Biology I

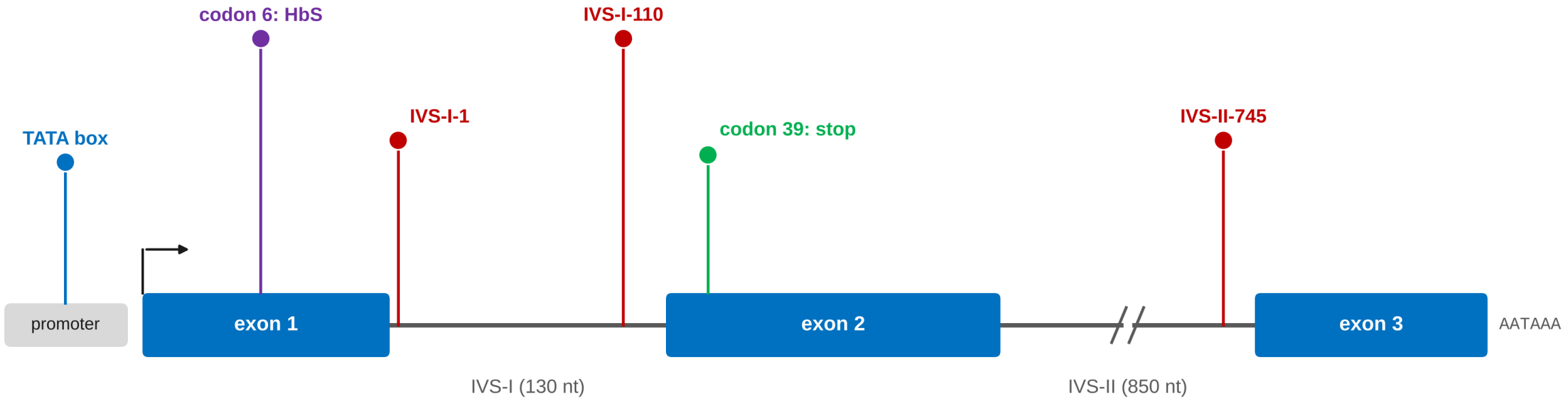
Basic Principles of Cell Biology

From DNA to proteins

Transcription, RNA processing and translation

Yiannis Drosos
Assistant Professor
Laboratory of Biology
Department of Medicine
University of Patras
idrosos@upatras.gr

One gene, the whole flow of information: HBB



transcription · splicing · translation · protein

HBB · chromosome 11p15.4 · not to scale

~8%
β-thalassemia carriers in Greece (national average)

42.5%
of Greek alleles are IVS-I-110, a splicing mutation

Every step from DNA to protein can break — and every step is a potential therapeutic target.

Papachatzopoulou et al., Hemoglobin 2010 · Kattamis et al., Br J Haematol 1990 · Thein, Cold Spring Harb Perspect Med 2013

Three questions for the next three hours

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

1

How is a gene transcribed into RNA — and who decides where to start?

→ Part A (1 hour): RNA, RNA polymerases, transcription

2

How does precursor RNA become mature mRNA — and what happens when splicing fails?

→ Part B1: cap, splicing, poly(A), diseases and therapies

3

How does the ribosome read the code — and what happens to the protein?

→ Part B2: genetic code, translation, folding, degradation

From previous lectures: DNA = double helix · gene · chromatin. Today: how the information is read. Alberts, Essential Cell Biology, Ch. 7.

The central dogma: the flow of information

NATURE VOL. 227 AUGUST 8 1970

Central Dogma of Molecular Biology

by
FRANCIS CRICK
MRC Laboratory of Molecular Biology,
Hills Road,
Cambridge CB2 2QH

The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein or nucleic acid.

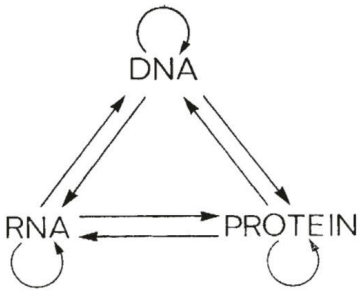


Fig. 1. The arrows show all the possible simple transfers between the three families of polymers. They represent the directional flow of detailed sequence information.

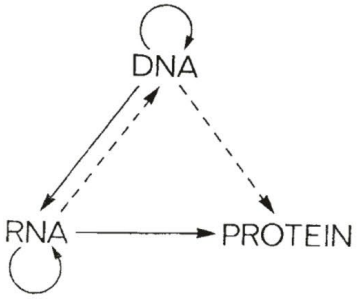


Fig. 2. The arrows show the situation as it seemed in 1958. Solid arrows represent probable transfers, dotted arrows possible transfers. The absent arrows (compare Fig. 1) represent the impossible transfers postulated by the central dogma. They are the three possible arrows starting from protein.

1958 · Crick

Information flows from nucleic acids to proteins — never back.

1970 · clarification

Reverse transcription (retroviruses; later telomerase, 1985) and RNA replication do not violate the dogma.

Today

DNA → RNA → protein

Transcription: DNA → RNA · Translation: RNA → protein · Both rely on base complementarity.

Crick F., Nature 1970;227:561–563

Key concepts

Gene

A DNA segment transcribed into a functional RNA — usually an mRNA that encodes a protein.

Transcription

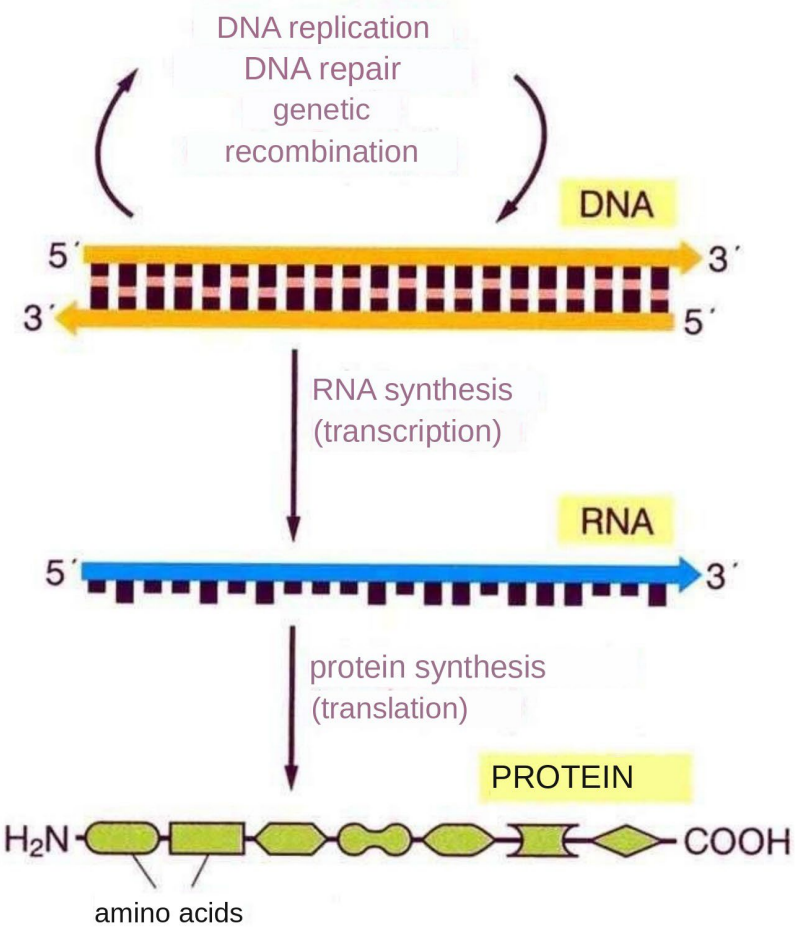
RNA synthesis using one DNA strand as template; enzyme: RNA polymerase (nucleus).

Translation

Protein synthesis directed by mRNA; machine: the ribosome (cytoplasm).

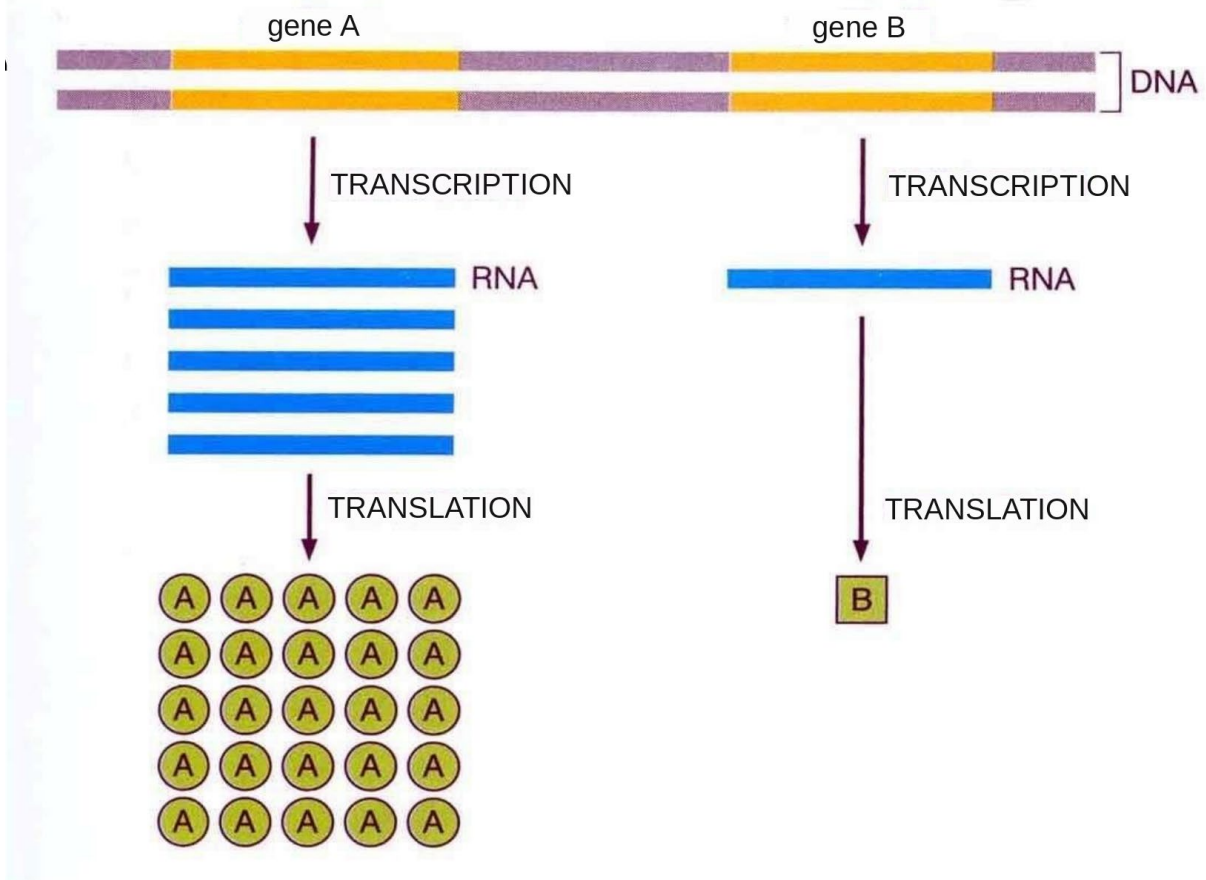
Gene expression

The whole path from DNA to final product; each gene is expressed at its own rate.



Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

Two steps, two chances to amplify



One gene → many mRNAs

Gene A is transcribed often, gene B rarely.

One mRNA → many proteins

Each mRNA is translated many times before it is degraded.

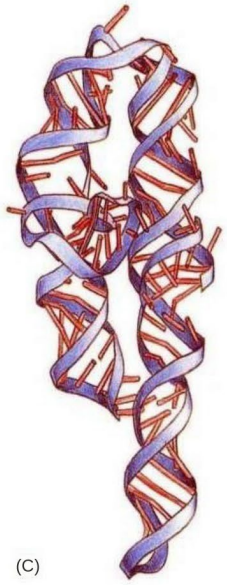
Not always a protein

For many genes the final product is RNA (rRNA, tRNA, snRNA, miRNA…).

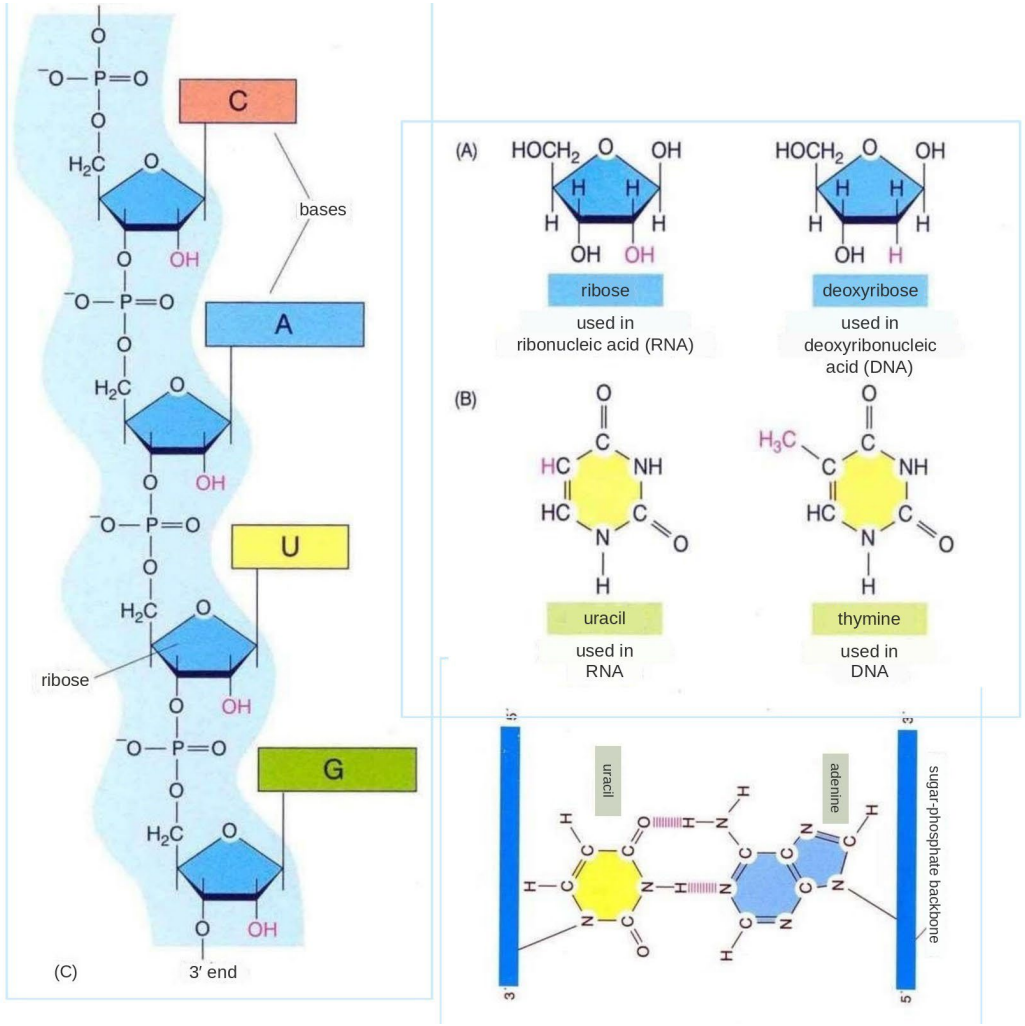
The amount of a protein is controlled at every step — topic of the next lecture.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

RNA and DNA: three chemical differences



snRNA



1. Ribose (2′-OH)

The extra hydroxyl makes RNA chemically more reactive and less stable.

2. Uracil instead of thymine

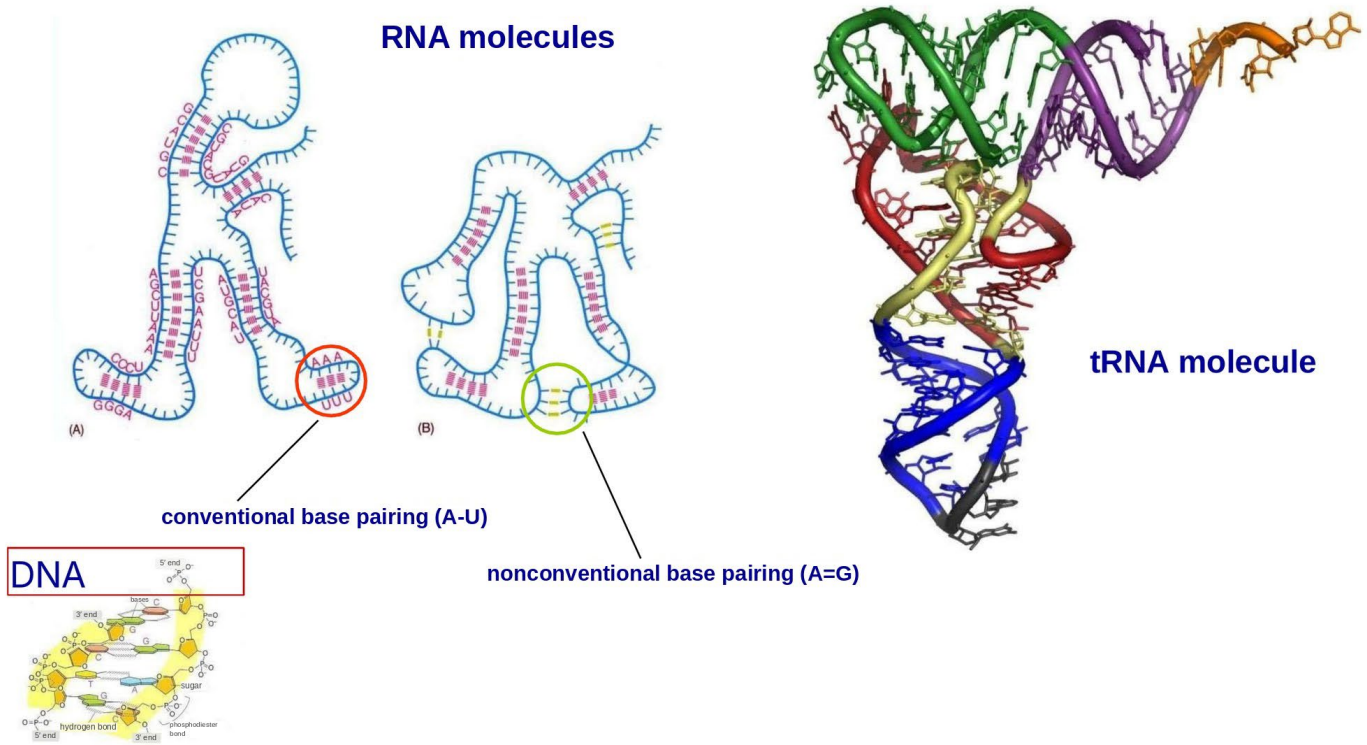
U pairs with A, like T.

3. Single-stranded

Made as a single chain, much shorter than DNA, which folds on itself.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

Single-stranded RNA folds into 3D structures



Intramolecular pairing

Conventional (A–U, G–C) and non-conventional pairs (e.g., G–U).

Structure = function

The shape of tRNA, rRNA and snRNAs determines what they do.

Ribozymes

Some RNAs catalyze reactions, like protein enzymes.

Alberts et al., Essential Cell Biology, 5th ed., Fig. 7-5

Cells make many types of RNA

RNA	Function	Polymerase
mRNA	Encodes proteins	II
rRNA	Structure and catalysis in the ribosome	I (28S, 18S, 5.8S) · III (5S)
tRNA	Adaptor between amino acid and codon	III
snRNA	Splicing of pre-mRNA	II (U6: III)
snoRNA	Chemical modification of rRNA	II (from introns)
miRNA, lncRNA	Regulation of gene expression	II

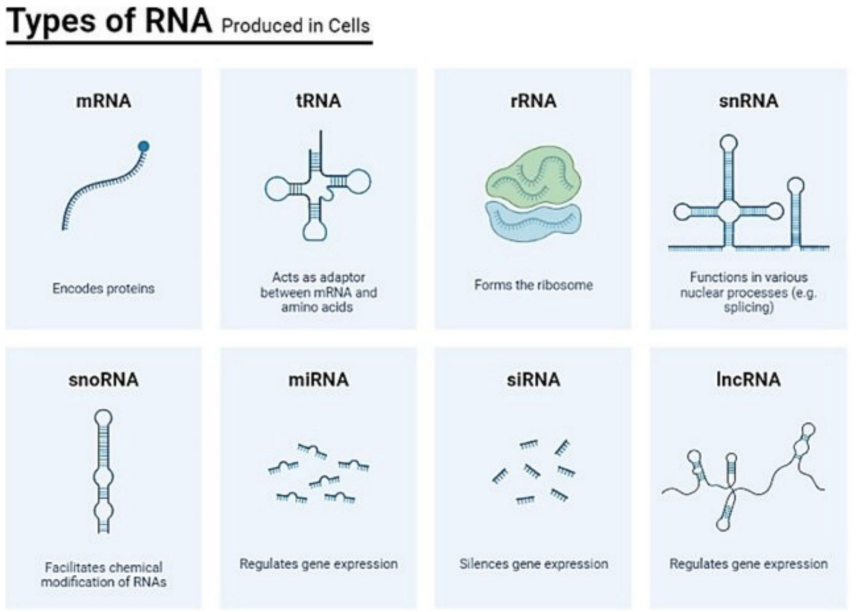
Most of the genome does not encode proteins; many genes give functional RNAs.

~80%

of RNA is rRNA

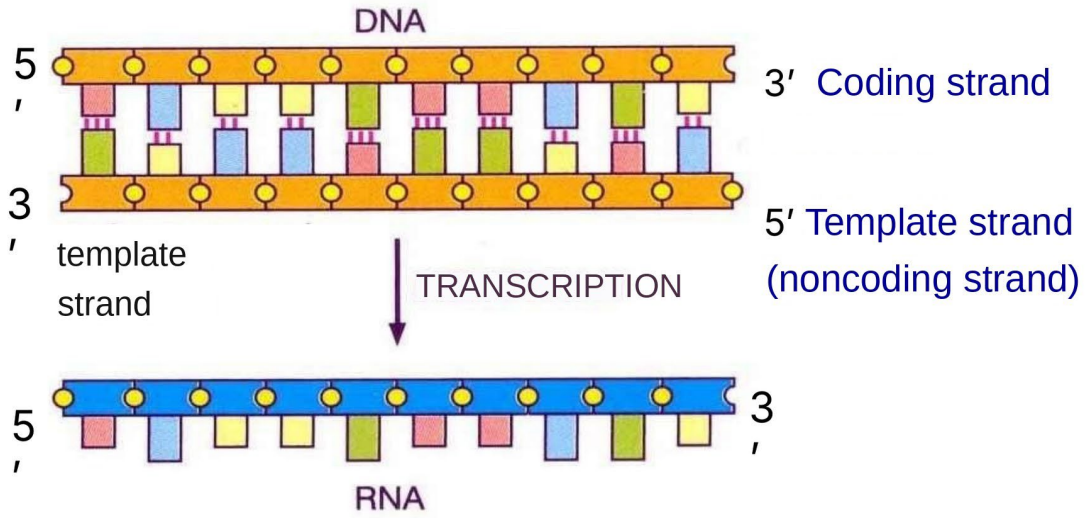
<5%

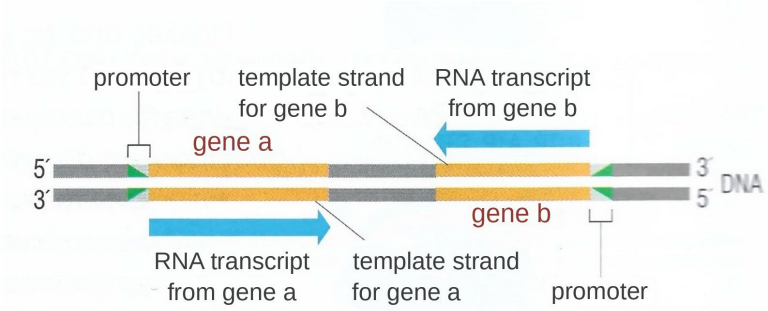
is mRNA



Alberts et al., MBoC, 7th ed., Ch. 6 · Image: Suleiman et al., Mol Biol Rep 2024

Transcription: template and coding strands





Template

The strand that is read (3′→5′); the RNA is complementary to it.

Coding strand

Same sequence as the RNA (T instead of U). This is what databases show.

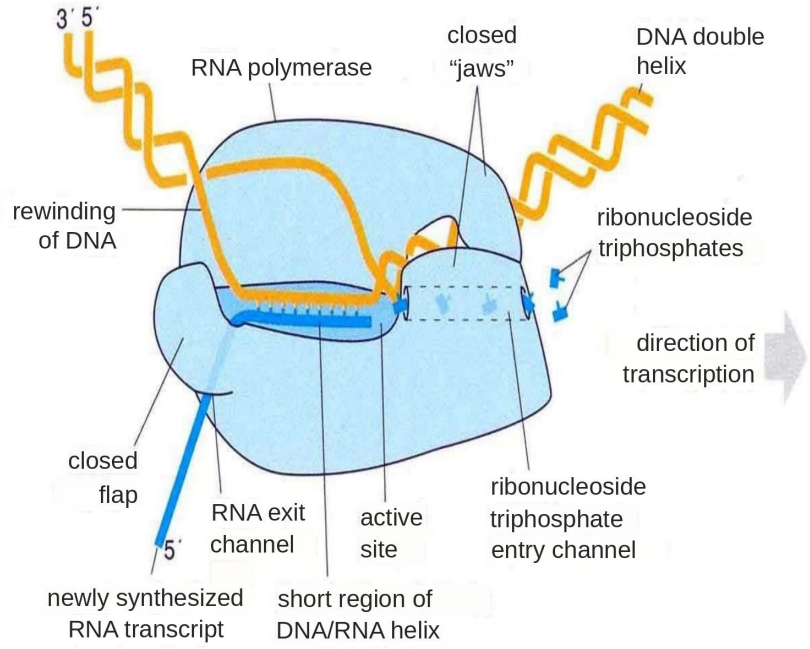
Each gene, its own strand

The promoter sets direction and template; neighboring genes may use opposite strands.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

RNA vs DNA polymerase: same chemistry, different rules

	RNA polymerase	DNA polymerase
Substrate	ribonucleotides (NTPs)	deoxyribonucleotides (dNTPs)
Primer	not needed	required
Direction	5′ → 3′	5′ → 3′
Errors	~1 per 10^4 nucleotides	~1 per 10^7 (with proofreading)
Product	released; many copies	stays paired; one copy



Why tolerate 1,000× more errors in RNA? It is a temporary copy — the error is not inherited.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7, Fig. 7-9

IN-CLASS EXERCISE #1 · 3 MINUTES

Write the RNA

RNA polymerase moves from left to right (arrow). The segment shown starts at the start codon.

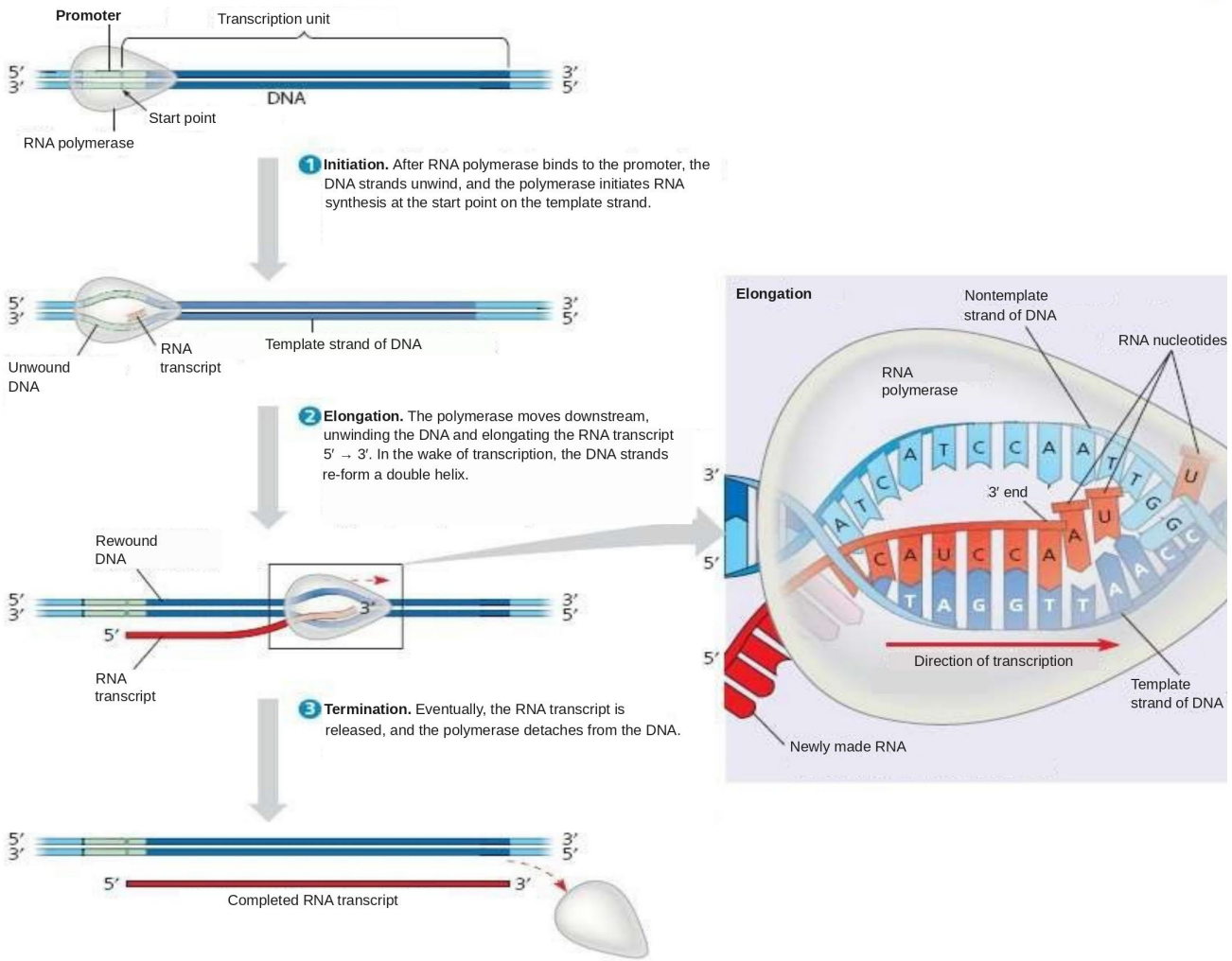
```
5′- A T G G T G C A T C T G A C T C C T G A G G A G -3′
3′- T A C C A C G T A G A C T G A G G A C T C C T C -5′
```

a Which strand is the template? Which is the coding strand?

b Write the RNA produced, with its 5′ and 3′ ends.

c Bonus: which human gene starts like this? (we will translate it in B2)

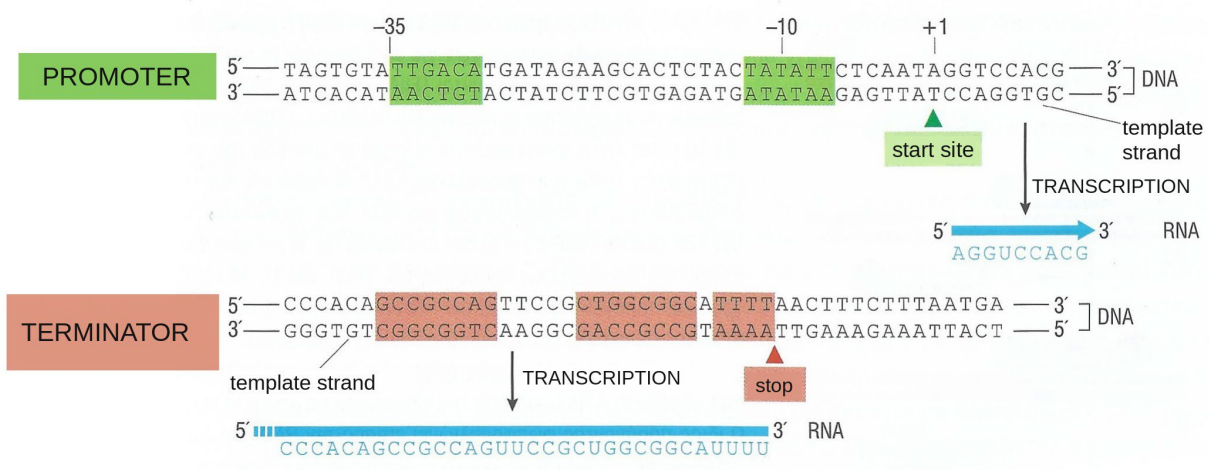
The stages of transcription

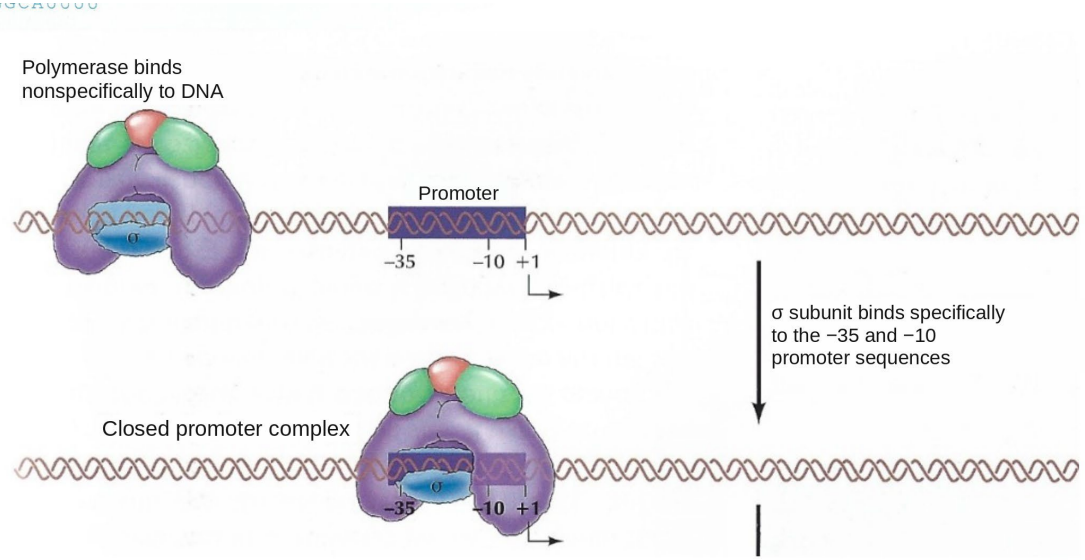


1. **Initiation:** binding to the promoter, local opening of the helix
2. **Elongation:** 5′→3′ synthesis; the helix re-forms behind the polymerase
3. **Termination:** release of RNA and polymerase

Alberts et al., Essential Cell Biology, Ch. 7 · Image: Campbell Biology, Fig. 17.7 (labels translated)

Bacteria: promoter, sigma factor and termination





Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

Promoter (−35, −10)

Conserved TTGACA and TATAAT boxes, read by the σ factor.

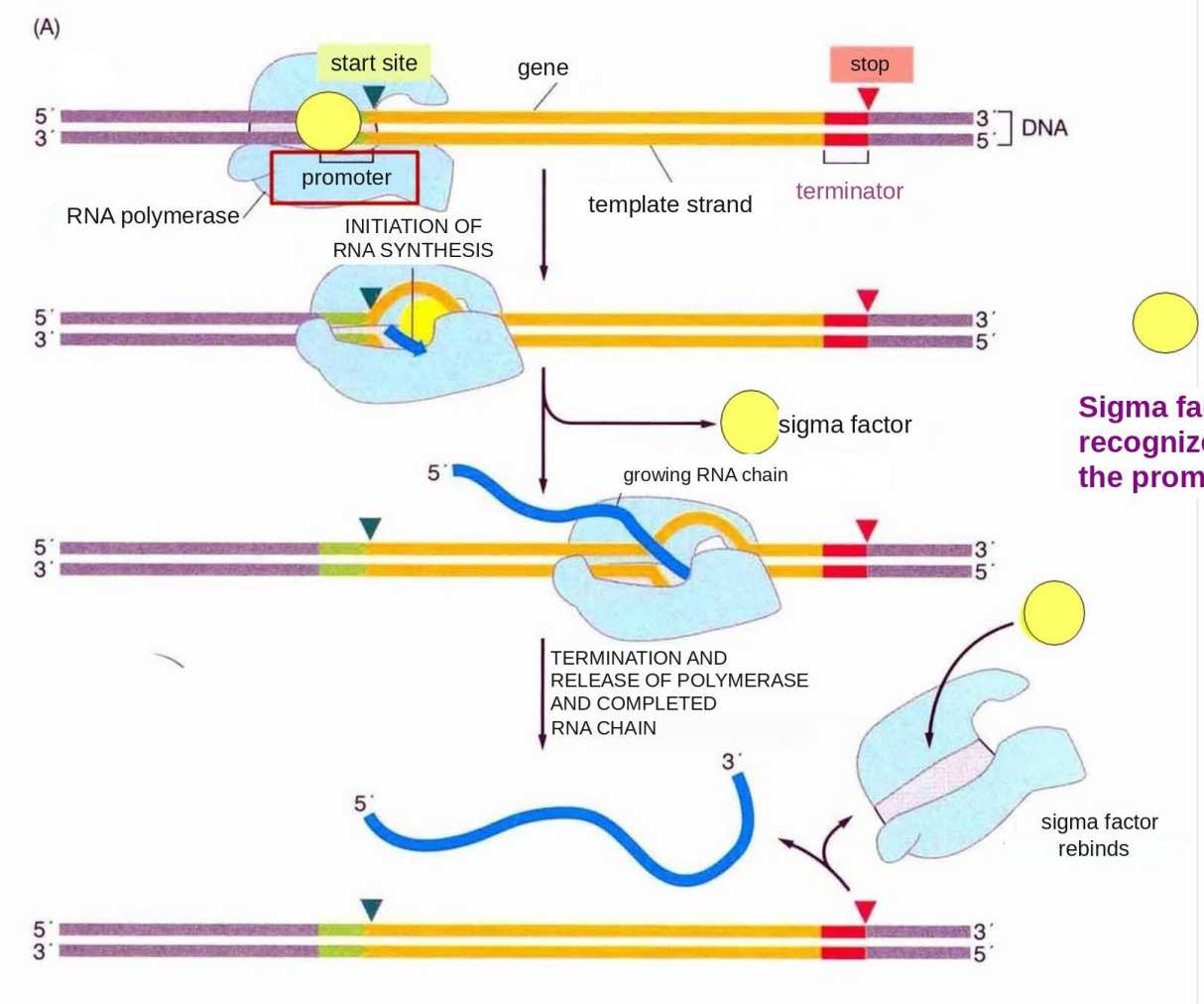
Sigma factor

Without σ the polymerase binds randomly; with σ it finds the promoter. It leaves after initiation.

Termination

A G–C-rich sequence forms an RNA hairpin, followed by Us → the polymerase falls off.

The bacterial RNA polymerase cycle



Sigma factor recognizes the promoter

One polymerase

Bacteria have one RNA polymerase for all genes.

Recycling

σ is released after the first ~10 nucleotides and is reused for a new start.

Alberts et al., Essential Cell Biology, 5th ed., Fig. 7-10

The three eukaryotic RNA polymerases

Polymerase	Transcribes	Where	α-amanitin
RNA Pol I	28S, 18S, 5.8S rRNA	nucleolus	resistant
RNA Pol II	all mRNAs; miRNAs, lncRNAs, most snRNAs	nucleoplasm	very sensitive
RNA Pol III	tRNA, 5S rRNA, U6 snRNA, other small RNAs	nucleoplasm	only at high dose

Bacteria

One polymerase + σ. Genes are tightly packed, often in operons.

Eukaryotes

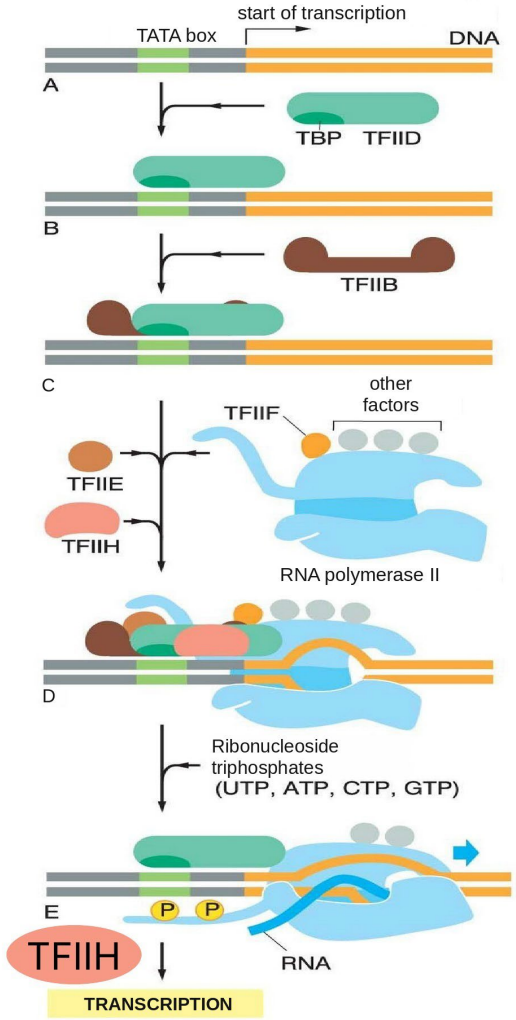
Three polymerases, many helper factors, genes packaged in chromatin.

Mitochondria

Their own single-subunit polymerase (POLRMT), related to phage T7 polymerase — not to the bacterial one.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Weinmann & Roeder, PNAS 1974

Eukaryotic initiation: the general transcription factors



1. **TFIID** (via TBP) recognizes the **TATA box** (~30 bp upstream of the start) and bends the DNA.

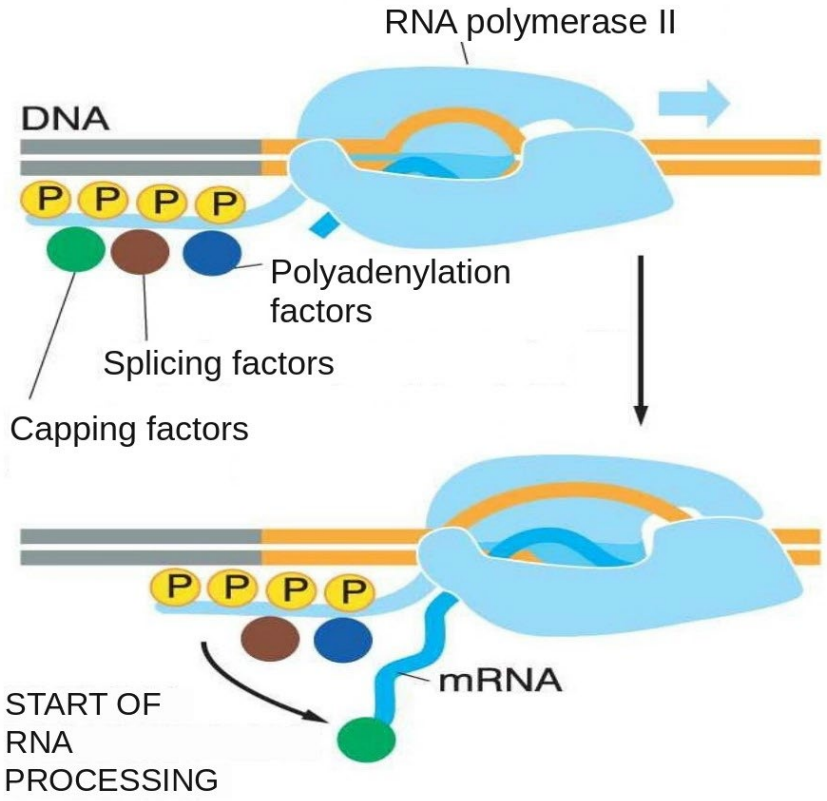
2. **TFIIB, TFIIF, TFIIE, TFIIH** and **RNA Pol II** assemble: the pre-initiation complex.

3. **TFIIH** opens the DNA (helicase) and phosphorylates the Pol II tail (CTD) → the polymerase takes off.

In vivo, activators, Mediator and chromatin modifiers are also needed → "Regulation of gene expression".

Alberts et al., Essential Cell Biology, 5th ed., Fig. 7-12 · Kornberg, Nobel Prize in Chemistry 2006

The Pol II tail: a platform for RNA processing



52×

YSPTSPS repeats in the human CTD

Ser5-P

at the gene start → capping

Ser2-P

along the gene → splicing, polyadenylation, termination

RNA processing happens while the RNA is being made (co-transcriptionally).

From research: PHF6 cooperates with SWI/SNF complexes to keep Pol II moving in rhabdoid tumor cells (Mittal et al., Nat Commun 2024).

Eick & Geyer, Chem Rev 2013 · Komarnitsky et al., Genes Dev 2000 · Mittal et al., Nat Commun 2024

Clinical link: drugs and toxins that stop transcription

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Rifampicin

Binds the β subunit of bacterial RNA polymerase and blocks the RNA after 2–3 nucleotides.

Tuberculosis · spares our Pol II

α-Amanitin

Toxin of the Amanita phalloides mushroom. Binds RNA Pol II and stops elongation.

Liver failure · hepatic uptake via OATP1B3

Actinomycin D

Intercalates into DNA (at GpC sites) and blocks polymerase movement.

Wilms tumor, rhabdomyosarcoma, trophoblastic tumors

TFIIH diseases

Subunits XPB and XPD are needed for both transcription and DNA repair.

Xeroderma pigmentosum, trichothiodystrophy

Campbell et al., Cell 2001 · Bushnell et al., PNAS 2002 · Letschert et al., Toxicol Sci 2006 · Sobell, PNAS 1985 · Compe & Egly, Annu Rev Biochem 2016

IN-CLASS EXERCISE #2 · 4 MINUTES · DISCUSS WITH YOUR NEIGHBOR

Clinical case: mushrooms from the forest

A 52-year-old eats wild mushrooms he picked himself. After 10 hours he has vomiting and diarrhea; on days 2–3 his transaminases rise sharply. Toxicology shows α-amanitin.

a Which enzyme is inhibited? Which class of RNA will fall first in hepatocytes: rRNA, tRNA or mRNA?

b Why is the liver mainly affected?

c Why does rifampicin kill bacteria but not our cells — and α-amanitin the opposite?

Hint: look at the table of the three polymerases and the previous slide.

Summary of Part A

1. Information flows DNA → RNA → protein; transcription and translation rely on base complementarity.
2. RNA differs from DNA (ribose, U, single-stranded) and folds; there are many types of functional RNA.
3. RNA polymerase reads the template 3′→5′ and makes RNA 5′→3′, without a primer, with ~1 error/10^4.
4. Bacteria: one polymerase + σ. Eukaryotes: Pol I, II, III and general transcription factors (TFIID–TFIIH).
5. The phosphorylated Pol II CTD couples transcription to RNA processing; drugs and toxins target the polymerases.

Part B: From RNA to protein

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

B1

1st hour · RNA processing and splicing

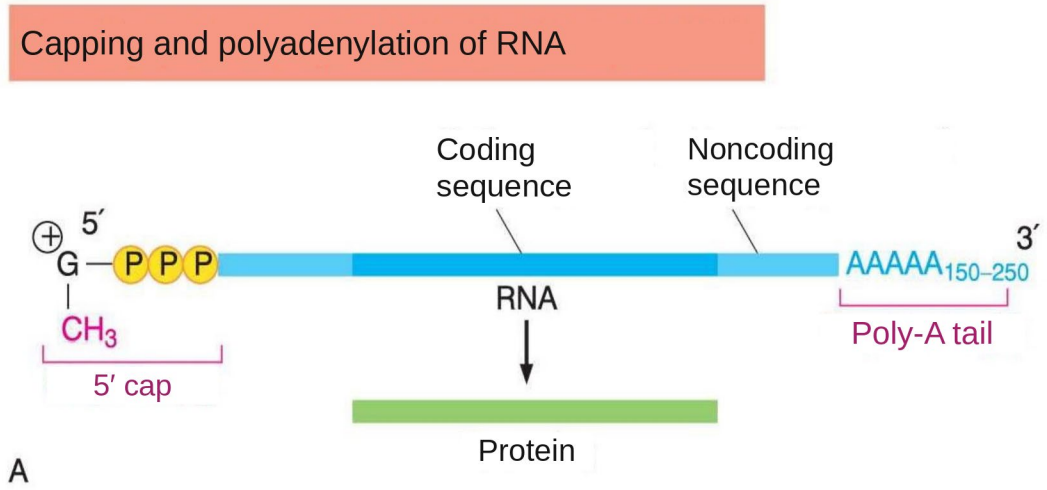
- Cap and poly(A) tail
- Exons, introns, the spliceosome
- Alternative splicing
- When splicing fails: thalassemia, progeria, SMA, DMD
- mRNA export · mRNA vaccines

B2

2nd hour · Translation and the fate of the protein

- The genetic code
- tRNA, aminoacyl-tRNA synthetases, ribosome
- Initiation, elongation, termination · NMD
- Antibiotics · folding · proteasome
- The RNA world

From precursor to mature mRNA: three modifications



1. **5′ cap:** 7-methylguanosine linked 5′–5′, added once the RNA is ~25 nucleotides long.
2. **Splicing:** introns removed, exons joined.
3. **Polyadenylation:** cleavage after the AAUAAA signal and addition of ~200 As.

Why?

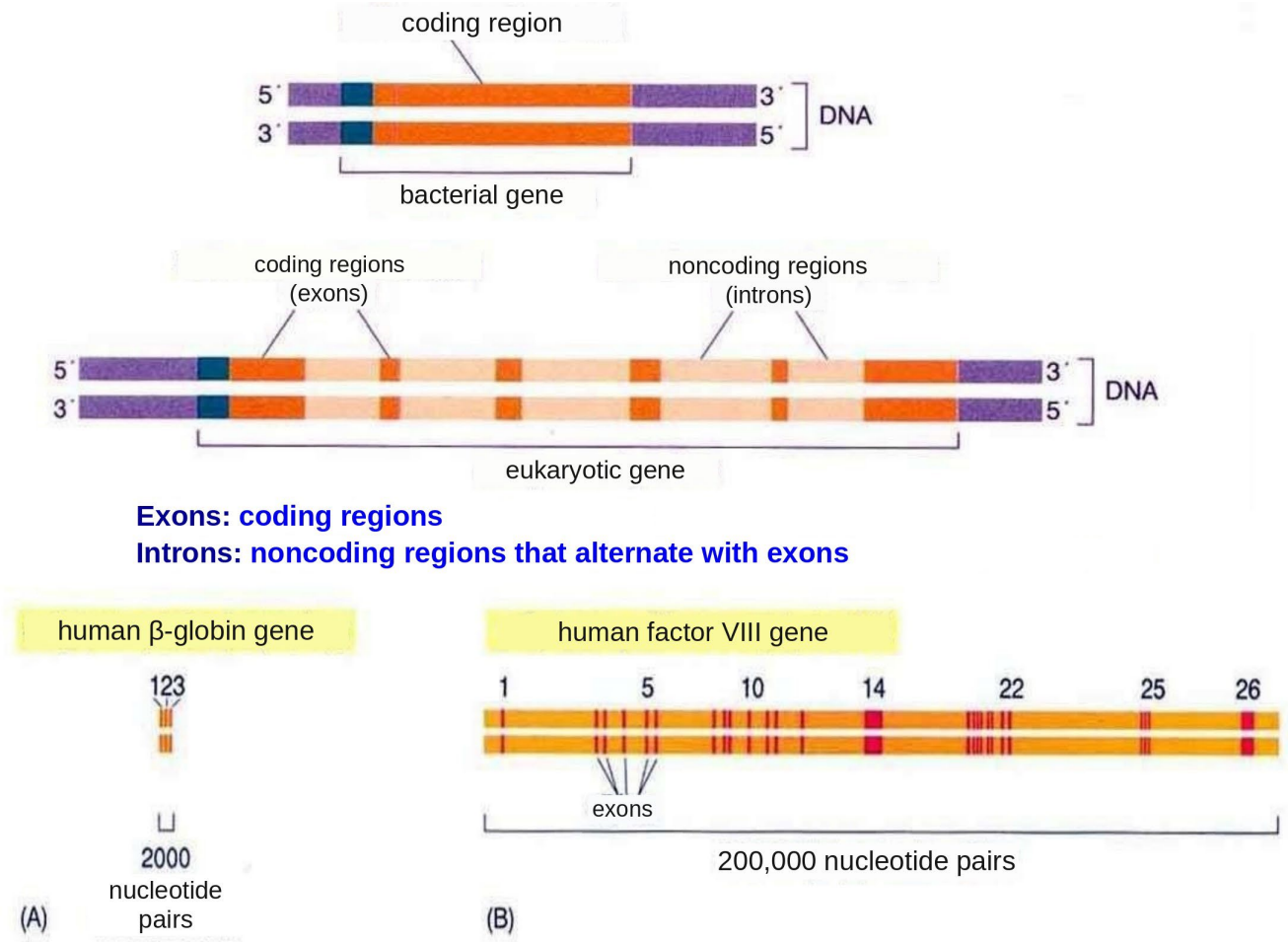
Cap and tail protect the mRNA from exonucleases, mark it for export and help start translation.

Clinical link

Mutations in the AAUAAA signal lower mRNA levels — e.g., in some α- and β-thalassemias.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Thein, Cold Spring Harb Perspect Med 2013

Our genes are split: exons and introns



~11

exons per transcript (mean)

~1.7 kb

median intron · exon ~130 nt

DMD: 79 exons

~2.3 Mb; transcribing it takes ~16 hours

Discovery (1977)

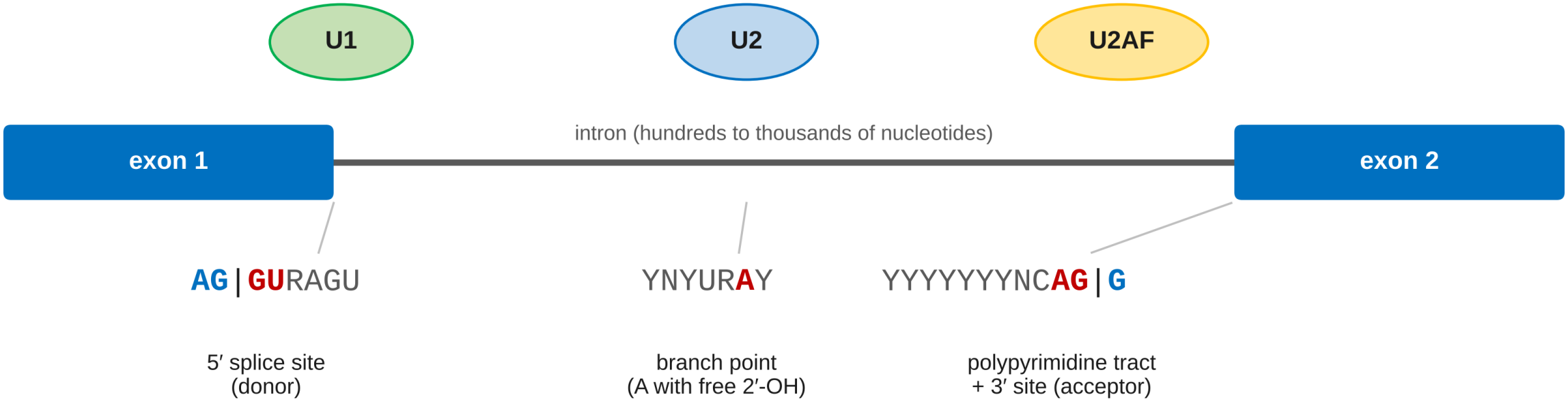
Sharp and Roberts, in adenovirus: mRNA is not a continuous copy of the gene. Nobel Prize 1993.

Piovesan et al., BMC Res Notes 2019 · Tennyson et al., Nat Genet 1995 · Berget et al., PNAS 1977 · Chow et al., Cell 1977

Where does the spliceosome cut? The signals in the RNA

R = A or G · Y = C or U · N = any base

GU–AG rule: ~99% of introns



U1 snRNP

base-pairs with the 5′ splice site (GU)

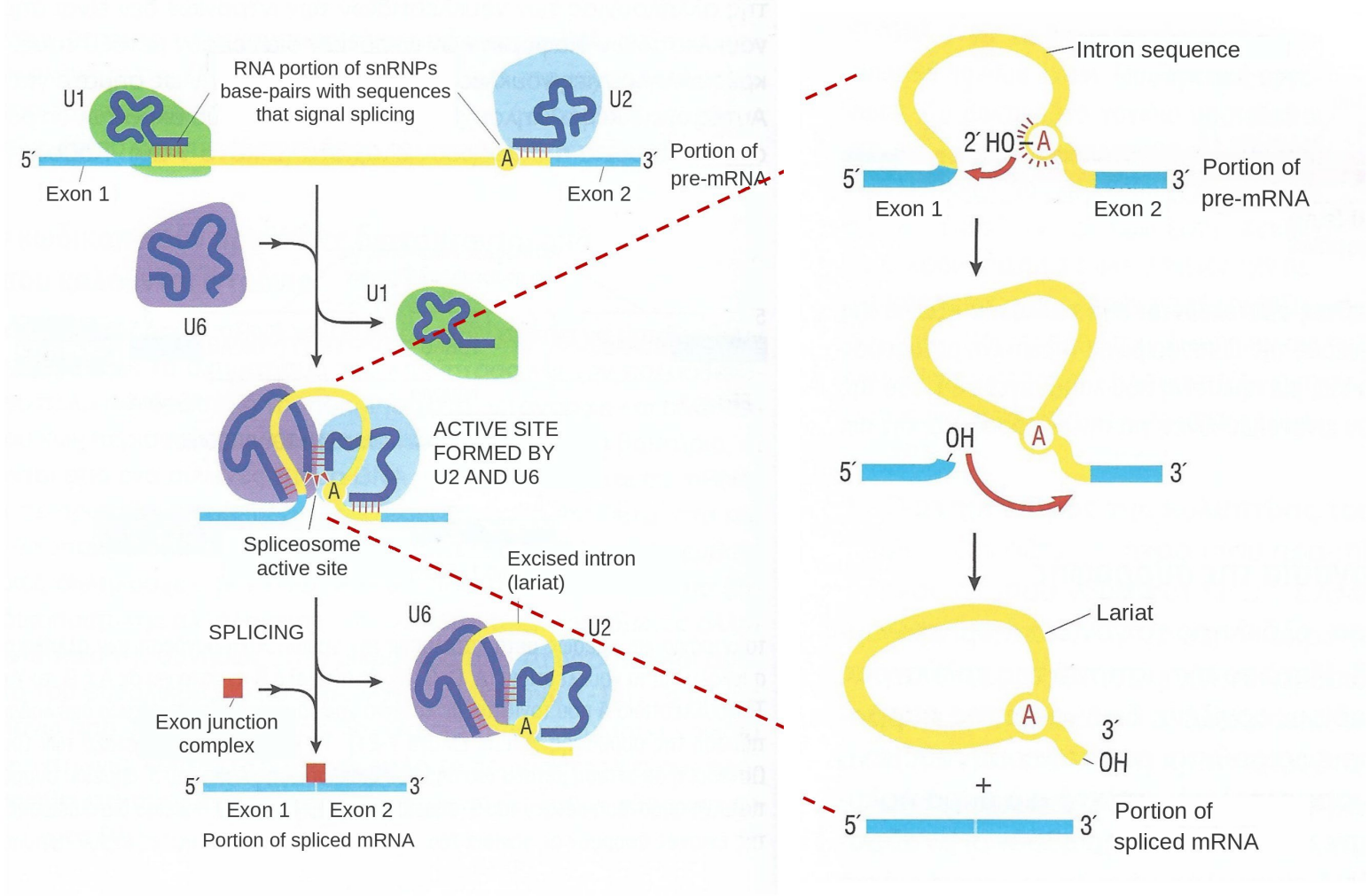
U2 snRNP

base-pairs around the branch-point A

Short, fuzzy signals → easily broken by a mutation.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Burset et al., Nucleic Acids Res 2000

The chemistry of splicing: two reactions and a lariat



1. The **branch-point A 2′-OH** attacks the 5′ site → the lariat forms.

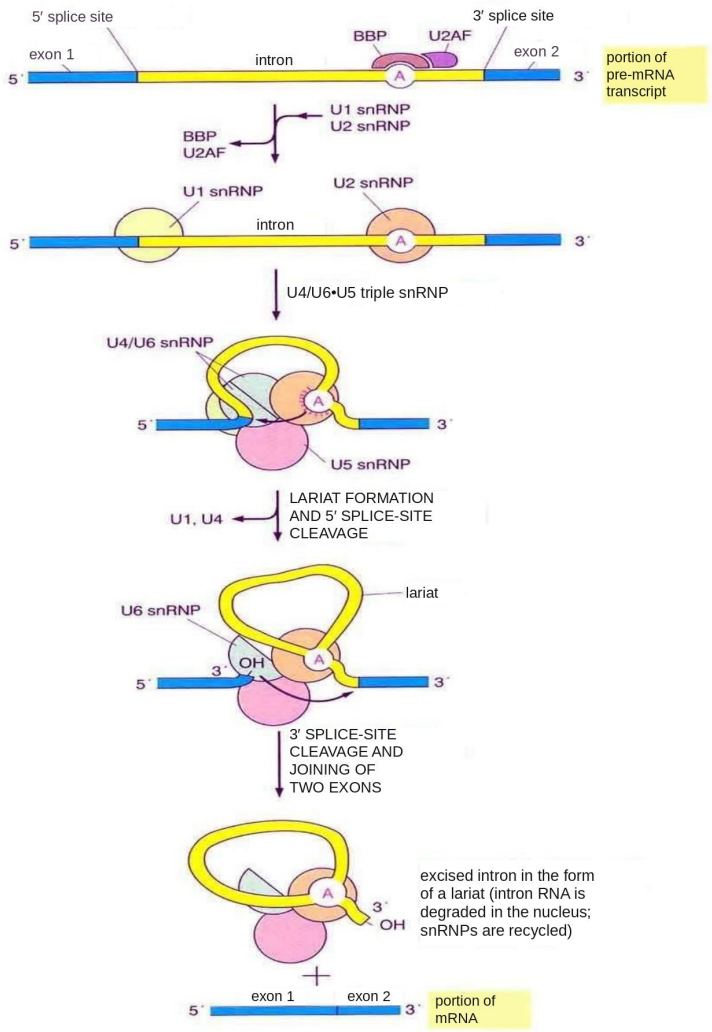
2. The **exon 1 3′-OH** attacks the 3′ site → exons are joined.

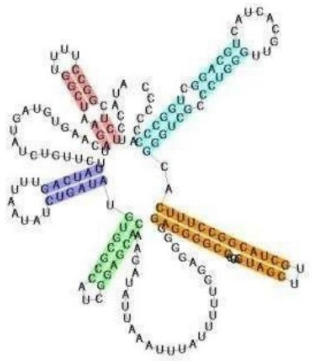
3. The lariat is released and **degraded**.

No net bond energy lost: two transesterifications.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

The spliceosome: a ribonucleoprotein machine





5 snRNA

U1, U2, U4, U5, U6

>100

proteins

RNA recognizes RNA

snRNAs pair with pre-mRNA signals and with each other; ATP-driven helicases rearrange the pairs at each step.

U6 snRNA positions the catalytic metal ions: the spliceosome is a ribozyme.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Fica et al., Nature 2013

IN-CLASS EXERCISE #3 · 10 MINUTES

Find the intron

A precursor mRNA (pre-mRNA) with two exons and one intron:

```
5′–  AUGGCCAAGGUAAGUAUCUUCACUAACUCUUUUCCCUUUUCAGGAAUGGUAA  –3′
```

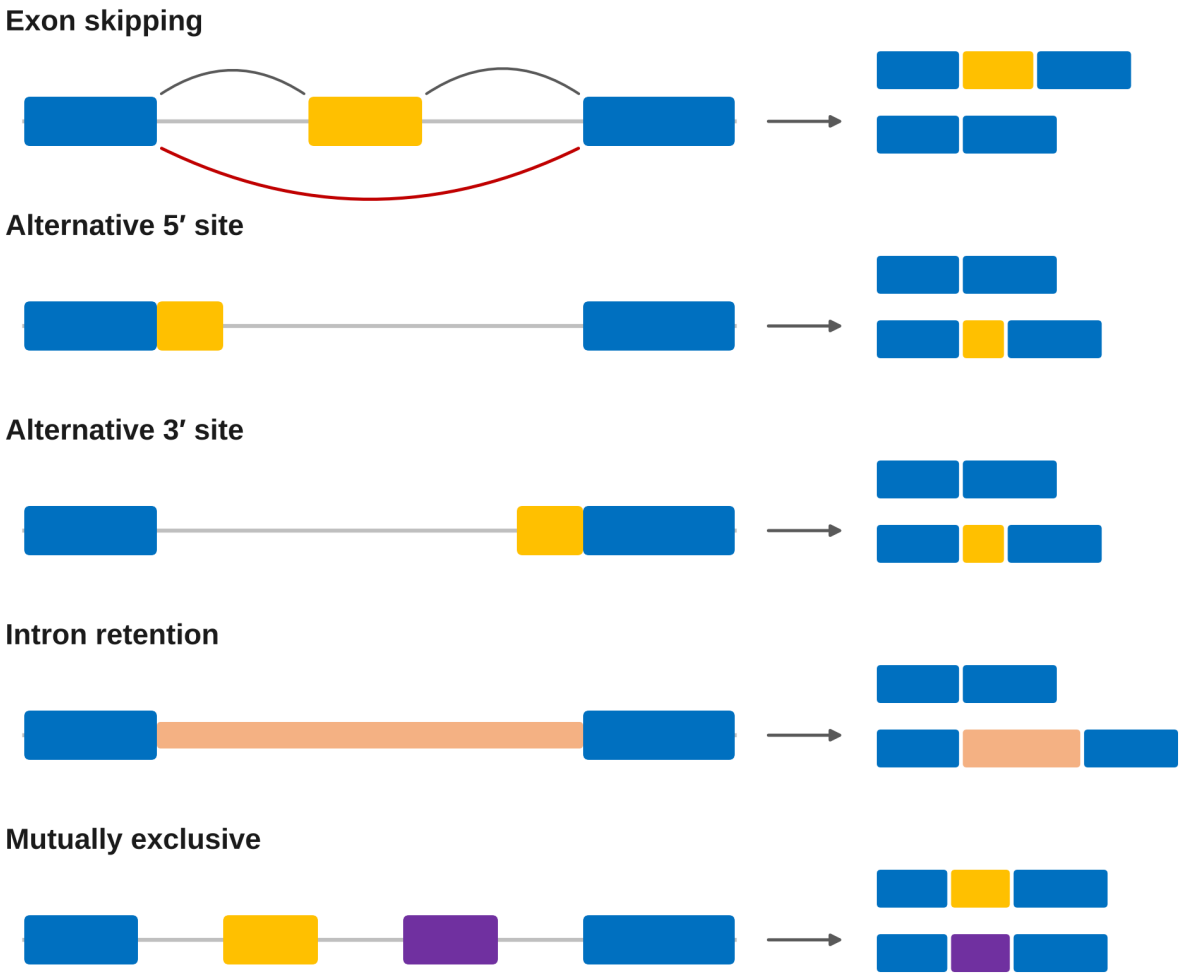
a Mark the 5′ splice site, the branch-point A and the 3′ splice site.

b Write the mature mRNA. How many nucleotides long is it?

c The G of GU becomes A (as in β-thalassemia IVS-I-1). Predict two possible outcomes.

Hint: GU … A … (pyrimidines) … AG

Alternative splicing: one gene, many proteins



~95%

of human multi-exon genes are alternatively spliced

~19,100 → ~47,000

protein-coding genes → protein-coding transcripts (RefSeq)

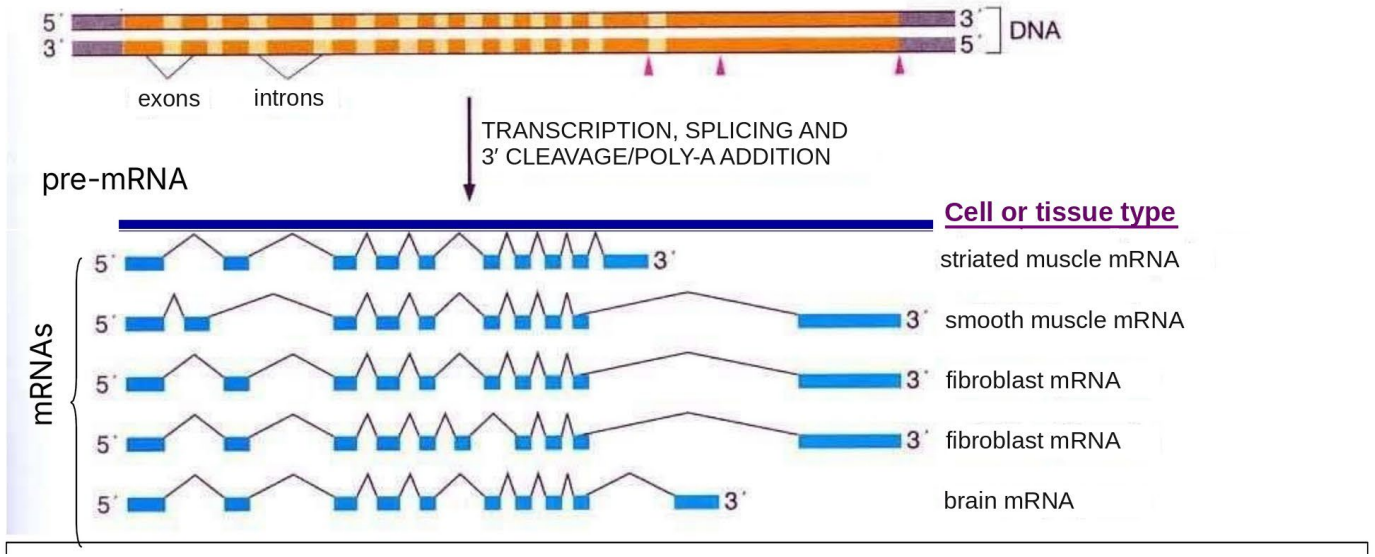
Who decides?

SR and hnRNP proteins bind enhancers and silencers in exons and introns — different in each tissue.

Pan et al., Nat Genet 2008 · Wang et al., Nature 2008 · Piovesan et al., BMC Res Notes 2019

Tissue-specific alternative splicing: examples

α-tropomyosin gene



Bcl-x

Bcl-xL blocks apoptosis · Bcl-xS promotes it. Same pre-mRNA.

Calcitonin / CGRP

Thyroid: the hormone calcitonin · neurons: the neuropeptide CGRP.

Dscam (Drosophila)

Up to 38,016 isoforms from one gene — an "ID tag" for each neuron.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Schmucker et al., Cell 2000

When splicing fails: four mechanisms

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

1. Site destroyed



GU→AU or AG→AA
→ exon skipping or
intron retention

2. New cryptic site



a new GU or AG appears
→ exon is lengthened
or shortened

3. Enhancers/silencers



change inside the exon
(even "silent")
→ exon is lost

4. Deep intronic



a "pseudo-exon"
is activated
inside the intron

≥10%

of inherited-disease mutations: at exon–intron junctions

~10%

of exonic disease mutations alter splicing

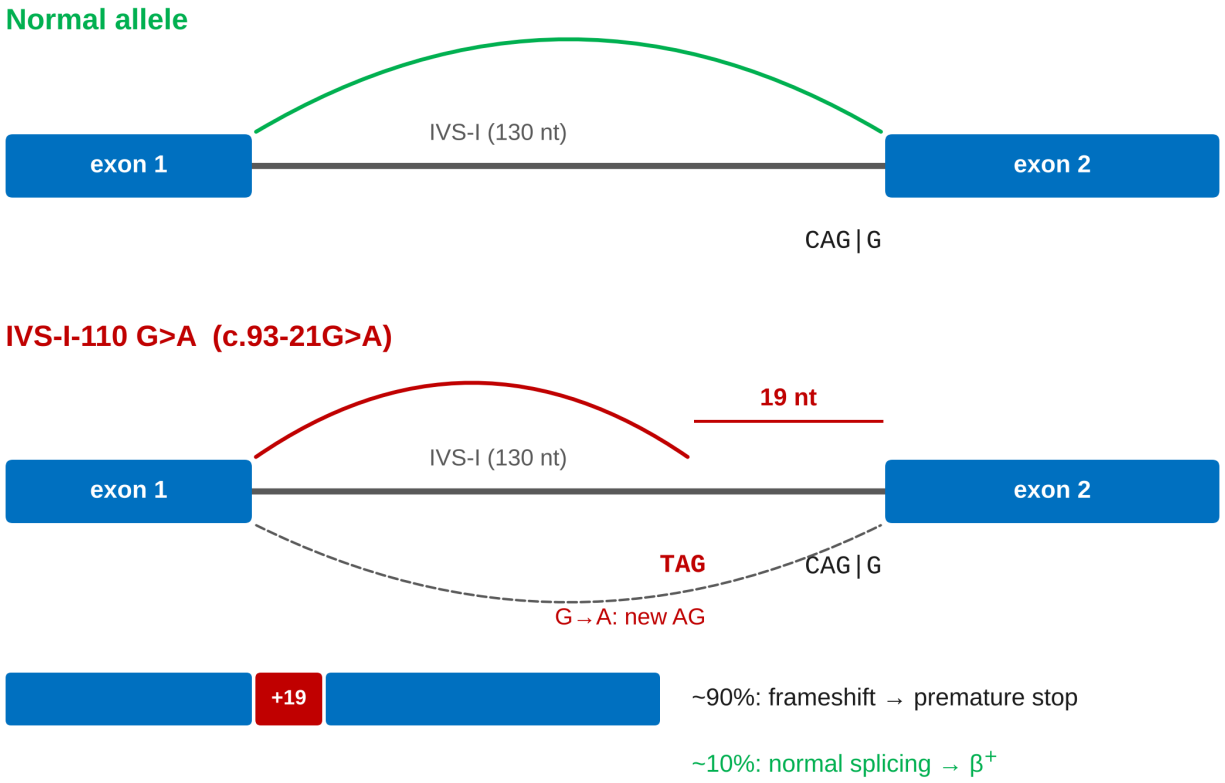
9–11%

of pathogenic mutations in rare disease: cryptic sites (SpliceAI)

This is why we now read introns too — not only codons.

Krawczak et al., Hum Mutat 2007 · Soemedi et al., Nat Genet 2017 · Jaganathan et al., Cell 2019

β-thalassemia in Greece: a disease of splicing



Allele	Step	%
IVS-I-110 G>A	splicing (new AG)	42.5
Codon 39 C>T	stop → NMD	17.0
IVS-I-1 G>A	splicing (GU lost)	13.2
IVS-I-6 T>C	splicing (weak 5′)	7.2
IVS-II-745 C>G	splicing (cryptic exon)	6.9

174 Greek patients with thalassemia major (Kattamis et al., 1990)

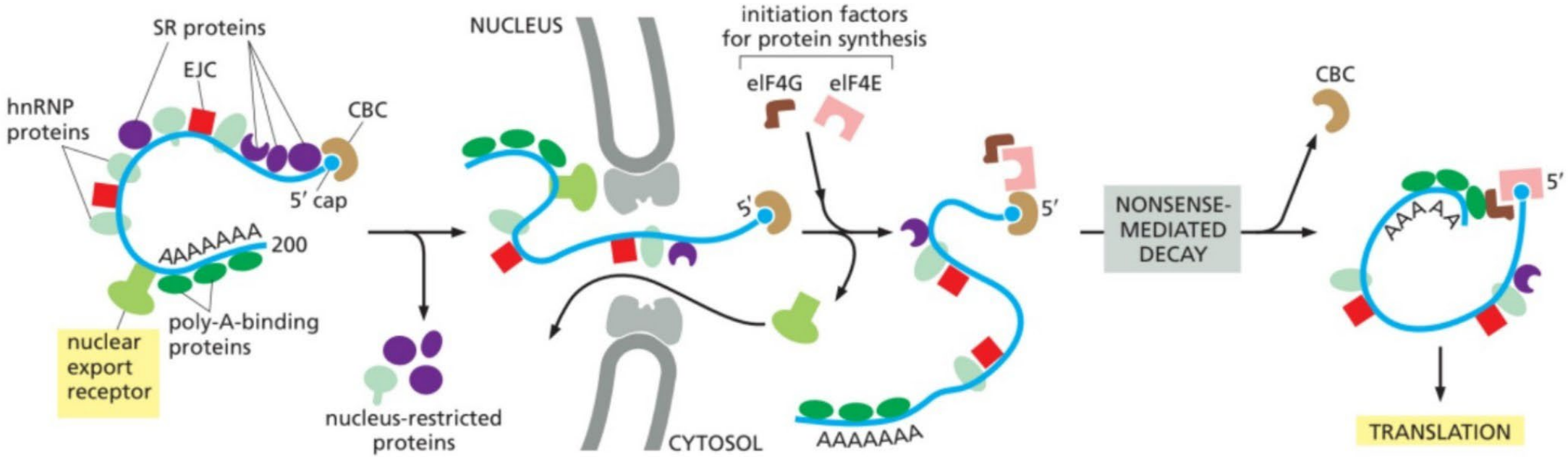
~8% carriers in Greece

Mutations in SW Greece: University of Patras study (Papachatzopoulou et al., 2010)

Busslinger, Moschonas & Flavell, Cell 1981: the first functional proof (expression of the cloned gene) that a single point mutation in an intron causes aberrant splicing. N. Moschonas is Professor Emeritus of Biology in our Department.

Busslinger et al., Cell 1981 · Fukumaki et al., Cell 1982 · Treisman et al., Nature 1983 · Kattamis et al., Br J Haematol 1990 · Papachatzopoulou et al., Hemoglobin 2010

Quality control and export of mRNA



Only mature mRNA

Incomplete transcripts are retained and degraded in the nucleus.

Nuclear pores

Export receptors recognize proteins on the cap, the tail and the junctions.

EJC mark

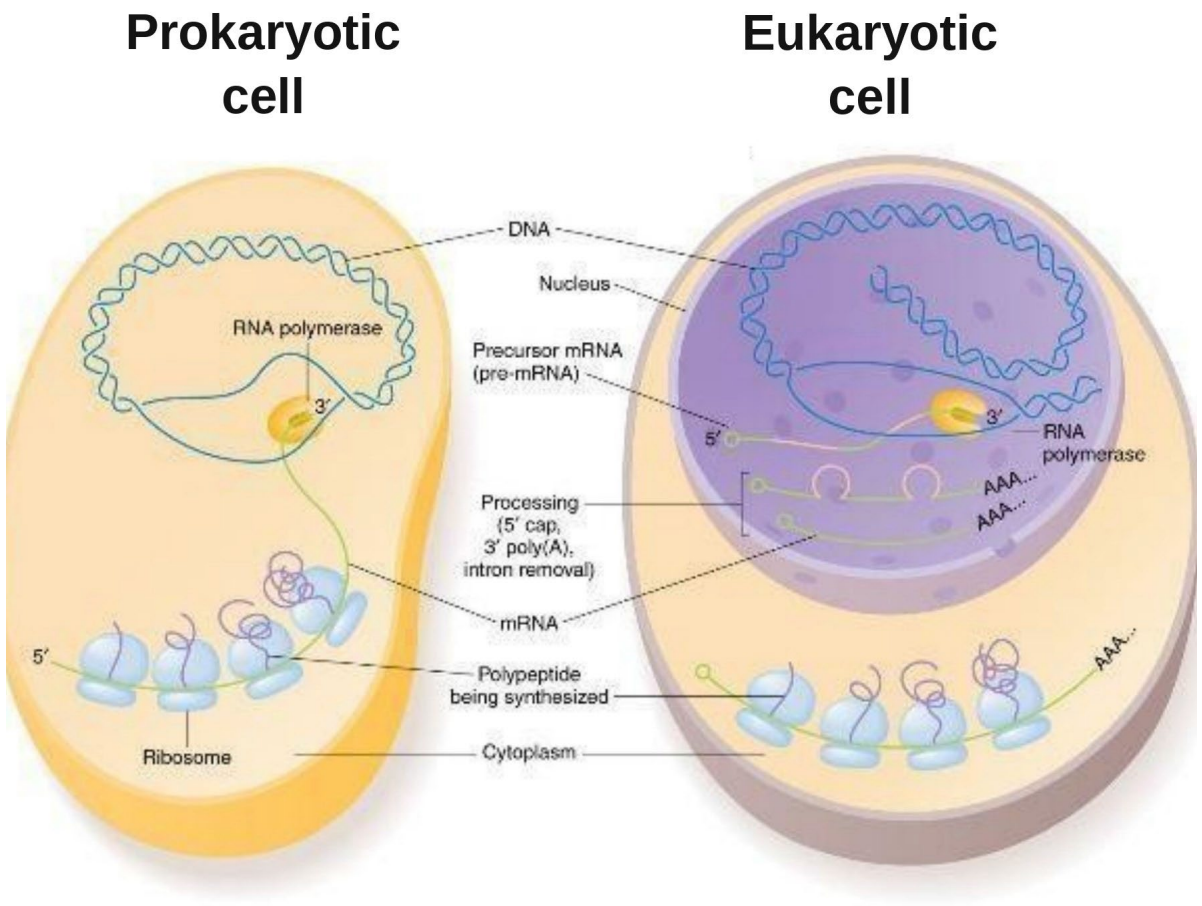
The exon junction complex sits ~20–24 nt upstream of each junction → we will meet it again in NMD.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Le Hir et al., EMBO J 2000

Bacteria vs eukaryotes: from gene to mRNA

Bacteria

- One RNA polymerase
- No introns, cap or tail
- Translation starts before transcription ends
- mRNA lives a few minutes



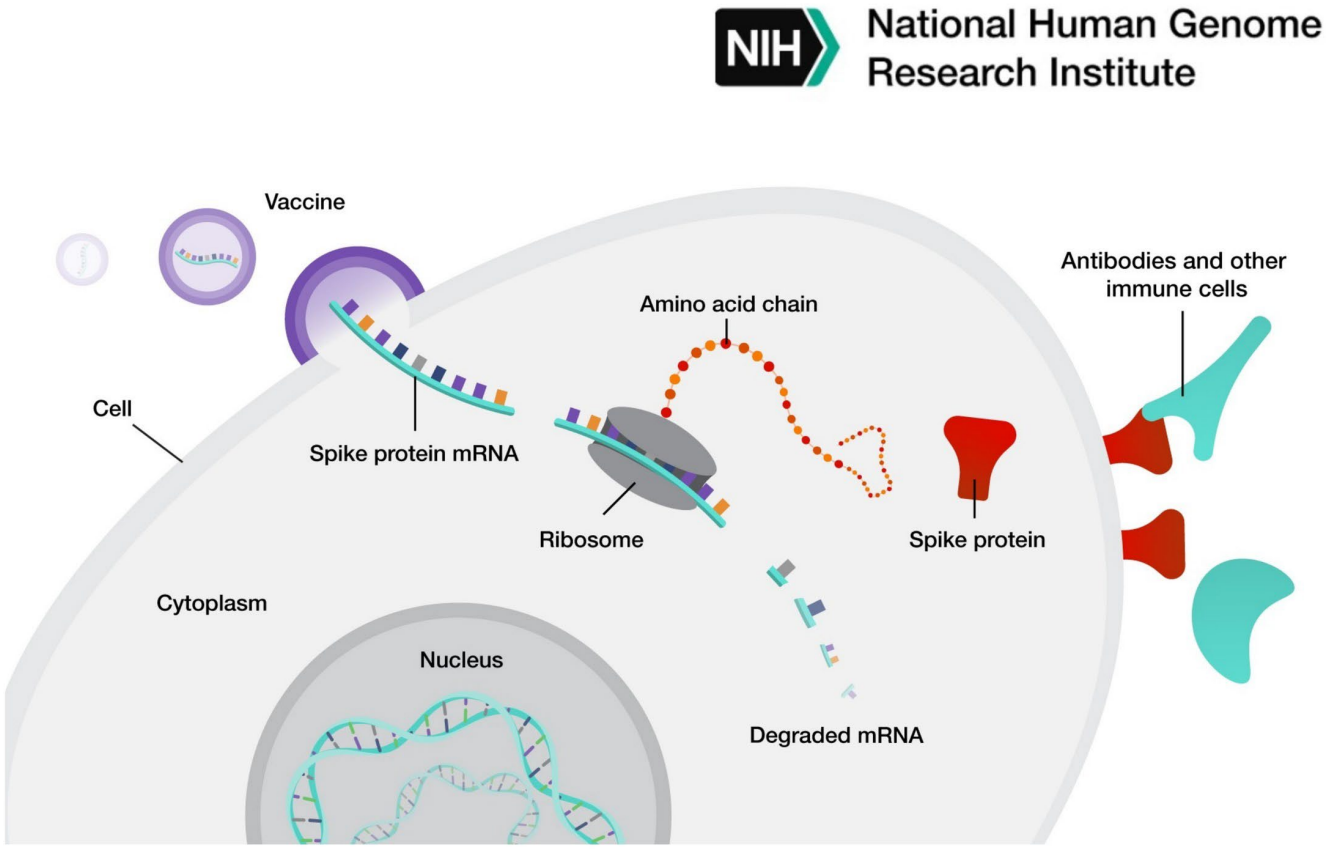
Eukaryotes

- Three RNA polymerases
- Cap, splicing, poly(A)
- Export to the cytoplasm
- mRNA lives minutes to >10 hours

The nucleus uncouples transcription from translation — time for processing and checks.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

mRNA vaccines: facts and myths



Nobel 2023: Karikó and Weissman

Modified nucleosides (pseudouridine) make mRNA invisible to innate immunity.

What it does

Packed in lipid nanoparticles; translated in the cytoplasm, then degraded.

What it does not do

It does not enter the nucleus or become DNA.

Mature mRNA → translation only. The vaccinated person's genome does not change.

Karikó et al., Immunity 2005 · Pardi et al., Nat Rev Drug Discov 2018 · CDC · Image: NHGRI

Summary of B1

1. Pre-mRNA gets a 5′ cap and a poly(A) tail while it is made; this protects it and allows export and translation.
2. Introns are removed by the spliceosome (snRNPs; U6 catalyzes) guided by GU … A … AG signals.
3. Alternative splicing (~95% of genes) gives many proteins from one gene, depending on the tissue.
4. Splicing mutations cause disease (β-thalassemia, progeria) — even "silent" or deep intronic ones.
5. Splicing is a drug target: ASOs and small molecules (SMA, DMD, milasen).

Cracking the genetic code

The Nobel Prize in Physiology or Medicine 1968

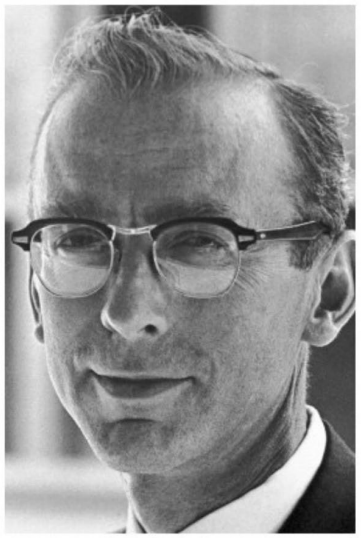


Photo from the Nobel Foundation archive.

Robert W. Holley

Prize share: 1/3



Photo from the Nobel Foundation archive.

Har Gobind Khorana

Prize share: 1/3



Photo from the Nobel Foundation archive.

Marshall W. Nirenberg

Prize share: 1/3

Holley: first tRNA sequence (alanine tRNA, 1965) · Khorana: chemical synthesis of RNA of defined sequence · Nirenberg: cell-free protein synthesis

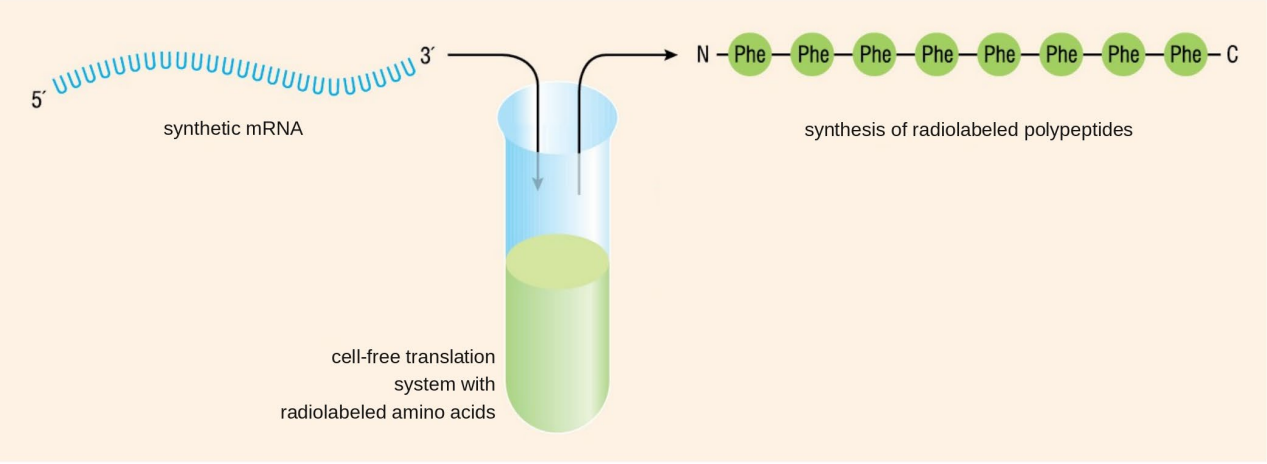
How many letters per "word"?

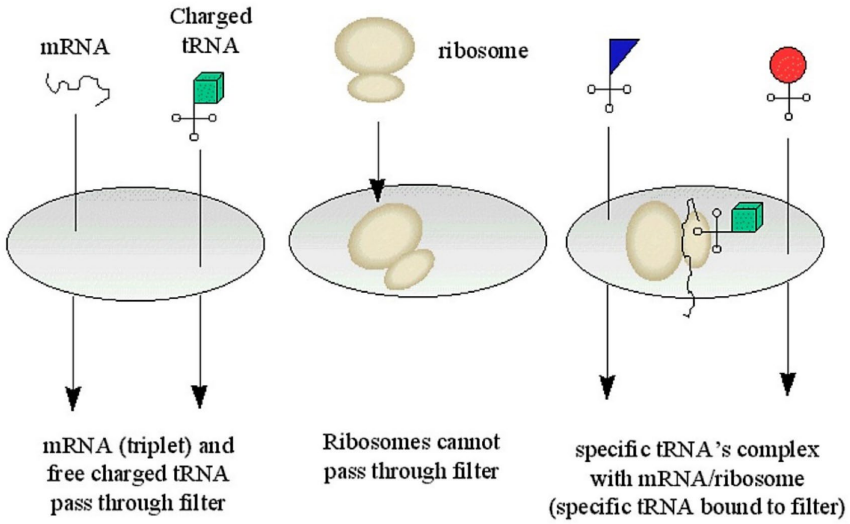
1 base	$4^1 = 4$	**< 20**
2 bases	$4^2 = 16$	**< 20**
3 bases	$4^3 = 64$	**≥ 20 ✓**

4 bases, 20 amino acids: the code must be at least a triplet. Nobel Prize 1968.

Image: nobelprize.org · Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

The experiments that read the code





1961 · Nirenberg & Matthaei

Synthetic poly-U in an E. coli extract → polyphenylalanine: UUU = Phe.

1964 · Nirenberg & Leder

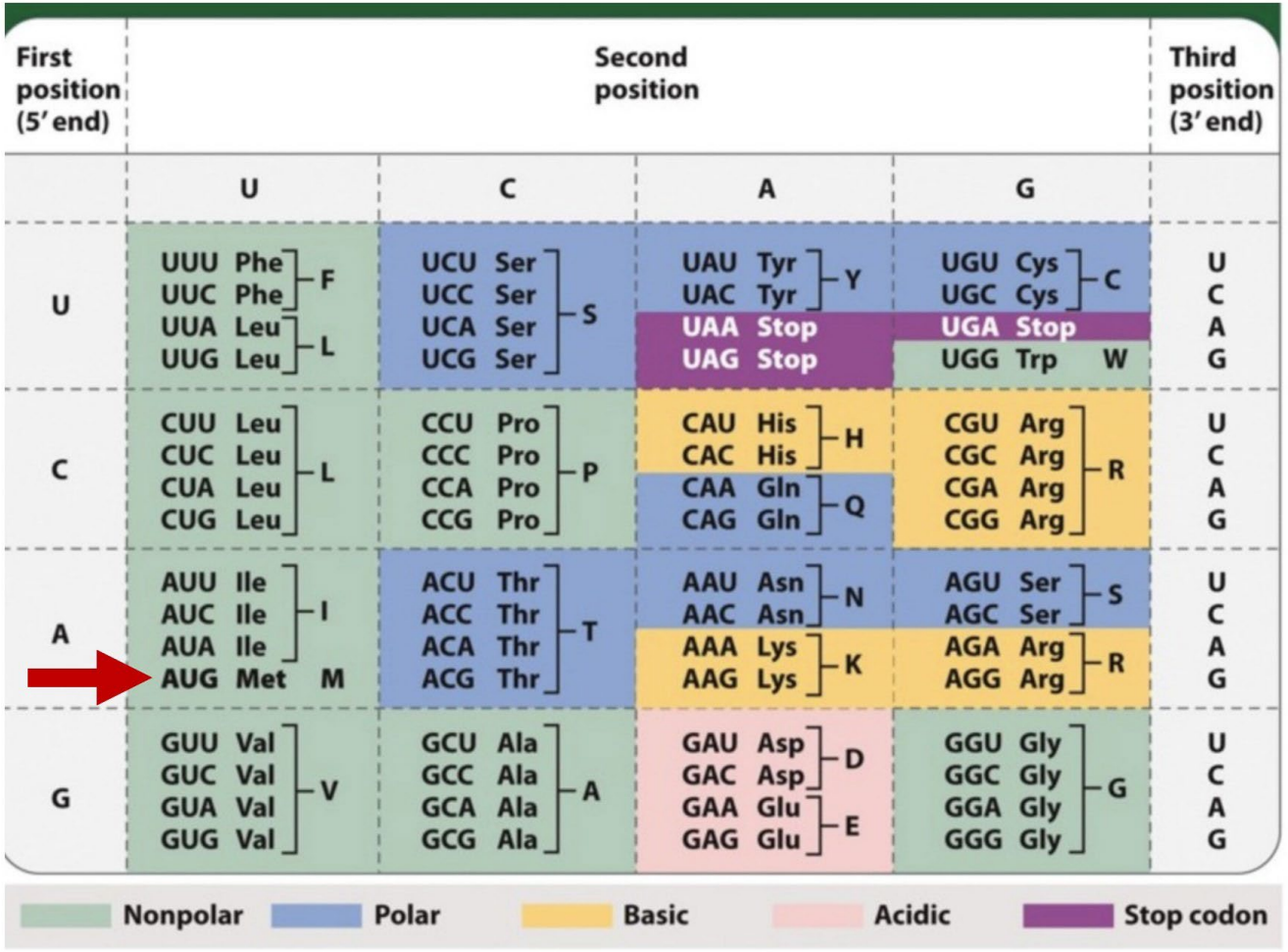
A single triplet binds the right charged tRNA to the ribosome → filter → nearly all codons.

Khorana

Repeating synthetic RNAs (e.g., UGUG… → Cys-Val-Cys…) confirmed the triplets.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

The genetic code



First position (5′ end)	Second position: U	Second position: C	Second position: A	Second position: G	Third position (3′ end)
U	UUU Phe (F)	UCU Ser (S)	UAU Tyr (Y)	UGU Cys (C)	U
	UUC Phe (F)	UCC Ser (S)	UAC Tyr (Y)	UGC Cys (C)	C
	UUA Leu (L)	UCA Ser (S)	UAA Stop	UGA Stop	A
	UUG Leu (L)	UCG Ser (S)	UAG Stop	UGG Trp (W)	G
C	CUU Leu (L)	CCU Pro (P)	CAU His (H)	CGU Arg (R)	U
	CUC Leu (L)	CCC Pro (P)	CAC His (H)	CGC Arg (R)	C
	CUA Leu (L)	CCA Pro (P)	CAA Gln (Q)	CGA Arg (R)	A
	CUG Leu (L)	CCG Pro (P)	CAG Gln (Q)	CGG Arg (R)	G
A	AUU Ile (I)	ACU Thr (T)	AAU Asn (N)	AGU Ser (S)	U
	AUC Ile (I)	ACC Thr (T)	AAC Asn (N)	AGC Ser (S)	C
	AUA Ile (I)	ACA Thr (T)	AAA Lys (K)	AGA Arg (R)	A
	AUG Met (M)	ACG Thr (T)	AAG Lys (K)	AGG Arg (R)	G
G	GUU Val (V)	GCU Ala (A)	GAU Asp (D)	GGU Gly (G)	U
	GUC Val (V)	GCC Ala (A)	GAC Asp (D)	GGC Gly (G)	C
	GUA Val (V)	GCA Ala (A)	GAA Glu (E)	GGA Gly (G)	A
	GUG Val (V)	GCG Ala (A)	GAG Glu (E)	GGG Gly (G)	G

Nonpolar · Polar · Basic · Acidic · Stop codon

64 codons

61 for amino acids + 3 stops (UAA, UAG, UGA). Start: AUG (Met).

Degenerate

Most amino acids have >1 codon — usually differing at the 3rd base.

Unambiguous

Each codon means one amino acid only.

Nearly universal

Exceptions: in mitochondria UGA = Trp; selenocysteine is encoded by UGA in ~25 human proteins.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Barrell et al., Nature 1979 · Kryukov et al., Science 2003

Three reading frames — only one makes sense

5′-ATGCAATGGAATGACGCCAGGTTAAAATAA-3′

+1	**ATG-CAA-TGG-AAT-GAC-GCC-AGG-TTA-AAA-TAA** THE-BIG-CAT-ATE-THE-FAT-RAT	**Met … Lys – stop: open reading frame**
+2	**A-TGC-AAT-GGA-ATG-ACG-CCA-GGT-TAA-AAT-AA** T-HEB-IGC-ATA-TET-HEF-ATR-AT	**meaningless "words" → wrong protein**
+3	**AT-GCA-ATG-GAA-TGA-CGC-CAG-GTT-AAA-ATA-A** TH-EBI-GCA-TAT-ETH-EFA-TRA-T	**stop (TGA) at the 4th full codon**

AUG sets the frame; from there the message is read in triplets, with no gaps or overlaps.

Point mutations and the reading frame

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Synonymous ("silent")

GGC → GGT **Gly → Gly**

Usually harmless — unless it alters splicing (progeria).

Missense

GAG → GTG **Glu → Val**

HbS: sickle cell disease (β-globin codon 6).

Nonsense

CAG → TAG **Gln → STOP**

β⁰-thalassemia, codon 39 — common in Greece.

Frameshift

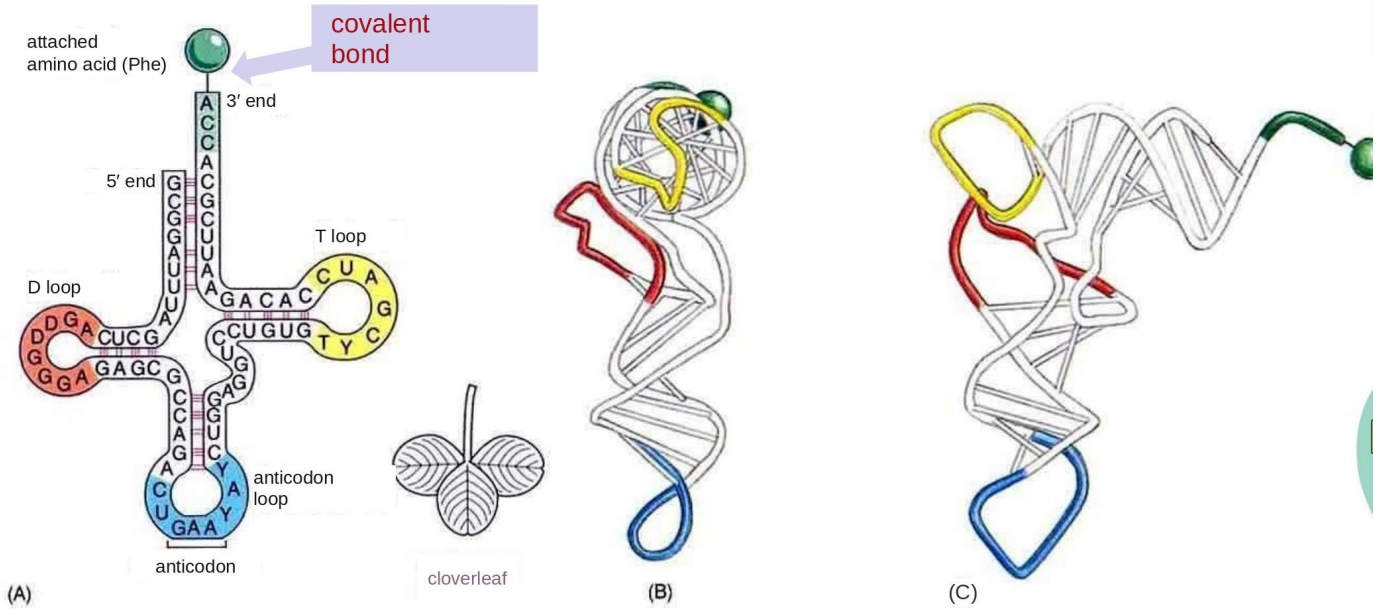
+1 / −1 nt **everything after changes**

THE-BIG-CAA-TAT-ETH…

Insertions or deletions of 3 (or a multiple of 3) keep the frame — as in DMD.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

tRNA: the translator with the anticodon



Two ends

Anticodon: pairs with the codon · 3′ CCA end: carries the amino acid.

Wobble

The 3rd codon position pairs less strictly (e.g., G–U, inosine).

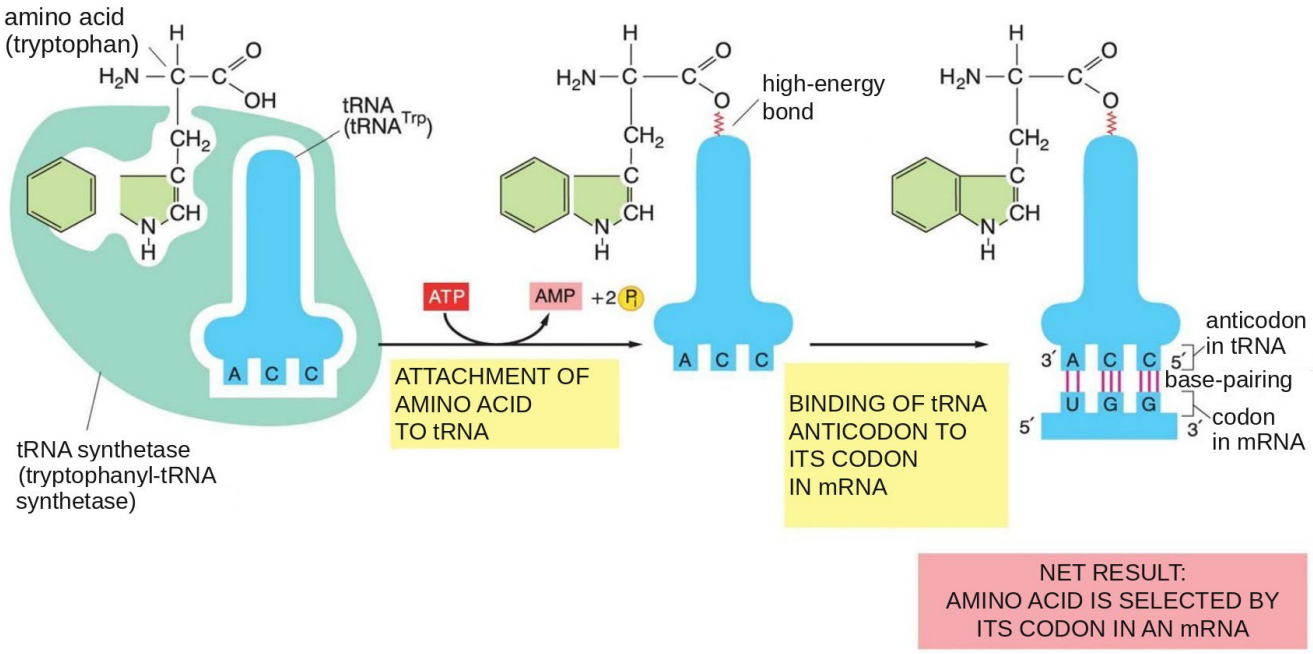
~420
tRNA genes

~48
anticodons, not 61

tRNA reads only the codon — it does not check which amino acid it carries. That is another enzyme's job…

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Chan et al., Nucleic Acids Res 2021

Aminoacyl-tRNA synthetases: the "second" genetic code



20 enzymes

One per amino acid; it recognizes the amino acid and all its tRNAs.

Two steps + proofreading

ATP → aminoacyl-AMP → aminoacyl-tRNA. Wrong amino acid? Hydrolyzed at the editing site.

Drugs

Mupirocin inhibits bacterial Ile-tRNA synthetase (staphylococcal ointment) · tavaborole: antifungal (Leu).

Disease

Mutations in GARS1, YARS1 (and others) → Charcot–Marie–Tooth disease: peripheral neuropathy.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Hughes & Mellows, Biochem J 1978 · Rock et al., Science 2007 · Antonellis et al., AJHG 2003 · Jordanova et al., Nat Genet 2006

IN-CLASS EXERCISE #4 · 3 MINUTES · USE THE CODE TABLE

Translate HBB

The mRNA from exercise #1 (β-globin). By convention, Met is not numbered: GUG is codon 1.

```
5′- AUG GUG CAU CUG ACU CCU GAG GAG AAG …
      0   1   2   3   4   5   6   7   8
```

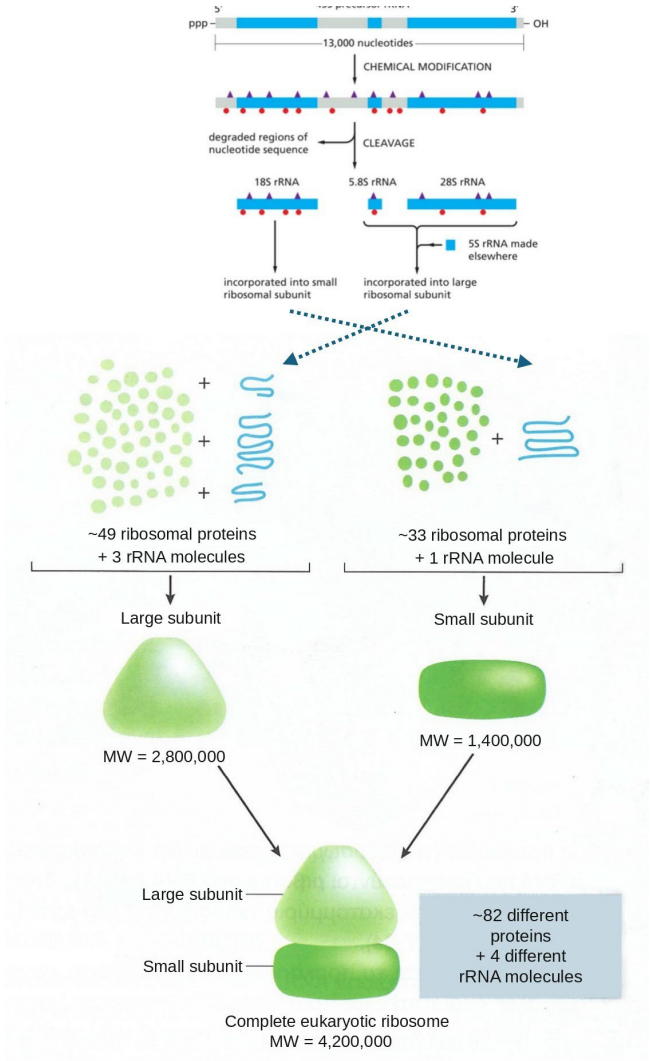
a Write the amino acid sequence (0–8).

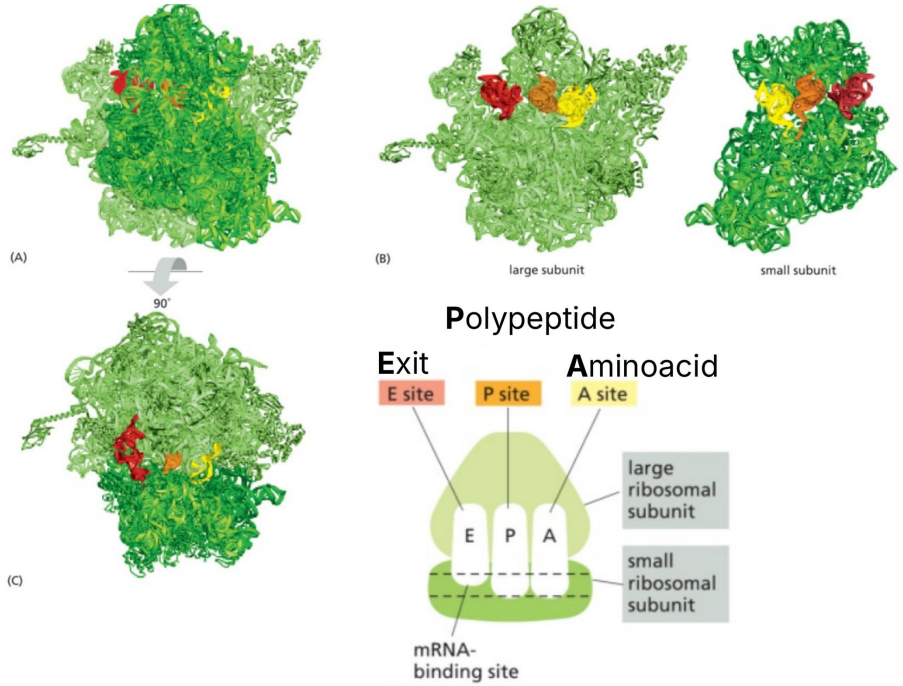
b Codon 6: GAG → GUG. What changes? What type of mutation? Which disease?

c Codon 39: CAG → UAG. What type of mutation? What do you expect for the protein?

d One nucleotide is lost in codon 8. What happens to the rest of the message?

The ribosome: a machine made of RNA and protein





Two subunits

40S: 18S rRNA + ~33 proteins · 60S: 28S, 5.8S, 5S rRNA + ~47 proteins.

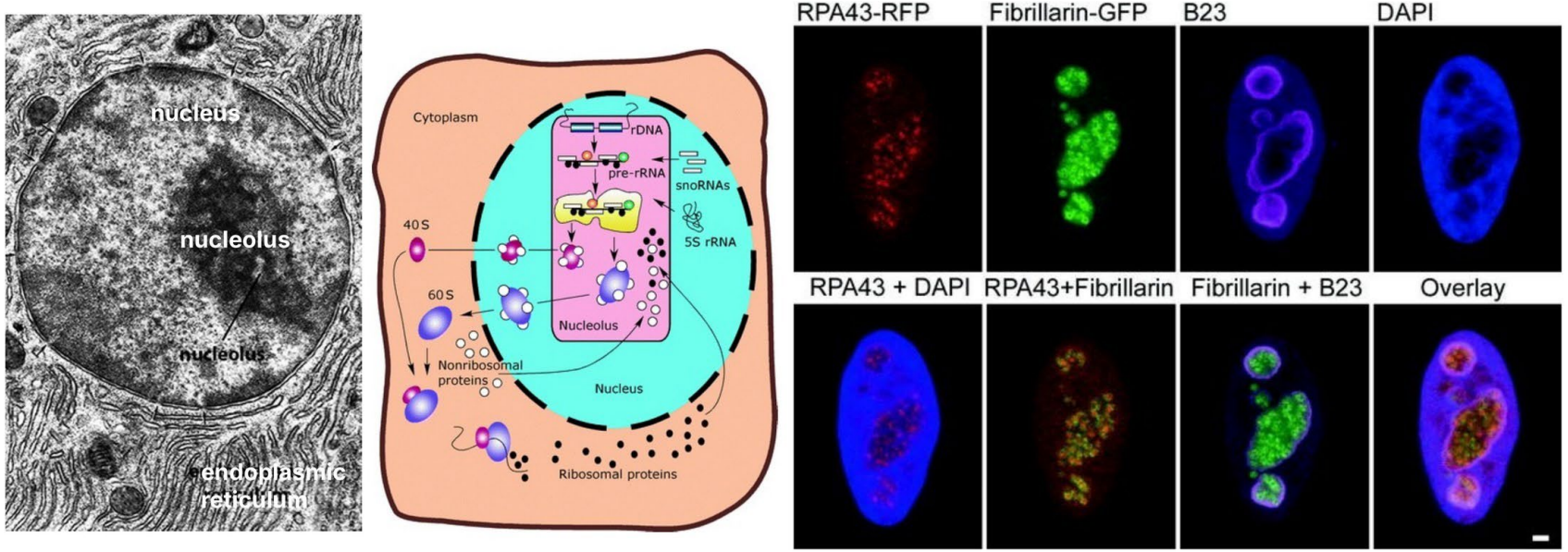
Three sites

A: aminoacyl-tRNA enters · P: peptidyl-tRNA · E: exit.

The peptide bond is made by rRNA, not protein: the ribosome is a ribozyme (Nobel Prize 2009).

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Ban et al., Science 2000 · Nissen et al., Science 2000

The nucleolus: the ribosome factory



Olson, M.O.J., Dundr, M. The moving parts of the nucleolus.
Histochem Cell Biol **123**, 203–216 (2005)

Where

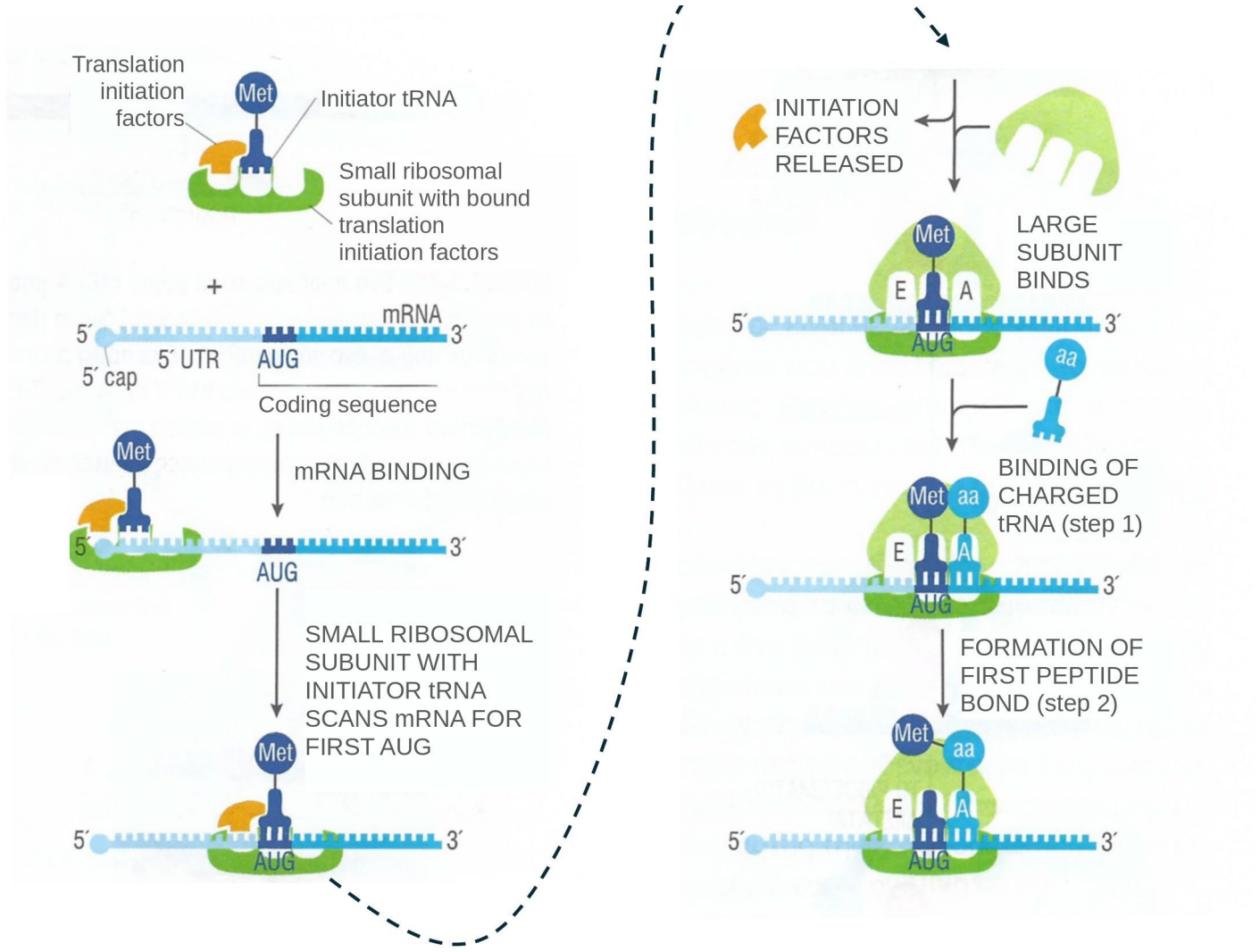
Around the rRNA genes on the short arms of the acrocentric chromosomes (13, 14, 15, 21, 22).

Ribosomopathies

Diamond–Blackfan anemia: mutations in ribosomal proteins (e.g., RPS19) → failure of red-cell production.

Olson & Dundr, Histochem Cell Biol 2005 · Draptchinskaia et al., Nat Genet 1999

Initiation: finding the first AUG



Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

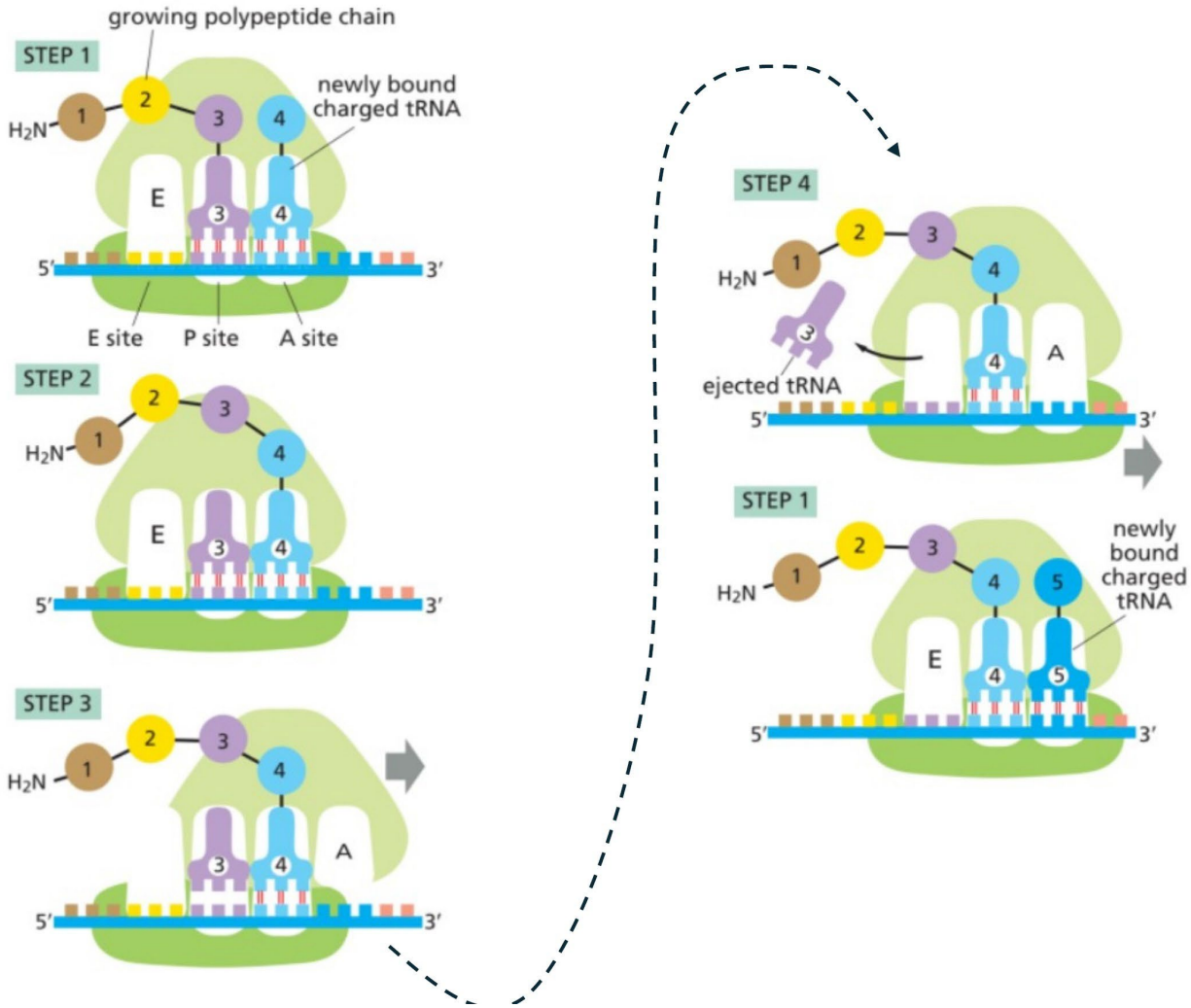
1. The small subunit, with **initiation factors** and **Met-tRNAi**, binds the 5′ cap.

2. It **scans** 5′→3′ to the first AUG (in a good "Kozak" context).

3. Factors leave, the **large subunit** joins; Met-tRNAi sits in the P site.

Bacteria: the Shine–Dalgarno sequence marks the AUG; one mRNA, several proteins.

Elongation: a four-step cycle



1. Aminoacyl-tRNA enters the A site
2. Peptide bond (catalyzed by rRNA)
3. The large subunit translocates
4. The small subunit follows; empty tRNA leaves from E

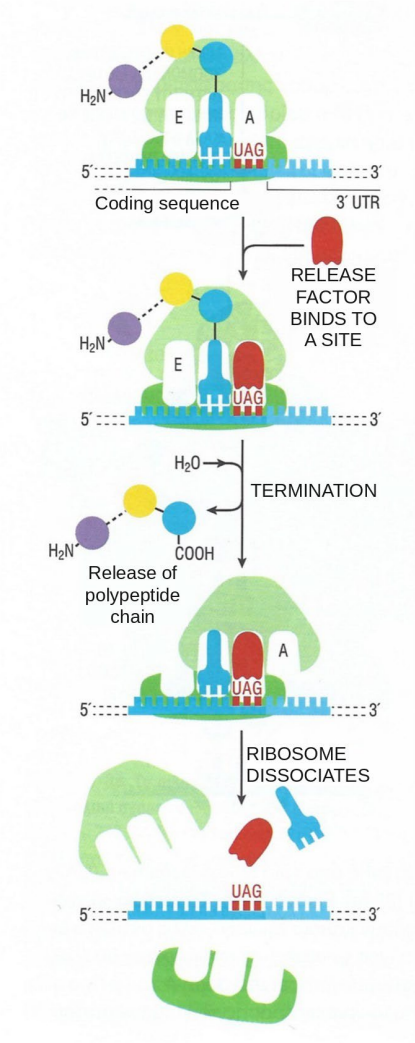
~5–6/s
amino acids, mammals

1/10³–10⁴
errors per amino acid

Polysomes: many ribosomes on one mRNA.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Ingolia et al., Cell 2011 · Zaher & Green, Cell 2009

Termination: no tRNA for stop codons



1. A stop codon (UAA, UAG, UGA) in the A site → a **release factor** binds (it mimics a tRNA).

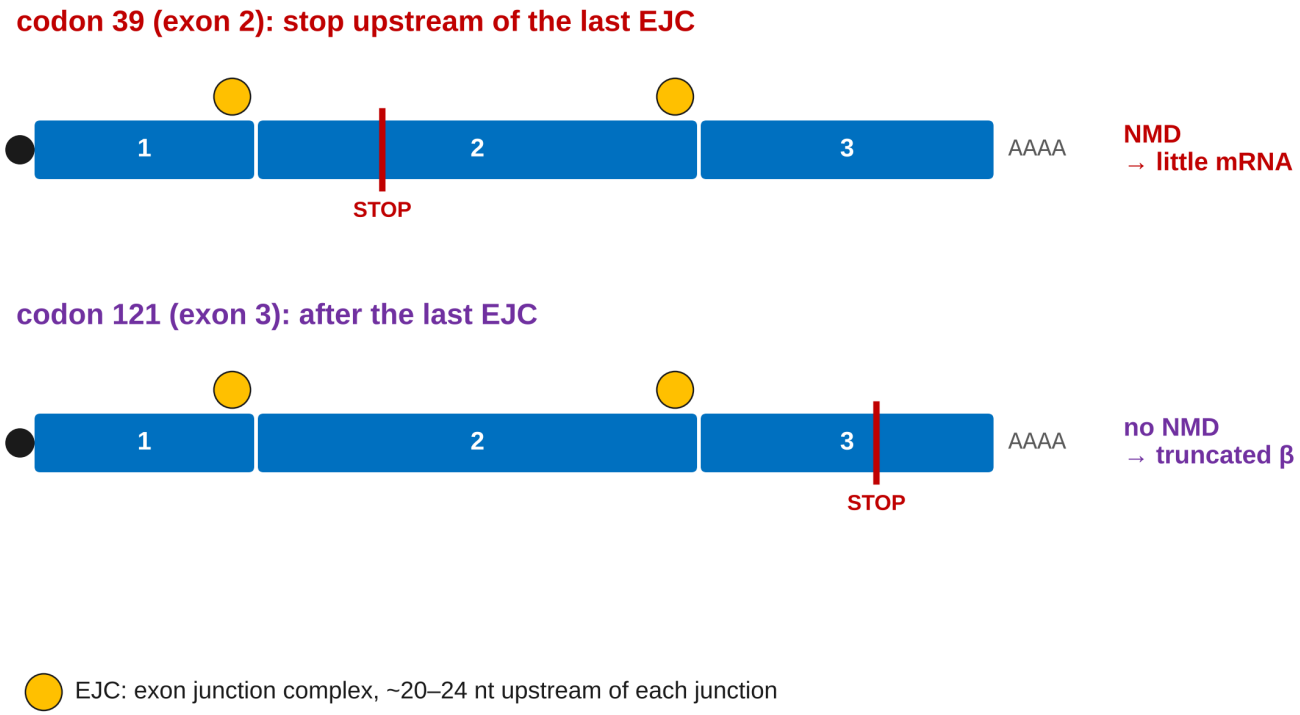
2. The ribosome adds **water** instead of an amino acid → the polypeptide is released.

3. The subunits separate and are **recycled** for a new start.

What happens when a stop codon appears too early? → next slide

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

Premature stop codons: nonsense-mediated decay



The rule

A stop >50–55 nt upstream of the last exon junction (with an EJC) → the mRNA is destroyed.

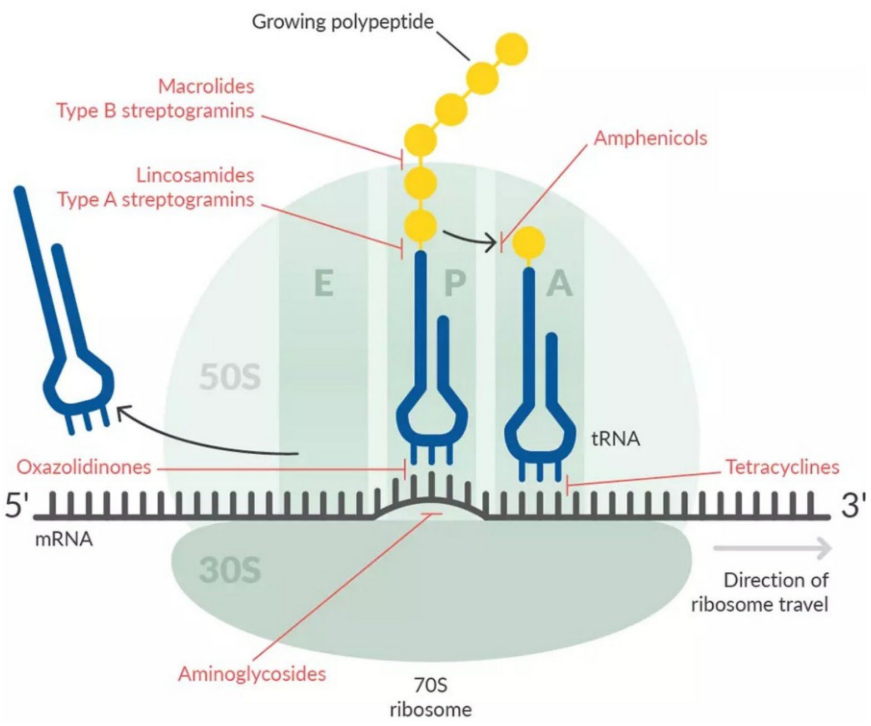
Two β-thalassemias

Codon 39 (exon 2): NMD → no β made → carriers are healthy (recessive). Stop in exon 3: no NMD → truncated, toxic β chains → dominant disease.

Stop-codon readthrough (ataluren): the EU license was not renewed (3/2025) — efficacy was not confirmed.

Nagy & Maquat, Trends Biochem Sci 1998 · Le Hir et al., EMBO J 2000 · Hall & Thein, Blood 1994 · EMA, Translarna 2025

Antibiotics that target the bacterial ribosome



Class	Target	Effect
Aminoglycosides	30S	codon misreading
Tetracyclines	30S	block the A site
Macrolides	50S	plug the exit tunnel
Chloramphenicol, linezolid	50S	inhibit peptidyl transferase

Why selective?

Bacterial 70S ≠ our 80S. But mitochondria have bacteria-like ribosomes.

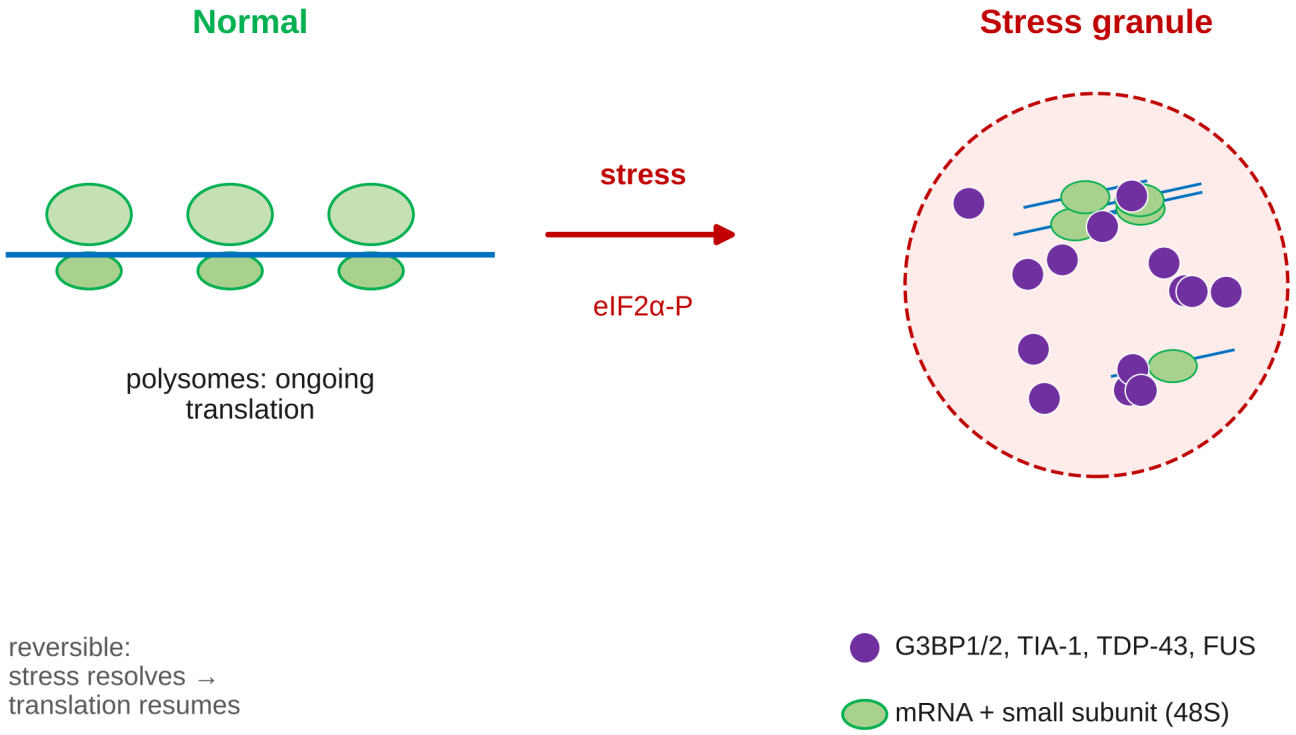
Toxins too

Diphtheria toxin inactivates eEF2 · ricin removes one base of 28S rRNA.

Question: if mitochondria resemble bacteria, who is at risk from aminoglycosides? →

Wilson, Nat Rev Microbiol 2014 · Van Ness et al., J Biol Chem 1980 · Endo et al., J Biol Chem 1987

When translation stops: stress granules



The signal

Under stress eIF2α is phosphorylated → translation initiation freezes (integrated stress response).

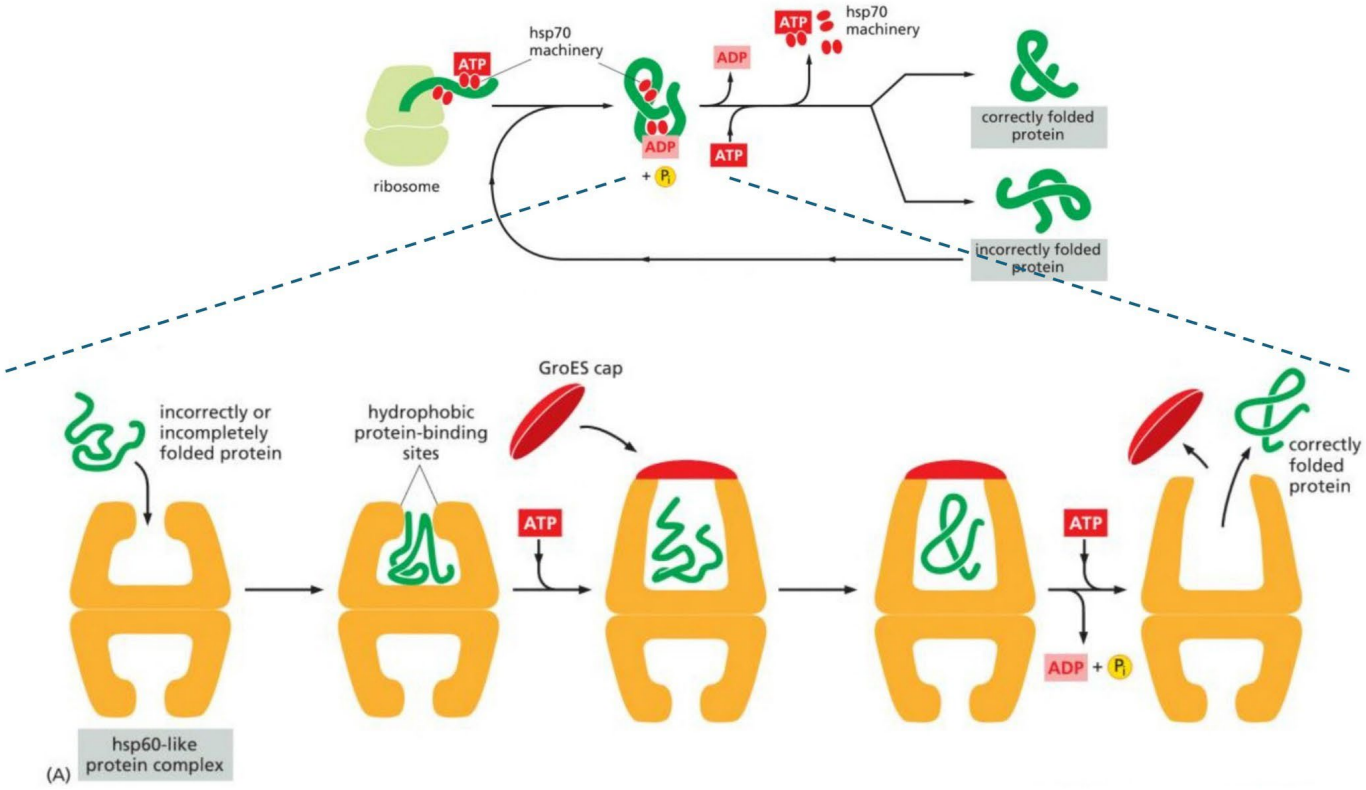
The granules

Stalled mRNAs condense with proteins such as G3BP1 into membraneless droplets (phase separation).

Disease and research: persistent stress granules with TDP-43 or FUS are linked to ALS and frontotemporal dementia. In our lab we study how they assemble and how they could be blocked.

Pakos-Zebrucka et al., EMBO Rep 2016 · Yang et al., Cell 2020 · Guillén-Boixet et al., Cell 2020 · Wolozin & Ivanov, Nat Rev Neurosci 2019

Folding: molecular chaperones



Hsp70

Binds hydrophobic stretches of the nascent chain while still on the ribosome (ATP-driven).

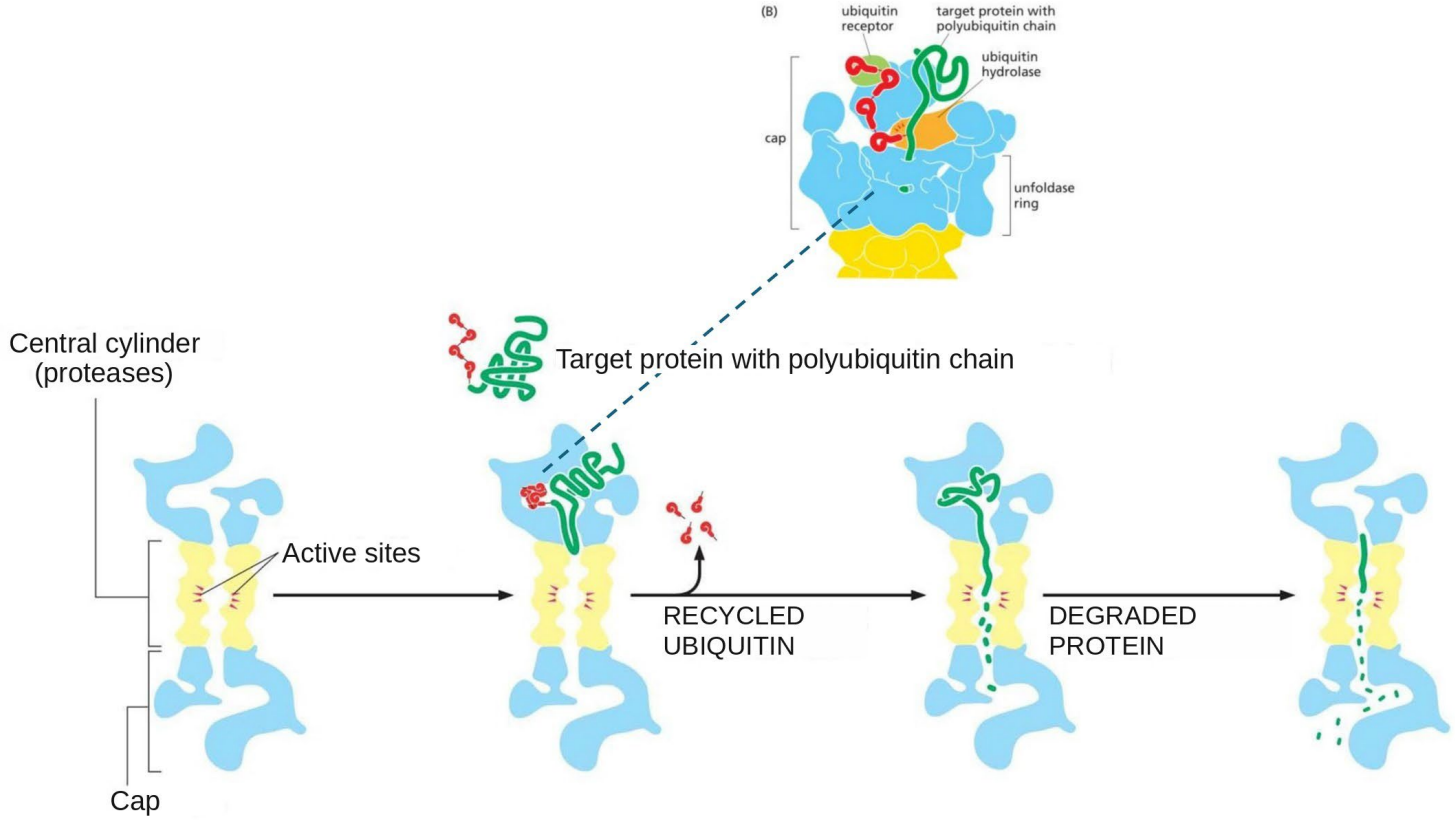
Hsp60 (chaperonin)

A chamber where the protein gets another chance to fold, away from other proteins.

Cystic fibrosis: F508del misfolds CFTR, which is destroyed before reaching the membrane. "Correctors" (elexacaftor/tezacaftor) + potentiator (ivacaftor): FDA 2019.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7 · Middleton et al., NEJM 2019

Ubiquitin and the proteasome: the tag for destruction

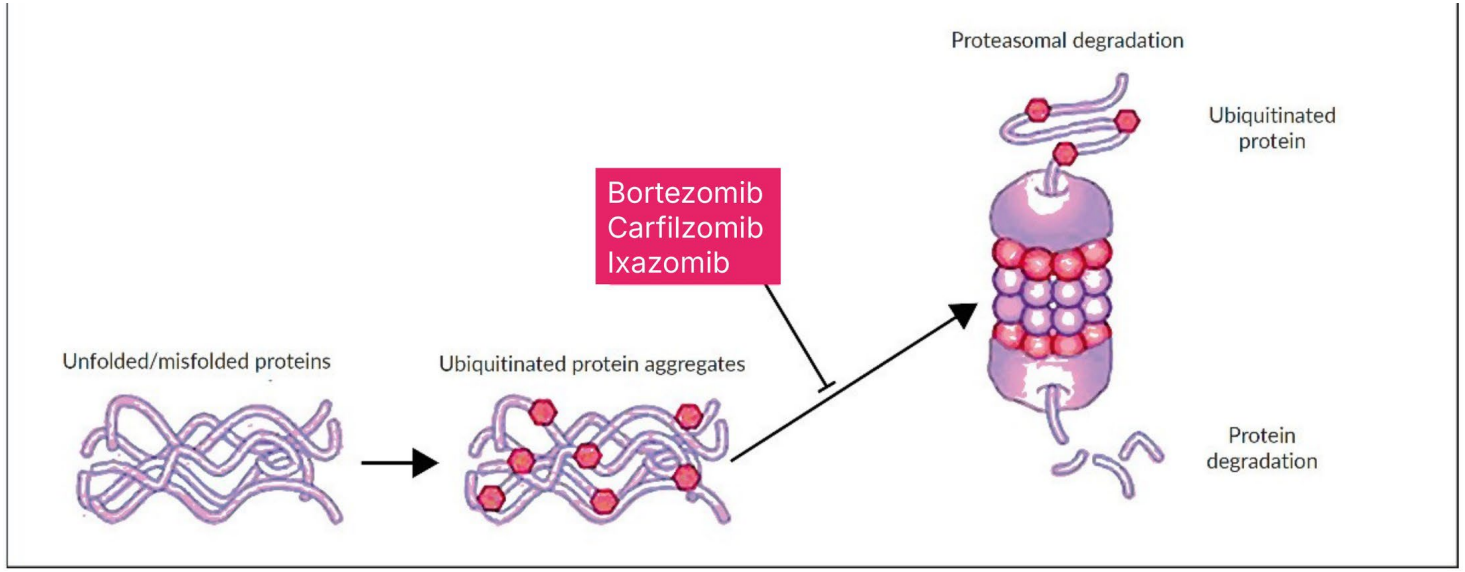


1. E1–E2–E3 enzymes attach a ubiquitin chain
2. The proteasome recognizes, unfolds (ATP) and chops
3. Ubiquitin is recycled

The ~600 E3 ligases decide which protein goes. Nobel Prize in Chemistry 2004: Ciechanover, Hershko, Rose.

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

Drugs that hijack protein degradation



Proteasome inhibitors

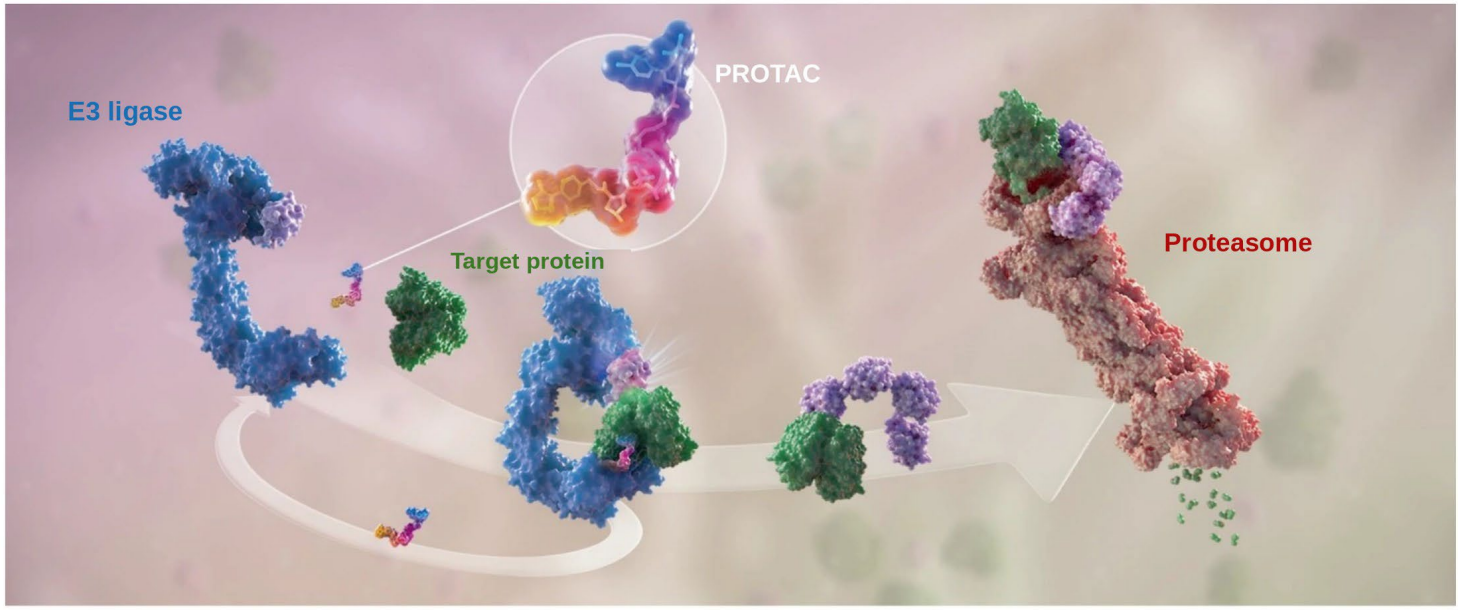
Bortezomib (2003), carfilzomib (2012), ixazomib (2015) for multiple myeloma: plasma cells make huge amounts of antibody and choke without the proteasome.

Molecular glues: from thalidomide to lenalidomide

They bind the E3 ligase cereblon (CRBN) and change its targets, sending the myeloma transcription factors IKZF1/IKZF3 to the proteasome. The same binding underlies thalidomide teratogenicity.

Ito et al., Science 2010 · Krönke et al., Science 2014 · Lu et al., Science 2014 · FDA

PROTACs: designed protein destruction



The idea

Two heads and a linker: one binds the target, the other an E3 ligase → ubiquitination → proteasome. Works even on proteins with no active site.

2026: the first approved PROTAC

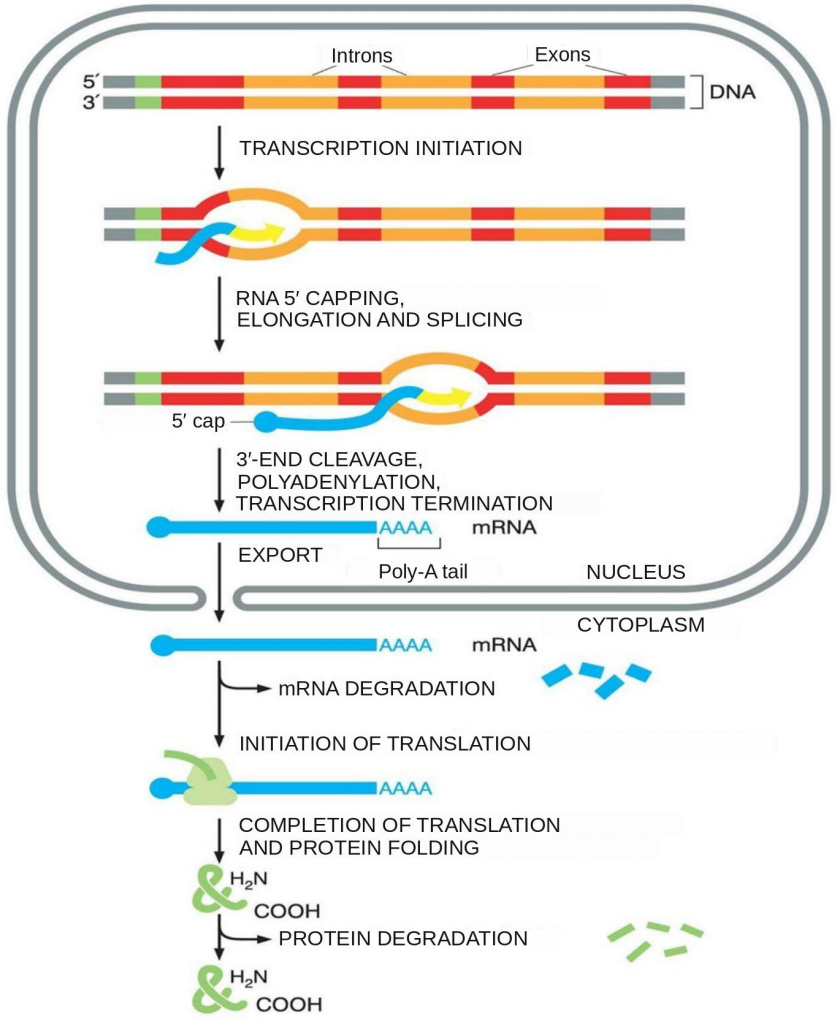
Vepdegestrant (FDA, 1 May 2026): ER+/HER2− breast cancer with an ESR1 mutation.
VERITAC-2: median PFS 5.0 vs 2.1 months (vs fulvestrant).

In clinical trials

Degraders of BTK (lymphomas, CLL; phase 3) and of the androgen receptor (prostate cancer).

Békés et al., Nat Rev Drug Discov 2022 · Arvinas/FDA, May 2026

Protein levels are controlled at many levels

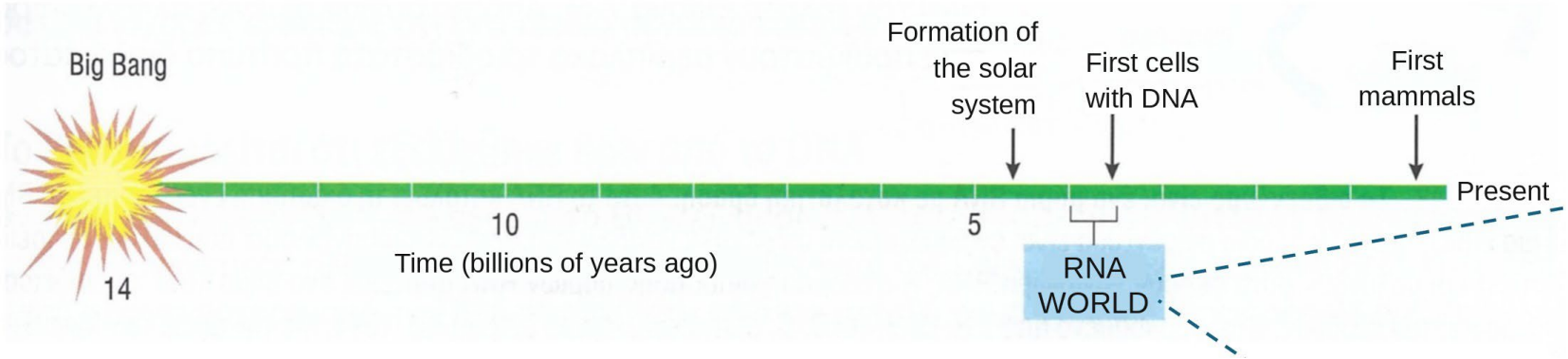


1. **Transcription**
   when and how often
2. **Processing**
   alternative splicing, poly(A)
3. **mRNA export and stability**
   NMD, miRNAs
4. **Translation**
   initiation, eIF2α, stress granules
5. **Protein**
   folding, modification, degradation

Next lecture: "Regulation of gene expression" (Ch. 8).

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

RNA as a catalyst and the "RNA world"

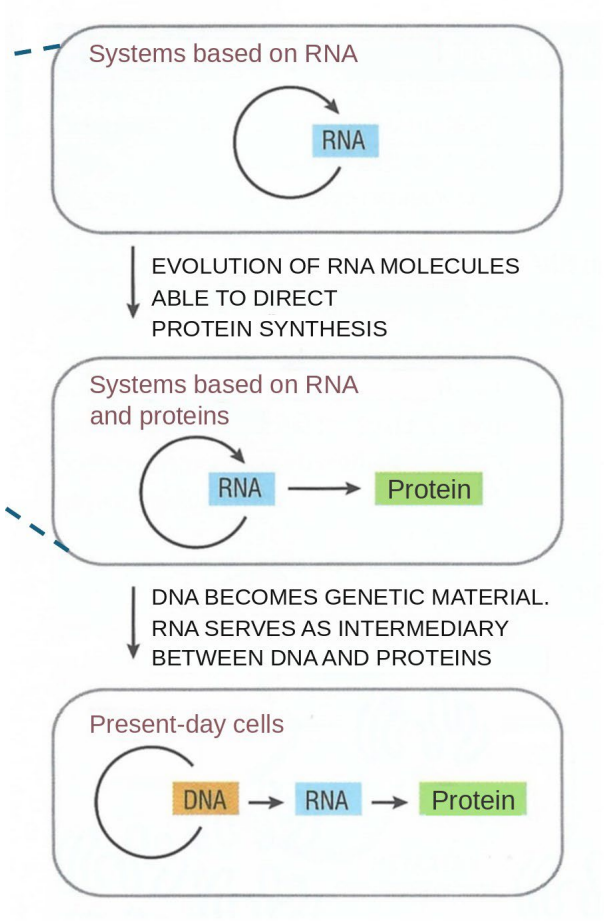


Ribozymes today

Ribosomal peptidyl transferase (rRNA) · U6 in the spliceosome · RNase P · self-splicing group I introns.

The hypothesis

Early on, RNA was both information and catalyst; later DNA (more stable) kept the information and proteins took over catalysis.



Nobel Prize in Chemistry 1989: Cech and Altman — "for their discovery of catalytic properties of RNA".

Alberts et al., Essential Cell Biology, 5th ed., Ch. 7

Summary of B2

1. The code is a triplet, degenerate, unambiguous and nearly universal; AUG sets the reading frame.
2. Aminoacyl-tRNA synthetases load the right amino acid; tRNA pairs with the codon (wobble at the 3rd position).
3. The ribosome (a ribozyme) reads mRNA 5′→3′: start at AUG, elongation in cycles, termination by a release factor.
4. NMD destroys mRNAs with premature stops; antibiotics and toxins target translation.
5. Chaperones, ubiquitin and the proteasome decide the fate of the protein — and are drug targets.

I. Comment briefly (1–3 sentences): true or false?

A For a given gene, either DNA strand can serve as the template.

B RNA polymerase II needs an RNA primer to start transcription.

C A synonymous mutation (e.g., GGC→GGT) can never cause disease.

D In the ribosome, the peptide bond is catalyzed by a protein of the large subunit.

E Each stop codon is recognized by a special tRNA that carries no amino acid.

III. Clinical case: β-thalassemia in a family

A child with β-thalassemia major is a compound heterozygote: IVS-I-1 G>A from the mother (the G of GU in intron 1 becomes A) and codon 39 CAG>TAG from the father. Both parents are healthy carriers.

a How does IVS-I-1 affect pre-mRNA maturation? Give two possible outcomes.

b What type of mutation is codon 39? Why do carrier cells contain very little mRNA from this allele?

c Why are the parents symptom-free, while a stop codon in exon 3 causes disease even in heterozygotes?

d Would a stop-codon readthrough drug help both mutations? Justify.

For study and revision

Animations & videos

HHMI BioInteractive: DNA Transcription (advanced)

HHMI BioInteractive: RNA Splicing

HHMI BioInteractive: Translation (advanced)

ITHANET (β-thalassemia, IthaGenes)

Reading

Alberts et al. Essential Cell Biology, 6th ed. (2023), Ch. 7 · Cooper & Adams, The Cell: A Molecular Approach, 9th ed. (2023) · Alberts et al. Molecular Biology of the Cell, 7th ed., Ch. 6

Busslinger, Moschonas & Flavell, Cell 1981 · Kattamis et al., Br J Haematol 1990

Pan et al., Nat Genet 2008 · Wang et al., Nature 2008 · Jaganathan et al., Cell 2019

Hua et al., Nature 2011 · Finkel et al., NEJM 2017 · Kim et al., NEJM 2019

Eriksson et al., Nature 2003 · Monaco et al., Genomics 1988

Nagy & Maquat, TIBS 1998 · Hall & Thein, Blood 1994

Wilson, Nat Rev Microbiol 2014 · Békés et al., Nat Rev Drug Discov 2022

Next lecture: Regulation of gene expression (Ch. 8) · Bioinformatics lab · Questions: idrosos@upatras.gr