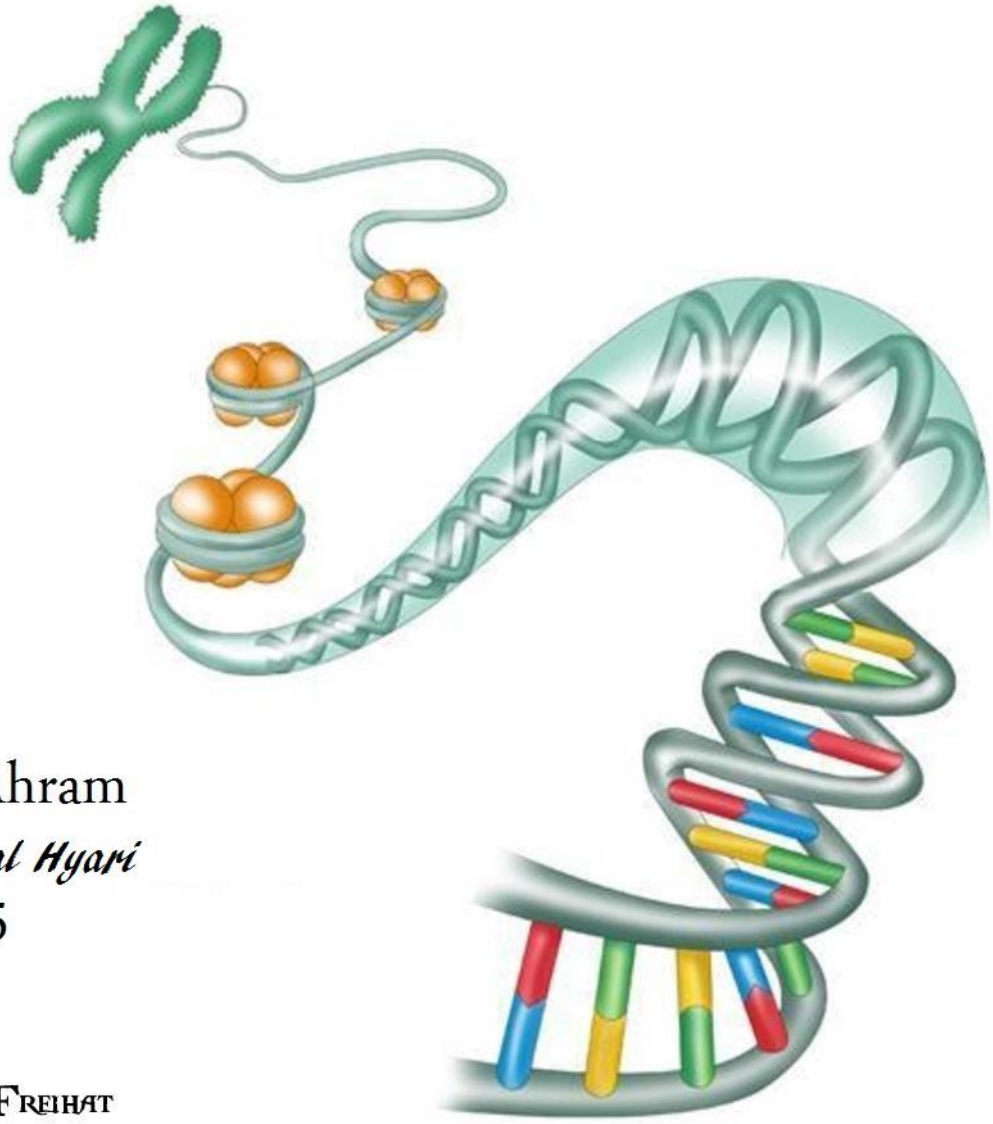
University of Jordan
Faculty of Medicine
Batch 2013-2019

Medical Committee
The University of Jordan

Genetics & Molecular Biology

○ Slides ● Sheet ○ Handout ○ other.....



Lecture # 3
Title:
Dr. Mamoun Ahram
Done By: *Ibtehal Hyari*
Date: 5/2/2015
Price:

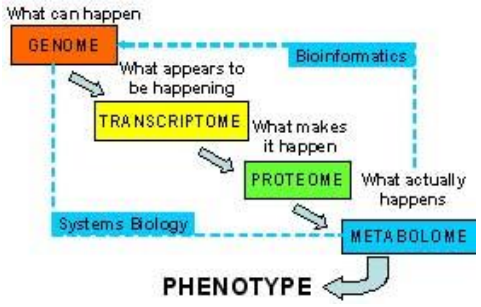
Designed by Nadeen Al-Freihat

Techniques to study molecules in cell biology

-The last thing we talked about was techniques, and we said that there are number of techniques that we can use to study molecules in cell biology and you will learn of different techniques later on in molecular biology.

Overall what we want to achieve is to understand the biology of the whole system and we can achieve this by studying the function of different cells and molecules whether they're intracellular or secreted, whether they're nuclear, mitochondrial or cytoplasmic proteins…etc

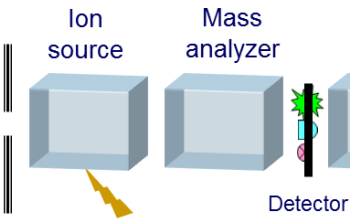
We have to understand how these proteins interact with each other as a whole, to control the whole system. As a result a number of fields have been created such as **genomics** which deals with studying hundreds to thousands of genes at the same time .We have **transcriptomics**: studying the expression of many genes at the same time .we have also **proteomics** (proteins) and we have **metabolomics** for metabolites such as glucose, pyruvate and acetyl-CoenzymeA and so on.

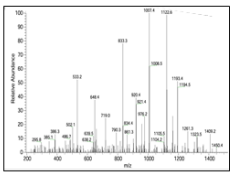


Now, one of these techniques is **two dimensional gel** and we talked about it before. Now there is another techniques, the doctor will not go into the details of this instrument he will give us the <u>basics</u> (only the purpose of the instrument and how it works) ☺.

This techniques basically is a **mass spectrometry** , what it does is that it analyzes a protein by converting it into an **ion** so first we have a protein or a peptide and this peptide once it go into the instrument it gets ionized. So let's say that we have three different proteins and they

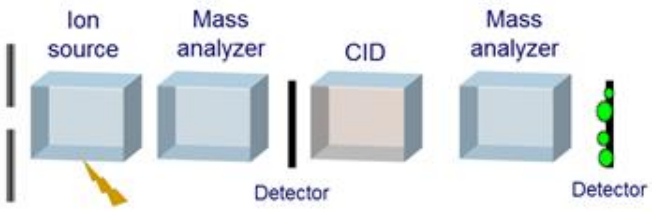
get into the system, they get ionized through *ion source* and then they travel through a *mass analyzer* into a *detector*, it hits the detector and the detector will detect that there is a peptide .The size of the peptide is inversely proportional to the speed by which it travels; so larger peptide moves slower through the mass analyzer and shorter peptide move at faster speed. Eventually we get certain peaks.

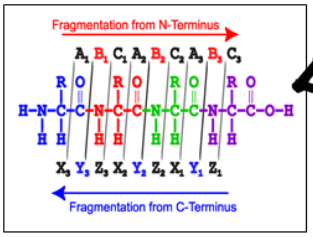


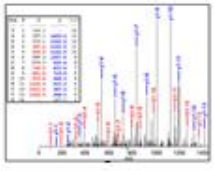


Every single protein would generate a **specific profile** (a specific pattern of peaks) and based on these peaks we can say that in our sample we have this protein or that protein and so on ..

This mass spectrometry has become more advanced to allow us to study hundreds and thousands of proteins at the same time; by integrating or attaching another mass analyzer (CID). So we select one peptide and this one peptide goes into another chamber and it gets dissociated into smaller pieces and these pieces then travel through the second mass analyzer into the detector. So you have a peptide that is cleaved→ generating another peptide →generating another profile or spectrum, with this new spectrum we get a more accurate sequencing of the peptide because now we have shorter peptides *(single amino acids)*, which makes it easier to really identify the peptide.







So what we do is peptide ionization and then we can detect different peptide speeds. We can also attach another mass analyzer, cleave the peptide into smaller peptides and that allows us to get another profile that gives us the sequence of the amino acids of the peptide itself, which give us a better or a more accurate identification of the peptide. Something else we do is that we can add a **liquid chromatography (HPLC)**; it goes into the mass analyzer 1 and mass analyzer 2 (some of them don't get into the second mass analyzer).

Eventually the point is that we can identify hundreds to thousands of proteins at the same time from one single sample and this is the basis of **proteomics** (the science of analyzing large number of proteins). What we do is that we take this profile and with certain informatics we can identify all the different proteins or peptides that exist in the sample .so the whole purpose is to identify the proteins in a sample.

** A student asked about the relation between the size and speed:

And the answer was: when the peptide gets ionized it travels .the larger the peptide the slower it travels.(it depends on both mass and charge).

Another technique that is used in proteomics is **proteins arrays**. With doctor Said Ismail you will learn about DNA microarrays which have been really useful in studying genes, in studying transcriptomics but the same concept has been applied in proteins where we can analyze large number of proteins at the same time using different types of arrays and these arrays are basically small slides and on these slides we can have different components so we can have antibodies or lysates.

We learnt from biochemistry and immunology that antibodies recognize specific proteins, we put the antibodies and then we add the sample that contain the proteins so if I have a single antibody it will

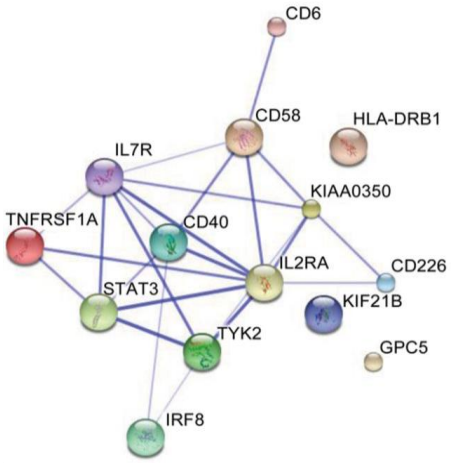
recognize just one single protein because they are specific. So with these slides or these arrays what we can do is we can analyze the presence of proteins in a sample using different methodology. We can identify for example protein–protein interaction or we can analyze interaction of protein with other molecules like carbohydrate or DNA or in different platform (in a different system) we can have an enzymatic microarray where we can analyze the enzymatic activity of different proteins on different substrates. So I can have hundreds enzymes and I want to know the substrate for these hundreds enzymes; I add the enzyme to the slide and I add the substrate. Now if the enzyme recognizes the substrate, if it catalyzes the reaction of this substrate→it will get a certain product of a certain color or florescent or a signal and I can say that on this part of the slide there is an enzyme that catalyzes a reaction with this substrate.

Let's say that on the next spot of the slide we have an enzyme but there's no signals; what does it mean? It means that the enzyme doesn't recognize the substrate; it doesn't catalyze a reaction with this substrate.

So basically the idea is that instead of taking only one single enzyme and trying to analyze it, we can do the same assay (same experiment) on large number of enzymes all at the same time. This is known as **high throughput** where we insert large amount of data and generate large amount of data and the whole purpose is to speed up discovery, instead of studying just one protein at a time we can analyze large number of proteins at the same time.

Finally we have another science known as **interactome** which is basically just like genomics or transcriptomics. Here you are studying *protein–protein interactions* and throughout the lectures we'd know that these proteins can interact with each other and this interaction:
1) specific. 2) The affinity is different of one interaction versus another.

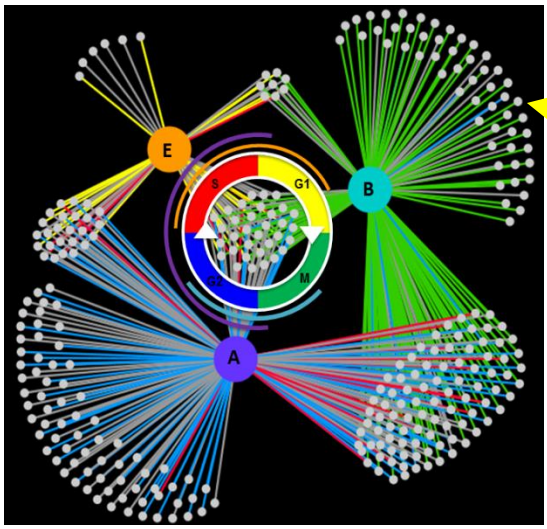
So you can basically study protein-protein interaction but this time instead of studying just protein A with protein B now you get the whole picture where you can create a **network**.



And we can say CD6 interacts with CD58 but CD58 can interact with four different proteins and so on...

In the picture below, which protein do you think if deleted would have a profound and terrible effect on the cell function?

Of course proteins A,B and E; because they can interact with large number of proteins so if we delete one of them, all the other proteins won't function.



Is it the protein the arrow is pointing at?

Of course no, because it is peripheral and hardly interacts with other proteins.

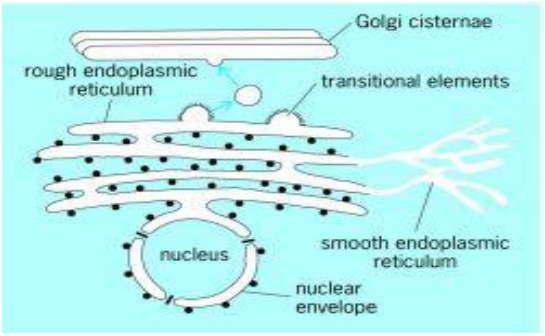
Let's remember the story of RAS, when we said that we have an inhibitor for RAS interaction with the plasma membrane and we said that when we add this inhibitor in humans it has no effect. One mechanism for the lack of effect is that we have other isoforms of the

same protein that can't be affected or there are other proteins that can replace RAS. So even though RAS is gone it doesn't matter because something else will take its place and can drive the signal transduction.

---

* **protein sorting** :

So basically the whole concept or what we want to focus on in cell biology is proteins functions. In biochemistry we learned about protein structure and the structure-function relationship. What we want to do now is talk about proteins and how they are differentially localized in different organelles; so you have ER proteins, mitochondrial proteins and nuclear proteins. We have to understand how these proteins are synthesized, how they travel (how they go into these organelles) and when they go there what's the function of these proteins. So we will start with protein sorting and will start with ER.



There is a nucleus, surrounding the nucleus we have endoplasmic reticulum that's where the protein synthesis takes place. So protein synthesis actually takes place in two places: it starts in the cytosol and some proteins stay in the cytosol or other protein synthesis starts in the cytosol but then it's completed in the ER and from the ER they go to other places. So the ER is basically a network of tubes and they form sacs "bags" and these bags are connected with the nucleus so they are part of the nucleus.

We have three types of ER:

1- Rough ER which is responsible for protein synthesis and modification and then sorting.

2- Smooth ER which is responsible for lipid metabolism.
3- Translational ER (at the end) which is responsible for moving of proteins from ER to the Golgi apparatus through vesicles.

There is something known as **the secretory pathway**, it tells us that you have a protein that is synthesized in the ER and from the ER all these proteins go to the Golgi apparatus and the Golgi does the sorting so in the ER you have protein synthesis and protein modification.

But in the Golgi you have the sorting, it's in the Golgi where proteins are moving to lysosomes, plasma membrane or go outside (released or secreted).

So this is the secretory pathway which is in brief:

ER → Golgi → lysosomes, plasma membrane or they're secreted

Proteins that are destined to other organelles like nucleus, peroxisome, mitochondria and chloroplast; they are synthesized in the **cytoplasm** and from the cytoplasm they go into the nucleus or to these organelles and some of them stay in the cytosol, so they don't enter the ER (unlike those of the secretory pathway).

* How are proteins distanced to go to the ER ?

The mRNA is processed in the nucleus and transported outside to the cytosol, it's recognized by ribosomes and these ribosomes interact with the mRNA molecule and synthesis starts in the cytosol.

As we know, the N-terminus of the protein-being-synthesized starts to come out of the ribosome first. So some proteins recognize it coming out of the ribosome during protein synthesis.

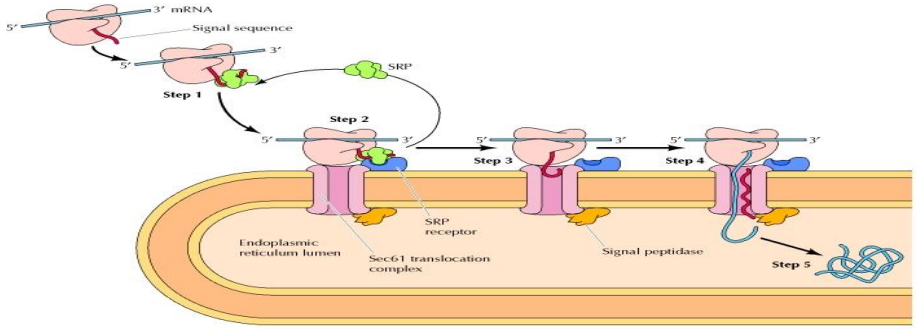
Now as it comes out, the new very few amino acids that come out of ribosome form a sequence known as **signal peptide.** This signal peptide or signal sequence is a signal that indicates that this protein should complete its synthesis inside the ER.

* So if we have a protein that is distanced to the ER and if we move the signal sequence what will happen to this protein ? If I have protein and I mutate it so that I move the signal sequence, what would happen to this protein ?

It will stay in the cytosol, it will be synthesized but it will stay in the cytosol, **it will not go to the ER.**

-What is the mechanism of translocation ? How does the protein gets into the ER?

You have synthesis then the polypeptide comes out (you have the signal sequence) and this signal sequence is recognized by a protein known as **signal recognition particle (SRP)**, it binds to the polypeptide and it takes the whole complex (polypeptide + ribosome) to the ER. Now the SRP binds to its receptor known as **SRP receptor**, and it allows for the ribosome to interact with a channel protein and this channel protein is known as **translocon** because it translocates the protein, so synthesis of the protein is completed. When this channel opens up, it allows the protein to be synthesized and synthesis is completed inside the ER as the peptide goes to the ER lumen.



* Roles of a number of proteins:

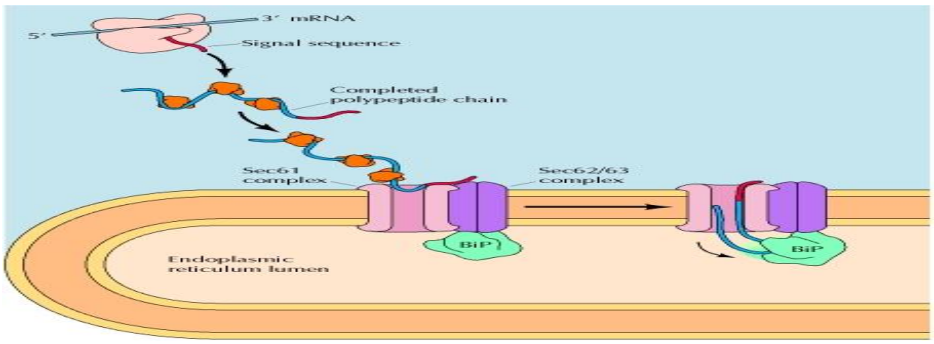
1- SRP
2- SRP receptor
3- Translocon; which is a complex of proteins
4- **Signal peptidase**; which *cleaves off the signal sequence* as the protein is translocated into the ER.

So the protein goes in <u>without the signal sequence</u>. Now in some proteins the signal sequence is known as the **pre-sequence** or the **pre-region PRE**. So some proteins are synthesized as pre-proteins which mean that these proteins contain the signal sequence.

The previous mechanism is known as **co-translational translocation,** because it's translocation of proteins to the ER as the proteins are translated.

We have another mechanism known as **posttranslational translocation.** Where the protein is totally synthesized in the cytosol, it binds to <u>chaperon proteins</u> and these chaperon proteins bind to proteins that contain a sequence that tells the protein you should go into the ER, and the protein then is translocated, it binds to the complex (translocon) and it allows the protein to be translocated into the ER.

Now we have another protein that is required in addition to translocon and this protein is known as **BiP** which is a chaperon.



-**Chaperon:** it is a protein that assess in protein folding by providing the environment that prevents the formation of improper hydrophobic interaction.

So what happens is that these chaperones help in translocation or in transporting the protein to the translocon and we have another chaperon "BiP" that facilitates the movement of the polypeptide into the ER through the translocon.

As it goes into the ER, the fate of protein is determined by its sequence, some proteins are secreted, and some go into the Golgi and then the lysosomes. These proteins are not membrane proteins , so the protein that get into the lumen of the ER then packaged in vesicles and then go to the Golgi and go to lysosomes and so on ..

But **membrane proteins** whether they are lysosomal membrane proteins, ER membrane proteins, Golgi membrane proteins, plasma membrane proteins; they must firstly be inserted into the ER membrane and from the membrane of the ER they are transported into the membranes of these other organelles. Now have to take in consideration a number of points:

1- Some proteins have a single transmembrane domain and other proteins have multiple transmembrane domains.

2- We have proteins with the C-terminus in and the N-terminus out and we have proteins with N-terminus in and C-terminus out; so we have different proteins with different orientations and different membrane spanning regions.

3- The lumen of the ER and the lumen of the Golgi are topologically equivalent to the exterior of the cell.

So when you have a protein that's inserted in the membrane of ER, which part will go into the outside of the cell?

It's the part of the protein that was exposed to the lumen of the ER, because translocation of membrane proteins from the ER to the plasma membrane needs vesicles, and during translocation the membrane protein becomes part of the vesicle, so when the vesicles fuses with the plasma membrane and opens up so the proteins inside the vesicles become out of the cell (exposed outside the cell).

Notice that the other part of the protein that became exposed to the cytosol (after translocation to the plasma membrane) was already exposed to cytosol (when the protein was still in the ER), which means that this part was originally cytoplasmic and will stay cytoplasmic, while the other part will be exposed to the outside of the cell.

* So the lumen of the ER resembles the exterior of the cell and the same thing with the Golgi membrane proteins.

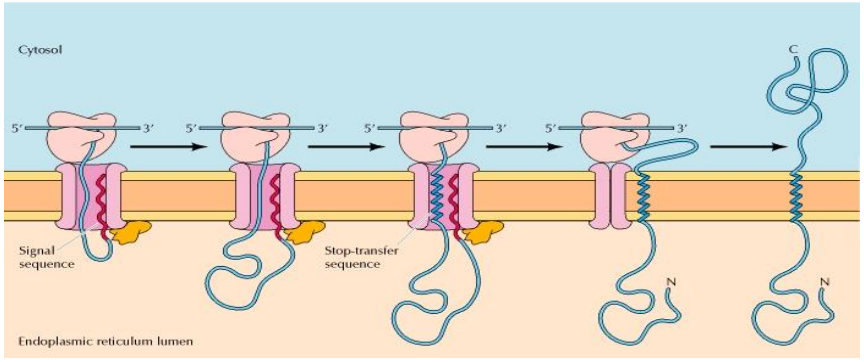
-Insertion of a membrane protein into the membrane of the ER:

* The first mechanism:

You have the signal peptide that is synthesized and the protein now is going into the ER lumen, the signal peptidase cleaves off the signal peptide.

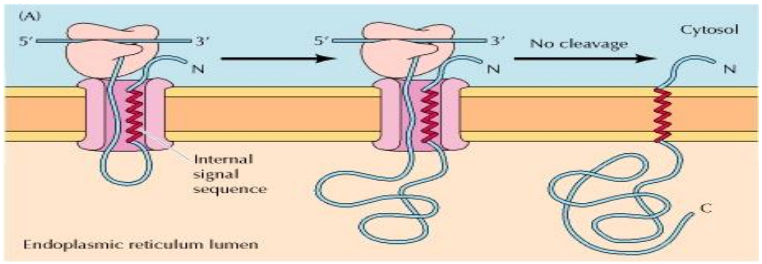
As the protein is synthesized, the channel is open and the protein goes into the ER lumen. There is a sequence known as **stop transfer sequence;** which tells the translocon to close the channel preventing the movement of the protein further into the ER lumen, so the protein continues to be synthesized outside (in the cytosol).

The N-terminus of the protein is inside, and the C –terminus is outside, and when the channel closes, protein synthesis continues. When it's completed, the protein is pushed laterally outside of the channel.

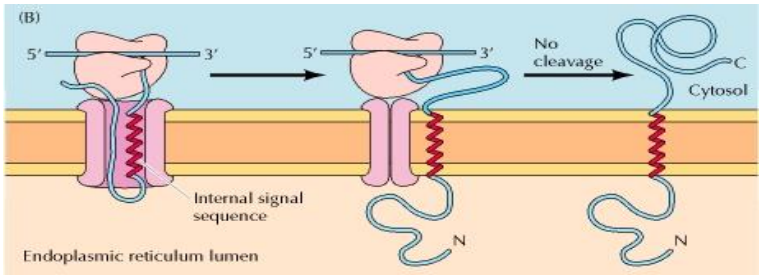


* The second mechanism:

In this case the signal sequence itself is the stop sequence, so you have synthesis of the protein with the N-terminus being outside and the C-terminus inside the cell and the protein continues to be synthesis.



Or we have the N-terminus inside and this depends on the orientation of the signal sequence, it depends on how it's located. So the N-terminus becomes inside and C-terminus outside.



So, to synthesize a single spanning protein with the C-terminus inside the ER lumen (C-terminus being eventually to the outside of the cell), the signal sequence acts as a stop transfer sequence, so it has two functions:

1- Targets the protein to the ER
2- Insert it into the membrane

Now for a protein to be oriented eventually so that the N-terminus will be to the outside there are two mechanisms:

1) When you have a stop transfer sequence that is different from signal sequence.

2) When the signal sequence acts as a stop transfer sequence.

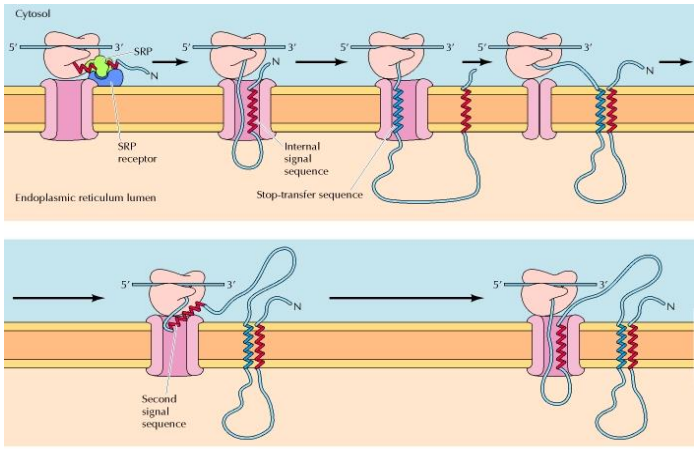
The difference in mechanism between this pathway and the other pathway is that it depends on the orientation of the signal sequence.

* To sum up:

You can have the signal sequence acting as a stop transfer sequence or not , if it doesn't act as stop sequence then the N-terminus of the protein would be to the inside of the ER and eventually will be exposed to the outside. If it acts as stop transfer sequence depending on the orientation of the signal sequence, the protein can be oriented where the N-terminus inside or the C-terminus inside.

Finally another mechanism is that we have **multiple spanning domain proteins**, we have a signal sequence and as the protein is synthesized you have a stop transfer sequence so the channel of the translocon closes. Let's say we have another sequence for the transmembrane domain, the channel opens up, and the domain gets inserted in and continues to be synthesized. Then the channel closes and so on ..

Till the protein is completely synthesized with multiple spanning domains, eventually what happen is that a protein is pushed outside of the translocon and now it's located in the ER membrane.



Good luck ☺