University of Zurich[UZH]

Swiss Institute of Allergy & Asthma Research

Proteomics

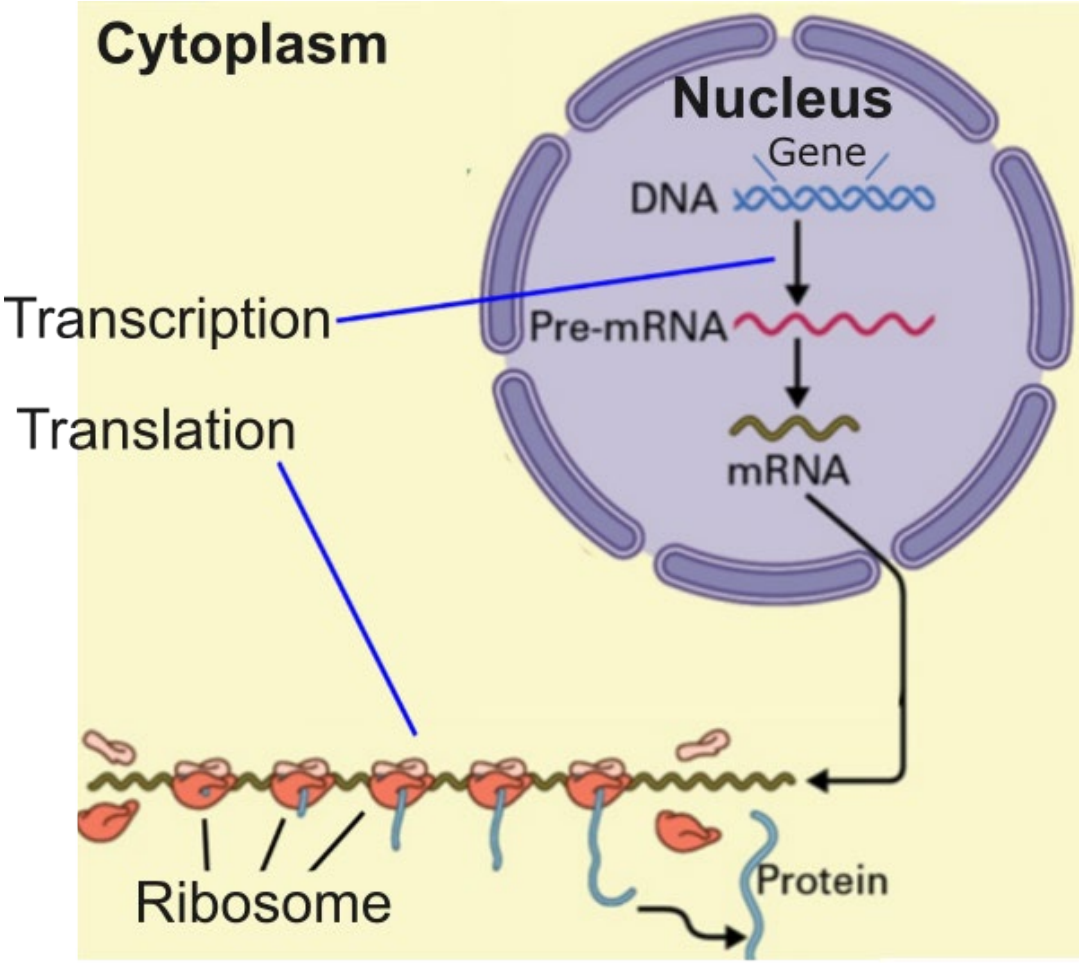
A Brief Glossary of Protein Terms

29.10.2024

PD Dr. Katja Bärenfaller
katja.baerenfaller@siaf.uzh.ch

SIB Swiss Institute of Bioinformatics

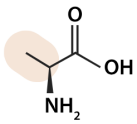
Gene expression – the biosynthesis of proteins encoded in genes



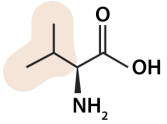
During gene expression, specific regions of genomic DNA are transcribed to produce pre-mRNA. This pre-mRNA is then spliced, capped, and polyadenylated to form mature mRNA. The mature mRNA is transported to the cytoplasm, where ribosomes translate it into proteins.

The building blocks of proteins are the 20 amino acids

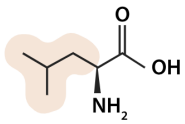
Non-polar side chains, uncharged, hydrophobic



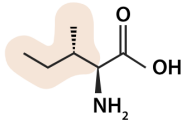
Alanine (Ala, A)
MW: 89,09
pI: 6,01
C3H7N1O2



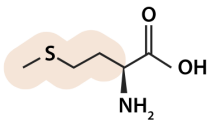
Valine (Val, V)
MW: 117,15
pI: 6,00
C5H11N1O2



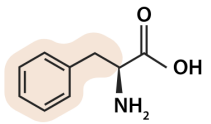
Leucin (Leu, L)
MW: 131,17
pI: 6,01
C6H13N1O2



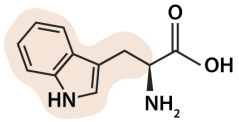
Isoleucine (Ile, I)
MW: 131,17
pI: 6,05
C6H13N1O2



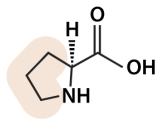
Methionine (Met, M)
MW: 149,21
pI: 5,74
C5H11N1O2S1



Phenylalanine (Phe, F)
MW: 165,19
pI: 5,49
C9H11N1O2



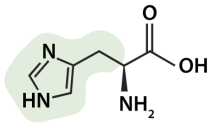
Tryptophan (Trp, W)
MW: 204,23
pI: 5,89
C11H12N2O2



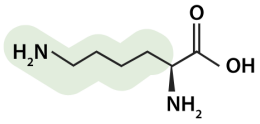
Proline (Pro, P)
MW: 115,13
pI: 6,30
C5H9N1O2

Electrically charged side chains

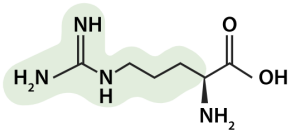
Basic



Histidine (His, H)
MW: 155,16
pI: 7,60
C6H9N3O2

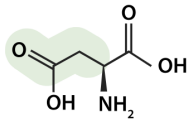


Lysine (Lys, K)
MW: 146,19
pI: 9,60
C6H14N2O2

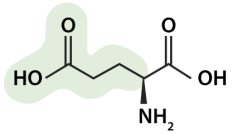


Arginine (Arg, R)
MW: 174,20
pI: 10,76
C6H14N4O2

Acidic

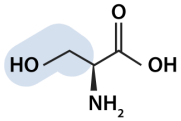


Aspartaic Acid (Asp, D)
MW: 133,1
pI: 2,85
C4H7N1O4

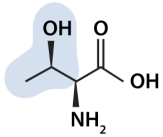


Glutamic Acid (Glu, E)
MW: 147,13
pI: 3,15
C5H9N1O4

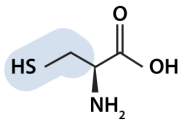
Polar side chains, uncharged



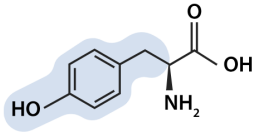
Serine (Ser, S)
MW: 105,09
pI: 5,68
C3H7N1O3



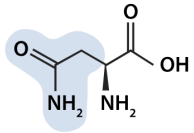
Threonine (Thr, T)
MW: 119,12
pI: 5,60
C4H9N1O3



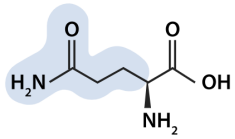
Cysteine (Cys, C)
MW: 121,16
pI: 5,05
C3H7N1O2S1



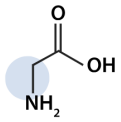
Tyrosine (Tyr, Y)
MW: 181,19
pI: 5,64
C9H11N1O3



Asparagine (Asn, N)
MW: 132,12
pI: 5,41
C4H8N2O3



Glutamine (Gln, Q)
MW: 146,15
pI: 5,65
C5H10N2O3

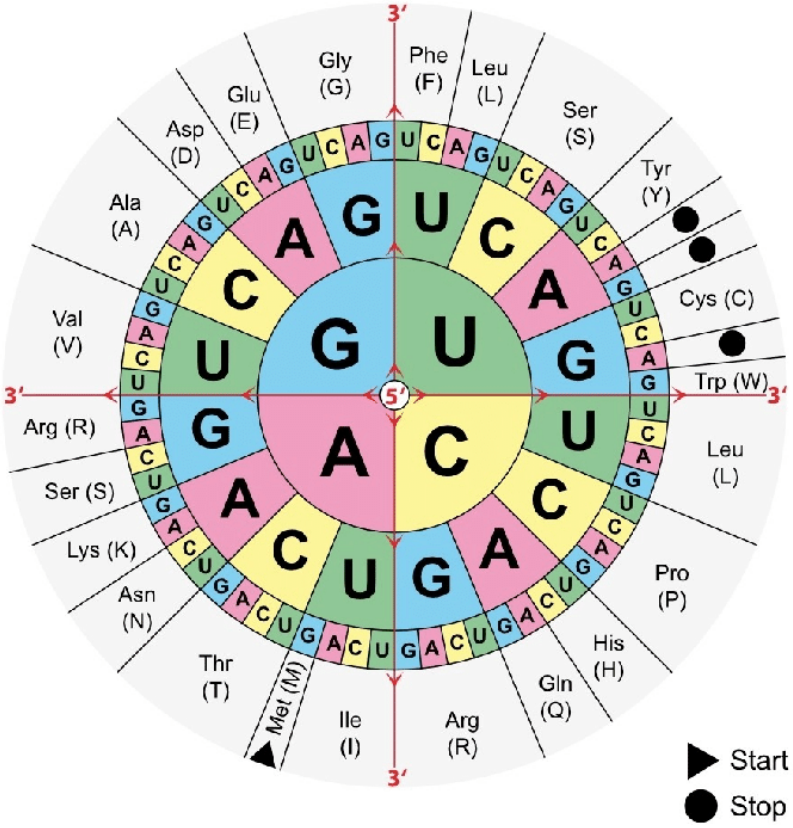


Glycine (Gly, G)
MW: 75,07
pI: 6,06
C2H5N1O2

Translation of the mRNA sequence into amino acid sequence

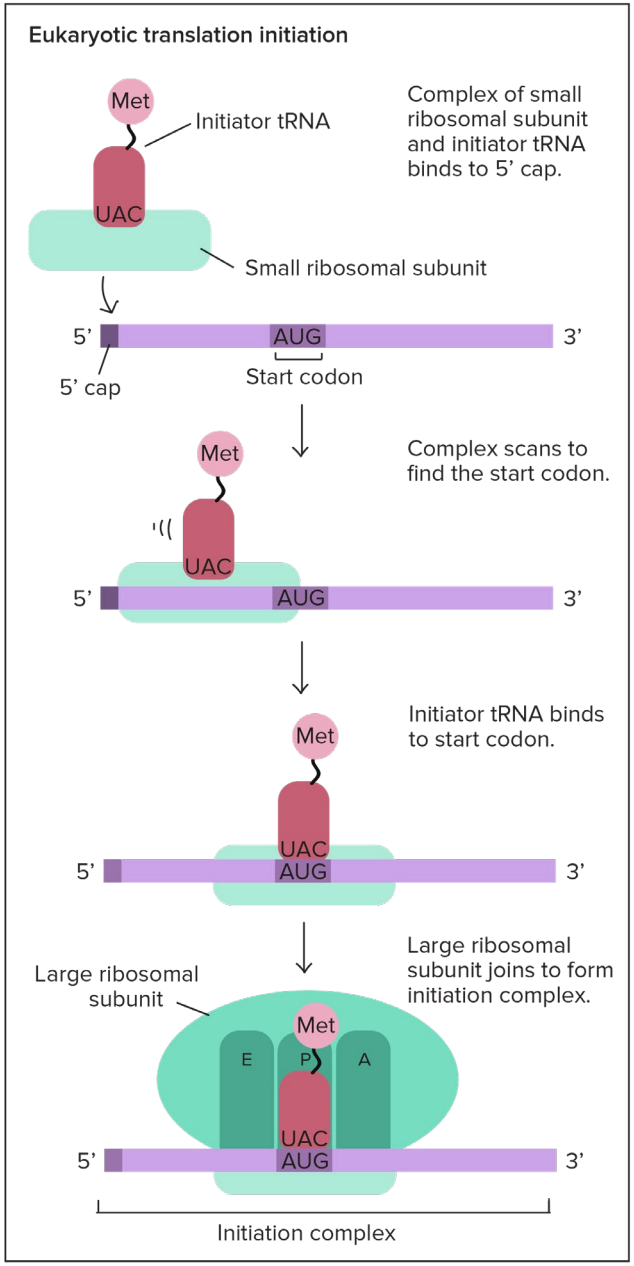
RNA ⟶ PROTEIN

To express the genetic information, the mRNA is translated to produce proteins. The genetic code serves as a translator between the nucleotide sequence and the amino acid sequence.



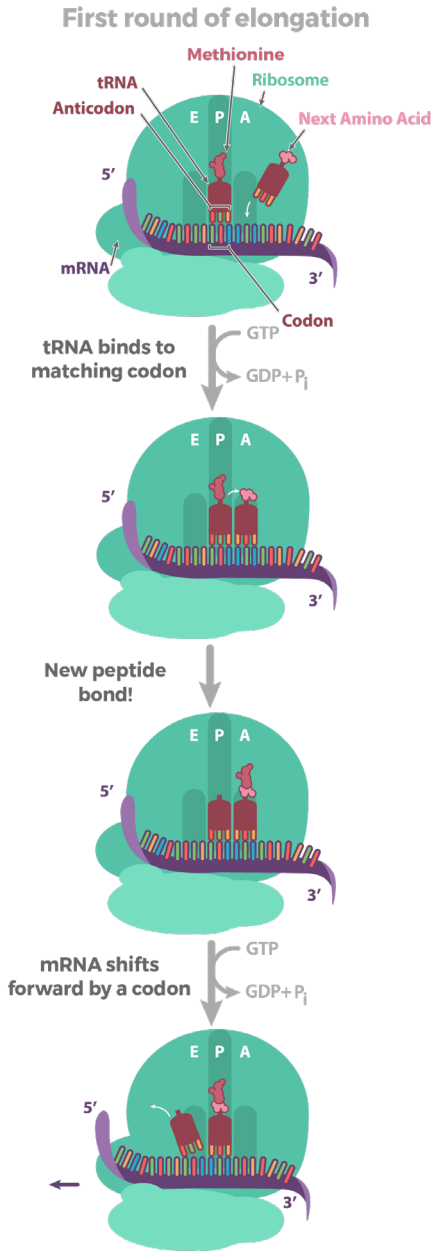
Amino acids	Three-letter code	Single letter code	Amino acids	Three-letter code	Single letter code
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

Translation – Initiation



1) A complex consisting of the small 40S ribosomal subunit, the initiator methionine-tRNA (Met-tRNAi), and various translation initiation factors starts at the 5'-cap of the mRNA and scans the mRNA in search of an AUG start codon.
2) Once the scanning complex finds the start codon, the anticodon of the Met-tRNAi binds to the start codon.
3) The initiation factors then dissociate, and the large 60S ribosomal subunit arrives.
4) The docking of the subunits and the formation of a complete, functional ribosome on the mRNA is completed, and elongation can begin

Translation – Elongation and Termination



Elongation is the extension of the polypeptide chain in the ribosome.

1) The first methionine-carrying tRNA starts in the middle slot of the ribosome, the P-site. Next to it, in another slot called the A-site, a new codon is exposed. The A-site is the landing site for the next tRNA, whose anticodon must be a perfect (complementary) match to the exposed codon.
2) Once the appropriate tRNA has landed in the A-site, the two amino acids are joined with a peptide bond. After the peptide bond is formed, the mRNA is pulled through the ribosome by exactly one codon.
3) This shift allows the first, now empty, tRNA to drift out through the E-site, and a new codon is exposed in the A-site, and so on.

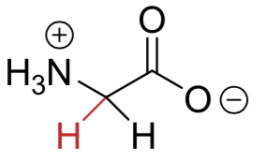
Termination ends the translation.

Termination occurs when a stop codon in the mRNA (UAA, UAG, or UGA) reaches the A-site. At this point, instead of a new amino acid, a water molecule is added to the last amino acid of the chain, and the newly formed protein is released.

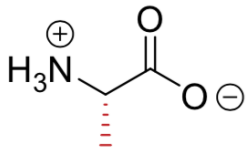
Protein N- und C-terminus

The protein N-terminus consists of an amine group (NH2)

The protein C-terminus consists of a carboxyl group (COOH)



Glycine
Gly
G



Alanine
Ala
A