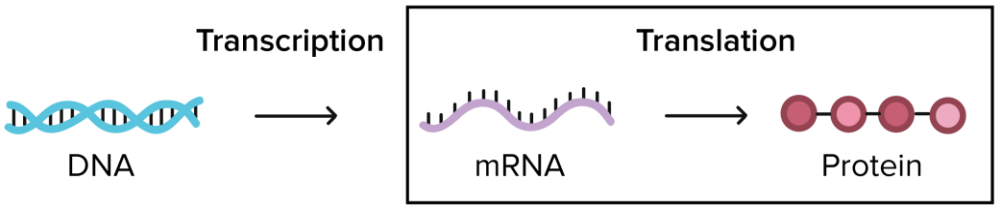
Overview of translation

Basically, a gene is used to build a protein in a two-step process:

Step 1: transcription! Here, the DNA sequence of a gene is "rewritten" in the form of RNA. In eukaryotes like you and me, the RNA is processed (and often has a few bits snipped out of it) to make the final product, called a messenger RNA or mRNA.

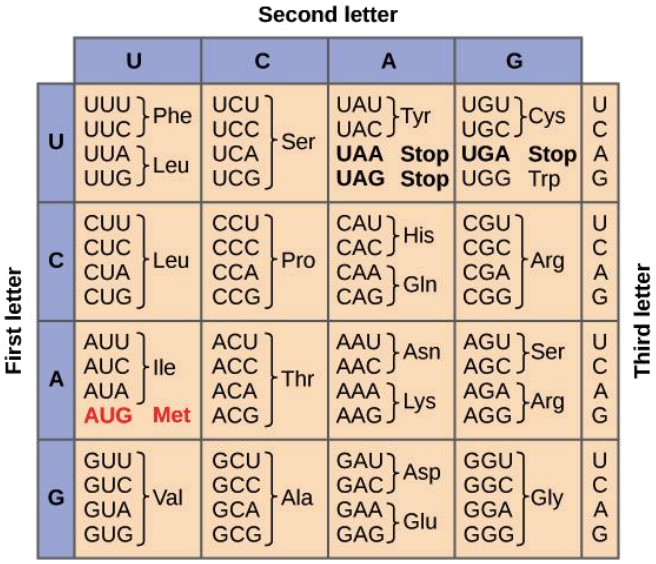
Step 2: translation! In this stage, the mRNA is "decoded" to build a protein (or a chunk/subunit of a protein) that contains a specific series of amino acids. [What exactly is an "amino acid"?]



"Central dogma of molecular biochemistry with enzymes(Opens in a new window),"

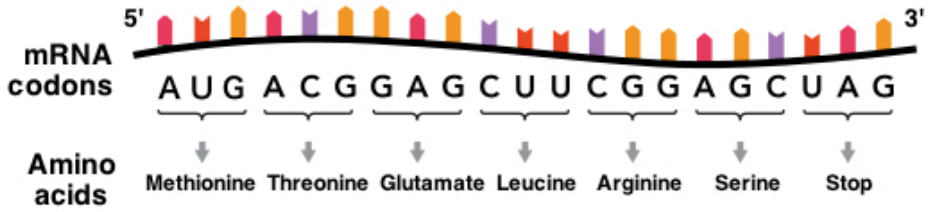
The genetic code

During translation, a cell “reads” the information in a messenger RNA (mRNA) and uses it to build a protein. Actually, to be a little more techical, an mRNA doesn’t always encode—provide instructions for—a whole protein. Instead, what we can confidently say is that it always encodes a polypeptide, or chain of amino acids.



In an mRNA, the instructions for building a polypeptide are RNA nucleotides (As, Us, Cs, and Gs) read in groups of three. These groups of three are called codons. There are 616161 codons

for amino acids, and each of them is "read" to specify a certain amino acid out of the 202020 commonly found in proteins. One codon, AUG, specifies the amino acid methionine and also acts as a start codon to signal the start of protein construction. There are three more codons that do not specify amino acids. These stop codons, UAA, UAG, and UGA, tell the cell when a polypeptide is complete. All together, this collection of codon-amino acid relationships is called the genetic code, because it lets cells “decode” an mRNA into a chain of amino acids.

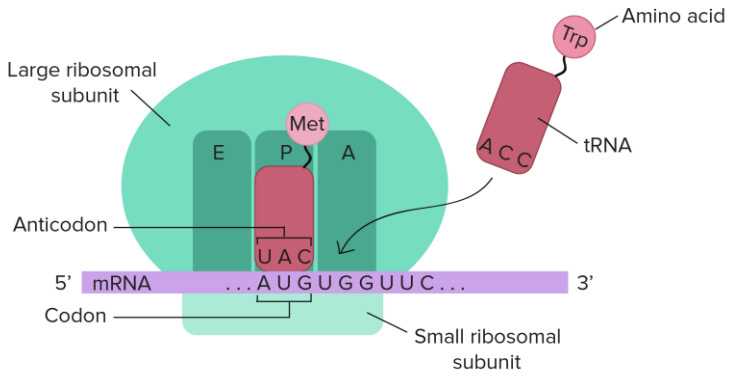


Overview of translation

How is an mRNA "read" to make a polypeptide? Two types of molecules with key roles in translation are tRNAs and ribosomes.

Transfer RNAs (tRNAs)

Transfer RNAs, or tRNAs, are molecular "bridges" that connect mRNA codons to the amino acids they encode. One end of each tRNA has a sequence of three nucleotides called an anticodon, which can bind to specific mRNA codons. The other end of the tRNA carries the amino acid specified by the codons. There are many different types of tRNAs. Each type reads one or a few codons and brings the right amino acid matching those codons.



Ribosomes

Ribosomes are the structures where polypeptides (proteins) are built. They are made up of protein and RNA (ribosomal RNA, or rRNA). Each ribosome has two subunits, a large one and a small one, which come together around an mRNA—kind of like the two halves of a hamburger bun coming together around the patty. The ribosome provides a set of handy slots where tRNAs

can find their matching codons on the mRNA template and deliver their amino acids. These slots are called the A, P, and E sites. Not only that, but the ribosome also acts as an enzyme, catalyzing the chemical reaction that links amino acids together to make a chain. Want to learn more about the structure and function of tRNAs and ribosomes? Check out the tRNA and ribosomes article!

Steps of translation

Your cells are making new proteins every second of the day. And each of those proteins must contain the right set of amino acids, linked together in just the right order. That may sound like a challenging task, but luckily, your cells (along with those of other animals, plants, and bacteria) are up to the job. To see how cells make proteins, let's divide translation into three stages: initiation (starting off), elongation (adding on to the protein chain), and termination (finishing up).

Getting started: Initiation

In initiation, the ribosome assembles around the mRNA to be read and the first tRNA (carrying the amino acid methionine, which matches the start codon, AUG). This setup, called the initiation complex, is needed in order for translation to get started.

Extending the chain: Elongation

Elongation is the stage where the amino acid chain gets **long**er. In elongation, the mRNA is read one codon at a time, and the amino acid matching each codon is added to a growing protein chain.

Each time a new codon is exposed:

- A matching tRNA binds to the codon
- The existing amino acid chain (polypeptide) is linked onto the amino acid of the tRNA via a chemical reaction
- The mRNA is shifted one codon over in the ribosome, exposing a new codon for reading

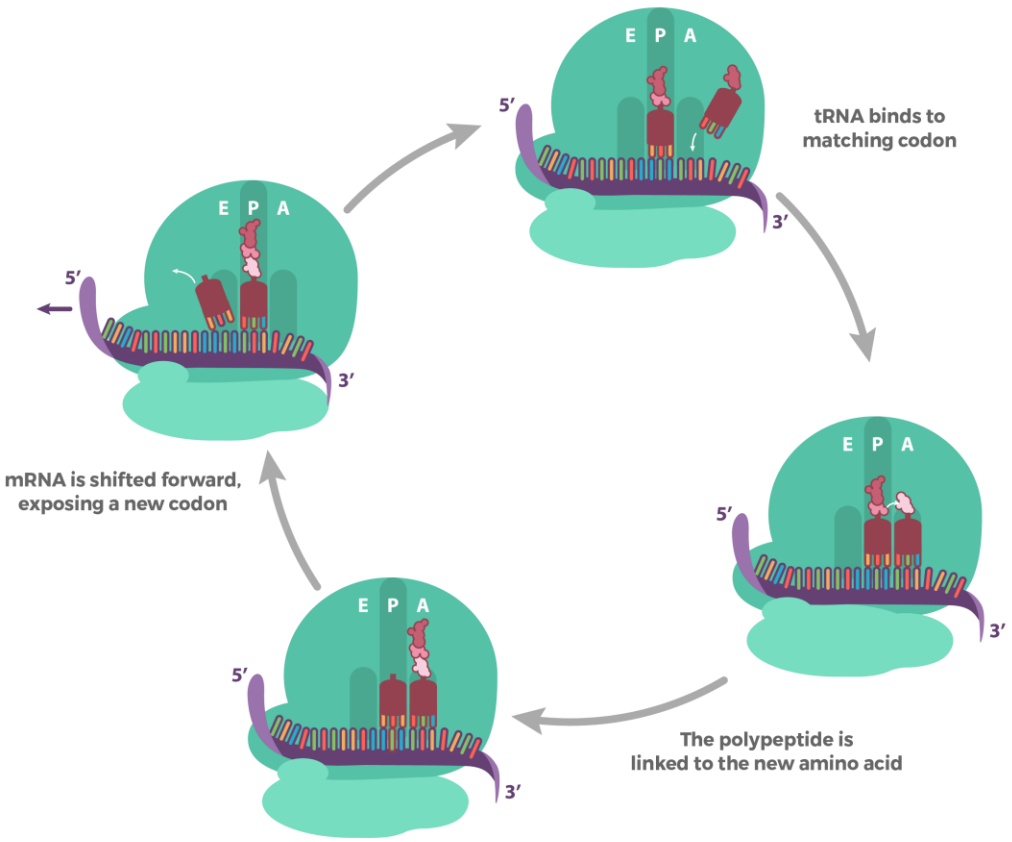


Image based on similar diagram in Reece et al.

During elongation, tRNAs move through the A, P, and E sites of the ribosome, as shown above. This process repeats many times as new codons are read and new amino acids are added to the chain.

For more details on the steps of elongation, see the stages of translation article.

Finishing up: Termination

Termination is the stage in which the finished polypeptide chain is released. It begins when a stop codon (UAG, UAA, or UGA) enters the ribosome, triggering a series of events that separate the chain from its tRNA and allow it to drift out of the ribosome. After termination, the polypeptide may still need to fold into the right 3D shape, undergo processing (such as the removal of amino acids), get shipped to the right place in the cell, or combine with other polypeptides before it can do its job as a functional protein.

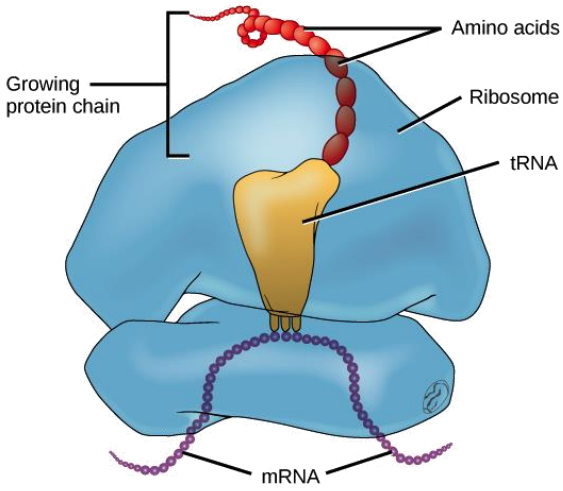
tRNAs and ribosomes

Introduction

Translation requires some specialized equipment. Just as you wouldn't go to play tennis without your racket and ball, so a cell couldn't translate an mRNA into a protein without two pieces of molecular gear: ribosomes and tRNAs. Ribosomes provide a structure in which translation can take place. They also catalyze the reaction that links amino acids to make a new protein. tRNAs (transfer RNAs) carry amino acids to the ribosome. They act as "bridges," matching a codon in an mRNA with the amino acid it codes for. Here, we'll take a closer look at ribosomes and tRNAs. If you're not yet familiar with RNA (which stands for ribonucleic acid), I highly recommend checking out the nucleic acids section first so you can get the most out of this article!

Ribosomes:

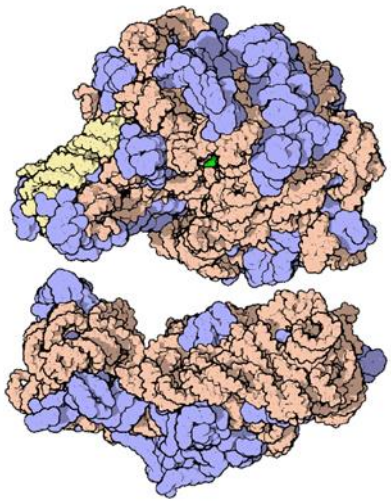
Translation takes place inside structures called ribosomes, which are made of RNA and protein. Ribosomes organize translation and catalyze the reaction that joins amino acids to make a protein chain.



Structure of the ribosome

A ribosome is made up of two basic pieces: a large and a small subunit. During translation, the two subunits come together around a mRNA molecule, forming a complete ribosome. The ribosome moves forward on the mRNA, codon by codon, as it is read and translated into a polypeptide (protein chain). Then, once translation is finished, the two pieces come apart again and can be reused. Overall, the ribosome is about one-third protein and two-thirds ribosomal RNA (rRNA). The rRNAs seem to be responsible for most of the structure and function of the ribosome, while the proteins help the rRNAs change shape as they catalyze chemical reactions.

Below, you can see a 3D model of the ribosome. Proteins are colored in blue, while strands of rRNA are colored in tan and orange. The green spot marks the active site, which catalyzes the reaction that links amino acids to make a protein. It surprised me to see that the ribosome is wrinkly, kind of like the surface of a brain!

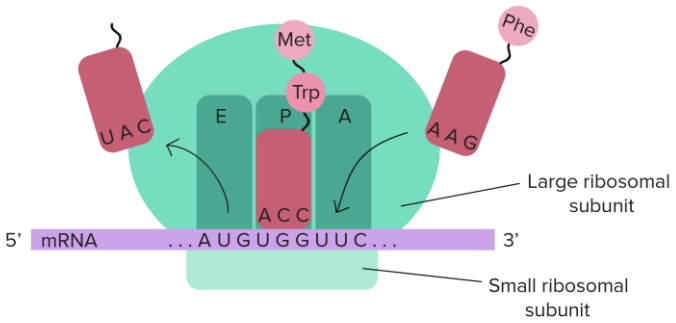


Ribosome," by Redondoself (CC BY 2.0(Opens in a new window)).

The ribosome has slots for tRNAs

As we saw briefly in the introduction, molecules called transfer RNAs (tRNAs) bring amino acids to the ribosome. We'll learn a lot more about tRNAs and how they work in the next section.

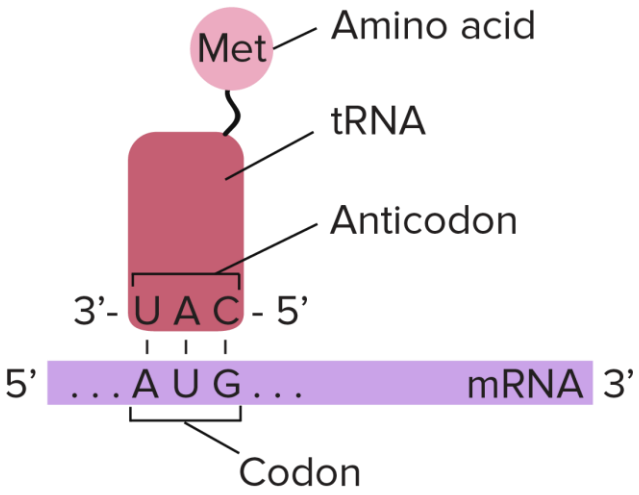
For now, just keep in mind that the ribosome has three slots for tRNAs: the A site, P site, and E site. tRNAs move through these sites (from A to P to E) as they deliver amino acids during translation.



What exactly is a tRNA?

A transfer RNA (tRNA) is a special kind of RNA molecule. Its job is to match an mRNA codon with the amino acid it codes for. You can think of it as a kind of molecular "bridge" between the two.

Each tRNA contains a set of three nucleotides called an anticodon. The anticodon of a given tRNA can bind to one or a few specific mRNA codons. The tRNA molecule also carries an amino acid: specifically, the one encoded by the codons that the tRNA binds.



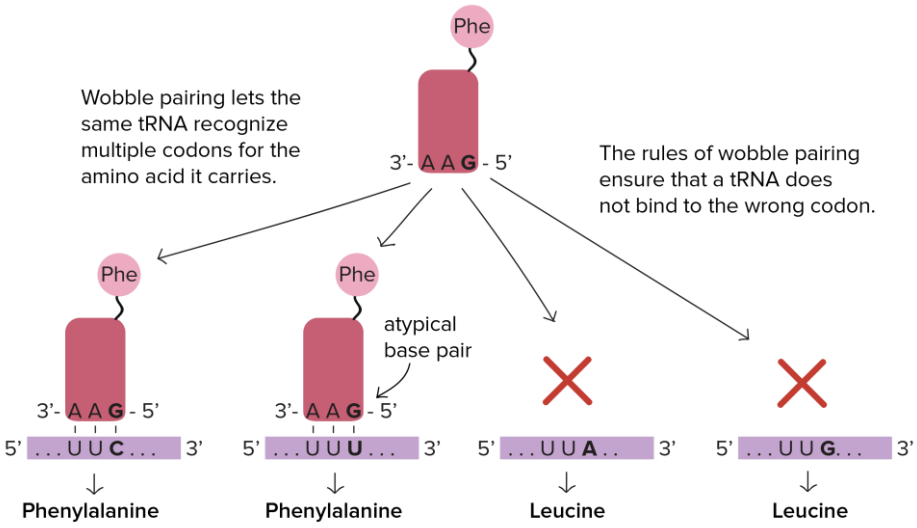
There are many different types of tRNAs floating around in a cell, each with its own anticodon and matching amino acid. In fact, there are usually 404040 to 606060 different types, depending on the species cubed. tRNAs bind to codons inside of the ribosome, where they deliver amino acids for addition to the protein chain.

Some tRNAs bind to multiple codons ("wobble")

Some tRNAs can form base pairs with more than one codon. At first, this seems pretty weird: doesn't A base-pair with U, and G with C?

Well...not always. (Biology is full of surprises, isn't it?) Atypical base pairs—between nucleotides other than A-U and G-C—can form at the third position of the codon, a phenomenon known as wobble.

Wobble pairing doesn't follow normal rules, but it does have its own rules. For instance, a G in the anticodon can pair with a C or U (but not an A or G) in the third position of the codon, as shown below. Rules like this ensure codons are read correctly despite wobble.



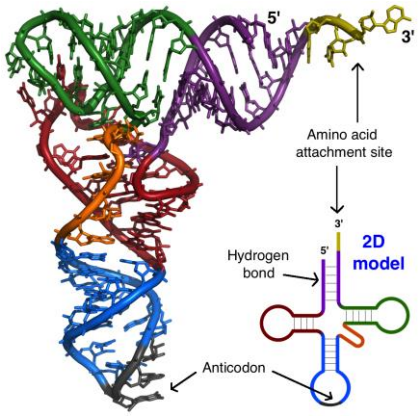
You may be wondering: why on Earth would a cell "want" a complicating factor like wobble? The answer may be that wobble pairing allows fewer tRNAs to cover all the codons of the genetic code, while still making sure that the code is read accurately.

The 3D structure of a tRNA

I like to draw tRNAs as little rectangles, to make it clear what's going on (and to have plenty of room to fit the letters of the anticodon on there). But a real tRNA actually has a much more interesting shape, one that helps it do its job.

A tRNA, like the one modeled below, is made from a single strand of RNA (just like an mRNA is). However, the strand takes on a complex 3D structure because base pairs form between nucleotides in different parts of the molecule. This makes double-stranded regions and loops,

folding the tRNA into an L shape.

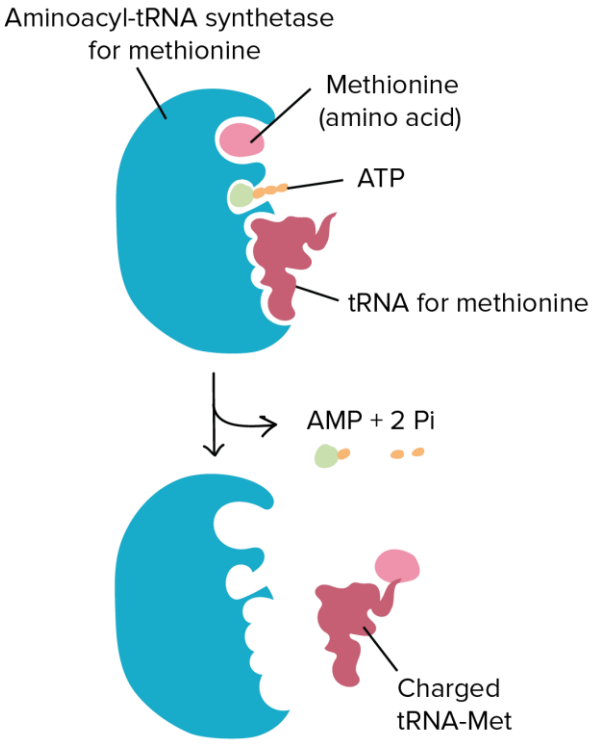


One end of the L shape has the anticodon, while the other has the attachment site for the amino acid. Different tRNAs have slightly different structures, and this is important for making sure they get loaded up with the right amino acid.

Loading a tRNA with an amino acid

How does the right amino acid get linked to the right tRNA (making sure that codons are read correctly)? Enzymes called aminoacyl-tRNA synthetases have this very important job.

There's a different synthetase enzyme for each amino acid, one that recognizes only that amino acid and its tRNAs (and no others). Once both the amino acid and its tRNA have attached to the enzyme, the enzyme links them together, in a reaction fueled by the "energy currency" molecule adenosine triphosphate (ATP).



Occasionally, an aminoacyl-tRNA synthetase makes a mistake: it binds to the wrong amino acid (one that "looks similar" to its correct target). For example, the threonine synthetase sometimes grabs serine by accident and attaches it to the threonine tRNA. Luckily, the threonine synthetase has a proofreading site, which pops the amino acid back off the tRNA if it's incorrect.

Stages of translation

Translation: Translation has pretty much the same three parts, but they have fancier names: initiation, elongation, and termination.

Initiation ("beginning"): in this stage, the ribosome gets together with the mRNA and the first tRNA so translation can begin.

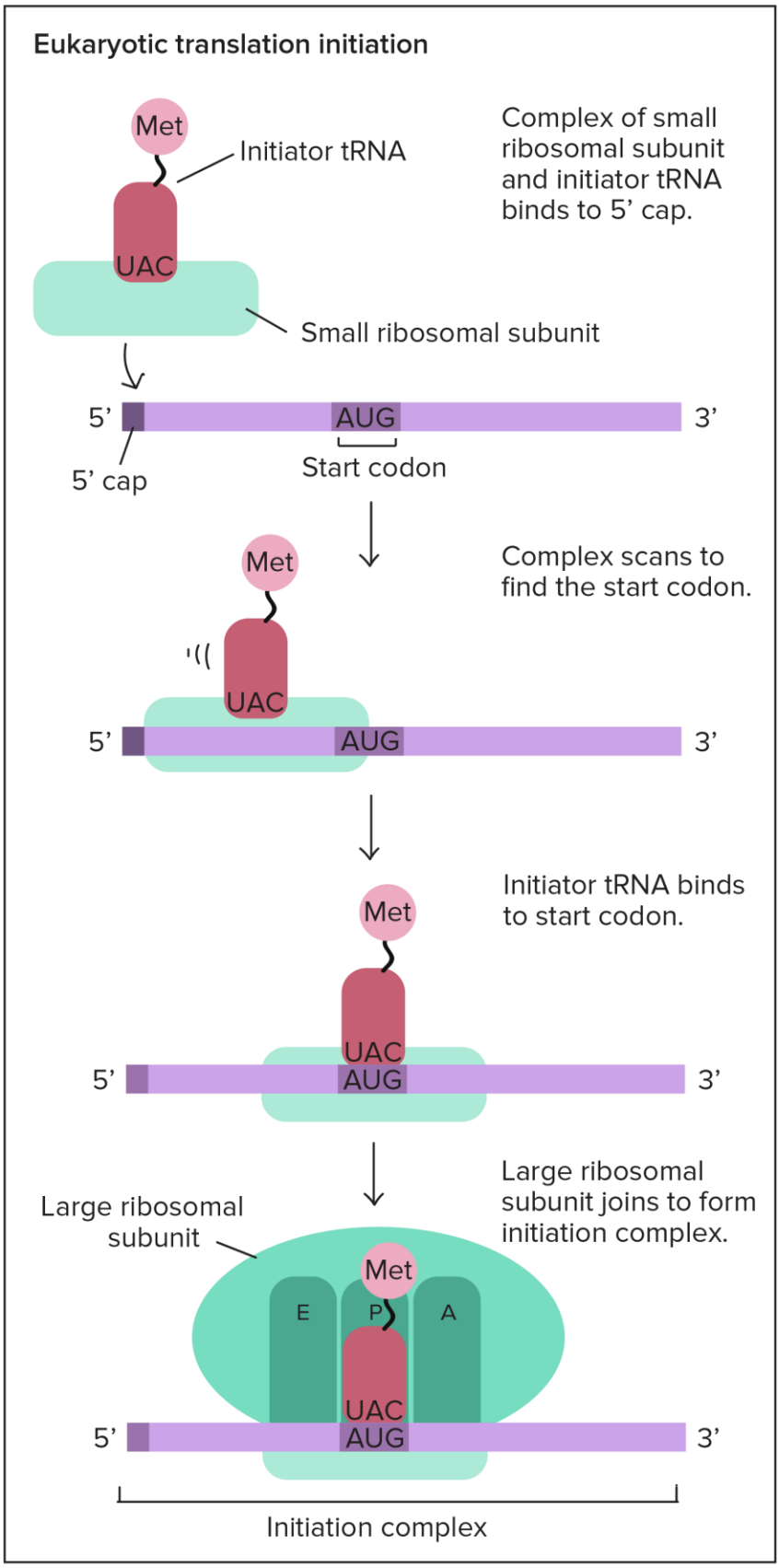
Elongation ("middle"): in this stage, amino acids are brought to the ribosome by tRNAs and linked together to form a chain.

Termination ("end"): in the last stage, the finished polypeptide is released to go and do its job in the cell.

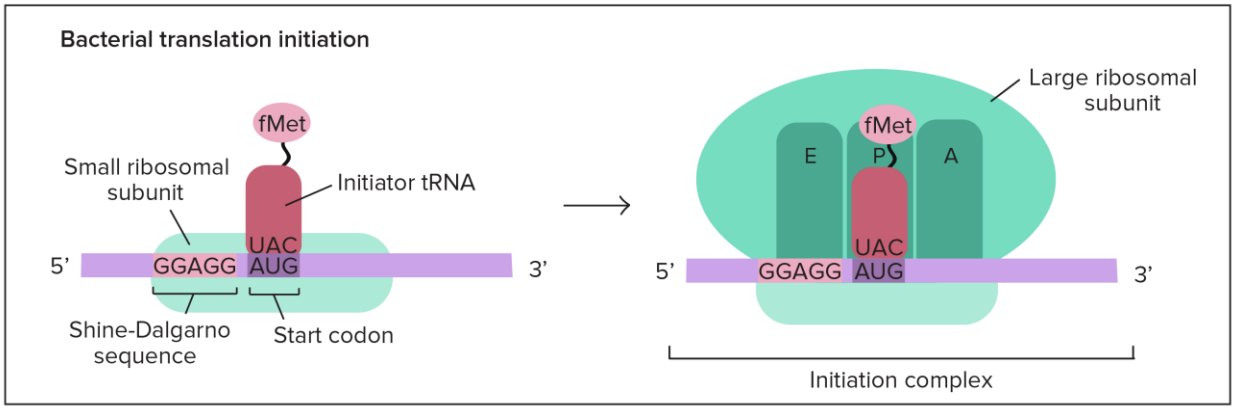
Initiation

In order for translation to start, we need a few key ingredients. These include: A ribosome (which comes in two pieces, large and small). An mRNA with instructions for the protein we'll build. An "initiator" tRNA carrying the first amino acid in the protein, which is almost always methionine (Met). During initiation, these pieces must come together in just the right way. Together, they form the initiation complex, the molecular setup needed to start making a new protein.

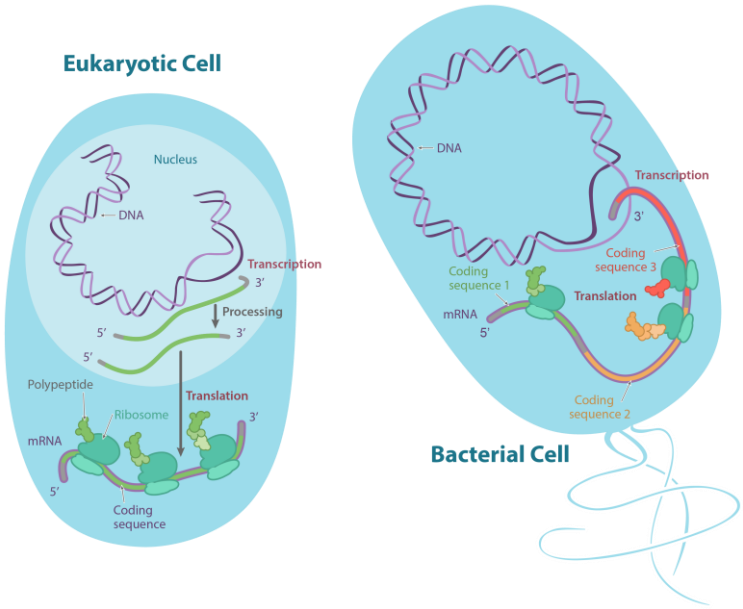
Inside your cells (and the cells of other eukaryotes), translation initiation goes like this: first, the tRNA carrying methionine attaches to the small ribosomal subunit. Together, they bind to the 5' end of the mRNA by recognizing the 5' GTP cap (added during processing in the nucleus). Then, they "walk" along the mRNA in the 3' direction, stopping when they reach the start codon (often, but not always, the first AUG).


Eukaryotic translation initiation
Met
Initiator tRNA
UAC
Small ribosomal subunit
Complex of small ribosomal subunit and initiator tRNA binds to 5' cap.
5'
AUG
3'
5' cap
Start codon
Met
UAC
5'
AUG
3'
Complex scans to find the start codon.
Met
UAC
AUG
5'
3'
Initiator tRNA binds to start codon.
Large ribosomal subunit
Met
E
P
A
UAC
AUG
5'
3'
Large ribosomal subunit joins to form initiation complex.
Initiation complex

In bacteria, the situation is a little different. Here, the small ribosomal subunit doesn't start at the 5' end of the mRNA and travel toward the 3' end. Instead, it attaches directly to certain sequences in the mRNA. These Shine-Dalgarno sequences come just before start codons and "point them out" to the ribosome.



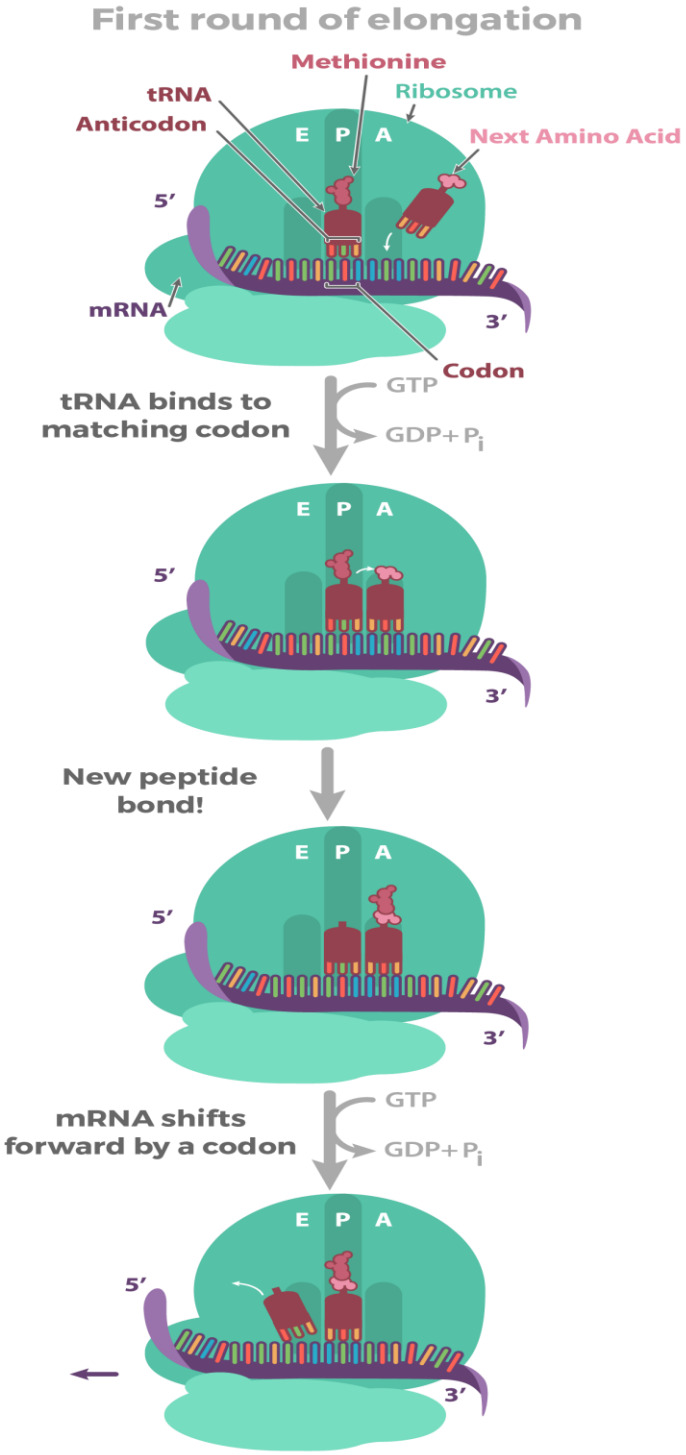
Why use Shine-Dalgarno sequences? Bacterial genes are often transcribed in groups (called operons), so one bacterial mRNA can contain the coding sequences for several genes. A Shine-Dalgarno sequence marks the start of each coding sequence, letting the ribosome find the right start codon for each gene.



Elongation

I like to remember what happens in this "middle" stage of translation by its handy name: elongation is when the polypeptide chain gets longer. But how does the chain actually grow? To

find out, let's take a look at the first round of elongation—after the initiation complex has formed, but before any amino acids have been linked to make a chain. Our first, methionine-carrying tRNA starts out in the middle slot of the ribosome, called the P site. Next to it, a fresh codon is exposed in another slot, called the A site. The A site will be the "landing site" for the next tRNA, one whose anticodon is a perfect (complementary) match for the exposed codon.



Once the matching tRNA has landed in the A site, it's time for the action: that is, the formation of the peptide bond that connects one amino acid to another. This step transfers the methionine from the first tRNA onto the amino acid of the second tRNA in the A site. Not bad—we now have two amino acids, a (very tiny) polypeptide! The methionine forms the N-terminus of the polypeptide, and the other amino acid is the C-terminus.

Termination

Polypeptides, like all good things, must eventually come to an end. Translation ends in a process called termination. Termination happens when a stop codon in the mRNA (UAA, UAG, or UGA) enters the A site. Stop codons are recognized by proteins called release factors, which fit neatly into the P site (though they aren't tRNAs). Release factors mess with the enzyme that normally forms peptide bonds: they make it add a water molecule to the last amino acid of the chain. This reaction separates the chain from the tRNA, and the newly made protein is released.

Epilogue: Processing

Our polypeptide now has all its amino acids—does that mean it's ready to do its job in the cell?

Not necessarily. Polypeptides often need some "edits." During and after translation, amino acids may be chemically altered or removed. The new polypeptide will also fold into a distinct 3D structure, and may join with other polypeptides to make a multi-part protein.

Many proteins are good at folding on their own, but some need helpers ("chaperones") to keep them from sticking together incorrectly during the complex process of folding.

Some proteins also contain special amino acid sequences that direct them to certain parts of the cell. These sequences, often found close to the N- or C-terminus, can be thought of as the protein’s “train ticket” to its final destination. For more about how this works, see the article on protein targeting.

Protein targeting

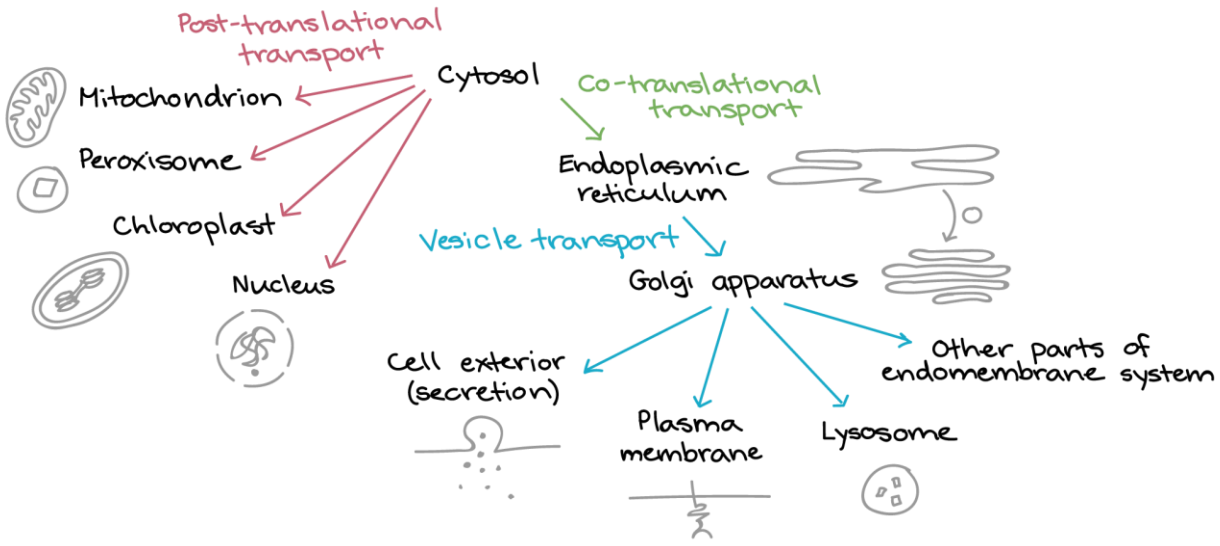
Introduction

Different proteins need to be sent to different parts of a eukaryotic cell, or, in some cases, exported out of the cell and into the extracellular space. How do the right proteins get to the right places?

Cells have various shipping systems, kind of like molecular versions of the postal service, to make sure that proteins arrive at their correct destinations. In these systems, molecular labels (often, amino acid sequences) are used to "address" proteins for delivery to specific locations. Let’s take a look at how these shipping systems work.

Overview of cellular shipping routes

Translation of all proteins in a eukaryotic cell begins in the cytosol (except for a few proteins made in mitochondria and chloroplasts). As a protein is made, it passes step by step through a shipping "decision tree." At each stage, the protein is checked for molecular tags to see if it needs to be re-routed to a different pathway or destination.



The first major branch point comes shortly after translation starts. At this point, the protein will either remain in the cytosol for the rest of translation, or be fed into the endoplasmic reticulum (ER) as it is translated.

Proteins are fed into the ER during translation if they have an amino sequence called a signal peptide. In general, proteins bound for organelles in the endomembrane system (such as the ER, Golgi apparatus, and lysosome) or for the exterior of the cell must enter the ER at this stage.

Proteins that do not have a signal peptide stay in the cytosol for the rest of translation. If they lack other "address labels," they'll stay in the cytosol permanently. However, if they have the right labels, they can be sent to the mitochondria, chloroplasts, peroxisomes, or nucleus after translation.

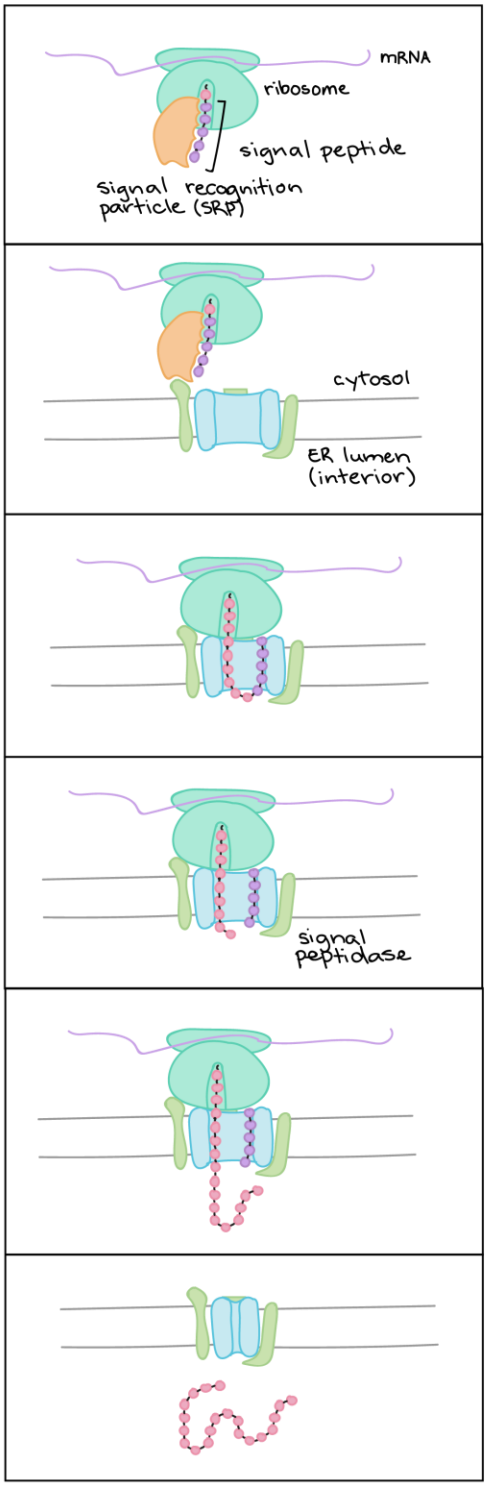
The endomembrane system and secretory pathway

Proteins destined for any part of the endomembrane system (or the outside of the cell) are brought to the ER during translation and fed in as they're made.

Signal peptides

The signal peptide that sends a protein into the endoplasmic reticulum during translation is a series of hydrophobic (“water-fearing”) amino acids, usually found near the beginning (N-

terminus) of the protein. When this sequence sticks out of the ribosome, it's recognized by a protein complex called the signal-recognition particle (SRP), which takes the ribosome to the ER. There, the ribosome feeds its amino acid chain into the ER lumen (interior) as it's made.



Signal recognition particle (SRP) binds to the signal peptide as it emerges from the ribosome.

SRP brings the ribosome to the ER by binding to a receptor on the ER surface. The receptor is associated with other proteins that make a pore.

The ribosome resumes translating, feeding the polypeptide though the pore and into the ER lumen (interior).

An enzyme associated with the pore (signal peptidase) snips off the signal peptide.

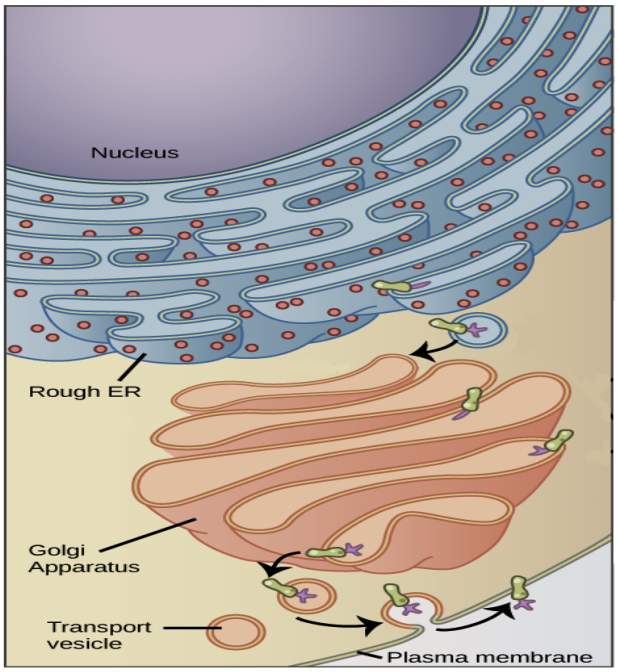
Translation continues, and the growing amino acid chain slides into the ER lumen.

The completed polypeptide is released into the ER lumen, where it floats freely.

In some cases, the signal peptide is snipped off during translation and the finished protein is released into the interior of the ER (as shown above). In other cases, the signal peptide or another stretch of hydrophobic amino acids gets embedded in the ER membrane. This creates a transmembrane (membrane-crossing) segment that anchors the protein to the membrane.

Transport through the endomembrane system

In the ER, proteins fold into their correct shapes, and may also get sugar groups attached to them. Most proteins are then transported to the Golgi apparatus in membrane vesicles. Some proteins, however, need to stay in the ER and do their jobs there. These proteins have amino acid tags that ensure they are shipped back to the ER if they "escape" into the Golgi.



In the Golgi apparatus, proteins may undergo more modifications (such as addition of sugar groups) and before going on to their final destinations. These destinations include lysosomes, the plasma membrane, and the cell exterior. Some proteins need to do their jobs in the Golgi (are "Golgi-resident), and a variety of molecular signals, including amino acid tags and structural features, are used to keep them there or bring them back cubed.

If they don't have any specific tags, proteins are sent from the Golgi to the cell surface, where they’re secreted to the cell exterior (if they’re free-floating) or delivered to the plasma membrane (if they’re membrane-embedded). This default pathway is shown in the diagram above for a membrane protein, colored in green, that bears sugar groups, colored in purple.

Proteins are shipped to other destinations if they contain the right molecular labels. For example, proteins destined for the lysosome have a molecular tag consisting of a sugar with a phosphate

group attached. In the Golgi apparatus, proteins with this tag are sorted into vesicles bound for the lysosome.

Targeting to non-endomembrane organelles

Proteins that are made in the cytosol (don't enter ER during translation) may stay permanently in the cytosol. However, they may also be shipped to other, non-endomembrane destinations in the cell. For instance, proteins bound for the mitochondria, chloroplasts, peroxisomes, and nucleus are usually made in the cytosol and delivered after translation is complete.

[**Don't mitochondria and chloroplasts have their own ribosomes?]**

To be delivered to one of these organelles after translation, a protein must contain a specific amino acid "address label." The label is recognized by other proteins in the cell, which help transport the protein to the right destination.

As an example, let's consider delivery to the peroxisome, an organelle involved in detoxification. Proteins needed in the peroxisome have a specific sequence of amino acids called a peroxisomal targeting signal. The classic signal consists of just three amino acids, serine-lysine-leucine, found at the very end (C-terminus) of a protein. This pattern of amino acids is recognized by a helper protein in the cytosol, which brings the protein to the peroxisome.

Mitochondrial, chloroplast, and nuclear targeting are generally similar to peroxisomal targeting. That is, a certain amino acid sequence sends the protein to its target organelle (or a compartment inside that organelle). However, the nature of the "address labels" is different in each case.