Structure of eukaryotic genome, its replication and gene expression

Biology

Doc. RNDr. Jan Hošek, Ph.D.
hosek@mail.muni.cz

Department of Molecular Pharmacy
FaF MU

Life of eukaryots - movie

https://www.youtube.com/watch?v=7Hk9jct2ozY&ab_channel=WEHImovies

Structure of eukaryotic genome

Genom of eukaryotic organisms

Animal cells: nucleus and mitochondria

Plant cells: nucleus, mitochondria and chloroplasts

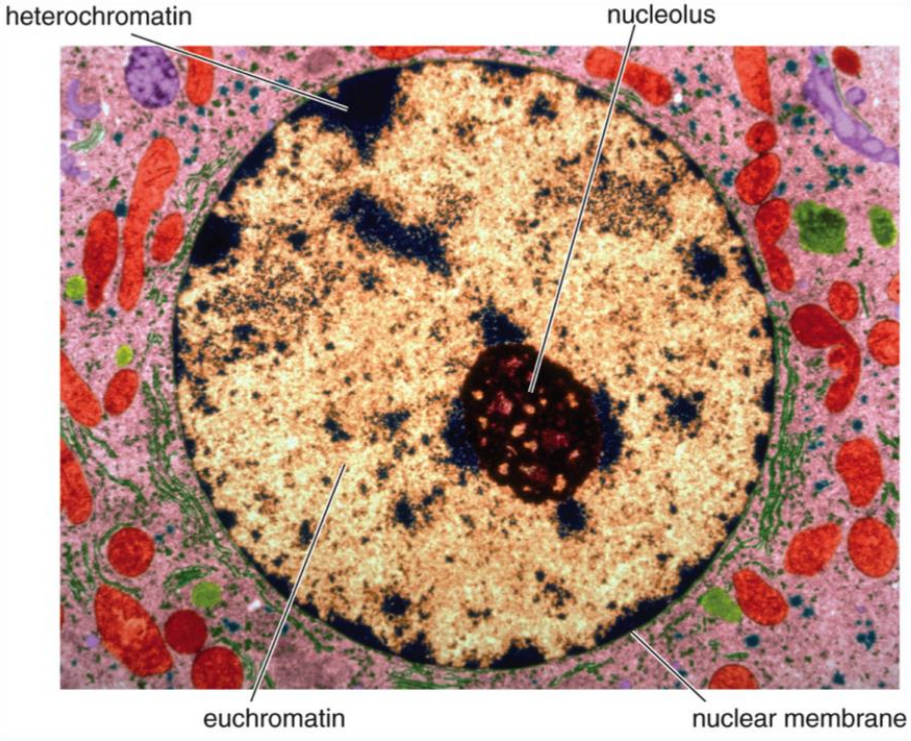
- **chromosomal (AKA nuclear) DNA (nDNA)**
- **mitochondrial DNA (mtDNA)**
- **chloroplast DNA (ctDNA)**
- **plasmids**

Chromatin

- Stainable material, which forms the nucleus of eukaryotic cells
- dsDNA, histones, nonhistones

According to the ability to be stained by basic dyes and degree of condensation we distinguish:

- **euchromatin** – weakly stainable, decondensed, "transcriptionally active"
- **heterochromatin** – strongly stainable, condensed, "transcriptionally inactive"



https://www.studyblue.com/notes/note/n/3-chromosomes/deck/4743713

Heterochromatin

Constitutive

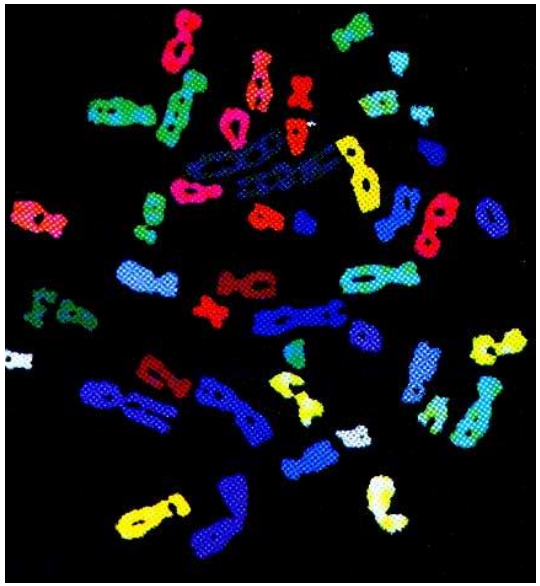
- Constantly in heterochromatin stage
- centromeres and telomeres
- One of X chromosome in women

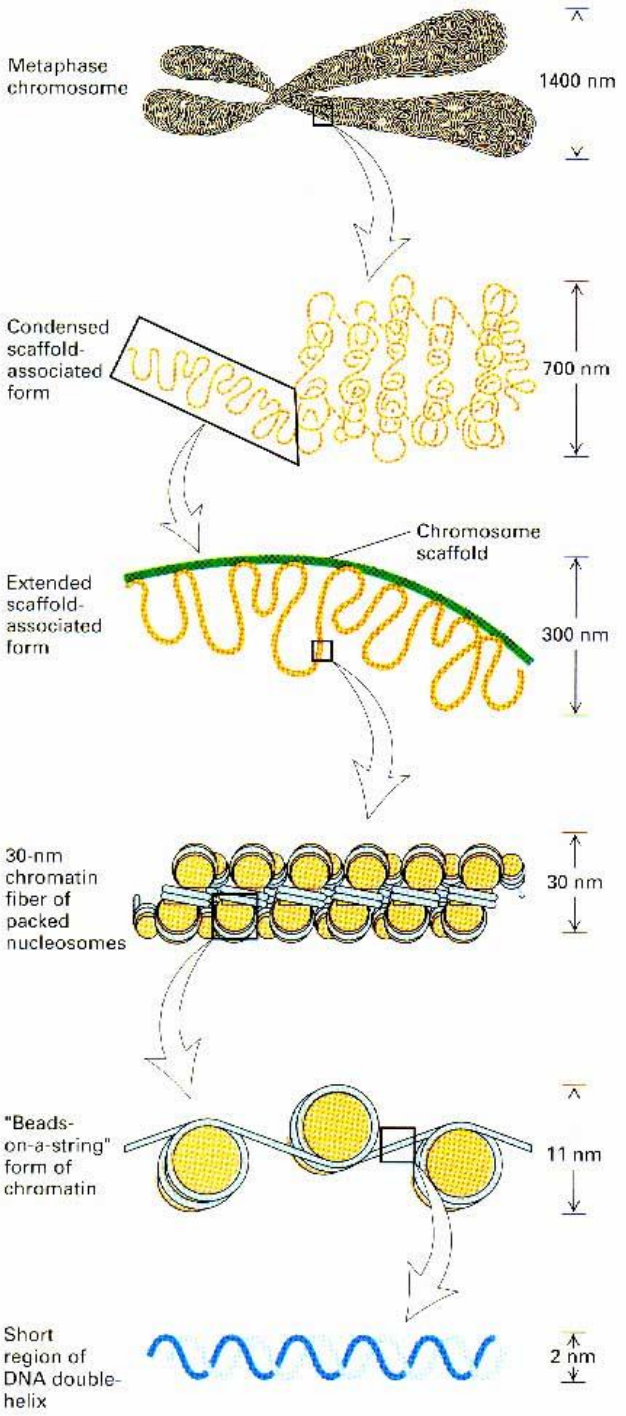
Facultative

- Switches between heterochromatin and euchromatin on the base of oncogenetic development of organism

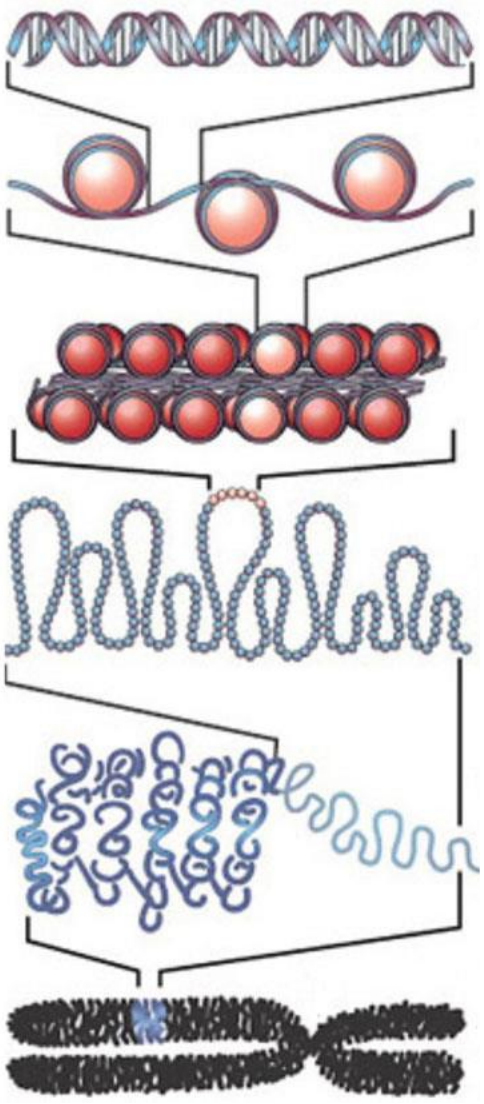
Chromatin condensation

1) Basic structure = interphase = decondensed 10-nm chromatin fiber (beads-on-a-string)
2) 30-nm chromatin fibre
3) Chromatin in mitotic phase = mitotic chromosomes





Chromatin condensation



„naked“ dsDNA

“beads-on-a-string“ form nucleosomes

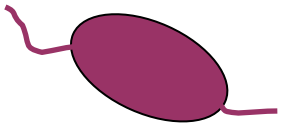
30 nm solenoid

relaxed form of chromosome

condensed region of chromosome

mitotic chromosome

Chromatin components



1) Histones

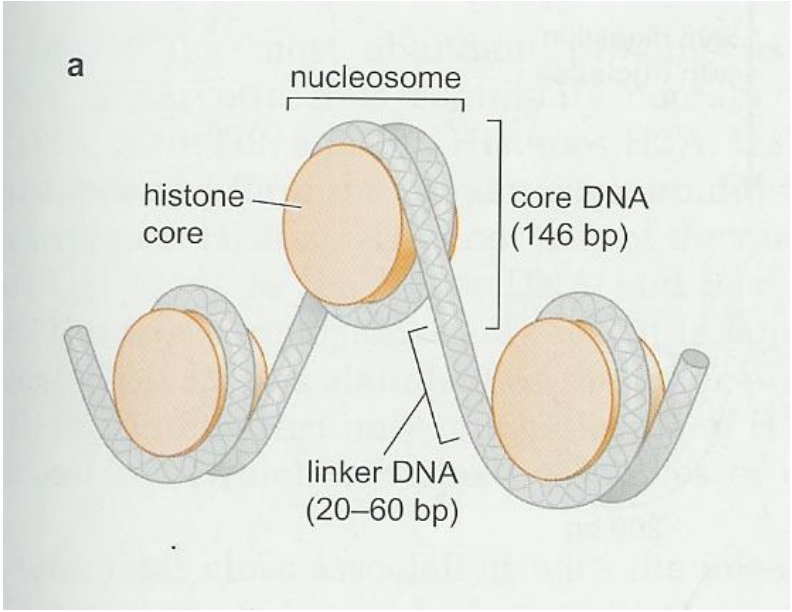
- Centre is globular, ends are flexible and filamentous
- High content of arginine and histidine
- 5 species = H1, H2A, H2B, H3 and H4

2) Nonhistones

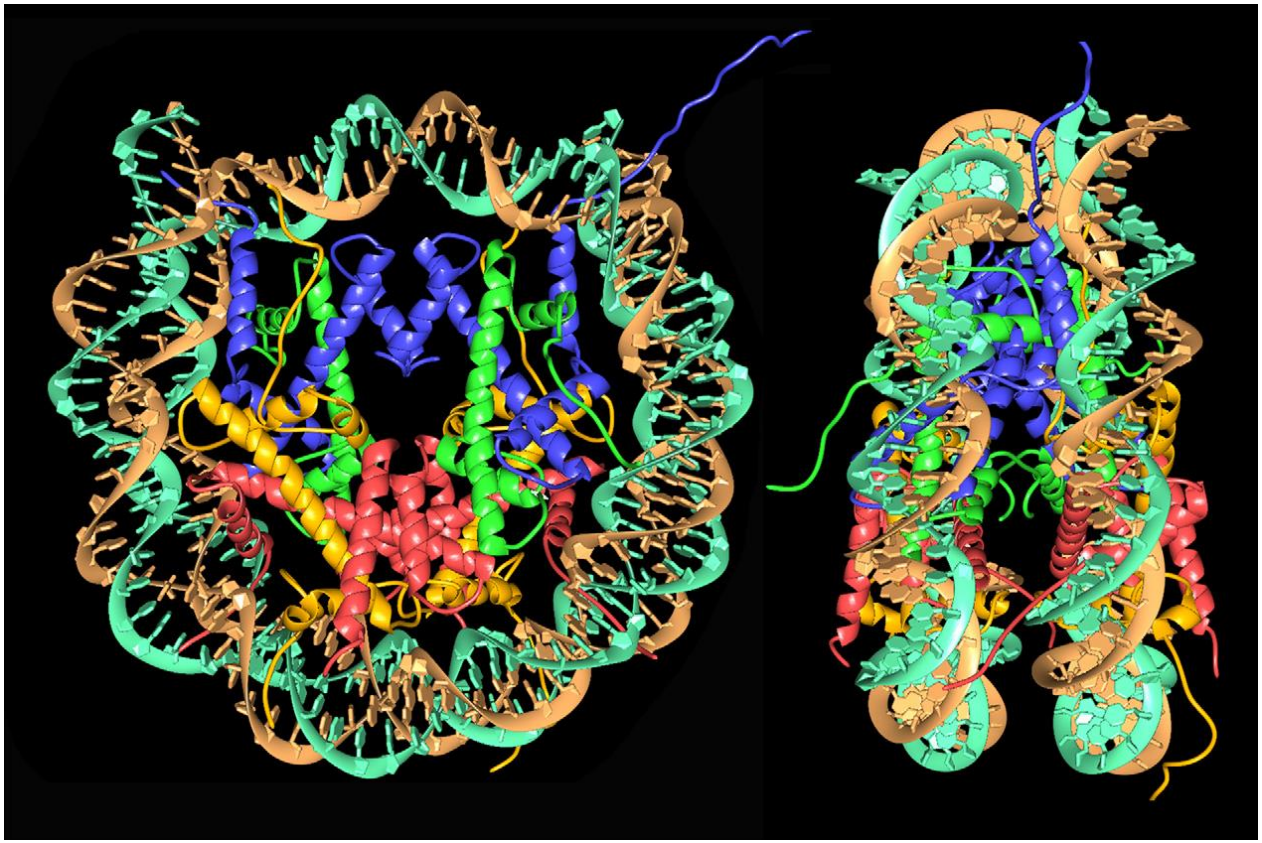
- RNA polymerase and other enzymes usable in transcription
- HMG1 and HMG2 – bind to unusual DNA structures
- HMG3 a HMG4 – bind to histone core especially in transcriptionally active regions

Nukleosome

- **The basic unit of chromatin**
- **octamer of histones $(H2A, H2B, H3, H4)_2$**
- **One molecule of histone H1**
- **DNA segment 200 bp long, which is wound about 2 times around the octamer of histones**



Nucleosome structure

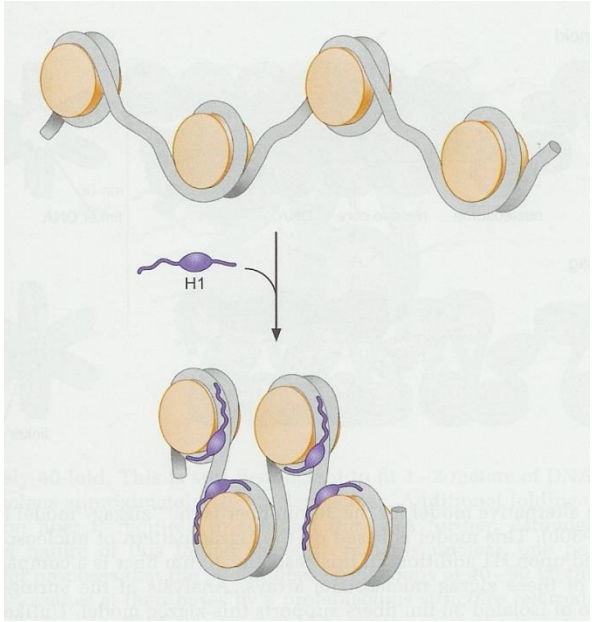


H2A – yellow, H2B red, H3 blue, H4 green

Nature 389: 251–260 (1997)

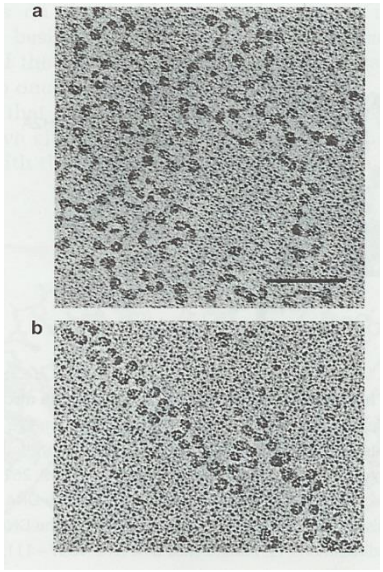
Nucleosome fibre

- 10nm chromatin fibre
- its individual items form nucleosome cores connected by long linear dsDNA
- visible by microscope



- H1

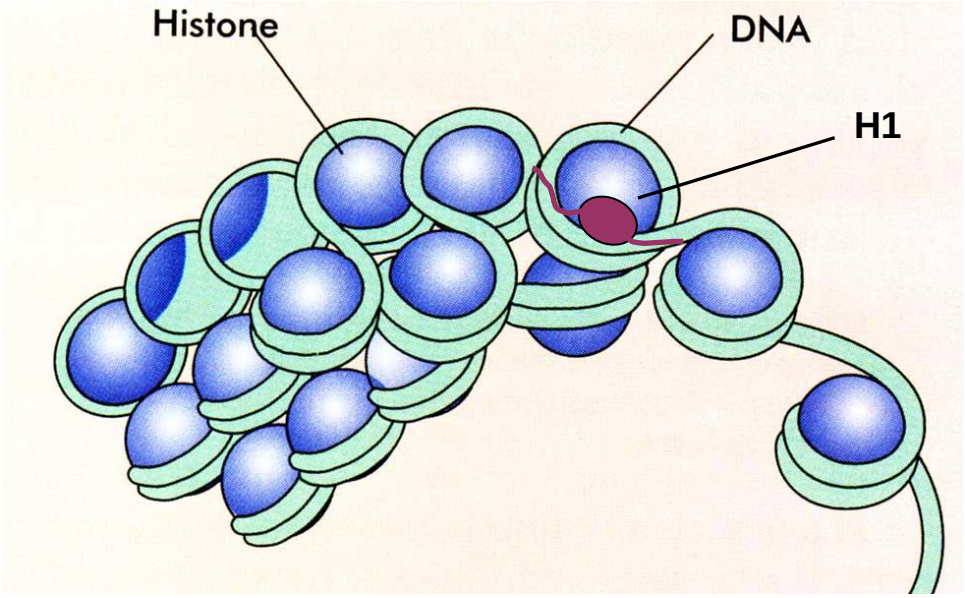
+ H1



30nm chromatin fibre

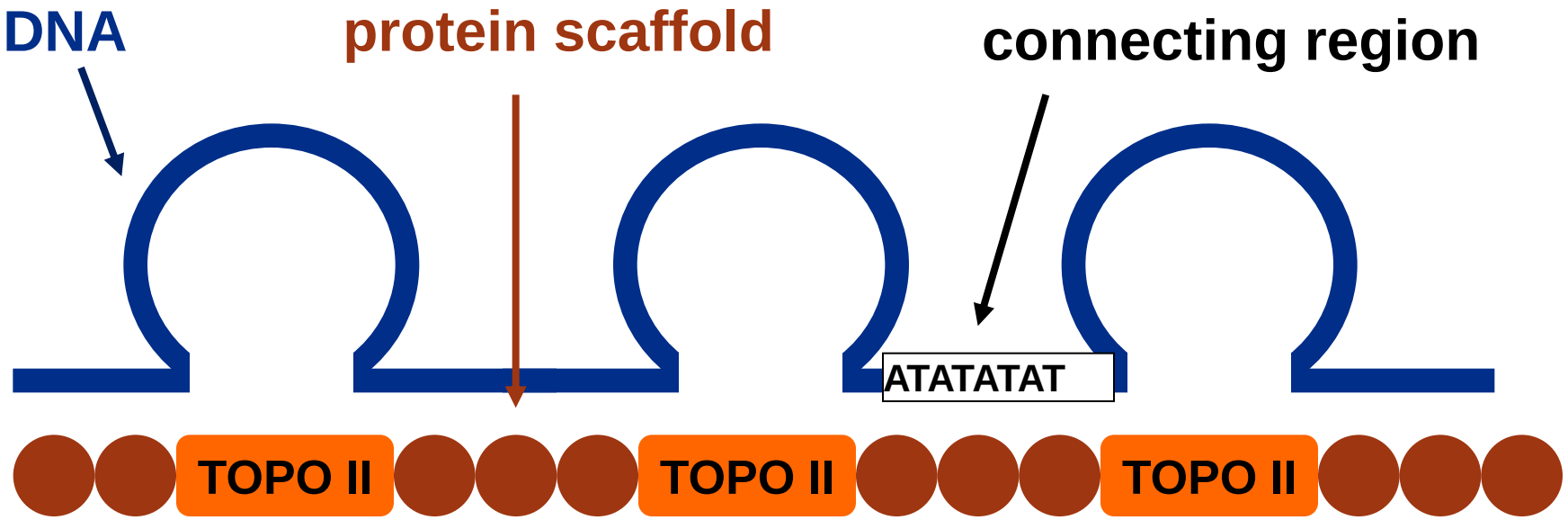
30nm chromatin fibre

- **It is created by the condensation of nucleosome fibre caused by histone H1**
- **It binds to protein scaffold (nonhistones, e.g. topoisomerase II)**



Chromatin domains

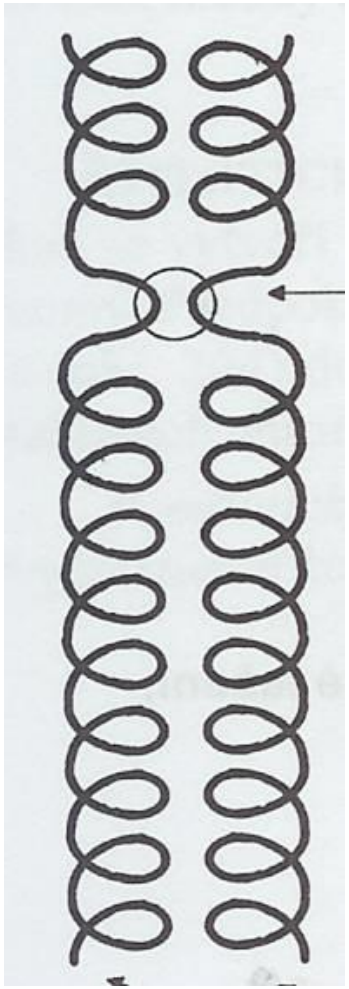
- loops of 30nm chromatin fibre attached to protein scaffold
- there is one molecule of topoisomerase II in base of each loop = change of topology during replication and transcription
- each domain has one *ori* locus

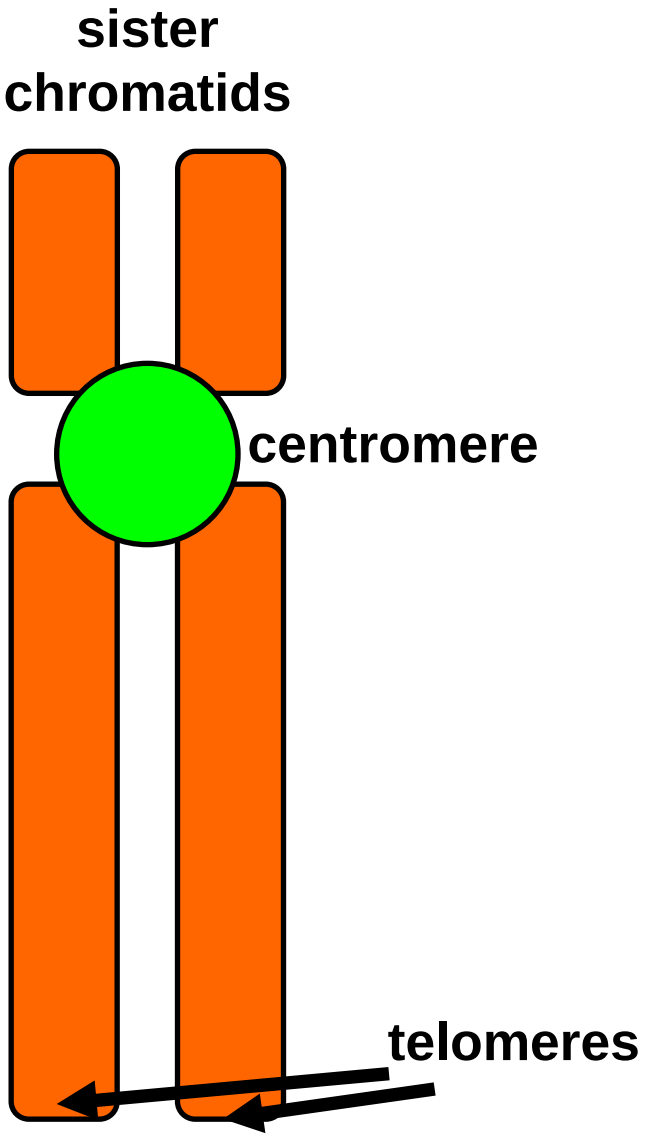


Mitotic chromosomes

- They originate by condensation of 30nm chromatin fibres
- They are formed during mitosis or meiosis
- Condensation of 30nm to 600-700nm chromatin fibres, which constitute the structure of chromosomes
- In chromosomes, the chromatin is in the stage of the highest condensation and is transcriptionally inactive

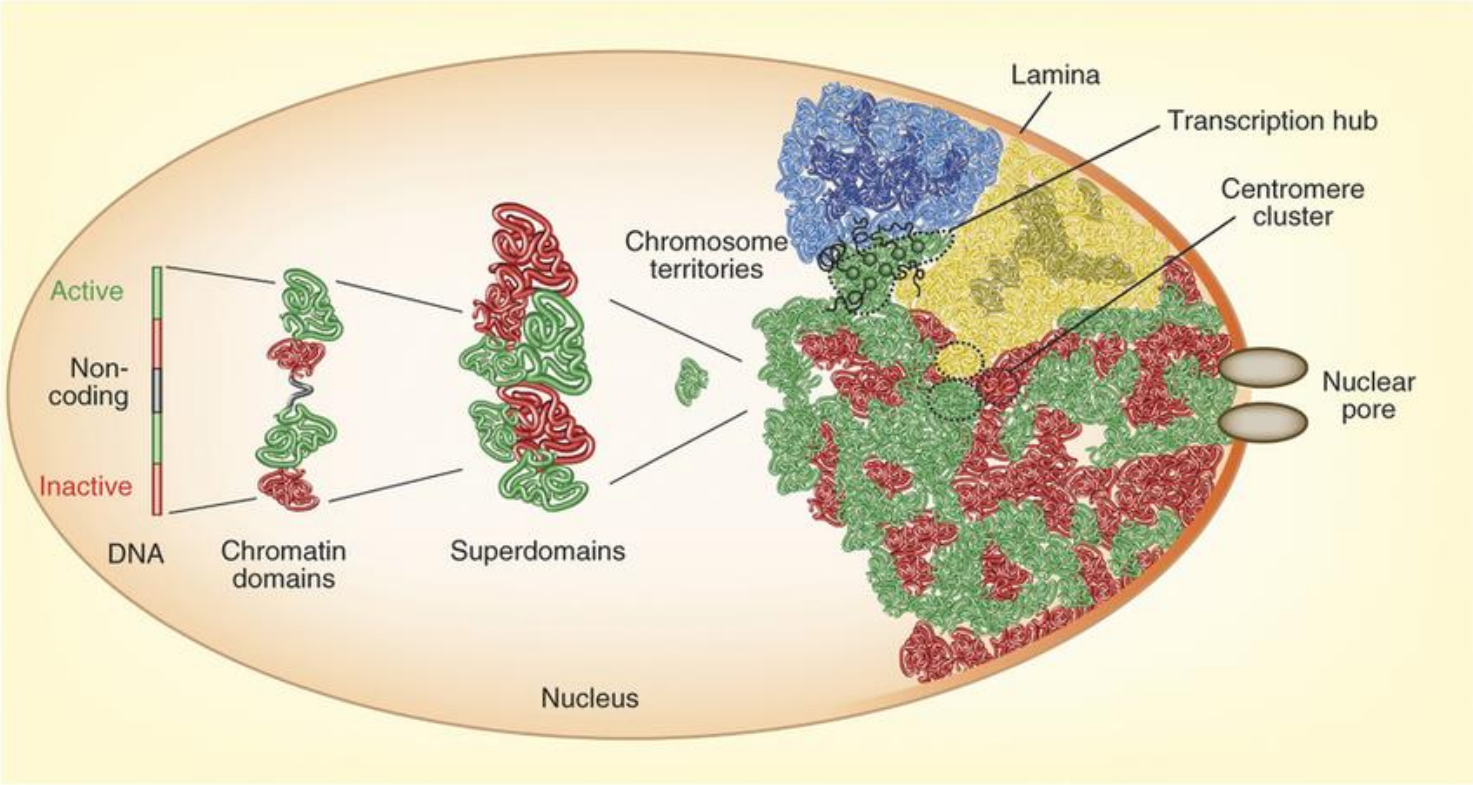
Mitotic chromosomes





Organisation of chromatine in nucleus

- Localisation of chromosomes is not accidental
- Crowding of region with similar function or activity

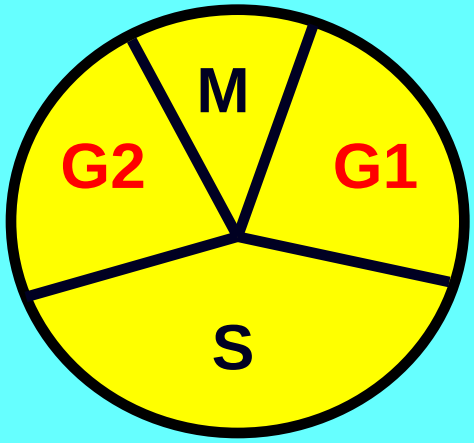


DOI: 10.1038/nsmb.2474

DNA replication in eukaryotes

Replication of eukaryotic genome

- replication of mitochondrial and chloroplat DNA
- replication of nuclear chromosomes
  - semiconservative and semidiscontinuous
  - initiation, elongation, and termination
  - only in S phase of the cellular cycle

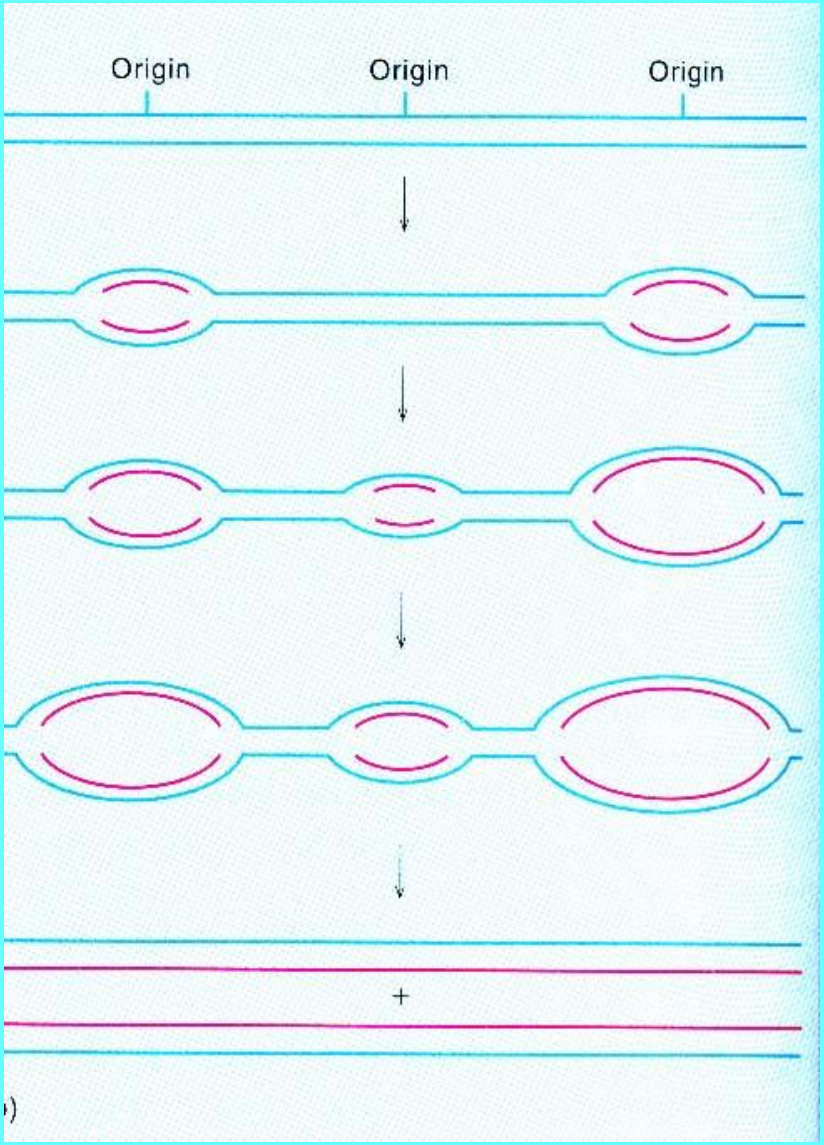


transcription,

translation,

metabolism

Replication of nuclear chromosomes

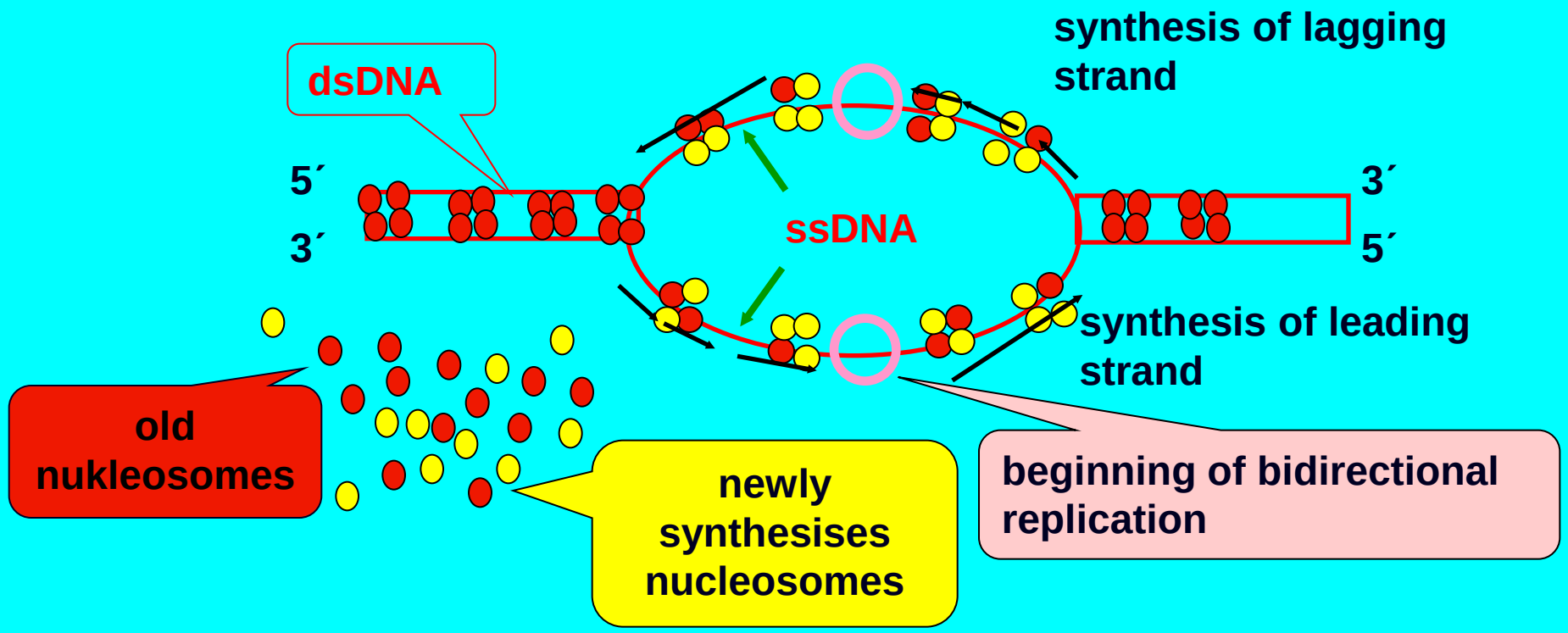


In contrast to prokaryotic cells the eukaryotic replication proceeds on several places in time

Chromosome is a couple of replicons, it has more *ori* sequences (mammals 30.000-50.000)

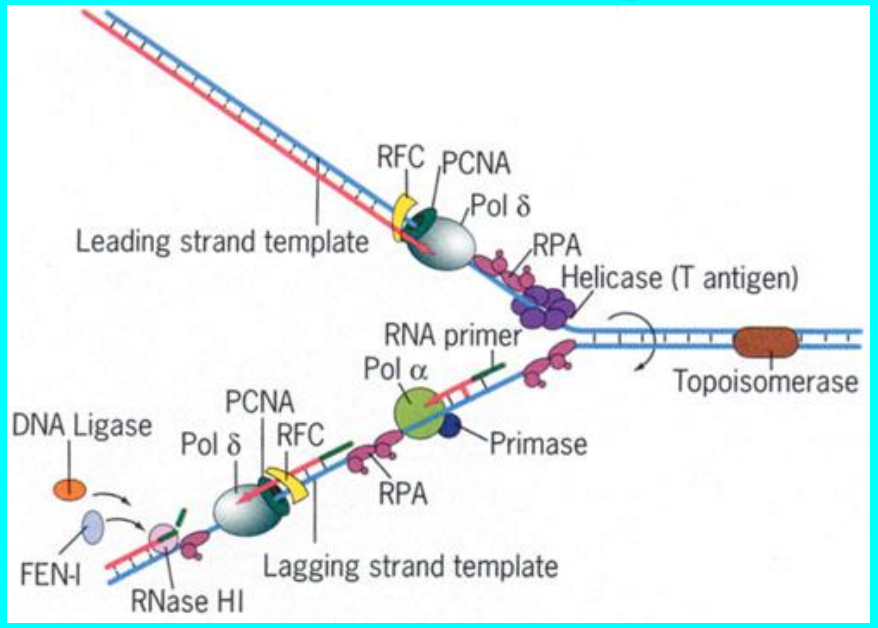
Euchromatin replicates earlier to heterochromatin

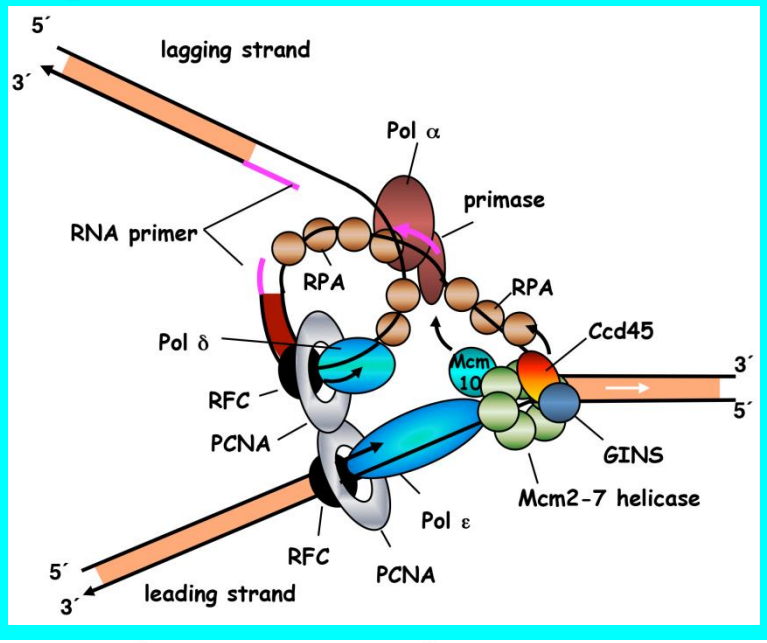
Process of replication



Eukaryotic DNA-polymerase: α, β, γ, δ and ε

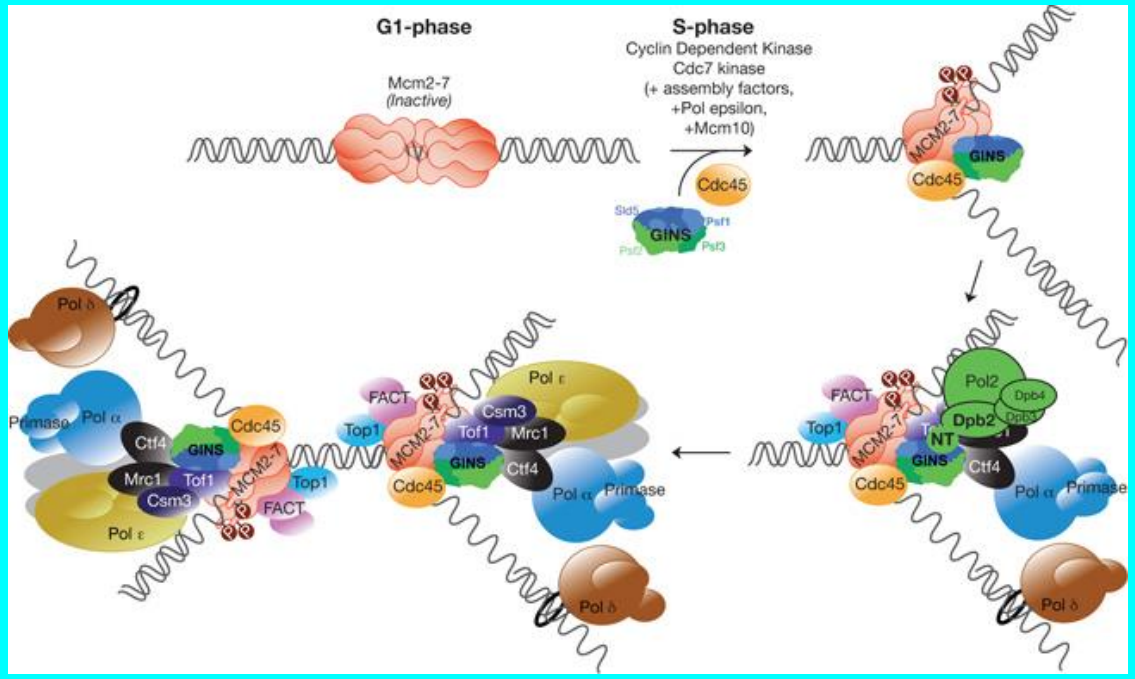
Eukaryotic replication fork





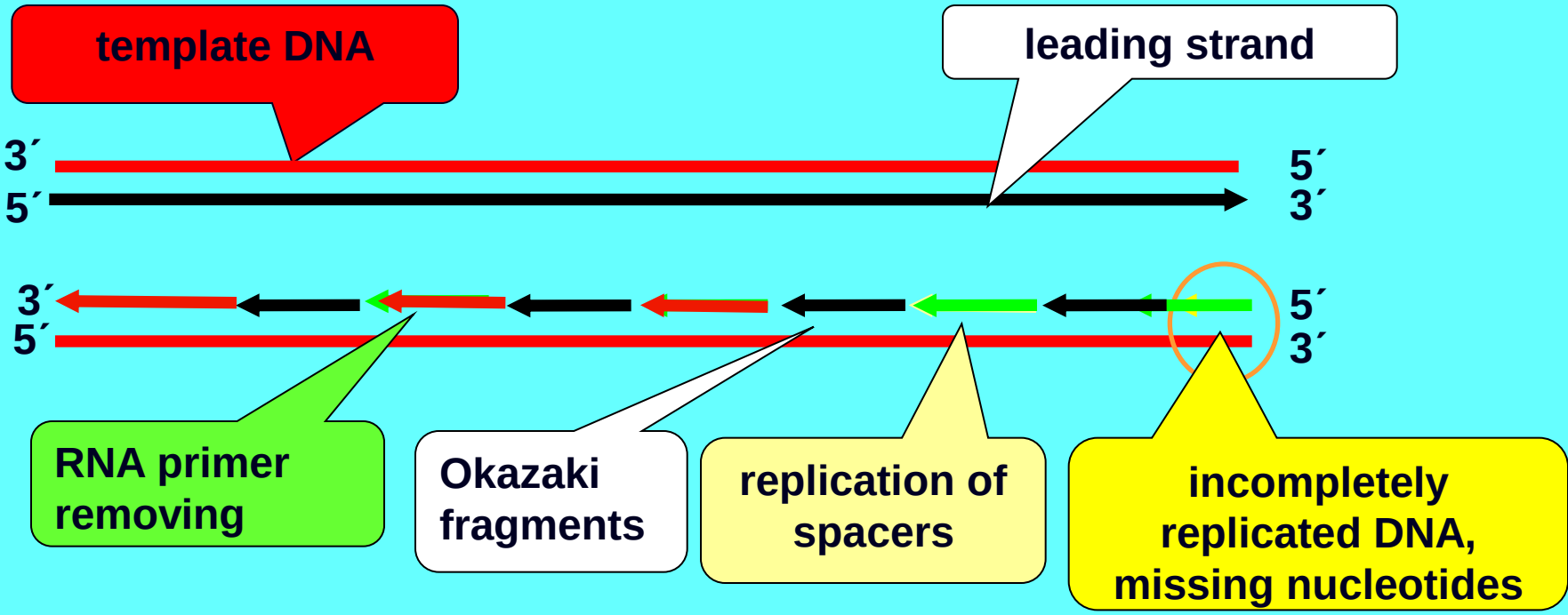
http://www.leibniz-fli.de/research/research-groups/grosse/

http://yxsj.baiduyy.com/



http://www.ppu.mrc.ac.uk/research/?pid=1012&sub1=research

Scheme of replication of linear molecules

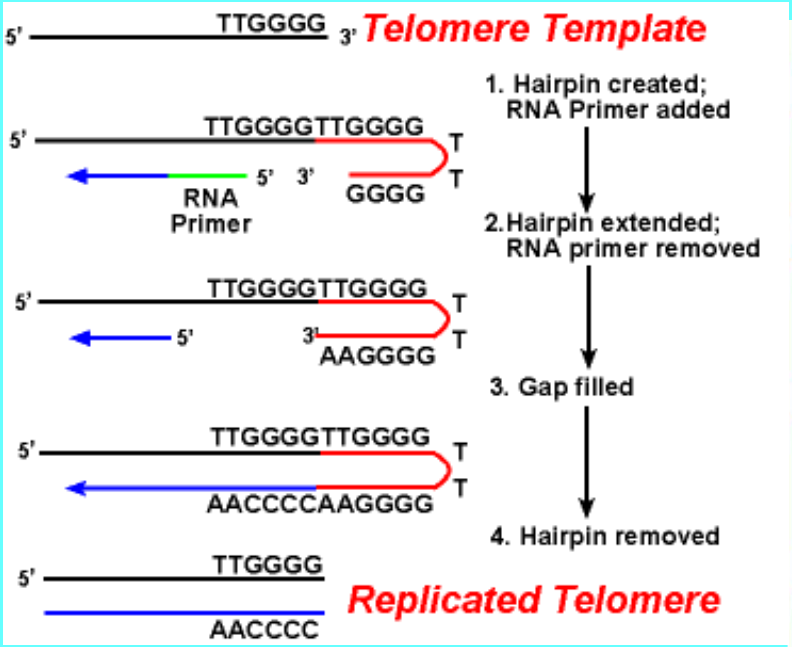


the end replication problem

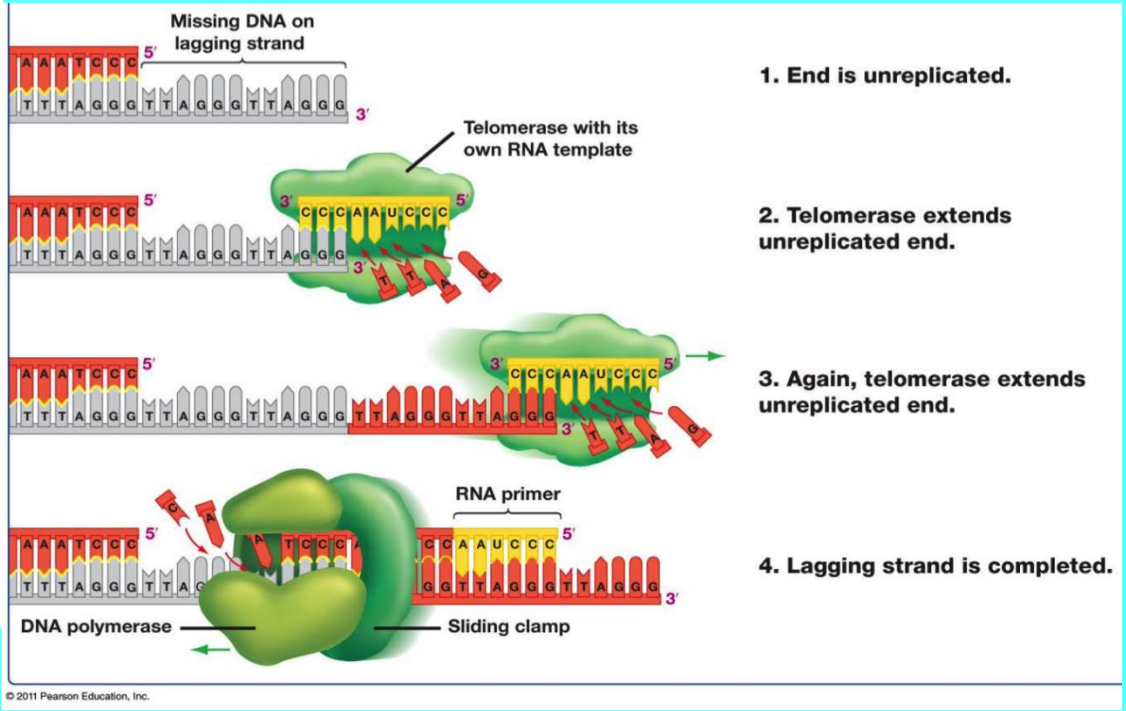
Telomerase = ribonukleoprotein - RNA acts as a template, protein has catalytic function

Filling of missing 3´-ends

- Telomerase elongates the 3´-end
- Formation of hairpin and RNA primer
- Replication of complementary strand and removing of hairpin



https://www.ndsu.edu/pubweb/~mcclean/plsc431/eukarychrom/eukaryo3.htm



http://masteringyourwaytomedschool.blogspot.cz/p/bio-1000-dna-shortening.html

Sequences of telomers

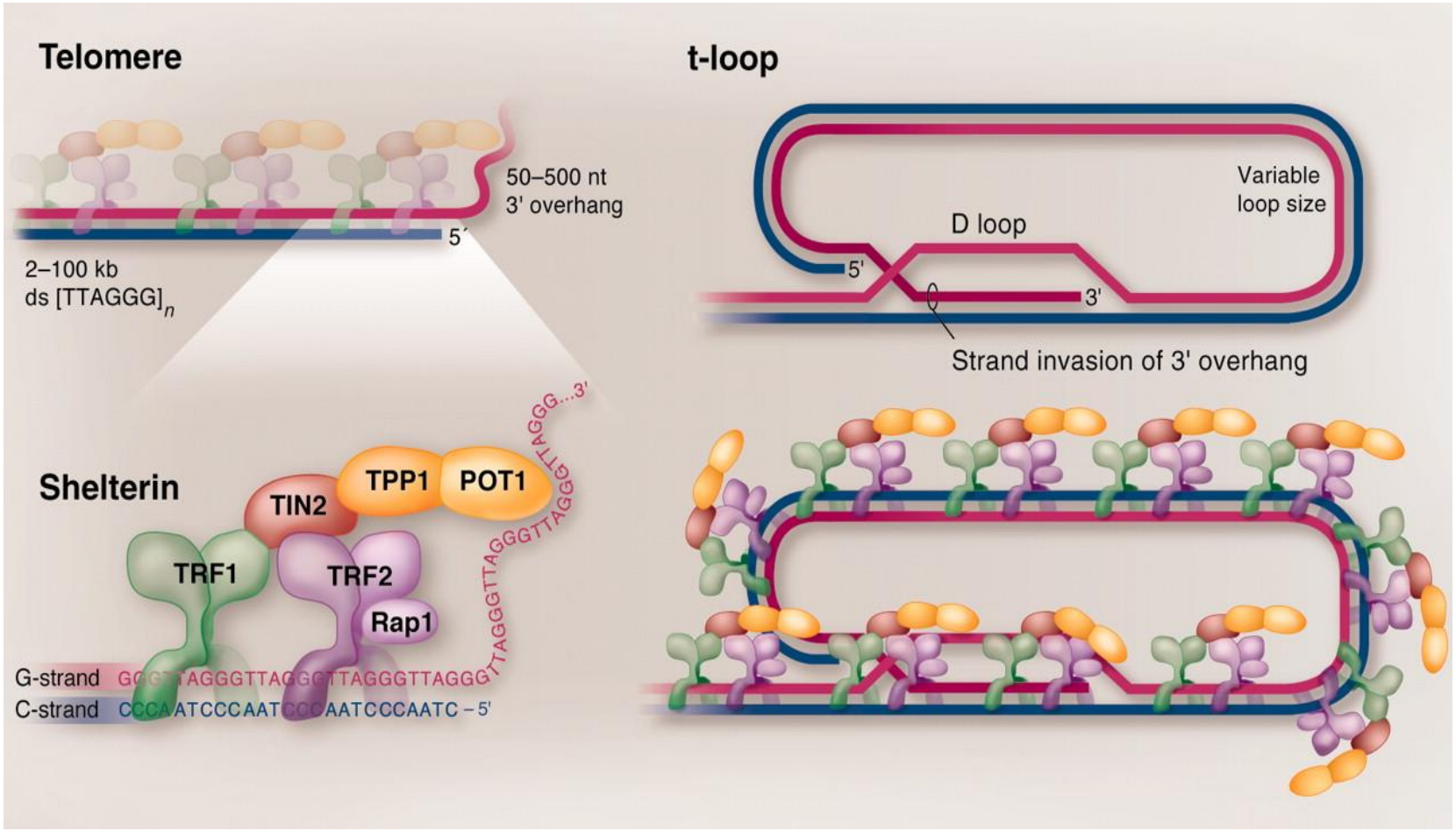
TABLE 11.5

Telomeric Repeat Sequences Within Selected Organisms

Group	Examples	Telomeric Repeat Sequence
Mammals	Humans	TTAGGG
Slime molds	*Physarum, Didymium*	TTAGGG
	Dictyostelium	$AG_{(1–8)}$
Filamentous fungi	*Neurospora*	TTAGGG
Budding yeast	*Saccharomyces cerevisiae*	$TG_{(1–3)}$
Ciliates	*Tetrahymena*	TTGGGG
	Paramecium	TTGGG(T/G)
	Euplotes	TTTTGGGG
Higher plants	*Arabidopsis*	TTTAGGG

http://reasonandscience.heavenforum.org/t2263-the-telomerase-enzyme

Fig. 2 Mammalian telomeres



T. de Lange Science 326, 948-952 (2009)

Transcription of eukaryotic genome

Transcription of eukaryotic genome

- **Primary transcripts**
  - **heterogeneous nuclear RNA (hnRNA) = pre-mRNA forming in nucleus**
  - **precursor ribosomal RNA (pre-rRNA)**
  - **precursor transfer RNA (pre-tRNA)**
  - **5S-rRNA**
  - **small RNA (snRNA, snoRNA, scRNA)**

- **Eukaryotic DNA-dependent RNA-polymerase**

  - RNA-polymerase I, II, III

- **Transkription factors**

Transcription unit of prokaryotes

polycistronic character

***trp* operon in *Escherichia coli* – 5 genes**

***trp* mRNA**

origins of translation

E

D

C

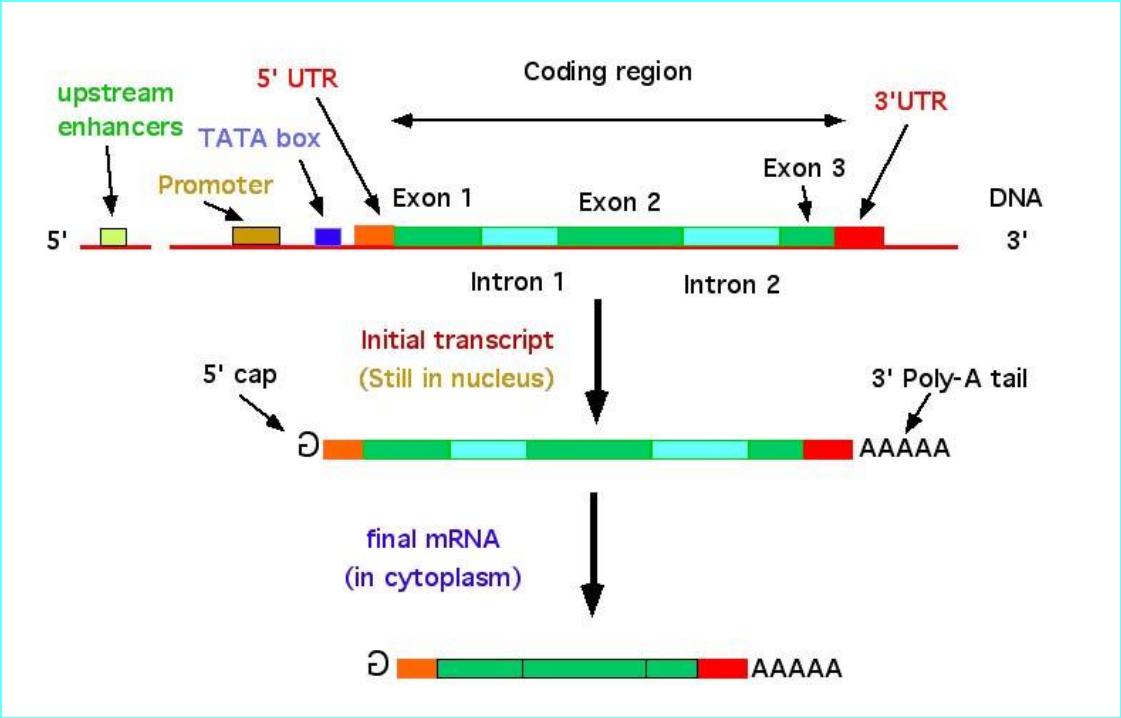
B

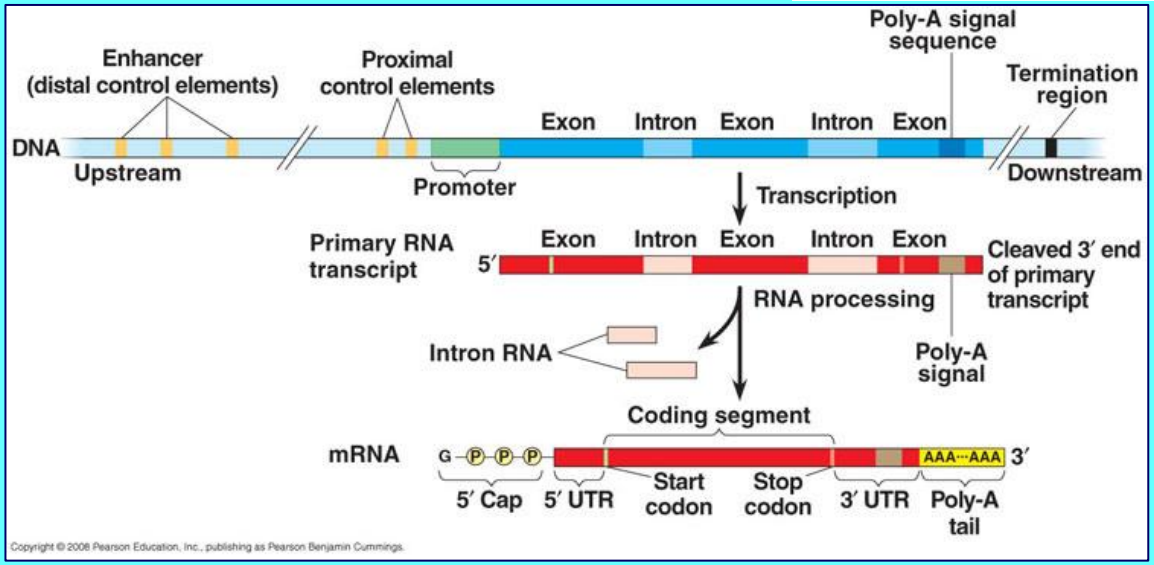
A

5 proteins for synthesis of Trp

Transcription unit of eukaryotes

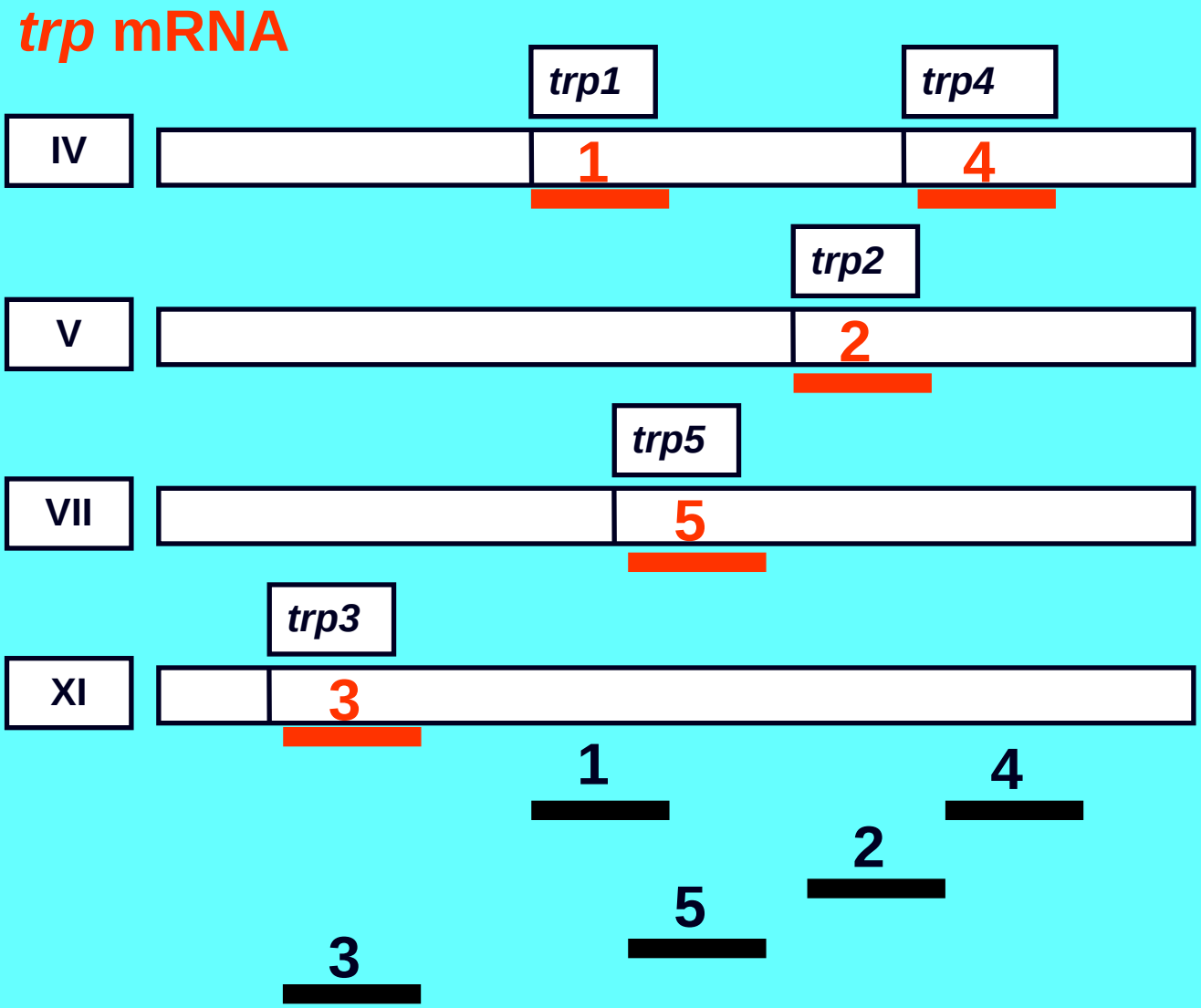
- Monocistroni character
- Contains:
  - Promoter
  - Leading sequence (5´-UTR)
  - Polyadenilation signal
  - Terminator





http://nitro.biosci.arizona.edu/courses/EEB600A-2003/lectures/lecture24/lecture24.html

The transcription unit of S. cerevisiae



Transcription unit for synthesis of Trp in *Saccharomyces cerevisiae*

= togehter 5 genes located on 4 chromosomes

Transcription factors

- **Regulatory elements necessary for transcription initiation**
- **Usually initiate transcription, rarely inhibit it**
- **Their different combination bind to the promoter, then the RNA polymerase bind to DNA strand**

Types of transcription factors

- **general TF**
  - **present in all and most types of cells**
  - **necessary to transcription initiation**
  - **basal – low activity, minimal cell requirements**
  - **constitutive – increase the basal activity according to cell type; basal cell requirements**
- **special TF**
  - **only in cells of specific tissues and in a certain time**
  - **applied in inducible transcription**

Transcription of hnRNA

1) Separation of transcription and translation
2) hnRNA is capped by the cap, and methylated (binding to ribosome)
3) In the 3´- region (after STOP codon) the sequence AAUAAA is present, in this location the hnRNA is digested
4) At 3´end is polyadenylated (stabilisation in cytoplasm)
5) After removing introns and joining exons it is transformed to mRNA

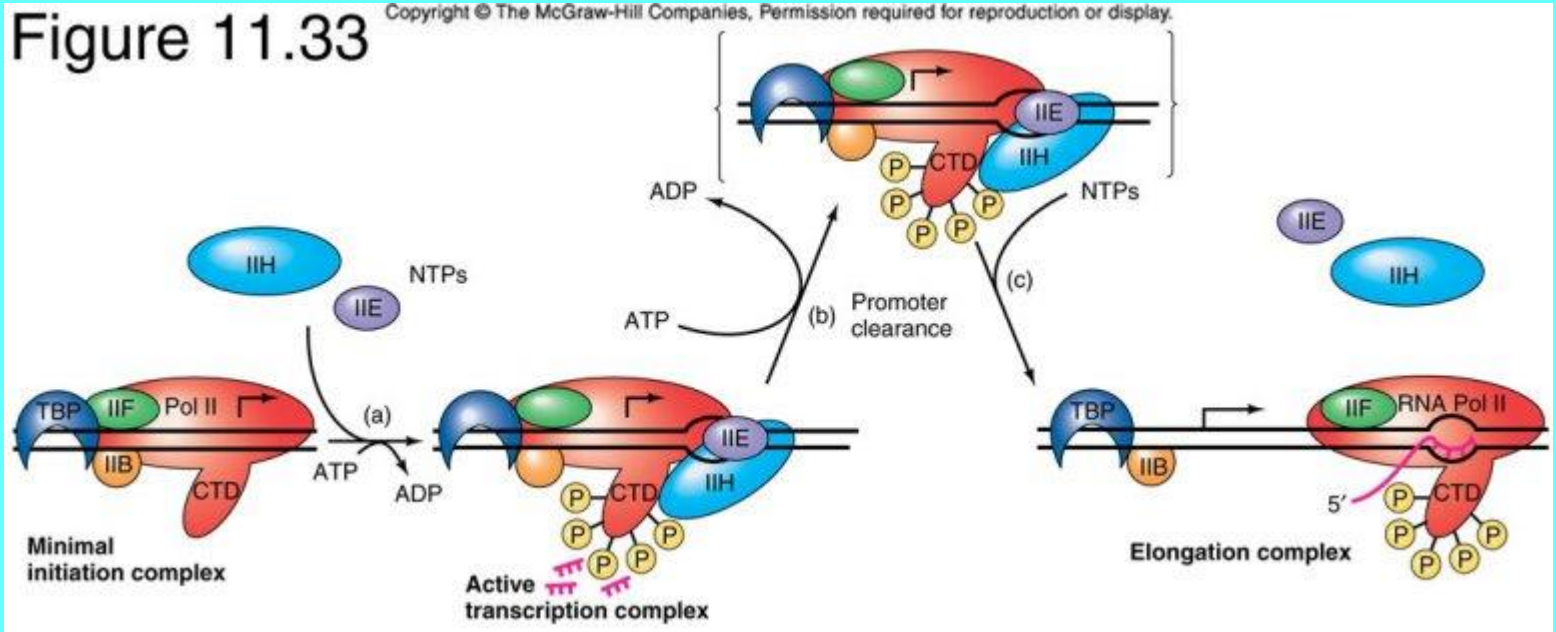
Initiation of transcription

1) Binding of transcription factors on TATA box and others regulation sequences = preinitiation complex

2) Binding of RNAP II on preinitiation complex = closed initiation complex

3) Phosphorylation of CTD domain of RNAP II by trancription factor TFIIH (halicase and kinase activities) → RNAP II activation and unwinding of dsDNA = open initiation complex

4) Disociation of RNAP II from TFs (except TFIIF) and start of RNA synthesis



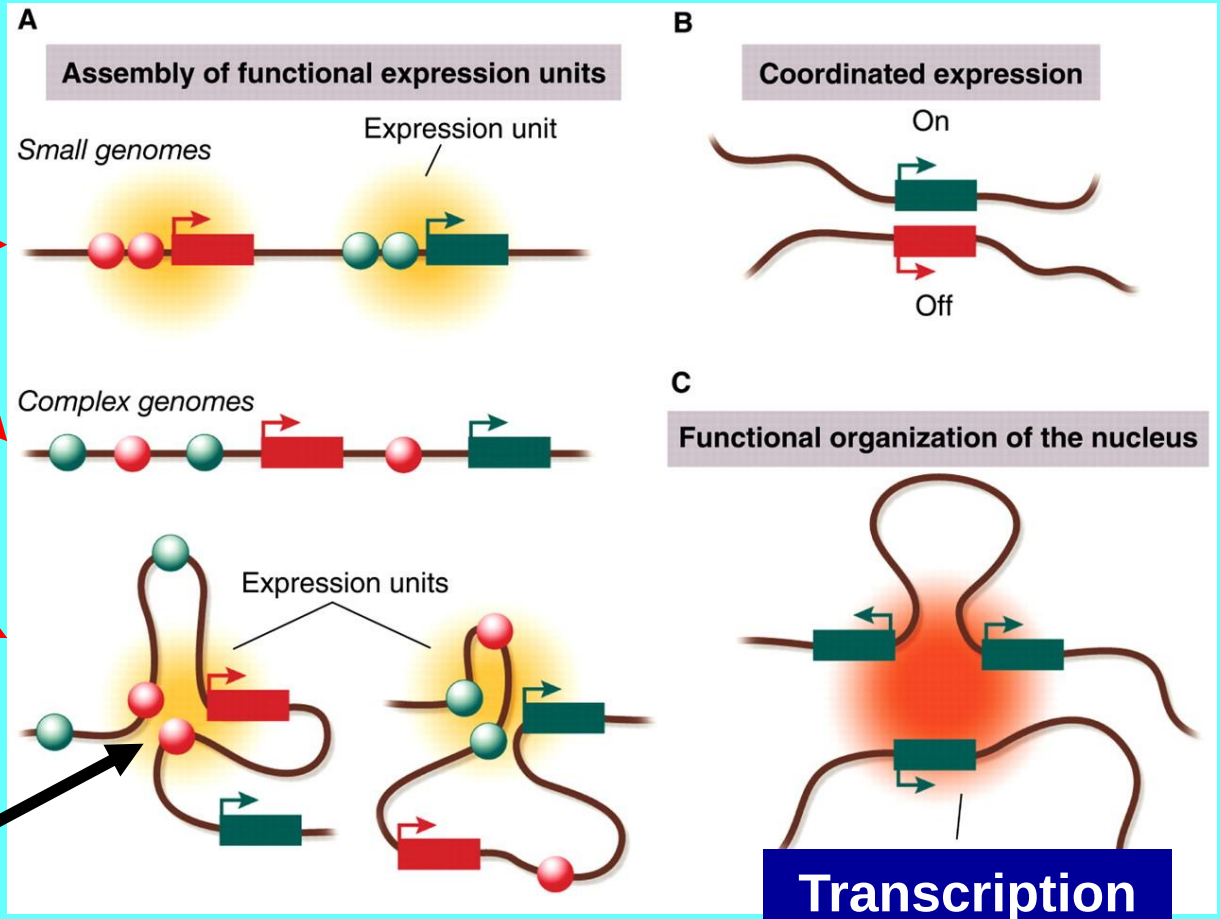
Eukaryotic transcription - video

https://www.youtube.com/watch?v=_Zyb8bpGMR0&ab_channel=ArmanHossain

Spatial assemblies of transcription

Linear transcription units

Spatial transcription units



Promoter assembled from 2 promoters?

Transcription factory

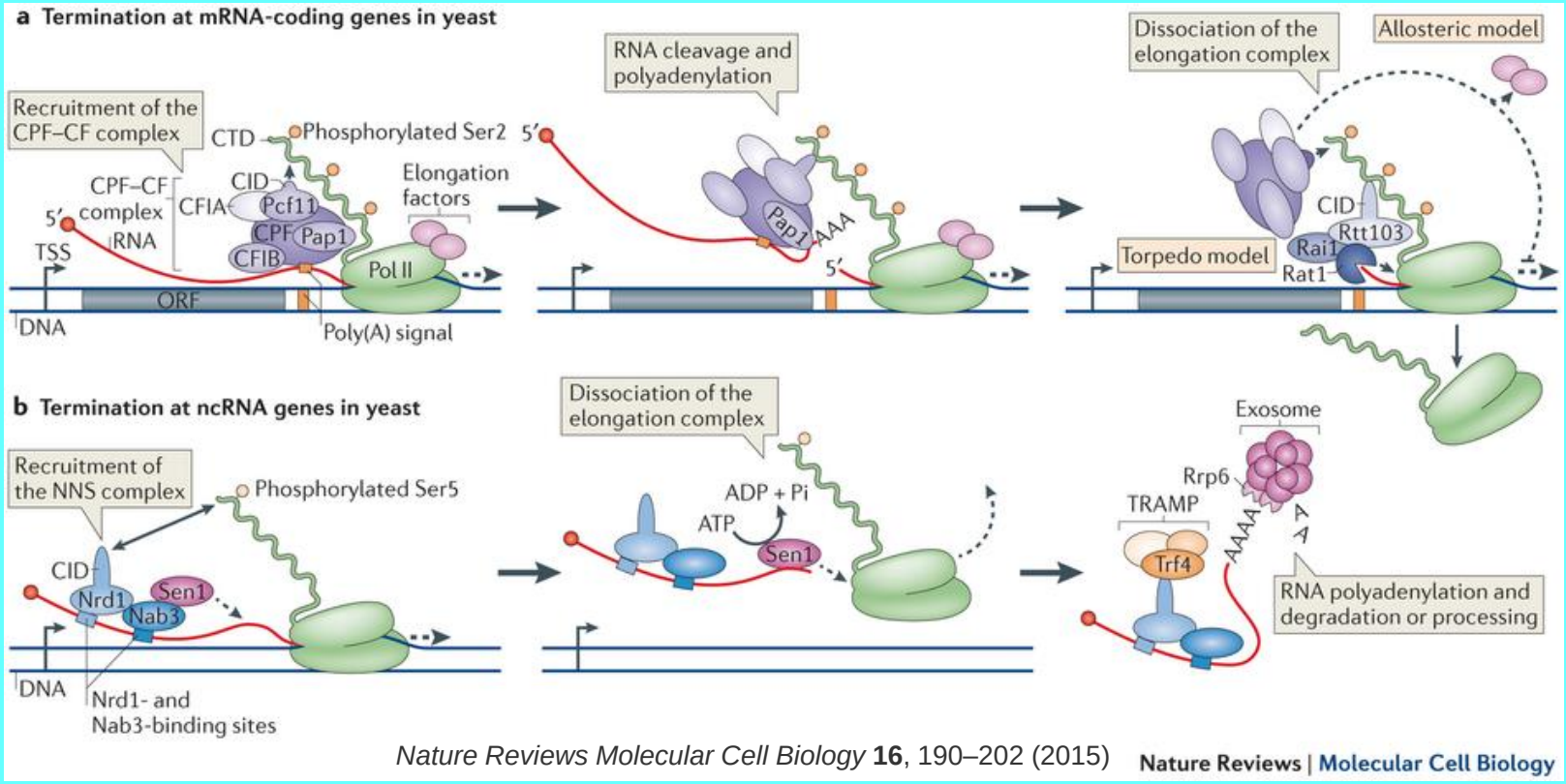
Dekker J.: Science 319, 1793 -1794 (2008)

Termination of transcription

1) Terminator contains AATAAA sequence = polyadenilation signal

2) Once polyadenilation signal is transcripted into hnRNA, it is recognised by protein complex, which cut hnRNA 10-30 nt towards 3´-end

3) Subsequently, RNAP II disociate from DNA and the rest of hnRNA behind the polyadenilation signal is degraded



Nature Reviews Molecular Cell Biology **16**, 190–202 (2015) Nature Reviews | Molecular Cell Biology

Posttranscription RNA processing

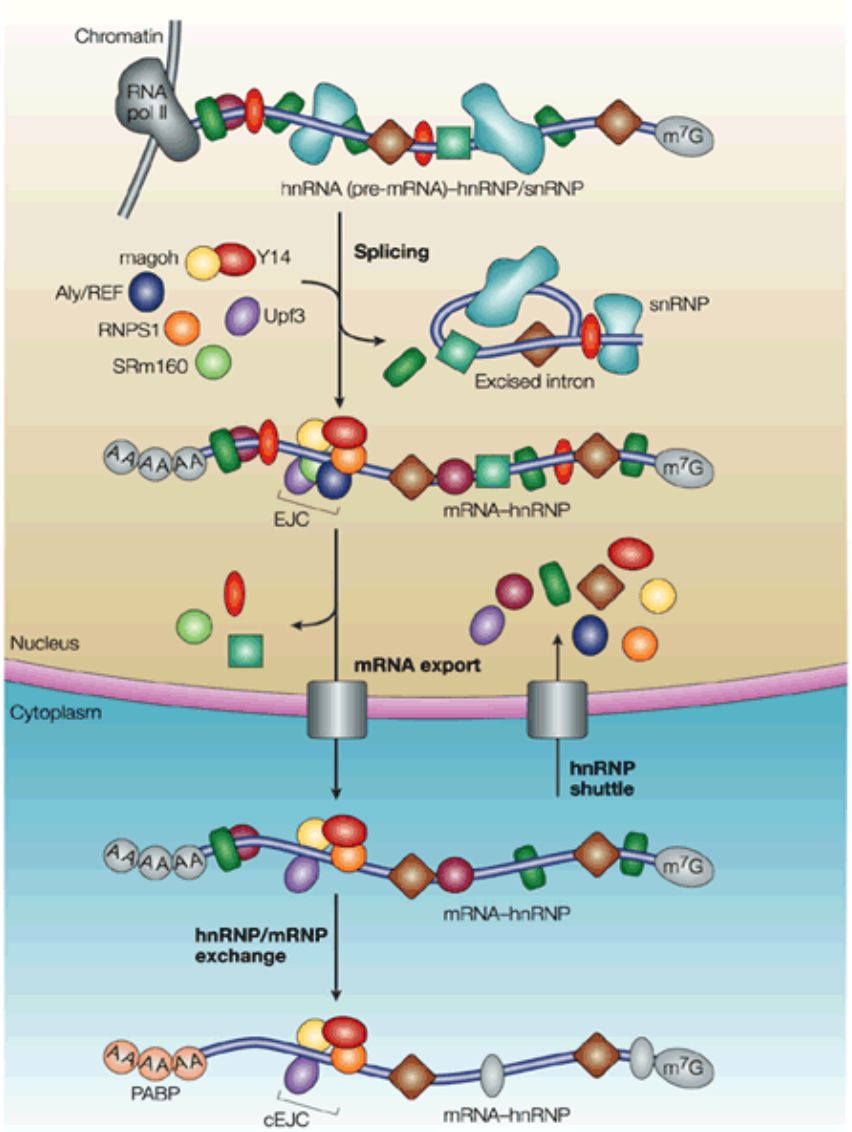
hnRNA modifications

- hnRNP-complexes forming
- adding cap to 5´- end
- polyadenylation of 3´- end
- splicing of hnRNA

hnRNP-complexes forming

- Proteins which specifically bind on hnRNA = **hnRNP-proteins**
- Proteins which specifically bind on small nuclear RNA (snRNA) = **snRNP-proteins**
- snRNP-proteins + snRNA = **snRNP-particles**
- hnRNA + hnRNP-proteins + snRNP-particles = **hnRNP-complex**
- snRNP-particles bind on introns and form **spliceosom**, which drive the hnRNA splicing
- hnRNP-proteins participate on transport of mRNA to cytoplasm

hnRNP-complexes forming

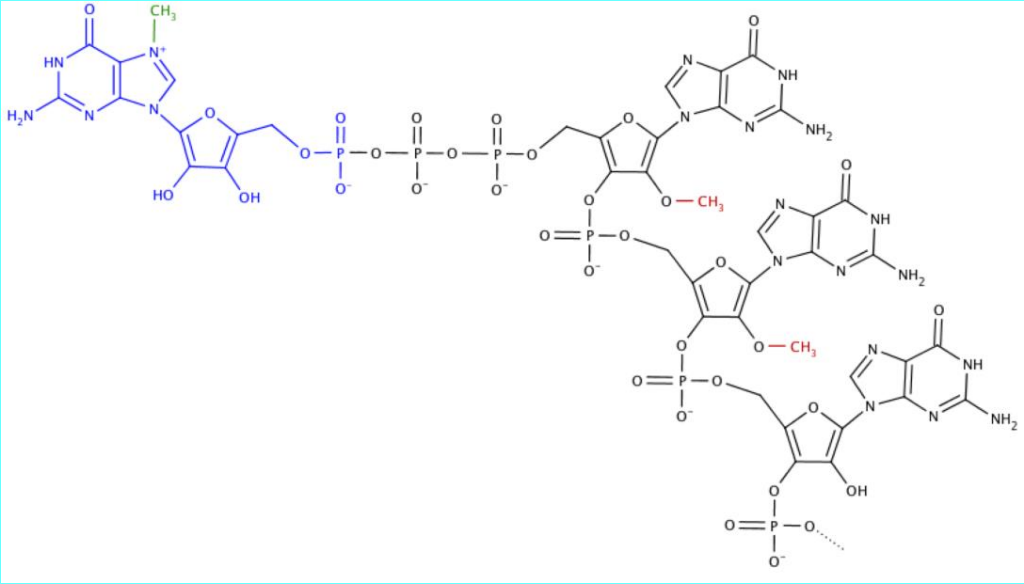


Nature Reviews Molecular Cell Biology **3**, 195-205 (March 2002)

Nature Reviews | Molecular Cell Biology

Adding cap to 5´- end

- Binding of 7-metylguanosine (**m^{7}G**) via three phosphate groups to 5´-end of hnRNA by 5´-5´ bound
- Last two 5´-end nucleotides could be aslo methylated
- m^{7}G plays important role during initiation of translation



pppNpRNA

Pi ↓ RNA Triphosphatase (RTPase)

ppNpRNA

GTP, PPi ↓ RNA Guanylyltransferase (GTase)

GpppNpRNA

SAM, SAH ↓ RNA (guanine-N7)-Methyltransferase (N7MTase)

m^{7}GpppNpRNA (Cap-0 RNA)

SAM, SAH ↓ RNA (nuceloside-2'-O-)-Methyltransferase (2'OMTase)

m^{7}GpppN$_m$pRNA (Cap-1 RNA)

DOI: 10.5772/56166

Polyadenylation of 3´- end

- Addition of of 50 – 250 adenosines to 3´-end of hnRNA = **poly(A) sequence**
- Catalyses by **poly(A)-polymerase**
- Poly(A)-polymerase is a subunit of complex, which binds on polyadenilation signal of hnRNA
- Poly(A) tail is important during transport of mRNA to cytoplasm and for its stabilisation

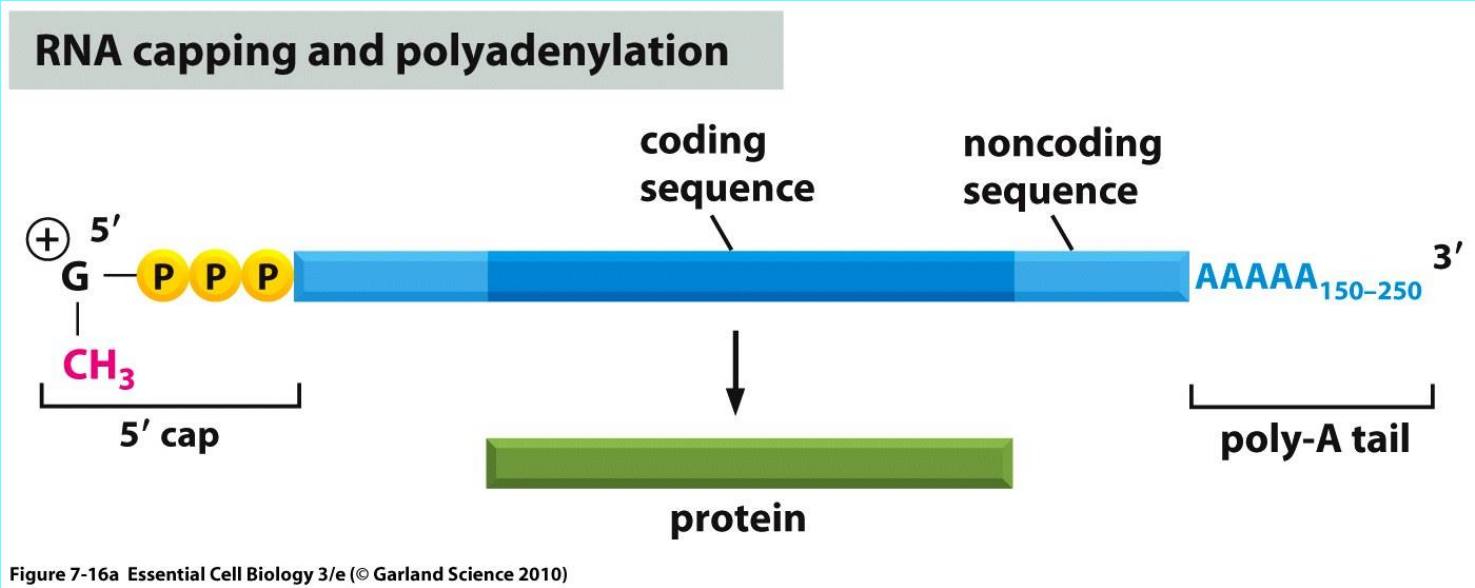


Figure 7-16a Essential Cell Biology 3/e (© Garland Science 2010)

Eukaryotic translation

Differences to prokaryotic translation

- It proceeds in 2-3 compartements, cytoplasm,
- mitochondria, and chloroplasts
- The first AA is not fMet, but Met, which binds to a specific initiator $tRNA_i^{Met}$, which recognize the AUG codon
- The number of initiation factors which are necessary to beginning of translation is higher in eukaryotes
- The number of initiation factors which are necessary to beginning of translation is higher in eukaryotes

http://www.ncbi.nlm.nih.gov Taxonomy Genetic Codes

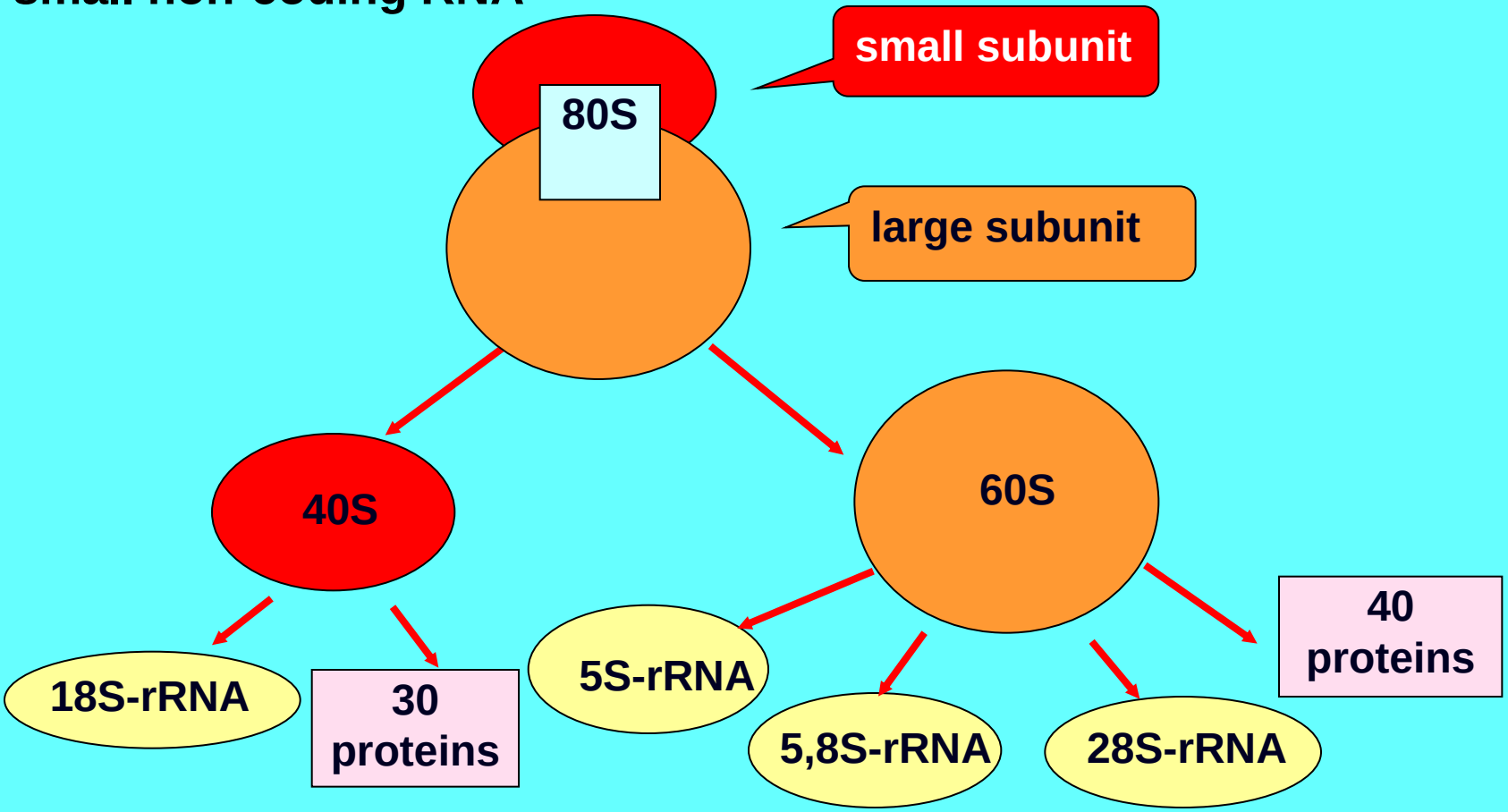
Course of translation

- similar as a translation in prokaryotes
- initiation, elongation, termination
- Particular complexes are more complicated
- More of translation factors
- Genetic code of mammalian mitochondria has different meaning of some codons, 22 tRNA
- Eukaryotic cell possesses 45 tRNA with different anticodons
- Speed of translation - 1-20 AA/s, depends on species and enviroment

The cytoplasmic ribosomes

Formation of ribosome structure involves also

- **150 non-ribosomal proteins**
- **100 small non-coding RNA**



Ferreira-Cerca, S. et al. (2007): Analysis of the In Vivo Assembly Pathway of Eukaryotic 40S Ribosomal Proteins, Molecular Cell 28, 446-457, November 2007

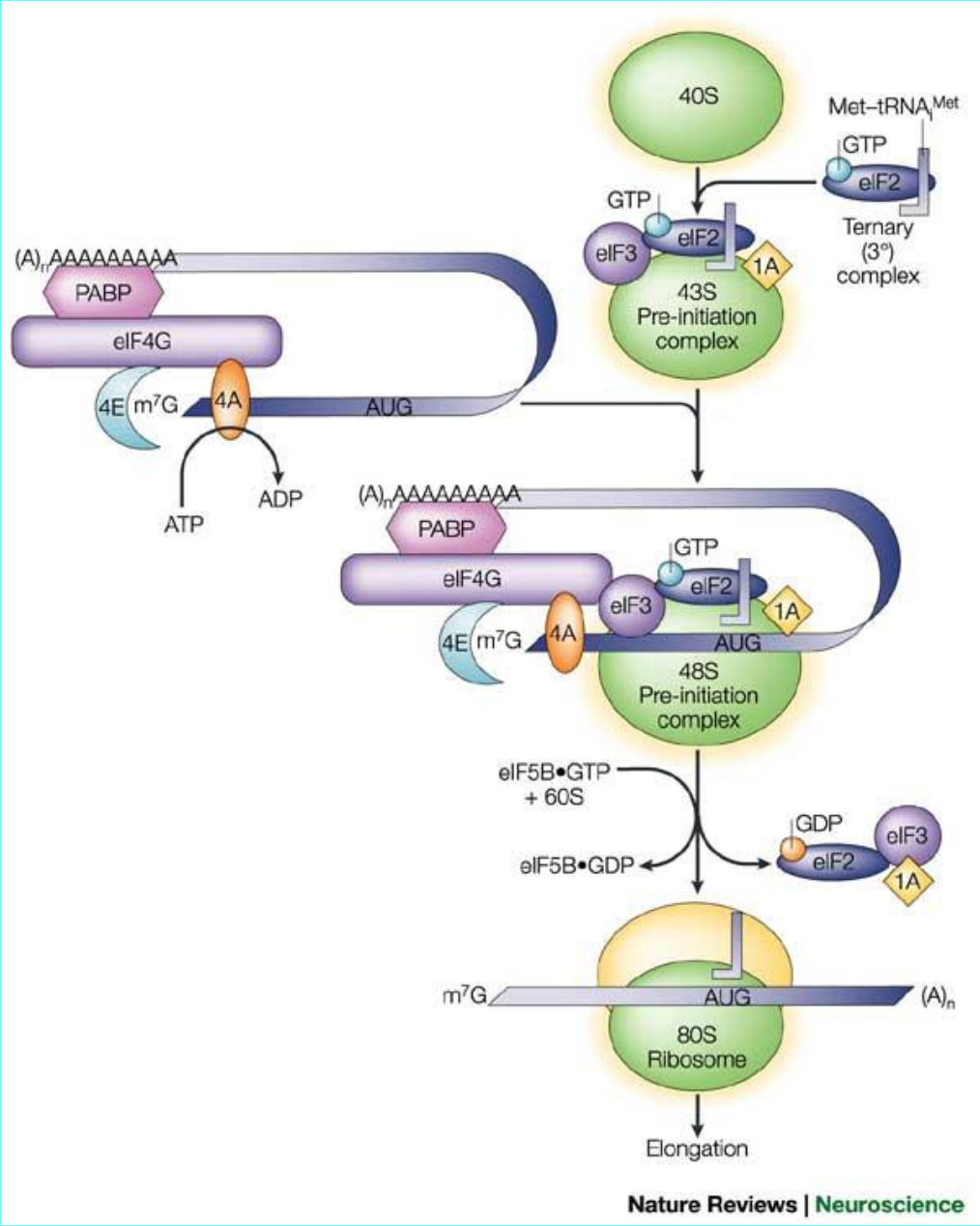
Free and bound ribosomes

- **Free ribosomes occur in cytoplasm**
  - **synthesis of intracellular proteins**

- **the rest is bounded to the endoplasmic reticulum**
  - **rough ER = covered by ribosomes**
  - **smooth ER = without ribosomes**
  - **synthesis of extracellular proteins**

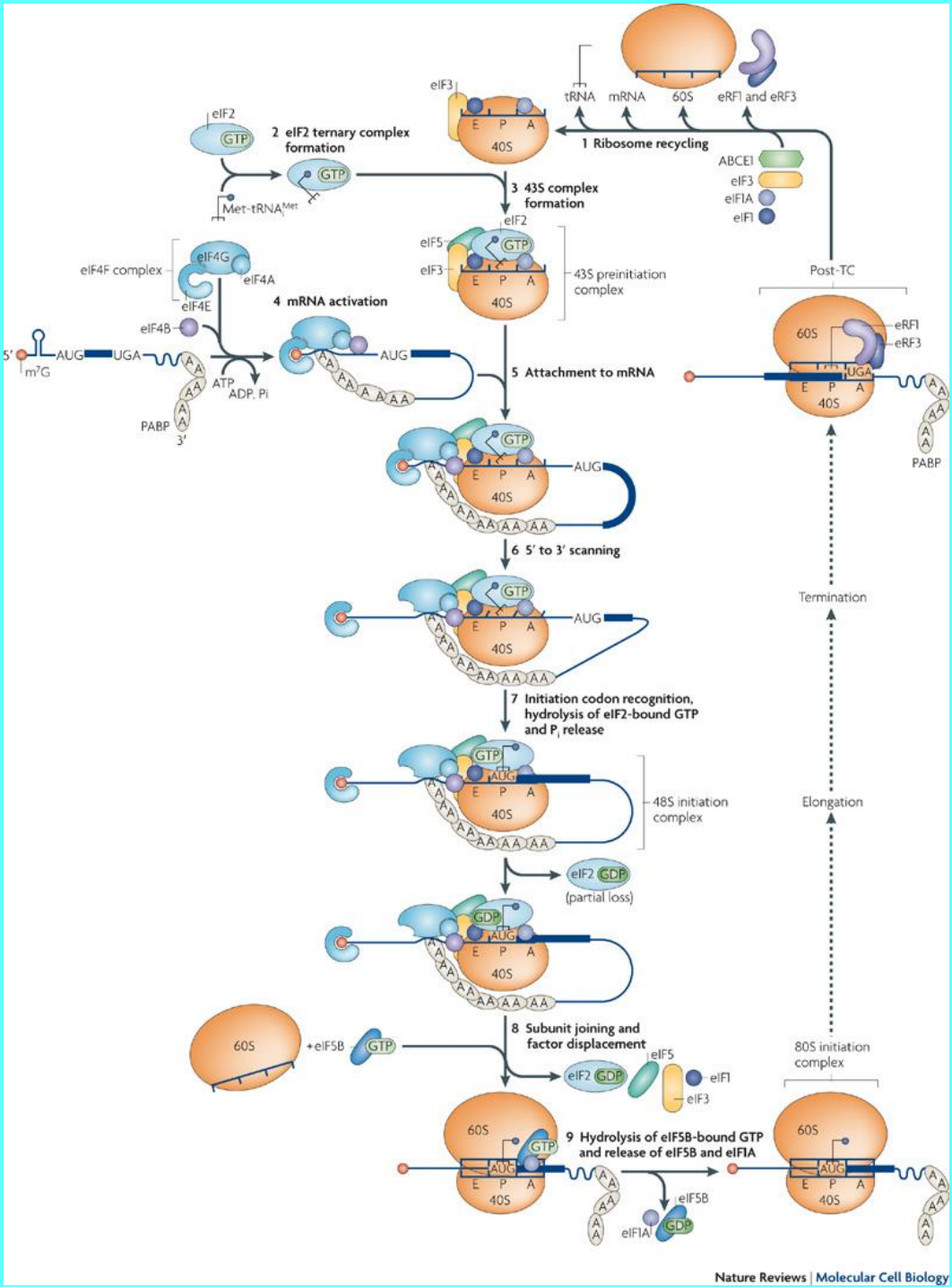
Initiation of translation

- 40S subunit with bound $tRNA_i^{Met}$ in P-site and initiation factors recognise m^7G cap of mRNA
- Subsequently, this complex moves to 3´-end until finds the initiation codon AUG
- Large 60S subunit binds to 40S subunit using the energy from hydrolysis of GTP



Nature Reviews Neuroscience **5**, 931-942 (December 2004)
doi:10.1038/nrn1557

Initiation of translation



Nature Reviews Molecular Cell Biology **11**, 113-127 (February 2010)
doi:10.1038/nrm2838

Initiation of translation – text to previous picture

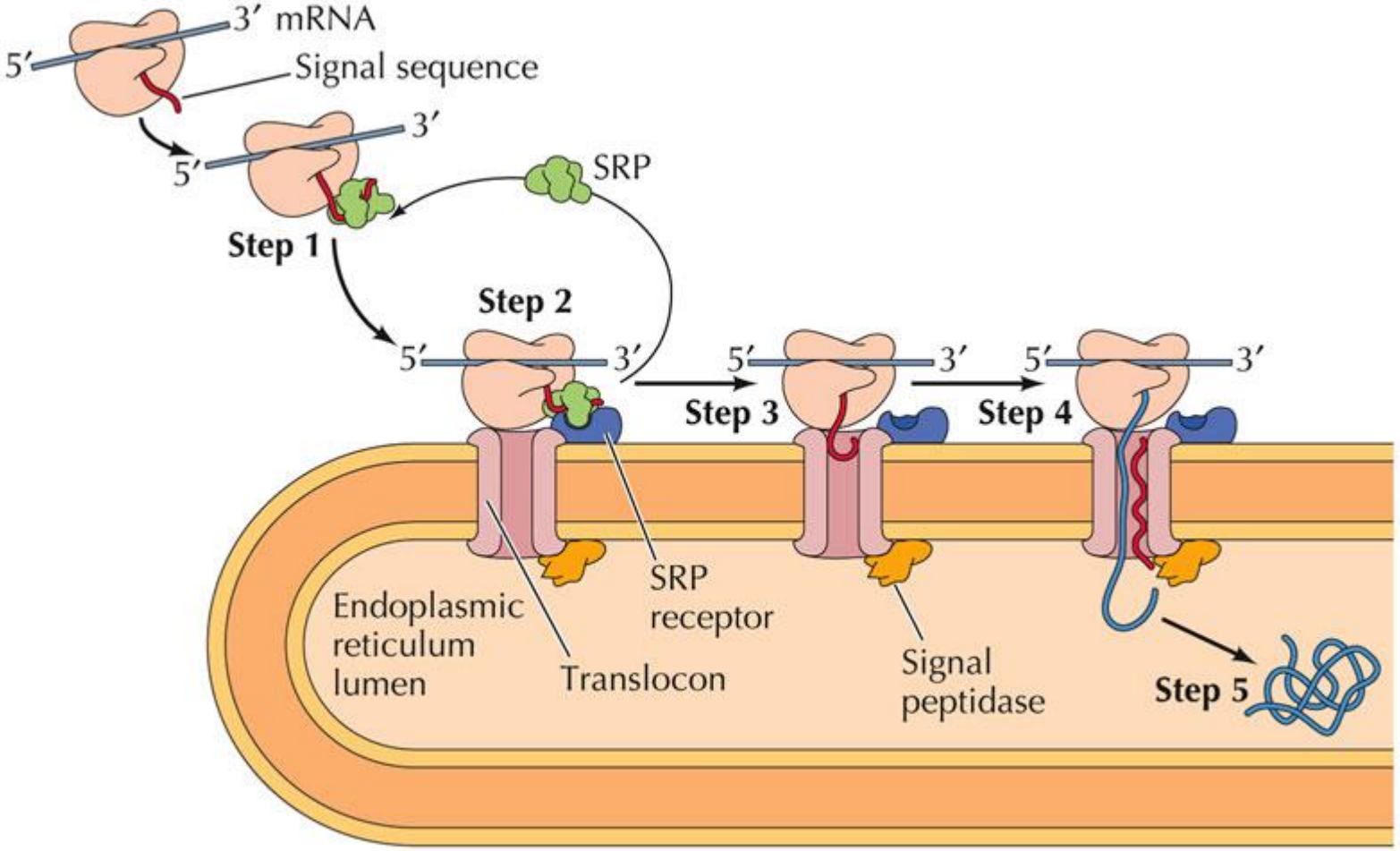
The canonical pathway of eukaryotic translation initiation is divided into eight stages (**2**–**9**). These stages follow the recycling of post-termination complexes (post-TCs; **1**) to yield separated 40S and 60S ribosomal subunits, and result in the formation of an 80S ribosomal initiation complex, in which Met-tRNA$^{Met}_{i}$ is base paired with the initiation codon in the ribosomal P-site and which is competent to start the translation elongation stage. These stages are: eukaryotic initiation factor 2 (eIF2)–GTP–Met-tRNA$^{Met}_{i}$ ternary complex formation (**2**); formation of a 43S preinitiation complex comprising a 40S subunit, eIF1, eIF1A, eIF3, eIF2–GTP–Met-tRNA$^{Met}_{i}$ and probably eIF5 (**3**); mRNA activation, during which the mRNA cap-proximal region is unwound in an ATP-dependent manner by eIF4F with eIF4B (**4**); attachment of the 43S complex to this mRNA region (**5**); scanning of the 5′ UTR in a 5′ to 3′ direction by 43S complexes (**6**); recognition of the initiation codon and 48S initiation complex formation, which switches the scanning complex to a 'closed' conformation and leads to displacement of eIF1 to allow eIF5-mediated hydrolysis of eIF2-bound GTP and P_i release (**7**); joining of 60S subunits to 48S complexes and concomitant displacement of eIF2–GDP and other factors (eIF1, eIF3, eIF4B, eIF4F and eIF5) mediated by eIF5B (**8**); and GTP hydrolysis by eIF5B and release of eIF1A and GDP-bound eIF5B from assembled elongation-competent 80S ribosomes (**9**). Translation is a cyclical process, in which termination follows elongation and leads to recycling (**1**), which generates separated ribosomal subunits. The model omits potential 'closed loop' interactions involving poly(A)-binding protein (PABP), eukaryotic release factor 3 (eRF3) and eIF4F during recycling (see Supplementary information S5 (box)), and the recycling of eIF2–GDP by eIF2B. Whether eRF3 is still present on ribosomes at the recycling stage is unknown.

Termination of translation

- **Only one termination factor = eRF**
- **Disociation of ribosome from mRNA needs the energy from GTP**

https://www.youtube.com/watch?v=qIwrhUrvX-k

Translocation of extracellular proteins



THE CELL, Fourth Edition, Figure 10.8 © 2006 ASM Press and Sinauer Associates, Inc.