Translation - Prokaryotes

Shine-Dalgarno (SD) Sequence
rRNA 3’-GAUACCAUCCUCCUUA-5’
mRNA ....GGAGG...(5–7bp)....AUG

Influences:

Secondary structure!! SD and AUG in unstructured region

Surrounding of SD and AUG!!!

Ribosomal protein S1: present only in Gram-negatives (not in Gram-positives):
→ binds to AU-rich sequences found in many prokaryotic mRNAs 15-30 nucleotides upstream of start-codon

Translational coupling

Start

AUG 91%
GUG 8
UUG 1
Translation - Eukaryotes

Start Codon

mRNA  5′-CAP......AUG

Influences:

Surrounding of AUG!!!

Kozak Consensus

........CC^A/GCC^AUGG...... mammalian

........^A/TAA^A/C^A^A/C^A^AUGTCT/C........ Yeast

........ gccgcc(A/G)cc^AUGG .......... Wikipedia
Translation elongation

- Codon usage
- Secondary structures
- Codon structure – translational frameshifting

```
AAAAAAA AAAA UCA
Lys  Lys  Lys  Ser
```

```
AAAAAAA AAAA UCA
Lys  Lys  Lys  Ile
```
**Figure 7.1** All the triplet codons have meaning: 61 represent amino acids, and 3 cause termination (STOP).

**Figure 7.2** The number of codons for each amino acid does not correlate closely with its frequency of use in proteins.
Figure 7.3 Third bases have the least influence on codon meanings. Boxes indicate groups of codons within which third-base degeneracy ensures that the meaning is the same.

Figure 7.4 Codon-anticodon pairing involves wobbling at the third position.

<table>
<thead>
<tr>
<th>Base in First Position of Anticodon</th>
<th>Base(s) Recognized in Third Position of Codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>A or G</td>
</tr>
<tr>
<td>C</td>
<td>G only</td>
</tr>
<tr>
<td>A</td>
<td>U only</td>
</tr>
<tr>
<td>G</td>
<td>C or U</td>
</tr>
</tbody>
</table>
Universial Triplet Code → rare exemptions

<table>
<thead>
<tr>
<th>Codon</th>
<th>Universal code</th>
<th>Mycoplasma</th>
<th>Paramecium Euplotes</th>
<th>Yeast</th>
<th>Protozoa</th>
<th>Mammals</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGA</td>
<td>Stop</td>
<td>Tryptophan</td>
<td>Stop</td>
<td>Cysteine</td>
<td>Stop</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>UAA/UAG</td>
<td>Stop</td>
<td>Stop</td>
<td>Glutamine</td>
<td>Stop</td>
<td>Stop</td>
<td>Stop</td>
</tr>
<tr>
<td>AUA</td>
<td>Isoleucine</td>
<td>Isoleucine</td>
<td>Isoleucine</td>
<td>Isoleucine</td>
<td>Methionine</td>
<td>Methionine</td>
</tr>
<tr>
<td>CUA</td>
<td>Leucine</td>
<td>Leucine</td>
<td>Leucine</td>
<td>Leucine</td>
<td>Threonine</td>
<td>Leucine</td>
</tr>
<tr>
<td>AGA/AGG</td>
<td>Arginine</td>
<td>Arginine</td>
<td>Arginine</td>
<td>Arginine</td>
<td>Arginine</td>
<td>Stop</td>
</tr>
</tbody>
</table>

The universal genetic code is used in the chromosomes of most cells, chloroplasts, plant mitochondria, and their viruses and plasmids. A few organisms use slightly different codes in their chromosomes (in the nucleus). The examples of these other nuclear codes are from Mycoplasma (Bacteria) and two different ciliated protozoa (Eukarya). All nonplant mitochondria use variations of the universal code, whereas plant mitochondria use the universal code. The examples here are only a few of the different types known.
**Figure 6.1** Ribosomes are large ribonucleoprotein particles that contain more RNA than protein and dissociate into large and small subunits.

<table>
<thead>
<tr>
<th>Ribosomes</th>
<th>rRNAs</th>
<th>Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterial</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50S</td>
<td>23S = 2904 bases</td>
<td>31</td>
</tr>
<tr>
<td>70S mass: 2.5 × 10^6 D</td>
<td>5S = 120 bases</td>
<td></td>
</tr>
<tr>
<td>30S</td>
<td>16S = 1542 bases</td>
<td>21</td>
</tr>
<tr>
<td><strong>Mammalian</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60S</td>
<td>28S = 4718 bases</td>
<td>49</td>
</tr>
<tr>
<td>80S mass: 4.2 × 10^6 D</td>
<td>5.8S = 160 bases</td>
<td></td>
</tr>
<tr>
<td>40S</td>
<td>18S = 1874 bases</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>5S = 120 bases</td>
<td></td>
</tr>
</tbody>
</table>
Genetik und Gentechnik

H. Schwab

Figure 11.8: Initiation of protein synthesis. (A) The initiation complex forms at the 5' end of the mRNA. (B) This consists of one 40S ribosomal subunit, the initiator tRNA^Met, and the eIF initiation factors. (C) The initiation complex recruits a 60S ribosomal subunit in which the tRNA^Met occupies the P (peptidyl) site of the ribosome. This complex travels along the mRNA until the first AUG is encountered, at which codon (translation) begins.

Figure 11.9: Elongation cycle in protein synthesis. (A) Pretranslocation state. (B) Posttranslocation state, in which an uncharged tRNA occupies the E site and the polypeptide is attached to the tRNA in the P site. (C) The function of EF-1α is to release the uncharged tRNA and bring the next charged tRNA into the A site, at which time a peptide bond is formed between the polypeptide and the amino acid held in the A site, in this case Glu. Simultaneously, the 60S subunit is shifted relative to the 40S subunit, re-creating the pretranslocation state. (D) The function of EF-2 is to translocate the 40S ribosome to the next codon, once again generating the posttranslocation state.
Figure 11.21 Termination of protein synthesis. When a stop codon is reached (A), no tRNA can bind to that site (B), which causes the release of the newly formed polypeptide and the remaining bound tRNA (C).
Regulation of Gene Expression

Prokaryotes

*Escherichia coli*

Lactose Metabolism

Absence of lactose → Only few molecules of β-galactosidase per cell

Presence of lactose → about 5000 molecules of β-galactosidase per cell

Not enzyme is inhibited, enzyme synthesis is affected

Detailed biochemical and genetic analysis

Jacob, Monod, Pardee → Nobel prize
$lac$ expression responds to inducer

Add inducer

Remove inducer

Level of $lac$ mRNA

Induced level

Basal level

0 2 4 6 8 10 12 min

Level of $\beta$-galactosidase

Lag

Induced level

Basal level

©virtualtext www.ergito.com
lac-Operon

The lac operon includes cis-acting regulator elements and protein-coding structural genes.

DNA

<table>
<thead>
<tr>
<th>lacZ</th>
<th>lacY</th>
<th>lacA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3510</td>
<td>780</td>
<td>825</td>
</tr>
</tbody>
</table>

Protein

- β-galactosidase
- Permease
- Transacetylase

Ort I

Ort O
Heterogenote analysis

Cis-configuration

Trans-configuration

inducible

constitutive
Heterogenote analysis

Cis-configuration

Trans-configuration

inducible
Model for behaviour of heterogenotes

\( \text{lacO} \rightarrow \) located adjacent to \( \text{lacZ} \), mutation in \( \text{lacO} \) results in loss of regulatory function when connected to \( \text{lacZ} \), no complementation by wt-allele in trans

\( \text{lacI} \rightarrow \) located upstream of \( \text{lacZ} \), mutation in \( \text{lacI} \) results in maintenance of regulatory function in both configurations to \( \text{lacZ} \) complementation by wt-allele

\( \text{lacO} \rightarrow \) DNA locus, mobile factor binds there and represses synthesis

\( \text{lacI} \rightarrow \) encodes a mobile factor (= protein) which binds at \( \text{lacO} \)
The lac operon includes *cis*-acting regulator elements and protein-coding structural genes.

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>lacI</th>
<th>PO</th>
<th>lacZ</th>
<th>lacY</th>
<th>lacA</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>1040</td>
<td>82</td>
<td>3510</td>
<td>780</td>
<td>825</td>
<td></td>
</tr>
<tr>
<td>mRNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>Repressor</td>
<td>β-galactosidase</td>
<td>Permease</td>
<td>Transacetylase</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

©virtualtext  www.ergito.com
Gene isolation
lac operon

Isolation of Lac Repressor
lacI<sup>q</sup> mutant

Binding studies
The promoter and operator overlap

Startpoint

Unwinding

Promoter binds RNA polymerase

Operator binds repressor

©virtualtext  www.ergito.com
A repressor tetramer binds the operator to prevent transcription

Monomer

Tetramer

\textit{lacI} gene synthesizes repressor monomer that forms tetramer

Tetramer binds to operator and blocks transcription

\textit{lacI} \quad \text{Promoter/Operator} \quad \textit{lacZ} \quad \textit{lacY} \quad \textit{lacA}
Inducer: β-1,6- allolactose (by product of β-galactosidase produced by transglucosylation)

Inducer inactivates repressor, allowing gene expression

RNA polymerase binds at promoter & transcribes mRNA

mRNA is translated into all 3 proteins

β-galactosidase Permease Transacetylasel
Mutant $O^c$

Mutation in $lacO$ prevents binding of LacI Repressor protein to Operator
Mutant $I^-$

Mutation in $lacI$ no binding capacity of LacI repressor protein
**Figure 10.9** Mutations map the regions of the *lacI* gene responsible for different functions. The DNA-binding domain is identified by *lacI*-d mutations at the N-terminal region; *lacI* mutations unable to form tetramers are located between residues 220–280; other *lacI* mutations occur throughout the gene; *lacI* mutations occur in regularly spaced clusters between residues 62–300.
The *lac* operator has dyad symmetry

```
```

Axis of symmetry

Mutations and chemical contacts identify key positions in the operator

**Constitutive mutations**

A T
ATA CAT
C G T

**Bases that contact repressor**

```
```

Protected by bound repressor

```
-10 -5 +1 +5 +10 +15 +20 +25
```

Axis of symmetry
Repressor is a tetramer of two dimers

DNA-binding sites

Core domains

1

2

Oligomerization

2

1

Two models for releasing repressor

Inducer binds to free repressor to upset equilibrium with bound repressor

Inducer binds directly to release repressor from operator
Lac repressor has several domains

- Helix-turn-helix
- α-helix
- β-sheet
- α-helix
- Core domain 1
- Inducer-binding site
- Core domain 2
- Oligomerization

Monomers → dimers → tetramers

- Inducer-binding cleft
- Hydrophobic core
- C-terminal helices

Two dimers make a tetramer

Mutations identify functional sites

- $I^S$ at dimer interface
- $I^S$ at inducer cleft
- Oligomerization

Oligomerization

www.ergito.com
**Figure 10.14** Inducer changes the structure of the core so that the headpieces of a repressor dimer are no longer in an orientation that permits binding to DNA. Photographs kindly provided by Mitchell Lewis.
DNA loops out between repressors

Repressor can make a loop in DNA
LacI repressor has general low affinity to DNA → Unspecific weak binding

LacI repressor has high affinity to specific operon Region on DNA → Specific strong binding
Operators for TrpR have related sequences

**aroH**

\[ \text{GCCG} \text{AATTGACTA} \text{AGAACTAGTGCA} \text{ATTAGGCTTTATTTTTTTGTATCATGCTAA} \]

mRNA

**trp**

\[ \text{AATC} \text{ATCGAAGTTAACATAGTGACGA} \]

mRNA

**trpR**

\[ \text{TGCTATCGTACTCTTTAGCGATACAACC} \]

mRNA

Operator region
Operators are close to the promoter

- gal
- aroH
- trp
- trpR
- lac

Promoter

Operator locations

Startpoint
Negative Regulation

**Figure 10.1** Overview: in negative control, a trans-acting repressor binds to the cis-acting operator to turn off transcription. In prokaryotes, multiple genes are controlled coordinately.
Figure 10.2 Overview: in positive control, trans-acting factors must bind to cis-acting sites in order for RNA polymerase to initiate transcription at the promoter. In a eukaryotic system, a structural gene is controlled individually.
Induction and repression can be under positive or negative control.

**NEGATIVE CONTROL**
- Repressor
- Inactive repressor
- Inducer
- Repressed

**POSITIVE CONTROL**
- RNA polymerase
- Active activator
- Inducer
- Repressed

**INDUCTION**
- Repressor
- Inactive repressor
- Active repressor
- Corepressor
- Induced

**REPRESSION**
- Inactive repressor
- Active activator
- Corepressor
- Repressed
Influence of Glucose on expression of lac Operon

Glucose controls import of lactose
and of other alternative carbon sources
Carbon Catabolite Regulation

Cyclic AMP acts as an inducer

CAP (CRP) protein is a positive acting regulator protein
**Figure 10.24** The CAP protein can bind at different sites relative to RNA polymerase.

- **gal**
- **lac**
- **ara**

**Figure 10.23** The consensus sequence for CAP contains the well-conserved pentamer TGTTA and (sometimes) an inversion of this sequence (TCANA).

**Figure 10.26** CAP bends DNA $>90^\circ$ around the center of symmetry.

- Center of dyad symmetry
- Bending angle $>90^\circ$
**Figure 10.30** A regulator protein may block translation by binding to a site on mRNA that overlaps the ribosome-binding site at the initiation codon.

**Figure 10.31** Proteins that bind to sequences within the initiation regions of mRNAs may function as translational repressors.

<table>
<thead>
<tr>
<th>Repressor</th>
<th>Target Gene</th>
<th>Site of Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>R17 coat protein</td>
<td>R17 replicase</td>
<td>hairpin that includes ribosome binding site</td>
</tr>
<tr>
<td>T4 RegA</td>
<td>early T4 mRNAs</td>
<td>various sequences including initiation codon</td>
</tr>
<tr>
<td>T4 DNA polymerase</td>
<td>T4 DNA polymerase</td>
<td>Shine-Dalgarno sequence</td>
</tr>
<tr>
<td>T4 p32</td>
<td>gene 32</td>
<td>single-stranded 5’ leader</td>
</tr>
</tbody>
</table>
**Figure 10.34** Translation of the r-protein operons is autogenously controlled and responds to the level of rRNA.

When ribosomal RNA is available, the r-proteins associate with it.

There are no free r-proteins, and translation of r-protein mRNA continues.

When no ribosomal RNA is available, the r-proteins accumulate.

One of the r-proteins binds to the mRNA and prevents translation.
Figure 10.39 The trp operon consists of five contiguous structural genes preceded by a control region that includes a promoter, operator, leader peptide coding region, and attenuator.
Figure 10.41 The \textit{trp} leader region can exist in alternative base-paired conformations. The center shows the four regions that can base pair. Region 1 is complementary to region 2, which is complementary to region 3, which is complementary to region 4. On the left is the conformation produced when region 1 pairs with region 2, and region 3 pairs with region 4. On the right is the conformation when region 2 pairs with region 3, leaving regions 1 and 4 unpaired.

Regions 3 & 4 pair to form the terminator hairpin

ALTERNATIVE STRUCTURES ARE POSSIBLE
Region 2 is complementary to 1 & 3
Region 3 is complementary to 2 and 4

Regions 2 & 3 pair; terminator region is single-stranded
**Figure 10.42** The alternatives for RNA polymerase at the attenuator depend on the location of the ribosome, which determines whether regions 3 and 4 can pair to form the terminator hairpin.

**TRYPTOPHAN ABSENT**

**TRYPTOPHAN PRESENT**

Ribosome movement disrupts 2:3 pairing

3:4 pairing forms terminator hairpin
Antisense RNA can be generated by reversing the orientation of a gene with respect to its promoter, and can anneal with the wild-type transcript to form duplex RNA.
Figure 10.44  Increase in osmolarity activates EnvZ, which activates OmpR, which induces transcription of micF and ompC (not shown). micF RNA is complementary to the 5' region of ompF mRNA and prevents its translation.
Regulator RNA binds RNA target

**Regulator RNA**

**Target RNA**

**Duplex region**

**Regulator excludes protein binding**

Protein binds single-stranded region in target

Protein cannot bind to target
Exonuclease cleaves duplex target

Target has alternative conformation

Secondary structure forms in absence of regulator
A loop at the 3' end of oxyS RNA pairs with the initiation site of flhA mRNA

flhA mRNA
5' UUUGC GGUGCUUUUCCUGGAAGAACAAA AUG ................. 3'

oxyS RNA
3' AGGACCU

©virtualtext  www.ergito.com
Genetik und Gentechnik IH. Schwab

*lin14 codes for a single protein*

*lin4 codes for an RNA that turns off *lin14*

---

**Figure 10.43** Antisense RNA can affect function or stability of an RNA target.

- Control of initiation of translation
  - Ribosome binds
  - Antisense RNA
  - Ribosome cannot bind

- Control of RNA stability
  - Antisense RNA
  - Cleavage & degradation

- Control of termination of transcription
  - Antisense RNA
  - Cleavage & termination
dsRNA is cleaved~22 bases from the 3' ends to generate siRNA

21-23 base siRNA with protruding 3' ends
RNAi works by generating siRNA

Nuclease cleaves dsRNA to siRNA

dsRNA

siRNA

siRNA base pairs with mRNA

mRNA

Helicase

Nuclease cleaves mRNA

©virtualtext www.ergito.com
dsRNA has both general and specific effects

- dsRNA >26 nucleotides
  - PKR
  - eIF2α

  ↓

- 2',5'AS
- RNAase L

  ↓

- Protein synthesis cannot initiate
- Degradation of all mRNA
- mRNA degraded
- siRNA targets complementary mRNA
25.11.14