Introduction to protein synthesis

- **Promoter**
- **Coding region**

**(Genomic DNA strand)**

Gene transcription by RNA-polymerase

Gene translation by ribosomes

Protein folding and transport

Message RNA

Nascent protein
Protein synthesis
How it works, and how much it costs

What is the ribosome?

What is tRNA?

How are proteins synthesized?
Some protein complexes contain RNA

Two major examples of ribonucleoproteins:

1) Ribosomes (rRNA, 5S, 16S and 23S)

   Over 60% of the ribosome is constituted by ribosomal RNA, which forms extensive secondary structures with many short duplex regions. The RNA plays an active catalytic role in ribosomes.

   Relevant i.e. the 16S rRNA is important for the recognition of the Shine-Dalgarno sequence.

2) Signal recognition particle (SRP, see later..) (mRNA transport to the rough endoplasmic reticulum)
mRNA can contain secondary structures

(Repetition)

Secondary structures are folded regions which provide 3-dimensional shapes as with proteins

Secondary structures have biological functions:

Interactions with proteins:
(i.e. ribosome binding, termination of RNA polymerase, RNA-protein complexes, a number of catalytic functions)

mRNA stability and translation efficiency
Important elements of prokaryotic genes
(Repetition)

- **Promoter**
- **Coding region**
- **Transcription terminator**

- Transcribed portion (mRNA) (contains non-coding region)
- Start codon (AUG) (sometimes GUG)
- Stop-codon (UAG, UGA, or UAA)

**Shine-Dalgarno sequence:**
purine-rich sequence that promotes ribosome binding

AUG
Important elements of prokaryotic genes
(Repetition)

- Promoter
- Coding region
- Transcription terminator

Start codon (AUG)

Transcribed portion (mRNA) (contains non-coding region)

Stop-codon (UAG, UGA, or UAA)

Terminator of transcription (Palindromic sequence creating hairpin loop)

UAG
Protein synthesis
How it works, and how much it costs

What is the ribosome?

What is tRNA?

How are proteins synthesized?
Example: transfer-RNA or tRNA

5’ AGCTGCTCTGATACAGATCTGTCAGATCGATAAC GAT CGATAAC 3’

TCGACGAGACTATGTCTAGACAGTCTAGCTATTG 5’

tRNA has a typical secondary structure

3’OH

5’P

CUA

5’ AGCUGCUCUGAUACAGAUCUGUCAGAUCGAUAAC 3’

mRNA

DNA

tRNA
Example: transfer-RNA or tRNA

```
5’ AGCTGCTCTGATACAGATCTGTCTGTCAGATCGATAAC 3’
   DNA
3’
```

```
TCGACGAGACTATGTCTAGACAGTCTAGCTATTG
```

tRNA has a typical secondary structure

```
5’ tRNA
```

```
AGCUGCUCUGAUACAGAUCUCA
```

```
GAU CGAUAAC
```

```
5’ P
```

```
5’ D (Aspartate)
```

```
3’
```

```
3’ tRNA
```

```
AGCUGCUCAUAACAGAUCUCAUGUCAGAUCGUAAC
```

```
3’ RNA
```

tRNA links the codon to a specific amino acid via its anti-codon
Charging a tRNA with an amino acid

Amino acids are first activated by adenylation, which costs an unusual amount of energy.

$$\text{Amino acid} + \text{ATP} + \text{tRNA} + \text{H}_2\text{O} \rightarrow \text{Aminoacyl-tRNA} + \text{AMP} + 2\text{P}_i$$

This is the equivalent of

$$2\text{ATP} + \text{H}_2\text{O} \rightarrow \text{ADP} + \text{P}_i \text{ reactions}$$
Charging a tRNA with an amino acid

It is a two-step reaction by the enzyme Aminoacyl-tRNA synthetase

Amino acid + ATP + H₂O → Aminoacyl-AMP + 2Pᵢ
Aminoacyl-AMP + tRNA → Aminoacyl-tRNA + AMP
Charging a tRNA with an amino acid

Amino acid + ATP $\xrightarrow{\text{Aminoacyl-\text{tRNA} synthetase}}$ Aminoacyl-AMP

Aminoacyl-AMP + tRNA (Terminal adenosine) $\xrightarrow{\text{Aminoacyl-\text{tRNA} synthetase}}$ Aminoacyl-\text{tRNA}
The question of specificity

There are only two classes of aminoacyl-tRNA synthetases (transfer of amino acid to either 3’OH or 2’OH of the terminal adenosine)

There are 61 codons specifying 20 amino acids and 3 codons specifying stop codons

The aminoacyl-tRNA synthetase must transfer the activated amino acid to the correct tRNAs

It needs to recognise amino acids, ATP and tRNAs,

It needs to make sure that only the right tRNA is used for a specific activated amino acid
Aminoacyl-tRNAs are used as substrates by ribosomes

GTP + H₂O \rightarrow GDP + P_i

Charged tRNA

tRNA(G) \rightarrow tRNA(W)

discharged tRNA (to be charged)

(tRNA(G) \rightarrow tRNA(W)
Proteins are synthesized from N to C
(Repetition)

Amino-terminus

Carboxy-terminus

This reaction costs too much energy. It will not happen spontaneously.
Protein synthesis
How it works, and how much it costs

What is the ribosome?

What is tRNA?

How are proteins synthesized?
A growing polypeptide is being synthesized from mRNA. Shown is a position in which 5 amino acids have already been linked together, forming a nascent chain. (Detailed explanation in our textbook: Alberts et al., 5th ed. From RNA to Protein, pages 366-400 or alternatively Biochemistry, Stryer, 5th edition, chapter 29, pages 813-844)
Protein synthesis by ribosomes

A new “charged” tRNA(E) binds to the next codon
In a sequence of reaction steps, the peptidyl bond is formed between the new amino acid (E) and the end of the growing chain (D), leaving a discharged tRNA (for D).
GTP-driven translocation of the mRNA is mediated via Elongation factor G, which displaces the ribosome one codon further downstream in the mRNA and the discharged tRNA(for D) is released from the ribosome. The next charged tRNA(F) can now bind to the next codon
The overall reaction:

Building proteins costs energy, depending on the number of amino acids. But this is little compared to the energy needed to build the amino acids in the first place.

\[ \text{Nx (Amino acid + ATP + GTP)} \]

\[ \text{Protein + Nx (AMP + GDP + 3P}_i\text{)} \]

\( (N=\text{number of amino acids in the protein}) \)

Although ATP is the energy carrier, GTP hydrolysis releases a similar amount of energy and is often used for regulated movement/shaping of proteins.
Introduction to protein folding

- promoter
- coding region
- (Genomic DNA strand)

Gene transcription by RNA-polymerase

Gene translation by ribosomes

Protein folding and transport

Nascent protein
Proteins are polymers which fold up in an aqueous environment (Repetition)

Protein folding
Involves hiding hydrophobic residues from the aqueous phase

- hydrophilic
- hydrophobic
What happens when hydrophobic regions are exposed?
Aggregates can be formed

They may be disruptive and will permanently fail to fold
Protein folding is assisted by chaperones
Chaperone action also costs energy (nothing in life is for free)

Substrate-release by the chaperone is ATP-dependent

\[ \text{H}_2\text{O} + \text{ATP} \rightarrow \text{ADP} + \text{P}_i \]
There are many chaperones, but the most important group is the Heat shock 70 family

Heat shock 70 family (hsp70, hsc70, BiP…..)

Members of this family are found in the cytosol and in all cellular compartments in which protein folding must occur.
Introduction to chaperones

- The HSP70 family (hsp70, hsc70, BiP, etc.)

- Molecular chaperone: assists in protein folding
  - Prevents aggregation
  - Dissolves and refolds aggregates (ATP-hydrolysis)
  - Members of this family are found in the cytosol and in all cellular compartments in which protein folding must occur
ATP hydrolysis & chaperone-ligand interactions

New folding attempt
If unsuccessful: → → Ligand binding

Weak initial binding

Ligand release

Nucleotide exchange & Conformational Change

Tight binding, energy transfer, unfolding

ATP hydrolysis & Conformational Change
Again we have learned a lot from mutants

T46G : Conformational change mutant

Deficient in Ligand release (Dominant negative mutant)
ATPase mutant: strong dominant-negative

G235D: Nucleotide binding mutant

Deficient in Ligand release (Dominant negative mutant)
Introduction to protein targeting

promoter  coding region

(Genomic DNA strand)

Gene transcription by RNA-polymerase

Gene translation by ribosomes

Protein folding and transport

Nascent protein

Where can it go?
Protein synthesis starts in the cytosol
(Except for chloroplast and mitochondrial genes)

With the help of chaperones, nascent proteins will fold up properly after synthesis by the ribosome and may or may not assemble to homo- or hetero-multimers
Protein synthesis can also occur on the rough endoplasmic reticulum (ER). Nascent proteins will fold up and assemble properly and are then exported from the ER in vesicles.
The endoplasmic reticulum is the entry point into the secretory pathway

Rough ER:

Studded with ribosomes which gives a “rough” appearance under the Electron microscope (EM). Often forms sheets, but it looks like tubules when sectioned for EM.

Smooth ER:

No ribosomes. Usually tubular in shape, but it looks like round and oval “vesicles” when sectioned for EM.
Protein synthesis, cytosol or ER surface?

Two synthesis routes for nuclear encoded genes:

1) Cytosol (soluble ribosomes)

2) Endoplasmic reticulum surface (membrane bound ribosomes)

Both synthesis routes are based on cytosolic mRNA.

How does the cell decide between these routes?
How are proteins targeted to the secretory pathway?

Answer:

They are not, the mRNA and the ribosomes are targeted to the rough ER

How does this work?

1) Translation in the cytosol
2) Translation on the rough ER
How are proteins targeted to the secretory pathway?

The “signal recognition particle” (SRP) recognises “signal peptides” when they emerge as nascent chains on the ribosome surface.

SRP stops or slows down translation by the ribosome.
After signal peptide recognition and translational arrest, SRP targets the ribosome/mRNA/nascent polypeptide complex to the translocation pore on the ER.

The SRP-receptor is close to the translocation pore and binds to SRP.
SRP and the SRP-receptor are being recycled, translation resumes and the signal peptide is cleaved off.

Cleavage of the signal peptide occurs well before the protein is fully translocated and is never part of the finally folded structure (= the mature protein).
Entry into the ER lumen:
The port of entry into the secretory pathway

The ribosome seals a translocation pore on the rough ER

aqueous channel

Nascent protein

Nascent proteins will fold up and assemble
Properly, but this is not spontaneous and requires chaperones
The ER contains a special chaperone: The binding protein (BiP)

BiP belongs to the Heat shock 70 family

Members of this family are found in the cytosol and in all cellular compartments in which protein folding must occur
The ER resident chaperone BiP plays a crucial role in the protein translocation and folding. Nascent proteins will fold up and assemble properly, but this is not spontaneous and requires chaperones. BiP masks exposed hydrophobic regions of the nascent chain, and when not occupied by a ribosome, BiP seals the translocation pore, resulting in a smaller pore diameter when not in use. The ribosome seals a translocation pore on the rough ER through the aqueous channel.
The ER resident chaperone BiP plays a crucial role in the protein translocation and folding. BiP continues to assist even after translocation until the protein is correctly folded and has no exposed hydrophobic domains left on its surface.
The ER resident chaperone BiP plays a crucial role in the protein translocation and folding.

When the protein is correctly folded, it can be exported from the ER to reach the Golgi apparatus.
Some proteins are multimeric. In those cases BiP associates with monomers.

In some cases, correct folding involves assembly. Monomers will then be bound to BiP because they have exposed hydrophobic regions, whereas correctly assembled multimers have those regions masked.
Example of multimeric proteins that have to assemble in the ER

Antibodies contain light chains and heavy chains which are synthesized independently in the ER lumen.

Light chain monomer

Heavy chain monomer

Assembled IgG tetramer

Disulfide bonds

Assembly is essential to permit secretion of antibodies
Antibody coding regions contain N-terminal signal peptides

“Nascent” polypeptides

After translocation, the signal peptides are cleaved

“Mature” poly-peptides

Assembly to tetramers
When less is more
How the ER chaperone BiP was discovered

Two independent lines of research led to the discovery of the ER chaperone BiP. One came from the immunology field and was based on the special properties of a cell line which produced only the heavy chain of an antibody. In this cell line, the ER filled up with heavy chains that could not assemble to fully functional tetramers due to the lack of the light chain. The cell line was characterised in detail and a new abundant protein was found to be induced and bound to these heavy chains (“Binding protein, BiP”). The other line of research originated from the physiology field and centered around the consequences of glucose starvation. In cells starved for glucose a similar protein was induced and called glucose-regulated protein 78 (“GRP78”). It took a while before it was realised that both proteins were actually the same. In the heavy chain producers, the chaperone BiP was needed to cope with all the heavy chains without a binding partner. In glucose starvation, glycoproteins in the ER were synthesized without the glycans (sugars, see further in the course) and some of these could not fold properly and constantly sequestered new chaperones, hence the induction of this chaperone (GRP78). Nowadays, the term BiP is used, simply because it is easier to say the word. The example illustrates that mutants or physiologically unfavourable conditions can lead to the induction of genes and result in the discovery of gene products that would remain unnoticed under normal conditions.
Some antibodies are even more complex

**IgA tetramer-dimer**

**IgM tetramer-pentamer**

Additional protein to join the tetramers