Steps at which eukaryotic gene expression can be controlled
Protein Variability and Protein Activity Control

Aminoacid sequence
Three-dimensional shape (conformation)
Function
Protein processing
Degradation
Four types of post-translational processing events

- Folding
  - Section 11.3.1
- Proteolytic cleavage
  - Section 11.3.2
- Chemical modification
  - Section 11.3.3
- Intein splicing
  - Section 11.3.4

Polypeptide

Position of the removed intein

New chemical groups
Protein Folding

The aminoacid sequence contains all the information needed to fold the polypeptide into its correct tertiary structure
The cellular mechanisms that monitor protein quality after protein synthesis
The co-translational folding of a protein

The cellular chaperone machinery is specifically recruited to bind to ribosomes and protects nascent chains and folding intermediates from nonproductive interactions.
Two main ATP-dependent classes of chaperons, Hsp70 chaperon and cylindrical chaperonin complexes, mediate protein folding.

Hsp70 chaperones bind to hydrophobic regions in unfolded polypeptides, including those that are still being translated, and hold protein in an open conformation until it is ready to be folded.

Structure of GroEL/GroES chaperonin (Hsp60)
Protein processing

Proteolytic Cleavage
Chemical Modification
(Intein splicing)
Protein processing by proteolytic cleavage

![Diagram showing the process of protein processing by proteolytic cleavage. The polypeptide is cleaved into an active protein and three active proteins, with a discarded end-segment.]
Ex. proteolytic cleavage: melitin and insulin

(A) Promelittin → Melitin

Extracellular protease

(B) Signal peptide - hydrophobic aa sequence that attaches preproinsulin to the membrane, before exporting the protein through the membrane to the extracellular environment

Proremittin

Preproinsulin

Proinsulin

Insulin

(A)

APEPEPEPEPEAEADPEAGIGAVLKVTGTPALISWIKRKRQQG

22 aa

(B)

N

24 aa

Set sites

A chain

21 aa

Cut

C chain

B chain

30 aa

Cut

Removal of the signal peptide

Disulfide bond formation

Cut

Removal of the B chain

Genomes 11.28
Converting preproinsulin to insulin

1. Translation and translocation

2. Folding, oxidation, and signal peptide cleavage

3. ER export, Golgi transport, vesicle packaging

4. Protease cleavage liberates C-peptide

5. Carboxypeptidase E produces mature insulin

human insulin (51 peptide)

CHAIN A

Gln-Cys-Cys-Thr-Ser-Ile-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cys-Asn
His-Leu-Cys-Gly-Ser-His-Leu-Val-Glu-Ala-Leu-Tyr-Leu-Val-Cys
Gln
Asn
Val
Glu
CHAIN B

Thr-Lys-Pro-Thr-Tyr-Phe-Phe-Gly

(Snustad et al. 1997)
Ex. proteolytic cleavage: the pro-opiomelanocortin polyprotein

POMC is glycosylated and then cleaved to give a number of neurohormones that can be cleaved enzymatically into the following peptides (neurohormones):

- γ2MSH
- ACTH
- β-LPH
- γ1MSH
- αMSH
- CLIP
- γ-LPH
- β-ENDO
- βMSH ME

In the intermediate lobe of pituitary gland:
- melanin production
- neurons in the arcuate nucleus of the hypothalamus - appetite

In anterior pituitary gland:

Each of these peptides is packaged in large dense-core vesicles that are released from the cells by exocytosis in response to appropriate stimulation.
Chemical Modification
Post-translational chemical modification of calf histone H3

Lysine acetylation and methylation of histone H3
<table>
<thead>
<tr>
<th>Modification</th>
<th>Amino acids that are modified</th>
<th>Examples of proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition of small chemical groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetylation (CH₃CooH)</td>
<td>Lysine</td>
<td>Histones</td>
</tr>
<tr>
<td>Methylation (CH₃)</td>
<td>Lysine</td>
<td>Histones</td>
</tr>
<tr>
<td>Phosphorylation (P)</td>
<td>Serine, threonine, tyrosine</td>
<td>Some proteins involved in signal transduction</td>
</tr>
<tr>
<td>Hydroxylation (OH)</td>
<td>Proline, lysine</td>
<td>Collagen</td>
</tr>
<tr>
<td>N-formylation (COH)</td>
<td>N-terminal glycine</td>
<td>Melittin</td>
</tr>
<tr>
<td>Addition of sugar side chains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O-linked glycosylation (hydroxyl groups)</td>
<td>Serine, threonine</td>
<td>Many membrane proteins and secreted proteins</td>
</tr>
<tr>
<td>N-linked glycosylation (amino group)</td>
<td>Asparagine</td>
<td>Many membrane proteins and secreted proteins</td>
</tr>
<tr>
<td>Addition of lipid side chains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acylation</td>
<td>Serine, threonine, cysteine</td>
<td>Many membrane proteins</td>
</tr>
<tr>
<td>N-myristoylation (miristic acid)</td>
<td>N-terminal glycine</td>
<td>Some protein kinases involved in signal transduction</td>
</tr>
<tr>
<td>Addition of biotin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biotinylation</td>
<td>Lysine</td>
<td>Various carboxylase enzymes</td>
</tr>
</tbody>
</table>

150 different aa already described
Some ways in which the activity of gene regulatory proteins is regulated in eucaryotic cells

Each of these mechanisms is typically controlled by extracellular signals which are communicated across the plasma membrane to the gene regulatory proteins in the cell - SIGNAL TRANSDUCTION
Protein Degradation
Protein degradation... when?
**Relation between N-terminal amino acid and half-life of E. coli β-galactosidase proteins with modified N-terminal amino acids**

<table>
<thead>
<tr>
<th>N-terminal Amino Acid</th>
<th>Half-life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met, Ser, Ala, Thr, Val, Gly</td>
<td>more than 20 h</td>
</tr>
<tr>
<td>Ile, Glu</td>
<td>30 min</td>
</tr>
<tr>
<td>Tyr, Gln</td>
<td>10 min</td>
</tr>
<tr>
<td>Pro</td>
<td>7 min</td>
</tr>
<tr>
<td>Phe, Leu, Asp, Lys</td>
<td>3 min</td>
</tr>
<tr>
<td>Arg</td>
<td>2 min</td>
</tr>
</tbody>
</table>

**PEST sequence**

<table>
<thead>
<tr>
<th>PEST sequence</th>
<th>Half-life less then 2 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Gln Ser Thr</td>
<td></td>
</tr>
</tbody>
</table>
Acerca da degradação de proteínas

• Várias vias de degradação
  - **Lisossomos**: contêm uma série de hidrolases e enzimas proteolíticas, degradando essencialmente proteínas transmembranares e do lúmen dos organitos
  
  - **Proteossoma**: complexo multiproteico que degrada proteínas ubiquitinadas, localizadas sobretudo no núcleo e no citosol. Ex. Factores de transcrição, proteínas da regulação do ciclo celular como as cinases e fosfatases etc.
  
  - ...

• Processos altamente selectivos e rápidos

• Eucariotas- descrito o Proteossoma
Ubiquitin and the marking of protein with multiubiquitin chains

Ubiquitin-mediated proteolytic pathway

(7-8 residues)