TRANSLATION 1: SYNTHESES, TRANSLATION CYCLE and FACTORS

Outline of the lecture:

1. Why is the translation machinery so complex? What are the challenges to produce functional proteins?
2. Tolerable error frequency and the discrimination problem
3. tRNAs: the adapters used during translation
4. Aminoacyl tRNA synthetases. Genetic dictionary used for translation (change of language)
5. Classes of synthetases, discrimination, proofreading and editing
6. Problems in synthetase fidelity…
7. Ribosome structure, elongation factors, structures and functions
8. Translation elongation cycle or “how all fit together”
9. Brownian movements and how the ribosome crawl along the mRNA
The translation machinery

The chemistry of peptide bond formation is fundamentally **SIMPLE**:

![Peptide bond formation diagram](image)

However the translation machinery is very **COMPLEX**?

**Challenges**

- Change of language but maintain the message/sequence
- Fidelity: selective amplification of Cognate vs. Non-Cognate
  
  aa:tRNA/mRNA:tRNA

- Reading frame maintenance

(Ramakrishnan, 2002)
Accuracy of Translation:
Throughout biology, genomes are maintained and expressed with remarkable fidelity. The overall accuracy of gene expression is not as high as is theoretically possible. The accuracy of each process involved represents a compromise that optimizes the evolutionary fitness of the organism. **Speed vs accuracy.**

Thus, replication of the genome, which transmits genetic information from one generation to the next, is extremely accurate, with error rates as low as $10^{-8}$ in bacteria or below $10^{-10}$ in eukaryotes. DNA replication has sophisticated error correction mechanisms, including editing and repair.

The error rate of transcription in vivo in *Escherichia coli* has been estimated to be $1.4 \times 10^{-4}$ per nucleotide and thus around $4 \times 10^{-4}$ per codon. It is considerably lower than that of replication.

### Major Constraint: tolerable error frequency

<table>
<thead>
<tr>
<th>Frequency of inserting an incorrect amino acid</th>
<th>Number of amino acid residues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>0.364</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>0.915</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>0.990</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>0.999</td>
</tr>
</tbody>
</table>

**Tolerable error frequency is $10^{-4}$!**

(Kurland, 1992)
The discrimination problem

- \( \Delta G \) of C vs. NC is only 10-100 x
  
  e.g.  
  
  Ile/AUU  
  Val/GUU  
  Leu/CUU  
  Phe/UUU

- Compounded by "Pool Bias": 20:1 / NC:C
  
  - a.a. @ synthetase site
  - amino acyl tRNA @ A site

tRNAs: the "adapters" used in decoding.

1) Sets of absolutely conserved nucleotides involved in stabilization of the structure.

2) tRNAs has to be recognized equally well by translation elongation factors/ribosomes,…but differentially by RS

3) All tRNA are subject to a variety of chemical modification 1 in 10 nucleotides are modified: dihydrouridine, pseudouridine, inosine, methylation, etc.). Some affect the overall structure of the tRNA and some target the functional elements of the tRNA (anticodon sequence and/or sequences important for aminoacylation), therefore having direct effects on decoding and protein synthesis.

4) A majority of modifications play more subtle structural roles that can affect tRNA stability and folding, including both rigidifying the overall structure as well as making tRNA more flexible
Several different codons can specify a single amino acid. Some amino acids have more than one tRNA and these tRNAs are constructed so that they require accurate base-pairing only at the first two positions of the codon and can tolerate a mismatch (wobble).

Codon usage could control translation elongation kinetics to facilitate co-translation proteins folding and processing.
Aminoacyl-tRNA synthetases

Synthetases act as the genetic dictionary and thus have a dual recognition problem: matching amino acids with their cognate tRNAs. There is a different synthetase for each amino acid. Synthetases attach amino acids to the 3’ end of the tRNA. The reaction is coupled to ATP hydrolysis and produces a high-energy bond between the tRNA and the aa. The energy of that bond is used at a later stage to link the amino acid to the growing polypeptide chain (Fersht and Dingwall, 1979; Ling et al., 2009).
Editing pathways against the noncognate amino acid (aa). (Pathway 1) Posttransfer editing. The misactivated aa is first attached to the tRNA and then translocated from the active site (AS) to the distinct editing site (ES) for hydrolysis. (Pathway 2) Translocation of aminoacyl-adenylate (aa-AMP). The misactivated aa-AMP directly translocates to the ES to be hydrolyzed. (Pathway 3) Selective release. After activation, the noncognate but not the cognate aa-AMP is expelled into solution and subjected to spontaneous hydrolysis. (Pathway 4) Active site hydrolysis. The noncognate aa-AMP is hydrolyzed at the AS before release. Pathways 2–4 collectively comprise pretransfer editing.
The tRNA synthetases have an editing activity that hydrolyzes incorrect products in a tRNA-dependent manner. Amino acids that are smaller than or isosteric to the cognate are edited by a hydrolytic editing mechanism (Ling et al., 2009). Distinguishing tRNAs requires only few "identity elements" (Ling et al., 2009).

### Table 1 Editing substrates and domains of aa-tRNA synthetases

<table>
<thead>
<tr>
<th>Class</th>
<th>aaRS</th>
<th>Editing substrates</th>
<th>Cis-editing domain</th>
<th>Trans-editing factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>IleRS</td>
<td>Val, Hcy&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CP1</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>ValRS</td>
<td>Thr, α-Abu&lt;sup&gt;a&lt;/sup&gt;, Hcy&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CP1</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>LeuRS</td>
<td>Hcy&lt;sup&gt;a&lt;/sup&gt;, γ-hLeu&lt;sup&gt;a&lt;/sup&gt;, Val, nLeu&lt;sup&gt;a&lt;/sup&gt;, Met, Ile</td>
<td>CP1/active site?</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>MetRS</td>
<td>Hcy&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Active site?</td>
<td>–</td>
</tr>
<tr>
<td>Class II</td>
<td>ThrRS</td>
<td>Ser</td>
<td>N2</td>
<td>ThrRS-ed</td>
</tr>
<tr>
<td></td>
<td>AlaRS</td>
<td>Gly, Ser</td>
<td>AlaX-like</td>
<td>AlaXp</td>
</tr>
<tr>
<td></td>
<td>ProRS</td>
<td>Ala, 4-hPro&lt;sup&gt;a&lt;/sup&gt;, Cys&lt;sup&gt;b&lt;/sup&gt;</td>
<td>INS&lt;sup&gt;c&lt;/sup&gt;</td>
<td>YbaK, ProX</td>
</tr>
<tr>
<td></td>
<td>PheRS</td>
<td>Tyr, Ile</td>
<td>B3B4</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>LysRS II</td>
<td>Hcy&lt;sup&gt;a&lt;/sup&gt;, Hse&lt;sup&gt;a&lt;/sup&gt;, Orn&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Active site?</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>SerRS</td>
<td>SerHX&lt;sup&gt;a&lt;/sup&gt;, Thr, Cys</td>
<td>Active site?</td>
<td>–</td>
</tr>
</tbody>
</table>
Case study: amino acid discrimination (ThrRS)

Achieved by monitoring differences in size and chemistry

(Dock-Bregeon et al., 2000; Fukai et al., 2000)
Errors in aminaocylation and neurodegeneration diseases

Misfolded proteins are associated with several pathological conditions including neurodegeneration. Is there a general mechanism that lead to misfolded proteins in neurons? Low levels of mischarged transfer RNAs (tRNAs) can lead to an intracellular accumulation of misfolded proteins mutant cells. These accumulations are accompanied by upregulation of cytoplasmic protein chaperones and by induction of the unfolded protein response. (c, e, g, and i are in fibroblast).

Mouse sticky mutation (sti), which causes cerebellar Purkinje cell loss and ataxia, is a missense mutation in the editing domain of the alanyl-tRNA synthetase gene that compromises the proofreading activity of this enzyme during aminoacylation of tRNAs. (Lee et al., 2006)
All tRNA must be recognized equally by EF-Tu and by the ribosomal A site. So, how can they be distinguished by RS??

THE SOLUTION: tRNA "Identity Sets"

1. Anti-codon is sole (major) determinant  
   e.g. met
2. Acceptor stem is sole (major) determinant  
   e.g. ala  (G-U)
3. Complex (anti-codon + acceptor + variable pocket)  
   e.g. phe

(Francklyn and Schimmel, 1990; Normanly et al., 1986; Saks et al., 1994)
Each large pre-rRNA contains 18S, 28S and 5.8S sequences which are separated by external and internal transcribed spacer sequences. During processing reactions, the 18S, 28S and 5.8S rRNA are released as individual molecules. Processing reactions involve exo- and endo-nucleolytic cleavages guided by snoRNA (small nucleolar RNAs) in complex with proteins. The genes for 5S rRNA are located outside the nucleolus and are transcribed into pre-5S rRNA by RNA polymerase III.

- Ribosome $\sim 3 \times 10^6$, 250 Å ($50S + 30S = 70S$)

- Factors: IF 1, 2, 3
- EF-Tu, EF-Ts, EF-G
- RF 1, 2, 3, RRF

- GTP hydrolysis
The structure of the ribosome was solved at 2.5 to 4.5 Å resolution. rRNA determines the overall structure, its ability to position tRNAs on the mRNA and its catalytic activity in forming peptide bonds. The rRNA is folded into a highly compact core, which determines the overall shape of the ribosome. The proteins are generally located on the surface and fill gap of the folded RNA and appear to stabilize the RNA and form ant interface with the environment. (Cate et al., 1999; Clemons et al., 1999; Korostellev et al., 2006; Petry et al., 2005; Yusupov et al., 2001)

Proteins that appear on the surface of the large ribosomal subunit. They are uniformly distributed except for the active site cleft and the flat surface that interacts with the 30S subunit.

Ribosme assembling must proceed in highly ordered manner. Chaperones may well be required to prevent the aggregation of the extended regions of these proteins and RNAs.
Domain I, lies in the back of the particle, behind and below the L1 region. Domain II is the largest of the six 23S rRNA domains, accounting for most of the back of the particle. It has three protrusions that reach toward the subunit interface side of the particle. Domain III is a compact globular domain that occupies the bottom left region of the subunit. Domain IV accounts for most of the interface surface of the 50S subunit that contacts the 30S subunit. Domain V, which is sandwiched between domains IV and II in the middle of the subunit, is intimately involved in the peptidyl transferase activity. Structurally, this domain can be divided into three: binding site for protein L1, the bulk of the central protuberance which interacts with 5S rRNA and domain II, and a domain that helps stabilize the elongation factor-binding region of the ribosome. Domain VI, which forms a large part of the surface of the subunit immediately below the L7/L12 stalk, an contains the sarcin-ricin loop (SRL) which is essential for factor binding,
Using biochemistry, footprinting, cryoEM and FRET experiments, it has been shown that a tRNA passes through six physically distinct states during its transit across the ribosome. In dramatic contrast to the classical, static two-site (A,P) model, the identification of hybrid sites provides an important link between decoding and translocation. (Frank and Gonzalez, 2010)
Elongation factors (EF-Tu and EF-G) enter and leave the ribosome in each cycle. The hydrolyze GTP to GDP and undergo conformational changes in the process. These changes in the elongation factors accelerate the transition between the different conformational states of the ribosome. (Kawashima et al., 1996; Rodnina et al., 1995; 1997)
Arrangement of tRNAs and EF-G in PRE, early POST, and late POST states. Using toolbox of antibiotics, non-hydrolyzable GTP analogs and EF-G mutants allowed the dissection of the translocation pathway into a number of intermediates. D: Overlay of tRNAs and EF-G from the structures displayed in panels A–C, aligned on domain 1 of EF-G. The positions of tRNAs are indicated. Note the different positions of domain 4 of EF-G in PRE (red), early POST (blue) and in late POST (green) (Holtkamp et al., 2014).
Major landmarks and mobile elements of the 30S subunit are the head (h), shoulder (s), platform (p), and spur (sp). The location of the decoding center (DC) active site is also denoted. Major landmarks and mobile elements of the 50S subunit are the L1 stalk (L1) and the L7/L12 stalk (L7/L12). The locations of the GTPase-associated center (GAC) and peptidyltransferase center (PTC) active sites are also denoted. The locations of all donor (D) and acceptor (A) fluorophore pairs that have thus far been used in single-molecule fluorescence resonance energy transfer investigations of translation elongation are labeled in green (D) and red (A).

Similar than for small enzymes, the ribosome is a highly dynamic molecular machine, specifically a processive Brownian motor. Evidence from smFRET and cryo-EM for spontaneous transitions between the macrostates has led to a revision of the role of EF-G in translocation, from instrumental to ancillary, accelerating a process structurally ingrained in the ribosome (Frank and Gonzalez, 2010).
References


