Molecular Biology of the cell

Protein Translation and PTMs

Lu Linrong (鲁林荣) PhD

Laboratory of Immune Regulation
Institute of Immunology
Zhejiang University, School of Medicine
Medical Research Building B815-819
Email: Lu.Linrong@gmail.com
Website: http://mypage.zju.edu.cn/llr
Why study proteins?

- The Final part of the central dogma
- Proteins are coded by genes
- They play crucial functional roles in almost every biological process (structural proteins, Enzymes, signaling)
The life cycle of a protein

• Where does a protein come from?
• How is a protein processed, modified, translocated to the proper place and degraded?
• How to describe the structure of a protein? What are the functions?
• How to study the protein(s)?

• Protein synthesis (Translation) 蛋白质翻译
• Protein maturation (folding, modification) and degradation 蛋白质成熟,降解
• Structure and function of protein 蛋白质的结构与功能
• Methods: protein-protein interaction et al 蛋白-蛋白相互作用
I. Protein Synthesis

Reference and further readings:


Chapter 8: Protein Synthesis, *Gene IX*, Lewin Benjamin

Internet Tools: Google, Wiki ......
Messenger RNA

Transcribed from DNA and carries the genetic information out of the nucleus into cytoplasm for protein synthesis.
Translation is a big issue for a living cell

In rapid growing bacterial cells, protein synthesis consumes
80% of the cell’s energy
50% of the cell’s dry weight

RNAs as both template and machinery:

1. mRNAs (~5% of total cellular RNA)
2. tRNAs (~15%)
3. aminoacyl-tRNA synthetases (氨酰tRNA合成酶)
4. ribosomes (~100 proteins and 3-4 rRNAs-~80%)
Genetic Code

The **genetic code** is the set of rules by which information encoded in genetic material (DNA or mRNA sequence) is translated into proteins (amino acid sequences) by living cells.

The code defines how sequences of three nucleotides, called **codons**, specify which amino acid will be added next during protein synthesis.

George Gamow (Russian-born **theoretical physicist**) postulated that a three-letter code must be employed to encode the 20 standard amino acids used by living cells to encode proteins. With four different nucleotides, a code of 2 nucleotides could only code for a maximum of $4^2$ or 16 amino acids. A code of 3 nucleotides could code for a maximum of $4^3$ or 64 amino acids.
The Crick, Brenner, Barnett, Watts-Tobin experiment of 1961 demonstrated that three bases of DNA code for one amino acid in the genetic code.

In the experiment, proflavin-induced mutations of the T4 bacteriophage gene, rIIB, were isolated. Proflavin causes mutations by inserting itself between DNA bases, typically resulting in insertion or deletion of a single base pair.

The mutants produced by Crick and Brenner could not produce functional rIIB protein because the insertion or deletion of a single nucleotide caused a frameshift mutation. Mutants with two or four nucleotides inserted or deleted were also nonfunctional. However, the mutant strains could be made functional by using proflavin to insert or delete a total of three nucleotides.

This proved that the genetic code uses a codon of three DNA bases that corresponds to an amino acid.
Genetic Code

The first elucidation of a codon was done by Marshall Nirenberg(sit) and Heinrich J. Matthaei (stand) in 1961 at the National Institutes of Health. They used a cell-free system to translate a poly-uracil RNA sequence (i.e., UUUUU...) and discovered that the polypeptide that they had synthesized consisted of only the amino acid phenylalanine.

Similar approaches were taken to decode other codes.
## Genetic Code

### TABLE 30.1 The Genetic Code

<table>
<thead>
<tr>
<th>First Position (5'-end)</th>
<th>Second Position</th>
<th>Third Position (3'-end)</th>
<th>Third-Base Degeneracy Is Color-Coded</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>UUU Phe</td>
<td>UCU Ser</td>
<td>UAU Tyr</td>
<td>UGU Cys</td>
</tr>
<tr>
<td>UCU Phe</td>
<td>UCC Ser</td>
<td>UAC Tyr</td>
<td>UGC Cys</td>
</tr>
<tr>
<td>UUA Leu</td>
<td>UCA Ser</td>
<td>UAA Stop</td>
<td>UGA Stop</td>
</tr>
<tr>
<td>UUG Leu</td>
<td>UCG Ser</td>
<td>UAG Stop</td>
<td>UGG Trp</td>
</tr>
<tr>
<td>CUU Leu</td>
<td>CGU Pro</td>
<td>CAU His</td>
<td>GGU Arg</td>
</tr>
<tr>
<td>CUC Leu</td>
<td>CCC Pro</td>
<td>CAC His</td>
<td>GCC Arg</td>
</tr>
<tr>
<td>CUA Leu</td>
<td>CCA Pro</td>
<td>CAA Gln</td>
<td>GCA Arg</td>
</tr>
<tr>
<td>CUG Leu</td>
<td>CCG Pro</td>
<td>CAG Gln</td>
<td>GCG Arg</td>
</tr>
<tr>
<td>AUU Ile</td>
<td>ACU Thr</td>
<td>AUA Asn</td>
<td>AGU Ser</td>
</tr>
<tr>
<td>AUC Ile</td>
<td>ACC Thr</td>
<td>AAC Asn</td>
<td>AGC Ser</td>
</tr>
<tr>
<td>AUA Ile</td>
<td>ACA Thr</td>
<td>AAA Lys</td>
<td>AGA Arg</td>
</tr>
<tr>
<td>AUG Met*</td>
<td>ACG Thr</td>
<td>AAG Lys</td>
<td>AGG Arg</td>
</tr>
<tr>
<td>GUU Val</td>
<td>GCU Ala</td>
<td>GAU Asp</td>
<td>GGU Gly</td>
</tr>
<tr>
<td>GUC Val</td>
<td>GCC Ala</td>
<td>GAC Asp</td>
<td>GGC Gly</td>
</tr>
<tr>
<td>GUA Val</td>
<td>GCA Ala</td>
<td>GAA Glu</td>
<td>GGA Gly</td>
</tr>
<tr>
<td>GUG Val</td>
<td>GCG Ala</td>
<td>GAG Glu</td>
<td>GGG Gly</td>
</tr>
</tbody>
</table>

**Third-Base Relationship**
- U, C, A, G
- U, A
- C, A
- A, G

**Third Bases with Same Meaning**
- U, C, A
- U, G
- A, C

**Number of Codons**
- 32 (8 families)
- 12 (6 pairs)
- 14 (7 pairs)
- 3 (AUX = Ile)
- 2 (AUG = Met)
- 1 (UGA = Stop)

* *AUG signals translation initiation as well as coding for Met residues.*

- **Start codon:** AUG-methionine
- **Stop codon:** UGA, UAA, UAG
- **Degeneracy:** 61 triplets code for 20 amino acids
- **Universal**
Protein Synthesize Machinery

- mRNA: the template
- tRNA: the amino acid carrier
- Ribosome (rRNA & ribosomal proteins): the translation apparatus with catalyzing activity
tRNA match AAs to codons in mRNA

- tRNA contains 60-95 nt, commonly around 80.
- Deliver amino acids to the translational complex.
- Binds to mRNA through 3-base anticodon complementary to codon in mRNA.
- In bacteria, there are 30-40 tRNAs with different anticodons. In animal and plant cells, about 50 different tRNAs are found. (different Coden preference in bacteria and human )
- "Wobble" during reading of the mRNA allows some tRNAs to read multiple codons that differ only in the 3rd base.
tRNA secondary structure (cloverleaf)

- tRNA secondary structure consists of **stem & loop domains**.
- Double helical stem domains arise from base pairing between complementary stretches of bases within the same strand.
- Loop domains occur where lack of complementarity or the presence of **modified bases** prevents base pairing.
tRNA tertiary structure (L-shaped)

- RNA tertiary structure depends on interactions of bases at distant sites.
- These interactions generally involve non-standard base pairing and/or interactions involving three or more bases.
- Unpaired adenosines predominate in participating in non-standard interactions that stabilize tertiary RNA structures.
Synthesis of Aminoacyl-tRNAs (adding AAs to empty tRNAs)

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

(2) Aminoacyl-AMP + tRNA $\xrightarrow{\text{Aminoacyl-tRNA synthetase}}$ Aminoacyl-tRNA

Figure 17.2
Two step decoding (accuracy)

Aminoacyl-tRNA synthetase ensure the right aa to tRNA

coden-anti-coden pair
Some Codons Are Used More Than Others

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>E. coli Gene Frequency/1000</th>
<th>Human Gene Frequency/1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>3.2</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>CUC</td>
<td>9.9</td>
<td>20.1</td>
</tr>
<tr>
<td></td>
<td>CUG</td>
<td>54.6</td>
<td>42.1</td>
</tr>
<tr>
<td></td>
<td>CUU</td>
<td>10.2</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>UUA</td>
<td>10.9</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>UUG</td>
<td>11.5</td>
<td>11.1</td>
</tr>
<tr>
<td>Pro</td>
<td>CCA</td>
<td>8.2</td>
<td>15.4</td>
</tr>
<tr>
<td></td>
<td>CCC</td>
<td>4.3</td>
<td>20.6</td>
</tr>
<tr>
<td></td>
<td>CCG</td>
<td>23.8</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td>CCU</td>
<td>6.6</td>
<td>16.1</td>
</tr>
<tr>
<td>Ala</td>
<td>GCA</td>
<td>15.6</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td>GCC</td>
<td>34.4</td>
<td>29.7</td>
</tr>
<tr>
<td></td>
<td>GCG</td>
<td>32.9</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>GCU</td>
<td>13.4</td>
<td>18.9</td>
</tr>
<tr>
<td>Lys</td>
<td>AAA</td>
<td>36.5</td>
<td>21.9</td>
</tr>
<tr>
<td></td>
<td>AAG</td>
<td>12.0</td>
<td>35.2</td>
</tr>
<tr>
<td>Glu</td>
<td>GAA</td>
<td>13.5</td>
<td>26.1</td>
</tr>
<tr>
<td></td>
<td>GAG</td>
<td>19.2</td>
<td>41.6</td>
</tr>
</tbody>
</table>

专利号 94112091.0
名称 合成的人γ-干扰素基因(cDNA)在大肠杆菌中的高效表达
公开（公告）号 1109506
公开（公告）日 1995.10.04
优先权
地址 200031上海市岳阳路320号
国际申请
专利代理机构 中国科学院上海专利事务所
代理人 衷诚宣
专利摘要
γ型干扰素（IFN-γ）具有免疫调节、抗病毒和抗肿瘤等多种生物功能，是世界上竞相开发的基因工程产品。本发明是将设计的化学合成γ型干扰素基因插入含有λ噬菌体P
专利主权项
一种合成的γ-型干扰素（IFN-γ）基因（cDNA）在大肠杆菌中的高效表达，其特征是：（a）使用人工设计，化学合成的基因，基因结构式如下： ATG CAG GAC CGG TAC GTT AAA GAA GCT GAA AAC CTG AAG AAA IAC TTC AAT GAC TTT GAG Gly Ala Gnu Asn Leu Lys Tyr Phe Asp Ala Gly His TCT GAC GGT TGT GTT GAC GAA ATT GTT ACC CTG TAC TTT ACT ACG GTT AAT TG TTT AAA CTG TCT CCG GAG GAA ACT ATC AAG GAA GAC ATG AAC GTC TTT Gys Asp Ser Ile Gln Lys Ser Val Glu Thr Ile Lys Glu Asp Met Asn Val AAA TTT TAC AAT ACG AAG AAA CGT GAC GAC TTC AAG CTG ACT AAT TCT CTT Lys Phe phe Asn Ser Asn Lys Lys Arg Asp Phe Glu Leu Thr Asn Tyr Ser GTT ACC GAT CTG AAC GTT CAA CGT AAA GCT ATC CAC GAG CTC ATC CAG GTT ATG GCT GAA Val Thr Asp Leu Asn Val Gln Arg Lys Ala Ile His Glu Leu Ile Gln Val Met Ala Glu CTG TCT GCC GCG GCT AAA ACT GGT AAA CGT AAT CTC CAT GCG TTC CGC GGT CTG Leu Ser Pro Ala Ala Lys Thr GLY Lys Arg Ser Gln Met Leu Phe Arg Gly Arg GTG GCT TCT CAG TAA TAG Arg Ala Ser Gln Term Term IFN-γ由143个氨基酸组成，合成的该基因有429bp，加上起始密码子ATG，两个终止密码子TAA、TAG
使合成的IFN-γ基因能在大肠杆菌中高效表达，产生大量IFN-γ，而不改变IFN-γ的氨基酸顺序，（b）将合成的IFN-γ基因克隆人带有PL启动子、Cts857温控阻遏蛋白基因、核糖体5SrRNA基因的终止信号t1t2和MCS区的表达载体PLY4后，构建成功IFN-γ的表达质粒pLY4-γ，转化大肠杆菌后获得高效表达，IFN-γ占菌体可溶性总蛋白的60～80%；本发明工程菌为大肠杆菌Escherichia coli SIB－203－2 pLY4－γ，保藏编号CCTCC NO（中国典型培养物保藏中心，中国．武汉．珞珈山．武汉大学校同内）
The Components of Ribosome

- **rRNA**
  - Prokaryotic: 23S (2906 bases) + 5S (120 bases)
  - Eukaryotic (mammalian): 28S (4800 bases) + 5.8S (160 bases) + 5S (120 bases)

- **Proteins**
  - Prokaryotic: L1, L2, L3 (Total: 31)
  - Eukaryotic (mammalian): S1, S2, S3 (Total: 33)

- **Subunits**
  - Prokaryotic: 5S + 23S + 50S
  - Eukaryotic (mammalian): 16S + 28S + 60S

- **Assembled ribosomes**
  - Prokaryotic: 70S
  - Eukaryotic (mammalian): 80S
Structure of Ribosome

http://www.molgen.mpg.de/~ag_ribo/ag_franceschi/franceschi-projects-50S-2.html

Ribosome from *D. radiodurans* (耐辐射球菌)
Ribosome Has Several Active Centers

**A site** (acceptor site): exposes the codon representing the next amino acid due to be added to the chain.

**P site** (donor site): nascent polypeptide chain lies. The codon representing the most recent added amino acid.

**E site**: a temporary tRNA-binding site.

**S1 site**: has a strong affinity for single-stranded nucleic acid, and is responsible for the initial binding of 30S subunits to mRNA and initiator tRNA.
Protein Synthesis Occurs by Initiation, Elongation, and Termination

- Protein synthesis falls into the three stages: Initiation, Elongation, Termination.

- Different sets of accessory factors assist the ribosome at each stage.

- Energy is provided at various stages by the hydrolysis of guanine triphosphate (GTP).
Initiation of Protein Synthesis

- Initiation Involves Base Pairing Between mRNA RBS (ribosome binding site) and rRNA
- Initiation Needs Accessory Factors (Initiation Factors, IFs)
- A Special Initiator tRNA (fMet-tRNAf) Starts the Polypeptide Chain
- Small Subunits Scan for Initiation Sites on Eukaryotic mRNA
- A rate-limiting step
Initiation of Protein Synthesis

Critical event:
Begin protein synthesis at the start codon, thereby setting the stage for the correct in-frame translation of the entire mRNA.

Main mechanisms:
Base pairing between mRNA and rRNA
Base pairing between mRNA and tRNA
fMet-tRNA_{i}^{Met} can only bind at the P site to begin synthesis

Participants:
- IFs
- mRNA
- fMet-tRNA_{i}^{Met}
- small subunit
- large subunit
Base Pairing Between mRNA and rRNA

- An initiation site on bacterial mRNA consists of the AUG initiation codon preceded with a gap of -10 bases by the Shine-Dalgarno polypurine hexamer (SD sequence).
- The rRNA of the 30S bacterial ribosomal subunit has a complementary sequence that base pairs with the SD sequence during initiation.
Various Shine–Dalgarno sequences recognized by *E. coli* ribosomes. These sequences lie about ten nucleotides upstream from their respective AUG initiation codon and are complementary to the UCCU core sequence element of *E. coli* 16S rRNA. G:U as well as canonical G:C and A:U base pairs are involved here.
Small Subunits Scan for Initiation Sites on Eukaryotic mRNA

- Eukaryotic 40S ribosomal subunits bind to the 5' end of mRNA and scan the mRNA until they reach an initiation site.
- The eukaryotic initiation site consists of a ten nucleotide sequence that includes an AUG codon.
- 60S ribosomal subunits join the complex at the initiation site.
**Cap vs IRES**

- **Cozak sequence**: NNN(A/G)NNAUGG
  
  ACCAUUGG
  
  GCCAUUGG

- **IRES (Internal ribosome entry site)**: Independent of 5' cap; co-expression of two proteins

  *Cap-dependent translation initiation in eukaryotes*

  ![Cap-dependent translation initiation](image)

  *Cap-independent translation initiation*

  ![Cap-independent translation initiation](image)

  *IRES (internal ribosome entry site)*

  ![IRES](image)
Functions of Initiation Factors in Bacteria

Initiation factors: IF-1, IF-2, and IF-3.

IF-3 is needed for 30S subunits to bind specifically to initiation sites in mRNA.

IF-2 binds a special initiator tRNA and controls its entry into the ribosome. has a ribosome-dependent GTPase activity

IF-1 binds to 30S subunit in the A site as a part of the complete initiation complex. Associates Prevents an aminoacyl-tRNA from entering. Modulates IF2 binding to the ribosome.
A Special Initiator tRNA Starts the Polypeptide Chain

The initiator N-formyl-methionyl-tRNA (fMet-tRNAf) is generated by formylation (甲酰化) of methionyl-tRNA, using formyl-tetrahydrofolate (四氢叶酸脂) as cofactor.

fMet-tRNAf has unique features that distinguish it as the initiator tRNA.
Protein Synthesis Elongation Essentials:

1) mRNA: 70S ribosome: peptidyl-tRNA complex

2) Aminoacyl-tRNAs

3) Elongation factors

3) GTP

Ribosome has three tRNA-binding sites. An aminoacyl-tRNA enters the A site. Peptidyl-tRNA is bound in the P site. Deacylated tRNA exits via the E site.

An amino acid is added to the polypeptide chain by transferring the polypeptide from peptidyl-tRNA in the P site to aminoacyl-tRNA in the A site.
The Process of Elongation

Elongation Steps:

1. Binding of incoming aminoacyl-tRNA at A site.

2. Peptide bond formation: transfer of the peptidyl chain from the tRNA bearing it to the -NH2 group of the new AA.

3. Translocation of the one-residue-longer peptidyl-tRNA to the P site
Elongation Factor Tu Loads Aminoacyl-tRNA into the A Site

- **EF-Tu** is a monomeric G protein whose active form (bound to GTP) binds aminoacyl-tRNA.

- The **EF-Tu-GTP-aminoacyl-tRNA complex** binds to the ribosome A site.
The Polypeptide Chain Is Transferred to Aminoacyl-tRNA

- The 50S subunit (mainly 23S rRNA) has peptidyl transferase activity.
- The nascent polypeptide chain is transferred from peptidyl-tRNA in the P site to aminoacyl-tRNA in the A site.
- Peptide bond synthesis generates deacylated tRNA in the P site and peptidyl-tRNA in the A site.
Translocation requires EF-G

- EF-G is a G protein which resembles the aminoacyl-tRNA-EF-Tu-GTP complex.

- Binding of EF-Tu and EF-G to the ribosome is mutually exclusive.

- EF-G binds to the ribosome to sponsor translocation.

- The hydrolysis of GTP is needed to release EF-G.
Polyribosomes

Active protein-synthesizing units consist of an mRNA with several ribosomes attached to it. Such structures are polyribosomes or polysomes.

Termination

- The codons UAA, UAG, and UGA terminate protein synthesis.
- Termination codons are recognized by protein release factors (RF).
- Release factors resemble aminoacyl-tRNA-EF-Tu and EF-G.
- Release factors respond to specific termination codons and hydrolyze the polypeptide tRNA linkage.
The Process of Termination

- The RF (release factor) terminates protein synthesis by releasing the protein chain.
- The RRF (ribosome recycling factor) releases the last tRNA.
- EF-G releases RRF, causing the ribosome to dissociate.

Movie Here!
Antibiotics

<table>
<thead>
<tr>
<th>INHIBITOR</th>
<th>SPECIFIC EFFECT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acting only on bacteria</strong></td>
<td></td>
</tr>
<tr>
<td>Tetracycline</td>
<td>blocks binding of aminoacyl-tRNA to A-site of ribosome</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>prevents the transition from translation initiation to chain elongation and also causes miscoding</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>blocks the peptidyl transferase reaction on ribosomes (step 2 in Figure 6–66)</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>binds in the exit channel of the ribosome and thereby inhibits elongation of the peptide chain</td>
</tr>
<tr>
<td><strong>Acting on bacteria and eucaryotes</strong></td>
<td></td>
</tr>
<tr>
<td>Puromycin</td>
<td>causes the premature release of nascent polypeptide chains by its addition to the growing chain end</td>
</tr>
<tr>
<td><strong>Acting on eucaryotes but not bacteria</strong></td>
<td></td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>blocks the translocation reaction on ribosomes (step 3 in Figure 6–66)</td>
</tr>
<tr>
<td>Anisomycin</td>
<td>blocks the peptidyl transferase reaction on ribosomes (step 2 in Figure 6–66)</td>
</tr>
</tbody>
</table>

![Diagram of ribosomal subunits with inhibitors](Image)
Eukaryotic Translation Initiation

**TABLE 30.9  Properties of Eukaryotic Translation Initiation Factors**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Subunit</th>
<th>Size (kD)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>eIF1</td>
<td></td>
<td>15</td>
<td>Enhances initiation complex formation</td>
</tr>
<tr>
<td>eIF1A</td>
<td></td>
<td>17</td>
<td>Stabilizes Met-tRNA(_i) binding to 40S ribosomes</td>
</tr>
<tr>
<td>eIF2</td>
<td></td>
<td>125</td>
<td>GTP-dependent Met-tRNA(_i) binding to 40S ribosomes</td>
</tr>
<tr>
<td>α</td>
<td></td>
<td>36</td>
<td>Regulated by phosphorylation</td>
</tr>
<tr>
<td>β</td>
<td></td>
<td>50</td>
<td>Binds Met-tRNA(_i)</td>
</tr>
<tr>
<td>γ</td>
<td></td>
<td>55</td>
<td>Binds GTP, Met-tRNA(_i)</td>
</tr>
<tr>
<td>eIF2B</td>
<td></td>
<td>270</td>
<td>Promotes guanine nucleotide exchange on eIF2</td>
</tr>
<tr>
<td>eIF2C</td>
<td></td>
<td>94</td>
<td>Stabilizes ternary complex in presence of RNA</td>
</tr>
<tr>
<td>eIF3</td>
<td></td>
<td>800</td>
<td>Promotes Met-tRNA(_i) and mRNA binding</td>
</tr>
<tr>
<td>eIF4F</td>
<td></td>
<td>243</td>
<td>Binds to mRNA caps and poly(A) tails; consists of eIF4A, eIF4E, and eIF4G; RNA helicase activity unwinds mRNA 2' structure</td>
</tr>
<tr>
<td>eIF4A</td>
<td></td>
<td>46</td>
<td>Binds RNA; ATP-dependent RNA helicase; promotes mRNA binding to 40S ribosomes</td>
</tr>
<tr>
<td>eIF4E</td>
<td></td>
<td>24</td>
<td>Binds to 5'-terminal 7-methyl-GTP cap on mRNA</td>
</tr>
<tr>
<td>eIF4G</td>
<td></td>
<td>173</td>
<td>Binds to PABP</td>
</tr>
<tr>
<td>eIF4B</td>
<td></td>
<td>80</td>
<td>Binds mRNA; promotes RNA helicase activity and mRNA binding to 40S ribosomes</td>
</tr>
<tr>
<td>eIF5</td>
<td></td>
<td>49</td>
<td>Promotes GTPase of eIF2, ejection of eIF2 and eIF3</td>
</tr>
<tr>
<td>eIF5B</td>
<td></td>
<td>175</td>
<td>Ribosome-dependent GTPase activity; mediates 40S and 60S joining</td>
</tr>
<tr>
<td>eIF6</td>
<td></td>
<td></td>
<td>Dissociates 80S; binds to 60S</td>
</tr>
</tbody>
</table>

**FIGURE 30.27** The characteristic structure of eukaryotic mRNAs. Untranslated regions ranging between 40 and 150 bases in length occur at both the 5' and 3' ends of the mature mRNA. An initiation codon at the 5' end, invariably AUG, signals the translation start site.
Eukaryotic 48s initiation complex
Regulation of Eukaryotic Translation Initiation

Biological Events related to translation

Different Cellular Functions: Red blood cell;

Different development Stages: Fast proliferation vs quiescent;

Tumor: Enhanced protein translation to support fast growth;

Nutrition Starvation: Cope with harsh conditions;

Viral Infection: Host cell defence against viral infection;

...
Regulation of Eukaryotic Translation Initiation

Regulation of initiation factor 4E

Regulation of initiation factor eIF-2
Regulation of initiation factor 4E

4E-BP binds to 4E as a limiting step for initiation.

Growth signals activate mTOR, which will in turn phosphorylate 4E-BP.

Phosphorylated 4E-BP dissociates from 4E.

Free 4E assemble with other eIF4 members and bind to the Cap of mRNA.

Translation initiated.

Stress/Starvation will block the activation of mTOR pathway.
Regulation of initiation factor 4E

- Closely related to metabolic and environmental status
4E-BP1 Is a Key Effector of the Oncogenic Activation of the AKT and ERK Signaling Pathways that Integrates Their Function in Tumors

Zhou X et al Clinical Cancer Research (intraductal hyperplasia (IDH), and ductal carcinoma in situ (DCIS))
Regulation of initiation factor eIF-2

GTP-form of Initiation factor eIF2 brings in Met-tRNAiMet

It convert into GDP form after provide energy for the assembly.

It recycled to GTP form by eIF2B GEF and reused in next round

Phosphorylation of eIF2 by eIF2a kinases prevent its recycle

Protein synthesis stopped due to the lack of GTP form of eIF2
eIF2a kinases activated under various stress conditions

eIF2a Kinases

Translation On                                Translation Off
HRI responsible for HS-induced translation inhibition

Translation Initiation Control by Heme-Regulated Eukaryotic Initiation Factor 2α Kinase in Erythroid Cells under Cytoplasmic Stresses

LINRONG LU,† AN-PING HAN, AND JANE-JANE CHEN*

Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Wild Type

Osmotic Stress

Arsenite

Heat Shock

Membrane

HRI-/-

PKR-/-

Cell Proliferation, Apoptosis and Differentiation

HRI-P

ROS

AA Starvation

ER Stress

eIF2αP & Protein Synthesis Inhibition

Heme Deficiency

eIF2αP

C As Os hs

C As hs Os

C As Os hs

eIF2αP

? ? ?
Activation of PKR upon viral infection

**Figure 4:** Activation of PKR by dsRNA leads to inhibition of protein synthesis.
In vitro Cell-Free protein translation

Rabbit Reticulocyte Lysate:
Rabbit reticulocyte lysate is a highly efficient in vitro eukaryotic protein synthesis system used for translation of exogenous RNAs (either natural or generated in vitro). In vivo, reticulocytes are highly specialized cells primarily responsible for the synthesis of hemoglobin, which represents more than 90% of the protein made in the reticulocyte.

Wheat Germ Extract:
Wheat germ extract is a convenient alternative to the rabbit reticulocyte lysate cell-free system. This extract has low background incorporation due to its low level of endogenous mRNA. Wheat germ lysate efficiently translates exogenous RNA from a variety of different organisms, from viruses and yeast to higher plants and mammals. The wheat germ extract is recommended for translation of RNA containing small fragments of double-stranded RNA or oxidized thiols, which are inhibitory to the rabbit reticulocyte lysate.

Human In Vitro Protein Expression System: Better modification

Cell-free bacteria expression system: Larger amount
In vitro Cell-Free protein translation

Retic: TAATACGACTCACTATAGGGAGAGCCACCAGTG
      T7 Promoter Sequence
      Translation Initiation Site

E. coli: TAATACGACTCACTATAGGGAGAGGAGGATATACAGTG
       T7 Promoter Sequence
       Translation Initiation Site

Circular or linear DNA template → 2-3μL

Transcription Reaction 1 hour, 32°C

1μg mRNA template from other source → 2-3μL

Translation Reaction 90 minutes, 30°C (28°C for glycosylation)
Definition of Protein Modification

Posttranslational modification
From Wikipedia, the free encyclopedia

Posttranslational modification (PTM) is the chemical modification of a protein after its translation. It is one of the later steps in protein biosynthesis for many proteins.

Posttranslational modification of amino acids extends the range of functions of the protein. It is estimated that the human proteome consists of ~300,000 different proteins, or about 10X more than the number of genes (!)

PTM includes the final trim of the peptide chain (remove amino acids from the amino end of the protein, or cut the peptide chain in the middle ) and attaching of other biochemical functional groups (such as acetate, phosphate, various lipids and carbohydrates), changing the chemical nature of an amino acid, or making structural changes.
Why are proteins modified?

- DNA $\rightarrow$ mRNA $\rightarrow$ Protein (nascent protein, precursor protein) $\rightarrow$ Mature protein $\rightarrow$ Biological active protein

- mRNA $\rightarrow$ translation
- Protein (nascent protein, precursor protein) $\rightarrow$ processing, post-translational modification
- Mature protein $\rightarrow$ folding
Why are proteins modified?

• Regulation of activity
  – turn activity on
  – turn activity off

• Protein-protein interaction
  – modification site may be a binding interface

• Subcellular localization
  – modification site may be a targeting signal
  – modification may be a membrane anchor

• Degradation
  – identify the protein for degradation
Post-transcriptional modifications includes formation or breakage of covalent bonds

- disulfide bond formation
- truncation
  - cleavage
  - cleavage of signal sequences - signal peptidases
  - protein splicing - proteolytic Processing
Proteolytic processing (蛋白酶解) upon activation

- Prevent premature activation of hydrolytic enzymes
  
Examples:

  Blood coagulation
  Caspases
Blood coagulation (凝血)

Tissue Factor Pathway

Contact Pathway

Blood coagulation involves several factors and enzymes, including Factor II, Factor VIIa, Factor X, Factor IX, Factor VIIIa, and Factor XIII. The process begins with the activation of Factor XII by tissue factor (TF) and factor XI (FXI). This leads to the formation of TF/FXIIa/FXIIIa, which in turn activates Factor V (FVa) and Factor VIII (FVIIIa). FVa and FVIIIa then activate Factor X (FXa), which converts Factor IX (FIX) to FIXa. FIXa then activates Factor X (FXa), forming a prothrombinase complex that converts prothrombin to thrombin. Thrombin then converts fibrinogen to fibrin, resulting in blood clot formation.

Key Figures:
- A1, A2, B, A3, C1, C2: Different fragments of blood coagulation factors.
- VIII precursor: Precursor form of Factor VIII.
- Proteolytic processing & Secretion: Proteolytic processing and secretion of blood coagulation factors.
- von Willebrand Factor: Essential for platelet adhesion and coagulation.
- Calcium (Ca++) activation: Critical for the activation of blood coagulation factors.

Blood clotting is a complex process involving multiple steps and factors, crucial for maintaining vascular integrity and preventing blood loss.
Caspase

(cysteine-containing aspartate-specific protease)

1. First identified in 1993
2. Total 14 members (11 in human), and caspase-2, 3, 6, 7, 8, 9, and 10 are known to be involved in apoptosis
3. A milestone in apoptosis study
4. Known as the execution machinery of apoptosis
Activation of Caspase by Proteolytic processing

N-peptide  Asp-X  large subunit  Asp-X  small subunit

Inactive monomer  Dimerization and interchain cleavage  Active caspase 8

Procasparse-3
Activated Caspase-3
Actin
### Post-transcriptional modifications includes modification of side chains

<table>
<thead>
<tr>
<th>Modification (中文翻译)</th>
<th>a.a. side chain (英文)</th>
<th>图例</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycosylation (糖基化)</td>
<td>Asn, Ser</td>
<td></td>
</tr>
<tr>
<td>Phosphorylation (磷酸化)</td>
<td>Ser, Thr, Tyr</td>
<td>![磷酸化示意图]</td>
</tr>
<tr>
<td>Acetylation (乙酰化)</td>
<td>Lys</td>
<td>![乙酰化示意图]</td>
</tr>
<tr>
<td>Methylation (甲基化)</td>
<td>Lys, Arg</td>
<td>![甲基化示意图]</td>
</tr>
<tr>
<td>Ubiquitination (泛素化)</td>
<td>Lys</td>
<td>![泛素化示意图]</td>
</tr>
<tr>
<td>SUMOylation (small ubiquitin-related modifier)</td>
<td>Lys</td>
<td>![SUMOylation示意图]</td>
</tr>
<tr>
<td>Myristylation (十四烷基化) and Farnesylation (法尼基化)</td>
<td>N-, C-, terminal</td>
<td>![Myristylation和Farnesylation示意图]</td>
</tr>
</tbody>
</table>
Methods to detect protein modifications

- Using antibodies:
  - Western Blot - Antibody against the modification mutation analysis to characterize modification sites

- 1D or 2D gel
- MS (mass spectrometry)  \( \text{Change in MW and charge} \)
- Other methods.....
Phosphorylation (磷酸化)

- Most common posttranslational modification to proteins in eukaryotes
- Phosphorylation usually occurs on serine, threonine, and tyrosine residues in eukaryotic proteins
- Enzymes and regulators are turned ‘on’ and ‘off’
- Energy from ATP
protein phosphorylation is the process in which a phosphoryl group (phosphate), donated by ATP, is transferred to an acceptor protein. The reaction is catalysed by a protein kinase.

___

Note the important conformational change that occurs in the cytosolic segment of the insulin receptor upon phosphorylation of its activation segment (on tyrosines). The change has important consequences for protein activity.
Phosphorylation and signaling transduction

- Stimulus
- Growth factor
- Integrin
- Oxidative stress
- IL-1

  - Activator
  - RasGTP
  - Rac1
  - Src
  - TRAF6-TAB1/2

    - MKKK
    - c-Raf1
    - MEKK1
    - MEKK2
    - TAK1

      - MKK
      - MKK1
      - MKK4
      - MKK5
      - MKK6

        - MAPK
        - ERK1
        - JNK1
        - ERK5
        - p38

          - Substrate
          - p90RSK
          - c-Jun
          - MEF2
          - MNK1
Protein phosphorylation is catalyzed by Kinases

A protein kinase is an enzyme that modifies other proteins by chemically adding phosphate groups to them (phosphorylation). Phosphorylation usually results in a functional change of the target protein by changing enzyme activity, cellular location, or association with other proteins.

The human genome contains about 500 protein kinase genes and they constitute about 2% of all human genes. Protein kinases are known to regulate the majority of cellular pathways, especially those involved in signal transduction.
Conserved Kinase Domain

Kinase Domain
Protein Kinase Facts

1. Serine/Threonine kinase vs Tyrosine kinase

2. Receptor tyrosine kinase vs receptor associated kinase

3. Dimerization and autophosphorylation

4. Protein kinase inhibitors

5. Protein kinase assay
Phosphorylation can be reversed (dephosphorylated) by phosphotase.

‘Classical’ PTPs

Methods to detect protein phosphorylation

- Western Blot, IF.
- 1D or 2D gel
- MS
Examples

Some protein will shift to the high molecular site when it was phosphorylated.

Phosphorylated protein can be detected with antibodies that recognize the phosphorylation sites.
Phospho-specific antibodies to specific phosphorylation sites

A

hGR

S113  S141  S224  S317  S508

DBD  LBD (AF-2)

S203  S211

L-Q-D-L-E-F-S-S-G-S-P-G-K-E

G-S-P-G-K-E-T-N-E-S-P-W-R-S

194  207  202  215

S203  S211

B

Immunoblot

U2OS-hGR: WT  203A  211A

Dex:  -  +  -  +  -  +

GR-P-203

GR-P-211

HA (Total GR)

1  2  3  4  5  6
**Choose** the right Ab and **mark** it clearly in your notebook/manuscript.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Applications</th>
<th>Reactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>4056</td>
<td>Phospho-Akt (Thr308) (28F9) Rabbit mAb</td>
<td>W, IF, IF-IC</td>
<td>H, M, R, Mk</td>
</tr>
<tr>
<td>2965</td>
<td>Phospho-Akt (Thr308) (C81E5E) Rabbit mAb</td>
<td>W, IF-IC, F</td>
<td>H, M, R, Hm, Mk</td>
</tr>
<tr>
<td>2918</td>
<td>Phospho-Akt (Thr308) (C81E5E) Rabbit mAb (Alexa Fluor® 488 Conjugate)</td>
<td>F</td>
<td>H, M, R, Mk</td>
</tr>
<tr>
<td>3378</td>
<td>Phospho-Akt (Thr308) (C81E5E) Rabbit mAb (Alexa Fluor® 647 Conjugate)</td>
<td>F</td>
<td>H, M, R, Hm, Mk</td>
</tr>
<tr>
<td>5086</td>
<td>Phospho-Akt (Thr308) (C81E5E) Rabbit mAb (Biotinylated)</td>
<td>W</td>
<td>H, M, R, Hm, Mk</td>
</tr>
<tr>
<td>9033</td>
<td>Phospho-Akt (Thr308) (C81E5E) Rabbit mAb (PE Conjugate)</td>
<td>F</td>
<td>H, M</td>
</tr>
<tr>
<td>5106</td>
<td>Phospho-Akt (Thr308) (L32A4) Mouse mAb</td>
<td>W</td>
<td>H, M, R, Mk</td>
</tr>
<tr>
<td>9278</td>
<td>Phospho-Akt (Thr308) Antibody</td>
<td>W, IF, F</td>
<td>H, M, R, Hm</td>
</tr>
<tr>
<td>9267</td>
<td>Phospho-Akt (Thr460) Antibody</td>
<td>W, IF, IF-IC</td>
<td>H, M, R</td>
</tr>
<tr>
<td>4060</td>
<td>Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb</td>
<td>W, IF, IF-IC, F, IF-IC-F</td>
<td>H, M, R, Hm, Mk, Dm, Z, B (C)</td>
</tr>
<tr>
<td>4071</td>
<td>Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb (Alexa Fluor® 488 Conjugate)</td>
<td>F</td>
<td>H, M, R</td>
</tr>
<tr>
<td>4075</td>
<td>Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb (Alexa Fluor® 647 Conjugate)</td>
<td>F</td>
<td>H, M, R</td>
</tr>
</tbody>
</table>
Phosphorylation changes PI of the protein, which can be easily detected on 2D gels.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEK + ERK + PO₄</td>
<td>silver</td>
</tr>
<tr>
<td>ERK + PO₄</td>
<td>silver</td>
</tr>
<tr>
<td>MEK + ERK + PO₄</td>
<td>PSer WB</td>
</tr>
<tr>
<td>ERK + PO₄</td>
<td>PSer WB</td>
</tr>
</tbody>
</table>

*MEK kinase*  
*ERK Substrate*  
*Aligned 2D gel sections*
MS analysis of C/EBPβ after in vitro phosphorylation by MAPK and/or GSK3β

Glycosylation (糖基化)

• **Major form of protein modification:** Most soluble and membrane-bound proteins made in the ER are glycoproteins, in contrast to cytosolic proteins.

• **N-linked glycosylation** (Asn, Asparagine 天冬酰胺): N-linked glycosylation is important for the folding of some eukaryotic proteins. The N-linked glycosylation process occurs in eukaryotes, but very rarely in bacteria. Initial glycosylation in the ER, processing in the Golgi.

• **O-linked glycosylation** (Ser, Thr) happened in the Golgi

• Many different forms and functions

• Glycoprotein synthesis is a 3-step process:
Why the "Glyco" Perspective?

- The cell surface is coated with sugar residues - **glycocalyx**
  - oligosaccharide chains attached to membrane protein (**glycoprotein**) and lipids (**glycolipids**)  
  - polysaccharide chains, free (**glycosaminoglycan** chains), or attached to a polypeptide core (**proteoglycan**)

- Why bother with the total **Glycosylation Profile**?

  - Just like the **proteome**, glycosylation profile is a characteristic of a cell at a particular physiological stage
  - Cancer cells are often marked by **aberrant glycosylation**
  - Cell-cell, host-pathogen, host-parasite interactions are governed by both protein-protein and **protein-glycan interactions**
N-linked glycosylation

- A precursor oligosaccharide is formed on dolichol (多萜醇) lipid

- This is transferred to the growing protein catalyzed by oligosaccharyl transferase.

- Attaches to NH₂ side chain of Asn but only in the context: Asn-x-Ser or Asn-x-Thr
Process of N-linked glycosylation

1. ER LUMEN
   - Asn
   - Glucosidase I, II
   - ER mannosidase

2. Golgi lummen
   - Asn
   - Golgi mannosidase I, II
   - N-acetylglucosamine transferase I
   - UDP

3. Asn
   - Endo H-sensitive
   - Endo H-resistant
   - next added here

4. Asn
   - high-mannose oligosaccharide

5. Asn
   - complex oligosaccharide

KEY:
- ▼ = N-acetylglucosamine (GlcNAc)
- ▲ = mannose (Man)
- □ = glucose (Glc)
- ▲ = galactose (Gal)
- ▼ = N-acetylneuraminic acid (sialic acid, or NANA)

乙酰葡萄糖胺 甘露糖 葡萄糖 半乳糖 乙酰神经氨酸（唾液酸）
多样性
Functions of glycosylation

- Proper folding 正确折叠
- Stabilise proteins against proteolysis 防止降解
- Form ECM (extracellular matrix) 胞外基质
- Modulation of immune response 免疫反应
  Selectins bind to oligosaccharides, cell-cell interactions
- Anchor proteins on membrane 锚定作用
  
  GPI (Glycosylphosphatidylinositol 糖基磷脂酰肌醇) anchored proteins
Protein engineering strategies (taken Glycosylation in consideration)

1. Some Recombinant proteins from E. Coli not functional;

2. Inserting additional N-linked sites has been proven to be clinically beneficial for recombinant glycoproteins.

Two new N-linked glycosylation sites were incorporated into recombinant human erythropoietin (rhEPO, 促红细胞生成素) via site-directed mutagenesis, which substantially increased in vivo activity (Elliott et al. 2003; Elliott, Chang, et al. 2004; Elliott, Egrie, et al. 2004).

In another study (Perlman et al. 2003), protein engineering was applied to FSH (促卵泡激素) by introducing additional N-linked glycosylation sites into the molecule. The resulting molecule (FSH1208) was found to have a 3- to 4-fold increased serum half-life, compared with wild-type recombinant FSH.

Incorporating new N-linked glycosylation site to improve activity and stability for recombinant antibody fragments (Stork et al. 2008).
GPI anchored proteins

Glycosylphosphatidylinositol: 糖基化磷脂酰肌醇

---

**A**

![Diagram of GPI anchored proteins]

**B**

<table>
<thead>
<tr>
<th>Protein</th>
<th>GPI signal sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylcholinesterase (<em>Torpedo</em>)</td>
<td>NQFLPKLLNATA<strong>C</strong></td>
</tr>
<tr>
<td>Alkaline phosphatase (<em>placenta</em>)</td>
<td><strong>D</strong>GELSSSSGTSSSKGIIFYVLSILYLYF</td>
</tr>
<tr>
<td>Decay accelerating factor</td>
<td></td>
</tr>
<tr>
<td>PARP (<em>T. Brucei</em>)</td>
<td>EPEPEPEPEPEPEPG</td>
</tr>
<tr>
<td>Prion protein (<em>hamster</em>)</td>
<td>QKESQAYYDEGRRS</td>
</tr>
<tr>
<td>Thy-1 (<em>rat</em>)</td>
<td>KTINVIRDKLVKC</td>
</tr>
<tr>
<td>Variant surface glycoprotein</td>
<td>ESNCKWENNA<strong>C</strong>KD</td>
</tr>
<tr>
<td>(<em>T. Brucei</em>)</td>
<td><strong>S</strong> SILVTKKFALTVSAAF<strong>V</strong>ALLF</td>
</tr>
</tbody>
</table>

*Boldfaced amino acid is the site of attachment of the GPI. Sequence to the right of the space is cleaved from the protein by the transpeptidase upon anchor addition.*
Methods to detect protein glycosylation

- Western Blot
- 1D or 2D gel
- MS (mass spectrometry)
- Pro-Q Emerald staining
Example

Pro-Q Emerald staining  Anti-Glycan Western Blot

Cell. 2009, 137(2):321-31
Acetylation (乙酰化)

Acetic Anhydride + Amino Terminus of Protein or Peptide → Acetylated Amino Terminus
Acetylation on Lysine
Histone acetylation and deacetylation

Histone acetyltransferases (HATs) catalyze the addition of an acetyl group to lysine residues in histones, while histone deacetylases (HDACs) remove the acetyl group, leading to reversible modifications that affect gene expression.
Histone acetylation and deacetylation

• A major part of epigenetics (表观遗传学)

• Acetylation enhances transcription

• Deacetylation represses transcription

• HAT (Histone acetyltransferase)

• HDAC (Histone deacetylase)
Acetylation of signalling molecules

- Type I Interferon Receptor signalling molecules, such as IFNaR2, IRF9, STAT2 (*Cell*. 2007 Oct 5;131(1):93-105)

- The acetylation of these molecules play a critical role in signal transduction.
Acetylation-dependent signal transduction for type I interferon receptor

Cell. 2007 Oct 5;131(1):93-105
Methylation (甲基化)

Lys

- Protein
  - Unmethylated lysine

- Protein
  - Monomethyl-lysine

- Protein
  - Dimethyl-lysine

- Protein
  - Trimethyl-lysine

Arg

- Protein
  - N<sup>G</sup>-monomethyl-arginine

- Protein
  - N<sup>ε,N<sup>G</sup></sup>-dimethyl-arginine (Type I)

- Protein
  - N<sup>ε,N<sup>G</sup></sup>-dimethyl-arginine (Type II)

- Protein
  - Citrulline

- Protein
  - Asymmetric

AdoMet
AdoHcy
HCHO
Histone methylation

**Histone methylation** is the modification of certain amino acids in a histone protein by the addition of one, two, or three methyl groups.

**Functions:**
A major part of epigenetics.

*Methylation at different sites of histones may have different effects on gene transcription.*

Histone is methylated by Histone methyltransferases (HMT)

Demethylation was mediated by Histone Demethylase
Histone methylation

Transcriptionally Inactive Chromatin

Pericentric Heterochromatin

Suv39h, G9a, L3MBTG1

H3 Lys9 Methyl

LSD1, JMJD1, JMJD2

Pericentric Heterochromatin

Inactive X Chromosome

Rb-Mediated Repression

Active X Chromosome

Hox Gene Repression

Nuclear Receptor Mediated Activation

Transcriptionally Active Chromatin

Transcription Initiation and Elongation

Dot1L

H3 Lys79 Methyl

Anti-Silencing
Histone modification and epigenetic regulation
Methods for histone methylation detection

- Western Blot
- 1D & 2D gel
- MS
- ChIP (Chromatin Immunoprecipitation, it can detect histone acetylation or methylation at specific genomic site)
Detecting histone modifications using ChIP method

Myc increases histone H4 acetylation at rDNA

Ubiquitination (泛素化)

Ubiquitin (泛素)

- Ubiquitin is a small regulatory protein that has been found in almost all tissues (ubiquitously) of eukaryotic organisms.
- 76 amino acids, 8.5 kDa:
  MQIFVKTLTGKITLEVEPSDITIENVK
  AKIQDKEGIPPDQQRLLFAGKQLEDG
  RTLSDYNIQKESTHLVLRLRGG
Protein degradation

Two Pathways

extracellular: digestive proteases

lysosomes
Non-specific, energy free

intracellular

proteasome mechanisms
Ubiquitin-mediated, needs ATP
Ubiquitination (泛素化)

History of the Finding of Ubiquitin:

Before 1970 – Lysosomal pathway
1977 – Etlinger JD, Goldberg
    Protein degradation in reticulocyte: no lysosomes; ATP dependent
1978 – Ciechanover & Hershko (BBRC)
    Identify protein degradation system in reticulocyte
    Two fractions – (ATP-dependent Proteolytic Factor; APF)
    APF-1 fraction conjugates to proteins
1980 – Wilkinson et al
    APF-1 was identified as ubiquitin
80s – Varshavsky
    Identification of E1, E2, E3 in yeast
1988 – “Proteasome”
Aaron Ciechanover  Avram Hershko  Irwin Rose

The Nobel Prize in Chemistry 2004 was awarded jointly to Aaron Ciechanover, Avram Hershko and Irwin Rose "for the discovery of ubiquitin-mediated protein degradation"

From nobelprize.org
Ubiquitination (泛素化)

1. Biochemical reactions of ubiquitination
Specificity of ubiquitination determined by different E3 ligases

脯氨酰羟化
Mono-Ubiquitination for transcription activation

- Mono-ubiquitination
  - Regulate endocytosis, gene expression, protein sorting, subnuclear trafficking
K63 Ubiquitination for signaling

a) Early TNF signaling
- Extracellular compartment
- TNF
- TNFR1
- TRADD
- TRAF2,5
- RIP
- Upregulation of A20
- IKK complex activation
- $\kappa B$ degradation
- NF-$\kappa B$ activation
- Transcription of target genes (such as A20)

b) Late TNF signaling
- TNF
- TNFR1
- TRADD
- TRAF2,5
- RIP
- K63 deubiquitination
- TAX1BP1
- Itch
- A20
- Degradation
- Termination of NF-$\kappa B$ activation

K63-linked ubiquitination
Multiple Forms and Functions of Ubiquitination

- **Unanchored**
  - Signal transduction

- **Endocytosis**
  - Trafficking
  - Gene expression/silencing
  - Signal transduction
  - Proteasomal degradation

- **Signal transduction**
  - Trafficking
  - DNA repair
  - Translation
  - Proteasomal degradation

- **Proteasomal degradation**
  - Signal transduction
  - Proteasomal degradation
  - Trafficking
  - Cell cycle control
  - DNA damage control
  - DNA repair
  - Proteasomal degradation

**Substrate** + E1/E2/E3 + Ub
Protein ubiquitination detection

• Western Blot
• 1D or 2D gel
• MS
The oncoprotein SS18-SSX1 promotes p53 ubiquitination and degradation by enhancing HDM2 stability.

SUMOylation (类泛素化)

SUMO (small ubiquitin-related modifier) proteins are small protein tags that are conjugated to proteins to modify their function.

Three different SUMO proteins are conjugated to proteins, SUMO-1, SUMO-2 and SUMO-3.

The ubiquitin system tags proteins for degradation by the proteosome but SUMO conjugation has a range of other functions, stabilizing some proteins and altering their subcellular localization.

Sumoylation may also influence ubiquitination and protein stability indirectly.
SUMOylation (small ubiquitin-related modifier)

- SUMO are around 100 amino acids in length and 12 kDa in mass.
- Three different SUMO proteins are conjugated to proteins, SUMO-1, SUMO-2 and SUMO-3.
- Function of SUMOylation: not degradation, but regulation.
- SUMOylation-Ubiquitination crosstalk.

*Cell.* 2010, 143
# Molecular consequences of sumoylation

## SUMO mechanisms

<table>
<thead>
<tr>
<th>MECHANISM</th>
<th>MODEL</th>
<th>SELECTED EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruiting factors</td>
<td><img src="image" alt="SUMO recruitment model" /></td>
<td>PCNA-Srs2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PML nuclear bodies assembly</td>
</tr>
<tr>
<td>Inhibition of interactions</td>
<td><img src="image" alt="Inhibition model" /></td>
<td>PCNA$_{K127}$-Eco1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2-25K-ubiquitin E1</td>
</tr>
<tr>
<td>Intramolecular conformational switch</td>
<td><img src="image" alt="Conformational switch model" /></td>
<td>TDG</td>
</tr>
</tbody>
</table>
Sumo-Ubiquitin Crosstalk

1. **SUMO prevents ubiquitylation at specific lysine residues.** PCNA, Smad4, and IκBα are all examples of substrates where a single lysine residue can be either sumoylated or ubiquitylated.

Hypoxia induces proinflammatory genes through Ub-dependent degradation of CREB (cAMP response element binding protein). CREB can also be sumoylated and is stabilized by SUMO overexpression.

2. **SUMO also appears to have a separate function in addition to preventing ubiquitylation.** The transcription factor Smad4 is protected from Ub-dependent proteolysis by attachment of SUMO at its ubiquitylation site, but there is also evidence that sumoylation separately promotes nuclear retention of Smad4 (26). In another example, sumoylation of PCNA inhibits Ub-dependent postreplication DNA repair, consistent with a function for SUMO in blocking ubiquitylation. However, sumoylated PCNA can itself promote spontaneous mutagenesis through the postreplication repair pathway, indicating an independent role for SUMO.
Signaling functions of SUMO

Higher-order chromatin structure

Antagonism of ubiquitination

Differential protein-protein interaction

Nuclear pore complex shuttling

Subcellular localization

Integrity of PML nuclear bodies

Transcription factor activity

Regulation of DNA binding
### Examples of SUMO substrates

<table>
<thead>
<tr>
<th>BIOLOGICAL PROCESS</th>
<th>SUMO SUBSTRATE</th>
<th>FUNCTION</th>
<th>MECHANISM</th>
<th>ORGANISM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytokinesis</td>
<td>Cdc3, 11, Shs1</td>
<td>Septin ring disassembly?</td>
<td>Not clear</td>
<td>Yeast</td>
</tr>
<tr>
<td>Ion channel activity</td>
<td>KCNA5</td>
<td>Modulation of activity</td>
<td>Not clear</td>
<td>Mammal</td>
</tr>
<tr>
<td>Receptor activity</td>
<td>GLUK2, TjRI</td>
<td>Promotion of internalization, Increased activity</td>
<td>Not clear, Recruitment of Smad3</td>
<td>Mammal</td>
</tr>
<tr>
<td>Glucocorticoid receptor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signal transduction</td>
<td>Smad4</td>
<td>Signal modulation</td>
<td>Not clear</td>
<td>Mammal</td>
</tr>
<tr>
<td>Transport activity</td>
<td>RanGAP1</td>
<td>Translocation</td>
<td>RanBP2 association</td>
<td>Mammal</td>
</tr>
<tr>
<td>Subcompartimentalization</td>
<td>PML</td>
<td>Formation of PML bodies</td>
<td>SUMO-SIM interactions</td>
<td>Mammal</td>
</tr>
<tr>
<td>Ribosome biogenesis</td>
<td>Nop7</td>
<td>Ribosome biogenesis and export</td>
<td>Not clear</td>
<td>Yeast</td>
</tr>
<tr>
<td>DNA replication</td>
<td>PCNA</td>
<td>Inhibition of recombination</td>
<td>Recruitment of Srs2</td>
<td>Yeast</td>
</tr>
<tr>
<td>Transcription</td>
<td>Mot1, Sp3</td>
<td>STUbL-mediated degradation, Repression of transcription</td>
<td>STUbL (Sx9/Sx8), Recruitment of histones methyltransferases</td>
<td>Mammal</td>
</tr>
<tr>
<td>Double-strand break repair</td>
<td>BRCA1, 53BP1, Rad52</td>
<td>Promotion of DSB repair, Regulation of Rad52 activity</td>
<td>Recruitment, increased ligase activity (in vitro), Stabilization, localization</td>
<td>Mammal</td>
</tr>
<tr>
<td>Base excision repair</td>
<td>TDG</td>
<td>Release from DNA</td>
<td>Conformational switch</td>
<td>Mammal</td>
</tr>
<tr>
<td>Mitosis</td>
<td>Nu10, Nu80, Aurora B</td>
<td>Accurate mitosis</td>
<td>Not clear, change of localization</td>
<td>Yeast, Mammal</td>
</tr>
<tr>
<td>Meiosis</td>
<td>Red1</td>
<td>Synaptosomal complex formation</td>
<td>Binding to synaptosomal complex protein Zip1</td>
<td>Yeast</td>
</tr>
<tr>
<td>Telomere maintenance</td>
<td>TRF1, TRF2</td>
<td>Telomere length homeostasis</td>
<td>Recruitment to PML nuclear bodies</td>
<td>Mammal</td>
</tr>
</tbody>
</table>
Myristylation (十四烷基化) and Farnesylation (法尼基化)

- Attaches cytosolic proteins to the plasma membrane.
- Protein usually involved in signal transduction.
Thanks for your attention!