now think that the actual feelings of pleasure result from the release of other neurotransmitters in different parts of the brain following dopamine releases. Instead, dopamine appears to be involved in setting up a conditioned craving for pleasure.

Repeated use of the drug and repeated floods of dopamine stimulate downstream pathways and lead to a learning process. And repeated communication between neurons results in physical as well as biochemical changes in the brain.

And because the amount of drug-induced dopamine is so much larger than the brain's normal little spurts, the brain responds by making fewer dopamine receptors, helping to reduce the effects of a flood of dopamine.

But a dialed-down level of dopamine binding creates other problems. The downstream pleasure circuits don't get stimulated to the same extent, so the same dose of drug will no longer produce the same effect. At the same time, even though the drug may no longer provide the pleasure it once did, drug craving then begins in the rewired brain, leaving the user unsuccessfully trying to achieve the original high.

Meanwhile, the normal, smaller elevations of dopamine from everyday feel-good activities no longer satisfy. This makes addicts lose interest in things that once mattered to them, while drug cravings drive them to focus on obtaining the drug. A variation on this theme leads to drug dependence on opioid drugs and prescription painkillers.

READINGS

Bloom, *Best of the Brain from* Scientific American.
Erickson, *The Science of Addiction*.
Linden, *The Compass of Pleasure*.

QUESTIONS

1 Given that signaling in neurons involves an ion "wave," why are neurotransmitters even needed at all?

2 Imagine a snake that has a venom that increases the activity of acetylcholinesterase. Predict the effect its venom might have on neuromuscular signaling.

CLICK HERE TO SEE THE ANSWERS.

23 THE BIOCHEMISTRY OF OUR SENSES

We can understand the power of molecules for the thousands of reactions in biochemistry, but most of those reactions largely occur outside our everyday awareness. The most important group of biochemical signals that we notice is the molecular reactions that give rise to our 5 senses—the signals that give us virtually all of our conscious experiences.

Taste

We detect 5 kinds of taste: sweet, salty, sour, bitter, and savory. The savory taste is sometimes called umami, derived from a Japanese word meaning "delicious."

1 Sweet tastes signal detection of energy-rich carbohydrates, which is why we are drawn to sweet food.

2 Salty tastes encourage us to obtain the sodium and potassium necessary to maintain water and salt balances.

3 Sour tastes may indicate spoilage for detecting foods that are past their prime.

4 Many poisonous substances are bitter, so detecting bitter tastes could be lifesaving.

5 Umami gets its savory taste from the amino acids glutamate and aspartate, alerting us to the nutrients in protein-rich foods.

But all the yumminess of food cannot be explained by a mere 5 or 6 tastes. Much of the sensory information we get about food comes not just from our taste buds, but also from our noses. The ability to distinguish smells is much more finely tuned than our ability to taste.

Other senses contribute to what we call "taste" as well. The feel of food, known as texture, is simply a sense of touch. The temperature of food also contributes to flavor.

The small bumpy structures scattered over your tongue's surface are called papillae, and they hold the taste buds, each of which may contain 50 to 100 sensory cells. In contact with the sensory cells are the dendrites of the sensory nerves that carry taste information to the brain.

A given taste bud can sense multiple tastes. But how do the sensory cells sense the taste molecules? In each case, the receptors, which are transmembrane proteins, are crucial.

Salty and sour tastes are the simplest ones. For detecting salt, sensory cells with a sodium channel depolarize when sodium enters them. This in turn opens calcium channels that initiate the action potential. As for all of the senses, an action potential sends a message to the brain, which responds by giving you the sensation—in this case, a salty taste.

When you have a cold, it's the temporary lack of smell that explains why everything "tastes" like Styrofoam.

Likewise, sour tastes, which result from acidity, are sensed when the acid's hydrogen ions traverse an ion channel and enter a sensory cell. The positively charged protons cause membrane depolarization when they enter the cell and set off an action potential.

The other 3 kinds of tastes—sweet, umami, and bitter—are sensed with the help of G-protein-coupled receptors (GPCRs), which are cell surface receptors that bind to signals and pass on the message with the help of G proteins.

Smell

The sense of smell, known as olfaction, is enormously important to animals—not just for enjoying food, but also for warning of danger and even selecting mates. It occurs as a result of the binding of odorant molecules to olfactory receptor cells in the nasal cavity.

Smells reach receptors by 2 routes. The primary one is via the nostrils. A second one is located where the roof of the mouth connects the throat to the nose. Chewing food releases odorants that can travel the second path to the olfactory receptors.

Mucus—a viscous fluid full of glycoproteins called mucins—covers the lining of the nose, the nasal epithelium, and serves as both a solvent for odorant molecules and an important protection against infection, because nasal neurons connect directly to the brain. Odorant receptor cells are located in the olfactory epithelium.

Each olfactory neuron expresses a single kind of receptor. But each receptor can bind several different odorants—some weakly and others more tightly. A particular odorant may bind more than one kind of receptor, again with different affinities. These combinations give humans the ability to distinguish many more than 350 smells. In fact, a 2014 study put the number of smells humans can tell apart at one trillion.

Mammals have about 1000 to 1500 olfactory receptor genes. In humans, only about 350 of them make active receptors, but rats have almost 1500 genes that make active olfactory receptors.

The sense of smell may play a role in the choice of sexual partners, even in humans. Female humans, as well as fish and mice, can apparently detect, by smell, a component of the immune system known as the major histocompatibility complex (MHC). And there is evidence that females prefer the smell of males with different MHC genes than their own. Children of such couples might have a more robust immune system, which would give them an evolutionary advantage.

Vision

The detection of light and the ability to distinguish detail by our eyes involves an amazing convergence of optics, light-sensitive proteins, nerve signaling, and brain processing. About 130 million photoreceptors in the retina absorb light and transmit visual signals to the brain.

The anatomy of the eye is such that light passes through the cornea, pupil, and lens, where it is refracted and focused onto the retina at the back of the eyeball. The retina is the location of photoreceptor cells, which contain proteins called opsins that detect light. Opsins are transmembrane proteins coupled to G proteins. This is similar to the taste receptors for sweet, bitter, and umami; the distinctive feature of opsins is that they hold a molecule of vitamin A that gets altered when exposed to light.

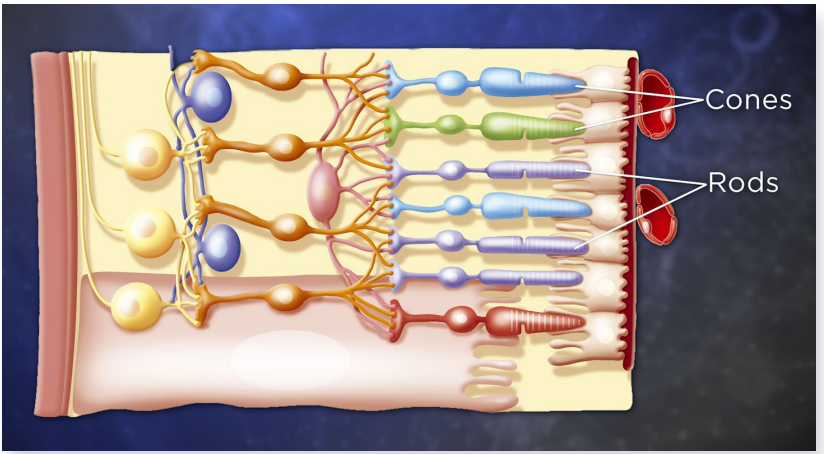
The bandwidth of the human retina is about 9 megabits per second. Standard DVDs have a bit rate that is almost the same, while a lot of videos posted on the web have a bit rate that is less than 1 megabit per second.

Thanks to opsins, photoreceptor cells detect light and pass signals on to other neurons, called ganglion cells. The axons of the ganglion cells are bundled into the optic nerve, which ultimately routes the information back to the visual cortex in the brain to process the signals and tell us what we are seeing.

There are 2 types of retinal photoreceptor cells involved in light detection: rods and cones. Rods are responsible mostly for detecting light under very low-light conditions. They provide little color information, because they are optimized to detect things in very dim light. But what rod cells lose in color detection, they gain in sensitivity; a rod cell can detect a single photon of light.

Color detection is the domain of the cone cells. Humans have 3 types of cone cells, each specialized to absorb wavelengths corresponding to red, green, or blue. Cone cells operate best in bright light.

Seeing color requires input from each of the red, green, and blue cone cells. Color blindness occurs if one of the 3 receptors is damaged or absent. Though damage to any of the sets of cones is possible, the most common type of color blindness is red-green color blindness, and it results from the lack of either a red or a green opsin.



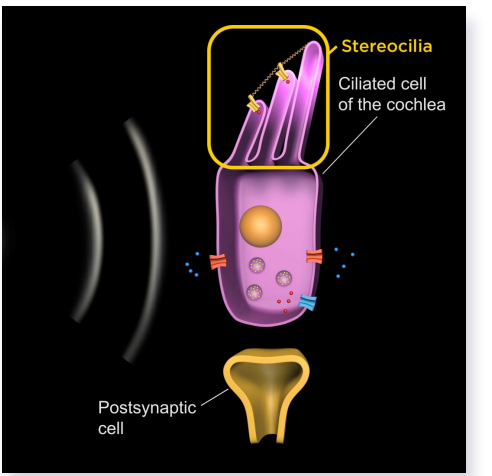
Genes encoding the opsins are on the X chromosome. Human males have only one X chromosome, while females have 2, so the loss of coding for an opsin on its single X is a bigger deal for men than for women. Consequently, 8% of males are red-green color blind, compared to only 0.5% of females.

Hearing

The sense of hearing is remarkable in its range. People with normal hearing can detect sounds of pitches, or frequencies, between 20 and 20,000 hertz. Normal, everyday speech has a mix of low- and high-frequency sounds that is generally in the 250 to 6000 hertz range. The length of time it takes our nervous system to respond to a sound signal is on the order of tens of microseconds.

Sound detection occurs when tiny hair cells called stereocilia move on the membrane of an ear structure called the cochlea. Each of the stereocilia is attached to an adjacent one by a filament that controls the opening of an ion channel.

When sound waves arrive at the cochlea, the cochlear fluid begins to move. The waves in the fluid move the hair cells on the membrane. And when this happens, the up-and-down motion of the basilar membrane sets up side-to-side movement of the fluid between the basilar membrane and the tectorial membrane. And this causes the stereocilia to tilt.



The tilting of stereocilia yanks the filament attached to the ion channel gate on the neighboring stereocilium and opens it, initiating the signaling process. When the brain receives these signals, it reports this as sounds.

The sensitivity of the mechanism is astounding: Movements as small as ½ the diameter of an atom can be detected.

Hearing loss is complex, but one thing that commonly happens is damage to the hair cells of the inner ear—whether due to age, infection, or exposure to loud noises. Damaged hair cells can't signal. Cells responding to high frequencies are located in the lower cochlea and are among the most easily damaged, thus explaining why high-frequency hearing loss is usually the first to occur.

A 2016 report shows that men are almost twice as likely as women to develop hearing loss.

Touch

The most mysterious sense is touch. Touch stimuli can be mechanical signals, where low-threshold mechanoreceptors sense contact with the skin. Low-threshold mechanoreceptors detect pressure, vibration, stretching of the skin, and movement of hair follicles in the skin. These receptors respond by triggering action potentials and communicate with sensory neurons to send information to the brain.

Thermoreceptors detect temperature signals that are lower or higher than your body's. Sometimes mechanical and temperature signals team up. For example, the "taste" of a meal is actually made by combining information from smell receptors, taste buds, and even texture or touch.

Although you might not expect it, the sense of touch is responsible for your ability to "taste" spicy foods.

One of the prime functions of the sense of touch is to help us distinguish innocuous stimuli from harmful ones. Pain receptors that detect dangerous stimuli and signal the spinal cord and brain are called nociceptors, related to the word *noxious*. These pain receptors respond to mechanical, thermal, or chemical stimuli that are outside the range that is innocuous and alert the brain to tissue damage.

Nociceptors for temperature are activated by dangerous heat or cold conditions with separate sensing cells for each. Their responses are accompanied by pain at both ends of the temperature spectrum. These are abnormal situations, and the ion channels involved signal the abnormality by opening and causing membranes to depolarize.

READINGS

DeSalle and Wynne, *Our Senses*.
Henshaw, *A Tour of the Senses*.
Rosenblum, *See What I'm Saying?*

QUESTIONS

1 The 5 taste receptors differ in the way their signals are initiated. Three use GPCRs, whereas the other 2 do not. Speculate why there is this difference.

2 Speculate on the effects of using a cyclic guanosine monophosphate (cGMP) phosphodiesterase inhibitor on vision.

3 The form of vitamin A used in retinas, called retinal, is the most easily oxidized form. How does this reconcile with what you learned previously about reactive oxygen species?

CLICK HERE TO SEE THE ANSWERS.

24 FROM BIOCHEMISTRY TO MOLECULAR BIOLOGY

You've now reached a point where you have a broad background in the major ideas in biochemistry. There is a whole world of examples where your biochemical knowledge can help you understand what is going on below the surface. And researchers are constantly uncovering new mechanisms and molecular events that impact our lives and enrich our views of the natural world. These are accessible to you now that you have a wide understanding of biochemistry.

Caffeine

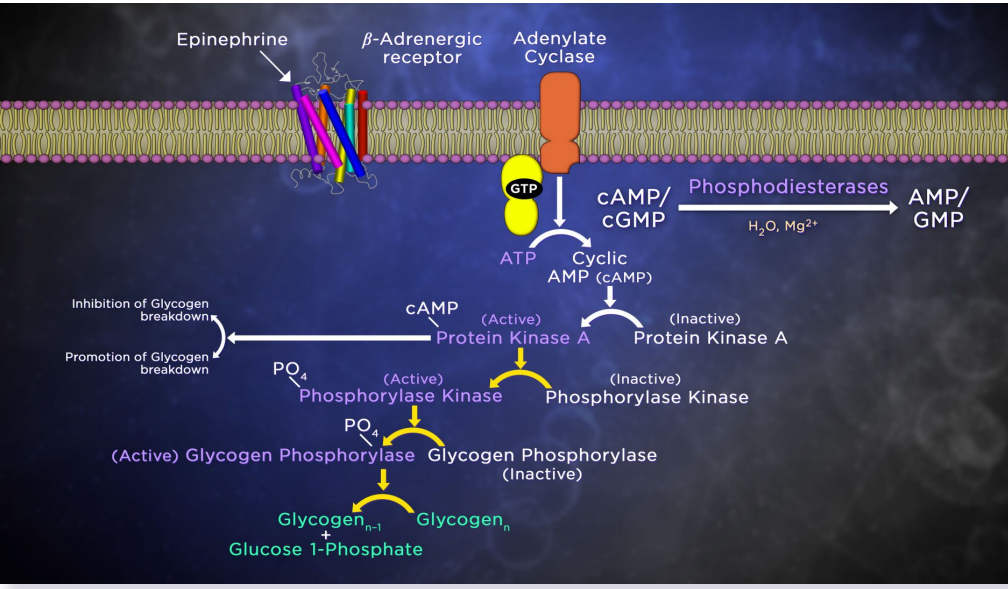
When you wake up in the morning, your body has already been preparing for the day. Levels of your fast-acting stress hormones epinephrine and norepinephrine are rising, your slower-acting cortisol is getting into the act, and your body is getting primed for action.

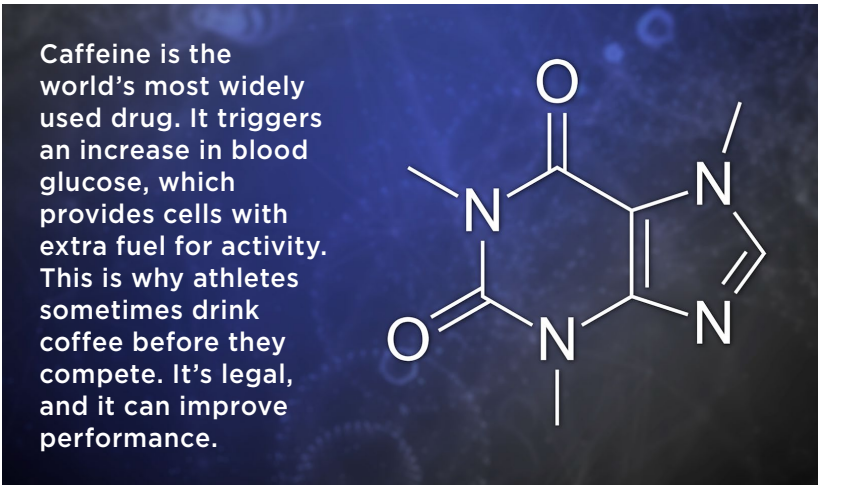
Your body seeks energy, having been asleep for a long time, during which you haven't eaten. So, you grab a cup of coffee, and your cells set in motion the pathway that will result in glycogen breakdown and release of glucose.

In the first step in the glycogen breakdown pathway, a protein in the signaling cascade activates an enzyme that makes cyclic adenosine monophosphate (AMP), known as cAMP. Second, cAMP activates an enzyme called protein kinase A (PKA) that starts a cascade of enzyme activations leading to the breakdown of glycogen. When PKA is active, glycogen gets broken down to release a form of glucose.

PKA and cAMP are the important players here. Normally, cAMP is broken down quickly by the enzyme phosphodiesterase and PKA becomes inactive. So, the activation of PKA is normally short-lived.

That's where caffeine comes in. It inhibits phosphodiesterase. When that happens, cAMP levels stay high and PKA stays stimulated and continues to activate enzymes that break down glycogen. That elevates blood glucose levels, providing cells with extra fuel for activity.





Fructose

Coffee is not the only thing that tinkers with sugar levels in your body. While the buzz from caffeine is fairly innocuous, the buzz from other foods and drinks—especially items containing high-fructose corn syrup—is not.

There are 2 issues to keep in mind: the overconsumption of sugar in general and the effects of high levels of fructose on the metabolism of sugars.

On the surface, it is surprising that fructose metabolism might be a problem. It has the same number of calories as glucose, and phosphorylated forms of them get interconverted in glycolysis and gluconeogenesis. But there is an alternate pathway for fructose metabolism.

The first difference between the metabolism of glucose and fructose is in the tissues that metabolize them. In general, glucose is used by the brain and muscles. Fructose, on the other hand, is metabolized almost entirely by the liver.

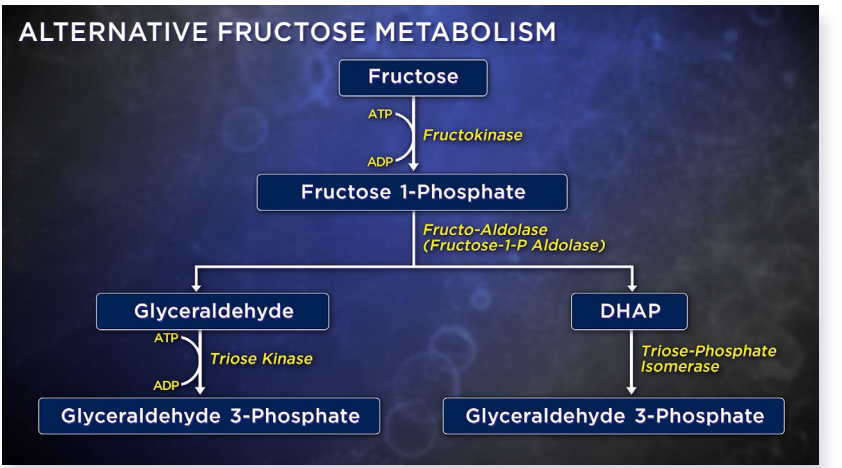
In glycolysis, glucose is made into glucose 6-phosphate, which is then converted to fructose 6-phosphate, which continues through the remaining steps of glycolysis. So, if glucose is converted to fructose 6-phosphate in glycolysis, how could fructose be a problem?

The answer is that fructose has an alternate entry pathway to glycolysis in the liver. It only involves 3 enzymes.

1 The first enzyme of this pathway is fructokinase. Like any kinase, it uses ATP to add phosphate to something. In this case, it's added to position 1 of fructose to create fructose 1-phosphate.

2 Fructose 1-phosphate acts as the substrate for the second enzyme, fructoaldolase, which acts like the aldolase enzyme of glycolysis, splitting the 6-carbon fructose 1-phosphate into 2 pieces: dihydroxyacetone phosphate (DHAP) and glyceraldehyde.

3 The third enzyme, called triose kinase, converts glyceraldehyde into glyceraldehyde 3-phosphate.



At this point in this alternate fructose metabolism pathway, the same intermediates have been made as are made in glycolysis after step 4: DHAP and glyceraldehyde 3-phosphate. Regardless of whether these molecules were made from glucose or fructose, they will, beyond this point, be broken down to pyruvate, which will feed the citric acid cycle.

With the alternate fructose metabolism pathway, fructose can be converted completely to pyruvate while bypassing the phosphofructokinase (PFK) reaction of glycolysis in step 3. It also bypasses the hexokinase reaction of step 1.

Hexokinase and PFK are 2 of the 3 regulated enzymes in the 10 steps of glycolysis. So, the alternate fructose pathway is bypassing 2 of the steps controlling glycolysis. What is the consequence of that?

Normally, when sugar is metabolized, glycolysis is slowed down when plenty of ATP and citrate are made, because these molecules inhibit PFK. Bypassing the PFK step, by using the alternate fructose pathway, allows pyruvate to keep on being made even when lots of ATP and citrate are present.

Why does that matter? Recall the metabolic fate of pyruvate. When oxygen is abundant, pyruvate gets oxidized to acetyl-CoA, which combines with oxaloacetate to make citrate in the first step of the citric acid cycle. However, when the citric acid cycle has too much citrate and ATP, pyruvate gets sent off to the cytoplasm, where it's broken down to acetyl-CoA to enter another pathway that makes fatty acids.

And because this is all happening in the liver, it would lead to accumulation of fat in the liver. Fat made there can be exported to adipose cells, leading to obesity. High levels of fat in the bloodstream also increase the likelihood of cardiovascular disease.

The use of high-fructose corn syrup as a sweetener increased in parallel with the obesity epidemic, but it's not clear if there is a causal link between them.

The alternate, unregulated pathway for fructose can lead to serious obesity-related health problems. Too much of any kind of sugar is not doing us any good, but high-fructose corn syrup has sugar in the easiest-to-metabolize form. Soft drinks and processed foods typically are sweetened with high-fructose corn syrup, so making a habit of these could especially increase total sugar intake. And while researchers continue to sort out all the nuances of these varying factors, there's no question about the benefits of restricting your intake of sugars of all types.

Alcohol

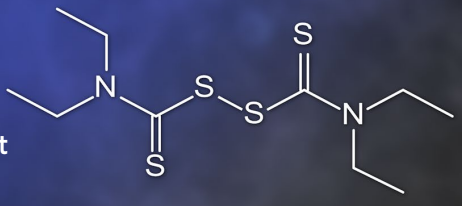
Alcohol is metabolized by alcohol dehydrogenase into acetaldehyde, which is toxic. So, your body summons up another dehydrogenase enzyme to convert acetaldehyde into acetate, which can give you a horrid headache. But unlike acetaldehyde, at least it won't give you palpitations, shortness of breath, blurred vision, and nausea.

But many East Asians have a variant form of alcohol dehydrogenase. The specific enzyme you have, like all the proteins you have, is always determined by information in your DNA. A gene is the region of DNA that encodes for a specific protein.

So, East Asians have a gene for an enzyme that is extremely good at converting alcohol to acetaldehyde. But the body accumulates large amounts of the toxic acetaldehyde quickly, which is when you really need your acetaldehyde dehydrogenase to help with disposing of the acetaldehyde.

But as many as ½ of all people of East Asian descent have a variant, and less effective, acetaldehyde dehydrogenase. If you have both variants, you make a lot of acetaldehyde and you can't convert it into something less innocuous—which means your body is going to be so miserable that you'll regret that alcoholic drink.

The drug Antabuse, prescribed to individuals trying to stop drinking alcohol, works by inhibiting acetaldehyde dehydrogenase. If you can't clear the acetaldehyde, it makes you feel awful if you do take a drink. Thus, you're less likely to drink.



Exercise

When you exercise, the harder you push yourself, the more heavily you breathe. Contrary to what you might think, when you're so out of breath that you can't speak, that is not aerobic exercise in a biochemical sense.

When your muscles are working hard, you are starting to go anaerobic because the muscles are using oxygen faster than the blood can deliver it. After a point, your blood will not be able to deliver the amount of oxygen you need. So, while it's pretty common to refer to challenging exercise as "aerobic," it is actually making you go anaerobic. Your heavy breathing is your body trying to catch up on oxygen supplies so that you can replenish ATP stores. Of course, your heavy breathing will catch up when you slow down or stop.

If you want to lose weight, will you lose more easily by going aerobic or by going anaerobic?

- When you are aerobic, you are getting sufficient oxygen to keep electron transport going.

- When you are anaerobic, you don't have sufficient oxygen, and you begin lactic fermentation.

Going anaerobic from exercise will trim your waistline more readily than aerobic exercise. When you consider that exercising hard is what makes your muscles go anaerobic, what's happening is you're burning more glucose when you're working the hardest.

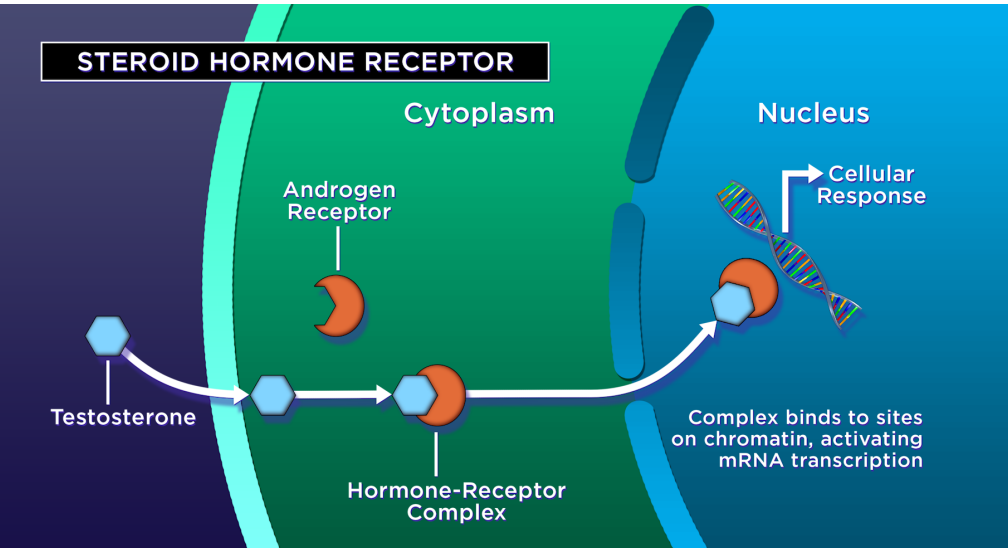
The field of molecular biology is centered on DNA, RNA, and proteins. But what we are finding increasingly is that knowledge from molecular biological research is informing us about the underpinnings of classical biochemistry.

Genetic Inheritance

The topic of how babies develop and how hormone receptors can affect genetic inheritance takes us from signaling molecules to chromosomes—the first intersection of traditional biochemistry and molecular biology.

In humans, individuals with 2 X chromosomes will be female, while those with one X and one Y chromosome develop as male. The default setting for all human embryos is female; that is, unless specific molecular events occur during development in the womb, the embryo will continue to develop as a female. Those events are controlled by a gene on the Y chromosome called *SRY* (sex-determining region Y). The protein encoded by the *SRY* gene is necessary to turn on genes that signal the embryo to develop testes. As that happens, the embryo begins producing testosterone, which then sets it on the path to fully developing as a male.

But, as you've learned in the biochemistry of signaling, cells can only respond to signals if they have receptors for them. Steroid hormone receptors, when bound by the appropriate hormone, control the synthesis of steroid-sensitive proteins in a cell. So, for the testosterone to communicate its signal, it must bind to a steroid hormone receptor called the androgen receptor, which controls the synthesis of proteins that are necessary for making male genitalia.



That's how testosterone typically binds in an XY embryo. But if the embryo has a mutation in the gene encoding the receptor, its cells have a receptor that cannot respond to the testosterone. Embryos that have a completely nonfunctional receptor like this develop as if there were no testosterone present—so, externally, when the child is born, the baby looks like a girl. But the baby still has embryonic testes within its body and also fails to develop ovaries, a uterus, or Fallopian tubes because XY embryos also produce a hormone that suppresses the development of some female reproductive structures.

Because this is not immediately evident, such children are raised as girls and at puberty appear to develop normally. But because they lack ovaries, they do not have periods; often, that's when they may first be discovered to be chromosomally XY.

Estimates indicate that androgen insensitivity is seen in about 1.7% of the population, making it about as common as red hair.

Aging

The bridging of classical biochemistry and molecular biology may help us understand—and perhaps one day slow—the process of aging.

Scientists have found that the molecule nicotinamide riboside, a form of vitamin B_3, may be a precursor molecule that can help slow the damaging effects of aging. It would help by increasing levels of NAD^+ in the body.

NAD^+ is a nicotinamide, a form of the niacin vitamin but with adenine dinucleotide attached. It plays important roles in the oxidative processes that help us get energy from food.

NAD^+ is also a required coenzyme for the activity of antiaging enzymes called sirtuins. We don't know completely why NAD^+ levels matter in this case, but there is one tantalizing clue: Sirtuin enzymes affect the chemical modification of proteins that bind to DNA in chromosomes. The sirtuins remove acetyl groups, and deacetylation has significant effects on when or whether many proteins that are coded by DNA actually get made or not.

As we get older, we are less efficient at making NAD^+, and low levels of NAD^+ lead to inactive sirtuins—both in the mitochondria and in the nucleus, where the chromosomes are located. These changes, in turn, result in the loss of healthy stem cells, so our tissues are unable to replace old and damaged cells.

Increasing NAD^+ levels with nicotinamide riboside seems to reverse these changes. However, don't dose yourself up on NAD^+, niacin, or even nicotinamide riboside, because there may be significant downsides to taking too much.

READING

van Holde and Zlatanova, *The Evolution of Molecular Biology*.

QUESTIONS

1 Imagine that a person has a defective aldolase enzyme from glycolysis. Prescribe a diet for the person that might help mitigate the effects of this loss based on information from this lecture.

2 Imagine the mouse described in the lecture that lacked fructokinase. Would you expect it to accumulate fructose as a result of lacking this enzyme? Why or why not?

3 You are part of a mountain rescue team and encounter a person in the bitter cold who quickly needs energy. Assuming glucose and fructose get absorbed equally fast, predict the effects of offering this person pure fructose, high-fructose corn syrup, table sugar (sucrose), or pure glucose. Assume that no pathway has a limit on the quantity it can handle.

CLICK HERE TO SEE THE ANSWERS.

25

DNA AND RNA: INFORMATION IN STRUCTURE

Molecular biology is the part of biochemistry that explains how living things reproduce. The fertilized egg that gives rise to each individual contains DNA, which is copied each time a cell divides, with a copy of the information passed on to every one of the trillions of cells in a human baby. In each generation, the DNA of 2 individuals mingles to create unique combinations that are carried by their children.

DNA

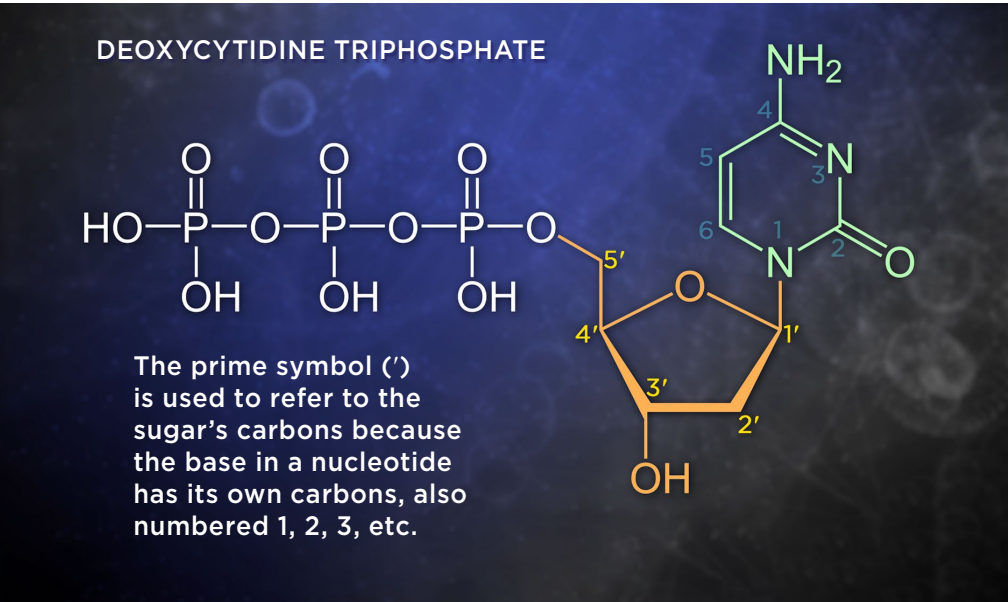
Our knowledge of how deoxyribonucleic acid (DNA) stores information came from its structure. The determination of the structure of DNA in 1953 by James Watson and Francis Crick, using data from Rosalind Franklin, was the most important biochemistry advance of the 20th century.

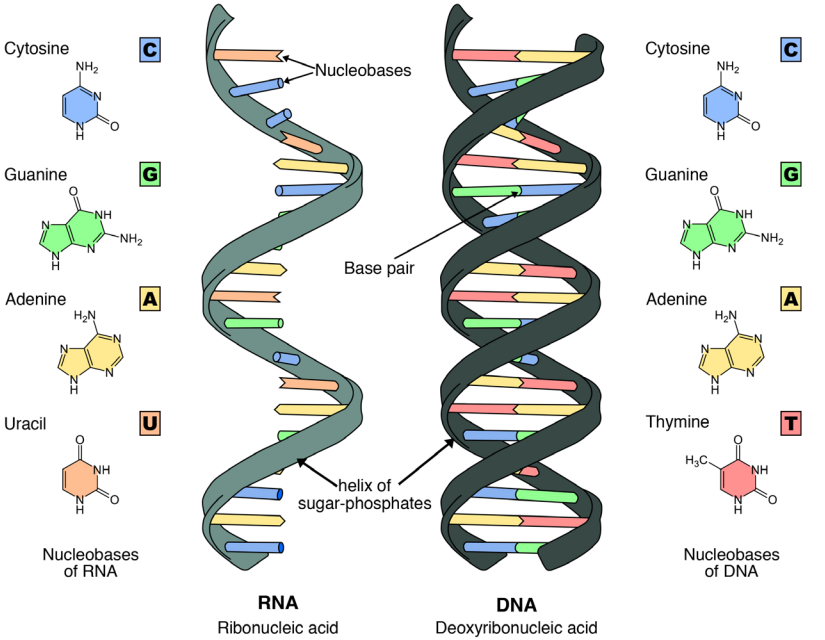
DNA is a single molecule that has all of the information for making all of the proteins in each of the body's trillions of cells.

The structure of a DNA molecule looks like a rope ladder twisted on itself. Two sides of the ladder serve as backbones of individual DNA strands, each of which is a long polymer. The repeating subunits comprising the polymer are called nucleotides, each of which is made up of 3 components: a 5-carbon sugar, a nitrogen-containing part called a base, and one or more phosphates. Without phosphate, the resulting molecule is called a nucleoside.

- Nucleotides that DNA is built from are called deoxyribonucleoside triphosphates (dNTPs). Deoxyribose, the sugar in DNA nucleotides, has 5 carbons, designated 1′, 2′, 3′, 4′, and 5′. The sugar in DNA would be ribose, except that its 2′ carbon lacks an oxygen and has only a hydrogen atom attached. So, this sugar is called deoxyribose or 2′ deoxyribose. Each nucleotide for building DNA brings energy to the job in the form of 3 phosphates attached to the carbon at the 5′ position.

- Also attached to the deoxyribose sugar is a structure called a base. There are 4 different bases in DNA nucleotides: adenine (A), guanine (G), cytosine (C), and thymine (T). The related molecule called RNA has the





same bases except that uracil (U) substitutes for thymine. Two of the bases, A and G, are structurally similar and are called purines. The other 3—C, T, and U—are called pyrimidines.

- Nucleotides get joined together in DNA strands by linking a phosphate to the sugar of an adjacent nucleotide. When the cell builds DNA from nucleotides, the hydroxyl on the sugar's 3′ position of one nucleotide is linked to the phosphate closest to the 5′ carbon on the next. During joining, the other 2 phosphates from the triphosphate starter material get chopped off, releasing energy to form the bond.

Each nucleotide in the finished DNA chain ends up with a single phosphate linking 2 sugars: one from its own nucleotide and the other from the next nucleotide.

DNA can contain thousands or even millions of nucleotides strung together in long, threadlike structures. Its backbone is made of sugars and phosphates, with the bases roughly perpendicular to the backbone. One end of a strand has phosphates at the 5′ carbon of the first nucleotide and is known as the 5′ end. The other end of the strand has a free hydroxyl group at the 3′ carbon, known as the 3′ end.

A double helix contains 2 such DNA strands oriented in opposite, or antiparallel, directions, with the 5′ end of one across from the 3′ end of the other. Bases in the 2 strands are oriented such that As and Ts always face each other and Gs and Cs are always across from each other.

The pairs of bases, then, form what look like the rungs of a ladder inside the DNA helix, with the sugar-phosphate backbones forming the sides. The ladder twists on itself, resembling a spiral staircase.

Bases pair as they do—and never switch partners—because the hydrogen bonds between the bases only properly align when G is adjacent to C and when A is adjacent to T. The As and Ts share 2 hydrogen bonds, while Gs and Cs have 3 hydrogen bonds between them.

Because A pairs with T and G pairs with C, it means that if the order, or sequence, of bases in one DNA strand is known, then the sequence of bases on the partner strand can be determined. It is in the order, or sequence, of these bases that information is encoded.

Instructions for Making Proteins

One major function of the information in DNA is to provide instructions for making proteins, which are built from amino acids. The information for amino acids arises directly from the sequence of the bases in DNA, so the sequence of bases in DNA stores instructions for making proteins.

Information that is securely stored in DNA is copied to RNA whenever the information needs to be accessed and used. RNA's nucleotide composition is the same as DNA's, except that U replaces T. So, whenever the genetic code is translated to protein, it's in the bases of RNA, which are A, U, C, and G.

There are only 4 bases in DNA or RNA, while proteins are built using 20 different amino acids. Using one base or 2 bases to represent one amino acid would not work. With 3 bases, though, there are enough possibilities. Cells use 3 bases, called a codon, to specify one amino acid in the genetic code.

With 3 bases, there are $4 \times 4 \times 4$ possibilities, so that's 64 different codons. But only 20 amino acids, not 64, are needed to make proteins. So, there are some things to consider.

First, in addition to coding for amino acids, the code has 3 punctuation marks, called stop codons, which function like periods to signal the end of a protein-coding sentence.

Second, there is more than one codon that specifies most amino acids. This is called redundancy. Generally, the more common an amino acid is in proteins, the more codons there are for it. Almost all the amino acids are specified by at least 2 codons. The codon that specifies methionine is also called the start codon because it is almost always the first codon in protein-coding sequences.

Not every region of DNA codes for proteins. But every protein-coding region in a DNA molecule starts with a start codon, next has a series of codons specifying amino acids, and then ends with one of the 3 stop codons.

The rules specifying which codons represent which amino acid are referred to as the genetic code. The genetic code is only stored in DNA, but it's read from RNA, where the base U replaces T.

The linear sequence of bases is read in groups of 3 to get instructions for making proteins. Any region of DNA that contains information needed to make a protein is called a gene. Every protein you've ever learned about—from hemoglobin to the enzymes of metabolic pathways—is encoded in genes.

In addition to storing information for making proteins, DNA also contains information that controls when and how much of each protein is made.

Storing and Copying Information

Double-stranded DNA molecules have 2 features that make copying it easy. First, the base sequences of the 2 strands are complementary: The As are always connected by 2 hydrogen bonds with Ts, while Gs are always connected by 3 hydrogen bonds with Cs. This is called base pairing.

When the 2 strands of a DNA molecule are separated, each strand contains information to act as a template for building a new strand. The building of new strands follows the base-pairing rules: Adenine always pairs with thymine, and vice versa. And the same goes for guanine and cytosine.

Hydrogen bonds between the bases are what hold the 2 strands of the double helix together. The stabilizing force of hydrogen bonds, summed together over thousands of base pairs, is strong, allowing the double helix to be stable over long stretches. Over short stretches, however, there are only a handful of hydrogen bonds, and those can be broken to separate the strands. This is important in the synthesis of both DNA and RNA.

So, thanks to the local weakness of hydrogen bonds, DNA's strands can be separated relatively easily for replication. This ease of strand separation is important for the enzymes that catalyze the copying of DNA.

DNA is also a remarkably efficient way to store information. Eukaryotes, including humans, store their DNA by employing a highly ordered form of spooling. Human DNA is divided into 46 pieces, each of which constitutes a single chromosome. The DNA of each chromosome, which still has millions of base pairs in length, is wound around spools made of proteins. The combination of DNA and proteins is called chromatin, and this combination is what makes up our chromosomes.

Since 2011, researchers have been working on using DNA to store the enormous flood of data that the world is churning out.

In 2013, a group of researchers at the European Bioinformatics Institute in England stored all of Shakespeare's sonnets in DNA and retrieved the information without errors.

RNA

The information for making a protein is encoded in the base sequence of DNA, but DNA does not serve as the direct template for building the protein. Instead, the information in the gene is copied into the closely related molecule, ribonucleic acid (RNA), which directs the synthesis of proteins.

Like DNA, RNA has a backbone of phosphates linking sugars, with each sugar attached to a base. In the case of RNA, the sugar is ribose instead of deoxyribose.

Three of the bases in RNA are the same as in DNA: guanine, cytosine, and adenine. In RNA, though, thymine is replaced by uracil. Like thymine, uracil can base-pair with adenine, forming 2 hydrogen bonds.

Unlike DNAs, RNAs are made as single strands, which gives them some interesting properties. They can, for example, fold on themselves, with base pairs forming between different sections of the same RNA that are complementary. This is similar to the tertiary folding of proteins. Such folding allows RNA molecules to take on complex 3-D shapes, which in turn confer on RNA the ability to carry out functions beyond encoding information.

There are 3 major types of RNA in cells:

1. Messenger RNA (mRNA) is responsible for carrying codon information for making proteins from the DNA to the complexes in the cell known as ribosomes, which are the site of protein synthesis.

2. Transfer RNA (tRNA) has 2 roles: carrying amino acids to the ribosomes and providing a decoder of the codons called anticodons. These functions are at opposite ends of the tRNA: One end has a covalent bond to the amino acid, and the other end has the decoder.

3. Ribosomal RNA (rRNA) is a component of the ribosomes that provides a scaffolding where ribosomal proteins bind and form the overall structures.

It is in ribosomes that amino acids are joined together with peptide bonds to make proteins. What catalyzes the formation of these bonds? In a surprise to biochemists, the formation of peptide bonds is not carried out by any of the proteins in the ribosome, but by a ribosomal RNA.

This shows that proteins are not the only molecules that can function as catalysts. And how is rRNA able to be a catalyst? RNA's ability to fold into a complex 3-D structure gives it some of the same characteristics as a folded enzyme. RNAs that function like an enzyme are called ribozymes.

So, RNA molecules have functions that include catalysis, carrying information, and synthesizing proteins. This diversity of RNA functions helps answer a question that had puzzled scientists for a long time: If enzymes are required to make DNA and the information for making the enzymes is in DNA, which came first, the enzymes or the DNA?

The RNA world hypothesis suggests that DNA was not the first genetic material. Instead, RNA could have originally performed both functions, replicating itself and catalyzing the formation of peptide bonds to MAKE proteins. In this view, RNA was the molecule from which all of these critical biochemical processes arose.

Using selection techniques that allow for what is referred to as "evolution in a test tube," catalytic RNAs capable of copying nucleic acid sequences have been created. Given the short laboratory time involved in creating these molecules, evolutionary time would clearly be sufficient for creating the variety of RNA catalysts that primitive cells might have required.

Over time, some of the functions were taken over by molecules that were better suited for particular roles. DNA, for example, is chemically more stable than RNA, and organisms that used DNA as their genetic material likely survived and reproduced more successfully as a consequence.

READINGS

Calladine and Drew, *Understanding DNA*.
Yarus, *Life from an RNA World*.

QUESTIONS

1 Given the complicated steps necessary to convert DNA to RNA to protein and the fact that RNAs can catalyze reactions, speculate on why cells ever evolved proteins at all. Why not simply use catalytic RNAs?

2 Cells copy DNA to make RNA to make protein. RNA is the messenger. Why use RNA at all? Why not make proteins directly from the DNA? Speculate on reasons for this.

CLICK HERE TO SEE THE ANSWERS.

26

DNA REPLICATION IN BACTERIA; PCR IN THE LAB

The DNA in our cells constitutes our genome, the set of instructions for making a human being. Almost every cell in your body has this exact DNA, copied faithfully from the original instructions in the fertilized egg you arose from. And each time a cell divides, its DNA gets copied, so progeny cells carry the same blueprint. This process is called DNA replication, and through it, genetic information is transmitted, ensuring that every new cell has a complete copy.

DNA Replication

In their 1953 paper, James Watson and Francis Crick—benefitting from Erwin Chargaff's observations and the unacknowledged work of Rosalind Franklin—described a model for DNA structure. They proposed that the DNA molecule had 2 intertwined strands with As across from Ts and Gs across from Cs. They noted that their structure suggested a means of copying the molecule.

We know today that their beautiful structure was correct. A parental DNA duplex gives rise to 2 daughter double helices, also called duplexes, each with one parental strand and one strand copied from it.

How is it possible to make a new copy of a DNA molecule? The original DNA is needed to serve as a template across from which the new strands can be assembled.

New DNA molecules are made by joining together free DNA nucleotides. When free nucleotides are joined to make DNA, the hydroxyl on the 3′ position of one nucleotide is linked to the phosphate closest to the 5′ carbon on the next. The result of such joining is that the remaining 2 phosphates are cleaved off, leaving the DNA chain with a single phosphate linking 2 sugars, one from each nucleotide.

Individual nucleotides start off as triphosphates; the chopping off of the 2 phosphates provides energy for the reaction to bond the nucleotides together. The covalent bond made in this way is called a phosphodiester bond. The formation of these bonds is catalyzed by enzymes called DNA polymerases.

Because DNA polymerase needs help getting started, then if we want to make a complementary strand to a template, we need a short little stretch, called a primer, to provide a starting point for DNA polymerase to extend from. This primer needs to be base-paired with the template. But how can we make a primer when DNA polymerase won't start on its own?

It turns out that cells have another way to get started; they use a short primer of RNA instead. The enzymes that make RNA are perfectly happy to start making a primer by copying DNA into RNA as long as they have a supply of RNA nucleotides.

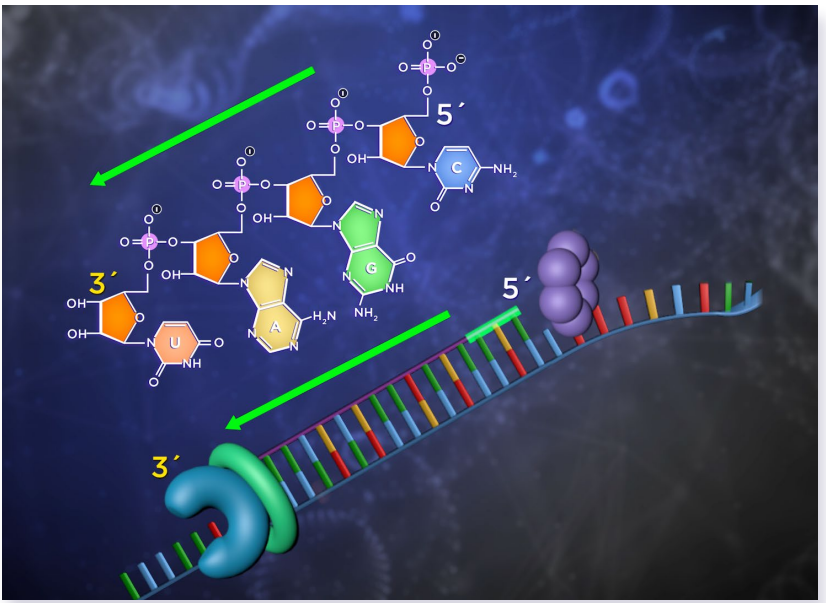
The enzyme doing this is called primase, and it makes an RNA primer of about 10 or 15 nucleotides long that is base-paired to the template. The primer provides a free end with a 3′ OH to which DNA polymerase adds more nucleotides for extending the strand of DNA.

RNA is only needed to start the process of replication and doesn't remain in the final DNA product.

At this point, we get an intermediate that's part RNA and part DNA. This is not a problem, though, because the RNA nucleotides can be removed later and replaced with DNA nucleotides.

The DNA polymerase can only add new nucleotides at a 3′ hydroxyl end; it can't extend the primer from the 5′ end. This has consequences for how all DNA replication works. Because synthesis of a strand has an unchanging 5′ end and new nucleotides are added by attachment to 3′ OH groups, DNA replication is said to proceed in the 5′ to 3′ direction.

So, we finally have the components for the synthesis of new DNA: the template, the dNTPs, the primer, and a DNA polymerase that assembles the new strands. In cells, numerous additional proteins assist in the replication process.



Replication in Bacteria

Bacteria like *E. coli* have a relatively small DNA genome, with 5 million base pairs wrapped in just one chromosome. And unlike our chromosomes, which are linear, *E. coli*'s single chromosome is circular.

DNA replication in all organisms is carried out by a large number of proteins that act together as a complex protein machine. The proteins of this cellular nanofactory work together in a coordinated fashion to replicate DNA.

DNA replication starts at sites in genomes called origins of replication. *E. coli* has a single origin of replication, called oriC, on its circular chromosome. OriC is the target for binding by special origin recognition proteins. When the origin recognition proteins bind oriC, a short section of the DNA double helix is forced open, allowing additional proteins to get into the act to begin replication.

Once a small region of the DNA duplex is opened up at each origin of replication, the 2 strands of DNA must be separated to allow the new strands to be built across from them. The hydrogen bonds in the DNA duplex get broken and the 2 strands unwind, thanks to the action of an enzyme called helicase.

Because there are about 10.5 base pairs per turn of DNA, it means that helicase unwinds the helix at the rate of approximately 100 turns per second, or 6000 turns per minute!

In *E. coli*, replication proceeds at the mind-blowing rate of 1000 base pairs per second. To support replication, helicase must unwind DNA at that same rate. This creates a problem: When strands are untwisted by helicase, the DNA ahead of it gets more twisted at the same rate. If nothing is done to relieve tension from that twisting, kinks and knots will form and stop DNA replication.

Enzymes called topoisomerases act to relieve this torsional stress. They do so by cutting the template DNA ahead of the replication fork and allowing the strands to swivel around each other. This relieves the tension, and then the ends are rejoined. In *E. coli*, the topoisomerase that relieves the twisting stress is called gyrase, which moves along the DNA, staying ahead of the replication fork and releasing the stress on the DNA as needed.

Once the 2 strands have been separated for replication, they need to be kept separated or they will come back together. Single-strand binding proteins (SSBs) ensure that unwound regions of the template DNA remain as single strands and available for copying. The separated strands of the original DNA are bound by many SSBs; wherever binding occurs, that portion of a strand is blocked from forming a duplex with its sister strand.

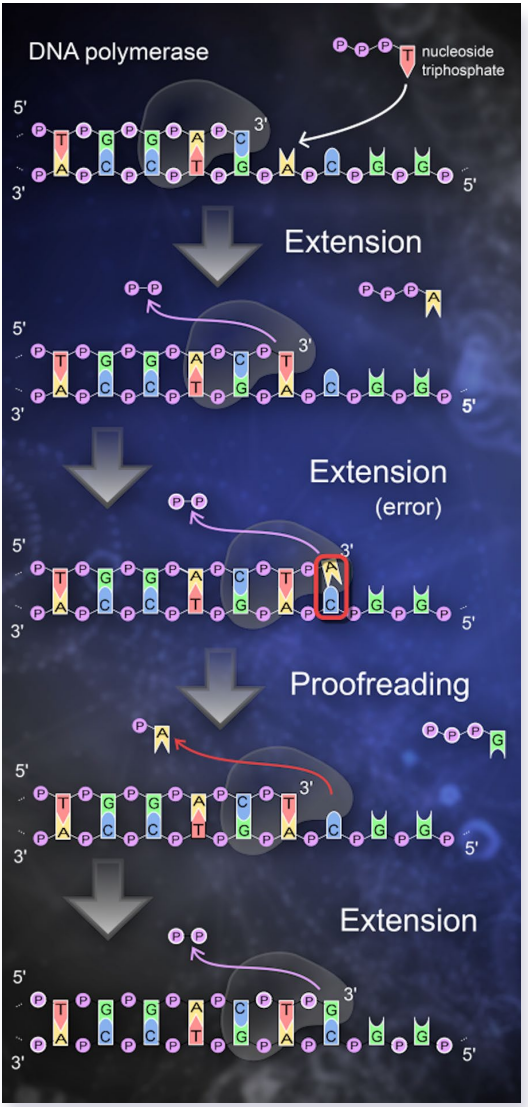
The synthesis of new DNA proceeds in both directions from the replication bubble opened up by helicase. And because new synthesis is bidirectional—one direction for each strand—each replication bubble has 2 replication forks, one that moves in each of the 2 opposite directions around the chromosome.

As a result, both strands are getting replicated at each replication fork, and the bubble grows larger with time. Helicase can keep separating the 2 strands of the template DNA as long as gyrase is relieving the torsional stress ahead of it.

Because the *E. coli* chromosome is circular, the 2 replication forks will ultimately meet at a specific sequence on the chromosome, called a termination site, where this all comes to an end and the job is done.

So, the concerted action of a large number of enzymes opens up the double helix and creates copies of each strand, resulting in the production of 2 identical copies of the original DNA.

In *E. coli,* the entire process of copying 5 million base pairs of DNA takes less than 20 minutes. But what is more impressive than the speed of DNA replication is its accuracy. In DNA replication, the error rate is roughly one wrong nucleotide put in for each 10 million that are strung together.



The error rate is so low because most DNA polymerases correct their work while they are zipping along.

As a DNA polymerase adds a new nucleotide to a DNA strand, it double-checks to see if it's the right one. If it isn't, it stops, backs up, removes the nucleotide, and then inserts the correct one before moving on. That's possible because in addition to being able to catalyze the addition of nucleotides, the polymerase has the ability to chop off a nucleotide at the end of a DNA chain. This process is called proofreading.

PCR in the Lab

Human ingenuity has harnessed the power and speed of DNA replication in a way that has revolutionized how genes are studied. Its inventor, Kary Mullis, won a Nobel Prize for it.

Called polymerase chain reaction (PCR), this procedure makes millions of copies of any DNA that we need in a few hours by replicating the desired sequence over and over in a test tube.

Replication doubles the amount of DNA each time. This is enormously useful for increasing the amount of DNA to work with. Samples like a tiny droplet of blood at a crime scene would not normally provide enough DNA for analysis, but with PCR, that miniscule amount can be scaled up easily.

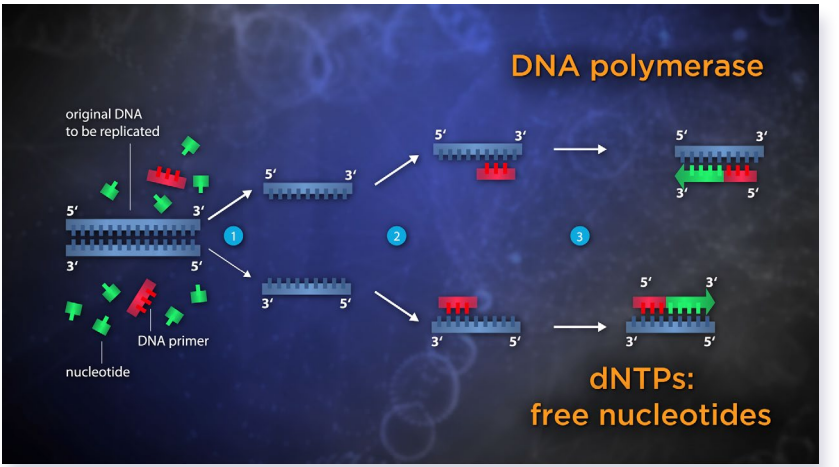
PCR uses a simplified form of DNA replication. For example, there are no replication forks, which is an improvement for lab purposes.

For PCR, we need a template to copy; in this case, that's the miniscule amount of DNA in a sample. We need nucleotides to build the DNA copies and a DNA polymerase to do the copying. And because DNA polymerase needs a primer, we have to supply primers that will base-pair close to the part of the DNA we want to copy.

Here, we have an advantage over the cell: Cells need to use RNA primers because there is no other way to start making a new strand. We, on the other hand, know how to make DNA primers, using organic chemistry. The process is automated and performed by machines.

Because the primers in PCR are DNA, we don't have to worry about getting rid of RNA primers in the process. We synthesize primers before starting the process. They are designed to be complementary to the start and end of the section we want to make copies of.

Next, we take the original DNA duplex apart and allow the primers to base-pair to the target regions in the DNA. The region in the middle between the primers is what we're aiming to amplify—by extending the primers from their 3′ ends, using the dNTPs and DNA polymerase.



To make things even easier, in PCR, we don't replicate all of the DNA, just the specific regions we want to focus on. That's easy, thanks to the primers. To identify an individual from a drop of blood, we choose only the regions of human DNA that vary between people, not the vast majority of the genome that is shared by everyone.

> The FBI uses a panel of several segments of the genome to amplify by PCR to create DNA fingerprints that can be used to identify individuals.

We mix the template with the dNTPs, primers, and the DNA polymerase and start. Although cells use many additional proteins to replicate the DNA, we don't need them.

The PCR process starts by separating strands with heat instead of enzymes. Then, to bind the primers to the sequences they are complementary to, the temperature is lowered to allow base-pairing. After the primers base-pair with the template, the primers are extended on both strands with DNA polymerase.

After one cycle, we have twice as many DNA molecules as we started with, and every round doubles the amount of DNA in the sample. Machines called thermocyclers do the work of changing the temperatures and can be programmed to do as many cycles as desired. When completed, we have millions of copies of the DNA we want, ready for analysis.

PCR can be used to identify individuals by their unique DNA profiles. Questions of paternity and genealogy can be settled. Patient samples can be tested for a virus or a genetic defect, and problems can sometimes be identified long before symptoms appear. Researchers can also find out about ancient human groups and their interactions, what they ate, whether they made beer or wine (or both), what they died of, and where they got the beads they were buried with.

READINGS

Mukherjee, *The Gene*.
Slack, *Genes*.

QUESTIONS

1. Predict the effect of a beta-clamp–inhibiting drug on the growth rate of *E. coli*.

2. Imagine that you give a cell a DNA ligase inhibitor during DNA replication. Predict what will happen during the next round of DNA replication with respect to helicase.

3. Proofreading is an important function for cellular DNA polymerases, but some viruses, including HIV, code for their own polymerases that lack proofreading. Speculate on that observation with respect to virus evolution.

CLICK HERE TO SEE THE ANSWERS.

27

CHROMOSOME REPLICATION, TELOMERES, AGING

Biochemical reactions, metabolic pathways, and processes have a great deal in common across all living cells. DNA replication in the cells of plants and animals is, in terms of the overall mechanism, quite similar to bacterial DNA replication. But DNA replication in plants and animals merits separate consideration. Plant and animal cells are more complex than bacteria, and there are distinct and significant features of their DNAs that pose challenges to DNA replication.

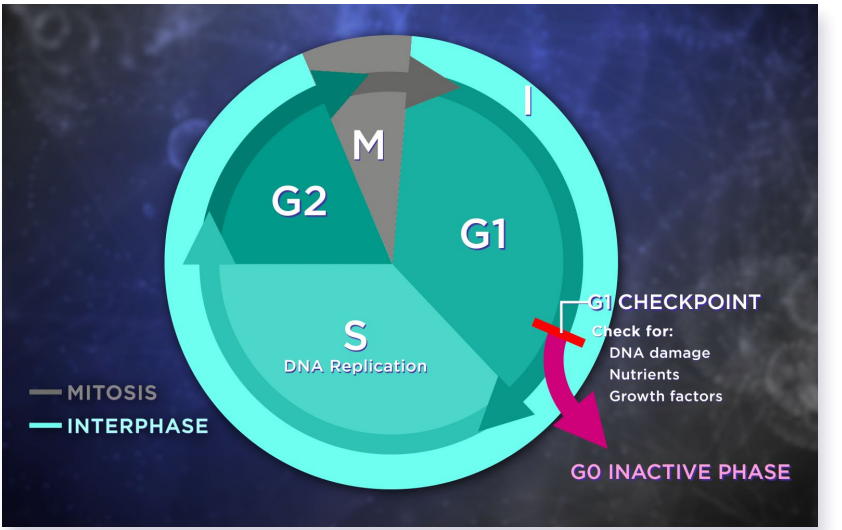
The Eukaryotic Cell Cycle

In eukaryotic cells, replication, or synthesis of new DNA, is intimately linked to the cell cycle, which serves to prepare the cell for division. The cell cycle is a series of stages that is stringently regulated to ensure that cell division occurs only when appropriate.

The cell cycle can be represented as a circle. The M phase of the circle indicates that a cell has undergone a brief period of cell division also known as mitosis, followed by a long I phase called interphase. Interphase is divided into 3 stages, during which cells take in nutrients and grow, replicate their DNA, and prepare for division.

During the first phase, called growth phase 1 (G1), proteins are synthesized and cellular components are assembled. Toward the end of G1, cells decide whether to enter an inactive, quiescent phase called G0 or proceed to synthesis in the S phase of DNA replication. Cells that proceed toward the S phase must prepare to replicate their DNAs so they can divide.

Once in the S phase, DNA is replicated, setting the stage for the next phase, G2, in which the replicated DNA is checked to ensure that it has been correctly and completely copied and the cell makes proteins that will be necessary for cell division. When everything checks out, the cell enters M phase of mitosis, in which the cell divides, after which the cycle begins again.



Replicating damaged DNA would pass mutations on to daughter cells, and this could contribute to the development of cancer.

The cell has various mechanisms, called checkpoints, that allow advancing from one phase to the next only after establishing that everything is in order.

But DNA replication happens only if the DNA passes damage inspection. If damage is detected, then a series of events arrests the cell cycle. First, the levels and activity of a protein, p53, increase, which in turn dials up the production of a protein called p21, which prevents progression into S-phase replication until the damage is addressed by cellular repair systems.

If the DNA damage is too extensive to fix, p53 will instead trigger cell death to prevent the proliferation of a cell with badly damaged DNA. However, cells that do not have functional p53 may continue to replicate their compromised DNA and proliferate.

More than 1/2 of all cancers have mutations that inactivate the *p53* gene.

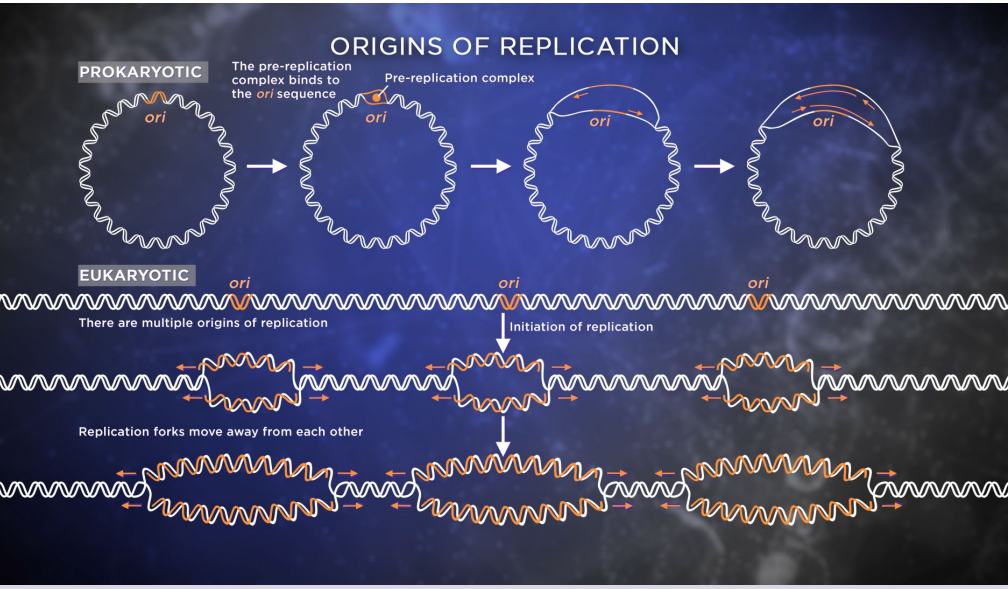
If the DNA passes inspection at the G1-S checkpoint, cells move into the S phase, where DNA replication occurs. As in bacterial cells, replication begins at sequences called origins of replication, which are sequences in the DNA that are bound by specific proteins that pry the 2 strands of the template apart.

One consideration for DNA replication in eukaryotes, compared to prokaryotes, is the vastly different sizes of genomes to be replicated.

Bacterial chromosomes have a single origin of replication, but this is not sufficient for larger eukaryotic genomes. A human cell must replicate 6 billion base pairs of DNA to divide. If the rate of replication is one nucleotide per second, it would take approximately 190 years to complete this task using only a single origin of replication.

But human DNA is divided into 46 chromosomes, so that helps—a little. However, even if we had one origin for each of those 46 chromosomes, it would still take 4 years to finish replicating the genome to allow the cell to divide one time.

So, how do our cells get all their DNA copied fast enough? Our DNA has 30,000 to 50,000 origins of replication! And each of these opens up a replication bubble, from which replication is initiated in both directions.



Replication forks in eukaryotic cells are almost identical to those in prokaryotic cells. There are many more origins in eukaryotes, but replication is bidirectional at every replication bubble, just as in bacteria. And eukaryotes have many of the same enzyme activities for replicating DNA. Eukaryotes also have DNA polymerases. But whereas *E. coli* and other bacteria have 5, eukaryotes have at least 14 of them, and they divide labor among themselves a bit differently from the polymerases in *E. coli*.

DNA replication stops when 2 replication forks meet each other or when the end of a chromosome is reached.

Unlike prokaryotic DNA, eukaryotic DNA is packaged together with proteins to make nucleosomes, which are tightly compressed to make chromosomes. These packing proteins get in the way of DNA polymerases that replicate DNA or RNA polymerases that make RNA from it. The resulting nucleosomes are building blocks of chromosomes, which contain DNA that is packed tightly, just like 500 feet of kite string can be packaged in a small container on its own spool.

When cells replicate their DNA, it must be unwound from the protein spools to allow access to the polymerases and other replication proteins. And as the original DNA is replicated and 2 new DNAs emerge, each of them must be repackaged into nucleosomes.

Telomeres

The chromosomes in eukaryotic cells cannot be replicated all the way to their ends. Unlike the circular bacterial chromosomes, human chromosomes are linear, and this presents a challenge.

When you get to the very end of a linear chromosome, the RNA primer can be built as usual and eventually removed, but this leaves a short section of the template unreplicated. The overhanging end of the template cannot be copied, because we've run out of room to place an RNA primer from which DNA polymerase can extend. So, the new DNA is shorter than the original template.

When the new DNA is copied, in the next round of replication, the same thing will happen again. And the DNA made in that round of replication will be even shorter. As a result of this end-replication problem, the loss of nucleotides at the ends of each chromosome with each replication cycle leads to physical shortening of the chromosomes.

The greater the number of cell divisions—and, thus, the more times DNA is replicated—the shorter the chromosomes get.

But if DNA is the information carrier of the cell and all of it doesn't get copied, aren't we losing genetic information every time our cells replicate their DNA? While this is a concern, over a span of years, our cells have evolved mechanisms to deal with this.

For starters, our chromosomes have DNA sequences at each end that do not encode proteins; that is, at the beginning and end of each DNA molecule that makes up a chromosome, there is no information that is necessary for making proteins the cell needs. Instead, there are thousands of repeats of a short sequence that is rich in Ts and Gs.

These repeated sequences fold over, forming a loop-like structure that is stabilized by the binding of proteins, effectively tucking the ends of chromosomes out of sight when replication is not happening. These structures at the ends of chromosomes are called telomeres, which comes from Greek and means "end part."

These repeated sequences at the chromosome ends serve as protection in the sense that chromosomes could lose a few of them in each round of replication without losing important information. Our chromosomes shrink with each round of DNA replication, but we're OK until our cells have divided so many times that further shortening of the chromosomes starts to be a problem.

This helps explain an observation, first reported in the 1960s by Leonard Hayflick, that human cells that are grown in culture cannot divide indefinitely. This suggests that the shrinking of our chromosomes is kind of a biological clock—one that keeps track of how many times our cells have replicated their DNA and divided. And once a certain number of rounds of replication has occurred, we begin to lose important information, and our cells will stop dividing and eventually die.

If our chromosomes shrink with age, how do parents pass along full-length chromosomes to their children? If each generation were born with successively shorter and shorter chromosomes, then the human race would have died out.

If our telomere sequences did not shrink, would we live forever—or maybe, at least, stay youthful longer?

The answer to both of these questions lies in a second mechanism our cells use to deal with the end-replication problem.

This is a job for telomerase, an enzyme that can add telomere repeats to the ends of chromosomes. When telomerase is active, it adds many copies of the repeated sequence to the ends of chromosomes, ensuring that the chromosome shrinkage doesn't affect important information.

Telomerase adds many repeats before it eventually dissociates from the end of the DNA, having extended it by many nucleotides. As a result, there is now plenty of room to place an RNA primer on the newly extended strand and to complete replication of it. So, no information is lost, and the chromosomes remain the same length at each round of replication.

Aging

But if there is an enzyme that maintains chromosomes at the same length, why do they get shorter? That's because the magic of telomerase is active primarily in our egg and sperm cells and in stem cells, which are a stock of undifferentiated cells. Our somatic cells, with few exceptions, make little to no telomerase, even though all of our cells have the gene for it.

Because telomerase is active in egg and sperm cells, they retain full-length chromosomes, ensuring that the next generation starts out with lots of the telomeric repeats at the ends of their chromosomes. But in all the somatic cells not making telomerase, chromosomes do shrink with age.

Telomere shortening has been linked to aging and lifespan, even though there are other factors, such as oxidative damage, that also affect aging.

So, could we turn the telomerase back on in our cells, recover our lost telomeres, and recover our youthfulness?

One of the useful things about cells eventually ceasing to divide is that this provides a built-in limit on cell division. In fact, telomerase that stays turned on, with no built-in limit on cell division, is often found in cancer cells. In rapidly dividing cells, telomeres would normally shorten swiftly to the point where they would have to stop dividing. It is the reactivation of telomerase that allows cancer cells to keep proliferating without anything to check their division.

Cancerous cells are immortal. Even after they kill their host, they can be kept alive in a lab indefinitely; noncancerous cells can't be kept alive like that.

In lab experiments, telomerase inhibitors have been shown to be effective in killing cancer cells, and some are in clinical trials.

Because cancer cells express significant amounts of telomerase, unlike most normal cells, drugs that inhibit telomerase would be useful against cancer cells. And because close to 90% of all cancers express telomerase, a telomerase inhibitor might be useful in treating a variety of different cancers.

Telomerase activity is also inhibited by compounds found in foods such as curcumin from turmeric and allicin from garlic and onions. Green and black tea also have a phenolic compound that inhibits telomerase. Populations that routinely include these in the diet may have a lower incidence of cancers.

But such preventive measures cannot replace cancer therapies for those who already have cancer because they do not provide sufficient inhibition of telomerase activity once cancer is under way.

Although the relationship of telomere length to aging is not fully understood, people with longer telomeres seem to live longer. Conversely, telomeres are much shorter than normal in children with a mutation that causes progeria, a disease of premature aging. Shorter telomeres have also been found in boys with a form of muscle degeneration called Duchenne muscular dystrophy.

Even stranger is the finding that stress can lead to shorter telomeres. The length of a baby's telomeres has been found to be related to maternal stress during the pregnancy.

Is there anything we can do to prevent telomeres from shrinking before their time, without creating conditions that are favorable to cancer?

In 2017, scientists from Houston Methodist Research Institute were able to dial up telomerase activity in cells from children with progeria. They found that even a temporary increase in telomerase was able to turn back the cellular clock, effectively reversing the signs of aging in the cells.

READINGS

Armstrong, *Borrowed Time*.
Blackburn and Epel, *The Telomere Effect*.

QUESTIONS

1 Why is cell cycle regulation much more important for plants and animals than it is for bacteria?

2 Two single characteristics of every DNA polymerase are at the heart of why eukaryotic linear chromosomes shorten with each round of replication. What are they?

CLICK HERE TO SEE THE ANSWERS.

28 DNA MISMATCH AND EXCISION REPAIR

If it weren't for mutation, life on Earth would still be exactly the way it was when the first cell arose; with no change in the genetic information, there can be no change in the kinds of organisms that exist. Mutations are said to be the raw material of evolution, the variations in a population that make it more likely that there will always be some individuals that can deal with unexpected changes in the environment. Yet cells go to an enormous amount of trouble to prevent mutation—to defend against unwanted changes or damage to the information in our DNA.

The Nobel Prize was awarded to Tomas Lindahl, Paul Modrich, and Aziz Sancar for unraveling 3 major pathways for DNA repair. These mechanisms work ceaselessly to maintain the integrity of the genome and ensure that information passed on to the next generation is a faithful copy of the original.

Sources of Mutation and Damage to DNA

Information can be degraded through mutation—that is, changes in the DNA sequence—or by damage to the DNA molecule.

One way that DNA sequence can be changed is if mistakes are made while replicating DNA. And while proofreading by DNA polymerases reduces the incidence of such errors greatly, not all mistakes that are made during copying are caught and fixed.

Replication in our cells entails copying 6 billion nucleotides in each round. At that rate, there could be 600 uncorrected mistakes per round of replication in each cell, even with proofreading. When you factor in the trillions of cells in the body that are dividing at any given time, the total number of potential errors becomes truly staggering.

And even when DNA is not being replicated, it suffers damage over time. Sources include sunlight with wavelengths in the ultraviolet range and cosmic radiation from the Sun and stars. Radioactive elements, such as radon, are another source of damage. And x-rays and CT scans add to our radiation exposure.

Chemicals in the environment also cause damage to DNA. Mutagenic compounds are found in tobacco smoke and auto exhaust as well as in such apparently innocuous things as grilled meats. Even chemical reactions that naturally occur within cells give rise to compounds that can damage DNA.

And if the mutations and damage weren't corrected, we would be in a pretty sorry state. In fact, defects do sometimes get past our repair systems, and when that happens, various kinds of cancer and a slew of other health issues can result.

Mismatch Repair

The integrity of information in DNA is compromised when DNA polymerases incorporate the wrong nucleotide into the new strand and the proofreading system does not detect it. One way this might happen is by the transient formation of isomers of the bases, in which some of the atoms mediating hydrogen bonding temporarily change position and stabilize base pairs that do not normally form.

Mispaired nucleotides won't convey correct genetic information, but they do get incorporated into the new strand of DNA that is being built. And they do not get removed by the DNA polymerase's proofreading activity that checks for incorrectly paired bases, because the isomer base pairs appear legitimate for the brief time they isomerize. The isomers are transient, lasting no more than seconds, but it's enough to fool the DNA polymerase.

Insertions and deletions can also happen when DNA replication occurs in regions of short, repeated DNA sequences. These can happen as a result of one strand sliding with respect to the other.

While the error rate of DNA replication is about one in 10 million nucleotides in the absence of mismatch repair, in the presence of mismatch repair, this rate is reduced 100-fold to fewer than one in a billion nucleotides.

Errors not caught by proofreading during replication can be corrected by a mechanism called mismatch repair. Such a system must scan the newly made DNA to see if there are any mispaired bases (for example, a G across from a T), identify and remove the nucleotides in the region of the mismatch, and correctly fill in the gap created by the excision of the mismatched region.

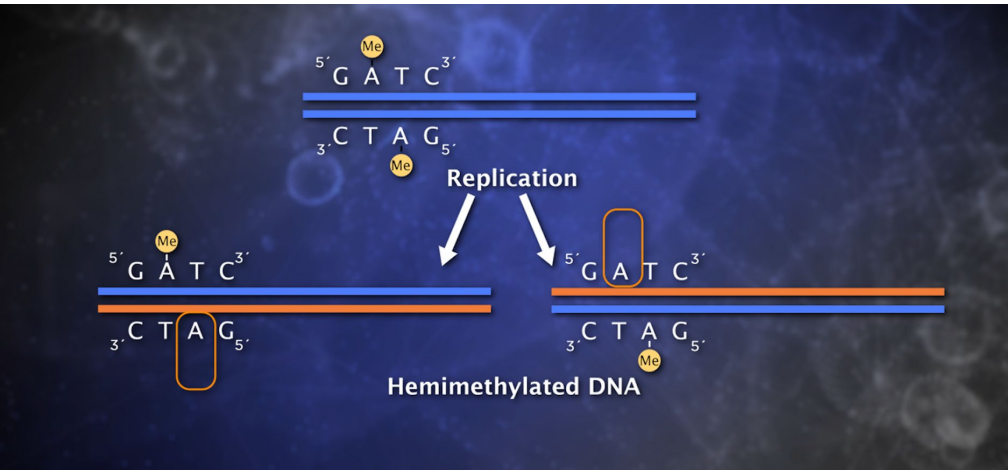
Crucially, the mismatch repair system must also have a means to distinguish the newly made DNA strand from the template strand if replication errors are to be fixed correctly. For example, when the mismatch repair system encounters an A-G mispair, it must know whether the A should be removed and replaced with a C or if the G should be removed and replaced with a T.

If mismatch repair randomly fixed problems, it will be wrong ½ of the time. So, how does the mismatch repair system distinguish between the original and the new strands of DNA and improve the error rate?

E. coli, for example, has an enzyme called DNA adenine methylase (DAM) that adds methyl groups to adenines in the GATC sequences of DNA. The sequence GATC reads the same way in the strand that is complementary to it.

When DNA replication copies one of the strands, the newly synthesized GATC will not be methylated at first, as this requires re-methylation by DAM to put the methyl group onto the adenine in the new strand. Until the methyl group gets put on, the GATC sequence is said to be hemimethylated: The old strand is methylated and the new strand is not.

The mismatch repair system in *E. coli* appears to assess the methylation status of GATC sequences nearest the mismatch/error. Newly made DNA has not yet undergone methylation and thus can be distinguished from the template strand, which is methylated at all GATC sequences.



The mismatch repair proteins selectively replace nucleotides in the strand that lacks methylation, thus ensuring that it is mistakes in the newly made strand that are removed and replaced.

Eukaryotes also have a mismatch repair system that repairs not only single-base mismatches but also insertions and deletions. Proteins corresponding to the *E. coli* system have been identified in other organisms, including humans. These, together with additional proteins, carry out mismatch repair in eukaryotic cells.

However, eukaryotes do not use DNA methylation to distinguish the new strand from the old. Instead, there is evidence that the newly made DNA may be recognized by the fact that it is nicked, or discontinuous. The presence of discontinuous strands may permit the new strand to be distinguished from the older continuous parental strand.

Defects in mismatch repair can spell trouble. In addition to leaving in mispaired bases, mismatch repair deficiencies can also lead to uncorrected errors in the replication of short, repeated sequences in DNA. Failure to handle short, repeated sequences is called microsatellite instability. This problem is usually prevented by functional mismatch repair, which can fix not only single-base mutations in DNA, but also small insertions or deletions of sequences during replication. When that ability is impaired, the repeated sequences vary in the number of repeats each strand contains.

Such variations are associated with a number of disease conditions. For example, mutations in the human mismatch repair genes are associated with the development of Lynch syndrome, also known as hereditary nonpolyposis colorectal cancer. People with Lynch syndrome have an 80% lifetime risk for colon cancer. They also have an elevated risk for other cancers, including those of the breast, prostate, bladder, and thyroid.

So, failures of mismatch repair have pretty severe consequences. But even with a perfectly good mismatch repair system, we cannot escape DNA damage. DNA is vulnerable, even when it isn't being replicated.

One source of DNA damage that surrounds us is exposure to ultraviolet (UV) radiation in sunlight. The same sunshine that lifts your mood and helps you make vitamin D in your skin can also wreak havoc as your skin cells steadily accumulate damage to your DNA, which will need to be repaired.

Cells with extreme levels of damage will be triggered to undergo programmed cell death so that they do not proliferate. Meanwhile, repair systems will do their best in the remaining cells. Proliferation of damaged cells can lead to skin cancer.

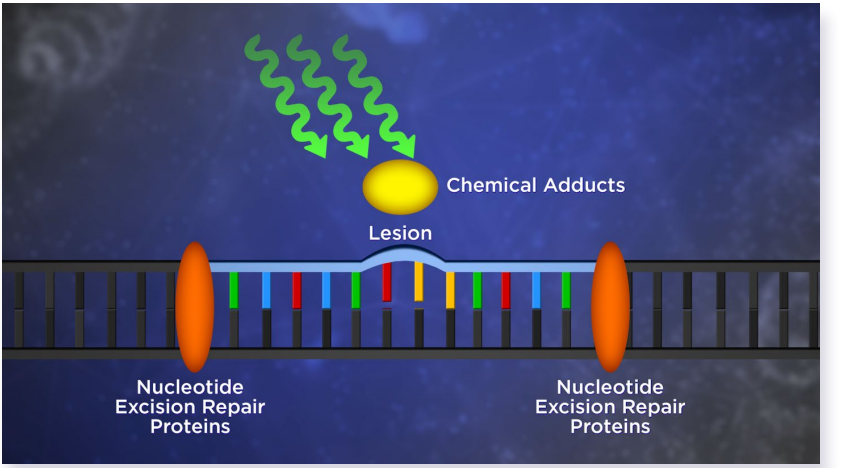
Using sunscreens or sunblocks that reduce the UV radiation that reaches your DNA is helpful. And being naturally dark-skinned provides some measure of protection because higher levels of melanin pigment in dark skin shields DNA from UV rays. But nothing can prevent a certain amount of damage that will need to be repaired by your cells.

Excision Repair

Cells have a set of proteins that deals with UV-damaged DNA. This system, called the nucleotide excision repair (NER) system, also fixes damage from environmental chemicals, such as those found in tobacco smoke and automobile exhaust.

These carcinogenic chemicals include molecules that can attach themselves to bases. These attachments of large chemical groups to bases in the DNA are referred to as chemical adducts or DNA adducts. Like the damage caused by UV radiation, chemical adducts can physically distort the DNA helix, causing DNA and RNA polymerases to stall when they attempt to copy those regions of DNA.

NER proteins cut the damaged strand on either side of the base that's been chemically altered, called a lesion. A short portion of the DNA strand containing the damage is then removed, and a DNA polymerase fills in the gap with the appropriate nucleotides.



NER has been extensively studied in bacteria. The damaged strand is recognized, and nucleases excise it while leaving its partner strand intact. The gap where the DNA was gets filled in by DNA polymerase and ligated together by DNA ligase.

NER is also important in eukaryotes. A number of proteins have been identified that function in ways similar to the bacterial repair proteins. The importance of these proteins is evident from the fact that mutations in the genes that encode them lead to genetic diseases.

Lifestyle interventions—such as scrupulously avoiding Sun exposure and consistent use of high-SPF sunscreen—can help reduce DNA damage. More recently, creams containing DNA repair enzymes have been formulated that show some promise in helping mitigate the consequences of accumulated UV damage.

One reason that humans are peculiarly susceptible to the harmful results of malfunctioning NER is that it is our only system for dealing with UV-induced lesions. Because we rely so heavily on our NER system, we're much more likely than plants and some other species to suffer from the effects of unrepaired UV damage. And we're out of luck if our NER does not work. But that may change in the future as science advances.

Other Sources of Genetic Mutation and Damage

In addition to environmental insults like radiation or chemicals, DNA can also be damaged by events inside the cells. For example, the DNA base cytosine is relatively unstable chemically and can sometimes spontaneously lose an NH_3 amine group and convert to uracil.

Other kinds of damage result from the normal chemistry of the cell. Reactive oxygen species (ROS) are molecules that are highly reactive because of their unpaired electrons. Such reactive species can react with guanine, for example, leading to an oxidized form of guanine. Methyl groups and other short carbon chains can be added to bases by chemicals called alkylating agents.

These modified bases do not actually change the physical structure of the DNA helix, but they can cause problems. The modified bases may end up with incorrect pairings compared to the unmodified forms, leading to mutations on the next round of replication and corresponding bulges in the DNA duplex.

For this reason, individual bases that have been modified or altered must be removed and replaced by repair systems. There are 2 ways this can happen. The first of these is the direct reversal of the modification with the help of a single-use enzyme that becomes nonfunctional after making the necessary fix.

Another repair system for individual bases is called base excision repair (BER), which can remove modified bases as well as uracils derived from cytosine deamination and is yet another multiprotein repair system. However, BER functions a little differently than NER. In BER, a single damaged base is first removed from the DNA, followed by the removal of a region of the DNA surrounding the missing base. The gap is then repaired.

READINGS

Branzei and Foiani, "Regulation of DNA Repair throughout the Cell Cycle."

Ralston, "Environmental Mutagens, Cell Signaling and DNA Repair."

QUESTIONS

1 Perfect DNA replication (with no errors) has enormous implications for evolution. Some say it would mean that cells would be identical to the way they originally were. Others disagree. Do you think this would be the case, based on information in this lecture? Why or why not?

2 Some wonder why DNA has thymine instead of uracil. Give one very good reason why, based on the information given in this lecture.

CLICK HERE TO SEE THE ANSWERS.

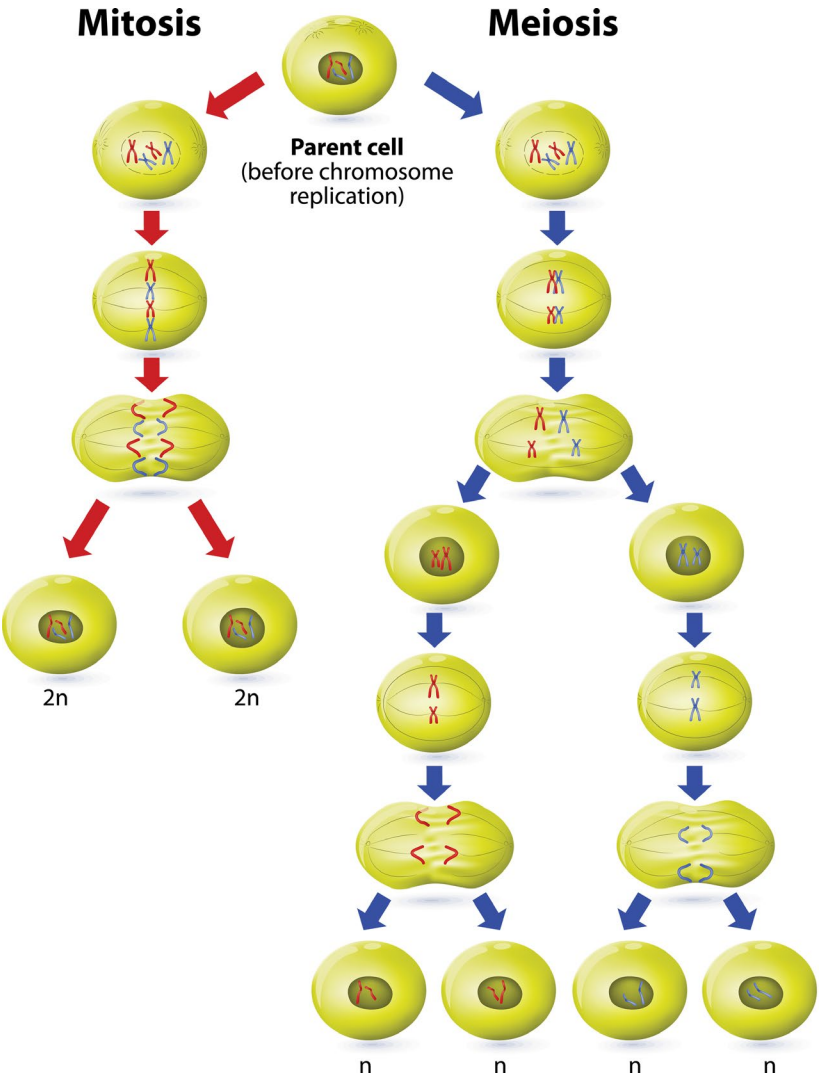
29 DNA RECOMBINATION, GENE EDITING, CRISPR

How is it that children of the same parents can end up so different from each other, and sometimes even from their parents? The answer involves a process called genetic recombination, which has the effect of assuring that no 2 individuals will have the same DNA, unless they are twins derived from a single fertilized egg.

Cell Division

Each individual has 2 sets of chromosomes, one inherited from each parent—23 chromosomes from the mother and 23 from the father. In both sexes, the cells that make the eggs or sperm undergo a special kind of cell division called meiosis. This reduces 2 sets of 23 chromosomes to one set of 23, resulting in what is called a gamete. Fertilization will unite the egg and sperm gametes, returning the chromosome number to 46.

In the regular cell division of mitosis, the chromosomes are duplicated and then partitioned to 2 cells so that each daughter cell gets a full set of 46. Gamete formation uses a different strategy for cell division called meiosis.


Mitosis
Meiosis
Parent cell
(before chromosome
replication)
2n
2n
n
n
n
n

In meiosis, in our cells, chromosomes are doubled initially, as in mitosis, to get 46 pairs of chromosomes. But after this, the cell goes through 2 cell divisions without any more chromosome replication. One division would leave us with 2 cells, each with a full set of 46 chromosomes, as in mitosis. But a second division, without further chromosome doubling, results in 4 cells, each with a single set of 23 chromosomes. This is how eggs and sperm come to have ½ as many chromosomes as the rest of our cells.

There are 2 features of meiosis that ensure that no 2 children who aren't identical twins will have the same DNA. First, the mother has 23 chromosomes each from her parents, so she has 2 each of chromosome 1, 2 of chromosome 2, and so on. When the mother's cells undergo meiosis to make eggs, the eggs randomly receive one of the 2 chromosomes from each pair.

So, an egg may receive the mother's father's chromosome 1 but the mother's mother's chromosome 2, and so on, for all 23 of the chromosomes. Theoretically, it is possible for an egg to receive all 23 of the mother's chromosomes that came from her father, or all 23 from her mother, or any combination in between.

The number of possible different combinations for 2 sets of 23 chromosomes is more than 8 million—from just the mother. So, although all eggs will have one complete set of 23 chromosomes, the likelihood of any 2 eggs having the exact same combination of chromosomes is vanishingly small.

The same math applies to the father's cells. So, one egg and one sperm are each drawn at random from their own 8 million possibilities. Together, there are 8 million possibilities times 8 million possibilities—or 64 quadrillion possible offspring!

That mind-boggling figure would be overwhelmingly more than enough to make every person's genetic inheritance unique, even if every single human being from all of human history had the same 2 parents. But there's even more.

This math assumes that the chromosomes you inherited from your mother and father were the exact same chromosomes they inherited from their parents. But that is not true.

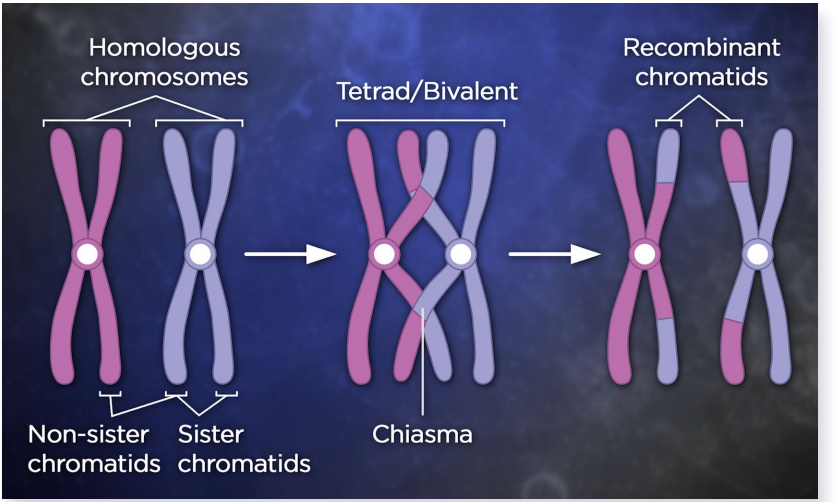
A second factor increases gene shuffling even more: During meiosis, to create an egg cell, the mother's chromosomes swap regions—or recombine—the genes she got from her parents. The same kind of swaps happen when sperm are created. These swaps within the maternal and paternal lines multiply the possibilities even more.

Swapping and Recombination

All the swaps and recombination take place within some very clear rules about what's permitted and what's not. The genes on each chromosome are all organized the same way, and that organization doesn't change. Chromosome 1 from your grandfather carries the same genes, in the same places, as chromosome 1 from your grandmother, only the versions of those genes can be different.

During meiosis, before the first round of cell division, there are 4 of every chromosome lined up so that all 4 copies lie alongside each other. This structure, called a tetrad, allows the chromosomes to swap sections without violating the rules about where genes are located.

When there's swapping during recombination, the organization of the genes remains the same. This kind of reciprocal exchange is called homologous recombination.



Because the chromosomes have different versions of the same genes, this sort of swapping shuffles the versions of certain genes present on each chromosome. So, the chromosomes that parents pass on to their offspring are not the exact same chromosomes they themselves carry; each parent passes on a mash-up of the chromosomes of the 2 grandparents.

Recombination can happen in every chromosome, so the egg and sperm have even more diversity than is generated simply by random assignment of the original chromosomes from the mother and father. So, layered on top of the 8 million times 8 million possibilities are now even more combinations.

Is it any wonder that no 2 humans except identical twins have identical DNA?

Information in 2 adjacent DNA strands can be swapped during meiosis. And 2 chromosomes can undergo recombination at multiple spots so that there's even greater diversity generated.

DNA Repair

Many of the same proteins that are responsible for recombination are used for another crucial cellular function: the repair of double-strand breaks in DNA.

Double-strand breaks are among the most dangerous forms of damage that happen to DNA. When a DNA molecule is broken on both strands, the 2 pieces of double-stranded DNA can separate. This double-strand break fragments the DNA and could lead to cell death. So, it's really important to ensure that breaks are promptly fixed.

Double-strand breaks can arise from chemical damage or cellular malfunction, but they must be fixed or they can be lethal. In a single-strand break, the problem is fixed by copying the undamaged strand. But when both strands are broken, this is not an option.

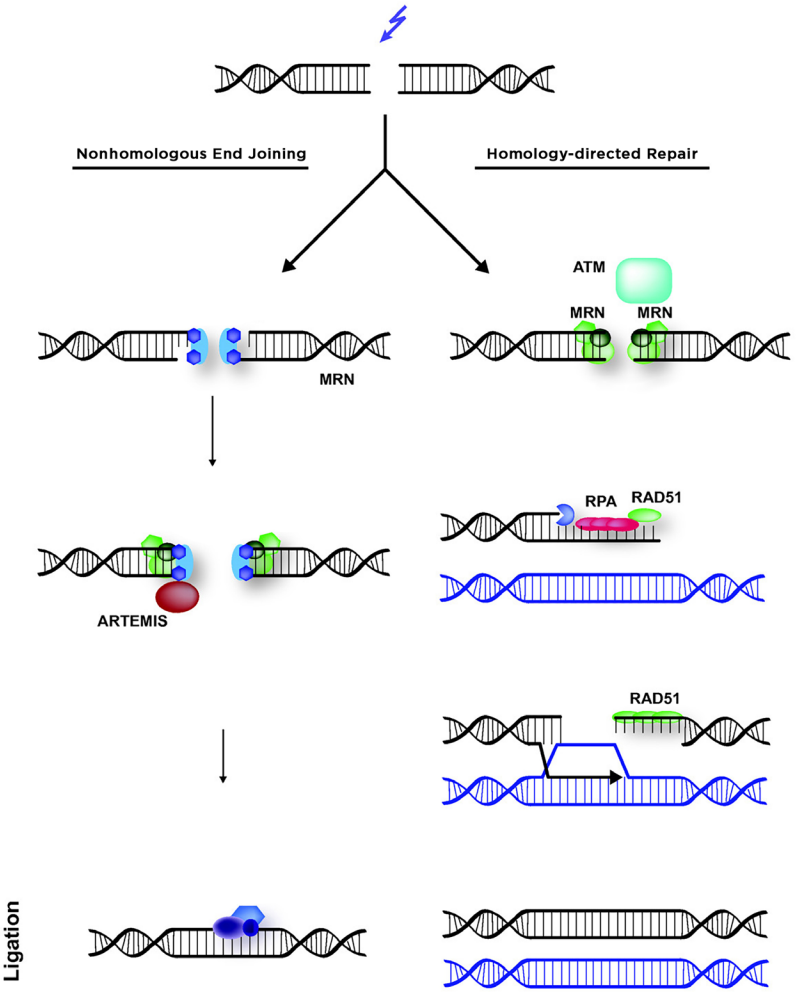
One solution for cells is to use a repair mechanism called nonhomologous end joining. It's nonhomologous because the ends of the DNA don't have to match. The idea is to grab the ends of the broken DNA and reconnect them using a ligase enzyme.

But this kind of end joining often chews back the ends of the broken strands before rejoining them. As a result, when the broken ends are ligated back together, there may be deletions or insertions of bases.

So, nonhomologous repair is fundamentally different from recombination: It does not return the DNA to its original state and is not an ideal way to make a repair. But that's where recombination can help.

Just like in homologous recombination during meiosis, chromosomes are deliberately broken by nucleases that introduce double-strand breaks. And we end up with perfectly good chromosomes again by using a matching chromosome, lined up perfectly, to provide a template to copy. Why not use this same mechanism to fix broken chromosomes?

Cells do, indeed, repair double-strand breaks by using a mechanism called homology-directed repair. This is very similar to meiotic recombination, and it allows double-strand breaks to be fixed perfectly. But chromosomes are not always conveniently aligned, so at other times, cells have to make do with nonhomologous repair.



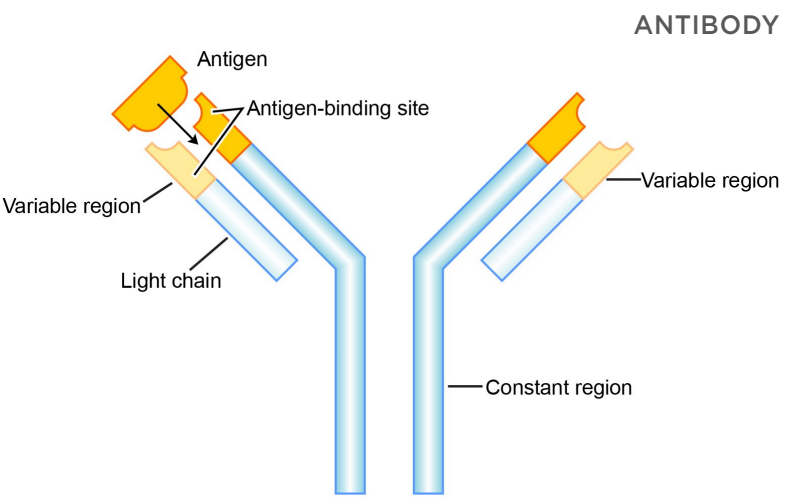
Antibody Creation

Recombination also helps generate the enormous variety of antibodies we need to fight off invaders like bacteria and viruses. Antibodies are defense proteins made in vertebrates that bind tightly and very specifically to targets like microorganisms or foreign molecules. The binding of antibodies to the intruders directly inactivates them or marks them for destruction.

Our bodies make billions of different antibodies, each specific for a particular target, or antigen. How can we make billions of proteins that are different with fewer than 25,000 genes in the entire genome?

There isn't a separate gene for each individual antibody. Instead, the antibody genes are broken up into chunks or segments, and there are many different versions of each segment. Using a special kind of recombination that is unique to somatic cells, immune cells create different combinations.

Each cell picks one segment from each category and joins segments together to make its own special variant antibody that recognizes and binds to one very specific target. Given the number of choices of each segment, immune cells could potentially make about 300 billion different antibodies!



Somatic recombination depends on the activity of a number of enzymes that help with the double-strand breaks and rejoining necessary for creating the hybrid antibody genes.

Gene Inactivation and Editing

For both recombination and repair, the similarity of sequences is what allows the initial base pairing of the broken strand with one of the strands of the intact DNA. Homologous recombination, then, is based on sequence similarity.

On the other hand, the cell's crude repair of double-strand breaks in DNA using nonhomologous end joining suggests a way to deliberately inactivate genes. After all, the process causes deletions and insertions, and that can be enough to inactivate the gene.

For researchers, inactivating a gene makes it possible to find out what role it plays in the organism. The challenge is finding a way to create double-strand breaks in the gene of interest at just the desired spot. This was the sticking point until the advent of a method that made it so easy to inactivate or edit genes that an entire genome's worth of mutants could be made in short order.

The name of this game-changing method is clustered regularly interspaced short palindromic repeats (CRISPR). This new tool to inactivate or edit genes started with a sort of immune system for bacteria, which can be attacked by viruses called bacteriophages.

To protect themselves against bacteriophages, bacteria have a clever system to recognize a virus that has attacked them before. They store pieces of the virus sequence so that they can ID the virus if they see it again. The pieces of virus sequence are stored in the bacterium's own genome, separated by short, repeated sequences of bacterial DNA.

The short, repeated sequences are palindromes, which in the world of DNA means that a sequence on the one strand is the same as the sequence on the complementary strand but read in the opposite direction.

GGATCC
CCTAGG

In plain English, a palindrome is a word that reads the same backward and forward, such as *radar* and *level*.

Thanks to the good filing system made possible with the short palindromic repeats, the bacteria are ready during a subsequent viral attack. The sequence of the invading virus can be compared with the collection of stored sequences, and an enzyme is dispatched to destroy any sequences that match with the stored sequences.

If the bacteria have stored sequences that came from a virus, the bacteria can copy the virus sequence into an RNA called a guide RNA. Because the sequence of the guide RNA matches part of the sequence of the attacking virus, the guide RNA can base-pair with the viral sequence. Attached to the guide RNA is an enzyme called Cas that can then cleave the virus genome.

Cas9 refers to the Cas enzyme discovered to have the desired effect. It came from the bacterium *Streptococcus pyogenes*, which was chosen for closer study because of its relatively simple CRISPR system. This is why you often see CRISPR referred to as CRISPR-Cas9.

Amazingly, we can use this same system, but paired with any guide RNA we can design. Because the guide RNA is what determines which sequence gets cut by Cas9, all we need is to design guide RNAs that base-pair with the gene sequence we want to target.

Basically, we steal the Cas9 enzyme and pair it with a guide RNA that will take it straight to the gene we want taken out. It works in any kind of cell—not just bacteria. This RNA-guided Cas9 missile is fast, easy to deploy, and amazingly effective. It makes a double-strand break in the DNA that the guide RNA leads it to. It can make double-strand breaks in any DNA—not just viruses.

And when there's a double-strand break, the cell rushes to fix it. If it's timed right, when the cell isn't preparing to divide, the double-strand break will be fixed by nonhomologous end joining, which results in deletions and insertions. Such sloppy repair has the benefit of rendering the target gene inactive.

Researchers can now create knockouts of every gene in a genome—one by one or in combination—to find out what role each plays. First-generation CRISPR works well to knock out genes, but sometimes there are off-target effects—for example, if a guide RNA mistakenly binds to a sequence that is similar but not identical to the intended target. Scientists are working on improving the specificity of the targeting by using Cas enzymes that are more accurate than the original Cas9.

Missile-strike knockouts that destroy a target are just the beginning. CRISPR and related techniques can and will do much more.

Researchers are also trying to perfect a CRISPR-like system that would deliver an enzyme that can edit a single base in the DNA. If this method is optimized, it would enable researchers to easily correct point mutations like the one that causes sickle cell anemia.

READING

Doudna and Sternberg, *A Crack in Creation*.

QUESTIONS

1 Recombination helps the immune system create the enormous numbers of antibodies it needs to protect the body. If the process of making antibodies were merely random, a given antibody would only be available in miniscule amounts and likely not when needed. Speculate on how the immune system avoids the problem of randomness.

2 Gene drives are remarkable technologies that bypass normal rules of inheritance. Speculate how they might be used to reengineer a species.

CLICK HERE TO SEE THE ANSWERS.

30 TRANSCRIBING DNA TO RNA

RNA is much more than simply a copy of the DNA blueprint. RNA not only carries the message to make a protein, but it also provides the decoding rings necessary for converting nucleotide sequences to amino acid sequences in proteins. And RNA is even an important component of the ribosomes that actually make the proteins.

Copying RNA

The copying of RNA is catalyzed by an enzyme called RNA polymerase. Unlike DNA polymerases, cells don't have so many different kinds of RNA polymerases. Prokaryotes have only one; humans have 3.

Transcription, like DNA replication, involves building a string of nucleotides across from a DNA template using the base-pairing rules.

To build an RNA, cells use RNA nucleotides, rather than DNA nucleotides. RNA nucleotides are very similar to DNA nucleotides, with 2 important differences: The sugar in RNA nucleotides is ribose, not deoxyribose; and instead of the bases A, T, G, and C, RNA uses A, U, G, and C (U pairs with A, just like T).

The word *transcription* is used to indicate the reading of DNA sequence to make RNA.

Regions of DNA that get transcribed into RNA are called genes. In a narrow sense, the word *genes* is used to indicate only the regions of DNA that code for protein, but in a broad sense, genes are DNA sequences that get copied into RNAs.

Transcription to RNA differs from copying DNA. First, when a sequence of DNA is transcribed to RNA, only one of the 2 DNA strands is copied into RNA. The DNA strand that is copied is called the template strand. The DNA strand that is not copied is called the nontemplate, or coding, strand, and it has the same sequence as the RNA that is made, except that it has a T wherever the RNA has a U.

Cells make several different kinds of RNA, and not all types of RNA code for proteins. There are 3 major types of RNA in cells:

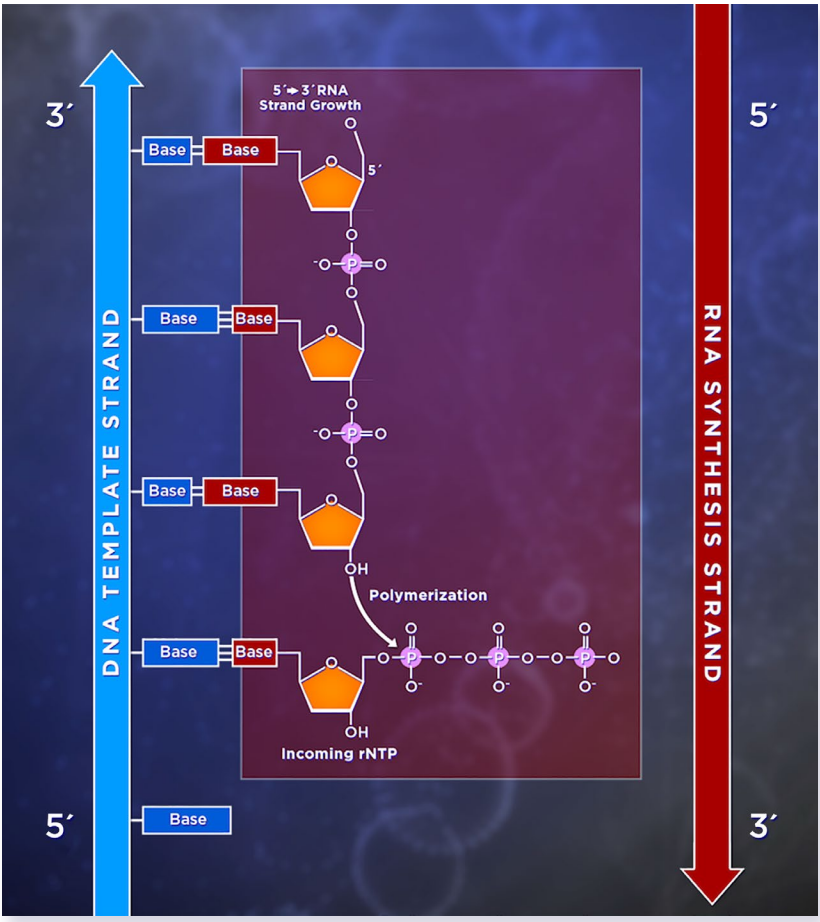
1. Messenger RNA (mRNA) is responsible for carrying coded information in the form of codons, which are groups of 3 nucleotides, for making proteins from the DNA to the complexes in the cell known as ribosomes, which are the site of synthesis of proteins.

2. Transfer RNA (tRNA) has 2 roles: carrying amino acids to the ribosomes and providing a decoder of the codons called anticodons. These functions are at opposite ends of the tRNA; one end has a covalent bond to the amino acid, and the other end has the decoder.

3. Ribosomal RNA (rRNA), a component of the ribosomes that makes proteins, provides a scaffolding where ribosomal proteins bind and form the overall structures.

Making RNA

The synthesis process is essentially the same for all types of RNA. It is also the same as the addition of nucleotides for making DNA. As with DNA replication, RNA synthesis only advances in the 5′ to 3′ direction.

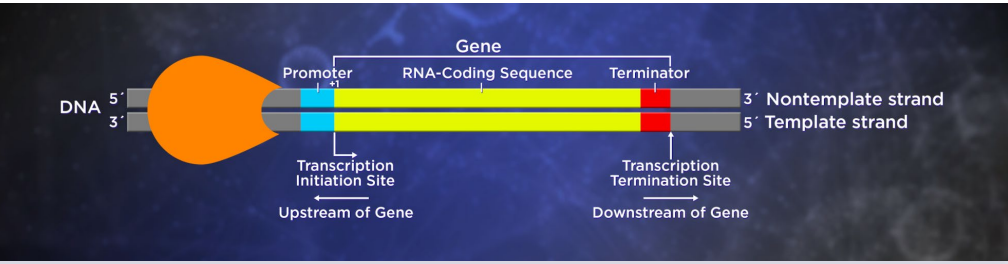


One important difference between DNA polymerases and RNA polymerases is that RNA polymerases do not require a primer to start making RNA. Once RNA polymerases are in the right place to start copying DNA, they just begin making RNA by copying the DNA template.

Unlike DNA replication, where every nucleotide of the parental DNA gets copied, transcription only copies selected regions of DNA. So, how do RNA polymerases know where to start copying on the DNA?

It turns out that cells have signs for RNA polymerase telling it where to start and stop copying. The start sign is a sequence in the DNA called a promoter, and it is located adjacent to a gene. Likewise, there are sequences called terminators in the DNA that act as stop signals.

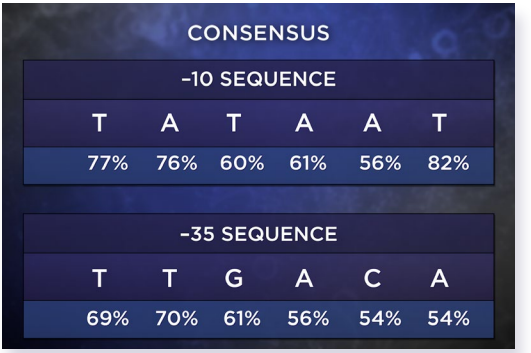
The promoter is said to control the gene it is associated with. This is because RNA polymerase physically binds to the promoter sequence to begin transcription. If the RNA polymerase does not bind at the promoter, the gene will not be transcribed.



Transcription in Prokaryotes

Much of what we know about transcription in prokaryotes comes from studies of *E. coli*. Prokaryotes have a single RNA polymerase that transcribes all of their genes. *E. coli* has only one DNA replication origin; by contrast, it has thousands of promoters, and their sequences are not identical.

The transcriptional start site, called the +1 site, is the place in the DNA across from which the first RNA nucleotide is placed to start synthesis. The promoter sequences are positioned before the start site. The first of these promoter sequences is found at –10. It is a sequence of 6 bases—TATAAT—and it's found in the majority of promoters. There is some variation of this sequence from gene to gene, but TATAAT is called the consensus sequence.



Farther upstream from +1, centered at position –35, is another sequence of 6 bases that is quite similar for most *E. coli* genes. Here, the consensus sequence, which is the most common base found at each position, is TTGACA.

The sequences centered at –10 and –35 are recognized and bound by RNA polymerase before transcription can begin. RNA polymerase of *E. coli* contains several polypeptide chains. One subunit that's temporarily attached is called sigma, and it can bind to the –10 and –35 sequences in the promoter.

The sigma subunit acts like an usher that guides RNA polymerase to its "seat" on the promoter. After transcription begins, the sigma subunit separates from the polymerase, which continues RNA synthesis. Meanwhile, the sigma subunit finds another RNA polymerase to escort to a promoter.

At one level, it is the job of promoter sequences to determine how much RNA is made in cells.

Once bound to the promoter, the RNA polymerase unwinds the DNA strands to create a transcription bubble. In contrast to the bubble found in DNA replication, the transcription bubble does not have replication forks at each end. Unwinding the strands to form the bubble is made easier by the fact that most promoters have a sequence that is rich in A-T base pairs. Unlike DNA polymerase, RNA polymerase requires no primer to start synthesis.

At first, the RNA polymerase simply strings together 8 or 9 nucleotides complementary to the template while still sitting at the promoter. Then, it begins to move along the template, unwinding the DNA ahead of it. No helicase enzyme is needed because RNA polymerase can handle the unwinding, too.

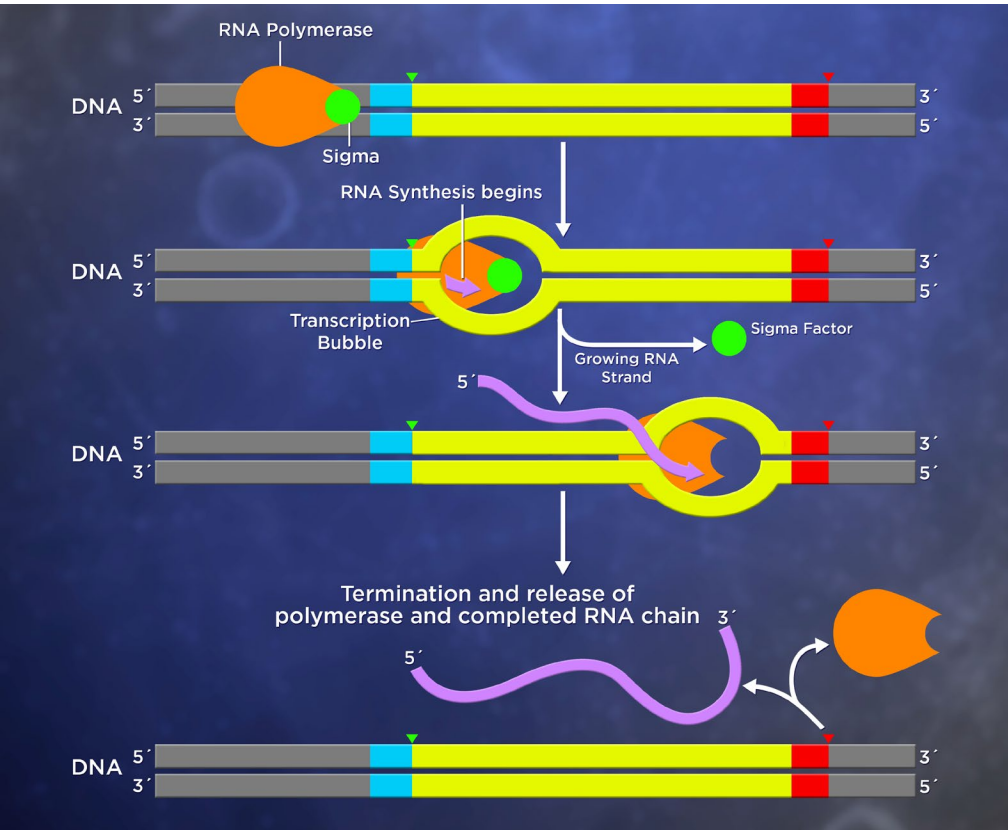
As it moves, RNA polymerase maintains a transcription bubble of 12 to 15 base pairs and synthesizes RNA complementary to the template strand at the rate of about 30 to 50 nucleotides per second. And behind the RNA polymerase, the DNA template is rewound, displacing the newly made RNA from its template strand.

Just as in replication, the opening up of the DNA strands for transcription creates overwinding ahead of the bubble. And just as in replication, DNA gyrase comes to the rescue to release the tension.

An RNA chain, complementary to the DNA template, is built by the RNA polymerase by the joining of the 5′ phosphate of an incoming ribonucleotide to the 3′ OH on the last nucleotide of the growing RNA strand.

How does the polymerase know where to stop? That's where the terminator sequence built into DNA gives the signal to the RNA polymerase to stop transcription and dissociate from the template.

Once the RNA polymerase is positioned correctly, it strings RNA nucleotides together, moving down the template until it has synthesized the sequence called the terminator, which helps signal the polymerase to stop transcribing and release the newly made RNA from the DNA template.



In bacteria, if the RNA being made encodes a protein, ribosomes will bind to the starter end at 5′ long before the entire RNA is made and released from the DNA. The ribosomes will begin translating the mRNA to make the protein it codes for, even as the RNA polymerase is still transcribing the rest of the RNA. Doing both at the same time is possible because in bacteria, there is no nucleus, so transcription and translation occur in a single compartment.

Transcription in Eukaryotes

The process of making RNA is essentially the same in eukaryotes as it is in prokaryotes. The differences are external to the synthesis process. Eukaryotic DNA exists as chromatin, where the DNA is tightly associated with histones and other proteins, increasing the complication of the process. Such packaging of DNA must therefore be loosened to allow the RNA polymerase access to the template in the region to be transcribed.

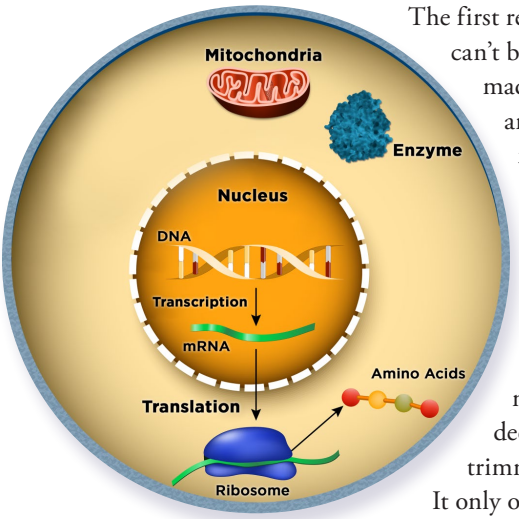
A second difference in eukaryotic transcription is the presence of 3 RNA polymerases, not one, as in bacterial cells. These different polymerases transcribe different subsets of RNAs. All 3 eukaryotic RNA polymerases need additional proteins to help them initiate transcription. In prokaryotes, RNA polymerase had the sigma factor helping it find and bind to promoters. Additional proteins are needed by eukaryotic RNA polymerases, and they are referred to as general transcription factors.

Finally, eukaryotic cells have a nucleus, and that is where transcription happens. RNAs produced there are processed further before they are sent into the cytoplasm. Protein synthesis, or translation, happens in ribosomes, as in prokaryotes, but these are located in the cytoplasm.

Eukaryotic promoters have some similarities to prokaryotic promoters. Eukaryotic RNA polymerases need general transcription factors to help the RNA polymerases find the promoter and initiate RNA synthesis.

Transcription in eukaryotes requires the general transcription factors and the RNA polymerase to form a basal transcription complex. The RNA polymerase is recruited to the promoter by a bunch of transcription factors that first bind there and is sent on its way down the template, when it gets a nudge from a general transcription factor, which phosphorylates the polymerase's tail. RNA synthesis proceeds, as in prokaryotes.

Though bacteria could begin using RNAs for protein synthesis even before mRNA synthesis was complete, several differences in eukaryotes preclude that possibility.

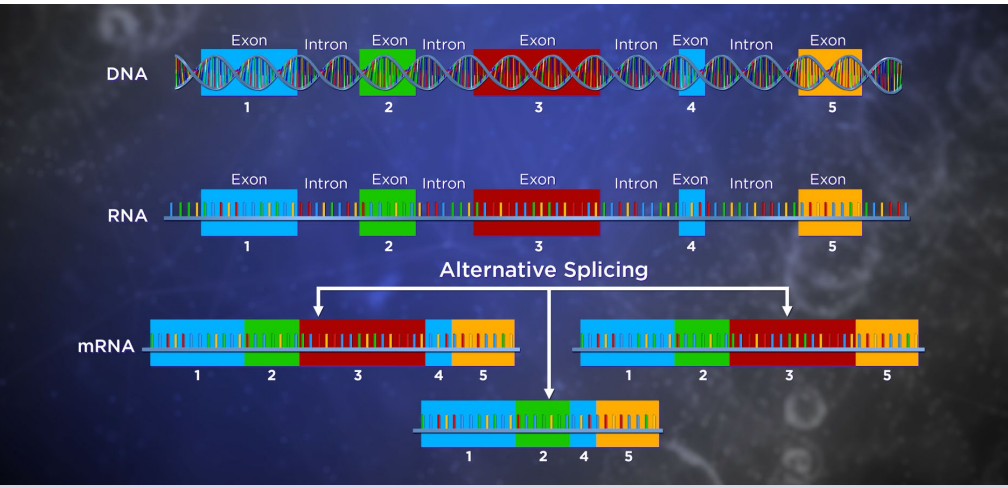


The first reason eukaryotic mRNAs can't be translated as they are being made is that eukaryotic mRNAs are made in the nucleus, but ribosomes for translating them to proteins are located in the cytoplasm. As a result, mRNAs must be moved to the cytoplasm.

Before they're moved, though, mRNAs get trimmed and decorated in other ways. The first trimming process is called splicing. It only occurs in eukaryotes.

In prokaryotic genes, the information for making proteins is contained in one uninterrupted block directly from the DNA. By contrast, eukaryotic genes have introns, noncoding regions that interrupt the coding information, which is contained in segments called exons. Plants, animals, and other eukaryotic cells cut out the introns and connect the bits together before the mRNA is sent out of the nucleus to be used for protein synthesis. The whole process of removing the introns and joining the exons is called splicing.

In eukaryotic mRNAs, some of the exons may also be removed from a transcript to make the spliced RNA. As a result, it becomes possible to create mRNAs that will be translated to make slightly different proteins. This phenomenon is called alternative splicing, and because of it, we can make more than 100,000 different proteins from fewer than 30,000 genes.



Splicing is a double-edged sword. Though it enables cells to increase the kinds of proteins they make from a given gene, mutations affecting splicing can have profound effects on human health.

It is estimated that as much as 22% of gene mutations affect splicing and cause diseases ranging from Duchenne muscular dystrophy to Hutchinson-Gilford progeria syndrome.

Splicing isn't the only alteration eukaryotic cells make to their mRNAs. They also decorate them—at both ends—with special structures. The 5′ end gets a cap, which contains a modified guanine nucleotide. The cap protects the 5′ end of the mRNA from degradation by nucleases, because it alters the normal 5′ end they would attack. The cap is also important in protein synthesis.

The transcription of eukaryotic mRNAs does not end at fixed sites. Instead, a special "this is the end" structure is added before termination occurs. This is done by cutting the mRNA to create a new, specific 3′ end. Then, an enzyme adds a tail of about 200 adenine-containing nucleotides to the new 3′ end. There is evidence that this tail, called a polyA tail, plays a role in the efficient translation of the mRNA as well as in the stability of the mRNA.

The 5′ cap and the 3′ polyA tail on an mRNA signal that the processing of an mRNA is complete. At this point, the transcripts are described as mature and the mRNAs can be moved to the cytoplasm, where they direct the synthesis of proteins.

READING

Witkowski, ed., *The Inside Story*.

QUESTIONS

1 Bacterial cells make several different sigma subunits. One of these is important in heat shock. Based on what you learned in previous lectures, describe how this might help cells recover specifically to heat shock.

2 Compared to DNA replication, which operates in *E. coli* at 1000 nucleotides per second, transcription only occurs at about 50 nucleotides per second. How can transcription operate so much slower and still serve the cell's needs?

CLICK HERE TO SEE THE ANSWERS.

31 TRANSLATING RNA INTO PROTEINS

The recipes for proteins are found in the base sequences of the genes that encode them. These base sequences are transcribed from storage in DNA into messenger RNA, just as individual recipes may be transcribed from a cookbook onto an index card. The cooks—large cellular complexes called ribosomes—translate those messenger RNA (mRNA) recipes from genetic code into proteins. And amino acids are the ingredients that are needed to build the proteins. The catch is that even the ribosome cooks cannot read the directions for making the proteins, nor can they recognize the amino acids. So, how do proteins ever get made?

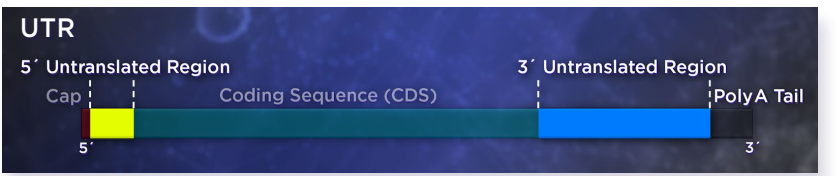
Translate can be defined as "changing something into a new form, especially turning a plan into something real."

In protein synthesis, the RNA is a plan for building a protein, and during translation, that plan is used to create something real: a protein.

Inventory for Protein Production

The first thing that's needed for protein production is the instructions in mRNA. Bacterial mRNAs come off the DNA template ready to use. In eukaryotic mRNAs, the 5′ end has a special nucleotide cap and the 3′ end has a polyA tail, both of which signal to the ribosome that this is an mRNA.

The part of the mRNA coding for the protein is sandwiched between noncoding sequences at both the 5′ and 3′ ends. These sequences, called untranslated regions (UTRs), have important roles to play in the translation process. At the 3′ end, the UTR sequences determine the stability of the mRNA—that is, how long the mRNA sticks around to be translated before it gets broken down.



In addition to the instructions and the recipe, a cook—the ribosome—is needed to mix the ingredients to make the protein. Ribosomes are very large complexes of proteins and ribosomal RNAs (rRNAs) that do the actual work of joining the amino acids together by making peptide bonds between them. Both prokaryotic and eukaryotic ribosomes contain 2 complexes: the small and large ribosomal subunits.

But the cook can neither read the recipe nor tell one ingredient from another, so it needs help. And help comes in the form of transfer RNAs (tRNAs), which are translators that speak RNA language and can tell one amino acid from another.

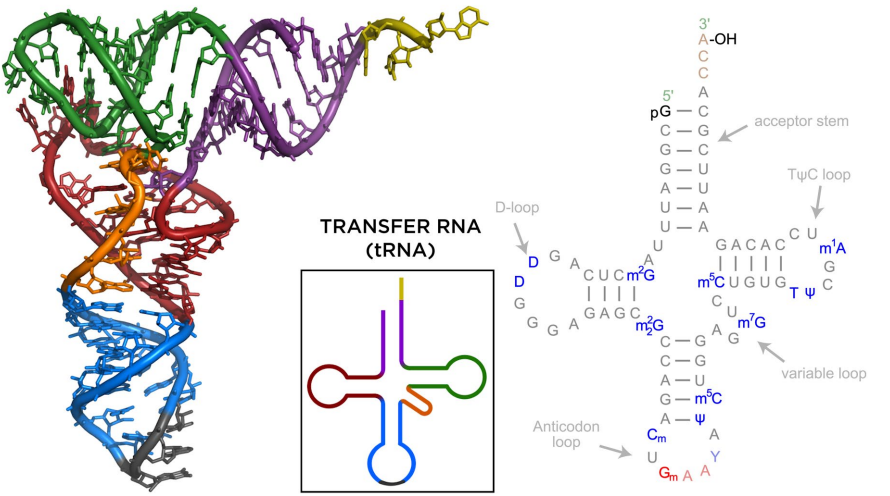
Transfer RNAs are about 75 to 80 nucleotides long, and they fold into an intricate 3-D shape with 3 major loops. In addition, tRNAs are distinctive in having their bases modified in many ways after they are transcribed from

DNA. As a result, tRNAs contain many unusual bases—in addition to A, U, G, and C—that play roles in folding and stability of the tRNAs as well as in their interactions with mRNA.

There are 20 different classes of tRNA, one for each amino acid. Members of each class carry and are specific for only one of the 20 amino acids in proteins. Each tRNA carries at its 3′ end a specific amino acid. At the other side of the tRNA, the anticodon loop has a 3-base sequence complementary to the codon found in mRNAs.

There are 20 enzymes, called amino acyl tRNA synthetases, which each read the anticodon loop and attach the appropriate amino acid at a tRNA's 3′ end. Each amino acid has its own unique synthetase enzyme for attachment to its corresponding tRNA(s). A tRNA linked to an amino acid is said to be charged.

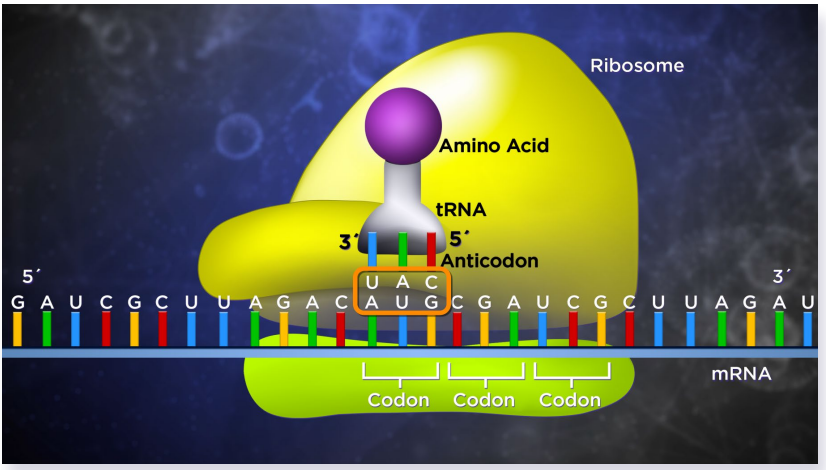
So, the tRNAs each carry a specific amino acid. But how do they know what order to line up in to make a protein? That information is in the mRNA, but it falls to the tRNAs to read the directions. The ability of tRNAs to read information in mRNA relies on specific molecular interactions.



The genetic code in messenger RNA is written in codons—3 bases that specify each amino acid. What tRNA uses to translate that message is an anticodon. The 3-base sequence of the tRNA anticodon is complementary to the codon in mRNA that specifies an amino acid. Complementarity is determined by the base-pairing rules for RNA: A goes with U, while G goes with C. For example, a messenger codon of GAG would have a transfer anticodon of CUC.

When each tRNA encounters the codon on the messenger RNA, the anticodon will base-pair with it, and the amino acid attached to the other end of the tRNA will be added to the growing protein chain.

The first codon in most mRNAs is AUG, which specifies the amino acid methionine. The sequence of this codon in mRNA, reading 5′ to 3′, is AUG. Many tRNAs are floating around, each linked to specific amino acids. The only one with the secret handshake that will allow it to pair with the AUG codon is the one whose anticodon sequence is complementary to AUG—which is CAU.



Other tRNAs, carrying other amino acids, will have other sequences in their anticodons, so they will not pair with the AUG.

This table shows which codons specify which amino acids to help predict what the anticodon sequences should be on the corresponding tRNAs. There are 64 codons, but not 64 tRNAs. How does that work?

Genetic Code

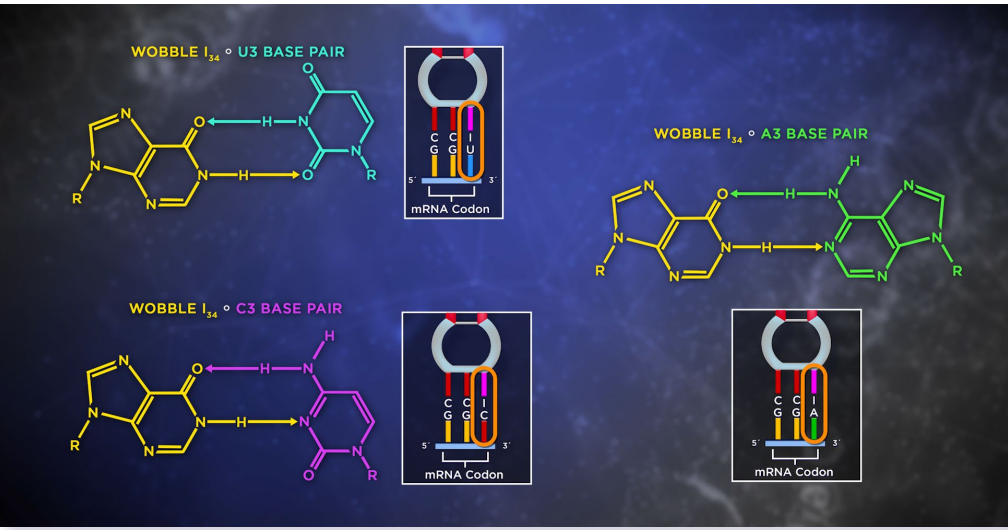
SECOND BASE OF CODON

FIRST BASE OF CODON	T		C		A		G		THIRD BASE OF CODON
T	TTT	Phenylalanine	TCT	Serine	TAT	Tyrosine	TGT	Cysteine	T
	TTC		TCC		TAC		TGC		C
	TTA	Leucine	TCA		TAA	STOP codon	TGA	STOP codon	A
	TTG		TCG		TAG		TGG	Tryptophan	G
C	CTT	Leucine	CCT	Proline	CAT	Histidine	CGT	Arginine	T
	CTC		CCC		CAC		CGC		C
	CTA		CCA		CAA	Glutamine	CGA		A
	CTG		CCG		CAG		CGG		G
A	ATT	Isoleucine	ACT	Threonine	AAT	Asparagine	AGT	Serine	T
	ATC		ACC		AAC		AGC		C
	ATA		ACA		AAA	Lysine	AGA	Arginine	A
	ATG	Methionine	ACG		AAG		AGG		G
G	GTT	Valine	GCT	Alanine	GAT	Aspartic acid	GGT	Glycine	T
	GTC		GCC		GAC		GGC		C
	GTA		GCA		GAA	Glutamic acid	GGA		A
	GTG		GCG		GAG		GGG		G

It turns out that a given tRNA can base-pair with more than one codon. There are rules about which base pairs with which, but tRNAs use a trick. The base-pairing of the anticodon with the codon starts at the 5′ end of the codon. While the first 2 base pairs must follow the base-pairing rules strictly, the third base—called the wobble base—may get away with a nonstandard pairing.

In the genetic code, many of the codons that encode the same amino acid vary, or wobble, in the third position. Given this, a trick can be used for those codons.

Remember that tRNAs have modified bases. One such modified base is part of the nucleotide inosine, which is often found at the anticodon's 5′ end. Inosine's modified base is a purine that can base-pair with either of the standard pyrimidines, uracil or cytosine, as well as with the purine it's derived from, adenine. This allows an anticodon ending with the modified base to pair with codons that have U, A, or C at their third position, as long as the first 2 bases of the codon make normal base pairs with the anticodon.



Stages of Protein Synthesis

The main steps of protein synthesis are very similar in prokaryotes and eukaryotes.

Translation happens in 3 stages.

The first stage, known as initiation, brings the mRNA and ribosome together with the first tRNA. Remember that the first amino acid in proteins is methionine, so this first tRNA—called the initiator tRNA—carries methionine. Complementary to the methionine codon, AUG, in the mRNA, the tRNA would have a CAU anticodon.

In bacteria, a modified methionine called formyl methionine is the first amino acid, but that makes no significant difference to the process.

The ribosome has 2 subunits, a small and a large. When assembled, the ribosome has 3 spots for a tRNA: the amino acyl (A) site, the peptidyl (P) site, and the exit (E) site. The coming together of the ribosome with the mRNA starts when the initiator tRNA binds to the small ribosomal subunit.

In eukaryotic cells, the small ribosomal subunit with the tRNA then binds the 5′ cap of the mRNA. The small subunit, with its initiator tRNA, then slides along the mRNA until the first codon, AUG, is positioned at the P site of the ribosome, allowing the anticodon of the initiator tRNA to base-pair with the AUG codon.

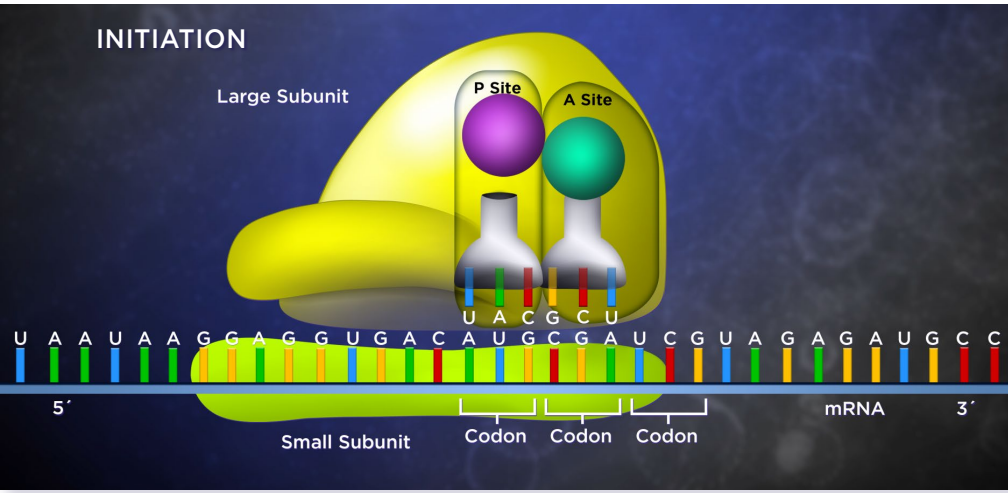
In bacteria, a ribosomal RNA component of the small subunit recognizes and binds a sequence in the 5′ UTR. This component is called the Shine-Dalgarno sequence, and it binds to a piece of ribosomal RNA and positions the ribosome and initiator tRNA correctly.

The large ribosomal subunit then joins the complex, and translation is set to begin. At this point, the P site of the ribosome is occupied by the initiator tRNA base-paired to the start codon. The next codon on the mRNA is positioned in the ribosome's A site, ready for the appropriate charged tRNA to base-pair with it.

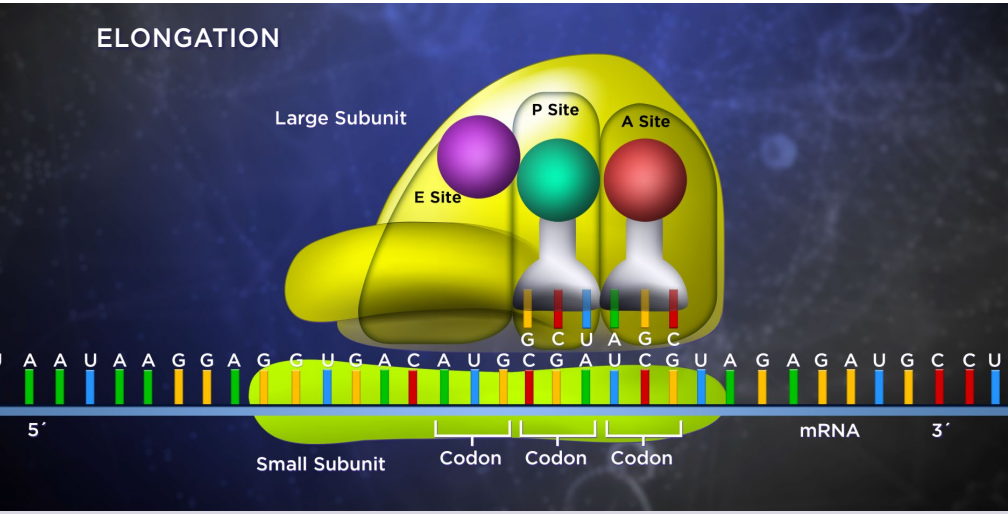
To bind at the A site, charged tRNAs need help from proteins called initiation factors, as well as energy from the hydrolysis of GTP. In this context, GTP functions like ATP, providing a high-energy molecule for cellular reactions.

When the correct tRNA corresponding to the second codon has base-paired with the codon in the A site, the amino acid it carries is positioned close to the methionine attached to the initiator tRNA. The ribosome then catalyzes the formation of a peptide bond between the 2 amino acids. Surprisingly, peptide bond formation in the ribosome is catalyzed not by the protein components of the ribosome, but by a ribosomal RNA.

The formation of the first peptide bond results in the tRNA in the A site having 2 amino acids attached, while the tRNA in the P site has none. With this step, protein synthesis has begun, and 2 amino acids have been joined together, making a dipeptide.



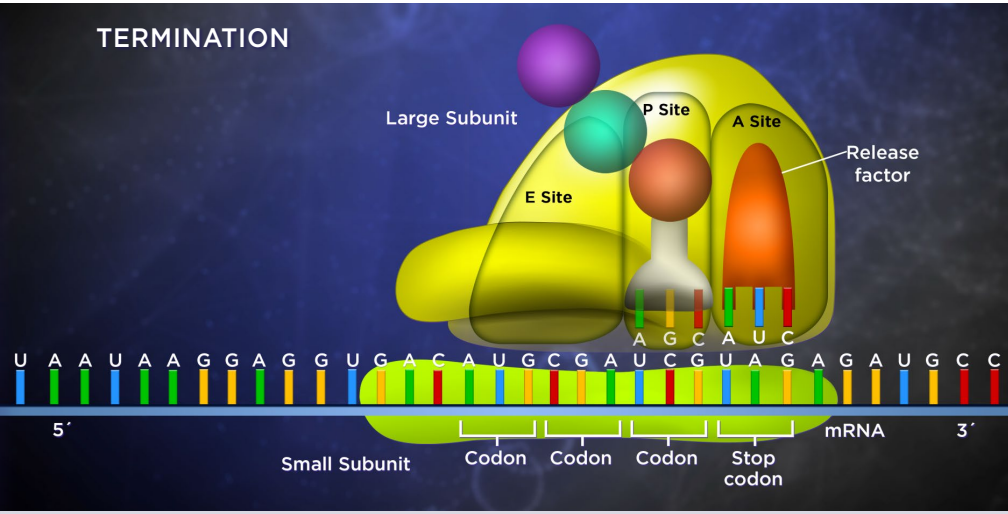
The next stage of protein synthesis, elongation, now begins. The ribosome moves along the mRNA so that the empty initiator tRNA is transferred to the E site, from which it can leave the ribosome. This movement requires the assistance of proteins called elongation factors as well as energy, once again, from GTP hydrolysis.



As the ribosome moves one codon forward, the tRNA carrying the dipeptide is moved into the vacated P site, leaving a new codon exposed in the A site, ready for the next charged tRNA to occupy.

The rest of the process is just a repeat of 3 steps: binding the proper tRNA in the A site, peptide bond formation, and sliding along the mRNA to move a new codon in the A site.

Eventually, a stop codon of the mRNA is reached in the A site, triggering the final stage, termination. No tRNA can base-pair with the stop codon. Instead, a protein called a release factor facilitates the separation of the completed protein from the ribosome. The ribosomal subunits then come apart for reuse when another mRNA is to be translated.



The speed with which translation occurs is impressive—about 15 to 20 amino acids per second. Because there are 3 bases per codon and one codon for each amino acid, the rate of movement down the mRNA by a ribosome is 45 to 60 nucleotides per second. Transcription occurs at about the same rate.

Protein Delivery

That's it for joining amino acids together. But functional proteins are more than strings of amino acids. The process of folding as well as covalent modifications may help them take on various 3-dimensional shapes. Covalent modifications can include the addition of phosphates or molecules like sugars or lipids before proteins can perform their functions.

In eukaryotic cells, where there are multiple cellular compartments, proteins must also be delivered to the correct location. Ribosomes are in the cytoplasm, so this is where all proteins start being synthesized. But histones, DNA, and RNA polymerases must be delivered to the nucleus, and enzymes of the citric acid cycle must be sent to the mitochondria.

Unless a protein is in the right spot, it is not of much use to the cell. To get a protein to where it's supposed to be, many have a built-in address.

For many proteins, a portion of their amino acid sequence, called a signal sequence, serves as an address label. Similar to how you need a postal worker to read address labels and deliver the mail to each house, there are carrier proteins that help deliver each addressed protein to its correct destination.

Membrane proteins are anchored within membranes with the help of stretches of hydrophobic amino acids within their sequences. Secreted proteins get packed up in membrane vesicles to be delivered outside the cell. These various mechanisms ensure that proteins get to where they are needed.

Most of the applications of our knowledge of protein synthesis, though, do not involve preventing it. Understanding the components needed for translation and how they work together has allowed researchers to develop novel therapies and technologies that promise to revolutionize the treatment of some diseases.

Because cells must make proteins to stay alive, interfering with protein synthesis is a great way to kill disease-causing organisms—or any cells you'd like to eliminate. Many antibiotic drugs work through targeting protein synthesis in bacteria. Subtle differences in the protein-synthesis machinery in our cells and those of bacteria allow us to kill off the bacteria without harming our own cells.

READINGS

Cobb, *Life's Greatest Secret*.
Jones, *The Genetic Code*.

QUESTIONS

1 The error rate of protein synthesis is considerably greater than that of DNA synthesis or transcription. Speculate why.

2 Given that inosine can pair with cytosine, adenine, and uracil, calculate for each amino acid how many different anticodons would be required.

CLICK HERE TO SEE THE ANSWERS.

32 PROTEIN-SYNTHESIS CONTROLS AND EPIGENETICS

Although all the cells in a person have the same DNA, each different cell type uses a different subset of the genes in that DNA to direct the synthesis of a distinctive set of RNAs and proteins. Cells can respond to change by increasing or decreasing the activity of particular proteins, but they can also respond by altering which proteins are made at all. In this case, cells are responding to change by regulating gene expression, which is a multistep process that provides many points at which cells can determine which genes are copied into RNA and translated to make proteins.

Prokaryotic Gene Expression

Bacteria are marvels of efficiency from a biochemical standpoint. Their genomes are compact. Genes are often clustered in convenient groups such that genes that need to be expressed at the same time are next to each other. This physical proximity allows them to be controlled as a single unit through a single promoter. Groups of genes operated on together like this are called operons.

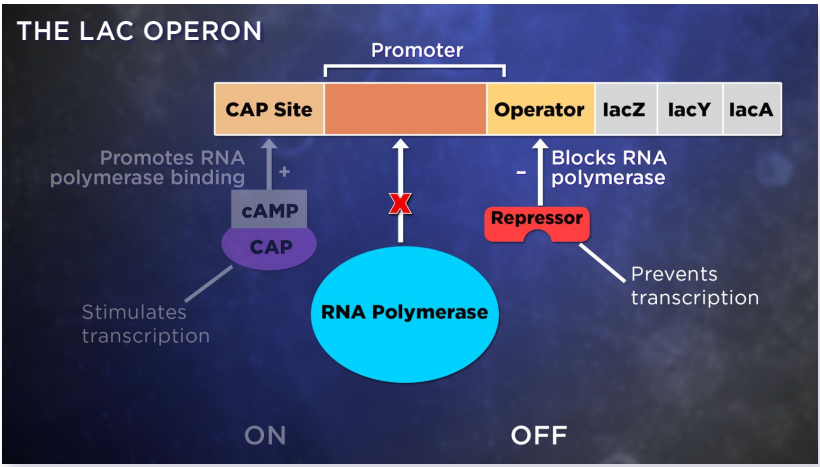
In 1965, the first operon was discovered in a group of bacterial genes called the lac operon, so named because they control and metabolize the sugar lactose. There are 3 genes in the lac operon under the control of just one lac promoter.

Bacterial gene expression is what drives the workings of our gut microbiome.

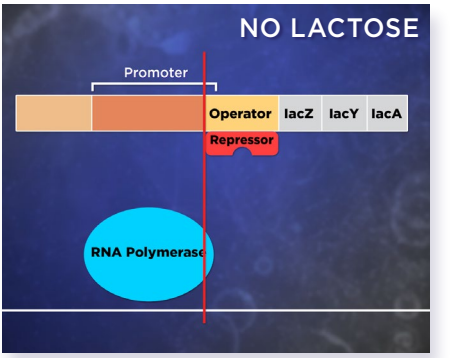
Lactose is the sugar in milk that some people are unable to digest. *E. coli* also avoids using lactose if its preferred sugar, glucose, is available. But if it runs out of glucose and finds lactose available, it switches on the lac operon so that it has the proteins needed to take up and use lactose.

In *E. coli* cells, managing the on-off switch for the lac genes is the work of 2 proteins: a repressor that prevents transcription when there's no lactose and an activator that stimulates transcription when glucose is low.

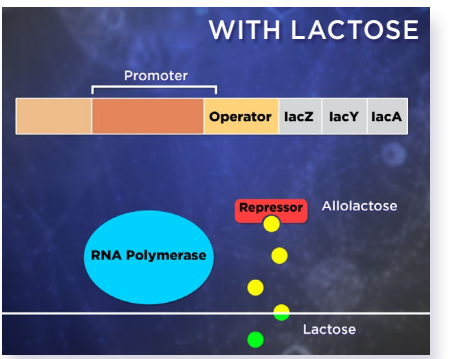
Because the bacteria don't normally use lactose, the default state for the lac operon is off. What keeps it off is the lac repressor protein, which is usually bound to a small region of the DNA, called the operator, right in front of the –10 sequence of the lac promoter.



The repressor is thus sitting squarely in the path of any RNA polymerase that might bind at the promoter. In fact, the repressor is hogging a bit of the space that RNA polymerase would occupy when it was bound at the promoter. The repressor prevents the polymerase from binding to its promoter, except if lactose is present.

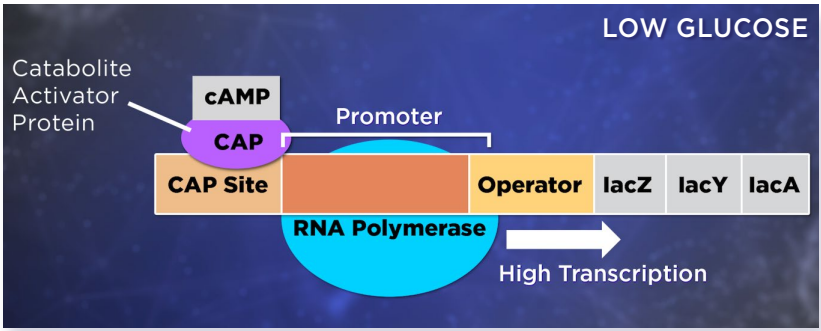


If lactose is present, a small amount enters the bacterial cell, where the lactose is converted into an isomer, allolactose. It is the allolactose, not lactose, that binds the repressor, loosening its grip on the operator. When this happens, the repressor comes off of the DNA. In biochemical lingo, the allolactose is called an inducer.



This means there isn't a barrier, but transcription still isn't going to happen unless RNA polymerase binds at the promoter to transcribe the genes.

It turns out that RNA polymerase can be lured to the lac promoter by a catabolite activator protein (CAP), which binds near the promoter. Upstream from the promoter on the DNA sequence is a spot called the CAP binding site. When glucose levels are low, CAP binds upstream of the promoter, and that recruits the RNA polymerase to come bind to the promoter and begin transcription of the lac genes. The proteins needed to use lactose are made; genes are turned on to make use of lactose.



When lactose levels drop, there's not much allolactose around. Without allolactose to ease its grip on DNA, the repressor goes back to sitting on the operator, and expression of the lac genes is shut off once again.

The binding of proteins to specific DNA sequences can increase or decrease transcription from a promoter.

Eukaryotic Gene Expression

The binding of proteins to specific DNA sequences is also one mechanism of regulation of transcription in eukaryotes, though it works a bit differently than in bacteria. In eukaryotes, genes commonly contain 2 groups of sequences that are important for transcription. First are promoters, whose job is to initiate transcription by assembling the basal transcription complex.

Additionally, in contrast to bacteria, eukaryotic cells also have regulatory sequences affecting transcription that can be thousands of base pairs away from the gene. Some of these are enhancers that work to increase transcription of a gene; others are silencers that reduce transcription.

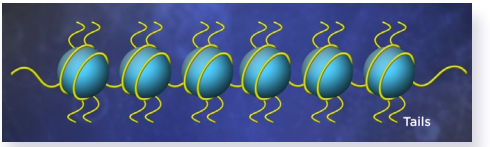
With these forms of remote control, first the sequences must be bound by specific proteins that either activate or repress transcription. This is a familiar theme, similar to how transcription in bacteria can increase or decrease. But for eukaryotes, the enhancer or silencer sequences are very far away on the DNA molecule from the gene they control.

What makes it possible for a regulatory sequence to influence what happens at a distant promoter? An enhancer or silencer, with its bound protein, can be brought close to the promoter of the gene being transcribed by looping of the DNA strand. In other words, when transcriptional activators bind at these distant sites, the DNA strand bends. This allows the proteins bound at enhancers to come close to the transcription initiation sites, where the activator proteins can interact with the basal transcription complex.

One effect of this interaction is to help in recruiting proteins necessary for transcription to the promoter. More of these proteins means the transcription initiation complex is made more efficiently and more frequently.

Another thing that transcriptional activators bound at enhancers can do is recruit enzymes to help open the chromatin structure. The chromosomes of eukaryotes have DNA packaged with proteins to form a complex called chromatin. This means that when a region of DNA has to be transcribed, it must be separated from the proteins temporarily to allow the RNA polymerase to transcribe it.

And how do the enzymes open the chromatin? Remember that DNA in our cells is wound around histone cores to make structures called nucleosomes. Histones, being proteins, are chains of amino acids that fold to take on 3-dimensional shapes. But one end of the amino acid chain making up each histone hangs out of the bead-like nucleosome. These tails are the targets of the enzymes, which add one or more small tag molecules onto specific amino acid side chains.



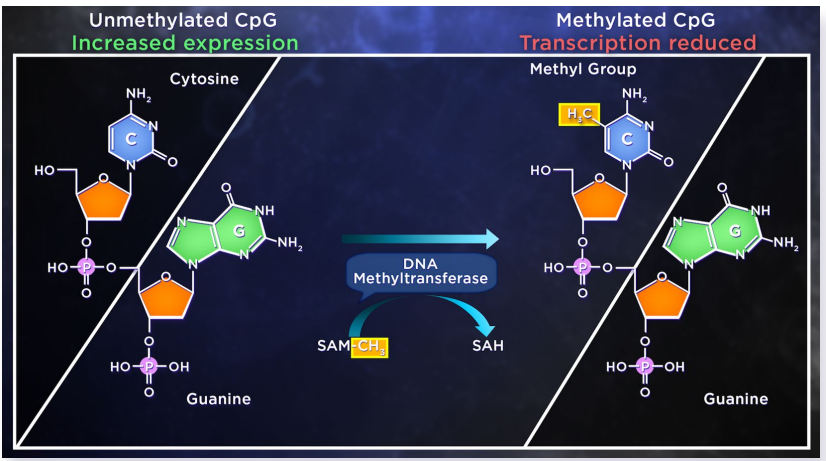
Depending on the particular combination of these chemical tags, the chromatin in that region will either be loosened up to allow transcription or packed tightly to silence the expression of genes in that part of the DNA.

Epigenetic Control

In addition to regulating transcription in eukaryotes by opening or tightly packing chromatin, there are enzymes that can modify the DNA itself, adding methyl groups at specific places.

In our DNA, for example, enzymes called DNA methyl transferases can add methyl groups onto cytosines that are next to guanines in a single strand—that is, where you would see a C followed by a G in the DNA. Not every C in a CG sequence is methylated, but methylation of DNA near transcription start sites is accompanied by silencing of the gene. Transcription is reduced.

By contrast, demethylation goes hand in hand with increased expression. We don't know for sure why, but it may be that methylation of the DNA at promoter regions prevents transcription factors from binding there.



Thus, although DNA is the information storage molecule, mechanisms like histone modification and DNA methylation control what cells do with the DNA information. This control over what cells do, which acts over and above the information in the DNA sequence, is called epigenetic control.

The trillions of cells in your body all have the same DNA, but they have differentiated into a variety of different cell types, each performing a specialized function.

For example, the overall locations of the chromosomes in relation to one another may vary by type of cell. Such variations in what are called chromosome territories may also play a role in how the genes for each type of cell are expressed.

Even more intriguing are the findings that patterns of epigenetic modification can be changed in response to environmental conditions. And there's some evidence that epigenetic changes due to environment may be inherited, at least for a few generations.

While it is clear that epigenetic modifications can change with an individual's experiences and exposure, researchers are working to determine how far-reaching the effects of these changes can be.

Gene Expression after Transcription

In addition to transcription, there are other steps in gene expression at which cells can regulate which proteins get made in cells.

Eukaryotic mRNAs must be spliced to remove introns, noncoding regions within the protein-coding information in the transcript, and join the exons to make a continuous coding sequence. But splicing does not always use all of the exons in a transcript.

Instead, different combinations of the exons may be joined together to make different mature mRNAs that will then direct the synthesis of different proteins. This is called alternative splicing, and it means that transcription of the same gene could result in the production of different proteins in different cells or even in the same cell at different times.

But there are even more possibilities for regulating and altering the expression of proteins. And strangely enough, one major way that the expression of a gene may be regulated is by a group of RNAs called small regulatory RNAs.

The number of available mRNAs—the bearers of information for making proteins—and whether they are translated is regulated by a whole other set of RNAs, known as regulatory RNAs.

Until the late 1990s, biochemists knew nothing about regulatory RNAs or the role they play in gene expression. More recently, regulatory RNAs have been explored as possible treatments for everything from cancer to viral infections.

The 2 types of regulatory RNAs that the most is known about so far are microRNAs and short interfering RNAs. Both are small noncoding RNAs that silence genes by base-pairing with target mRNAs and marking them for degradation or by blocking their translation. In either case, the protein encoded by the mRNA will not be made. This sort of regulation is called RNA interference.

Laboratory studies have shown that RNA interference can successfully be used against viruses like HIV and HPV and to target specific cancer-related genes in melanoma, leukemia, and pancreatic carcinomas.

Cells have clever ways to regulate gene expression at the point of translation, too. One such mechanism regulates the production of 2 proteins we need for the uptake and storage of iron in our cells: transferrin receptor and ferritin.

The transferrin receptor is a membrane protein that's involved in importing iron atoms into cells, whereas ferritin packs away iron atoms into a ball-like structure within cells so that they don't create dangerous free radicals.

When levels of iron in a cell are low, the cell needs to take up more iron and needs more of the transferrin receptor. When iron is abundant in the cell, more ferritin is needed to safely store the iron.

Interestingly, the expression of both these proteins is regulated by a remote sequence in the untranslated regions (UTRs) of their mRNAs. This regulating sequence, called the iron response element (IRE), folds on itself to form a hairpin structure and can be bound by a binding protein. The IRE and its binding protein respond to the levels of iron and control the production of transferrin receptor and ferritin.

READINGS

Arney, *Herding Hemingway's Cats*.
Carey, *The Epigenetics Revolution*.
Latchman, *Gene Control*.

QUESTIONS

1 Given what you learned previously about how closely a given promoter matches the consensus sequence for a bacterial promoter, predict how closely the lac promoter matches the consensus sequence and explain your reasoning.

2 Given that it requires a lot of energy to make mRNAs, speculate on why cells have RNA interference systems to break them down.

CLICK HERE TO SEE THE ANSWERS.

33

HUMAN GENETIC DISEASE AND GENE THERAPY

Genetic diseases arise from mutations in our DNA. Sometimes, they affect a single gene, as in sickle cell anemia. Other times, a combination of mutations in different genes may be the cause, as in familial tendencies to develop heart disease or mood disorders. In other cases, because mitochondrial DNA is different from the nuclear genome, genetic diseases also arise from mutations in mitochondrial DNA. In each case, changes in the DNA result in the production of an altered protein that misbehaves in some way to cause the disease.

- Humans have an estimated 20,000 genes, and more than 10,000 gene-specific disorders are known.
- Roughly 10,000 human diseases are thought to be caused by a single gene.
- The World Health Organization estimates that genetic disorders, taken together, affect one person in 100 across the globe.

Gene Mutations

Point mutations change a single base. Most of the amino acids are coded for by more than one codon, and these often only differ in the third base. Mutating the third base of a codon often results in an alternative codon for the same amino acid, so there's no effect on the protein.

Other changes, however, likely will alter the amino acid that's specified. Such replacements of one amino acid for another in a protein are called missense mutations. They have the potential to change the protein's activity.

Another significant mutation occurs when a point mutation in a gene changes a codon from one that specifies an amino acid to one of the 3 stop codons. Ribosomes making protein stop when they encounter such a stop codon, resulting in a shorter, partial protein.

Other mutations arise from deletions or insertions of bases within a coding region, changing how the sequence is read by the protein-synthesis machinery.

Hemophilia

Hemophilia is a sex-linked bleeding disorder that affects males disproportionately. Some people with hemophilia bleed internally even without any injury.

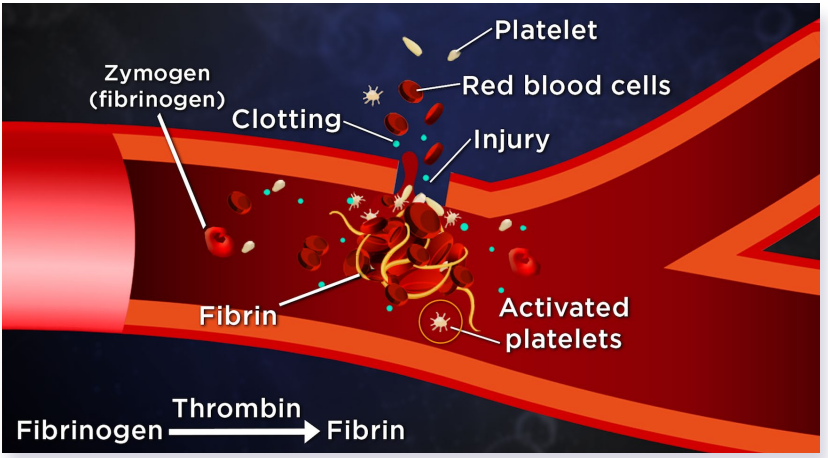
Queen Victoria of England famously passed hemophilia to one son and through 2 daughters to no fewer than 9 additional male descendants.

Hemophilia comes mostly in 2 main types—A and B—both of which have the same symptoms. Hemophilia A is more common and results from mutations in the gene for a clotting factor called factor VIII. Hemophilia B is caused by mutations in the gene for a different clotting factor, factor IX.

In both forms, hemophilia is X-linked, meaning that the genes for these clotting factors are found on the X chromosome. This is why males, who have only one X chromosome, are disproportionately affected: They only have one copy of the gene.

Blood clotting is a process that involves the sequential activation of a series of protein clotting factors. There are more than a dozen such factors present in the blood as inactive precursors, or zymogens. The blood clotting sequence sets off a series of reactions in which the inactive zymogens are activated, one after the other, until the zymogen, prothrombin, is converted to thrombin.

Thrombin, which is a protease, cuts up a zymogen called fibrinogen, which is also dissolved in the blood, to make fragments called fibrin. Fibrin fragments then assemble with each other to form a mesh that adheres to the wound and helps form a clot, along with small blood cells called platelets. Not having one of the clotting factors interferes with this sequence of events.



One of the first treatments attempted for hemophilia was to give patients a transfusion of plasma from healthy donors. Unfortunately, this sometimes triggered a response against the infused proteins, which were recognized as foreign, resulting in outcomes as bad as, if not worse than, the ailment itself.

Treatments improved dramatically when recombinant clotting factors were made in labs by inserting the relevant gene into cells grown in culture. Clotting factors can be produced by inducing cells grown in the laboratory to express the genes encoding those proteins. For hemophilia, the clotting factor is purified from these cells and provided to those who need it.

Powerful new molecular biological techniques provide hope for gene therapy to permanently fix some genetic diseases. One such approach is to introduce a good copy of the clotting factor gene into a harmless virus that has been stripped of all its own genes and inject it into the patient's muscle. There, the virus would ferry the gene into the cells, which would express it and secrete the clotting factor into the blood. This technique has been successfully demonstrated in laboratory animals.

Another approach is to obtain stem cells from the patient, introduce the clotting factor gene into the stem cells in a form that they can express, and then return the expressing cells to the patient's body to do their work. This, too, has successfully been accomplished in animal models.

Cystic Fibrosis

People with cystic fibrosis (CF) suffer from a buildup of thick mucus in various parts of the body. Mucus is normally produced by the body to protect and lubricate airways and the lining of the digestive system. But in cystic fibrosis, the mucus becomes so thick that cysts and fibrous material form in the body, giving the disease its name.

This very thick mucus clogs the lungs. It makes breathing difficult and also traps bacteria, raising the risk of respiratory infections and scarring. The mucus can also block the release of pancreatic digestive enzymes and insulin, causing digestive problems and even diabetes.

Cystic fibrosis is the most common inherited disorder in the US among people of European descent—who have about a one in 25 chance of carrying a copy of a mutant *CFTR* gene.

As early as the 1600s, physicians knew that babies whose sweat was salty would not thrive and soon would die. The saltiness was the clue that led to the discovery in 1989 of the gene responsible for the problem. It encoded a membrane transport channel for chloride ions called cystic fibrosis transmembrane regulator (CFTR).

Chloride channels control the salt and water balance of cells, and when disrupted, the disease results. Comparing the gene for the CFTR regulator between healthy individuals and CF patients, researchers found a deletion of one codon! The deleted codon specifies the amino acid phenylalanine at position 508 of the protein.

The name of this mutation is delta F508. Other mutations have also been identified in the *CFTR* gene, but the deletion at position 508 is found in more than 70% of cases worldwide.

The molecular defect that this deletion causes is aberrant protein folding. Normally, the CFTR protein is folded in the endoplasmic reticulum of the cell and is sent to the cell membrane, where it functions. The delta F508 CFTR protein, which doesn't fold properly, is instead marked for destruction. The misfolded protein is released into the cytoplasm, where it is degraded by the proteasomes.

As a result, the mutated chloride channel proteins never make it to the cell membrane. Without the chloride channels in the cell membrane, cells cannot maintain the proper salt and water balance. This leads to the excessive salt in the skin and the thick mucus that leads to the problems of cystic fibrosis.

Conventional treatments for cystic fibrosis have focused on preventing lung infections and inflammation as well as ensuring proper nutrition. Screening and conventional treatment by themselves have increased life expectancy significantly and have provided modest improvements in the quality of life.

Gene therapy is the new frontier for cystic fibrosis. As for other genetic diseases, it offers the possibility of a cure, rather than treatments to manage symptoms. In 2015, researchers in the UK carried out the first gene therapy trial and had some success.

Another approach involves gene editing using stem cells from the person with the disease. These stem cells are grown in the laboratory, and the mutation in the *CFTR* gene is corrected using a technique such as CRISPR. Then, the corrected cells are introduced back into the patient's lungs, where the stem cells could give rise to new lung cells that are free of the mutation. This method has the potential to solve the problem at its source.

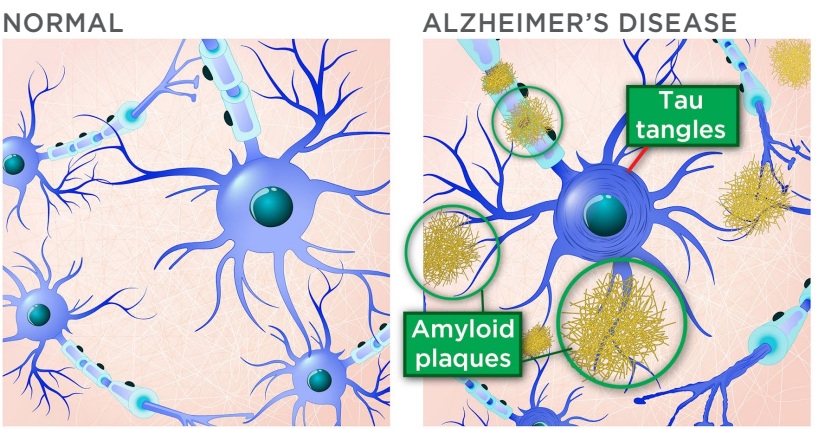
Alzheimer's Disease

In 2018, Alzheimer's disease affected 5.7 million people in the US. Of these, 5.5 million had the more common late-onset Alzheimer's, which is seen in people past the age of 65, while the remaining 200,000 had the early-onset form, which strikes before then. Early-onset Alzheimer's is associated with mutations in genes for making proteins called presenilin 1, presenilin 2, and amyloid precursor protein.

The risk of developing late-onset Alzheimer's is higher in people who have a particular variant of the gene for apolipoprotein E (ApoE), which has 3 common forms: E2, E3, and E4. The E4 variant is associated with increased risk for late-onset Alzheimer's, though not everyone with this variant will develop the condition. The normal function of the protein encoded by all the variants of the *APOE* gene is in the transport of fats and cholesterol in the body.

Why does a mutation in *APOE* increase the risk of Alzheimer's? Two characteristic features that are seen in Alzheimer's brains are clumps of proteins called amyloid plaques outside the neurons and tangles of a protein called tau inside the neurons. Both the plaques outside and tangles inside seem to kill neurons.

In 2018, a team of researchers from MIT helped connect the mutation in *APOE* with the changes seen in the brain. In the brain, glial cells make lipids and cholesterol and deliver them to neurons with the help of ApoE. Glial cells also play an important role in clearing away waste and dead neurons.



To find out what was different in brain cells with ApoE4 compared to those with ApoE3, scientists grew both kinds of cells in culture and found that neurons with ApoE4 made and secreted much higher levels of the amyloid protein that forms the plaques. They also seemed to phosphorylate the tau protein more, which favors forming the tangles of tau within the neurons. Glial cells with ApoE4 made twice as much cholesterol as those with the *APOE3* gene and were not very efficient in clearing away plaques and cell debris.

So, is there some way to get ApoE4 to behave like ApoE3? Scientists have designed a small molecule that could be provided as a drug that travels to neurons and blocks the problems of the mutated *APOE* gene. While there are technical hurdles to overcome to ensure that the right amount of the molecule can cross the blood-brain barrier and enter into the neurons, such experiments are the first to demonstrate that a drug could persuade cells that express ApoE4 to behave normally instead.

How to Reduce the Risk of Alzheimer's Disease

The advice for how to reduce the risk of developing Alzheimer's disease is right in line with the recommendations for staying healthy in general:

- ◇ Get enough sleep.
- ◇ Floss.
- ◇ Control blood sugar.
- ◇ Exercise.

Leber Hereditary Optic Neuropathy

You get all of your mitochondrial DNA from your mother, which was passed down from her mother, etc.—all the way back up the maternal lineage, with the only changes due to mutation.

Most people don't think of mitochondria in connection with genetic disease, but mitochondria have their own genomes. In humans, it's a double helix of 16,589 base pairs.

This is tiny compared to the nuclear genome. Even inside the mitochondria, most of the proteins found there are coded in the nuclear genome and get imported into the mitochondria from the cytoplasm. Mitochondrial DNA mostly codes for proteins for the electron transport chain, which is the mitochondrial system for generating ATP.

Although mitochondrial DNA is tiny, most of it does get transcribed into mitochondrial RNA, unlike the merely 1% of nuclear DNA that gets turned into proteins. This means that any mutations in mitochondrial DNA are more likely to become the cause of some serious diseases.

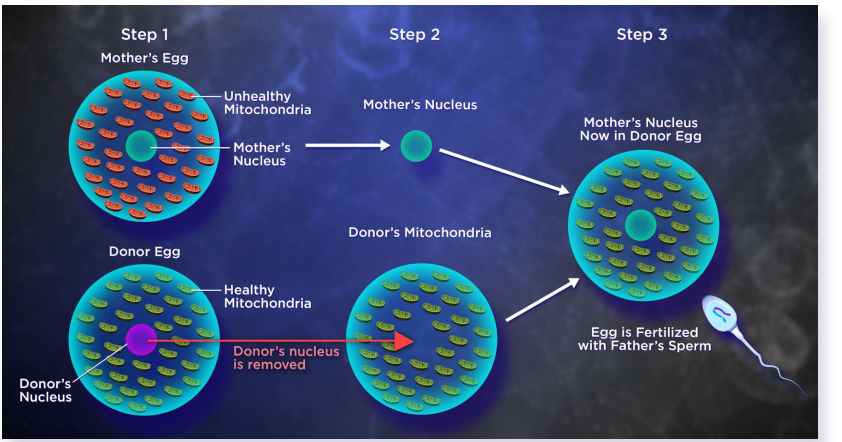
One disease of the mitochondrial DNA is a form of vision loss that begins in young adulthood called Leber hereditary optic neuropathy. In some patients, there are other symptoms, including dysregulation of the signals that control heartbeat, muscle weakness, and tremors.

The culprit is a mutation in one of the mitochondrial genes that encodes the enzyme NADH dehydrogenase, or complex 1, which is needed to pass electrons from NADH to the electron transport chain in mitochondria—without which cells cannot make ATP through oxidative phosphorylation.

In the past, treatment for mitochondrial genetic disorders was limited to providing support to the patients to help them deal with blindness and other impairments that came with this disease. But 2 developments have provided some hope.

The first is that researchers have successfully used DNA-editing enzymes to reduce the number of mitochondria with defective DNA in live mice. In fact, the researchers were able to reduce the number of mitochondria with the mutation by ½, which would be enough to prevent symptoms of disease.

There is also the option of preventing mitochondrial disease altogether. To avoid passing on mitochondrial disease from a mother, scientists have successfully used a 3-parent technique to produce a healthy baby. In this method, a donor egg from a woman with healthy mitochondria had the nucleus removed and replaced with the nucleus from the mother's egg. This egg, with the mother's nuclear genome and mitochondria from the donor, was then fertilized with a sperm from the father. The resulting baby was free of the mother's defective mitochondria and all the attendant health issues, while still having the full contribution of the birth mother's nuclear genome.



READINGS

Gillham, *Genes, Chromosomes and Disease*.
Kolata, *Mercies in Disguise*.
Leroi, *Mutants*.
Samuelsson, *The Human Genome in Health and Disease*.

QUESTIONS

1 The CFTR mutation is a 3-base deletion of a single codon that leads to the improper folding of the protein, though the protein still functions. Speculate on what might happen to the protein if only a 1-base deletion were made. Would the resulting protein fold properly and/or function? Justify your answer.

2 In the mitochondrial disease described in the lecture, only some of the mitochondria in the person's cells had the problem. Given that cells can have upward of 1000 mitochondria, speculate how this disease might arise and be transmitted genetically, and consider whether or not you would expect that it would eventually affect all the mitochondria in a person's cells.

CLICK HERE TO SEE THE ANSWERS.

34 CANCER MECHANISMS AND TREATMENTS

Cancer cells proliferate rapidly, ignore biochemical signals to stop reproducing, and ignore signals telling them to die. They mutate genes and rearrange chromosomes. They set up their own blood flow networks, invade nearby tissues, and metastasize. Instead of maintaining connections with neighbors, they spread throughout the body. In short, cancer undermines many forms of cell cooperation we take for granted in the body.

Unregulated Cell Division

There are many different kinds of cancer, but one thing all types of cancer have in common is unregulated cell division. That contrasts with regulated cell division, which converted us from a single-cell fertilized egg to adulthood. Regulated cell division also helps replace worn-out cells and is essential to healing from injury.

Cells monitor signals that tell them to divide. They don't proceed unless they get the green light from a relevant growth factor, such as epidermal growth factor or nerve growth factor.

Growth factors bind to receptors on the cell membrane. The directive for the cell to divide passes through multiple signaling steps that make their way into the nucleus. These signals travel complicated pathways, but the result is to set in motion the process leading to cell division.

Along the path to cell division, cells check their DNA for damage before they split in 2. If there is damage, they stop and fix it. If there is too much damage, the cell undergoes programmed cell death. Special parts of the cycle called checkpoints ensure that each step in the cell cycle is completed properly before proceeding with the next.

On top of these considerations, cells have a natural limit on the number of times they can divide. Our chromosomes shrink with every DNA replication, and after a certain number of replications, they stop dividing and die.

With these safeguards, how do cells still manage to get off track?

All cancers start with a mutation in a single cell. This mutant cell begins to divide more than it normally would. It may initially produce a small abnormal clump of cells called a tumor, but that in itself does not result in cancer. Tumors may be benign, or harmless, at first. If additional mutations occur, cells may proliferate faster. And the more they divide, the greater the chances for mutations are when their DNA is copied. And the more mutations there are, the greater the dysregulation of cellular controls becomes.

As more control is lost, cancerous cells spread outside of their home tissue and are said to be malignant. They can then migrate via blood or lymph to distant sites in the body, where they form secondary tumors, a process called metastasis.

Usually, it takes many mutations to make a normal cell cancerous. This is fortunate, because mutations occur in our DNA every time cells replicate. Although DNA polymerase is very accurate, mammalian cells have several billion base pairs of DNA to copy, and mistakes cannot be avoided. The more times cells undergo cell division (and DNA replication), the higher the chances are that they will have accumulated enough mutations to become cancerous. Consequently, cancer risk grows with age.

Though there are mechanisms for fixing replication errors, a few escaping each time can add up over the years.

In many cancer cells, telomerase, the enzyme that extends chromosome ends, is reactivated. While telomerase is normally inactive in most adult somatic cells, it's switched on in cancer cells. And this means that cancer cells go on dividing, because their chromosomes never shrink enough to wind down activities.

Why aren't there better repair mechanisms? By the standards of Mother Nature, once you've lived long enough to have a chance to pass on your genes, you've served your biological purpose. So, cancer is, in some ways, a side effect of planned obsolescence.

But obviously not everyone who lives a long life gets cancer. This is because the development of cancer depends on factors that are beyond the natural pileup of mutations. Genetics matters—as do environmental factors, including exposure to chemical carcinogens, radiation, and even certain viral infections, such as hepatitis B and C. Lifestyle factors play roles, too, such as smoking, excessive alcohol use, diet, lack of exercise, and too much Sun exposure.

Carcinogens cause damage or mutations in DNA. Any mutations that interfere with cells' ability to correct DNA damage make it more likely that other mutations will accumulate. Control of cell division can be lost due to defects in signaling pathways. Mutations can also prevent cell cycle checkpoints from functioning normally. Genes that normally suppress tumors can be inactivated.

Mutations in 2 major classes of genes are crucial in the development of cancer. They are grouped based on their roles in controlling cell division.

1. Proto-oncogenes are genes that encode proteins that promote cell division. These genes code for normal cell division but through mutation can become oncogenes—genes responsible for cancer.

2. Tumor suppressors normally encode proteins that can halt cell division, if necessary, to prevent cancerous growth.

Together, these 2 types act as green lights or red lights, respectively, for cell division. Too much green lighting or too little of the red lights and cells will divide too much.

Proto-Oncogenes

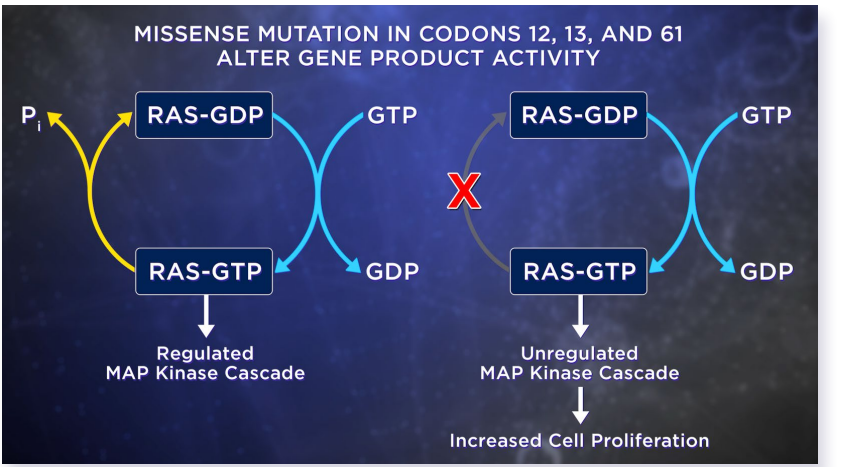
More than one kind of gene encodes green-light proteins for cell division. The green lights include

- genes for growth factors that signal cells to divide,
- genes for receptors of the growth factors, and
- genes for signaling proteins that pass on the directive to divide.

Any of these genes that encode green-light proteins can become a proto-oncogene.

Many cancers have mutations in the gene for *ras*, which encodes ras proteins that pass on signals telling cells to divide. The most common of the *ras* mutations are in positions that change amino acids 12, 13, or 61 in a ras protein. Each of these amino acids is in the GTP-binding region of the protein, and the mutations impair ras's ability to remove a phosphate from the GTP to become GDP.

If ras can't break down the GTP, ras gets stuck in the active form of growth, even in the absence of external growth signals. So, if the green lights for cell division are on when they should be off, this results in uncontrolled cell proliferation.



Different from ordinary traffic signals, each cell has 2 copies of each gene, so the green light can be thought of as being lit by 2 bulbs. When a green light is supposed to be off but one of the 2 bulbs is on, that still looks like a go signal, even if it's not as bright. So, if even one copy of a proto-oncogene is active when it's supposed to be off, cells will divide when they shouldn't.

Tumor Suppressors

Tumor suppressors are genes that encode cellular red lights that stop cell division. A good example is *p53*, whose job is to stop the cell cycle, providing time for the cell to fix DNA damage. If the damage is too extensive to fix, *p53* triggers programmed cell death.

Mutations inactivating *p53* remove the brakes on cell proliferation. In this case, a red-light signal that should prevent cell division is off when it should be on.

More than 1/2 of all cancers display mutations in the gene encoding the protein p53.

Given the red-light function of *p53*, the 2 copies of it in the cell mean that it's a red light with 2 bulbs. If one of the bulbs burns out, but the other is fine, it's still a red light. The red light won't be off unless both bulbs burn out. In this case, both copies of a tumor suppressor gene must be knocked out for activity to be lost.

The other job of *p53* is triggering apoptosis, or programmed cell death, if the damage is severe. But if the *p53* in a cell is inactivated, then that cell could continue copying badly damaged DNA and proliferate, instead of either fixing the DNA or undergoing apoptosis. The continued division of cells with damaged or mutated DNA favors the progression of cancer.

Mutations that inactivate *p53* also complicate cancer treatment. Radiation and some forms of chemotherapy, for example, work by causing damage to the cancer cells' DNA—in hopes of triggering apoptosis. But if cancer cells have a fully defective *p53*, they may not undergo apoptosis. The result can be a lot of cells that will not die.

Why do elephants so rarely get cancer, despite living as long as humans do?

Where humans have just one pair of *p53* genes, elephants have a whopping 20 pairs! They also have extra copies of another tumor suppressor gene called *LIF6*. Between them, the proteins encoded by these genes make it more likely that cells with DNA damage self-destruct. By eliminating cells with DNA damage on a regular basis, most elephants can apparently dodge bullets that would have become cancer.

Tumor suppressor genes help answer the question of why cancer sometimes seems to run in families. It's a matter of how those genes are mutated. Remember, both copies of genes that encode tumor suppressors, like *p53*, must be mutated for there to be no suppression of tumors. The probability of a mutation in both copies of the exact same gene is low. For most people, it would take the accumulation of a whole lot of mutations before both copies were inactivated.

But if a parent or grandparent had already acquired a mutation in a tumor suppressor gene that he or she passed on to you, then you start life with only one good copy in each cell of your body. This doesn't mean you are doomed; you still have one good copy. But this means that your cells are much, much more likely to develop 2 mutated *p53* genes than someone who started with 2 good copies.

Most cancers don't arise from just one mutation, but rather, from a combination of them. As the mutations accumulate, cell division becomes increasingly dysregulated.

Approaches to Treating Cancer

One way to treat cancer is to use agents that target rapidly dividing cells. Dividing cells need large amounts of DNA nucleotides to replicate their DNA, so drugs that interfere with nucleotide biosynthesis can starve these cells of the building blocks they need.

The downside of such treatments is that cancer cells are not the only ones in your body that are dividing. For example, your hair and nails grow all the time. Invisible to you, but also dividing regularly, are the cells that make up your intestinal lining and your skin, among others. All of these cells are affected by treatments that target rapidly dividing cells, and it is this inhibition of cell division that explains why chemotherapy can cause sudden hair loss. Because there's no way for current drugs to tell healthy cells from sick ones, there's no way to avoid collateral damage.

Devising ways to target cancer cells specifically is the holy grail of cancer drug researchers. Such treatments would rely on the differences between cancer cells and normal cells.

One of those differences is that cancer cells turn on telomerase so that they can keep dividing. So, to stymie them, we could inhibit telomerase. This would not be a problem for most of the rest of our cells, but it would hit cancer cells. Clinical trials of telomerase inhibitors show that this may be a good option.

Another interesting and promising approach—which won the 2018 Nobel Prize in Medicine and has received FDA approval for some very specific treatments—is called immune checkpoint therapy, which relies on harnessing the patient's own immune system to attack and clear cancer cells.

T cells are types of immune cells that do not normally attack a patient's own cells. Some cancer cells get eliminated, but others dodge that system. This is possible because the T cells have on their cell surface a few blocking proteins that must be disabled before the T cells will attack the cancer cells.

When you have something powerful that can kill, it makes sense to keep it from doing unintentional damage to your own cells. To allow the T cells to attack the cancer cells, the blocking proteins must be disarmed.

Targeted therapies have been developed for specific kinds of cancers, such as chronic myeloid leukemia and some breast cancers, based on the understanding of the molecular events involved.

Immune checkpoint therapy uses antibodies to disarm the blocking proteins, freeing the T cells to do their work. And once the blocking proteins are disarmed, they can no longer prevent the T cells from going to town on the cancer cells.

There's still a lot to be worked out to ensure that normal cells are not also attacked by the T cells once the blocking proteins are disengaged. That could trigger unwanted autoimmunity, similar to lupus or psoriasis. But clinical trials have shown that immune checkpoint therapies are a game changer for the treatment of several different kinds of cancer, including lymphoma, melanoma, and lung and kidney cancers. The method is also being tested for other kinds of cancers and may completely alter the way physicians approach cancer treatment.

READINGS

Armstrong, p53, *the Gene That Cracked the Cancer Code*.
Graeber, *The Breakthrough*.
Mukherjee, *The Emperor of All Maladies*.

QUESTIONS

1 Imagine that you could design a single RNA interference molecule to inactivate a gene discussed in this lecture to increase the chance of it becoming a tumor. What would you target?

2 Imagine that you could design a single RNA interference molecule to inactivate a gene discussed in this lecture to reduce the chances of a cell becoming a tumor. What would you target?

CLICK HERE TO SEE THE ANSWERS.

35 BIOTECHNOLOGY, STEM CELLS, SYNTHETIC BIOLOGY

The term *biotechnology* loosely refers to any combination of biology and technology, but modern biotechnology usually refers to the more specific application of biomolecular knowledge about the underpinnings of life. Molecular biology has given scientists and engineers the ability to manipulate the recipes written in our genes in ways that could not even have been imagined in 1953, when the structure of DNA was determined.

Recombinant DNA

Biotechnology in its current form began in the early 1970s when scientists figured out that DNA from organisms as diverse as elephants and amoebas have the same structure and use the same genetic code. As a result, one can insert an elephant gene into amoeba DNA and, seamlessly, the cut-and-pasted gene becomes a part of the DNA it is inserted into.

DNA molecules can be cut using highly specific enzymes and joined to other DNAs using DNA ligase. Because this combines DNA from 2 different sources, it is said to be recombinant DNA.

And if the elephant gene is inserted so that it is near a promoter that can regulate its expression, the gene can be transcribed, processed, and expressed in the amoeba, just like the amoeba's own genes. It's basically using the amoeba's RNA and protein-synthesis machinery to build a protein to elephant specifications. The protein made in this way is called a recombinant protein.

Researchers developed this method because it's useful to be able to produce proteins that are otherwise hard to come by.

Human proteins, for example, can be made in yeast or bacteria, which serve as factories for building proteins according to the instructions provided by the human genes inserted into these cells. Why not get the proteins from the cells that actually make them? Well, it can be hard to lay your hands on a few kilograms of human brains, for example, to study a specific brain protein. Making recombinant protein gets around this problem.

The same method also allows us to make proteins for medical treatments. Dozens of medical-grade proteins have been made using recombinant DNA technology, including human growth hormone, clotting factors, and anticoagulant proteins. But where the initial techniques used bacteria or other cells grown in culture, some of these proteins are now made in animals.

The first recombinant protein made for therapeutic purposes was insulin, which was approved for use in 1982.

Recombinant DNA technology revolutionized biomedical research. It also made it possible to improve important crop plants. One of the biggest success stories in agriculture was the production of a virus-resistant papaya that saved the multimillion-dollar papaya industry in Hawaii.

Stem Cells

There are 2 things that are special about stem cells: They are capable of dividing indefinitely (that is, as long as the organism is alive), and they are undifferentiated.

When a stem cell divides, the new cells may differentiate, or they may remain as part of the stock of stem cells.

In the earliest stages of embryonic development, the fertilized egg must divide to produce a large number of unspecialized cells before those cells are assigned a specific function. These embryonic stem cells undergo specialization over time, turning into all the different cell types we have in our bodies, depending on the signals they receive.

This means that if we know what signals to provide, we could use stem cells to create new cells to replace cells that are injured or lost due to disease or trauma. For example, they could be used to replace neurons in people with spinal cord injuries or provide new heart cells to people who have suffered damage to their heart muscles following a heart attack.

Because of ethical concerns surrounding the use of cells obtained from a human embryo, alternative strategies have been explored.

One solution is the production of induced pluripotent stem cells (iPSCs). Pluripotent refers to the ability of a stem cell to give rise to many other differentiated cell types. To do this, yet avoid working with cells from a human embryo, scientists begin with differentiated somatic cells—that is, cells that are not reproductive cells (such as skin cells).

Scientists must then reverse the differentiated state of the cells in the laboratory to return them to the condition in which they are once again capable of becoming any type of cell. By comparing embryonic stem cells to differentiated cells, researchers were able to identify a handful of genes that were expressed at high levels only in embryonic stem cells.

When scientists increased the expression of these genes in differentiated cells, the cells reverted to a state similar to that of the undifferentiated embryonic stem cells; that is, once created, these iPSCs can, if given the right signals, differentiate into a variety of different tissues. iPSCs have been used in labs to produce beating heart cells, motor neurons, light-sensing photoreceptor cells, insulin-producing pancreatic cells, and even egg and sperm precursor cells.

In 2018, clinical trials with humans were begun using iPSCs to treat Parkinson's disease, heart disease, and macular degeneration.

Therapeutic Cloning

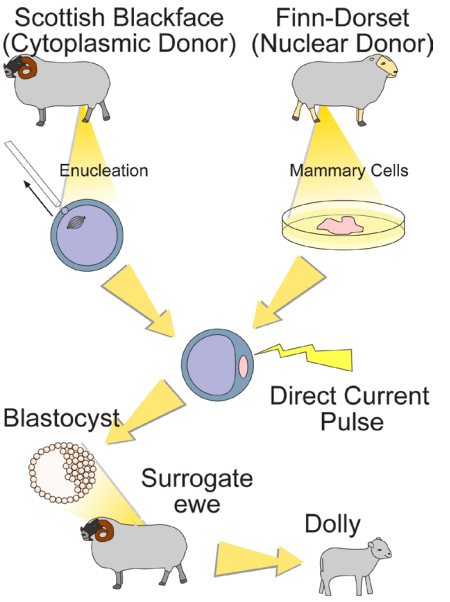
In a related technique, somatic cell nuclear transfer, sometimes called therapeutic cloning, the aim is not just producing undifferentiated stem cells, but an entire organism. As the name suggests, this method uses the nucleus from a somatic cell.

The nucleus of the somatic cell is extracted for placement inside an unfertilized egg, which has had its own nucleus removed. The inserted nucleus then provides the genetic material for the unfertilized egg, known as an oocyte, which is induced to develop as if it were a fertilized egg.

The idea here is that the somatic cell genes in the oocyte environment will be activated or inactivated as appropriate for a fertilized egg. This will restore the implanted genome to behaving once again like an embryo.

Somatic cell nuclear transfer is especially useful in making donor-matched stem cells that can be transplanted back into the patient. Because the DNA in these cells is a therapeutic clone from the patient, the cells would not be seen as foreign by the patient's immune system.

Somatic cell nuclear transfer has been used to make entire animal clones. This was the technology that was used to make the famous cloned sheep, Dolly, in 1996. To make Dolly, scientists used an oocyte from one sheep and placed into it the nucleus from a mammary gland cell of another sheep. The resulting embryo was implanted back into an ewe and allowed to develop.



Cloning in the plant world is commonplace. From common potatoes to the Cavendish banana found in supermarkets all over the world, plant cloning is easy because plants can be completely regrown from a tiny amount of tissue. With somatic cell nuclear transfer, cloning of animals is now possible, too.

The practicality of any cloning project is colored by a success rate of somatic cell nuclear transfer that is very low; in some areas, only 2% succeed, and even a very high success rate might be only about 40%.

Why would someone clone an animal, when animals are quite capable of producing offspring?

One reason is to create an exact replica of a particular animal—perhaps a prize bull or a superior racehorse.

Cloned animals can also reproduce, just like their conventional counterparts.

But some people are willing to pay big money to clone their pets. In 2005, scientists in South Korea successfully cloned a dog they named Snuppy, short for Seoul National University Puppy. Though costs are declining over time, the biggest factors are the technology required and the low success rates. Until these factors change significantly, pet cloning will tend to be limited to the wealthy.

Animal cloning might help save endangered species. Scientists have taken nuclei from somatic cells of an endangered animal and inserted that into an oocyte of a closely related nonendangered species. An endangered wild cow, for example, was cloned successfully this way—and displayed in the San Diego Zoo.

There is also the possibility of de-extinction of animals that have already died out. Now dinosaurs died out 65 million years ago, and it's unlikely we can recover stable DNA that's more than thousands, or possibly hundreds of thousands, of years old. But more recently extinct animals, such as woolly mammoths, are excellent possibilities. Mammoth remains have been found frozen and excellently preserved, and an elephant surrogate mother might be successfully used to bring a woolly mammoth calf into the world.

Designer Genomes

Let's say that instead of moving a single gene, we want the output of an entire metabolic pathway. DNA and the genetic code happily supply us with a single protein, but when the desired output is not a single protein, we are likely to need an entire metabolic pathway of enzymes to make molecules so that we can collect the molecules.

How do we copy all of the enzymes of an entire pathway? Computer algorithms can analyze genome sequences and help us figure out what enzymes a particular organism makes and what additional ones would be needed for it to carry out a particular metabolic pathway.

For example, scientists have engineered yeast cells to help synthesize an antimalarial drug called artemisinin. Because the amount of artemisinin made in plants is small and producing it in the laboratory is too complicated and too expensive, researchers decided to use cells, which are already capable of carrying out complex chemical transformations and can handle a new reaction as long as they have the necessary enzymes to help out with the difficult part of synthesis. The result is the efficient production of artemisinin and at a price point that is realistic for the developing countries that most desperately need antimalarials.

Stealing a pathway from a cell is a pretty cool advance, but what if you wanted to create a pathway that doesn't fully exist in one cell but individual reactions of it might occur in 30 cells? How would you do it?

One strategy is to make an entirely new cell by creating designer genomes. This is possible thanks to 2 advances: We have knowledge of the contents of thousands of genomes, and we have developed very efficient and sophisticated chemistry to chemically synthesize DNAs in the laboratory. We can make any sequence we desire, and best of all, machines do all of the work.

At first, the chemical process was relatively inefficient, so only short oligonucleotides on the order of 20 to 30 bases could be made. But improvements in technology soon made it possible to cheaply and accurately synthesize not just individual genes but entire genomes.

In 2010, a research group headed by J. Craig Venter reported that they had created the entire genome of the bacterium *Mycoplasma mycoides* from the known sequence stored in a computer. They then removed the genome from a

bacterial cell, inserted the synthetic genome instead, and showed that the cell now was controlled by the synthetic genome—the first time ever that a cell was running on synthetic instructions.

While Venter made the DNA according to the sequence that we already knew for the bacterial genome, there's nothing that says that we can't create genomes that don't already exist. Genomes could be designed for bacteria so that they could make biofuels based on metabolic pathways that don't exist in any single cell. Genomes could be designed for plants to enable them to synthesize brand-new drugs, instead of fabricating the drugs with the much slower and more expensive reactions elsewhere in chemistry.

We could even rewrite our own human genome—making artificial human cells, not human beings. Advantages of making artificial human cells include creating customized human organs for transplant into real human beings to eliminate the issue of rejection, creating custom stem cells that could be used to regenerate an immune system for a person who had theirs destroyed by radiation therapy to kill a cancer, custom-making retinal cells to replace ones lost by macular degeneration, and replacing pancreatic cells responsible for making insulin that are damaged in diabetics.

Virtually every technology that has been proposed for stem cell use could be accomplished with artificial human cells. And although there are enormous technical hurdles to overcome with making artificial human cells, the many benefits may well justify working to overcome them.

Scientists agree that making human beings needs to await a full discussion of the enormous ethical implications. But growing human cells in cultures in laboratories, independent of human beings, is already commonly done, accepted, ethical, and not harmful in any way.

READINGS

Church and Regis, *Regenesis*.
Davies, *Synthetic Biology*.
Knoepfler, *Stem Cells*.
Shapiro, *How to Clone a Mammoth*.
Stockwell, *Quest for the Cure*.

QUESTIONS

1 Copying DNA from a eukaryotic cell into a bacterium doesn't always give the same protein, even though they have the same genetic code. Why?

2 Consider the mouse endothelial tumor in the lecture targeted by Chinese researchers using a nanobot. The tumor contained a protein called nucleolin on its surface that was not found on normal cells. Describe another technology from these lectures that might similarly be employed to target and potentially destroy such a tumor.

CLICK HERE TO SEE THE ANSWERS.

36

OMICS: GENOMICS, PROTEOMICS, TRANSCRIPTOMICS

The omics are the leading edge of an emerging new approach to medical care sometimes called P4 medicine—an approach that is predictive, preventive, personalized, and participatory. It contrasts with traditional care, which is mostly about reacting to problems and treating symptoms.

Genome Sequencing

The arrival of omics started with genomics, which was itself spawned by the completion of the Human Genome Project in 2003. That massive undertaking determined the order, or sequence, of every base in the human genome.

One of the benefits of this project was the development of ways to sequence DNA that are a million times cheaper and 5000 times faster—to the point that today you can get your entire genome read in about a day. And in this case, it would be your specific genome, with its own tiny quirks.

Of course, this is a lot of information to have about yourself. Theoretically, storing 3 billion base pairs of information would take up most of a good-sized thumb drive, depending on how much peripheral information you included.

Moreover, accurate sequencing is not a "one-and-done" sort of process. The quality of your data strongly depends, for example, on the number of times each base in the genome is read during sequencing. A sequence that has not been read and checked multiple times is called a draft. Fortunately, costs have fallen dramatically.

Starting in 2001, researchers had called for the cost of sequencing a human genome to decline to less than $1000, which is less than an MRI at some hospitals. That goal was achieved in around 2014 or 2015. The next goal is whole-genome sequencing in about an hour for $100.

Genome sequencing is not just for humans. The entire DNA sequence for thousands of other organisms has also been obtained. There are now huge databases filled with the sequence information generated by these projects, and more information is added each day. Side by side with getting sequence data, scientists have also been determining and cataloging the structure of thousands of proteins.

The easy availability of so much genome data allows for a whole different level of analysis, called genomics. Entire new fields of study have sprung up to analyze the flood of data—for example, bioinformatics and systems biology.

Besides genomics for DNA, there are several other omics, including transcriptomics for RNA, proteomics for proteins, and metabolomics for the metabolites in a cell. These fields aim to look at all the genes, all the RNAs, all the proteins, and all the metabolites in a cell simultaneously in order to obtain a complete picture of what's going on and when.

For genomics, researchers are examining DNA sequences to determine not just protein-coding locations, but also the functions of encoded proteins and how they might be regulated. With information from so many species available, comparative genomics, where processes central to evolution may be revealed, also becomes possible.

Central to understanding information in DNA is genome annotation. The sequence of a genome is just a string of billions of bases. Scientists use algorithms that look for features such as protein-coding regions, potential regulatory elements, and so on, so that function can be assigned to each sequence in the genome. This gets noted in the form of an annotation that helps other researchers use the information.

All of the genome, RNA, and protein sequence info and annotations are stored in databases across the world. While the data can be accessed by anyone with an internet connection, the raw information is pretty incomprehensible without specialized knowledge and understanding. At some point in the future, friendlier interfaces may permit nonexperts to interact more meaningfully with the data.

Researchers are steadily identifying the variations in the DNA sequences of different individuals and the significance of those differences. These patterns of differences can be used in the diagnosis and personalized treatment of disease. Normal cells and cancer cells can be compared to determine which proteins are involved, how those proteins interact with each other, and what cellular processes are altered.

Meanwhile, genome analysis is adding to our understanding of biochemistry as an integrated whole. And the study of the small genetic variations among humans has already provided quick returns. Only about 0.1% of our genomes varies from one human to another. But those small, subtle differences can have consequences for our health. They can affect whether, and to what extent, we get sick and even how we respond to treatments. Some differences are in protein-coding regions, so they affect the function of the encoded protein. And even the mutation of a single base can have negative implications for health.

Cataloging Variations in Genomes

Scientists are also cataloging single-nucleotide polymorphisms (SNPs), which are specific places in genome sequences where a single nucleotide differs between groups of people—for example, an A is replaced by a C. An SNP is a point mutation, but one that is found in many individuals, not just one person. SNPs are dotted all over the genome, about once in a thousand nucleotides, on average.

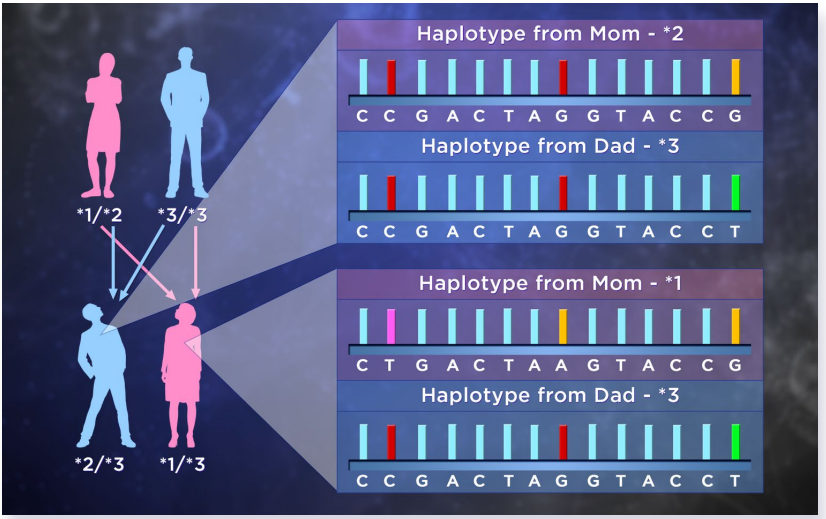
Of the millions of SNPs in our genomes, many are not in protein-coding regions. And a single SNP is not very informative. But researchers have done genome-wide association studies that look across the genome to see whether some SNPs are more prevalent in people with a particular disease compared to those who are healthy.

Such studies have shown that some patterns of SNPs in a stretch of DNA are found to be associated with susceptibility to a given disease or with response to a medication. This could be because variations in SNPs might change how gene expression is regulated, for example.

A particular combination of SNPs that is inherited together from one parent—or that tend to be inherited together across many people—is known as a haplotype.

A person gets ½ of his or her chromosomes from the mother and the other ½ from the father, so his or her full SNP profile will have 2 haplotypes. The mother might have 2 different haplotypes to offer, while both of the father's haplotypes might be the same. Whatever the full SNP profile is could make one child more likely to benefit from, or have an adverse drug reaction to, a particular treatment than someone with a different SNP profile.

SNP profiles can help in decisions about whether an individual will benefit from the asthma drug albuterol and in determining the dosage of the blood thinner warfarin for a particular patient.



The use of genomic profiles in deciding how to treat patients is called pharmacogenomics. The FDA has begun to provide information based on this knowledge and, in some cases, recommends patient testing before the drug is prescribed.

With more genomic information available, doctors will be able to choose medications that are best suited to each patient, rather than a one-size-fits-all drug.

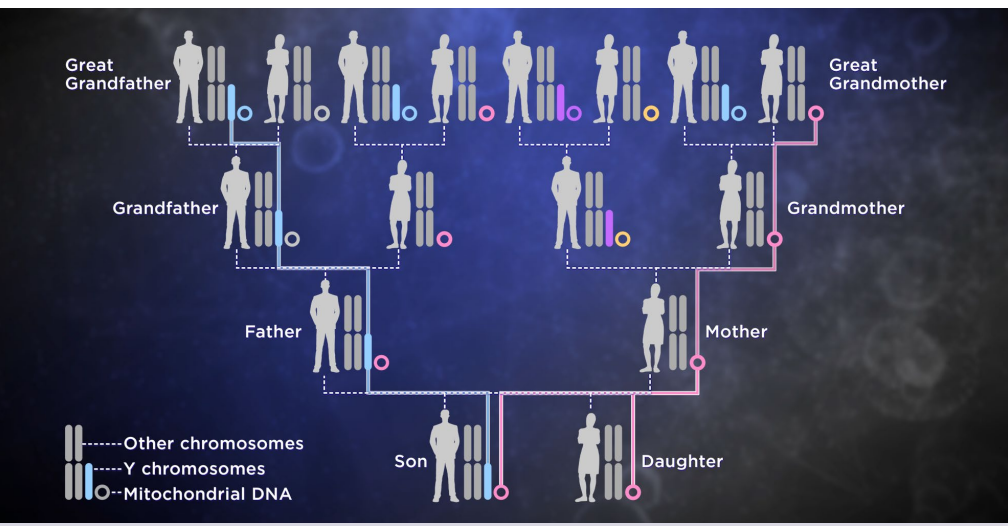
For some conditions—such as cardiovascular disease, inflammatory bowel disease, and breast cancer—genome-wide studies of SNPs have been very useful in risk prediction. For other multifactorial diseases, SNP profiles are not yet sufficient to provide clear-cut answers. This is partly a function of the complexity of the diseases; there are often many genetic factors, which interact with each other and with the environment. Scientists are still determining what all the pieces are.

Because of this, people who send away to companies offering personal SNP profiles, or sequencing of their genomes, should be aware that information is not available to provide a complete picture of their disease risk.

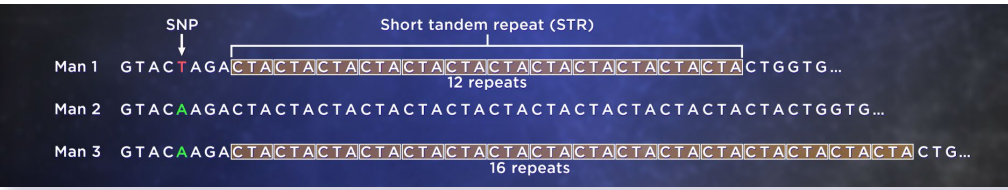
Of course, the cataloging of small variations in the genomes of individuals is also useful for tracing human migrations and ancestry. The patterns of variations used in such studies focus on mitochondrial DNA to trace the maternal line and Y chromosomes to trace the paternal line.

The DNA in our mitochondria is inherited from our mothers. This is because, in the production of a fertilized egg, the egg is usually the source of mitochondria. Except in rare cases, sperm do not contribute mitochondria. This means that we inherit our mitochondrial DNA from our mothers, who got it from their mothers, and so on, down the female line.

Likewise, the Y chromosome is passed down from father to son, down the male line. In the lines from both parents, mutations in the DNA are passed on, and sets of common SNPs can be used to trace a maternal or paternal line of inheritance.



In addition to SNPs, Y chromosome analyses also include variations in the number of short repeated sequences called short tandem repeats (STRs), which present in varying numbers in different individuals. The numbers of these repeats, like the SNPs, are inherited by sons through the Y chromosomes.



Particular combinations of variations are used to define haplogroups, groups of people with a shared haplotype. If you share the same mitochondrial haplogroup as someone else, you share a maternal ancestor somewhere along the line. Likewise, if you are in the same Y chromosome haplogroup as another person, you have a male ancestor in common with that person.

Still, it's important to remember that these haplogroups represent ancient branches of the human family tree, going back tens of thousands of years. Everyone in a given haplogroup can trace their ancestry back to a single individual in the deep past.

It's entirely possible to, for example, be from Korea and fall in the same mitochondrial haplogroup as someone who is Native American. People migrated across the land bridge from Asia to North America in the past, so both Koreans and the descendants of those who migrated from Asia to North America could have had a common female ancestor who lived in Asia a very long time ago.

Genghis Khan's Y chromosome haplotype is seen in more than 16 million men today.

Transcriptomics

Transcriptomics analyzes transcripts—that is, RNAs—on a genome-wide level. This allows us to know exactly which genes are being copied into RNA under different conditions. For example, it is possible to identify all the genes transcribed in response to treatment of cells with a hormone or to compare the global transcription patterns in normal and cancerous cells.

This is made possible by the development of technologies for the rapid sequencing of thousands of RNAs. A technique called RNA-seq is used to detect and measure the levels of every transcript in a given sample.

Transcriptome sequencing can provide information about all types of RNAs, including regulatory RNAs. The sequencing of regulatory RNAs can tell us about when certain genes might be silenced, while sequencing of the other RNAs provides a complete profile of all the protein-coding genes that are expressed. For example, transcriptomics has shown that in some cancer cells, regulatory RNAs are made that shut off production of tumor suppressor proteins, allowing cancer cells to proliferate unchecked.

Getting your genome sequenced if you want to understand what's happening with many ordinary ailments, such as a cold or pneumonia, will have almost no meaning—your genome doesn't change during an illness. But the transcriptome changes in response to everything! Knowing how the transcriptome changes in response to drugs, infections, cancer, and so on is going to be a treasure trove of information for better diagnosis and treatment.

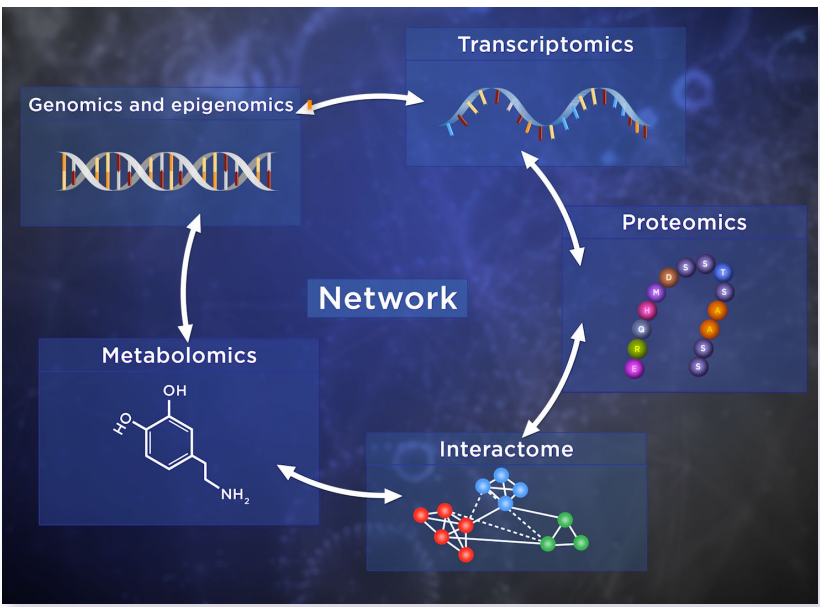
Other Omics

Besides studying genomes and transcriptomes, researchers also analyze complete collections of proteins with proteomics. All of the proteins that are in a particular cell type or tissue can be extracted, separated, and identified.

Differences in protein profiles between normal and cancerous cells may provide clues to the mechanisms by which the cancer arose or suggest ways to treat the cancer. Such information is useful in understanding the underlying causes of cellular dysregulation.

Lipids are another group of molecules that give us insights into many medical conditions. Advances in the techniques for studying global lipid profiles have opened up the field of lipidomics. Information about how lipids vary in cells under different conditions help in predicting the risk of metabolic diseases, such as diabetes, nonalcoholic fatty liver, and cardiovascular disease.

Lipidomics can also show us how diets affect cellular lipid profiles. And lipid profile changes can tell us how effective drugs are. Lipidomic studies can help screen out treatments that have unexpected and severe side effects and identify those that normalize biochemical profiles.



Lipidomics is complemented by global measures of molecules like carbohydrates and other classes of cellular molecules. All of these are grouped together in a study of metabolites known as metabolomics.

Together, these fields are allowing biochemists to have, for the first time, a big-picture view of the activities of cells. Integration of information from the genome, transcriptome, proteome, and metabolome will show how these components are interconnected and how these interactions determine biological functions.

READINGS

Dudley, et al., *Exploring Personal Genomics*.
Field and Davies, *Biocode*.
Quackenbush, *The Human Genome*.
Rutherford, *A Brief History of Everyone Who Ever Lived*.

QUESTIONS

1 It is often noted that although we know the complete sequence of the human genome, we are many years away from knowing what it all means. Based on what you have learned in these lectures, what meaning of the genome is missing when you only look at DNA sequence?

2 Your friend asks you why both proteomics and transcriptomics exist, given that the transcriptome leads to the proteome. Based on what you've learned in these lectures, how would you respond?

CLICK HERE TO SEE THE ANSWERS.

Answers

LECTURE 1

1 Elements needed for making biomolecules must be able to make multiple bonds and/or share electrons to make covalent bonds. The 6 most abundant ones—carbon, hydrogen, oxygen, phosphorus, sulfur, and nitrogen—meet at least one of these criteria. Other abundant elements on Earth, such as iron and aluminum, cannot do what these atoms do. Silicon is a notable exception, but carbon is probably a better atom for building larger molecules. Some people wonder, though, if silicon-based life might occur somewhere else in the universe.

2 Enzymes are the proteins that bridged the gap. They do not require a cellular environment in which to function, yet they perform the magic of catalyzing reactions rapidly. Eduard Buchner was able to demonstrate this clearly when he showed that the reactions of fermentation could be performed in a test tube by extracting the relevant materials from cells undergoing fermentation. The relevant materials turned out to be enzymes.

LECTURE 2

1 The other feature, which is not commonly mentioned about water, is its V shape, which gives it a polarity: One end is negatively charged, and the other end is positively charged. If water were linear like carbon dioxide, it would not have polarity and would probably behave quite differently than it does.

2 The common feature that nonionizing substances have that allows them to dissolve in water is the ability to form hydrogen bonds. Sugars readily form hydrogen bonds due to their numerous hydroxyl groups. Fats and cholesterol do not form significant hydrogen bonds in water and do not dissolve.

3 a pH < pKa = less salt than acid

b pH = pKa = equal amounts of salt and acid

c pH > pKa = more salt than acid

For example, when salt < acid, the pH = the pKa + a negative number. That means the pH must be less than the pKa.

LECTURE 3

1 Hydrophobic side chains (methionine, for example) avoid water and will associate with other hydrophobic side chains to avoid water together.

Bulky side chains (tryptophan, for example) cause bulges in polypeptide backbones and interrupt regions of repeating structure when near other bulky side chains.

Charged side chains (aspartic acid, for example) form ionic bonds with oppositely charged side chains (lysine, for example). They may be repelled by similarly charged side chains and will interact well with water.

Polar side chains (serine, for example) interact well with water by forming hydrogen bonds.

Reactive side chains (cysteine, for example) form covalent bonds with other cysteine side chains when brought into close proximity.

Inflexible side chains (proline, for example) restrict the twisting motion of peptide backbones. They often interrupt the structure of repeating structures and may cause bends in the structure.

2 Acids plus alcohols yield esters. Phosphate bonds with serines, threonines, and tyrosines are ester bonds.

3 Any health issue producing nasal congestion arises from the production of histamines, which are made ultimately from the amino acid histidine. Because histidine is an essential amino acid and is needed for making proteins on an ongoing basis, it would not be a good idea to restrict histidine intake.

LECTURE 4

1 Most alpha helices and beta sheets have small structures, so only a few hydrogen bonds are involved. The energy needed to break bonds is partly related to the number of them. DNA, by contrast, has thousands to millions of hydrogen bonds holding together a duplex. So, the smaller number of hydrogen bonds in proteins makes them easier to break, and thus lower temperatures are required than is needed for breaking hydrogen bonds in DNA.

2 Fibrous proteins like silk often have repeating structures. Plastics are a nonbiological example, and they have the structure they do thanks to polymers—millions of repeats of the same thing. Repetition is good for uniformity, so using a small subset of the 20 amino acids makes for easier structural uniformity. Globular proteins are focused on function, so they do not need repeating structures, though they do need to have specific structures. The 20 amino acids allow them to achieve their specific nonrepeating structures.

3 Membrane proteins often have polypeptide chains that cross the layer multiple times. The amino acids in membrane proteins are arranged so that portions of the protein that are touching water, such as the cytoplasm or exterior of the cell, have hydrophilic amino acids, whereas the amino acids in the nonpolar part of the bilayer have side chains that are hydrophobic.

LECTURE 5

1 Folding in solution, as opposed to folding in a cell as a protein is being made, can result in cysteines of different proteins bouncing into each other and reacting, thus causing bonds to form that might otherwise form internal to a protein. These will prevent the proteins involved from folding properly. A small amount of mercaptoethanol allows them to disengage and fold properly.

2 Exposure of hydrophilic side chains to the unfolded protein in the chamber mimics the polarity of the aqueous environment of the cell. Consequently, hydrophobic amino acids avoid the polarity and associate with each other. This helps drive the folding process that leads to hydrophobics inside a folded protein and hydrophilics on the outside.

3 Hydrophobic interactions may be necessary to help stabilize a folded protein's structure. Without enough hydrophobic amino acids to form a core, that element of stability is lost, and the protein may be much less restricted and free to adopt many shapes.

LECTURE 6

1 Such a protein would be a single subunit (thus lacking quaternary structure) that would have multiple binding sites for a molecule.

2 Myoglobin acts like an oxygen battery backup: It provides oxygen when cells run out of it. But batteries get drained if they are used too often. Myoglobin's tendency to bind oxygen tightly means it will only release oxygen when the oxygen concentration is very low. That's the other important role for its high affinity.

3 The oxygen demands of a fetus are not nearly as drastically changing as the mother's. The mom has to move, work, and be active during the day but also sleep at night. Thus, the needs range from high to low. Because the

fetus is mostly immobile, except for a little kicking, its oxygen needs don't vary much, so a relatively constant oxygen supply is all that is needed. This can be obtained from hemoglobin in the R state.

LECTURE 7

1 The model in the lecture regarding the bending/breaking of a stick is the relevant one. The flexibility of the enzyme can place a strain on the bound substrate, and this strain can be a factor in favoring the breakage of bonds that is necessary to split a molecule into 2 pieces.

2 The active site of serine proteases is internal to the protein and protected from exposure to outside molecules. As a result, the only thing that gets exposed to it is the peptide bond of the bound substrate.

3 Perfect enzymes sometimes evolve to reduce the release of intermediates that are toxic. The faster a reaction occurs, the less likely the intermediate can be released. That is the case with triose phosphate isomerase, whose intermediate, known as methylglyoxal, is toxic.

LECTURE 8

1 By controlling the first enzyme of a pathway, the first product is controlled. If the enzyme is activated, then the first product is available to be the substrate for the second enzyme, which can catalyze and pass the product to the third enzyme, etc. Because the second enzyme is always on, it is ready to act on the product of the first reaction when it is made. On the other hand, if the first enzyme is turned off, then there is no product made, and the second enzyme has nothing to act on, even though it is active. Thus, the activity of the second enzyme in both cases has no effect. Regulating the first enzyme is all that is necessary.

Those who say that cells do regulate all of the enzymes in the pathway are referring to the secondary effects resulting from the presence or absence of substrate. With no substrate, other enzymes are essentially inactive, because there is nothing to act on, but when substrate is present, enzymes are actively converting it to product.

2 Though PALA is an inhibitor of the enzyme, the binding of PALA converts the enzyme from the T state to the R state. If one has sufficient PALA, all of the substrate binding sites get occupied, so the enzyme is blocked from acting. When only tiny amounts of PALA are available, some of the substrate binding sites of the enzyme are left open. Thus, substrate can bind to those sites, and the enzyme can act on them, because it has been converted to the R state.

3 Not surprisingly, the person's health outcomes are not positive. Like the smokers who inactivate their alpha-1 antitrypsin, this person will have an overly active elastase enzyme and will be much more likely to develop emphysema due to its damage.

LECTURE 9

1 Fatty acids are the components of soaps, and like soaps, they interact with nonpolar substances. Many proteins have interiors that are nonpolar. If fatty acids interacted with them randomly, some proteins would denature due to the interaction with the soap. Soaps and detergents readily do denature proteins. That, in fact, is why you kill germs by washing your hands with soaps: You denature proteins they need to survive.

2 The inaccuracies are that they often do not show the unsaturated fatty acids, which have bends in them that break up the regular structures often depicted.

3 Skin color is a factor. People with darker skin who live in far northern or far southern latitudes may have trouble making sufficient vitamin D in winter months. Melanin, which gives skin its color, absorbs ultraviolet light, which is needed for making vitamin D. In the tropics, sufficient ultraviolet light escapes the melanin, but in other areas, there may not be enough available to make proper amounts of vitamin D.

LECTURE 10

1 All nonhuman mammals are lactose intolerant due to the fact that they never drink milk after infancy. There is no selective advantage for them to be able to digest lactose after they are weaned, so their lactase enzymes get turned off shortly after that time.

2 With a grazing diet, insulin levels are continually rising with a meal and falling after it. Insulin levels fall after glucose levels fall from insulin action. Meanwhile, rising and falling glucose levels change appetite from low, when glucose levels are high, to high, when glucose levels are low. These roller coaster effects can lead to insulin resistance—a precursor to diabetes.

3 Diet drinks, particularly when consumed in large quantities on a regular basis, may induce insulin production that is similar to the consumption of sugar. This is because sweet receptors can alert the brain that food may be coming. The problem with insulin released like this is that insulin causes glucose to be taken up by cells, thus reducing blood glucose levels further. In this case, no glucose is in the intestines to actually be absorbed into the blood. As a result, the diet drink consumer may experience hunger, something counter to what the diet drink was designed to avoid.

LECTURE 11

1 At equilibrium, the ΔG equation is

$$\Delta G = 0 = \Delta G^{\circ\prime} + RT\ln\{[B]/[A]\}$$

When B is doubled, the ratio of B to A doubles. This increases the positive value of the $RT\ln\{[B]/[A]\}$. Because that term is getting more positive and the $\Delta G^{\circ\prime}$ is a constant, the ΔG is no longer 0, but instead has a positive value. Positive ΔG values mean the reaction runs backward. Thus, adding product B upsets the equilibrium, and the system responds by converting B into A by reversing the reaction. When A is doubled instead of B, the forward reaction is favored, because ΔG is negative. The Gibbs free energy equation is thus reflecting Le Chatelier's principle.

2 Let C stand for creatine and CP stand for creatine phosphate:

$$\Delta G = \Delta G^{\circ\prime} + RT\ln\{[CP][\text{ADP}]/[C][\text{ATP}]\}$$

If you increase C, the ratio of products to reactants will become smaller, and thus the $RT\ln\{[CP][\text{ADP}]/[C][\text{ATP}]\}$ will become more negative. A more negative term here means that the overall ΔG will become more negative. A negative ΔG means that the reaction will move forward. This means that an athlete who takes creatine before a race will be decreasing his or her level of ATP—which is not a good idea at the start of a race.

LECTURE 12

1 Rapidly respiring cells, such as muscles, have greater oxygen needs than more slowly respiring cells. The body needs a mechanism to determine what a rapidly respiring cell is; it can't simply target muscle cells. You might, for example, be running, so your leg muscles would need more oxygen than your neck muscles at that time.

One signal is the amount of energy-producing reactions a cell is running. The more of these there are, the greater the oxygen needs of the cell will be. The molecule 2,3-BPG is an excellent indicator of the extent of energy-producing reactions. The more of it is released by cells, the more glycolysis is going on, and the greater the oxygen needs of the cell are.

Importantly, 2,3-BPG also facilitates the oxygen release by binding to the hemoglobin and favoring oxygen release. Thus, 2,3-BPG is both an indicator and an activator of oxygen release at the point where the body most needs it.

2 Because lactic acid is an acid and releases protons, when it is released by cells in fermentation, the indication is that the cell doesn't have enough oxygen and needs more. Protons binding to hemoglobin facilitate oxygen release to help ease the problem.

3 Because lactose contains glucose and galactose, a dietary modification might involve the reduction of the consumption of dairy products containing lactose. Note that some foods lack lactose because they are treated with the enzyme lactase, which breaks lactose down to glucose and galactose. The consumption of lactase-treated foods would do no good for sufferers of this disease.

LECTURE 13

1 Yes, it would, at least from the perspective of having the necessary substrates. With aconitase inactive, citrate would accumulate, and when that happened, the conversion of oxaloacetate to citrate would stop. Thus, oxaloacetate would increase in concentration, and that is the substrate for making glucose. One caveat, though, is that a cell lacking aconitase would likely not be viable, because the citric acid cycle wouldn't operate to provide much-needed energy.

2 The citric acid cycle works because it has inputs and outputs. Left to itself, the cycle only requires an input of acetyl-CoA with each turn, and left to itself, the removal of succinyl-CoA would indeed break the cycle if nothing were added to make up the difference.

Fortunately, molecules from other pathways can contribute. Alpha-ketoglutarate, for example, can be made from glutamate and enter the cycle. If that happened, glutamate could be a source of making succinyl-CoA and ultimately heme. That's the beauty of a cellular interchange like the citric acid cycle: Molecules from seemingly unrelated pathways can provide materials for or take materials from other pathways. The pentose phosphate pathway is similar in this regard.

3 Running the citric acid cycle backward provides cancer cells with a means of using acetyl-CoA to make fatty acids for making membrane lipids. Making fatty acids requires NADPH, which is made by the PPP. Blocking the oxidative reactions of the PPP would stop NADPH production, which in turn would inhibit membrane-lipid synthesis. Without membrane lipids, cells couldn't divide.

LECTURE 14

1 The most poisonous of the electron transport inhibitors are the ones inhibiting complex IV, because there is no way around that complex. If there are alternative points of entry when complex I is blocked, then the system could still function using the alternate routes, but there is no way around complex IV, so molecules inhibiting it have the most poisonous effects.

2 It is energetically favorable for electrons to get passed from NADH and $FADH_2$ to complexes I and II, respectively. No push is required, so these processes are downhill. The light cycle of photosynthesis, on the other

hand, requires electrons to be excited by sunlight to start the process and is consequently an uphill process. Note, though, that once excited by the sunlight, electrons flow downhill much like they do in mitochondrial electron transport. Only the first part of their journey is uphill.

3 Until we try the experiment, we won't know, of course, but that doesn't stop us from speculating. The fish probably would remain the same size as it was when you took away its carbon source. There is a good chance, though, that it would die for a different reason, because carbon sources also often serve as nitrogen sources (think amino acids), and without a nitrogen source, the fish would not be viable.

LECTURE 15

1 This would likely arise from consuming a diet rich in fats but poor in proteins and carbohydrates. Under these conditions, there would be few carbs for the body to use, so when this happens, ketone bodies are made to feed the brain and other tissues, because the body would not be able to make much in the way of carbs without amino acids from proteins.

2 The 2 pathways occur in different cellular locations. Fatty acid oxidation occurs in mitochondria, whereas fatty acid synthesis occurs in the cytoplasm. Physical separation is one way of avoiding futile cycles.

LECTURE 16

1 Digoxin affects the sodium/potassium ATPase, which makes the sodium/potassium gradients across cell membranes that are critical for the function of nerve cells. All of the processes described rely on nerve cell function.

2 LDL receptors on cell surfaces bind to LDLs and take them into the cell so that the cholesterol and other materials can be used. As a result of this update of LDLs, the concentration of LDLs in the bloodstream decreases. When enzymes that break down LDL receptors are active, they decrease the number of receptors and thus decrease the number of LDLs removed from the blood, allowing LDL levels to be higher. If the enzymes that break down the receptors are blocked, then there are more receptors to take in LDLs, so LDL levels in the blood fall. Thus, blocking enzymes that break down LDL receptors decreases LDL levels in the blood.

LECTURE 17

1 Glucose is a source of quick energy for cells and occurs in the cytoplasm of the cell. Gluconeogenesis also occurs mostly in the cytoplasm. Because they both occur in the same cellular location, for the most part, they will tend to occur faster than futile cycles that are separated spatially in cells. That is the case with fatty acid synthesis (cytoplasm) and fatty acid oxidation (mitochondria). Finally, glycolysis has the big bang reaction catalyzed by pyruvate kinase that releases a lot of heat energy. Fatty acid oxidation has no such large heat-generating reaction.

2 ATCase has 2 binding sites for ATP, whereas enzymes that are activated by a substrate use one site for both activation and catalysis. ATCase is in this category with respect to aspartate. ATCase does have separate regulatory sites for ATP and cytidine triphosphate (CTP), but aspartate acts through its substrate binding site at the active site. When aspartate levels are high, binding results, and the enzyme gets activated and catalysis results.

PFK, on the other hand, has one allosteric binding site for ATP and a separate substrate binding site for ATP. The interesting thing here is that the allosteric binding site for ATP has low affinity for ATP, meaning that only when ATP levels are high will the allosteric binding site get bound by ATP to turn the enzyme off. When ATP levels are low, the only site it can bind to is the substrate binding site.

Thus, low ATP levels leave the enzyme active (no allosteric binding) and able to catalyze its reaction to keep glycolysis going. When ATP levels are high, the allosteric binding site gets bound, and the enzyme gets turned off.

3 It wouldn't have any drastically different effects. Uncoupling protein is the protein that permeabilizes the mitochondria and causes brown fat to burn through energy stores. The primary factor affecting uncoupling protein is the fatty acid palmitate. When palmitate is present, the protein will be blocked, but in its absence, it will be active. Exercise versus nonexercise won't have a drastic difference in palmitate concentration, at least in the short term.

LECTURE 18

1 ATP to make glucose = 18 + 3 × 12 = 54 ATPs.

ATP from glucose breakdown with oxygen present = 36.

ATP from glucose breakdown with no oxygen present = 2.

Efficiency of energy capture from glucose with O_2 = $^{36}/_{54}$ × 100 = 66.7%.

Efficiency of energy capture from glucose with no O_2 = $^{2}/_{54}$ × 100 = 3.7%.

2 Assuming that C_3 and C_4 plants are each capturing sunlight at the same rate and there are no limiting factors, such as ADP or $NADP^+$, the 2 will be the same. The differences between C_3 and C_4 plants are in the dark phase, not the light phase, of photosynthesis.

3 No, it is more efficient for the dark cycle to be operating in the light, because the dark cycle uses ATP and NADPH and generates ADP and $NADP^+$, both of which can then be "recharged" with another light reaction. If the plant ran out of ADP or $NADP^+$, then it would waste energy from the Sun, because it wouldn't be able to start electron transport. For this reason, plants run the dark cycle in the light, unless they are CAM plants.

LECTURE 19

1 Carbon dioxide: poisonous as is (can also be released by decarboxylation).

Oxygen: poisonous when converted to reactive oxygen species.

Ammonia: poisonous when released freely from amines in amino acids.

2 Nitrogenase requires an absence of oxygen. Plants and animals live in oxygen-rich environments. Consequently, places have to be created that lack oxygen for cells using nitrogenase to function. These are often in plant root nodules and are not common; hence, nitrogenase is rather rare.

3 The elimination of nitrogen in higher animals occurs via urea that comes from the urea cycle. Amine groups enter the cycle in 2 ways: by combining ammonia with carbon dioxide and phosphate (no transamination involved) and by transferring the amine from aspartate through arginine to make urea (also no transamination involved). Transamination is important for getting the amine to make aspartate and to get it to glutamate so that ammonia can be cleaved from it, but transamination itself is involved in neither of the reactions that gets amine groups together to make urea.

LECTURE 20

1 superoxide dismutase = SOD
copper = Cu
superoxide = O_2^-
oxygen = O_2
$SOD\text{-}Cu^{++} + O_2^- \Leftrightarrow SOD\text{-}Cu^{+} + O_2$

In this reaction, moving to the right, $SOD\text{-}Cu^{++}$ is getting reduced to $SOD\text{-}Cu^{+}$ and O_2^- is getting oxidized to O_2.

2 The key to understanding reactions involving molecules with unpaired electrons is that they can start chain reactions. As a result, a hydroxyl radical can take away an electron from another molecule, creating a hydroxide ion (no unpaired electron), but the molecule that lost the electron now becomes a radical on its own, because it now has an unpaired electron and will repeat the process. A hydroxide ion has no unpaired electrons, so it will not take electrons away from anything; instead, it will seek a proton to react with and create water.

LECTURE 21

1 Bacteria use molecules in quorum sensing to detect if there are enough of them present to be effective. If quorum-sensing molecules send the signal that there are more bacteria present than are actually there, they may mount an attack that has a lower likelihood of effectiveness.

2 Exocrine signaling in animals occurs through sweat glands, though it may come from other sources in some animals. Such signaling most commonly has to do with mating—the detection of mates or warding off competition.

3 Remember that GTP is important for G proteins. The binding of GTP by the G proteins results in their activation and the subsequent stimulation of adenylate cyclase to make cAMP. G proteins also slowly convert GTP to GDP, which inactivates the G protein, which would stop synthesis of cAMP. If you had an inhibitor of that reaction, GTP would not get broken down to GDP, so the G protein would remain active and adenylate cyclase would continue to make cAMP. This would activate protein kinase A, which would phosphorylate a bunch of target proteins and affect their activities.

LECTURE 22

1 The wave is largely occurring inside a neuron and won't affect much outside of it. Neurotransmitters are therefore needed to move the signal out of the first neuron to the second one. The wave won't do that.

2 This would probably be a relatively mild venom, depending on the amount of activation of acetylcholinesterase. The enzyme would decrease the level of the signaling molecule, so instead of tetany resulting from inhibition of the enzyme, less muscle contraction would occur due to less signal being received by the target muscle, due to destruction of the neurotransmitter.

LECTURE 23

1 The 2 that do not use GPCR are ions (salt and protons from acids), which can directly affect ion-gated channels to initiate the signaling process. The other 3 tastants are not ionic and cannot directly open ion-gated channels. Consequently, they must signal through binding to a specific receptor that initiates the signaling for them.

2 In retina cells, cGMP functions to bind to sodium/calcium channels to keep them open in the absence of detection of light. When light is detected, a phosphodiesterase is activated that breaks down cGMP, thus closing the channels and initiating a signal to the brain. Blocking phosphodiesterase activity would result in cGMP levels remaining high, sodium/potassium gates remaining open, and no signal being sent to the brain. Effectively, this could result in blindness, because retinal cells would not respond to light.

3 Cells in the retina act to protect retinal from oxidation using carotenoids like lutein and zeaxanthin.

LECTURE 24

1 For such a person, a high-fructose diet would be of benefit, as it would be using the alternate pathway of fructokinase described in the lecture.

2 No, you would not expect it to accumulate fructose, because fructose can be converted to fructose 6-phosphate by hexokinase, as noted in the lecture. Fructose 6-kinase would freely enter glycolysis, and fructose would not accumulate.

3 Assuming no limit on pathway quantity (glycolysis or the alternate fructose pathway), you would not expect any difference for the 3 different sugar offerings. If there were a limit on the capacities of these pathways, then mixtures of fructose and glucose, such as high-fructose corn syrup or sucrose, would be better than either pure fructose or pure glucose, depending on which pathway had the capacity limit.

LECTURE 25

1 Although catalytic RNAs can catalyze some reactions, they probably cannot catalyze the diversity of reactions needed in metabolism (and other cell reactions) due to their relatively simple set of 4 building blocks—compared to the 20 amino acids used to make proteins. Further, the chemistries of the amino acids range from hydrophobic to polar to ions, and this is important not only for structure, but also for catalytic mechanisms. RNAs lack those properties.

2 DNA is extremely large and unwieldy. By making a subset of it in the form of RNA, the management of these smaller pieces is easier, and the pieces are relatively easily handled by the protein-making complexes. Further, in eukaryotic cells, DNA is coated with histone proteins that would prevent easy access to coding sequences by protein-making complexes.

LECTURE 26

1 The beta-clamp protein functions to hold DNA polymerase III onto the DNA so that it does not fall off readily, thus facilitating long, productive, and rapid replication of the DNA. In the absence of beta clamp function, the DNA polymerase would fall off of the DNA more frequently, and DNA replication would take much longer to complete.

2 If DNA ligase were inhibited, Okazaki fragments would not be joined together, so when helicase came along on the next round of replication and separated strands, nothing would hold them to the rest of the strand they would normally be attached to, and they would go flying off as helicase unwound them.

3 Poor error correction is enormously beneficial for viral evolution and may be why many viral polymerases lack proofreading. Mutations are the driving force of evolution. For functioning cells, most mutations are harmful, and death can result.

The viral strategy, though, is to replicate as rapidly as possible, and if a few viruses are made that are not functional, that is not a big deal. Those few that are nonfunctional might turn out to have mutations that allow them to function in an environment where nonmutated viruses cannot replicate.

Thus, the more mutant viruses exist, the more likely a virus will persist. This, in fact, is the way that the flu virus evolves, and it is also HIV's strategy for evolving resistance to antiviral drugs. Many other viruses do the same thing.

LECTURE 27

1 Differentiated organisms must manage the growth of individual cells. What if your 2 legs were not the same length? What if one eye were twice the size of the other? Cell cycle regulation allows for the control of division, which in turn allows for properly functioning organisms. Loss of control can also lead to cancer.

2 The 2 characteristics are the requirement for a primer to start replication and the absolute requirement to replicate from the 5′ end to the 3′ end. If either of these were not necessary, chromosomes could easily be replicated all the way to their linear ends.

LECTURE 28

1 Perfect DNA replication would, in fact, mean no evolution if replication were the only way mutation can arise. However, mutation can arise as a result of chemical damage, and changes in DNA can also occur as a result of recombination. In addition, the base cytosine is chemically unstable. So, while evolution would likely be slowed with perfect DNA polymerases, it still would occur at a slower rate.

2 Cytosine is chemically unstable. When it deaminates, it produces uracil. If uracil were a common component of DNA, the cell would have no way of knowing if a guanine-uracil base pair was there because of a DNA polymerase error (meaning either base could be the wrong one) or by deamination of cytosine (meaning the uracil is the wrong base). Due to the inability to distinguish these 2 possibilities, more mutations would make their way into DNA. Having thymine instead of uracil as a base means that the presence of uracil in DNA signals to the cell that it is the wrong base and needs to be removed, because it came from a deaminated cytosine.

LECTURE 29

1 This turns out to be an important consideration for the immune system. Once recombination has occurred in a cell to code for a specific antibody, it does not occur further, thus locking the cell into production of the same antibody as the cell divides over and over.

2 Because gene drives alter inheritance patterns and can be readily spread in a rapidly replicating population, it would be possible to create a species with a desired trait quickly by introducing it, as was described for mosquitoes in the lecture. In that case, the interest was in eliminating fertility, but it could just as easily be applied to using a desired trait.

LECTURE 30

1 A sigma factor used for heat shock response would be triggered only by heat shock and would activate the transcription of genes near promoters that are only activated by the heat shock sigma factor. These, for example, could include proteins that help the cell respond to the shock, such as those involved in protein folding or refolding. In the absence of the heat shock, no heat shock sigma factor would be made, and the associated proteins also would not be made.

2 In prokaryotic cells, transcription occurs at about the rate of translation, and this is important because the 2 processes occur simultaneously there. In eukaryotic cells, RNAs have much longer lifetimes and, once made, can be used for a long time, if desired or necessary. Their speed of synthesis is not as critical as the speed of DNA replication, which is needed for a cell to divide. Further, in both cases, mRNAs can be used over and over, so speed is not really needed in synthesis.

LECTURE 31

1 Errors in protein synthesis are not transmitted genetically, so the likelihood of a problem from a translation error isn't as great as that of DNA errors. Further, one protein with an error is unlikely to cause much in the way of problems. Because RNAs can be used repeatedly to make a protein in a cell, the more times a problem protein is made, the larger the potential impact might be.

2 phenylalanine = 1
leucine = 4
isoleucine = 1
methionine = 1
valine = 2
serine = 3
proline = 2
threonine = 2
alanine = 2
tyrosine = 1
histidine = 1
glutamine = 2
asparagine = 1
lysine = 2
aspartic acid = 1
glutamic acid = 2
cysteine = 1
tryptophan = 1
arginine = 4
glycine = 2

LECTURE 32

1 The lac promoter sequence is not close to the consensus sequence. If it were, you would predict that the effect of the CAP would not be seen or needed, because the promoter would always be very active whenever the lac repressor was not bound. Having CAP as the activator allows the cell to better control when the operon is transcribed.

2 Gene expression is controlled at several levels to ensure cells have the right proteins in the right amounts at the times they are needed. Synthesis of unneeded proteins wastes energy and may have undesired effects/products if a protein is made when it shouldn't be.

Consequently, there are systems for controlling the synthesis of proteins. These include the system described for the iron proteins, and RNA interference is yet another system. Breaking down an RNA using RNA

interference is the surest way to stop a protein from being made once the mRNA has been synthesized. Without RNA interference systems, an mRNA might hang around much longer than needed and exacerbate problems of unneeded proteins. RNA interference thus gives much more precise control of the timing of protein synthesis.

LECTURE 33

1 If the CFTR mutation were one or 2 bases instead of 3, the protein made would almost certainly not fold properly and would not function. The deletion of one or 2 bases changes the codon arrangement, whereas the deletion of 3 bases keeps the codon arrangement. For example, consider the following sequence:

AAG	ACA	GCC	UGG	UAC	GUG	CAU	(no deletion)
LYS	THR	ALA	TRP	TYR	VAL	HIS	

Deletion of the G yields

AAG	ACA	~~G~~CCU	GGU	ACG	UGC	AU	(1-base deletion)
LYS	THR	PRO	GLY	THR	CYS		

Deletion of the G and C yields

AAG	ACA	~~GC~~CUG	GUA	CGU	GCA	U	(2-base deletion)
LYS	THR	LEU	VAL	ARG	ALA		

Deletion of the G, C, and C yields

AAG	ACA	~~GCC~~UGG	UAC	GUG	CAU	(3-base deletion)
LYS	THR	TRP	TYR	VAL	HIS	

Note that the 3-base deletion keeps the amino acids after the deletion the same as the original sequence, whereas the deletion of one or 2 bases changes the codons after the deletion so that different amino acids are used to make the protein. This would continue for all codons after the point of the deletion. The deletion of one or 2 bases results in what is called a frame-shift mutation, because the codons change after the point of the deletion.

2 Because dividing cells start with multiple mitochondria, this is a different type of inheritance than occurs with nuclei, where gametes have one copy of each chromosome. Multiple mitochondria can sort in many ways, so it is unlikely that all mutant mitochondria will migrate to the same cell. Consequently, cell division of mutant mixes of mitochondria will favor continued mixes being passed down, unless there is a selective advantage/disadvantage to the mutant mitochondria.

Given that the mitochondrial disease described in the lecture inhibited the function of complex I of electron transport, this would be a serious mutation (as would most mitochondrial mutations), so it would be unlikely to eventually take over all mitochondria in a given cell, because it would likely lead to death of the cell.

LECTURE 34

1 To inactivate a gene and increase the chance of it becoming a tumor, you would pick a stop-signal gene, such as *p53*. With no stop signals, cells with damaged DNA would have a greater chance of avoiding apoptosis and would more likely mutate and progress to the metastatic stage.

2 This is a much harder question, and it doesn't really have an answer. You might be tempted, based on the previous answer, to say "pick a go-signal gene," but there are many go-signal genes, and picking one wouldn't stop the other ones from causing problems. If you had a cell that was cancerous and you knew that it arose from a go-signal gene mutation and you knew what the mutation was, you could try targeting that mutation. But by

the time you've detected it as a cancer, many other mutations may have occurred, so stopping that one mutation might have little effect. So, it is easier to design a way to create cancer than it is to find a way to treat it, at least using RNA interference.

LECTURE 35

1 The big difference between gene organization in prokaryotes versus eukaryotes is that prokaryotic genes are intact and do not require splicing to put the exon coding sequences together. Prokaryotes have no introns in their genes, so each coding region is one giant exon. As a result, prokaryotes have no need for, or ability to perform, the splicing that eukaryotic cells do to remove the introns they have in their genes.

Thus, if you take eukaryotic DNA coding for a protein and put it directly into bacteria, the protein that is made will include codons from the introns, which are not part of the eukaryotic protein, so the protein made will not be the same as the one made in eukaryotes. When you want bacteria to make eukaryotic proteins, you have to give them DNA that has had the introns removed from it; otherwise, the protein will not be correct.

2 A simple one would be to make an antibody against the nucleolin protein and inject it into the bloodstream. It could carry toxic substances that it injects into the target cell, for example. One problem that might arise is that the injected antibody might stimulate the body's own immune response, and the attempt would fail because the immune system would attack the nucleolin-directed antibody. Another problem could arise from accidental side effects of the toxic substance that was carried.

LECTURE 36

1 Probably the 2 biggest meanings that are lacking are how the sequence translates into levels of gene expression in different cells and how the mass of enzymes, proteins, and metabolites in the cell function from a system-wide perspective. There are undoubtedly many other considerations as well.

2 Cells have many means of controlling how much of a protein is made. This means that just because an RNA is synthesized, it says nothing about how much protein is made from it. This can be modified by many factors, including the half-life of mRNAs and proteins like the iron response element (IRE) binding protein that can block translation. In addition, proteins may be made, but their lifetimes vary as well, thanks to proteasomal degradation. As a result, to get the big picture of gene expression, it is necessary to track both the RNAs made and the proteins made from them.

Bibliography

Ahern, K., and I, Rajagopal. *Biochemistry Free and Easy*. DaVinci Press, 2012. http://www.davincipress.com/freeforall.html.

Ahern, K., I. Rajagopal, and T. Tan. *Biochemistry Free for All*. DaVinci Press, 2016. http://www.davincipress.com/freeforall.html.

Al-Karadaghi, S. "Protein 3D Structure: Structural Levels, Motifs and Folds." *Structural Bioinformatics: Practical Guide*. https://proteinstructures.com/Structure/protein-structure1.html.

American Society for Biochemistry and Molecular Biology. "A Spring-Loaded Sensor for Cholesterol in Cells." December 7, 2017. https://www.eurekalert.org/pub_releases/2017-12/asfb-ass120717.php.

Armstrong, S. *Borrowed Time: The Science of How and Why We Age*. Bloomsbury Sigma, 2019.

———. p53*, the Gene That Cracked the Cancer Code*. Bloomsbury Sigma, 2016.

Arney, K. *Herding Hemingway's Cats*. Bloomsbury Sigma, 2016.

Ball, P. *Water*. Orion Publishing Co, 2000.

Blackburn, E., and E. Epel. *The Telomere Effect*. Grand Central Publishing, 2018.

Bloom, F. E. *Best of the Brain from* Scientific American. Dana Press, 2007.

Branzei, D., and M. Foiani. "Regulation of DNA Repair throughout the Cell Cycle." *Nature Reviews in Molecular Cell Biology* 9 (2008): 297–306.

Bröer, S., and A. Bröer. "Amino Acid Homeostasis and Signaling in Mammalian Cells and Organisms." *Biochemical Journal* 474 (2017): 1935–1963.

Brown, G. *The Energy of Life*. The Free Press, 2000.

Calladine, C., and H. Drew. *Understanding DNA*. Academic Press, 2004.

Carey, N. *The Epigenetics Revolution: How Modern Biology Is Rewriting Our Understanding of Genetics, Disease, and Inheritance*. Columbia University Press, 2011.

Chandel, N. S. *Navigating Metabolism*. Cold Spring Harbor Laboratory Press, 2014.

Church, G., and E. Regis. *Regenesis*. Basic Books, 2014.

Cobb, M. *Life's Greatest Secret: The Race to Crack the Genetic Code*. Profile Books, 2015.

Craig, N., et al. *Molecular Biology*. 2nd ed. Oxford University Press, 2014.

Crow, J. M. "Go with the Fold." *Chemistry World*. https://www.chemistryworld.com/features/go-with-the-fold/3008748.article.

Davies, J. A. *Synthetic Biology: A Very Short Introduction*. Oxford University Press, 2018.

DeSalle, R., and P. Wynne. *Our Senses: An Immersive Experience*. Yale University Press, 2018.

Doudna, J., and S. H. Sternberg. *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution*. Houghton Mifflin Harcourt, 2017.

Dudley, J., et al. *Exploring Personal Genomics*. Oxford University Press, 2013.

Erickson, C. *The Science of Addiction*. W. W. Norton and Co., 2007.

Field, D., and N. Davies. *Biocode: The New Age of Genomics*. Oxford University Press, 2015.

Foldit: Solve Puzzles for Science. https://fold.it/portal/info/science.

Genetic Science Learning Center. "Spotlight on Sugar." *Learn.Genetics*. September 1, 2015. https://learn.genetics.utah.edu/content/metabolism/sugar/.

Gillham, N. *Genes, Chromosomes and Disease*. FT Press Science, 2011.

Gluckman, P., and M. Hanson. *Mismatch: The Time Bomb of Lifestyle Disease*. Oxford University Press, 2008.

Goodsell, D. *Our Molecular Nature*. Springer Science, 2012.

Graeber, C. *The Breakthrough: Immunotherapy and the Race to Cure Cancer*. Twelve, 2018.

Gray, T. *Molecules*. Blackdog and Leventhal, 2018.

———. *Reactions*. Blackdog and Leventhal, 2018.

Gutiérrez-Preciado, A., H. Romero, and M. Peimbert. "An Evolutionary Perspective on Amino Acids." *Nature Education* 3, no. 9 (2010): 29.

Hancock, J. *Cell Signaling*. Oxford University Press, 2017.

Henshaw, J. M. *A Tour of the Senses*. Johns Hopkins Press, 2012.

Jones, P. *The Genetic Code*. Chelsea House, 2010.

Jonsson, A., et al. "Essential Chemistry for Biochemists." *Essays in Biochemistry* 61 (2017): 401–427.

Knoepfler, P. *Stem Cells*. World Scientific, 2013.

Kohli, A. G., et al. "Designer Lipids for Drug Delivery." *Journal of Controlled Release* 190 (2014): 274–278.

Kolata, G. *Mercies in Disguise*. St. Martin's Press, 2017.

Kornberg, A. *For the Love of Enzymes*. Harvard University Press, 1989.

Lane, N. *Oxygen: The Molecule That Made the World*. Oxford University Press, 2002.

———. *Power, Sex, Suicide: Mitochondria and the Meaning of Life*. Oxford University Press, 2005.

———. *The Vital Question: Why Is Life the Way It Is?* Oxford University Press, 2015.

Latchman, D. *Gene Control*. Garland Science, 2010.

Leroi, A. M. *Mutants*. Penguin Random House, 2005.

Linden, D. *The Compass of Pleasure*. Viking Penguin, 2011.

Milo, R., and R. Phillips. *Cell Biology by the Numbers*. Garland Science, 2015.

Morton, O. *Eating the Sun: How Plants Power the Planet*. Harper Collins, 2007.

Mukherjee, S. *The Emperor of All Maladies*. Scribner 2011.

———. *The Gene: An Intimate History*. Scribner, 2017.

Neitzel, J. J. "Enzyme Catalysis: The Serine Proteases." *Nature Education* 3, no. 9 (2010): 21.

Niehoff, D. *The Language of Life: How Cells Communicate in Health and Disease*. Joseph Henry Press, 2005.

OpenStax. "Carbohydrate Metabolism." *Anatomy and Physiology*. Rice University. https://opentextbc.ca/anatomyandphysiology/chapter/24-2-carbohydrate-metabolism/.

Pond, C. *The Fats of Life*. Cambridge University Press, 2010.

Pross, A. *What Is Life? How Chemistry becomes Biology*. Oxford University Press, 2012.

Quackenbush, J. *The Human Genome: Book of Essential Knowledge*. Random House, 2011.

Ralston, A. "Environmental Mutagens, Cell Signaling and DNA Repair." *Nature Education* 1, no. 1 (2008): 114.

Robinson, P. K. "Enzymes: Principles and Biotechnological Applications." *Essays in Biochemistry* 59 (2015): 1–41.

Rogers, A. *The Science of Booze*. Houghton Mifflin Harcourt, 2014.

Rosenblum, L. *See What I'm Saying?* W. W. Norton and Co., 2010.

Royal Society of Chemistry. "Photosynthesis." *Chemistry for Biologists*. https://www.rsc.org/Education/Teachers/Resources/cfb/Photosynthesis.htm.

Rutherford, A. *A Brief History of Everyone Who Ever Lived*. Harvard University Press, 2017.

Samuelsson, T. *The Human Genome in Health and Disease*. CRC Press, 2019.

Shapiro, B. *How to Clone a Mammoth*. Princeton University Press, 2015.

Slack, J. *Genes: A Very Short Introduction*. Oxford University Press, 2014.

Snape, A., et al. *Biochemistry and Molecular Biology*. Oxford University Press, 2018.

Stockwell, B. *Quest for the Cure*. Columbia University Press, 2017.

Storz, J. F. *Hemoglobin: Insights into Protein Structure, Function and Evolution*. Oxford University Press, 2018.

Sumner, J. *The Natural History of Medicinal Plants*. Timber Press, 2008.

SWISS-MODEL. "Part 1: Introduction to Protein Structure." *Principles of Protein Structure, Comparative Protein Modeling and Visualisation*. https://swissmodel.expasy.org/course.

Thom, C. S., et al. "Hemoglobin Variants: Biochemical Properties and Clinical Correlates." *Cold Spring Harbor Perspectives in Medicine* 3, no. 3 (2013).

van Holde, K., and J. Zlatanova. *The Evolution of Molecular Biology*. Academic Press, 2018.

Voet, D., J. Voet, and C. Pratt. *Fundamentals of Biochemistry*. 4th ed. Wiley, 2013.

von Cammaerer, S., et al. "C_4 Photosynthesis: 50 Years of Discovery and Innovation." *Journal of Experimental Botany* 68 (2017): 97–102.

Widmaier, E. *Why Geese Don't Get Obese (and We Do)*. W. H. Freeman, 1998.

Witkowski, J., ed. *The Inside Story: DNA to RNA to Protein*. Cold Spring Harbor Laboratory Press, 2005.

Yang, T., et al. "Cellular Cholesterol Homeostasis and Alzheimer's Disease." *Journal of Lipid Research* 58, no. 12 (2017): 2239–2254.

Yarus, M. *Life from an RNA World*. Harvard University Press, 2011.

Yong, E. *I Contain Multitudes*. Ecco Press, 2016.

Image Credits

26: Muffet/Flickr/CC BY 2.0; 48: Ramin Herati/Wikimedia Commons/Public Domain; 101: karandaev/Getty Images; 193: luza studios/iStock / Getty Images Plus; 218: ooyoo/E+/Getty Images; 242: FredFroese/iStock/Getty Images Plus; 249: Oleh_Slobodeniuk/Getty Images; 257: Roland1952/Wikimedia Commons/CC BY-SA 3.0; 294: ttsz/Getty Images; 299: Hannes Lans, et al., BioMed Central/Wikimedia Commons/CC BY 2.0; 300: Fvasconcellos/Wikimedia Commons/Public Domain; 318: Yikrazuul/Wikimedia Commons/CC BY-SA 3.0; 343: Designua/Shutterstock.com; 362: Squidonius/Wikimedia Commons/Public Domain