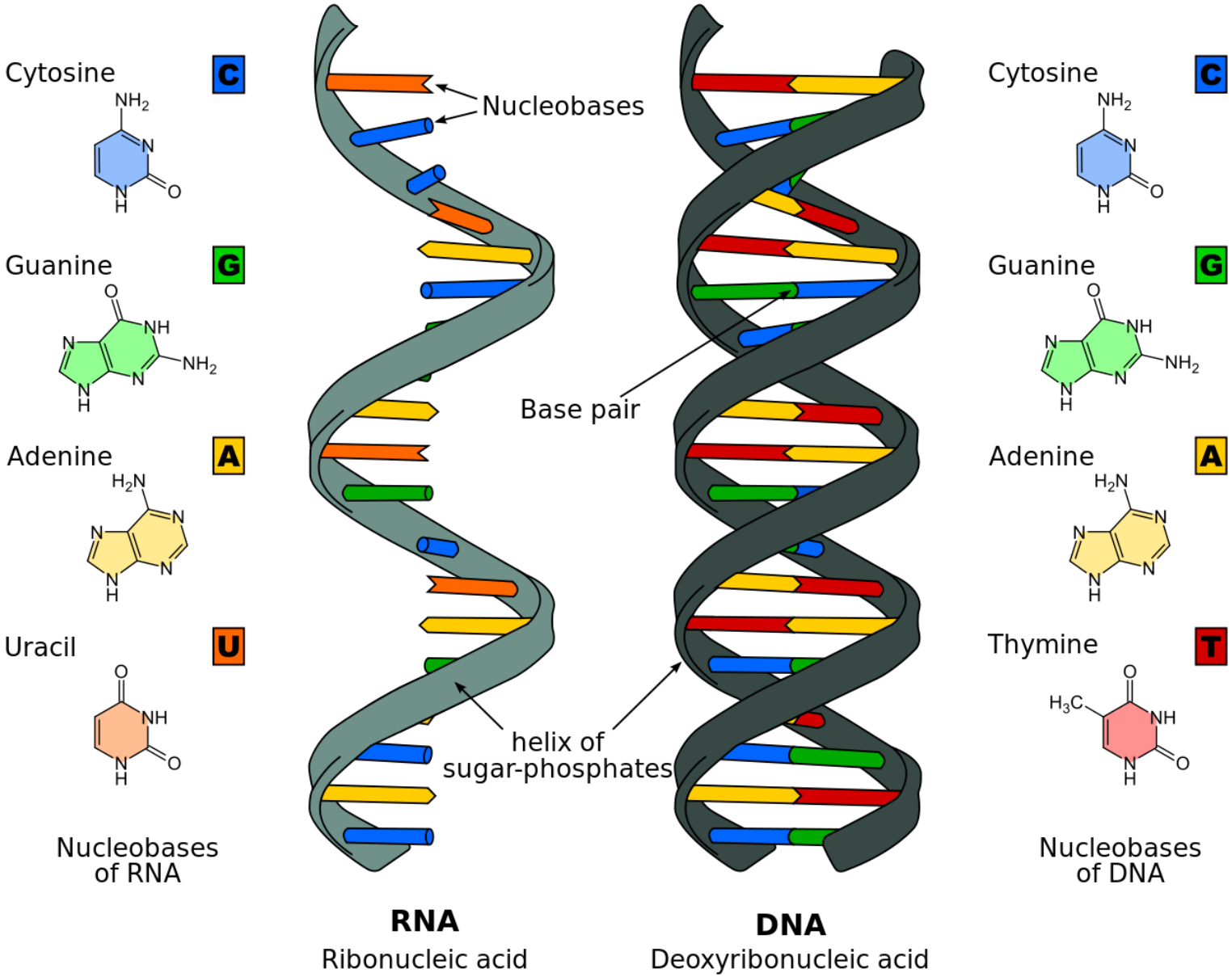
Lecture 3 – DNA and RNA

Prof. Sebastian Maerkl

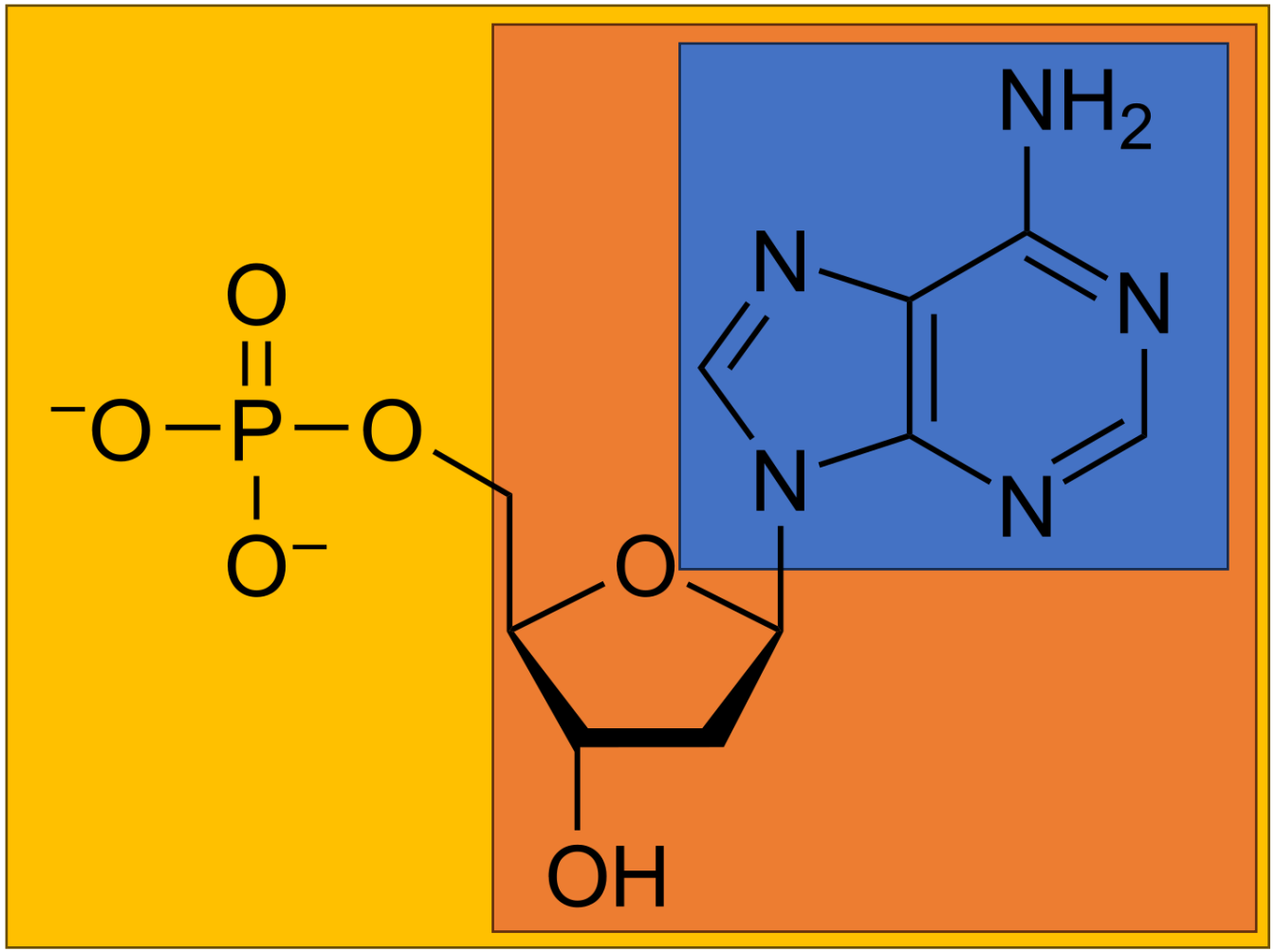
DNA Structure and Function

RNA and DNA

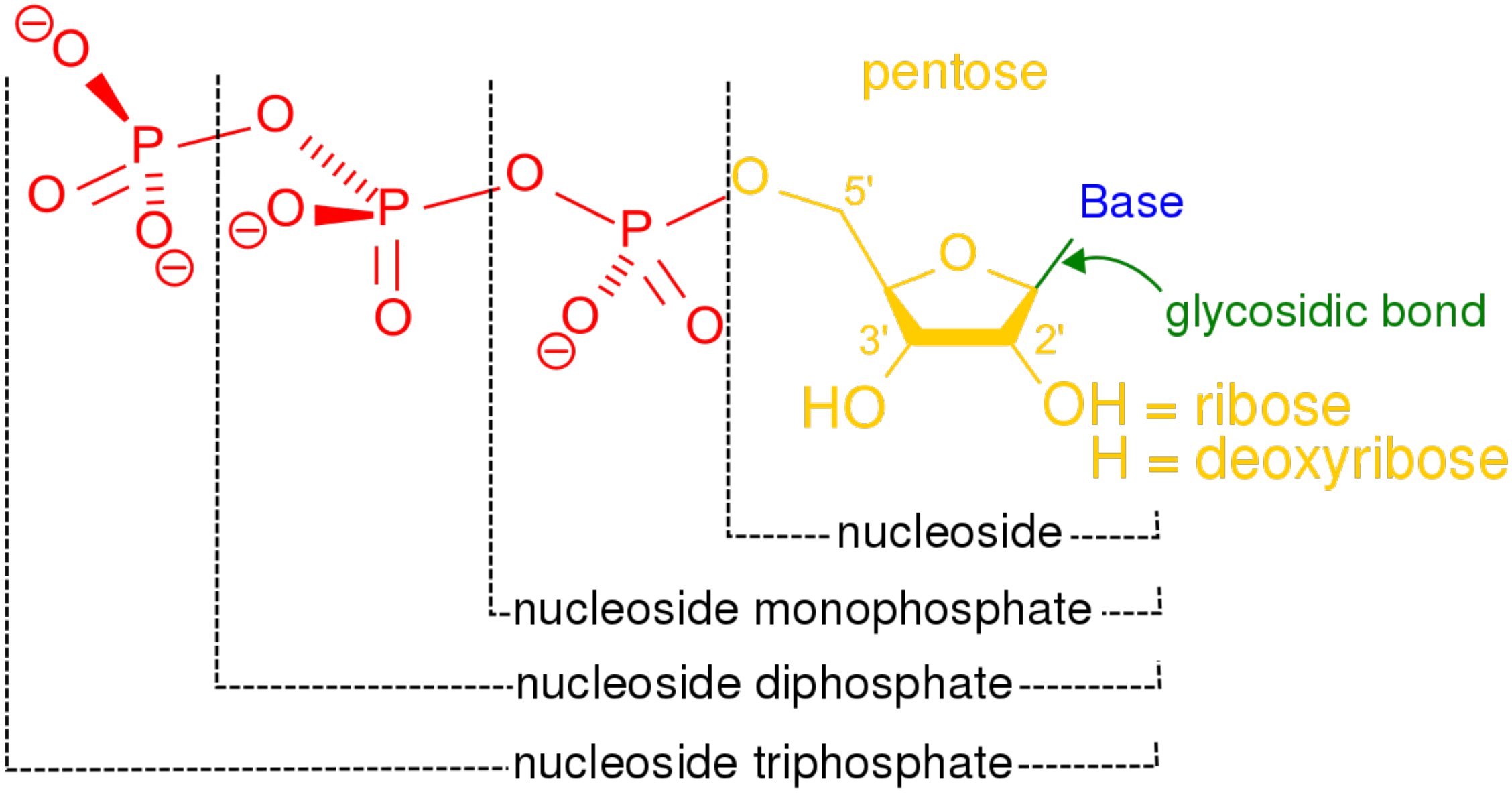

Cytosine
C
Guanine
G
Adenine
A
Uracil
U
Nucleobases of RNA
Nucleobases
Base pair
helix of sugar-phosphates
RNA
Ribonucleic acid
DNA
Deoxyribonucleic acid
Cytosine
C
Guanine
G
Adenine
A
Thymine
T
Nucleobases of DNA

Terminology

Nucleobase
Nucleoside
Nucelotide

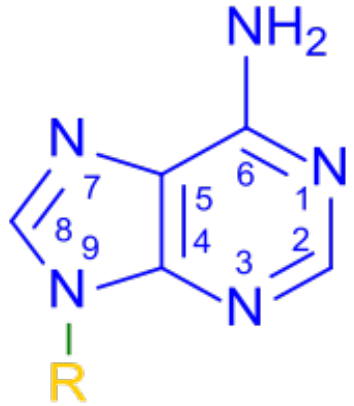


RNA and DNA

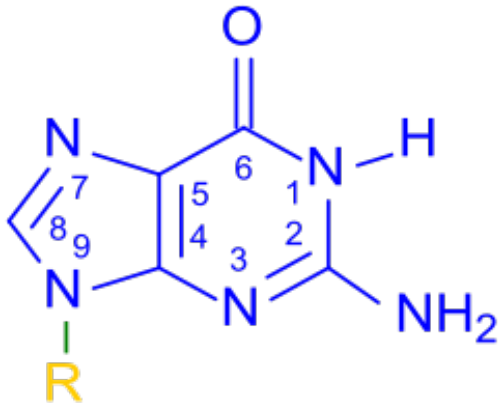

pentose
Base
glycosidic bond
5'
3'
2'
HO
OH = ribose
H = deoxyribose
nucleoside
nucleoside monophosphate
nucleoside diphosphate
nucleoside triphosphate

RNA and DNA

Purines

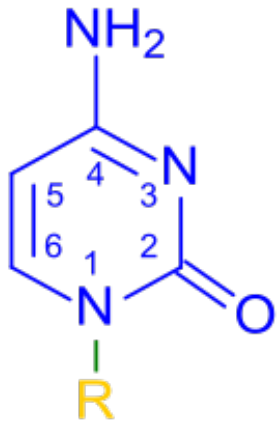


Adenine

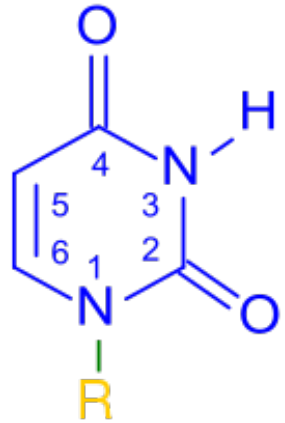


Guanine

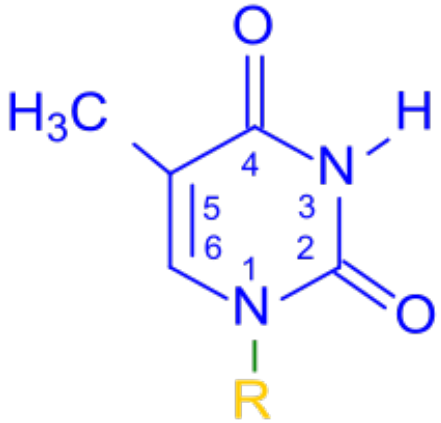
Pyrimidines



Cytosine

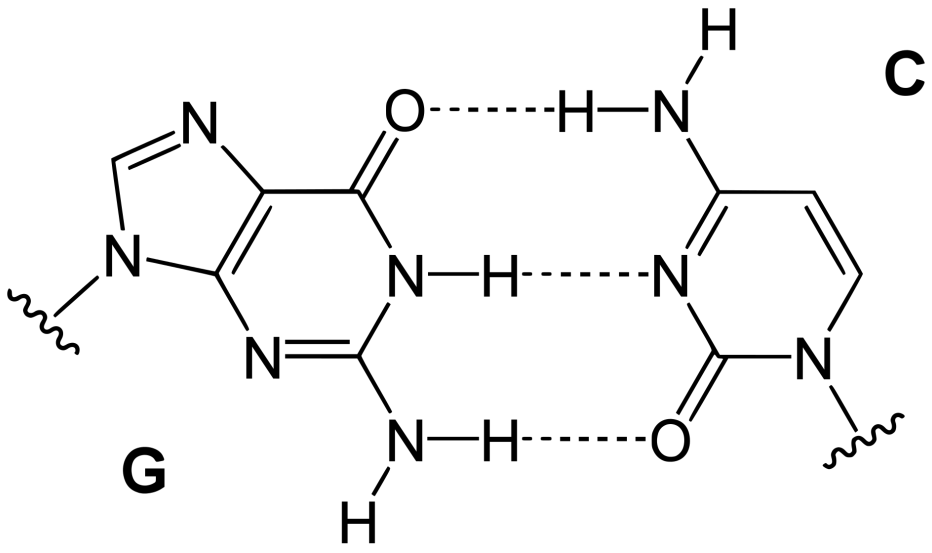


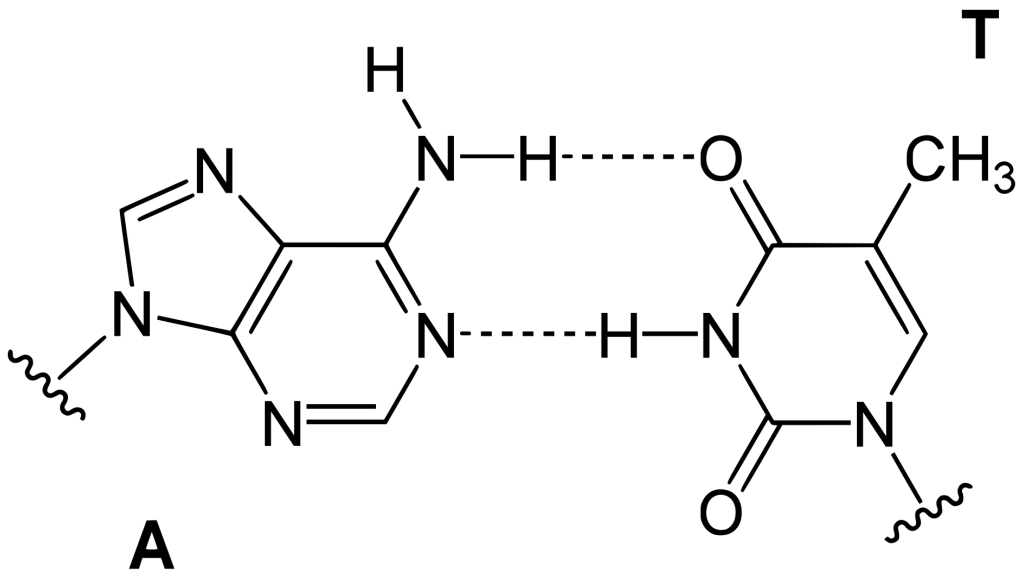
Uracil



Thymine

Base pairing via hydrogen bonds



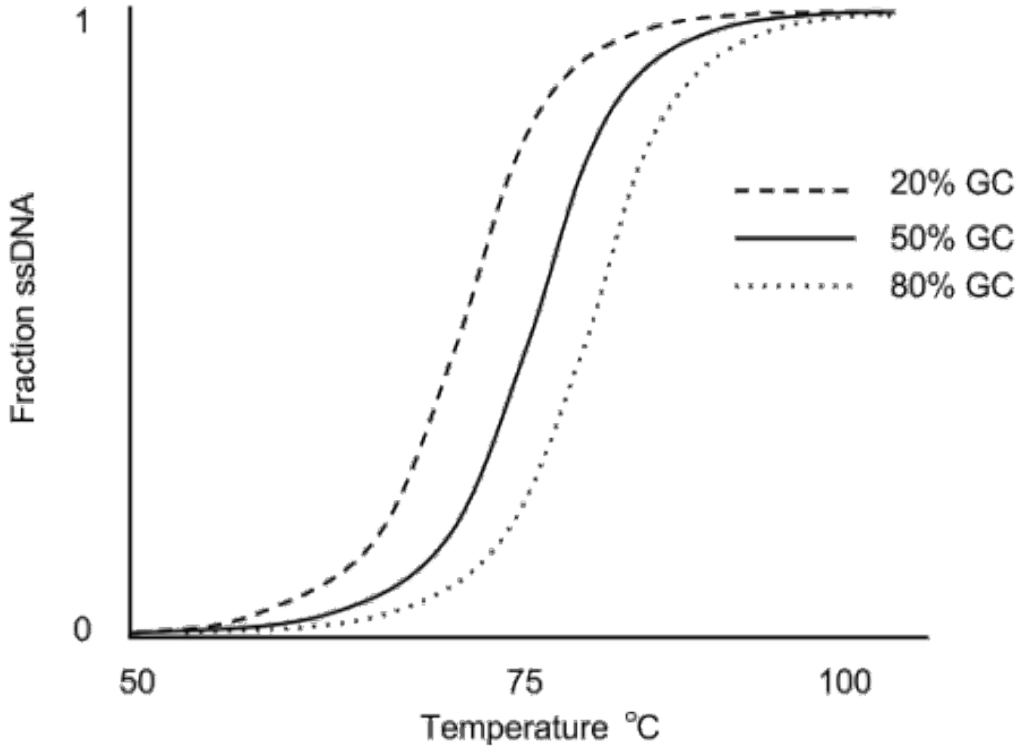


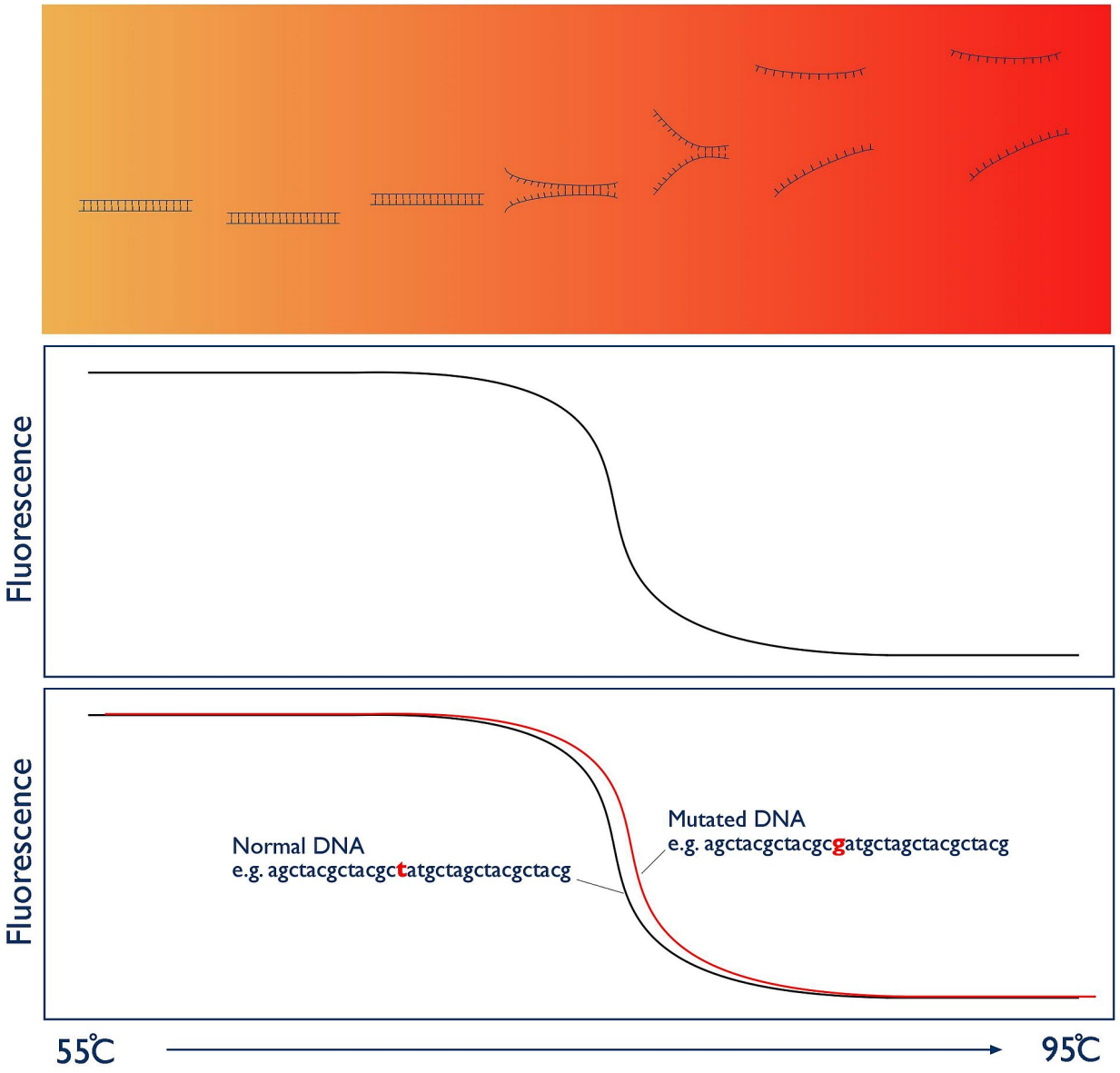
H-bond strength: ~1-5 kcal/mol

A-T: 12-14 kcal /mol

C-G: 21-28 kcal/mol

Melting Temperature





Melting Temperature

Product Group

Taq DNA Polymerase

Polymerase/Kit

Taq DNA Polymerase with Standard Taq Buffer

Primer Concentration (nM)

200

Reset concentration

Primer 1

ATCGATTGAGCTCTAGCG

Primer 2

ATCGATTGAGCTCTAGCG

Switch to batch mode

Clear

Use example input

Anneal at

48 °C

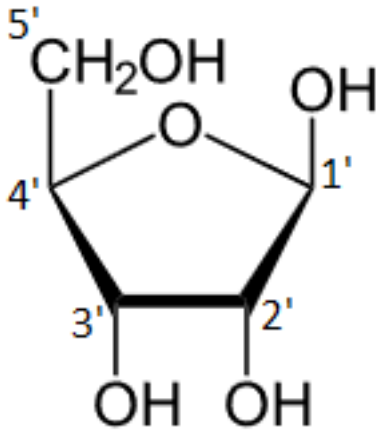
Primer 1
18 nt
50% GC
Tm: 53°C

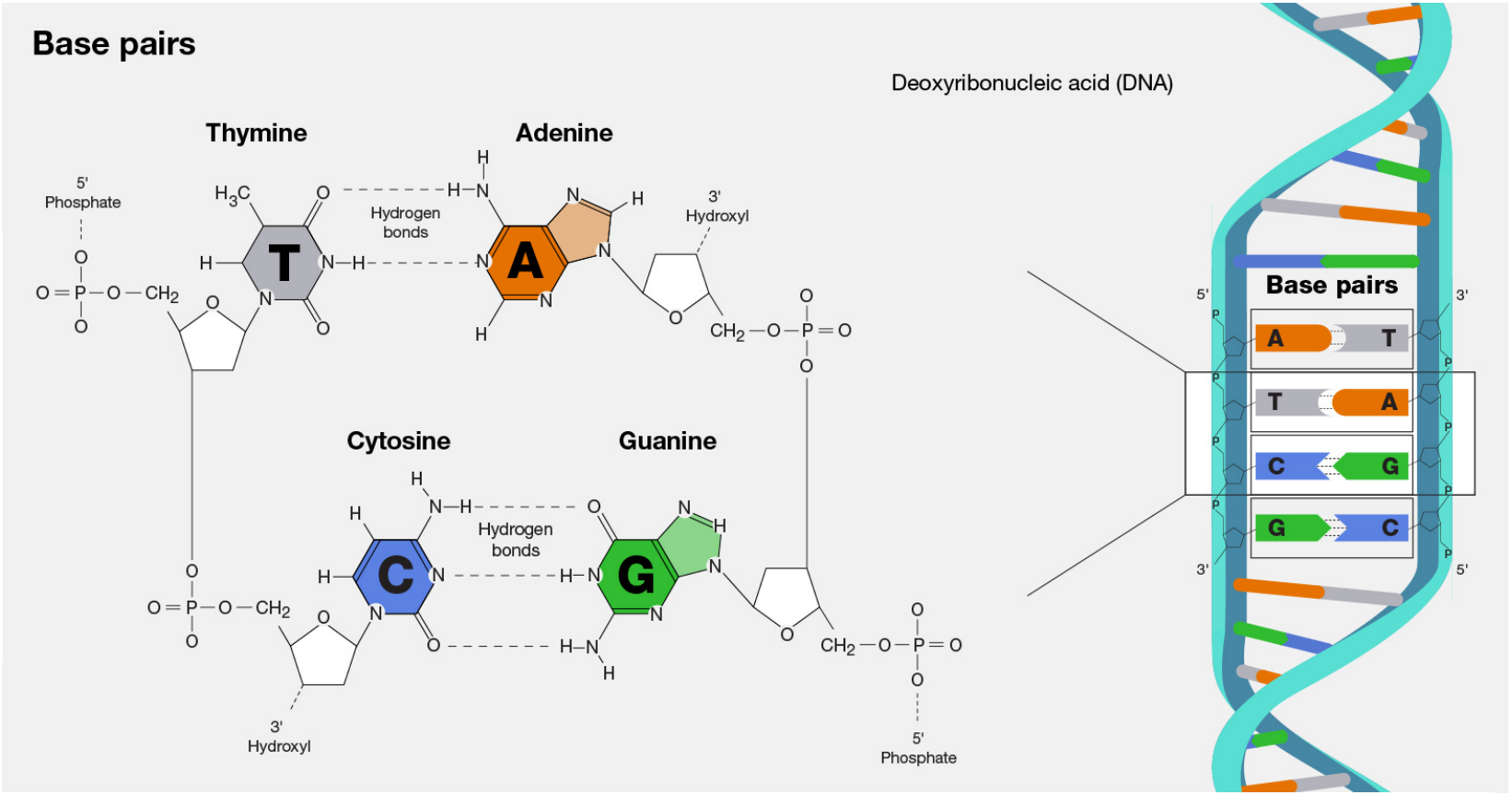
Primer 2
18 nt
50% GC
Tm: 53°C

Oligo Tool Live Demo

https://www.thermofisher.com/ch/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html

Base pairing and 5’ / 3’ Terminology





The following DNA sequences illustrate pair double-stranded patterns. By convention, the top strand is written from the 5′-end to the 3′-end; thus, the bottom strand is written 3′ to 5′.

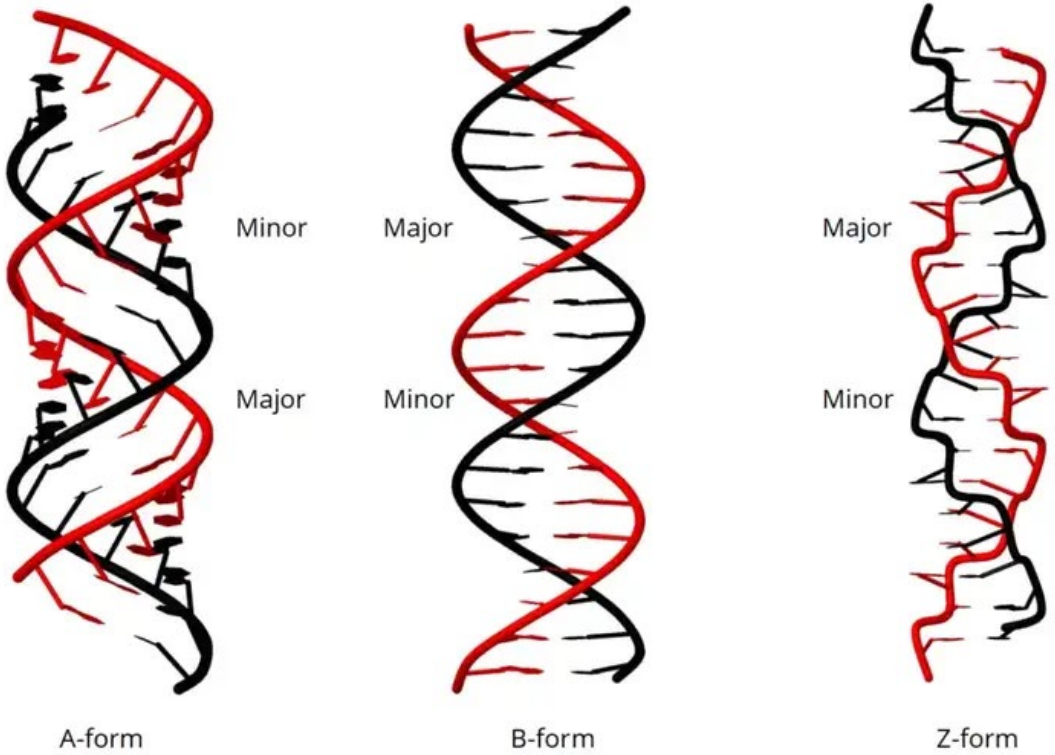
A base-paired DNA sequence:

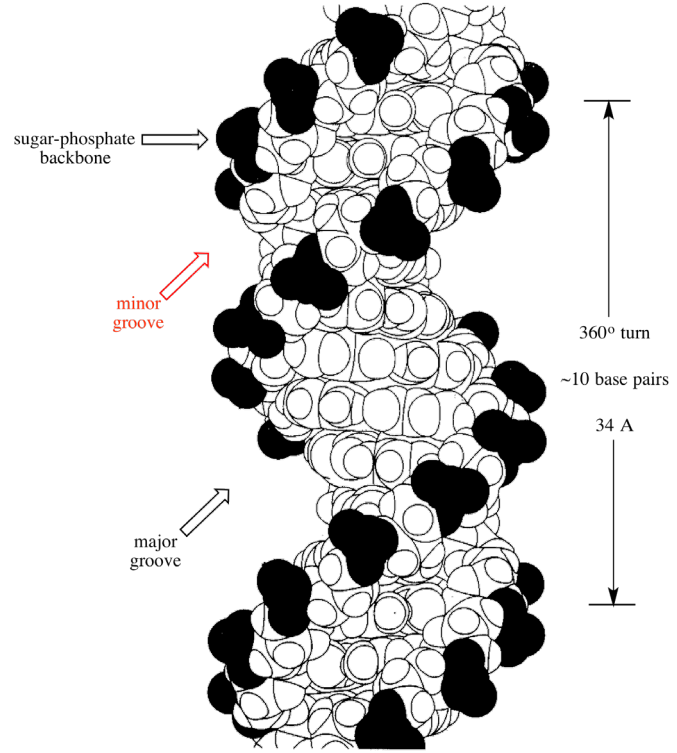
```
ATCGATTGAGCTCTAGCG
TAGCTAACTCGAGATCGC
```

The corresponding RNA sequence, in which uracil is substituted for thymine in the RNA strand:

```
AUCGAUUGAGCUCUAGCG
UAGCUAACUCGAGAUCGC
```

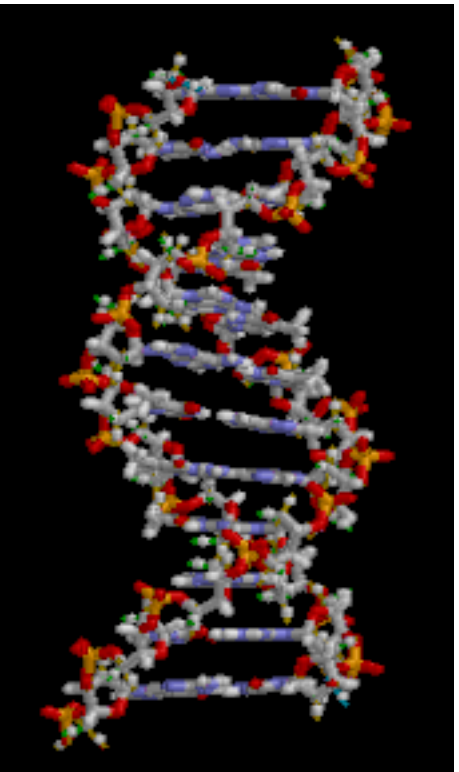
DNA Structure

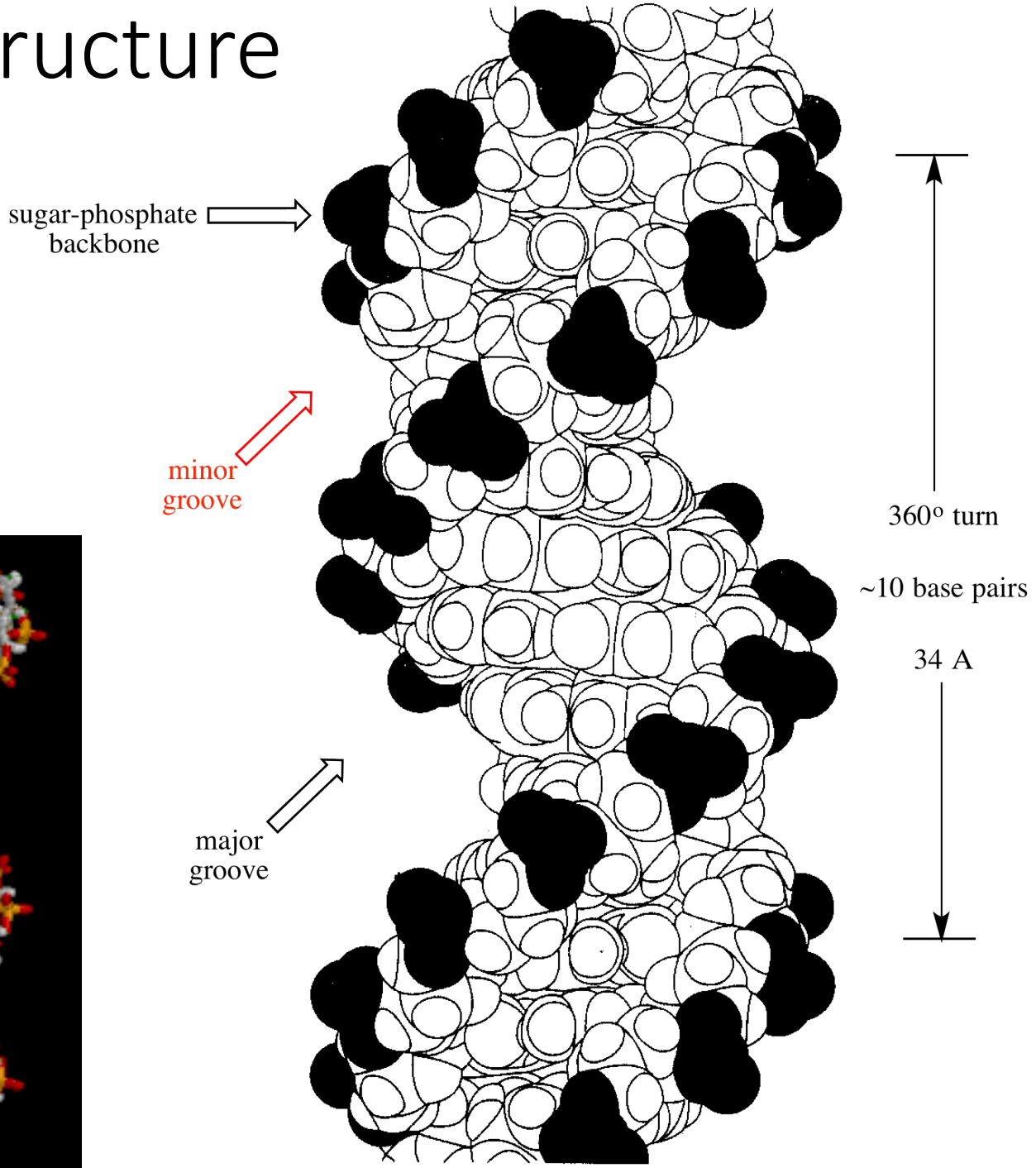


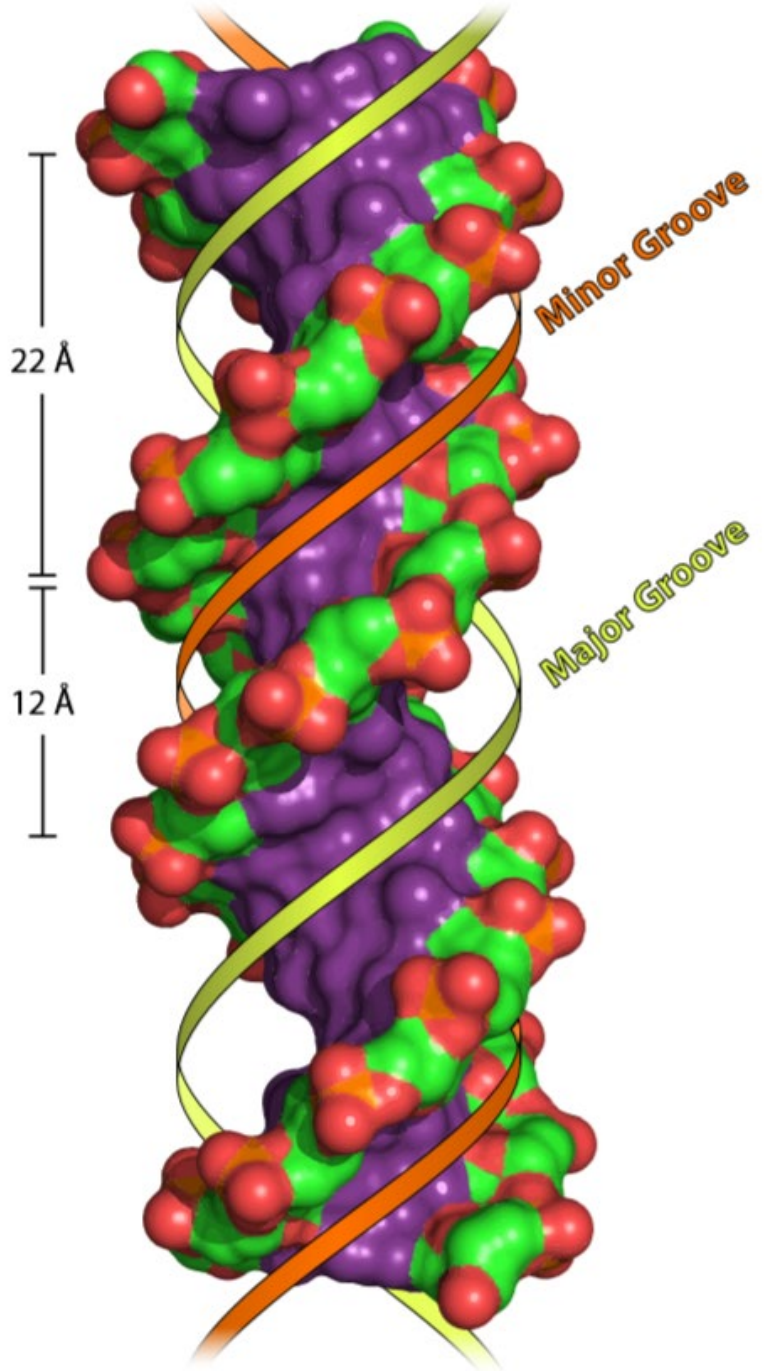


Prop|Form	A	B	Z
Spiral type	Right	Right	Left
Step, Å	28.03	33.75	43. 5
Bases per coil	11	10	12
Major groove width, Å	7.98 5:A.P — 30:B.P 4:A.P — 31:B.P	17.91 8:A.P — 29:B.P 9:A.P — 28:B.P	15.17 14:A.P — 27:B.P 13:A.P — 28:B.P
Minor groove width, Å	16.97 31:B.P — 13:A.P 30:B.P — 14:A.P	11.69 34:B.P — 11:A.P 35:B.P — 10:A.P	9.87 38:B.P — 7:A.P 37:B.P — 8:A.P

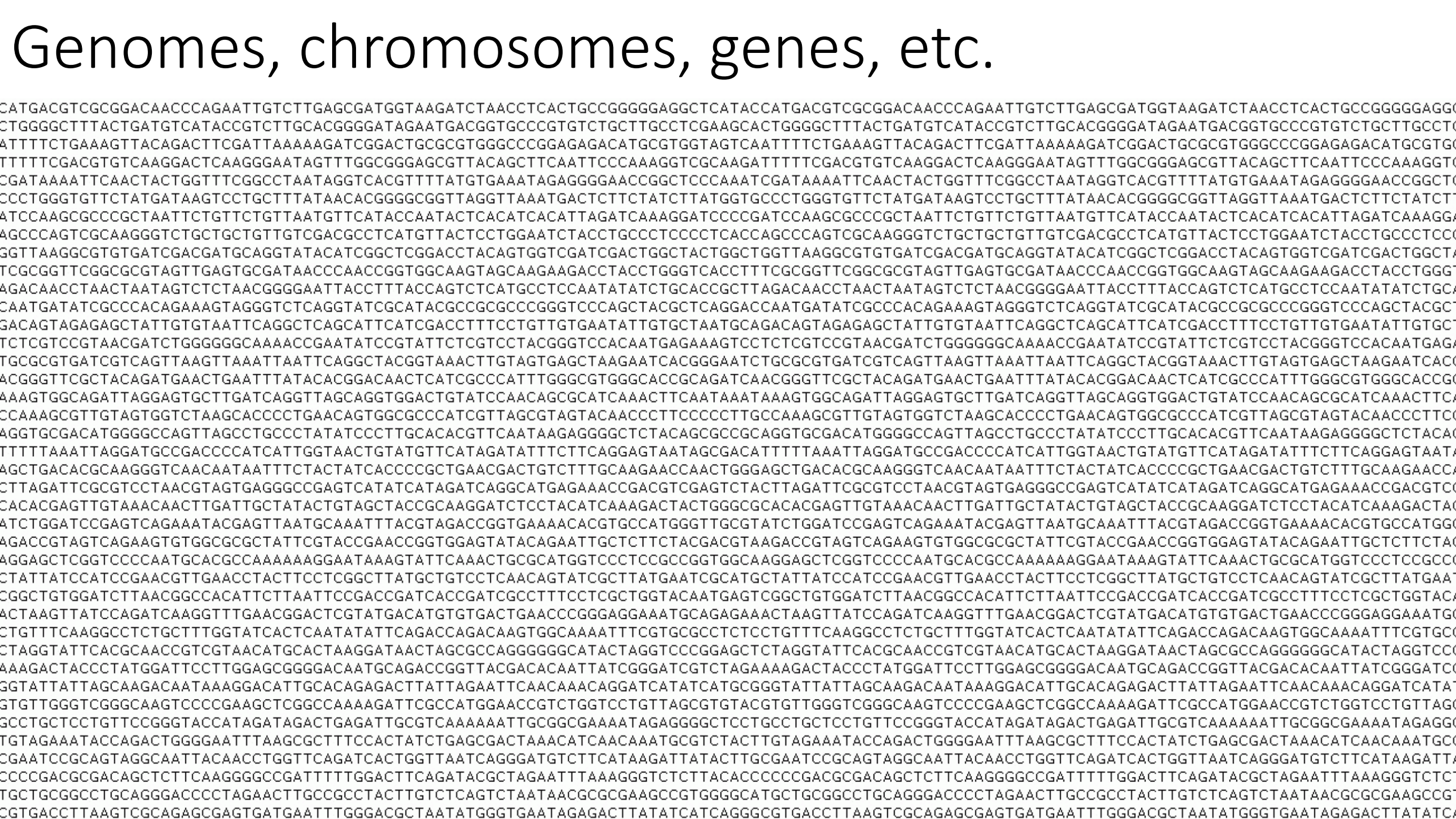
DNA Structure



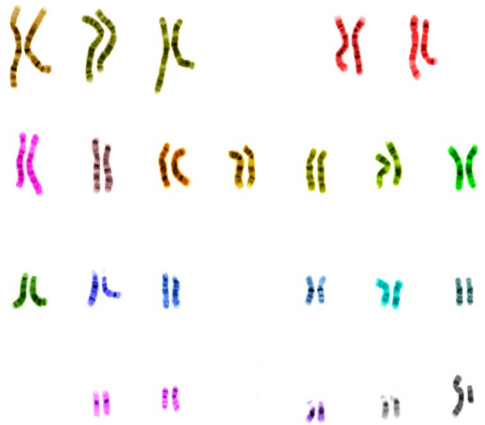




Genomes, chromosomes, genes, etc.



Terminology



Genome: is all the genetic information of an organism.

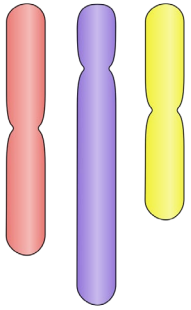
Chromosome: is a long DNA molecule with part or all of the genetic material of an organism.

Chromatin: is a complex of DNA and protein found in eukaryotic cells.

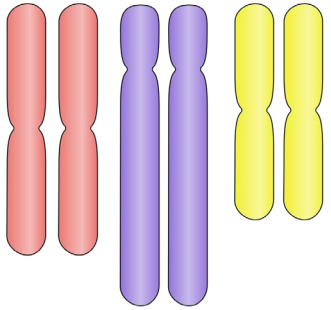
Ploidy: is the number of complete sets of chromosomes in a cell -> most common are haploid, diploid, polyploid

Karyotype: is the general appearance of the complete set of chromosomes in the cells of a species or in an individual organism, mainly including their sizes, numbers, and shapes.

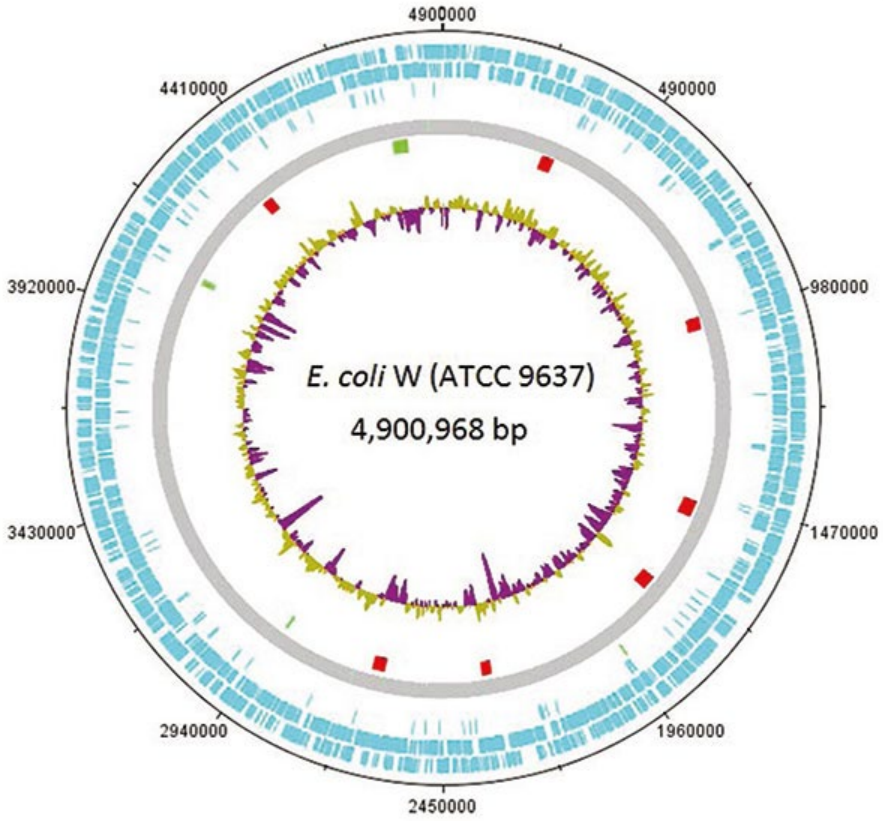
Haploid (N)

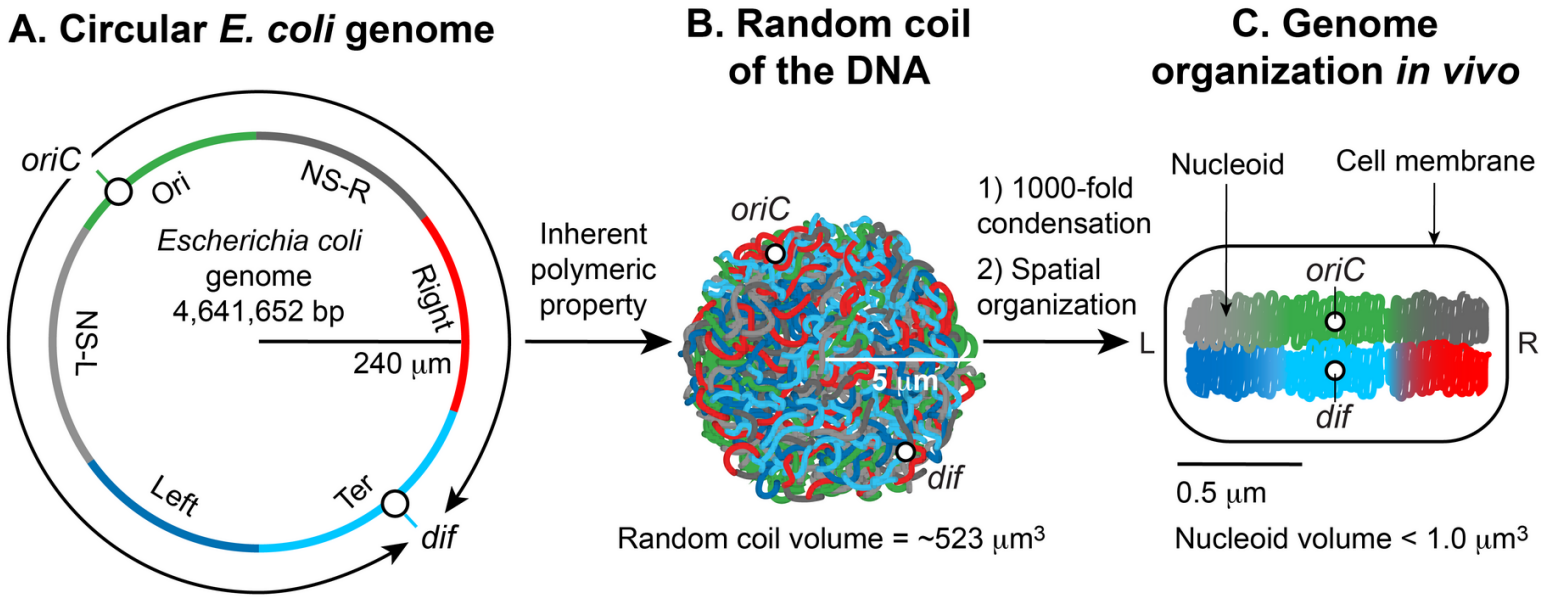


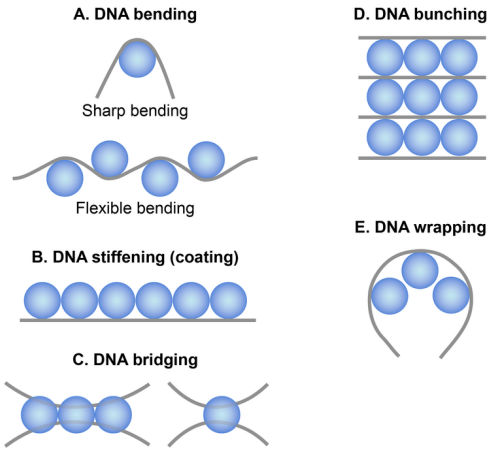
Diploid (2N)



Prokaryotic Genomes and DNA condensation

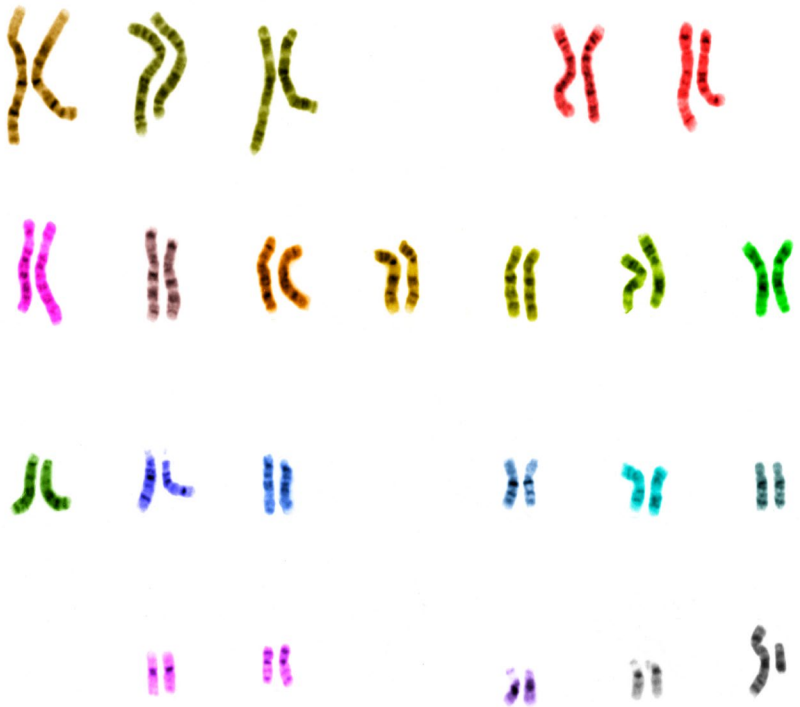


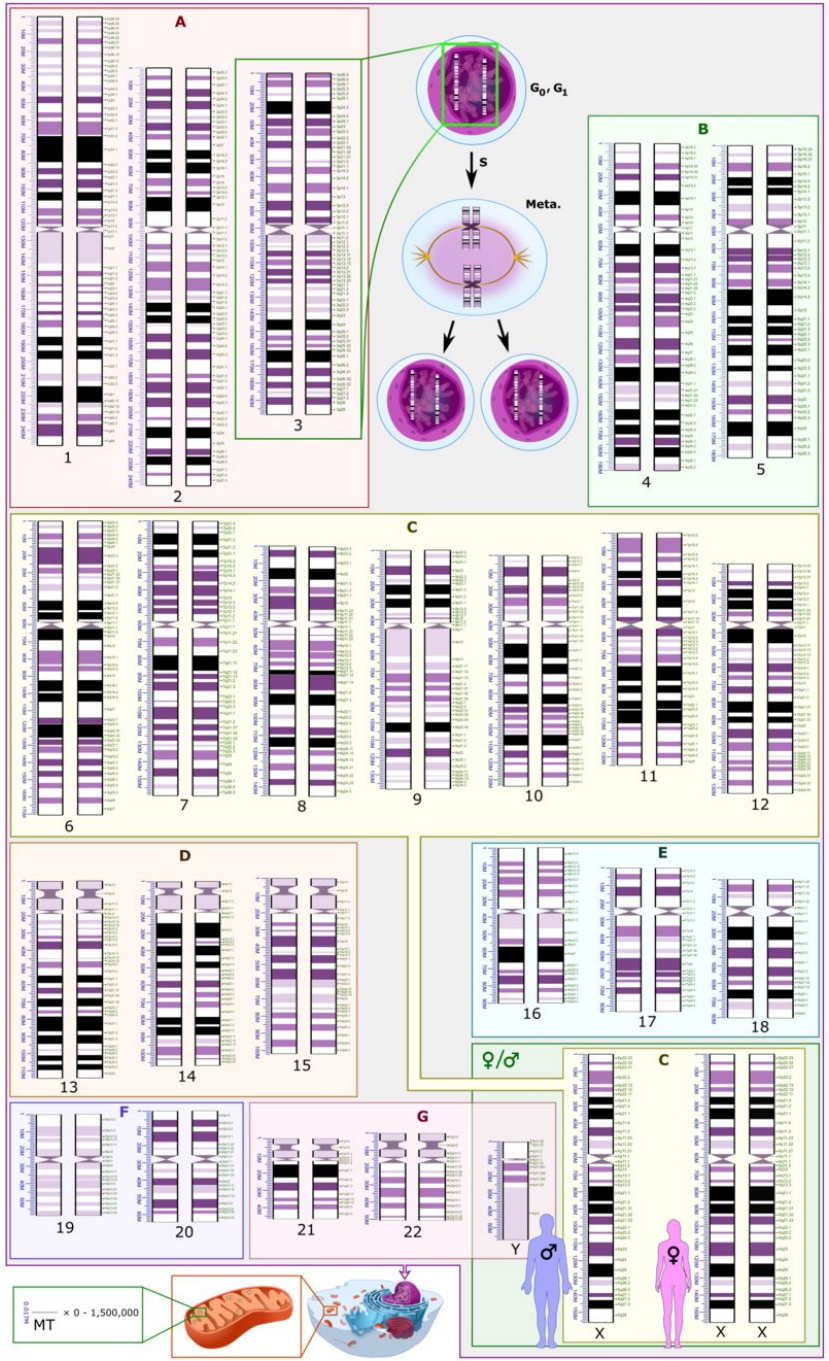




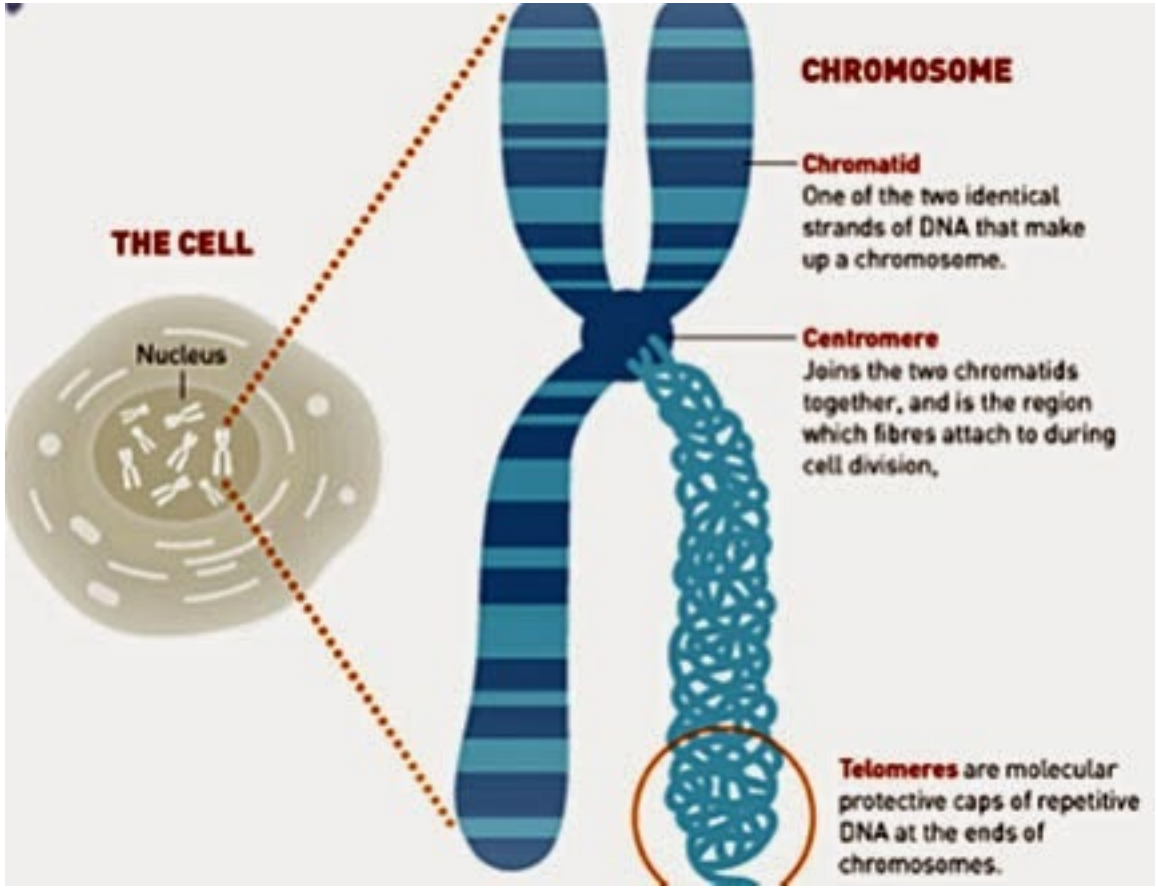
NAP	DNA organization
HU	DNA bending, bunching, and stiffening
IHF	DNA bending
H-NS	DNA stiffening and bridging
Fis	DNA bending and bridging

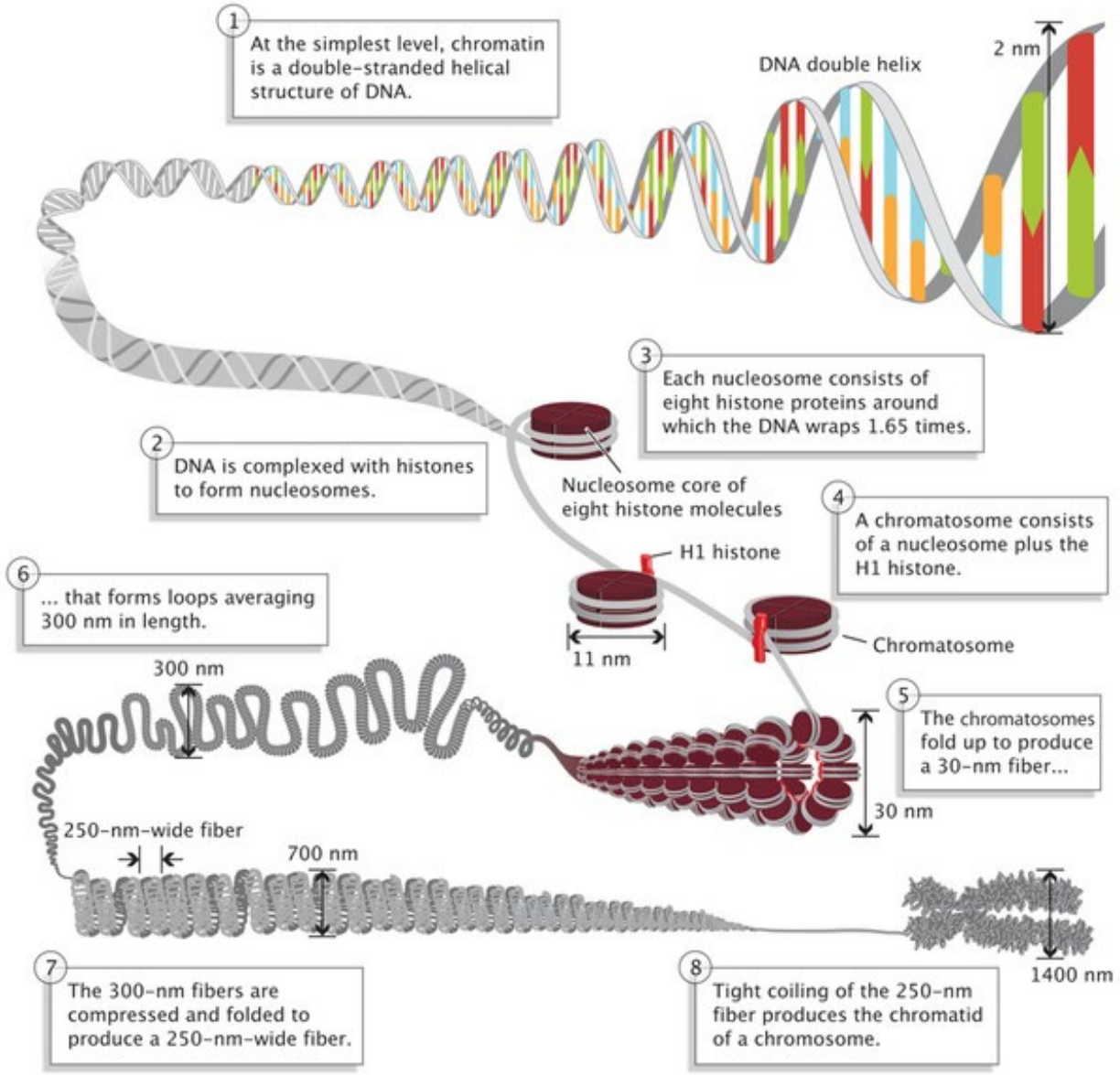
Eukaryotic Genomes



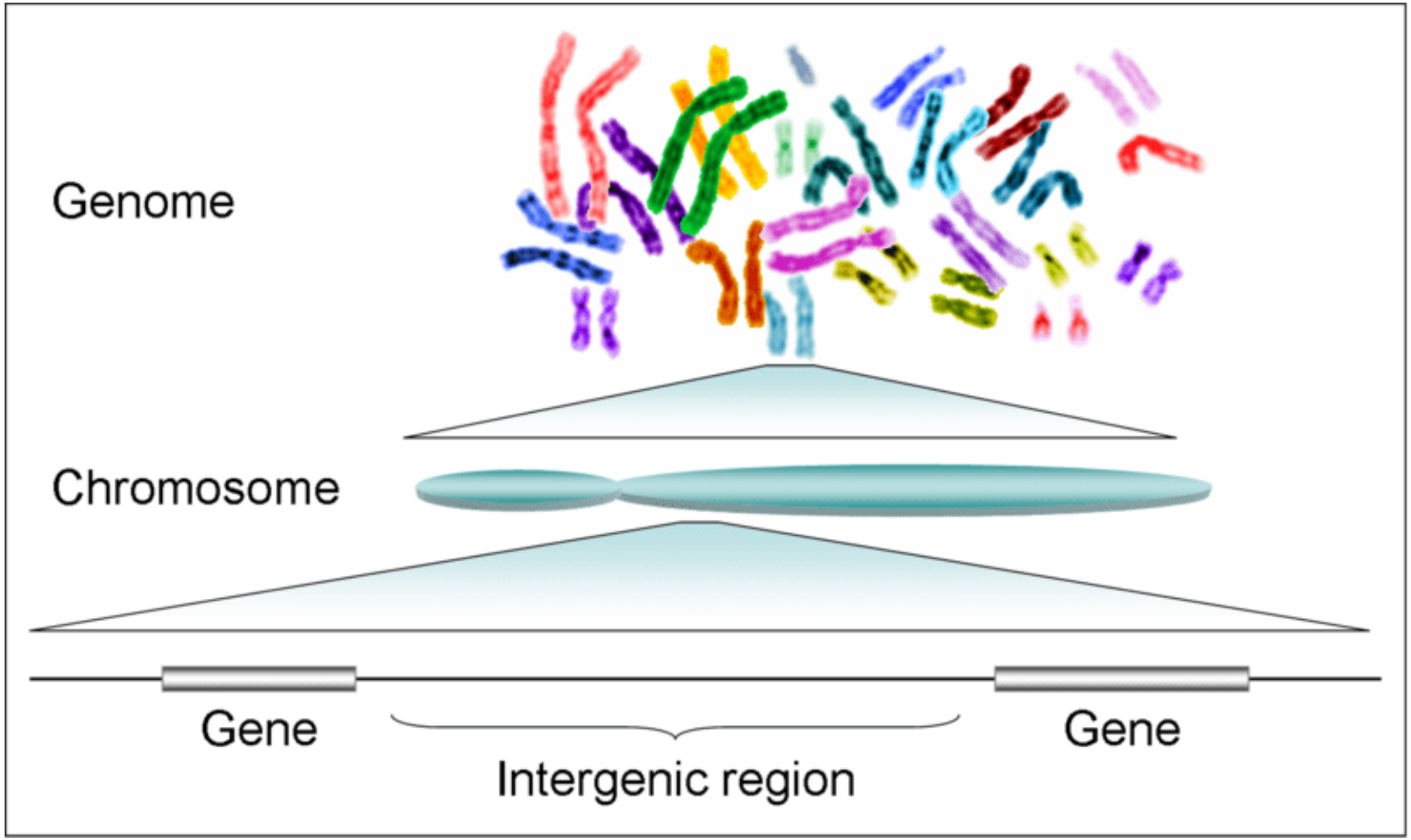


DNA condensation in eukaryotes





From Genomes to Genes



Terminology

Gene: is a sequence of nucleotides in DNA that is transcribed to produce a functional RNA.

Promoter: is a sequence of DNA to which proteins bind to initiate transcription of a single RNA transcript from the DNA downstream of the promoter.

Operator (prokaryotes): is a place of an operon where activators or repressors bind to regulate transcription.

Enhancer/Silencer (eukaryotes): is a short (50–1500 bp) region of DNA that can be bound by proteins (activators or repressors) to increase or decrease the likelihood that transcription of a particular gene will occur.

Terminator: is a section of nucleic acid sequence that marks the end of a gene or operon in genomic DNA during transcription.

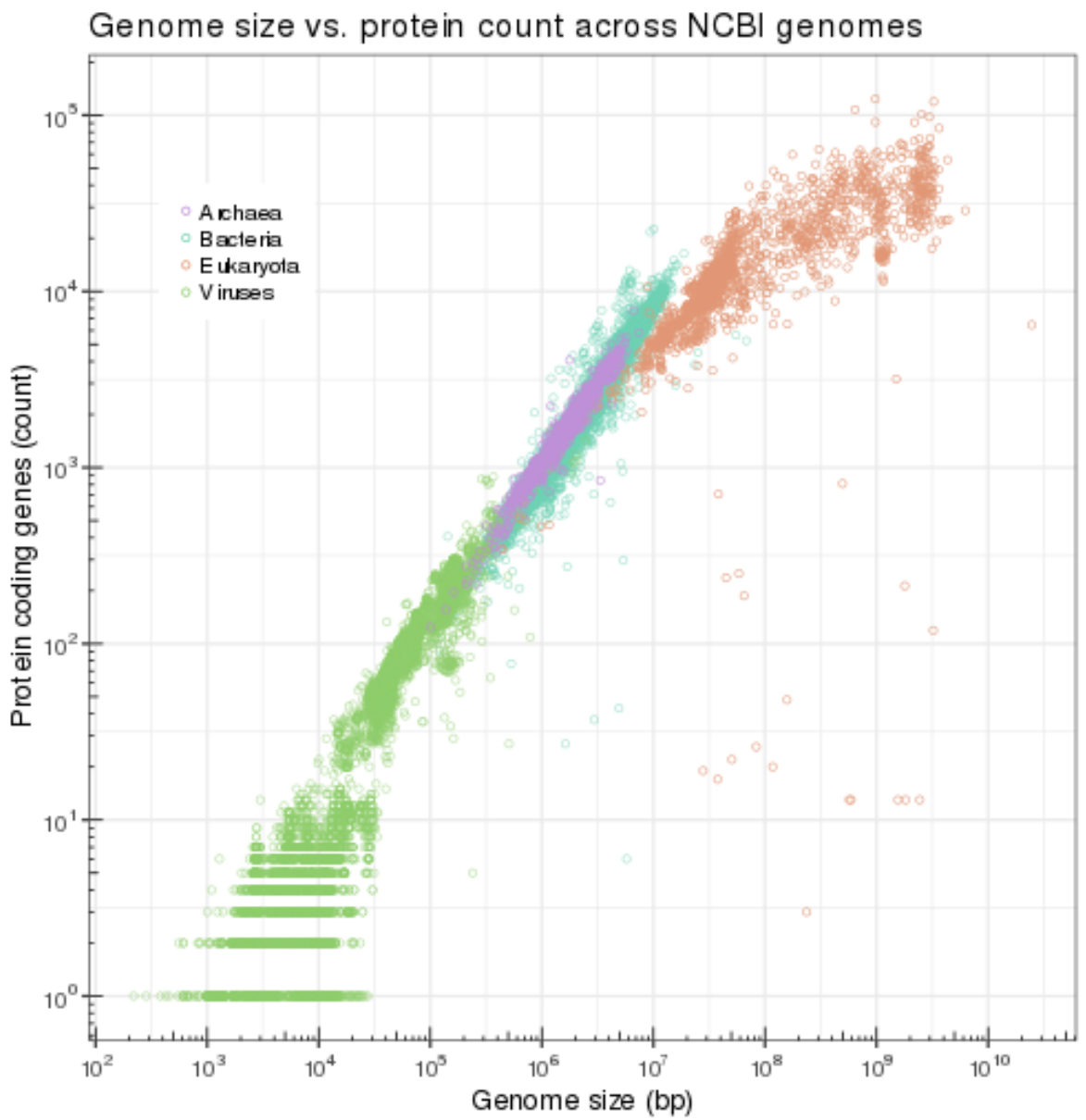
5’ / 3’ UTR: refers to either of two sections, one on each side of a coding sequence on a strand of mRNA

Exon: is any nucleotide sequence within a gene that is expressed or operative in the final RNA product.

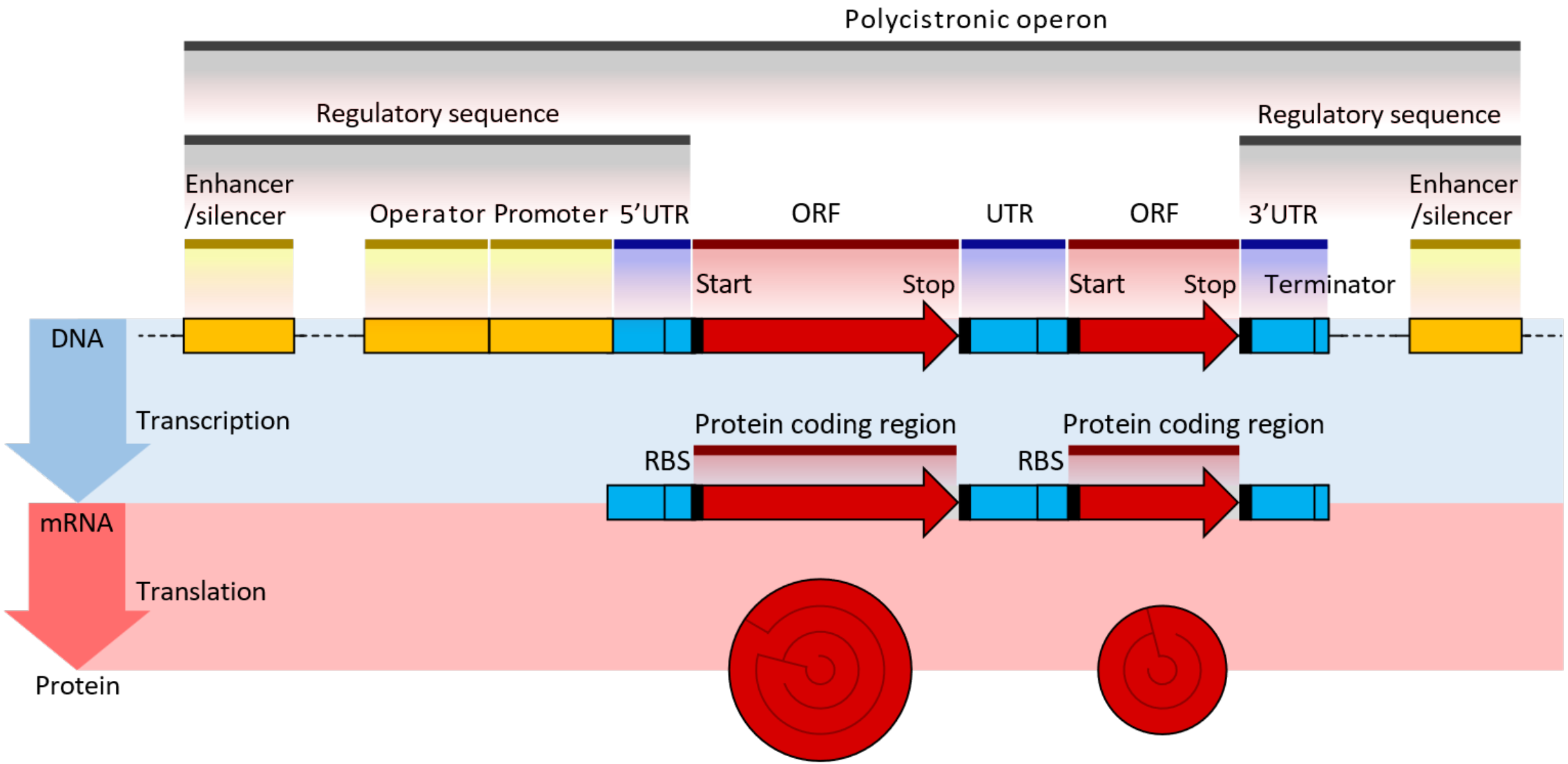
Intron: is any nucleotide sequence within a gene that is not expressed or operative in the final RNA product.

Poly-A tail: is a stretch of RNA that has only adenine bases at the 3’ UTR end of mRNA

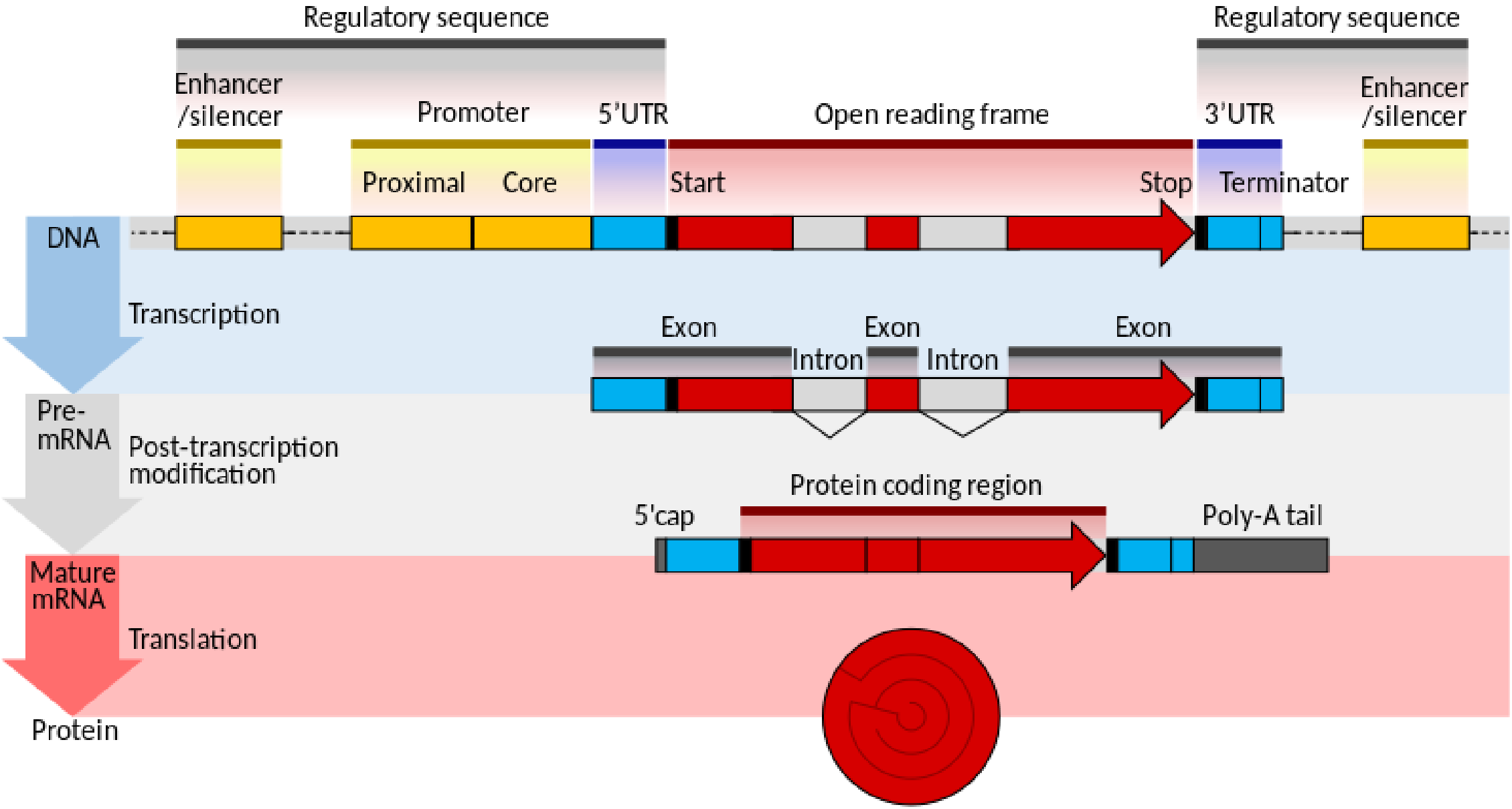
Genome sizes



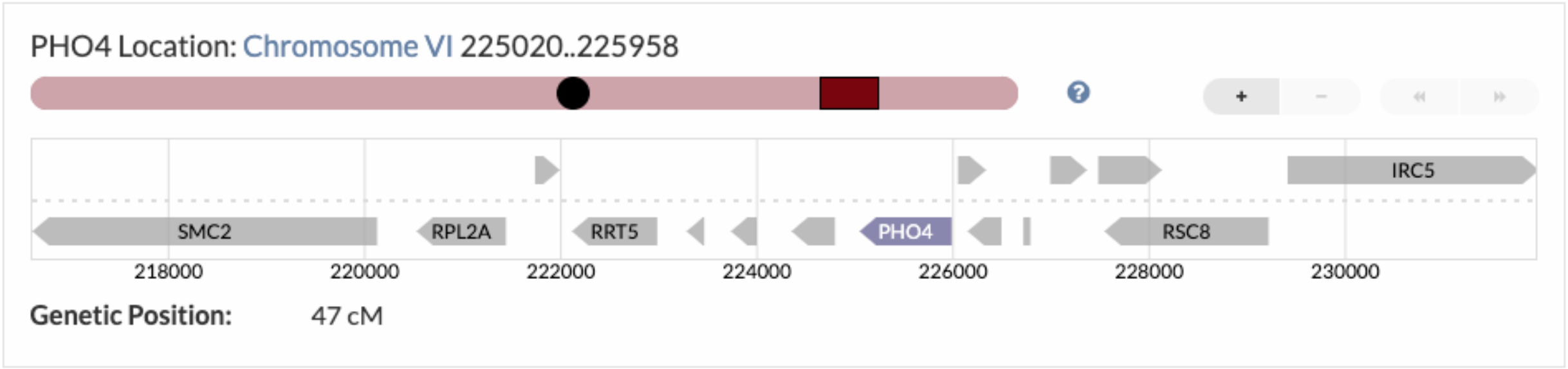
Genomic gene and regulatory structure (prokaryotes)



Genomic gene and regulatory structure (eukaryotes)



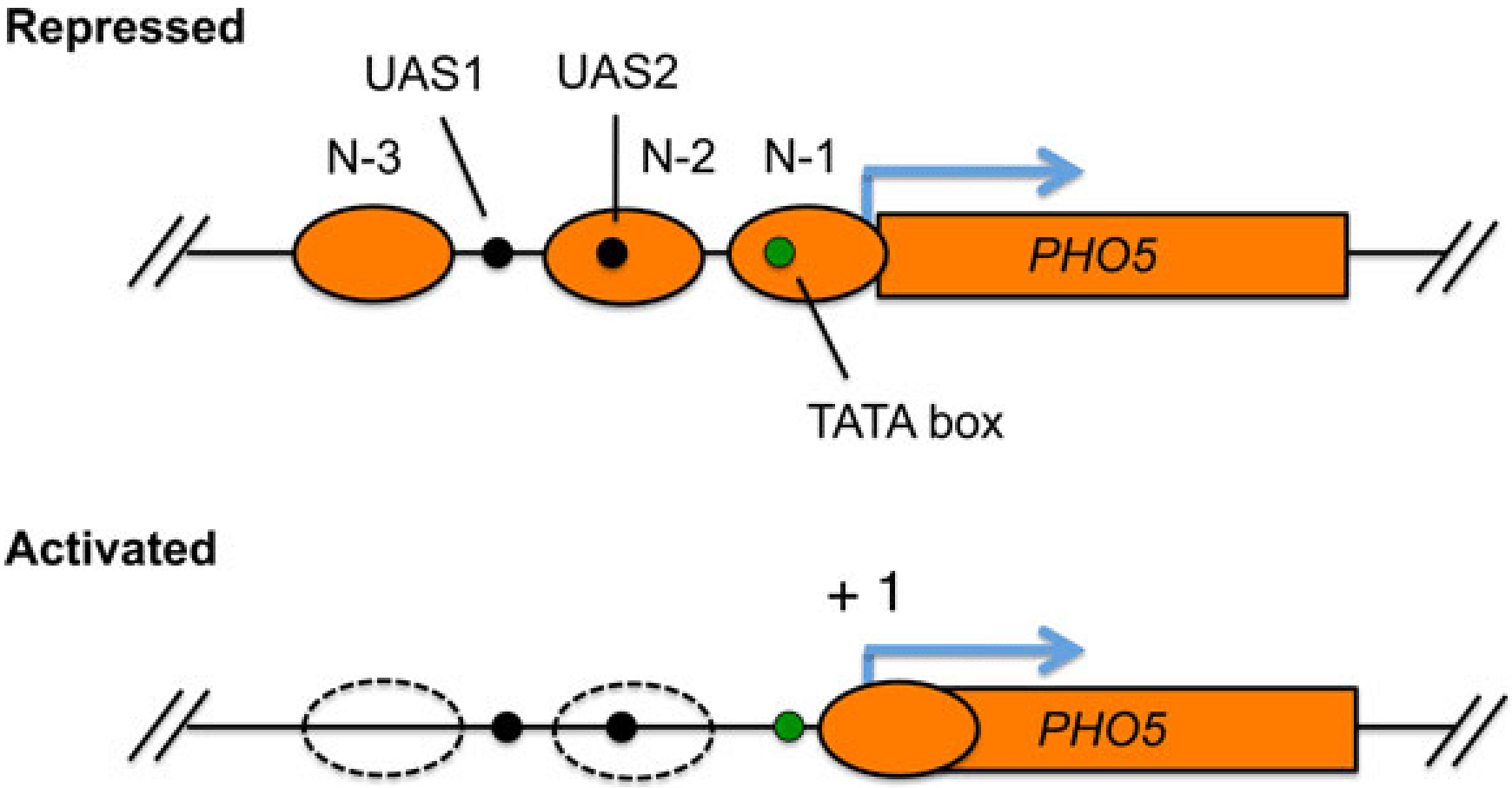
Genes, Promoters, Terminators, Gene regulation



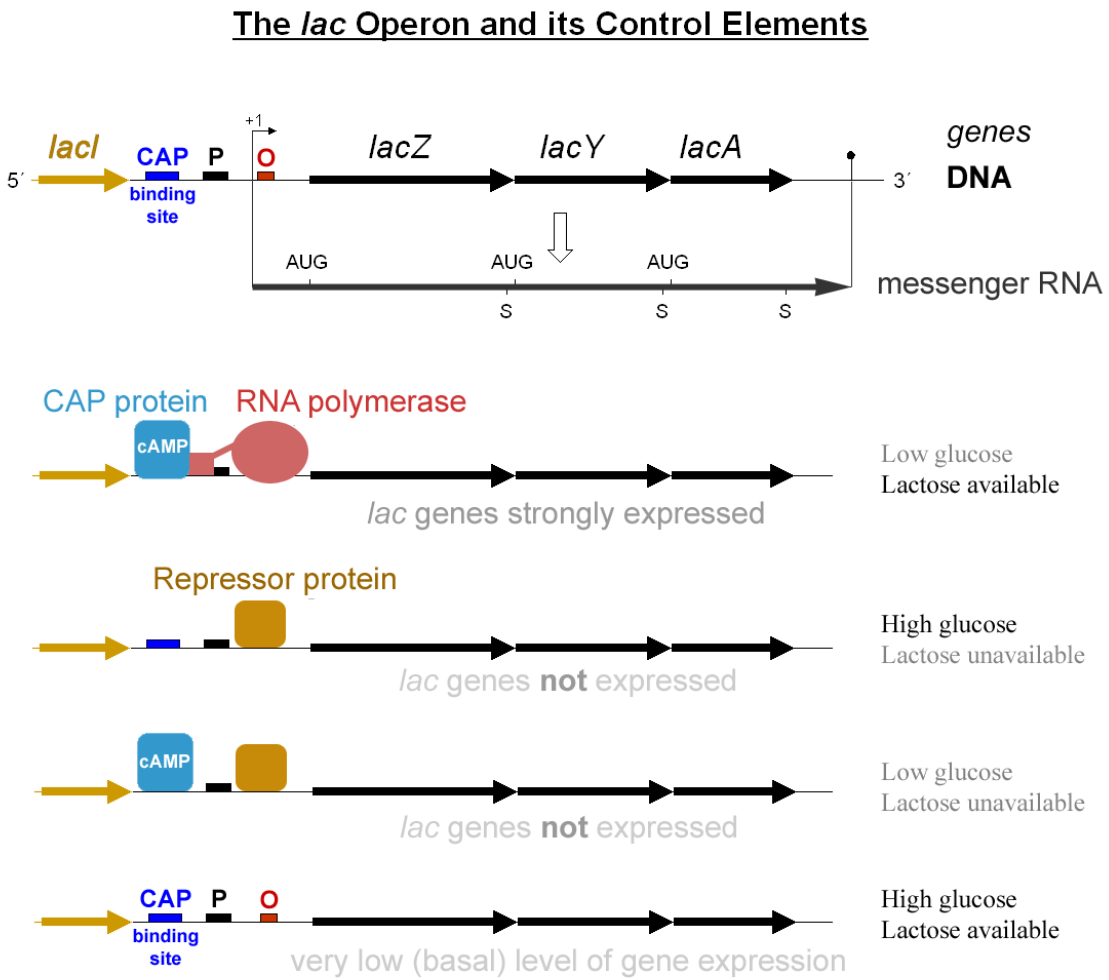
Yeast Genome Database Live Demonstration

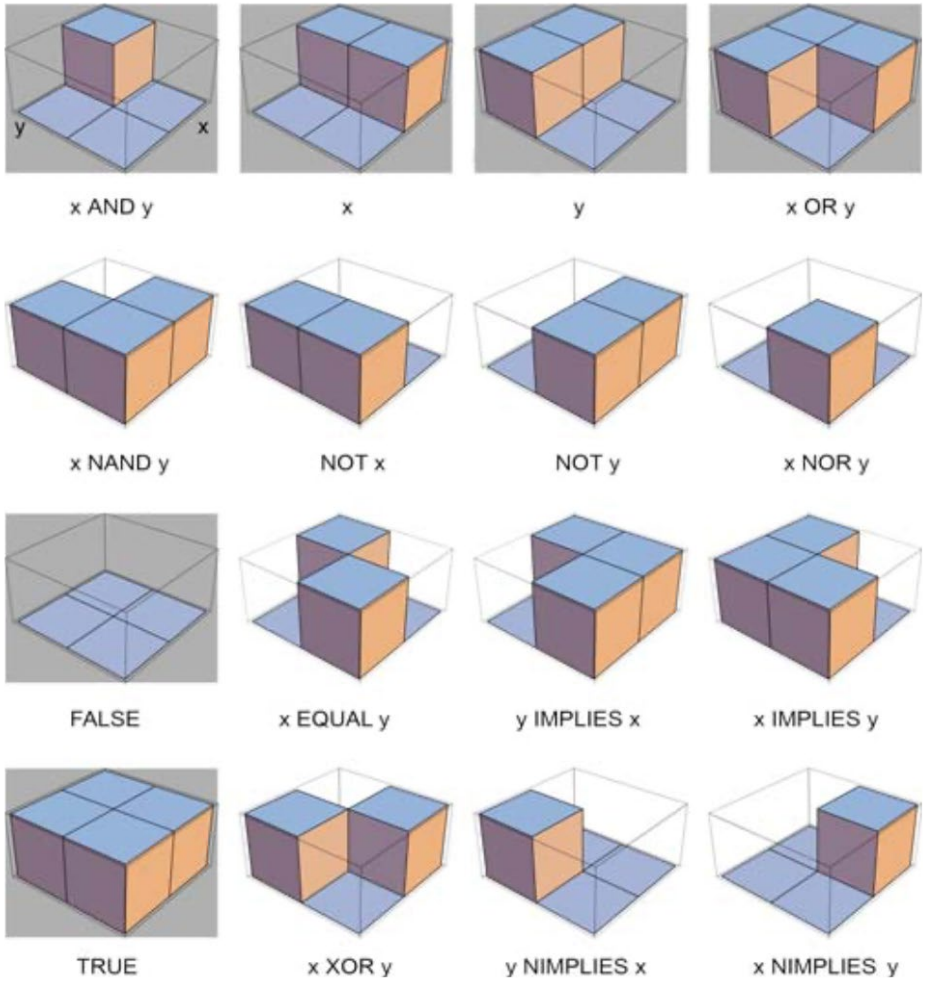
https://www.yeastgenome.org/

Pho5 Promoter Example



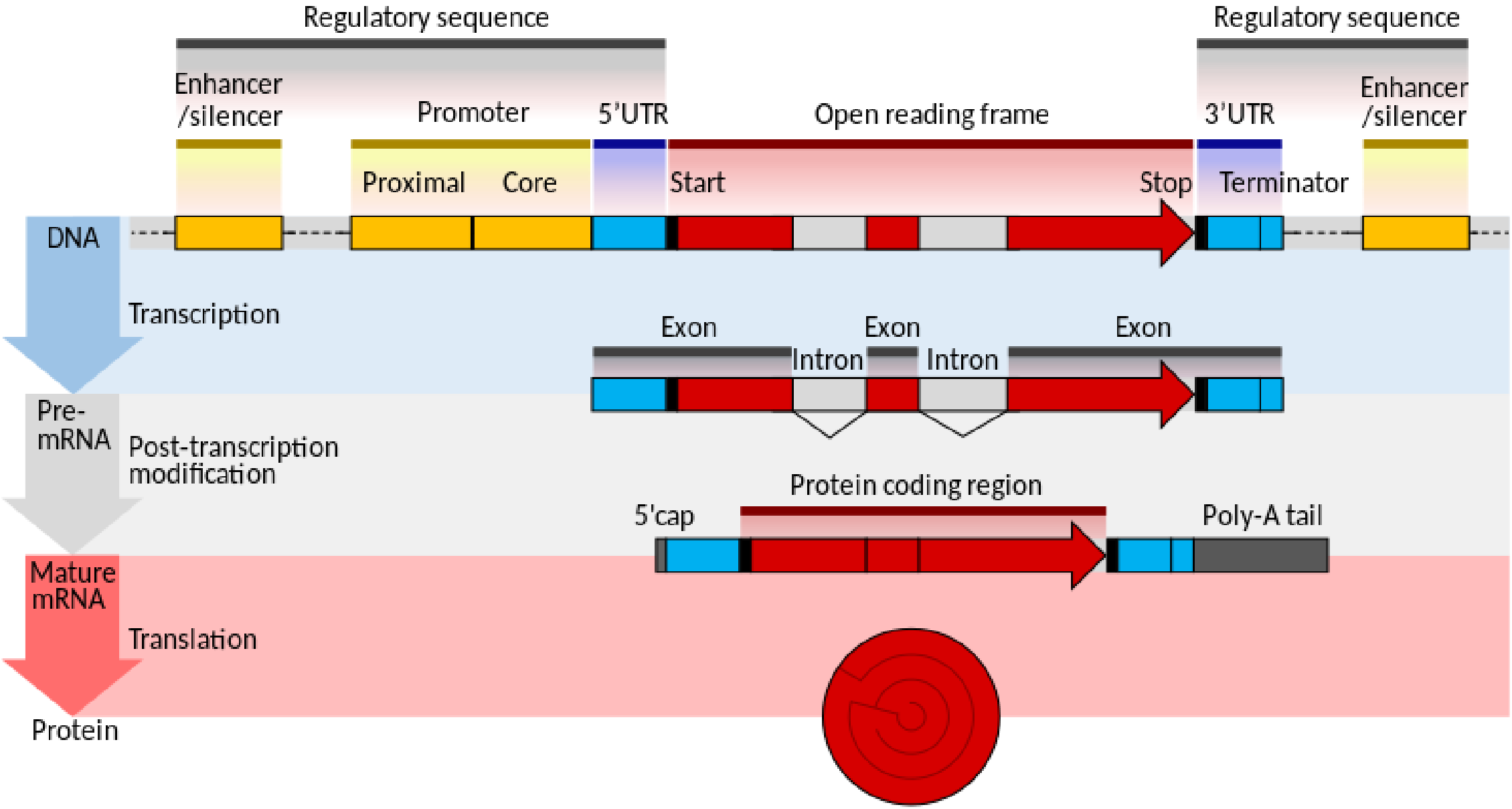
Genes, Promoters, Terminators, Gene regulation





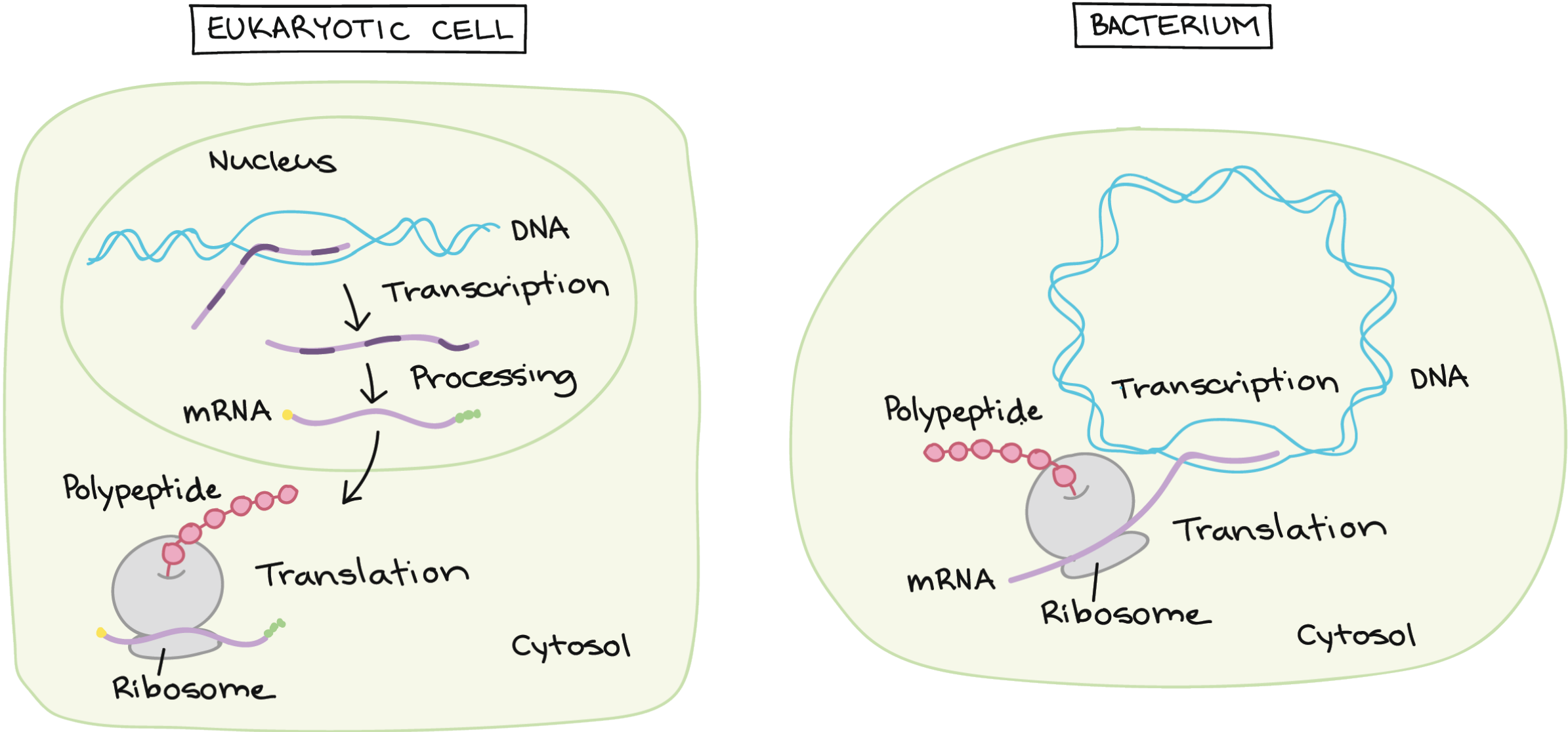
RNA Structure and Function

Genomic gene and regulatory structure (eukaryotes)

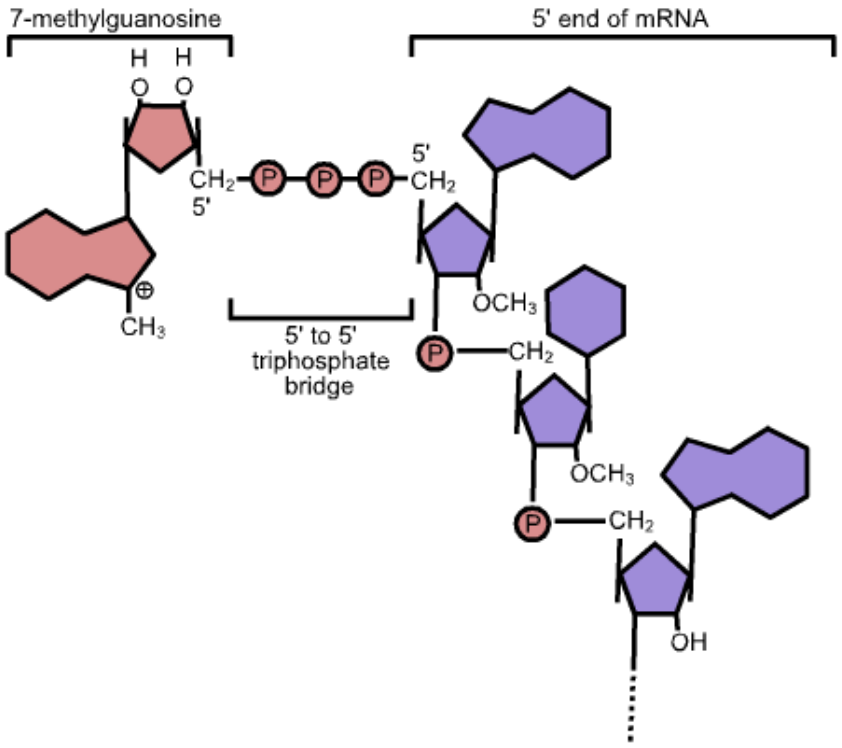


mRNA

mRNA (messenger RNA)



mRNA 5’cap



5’ Cap functions:

1. **Regulation of nuclear export**
2. **Prevention of degradation by exonucleases**
3. **Promotion of translation**
4. **Promotion of 5’ proximal intron excision**

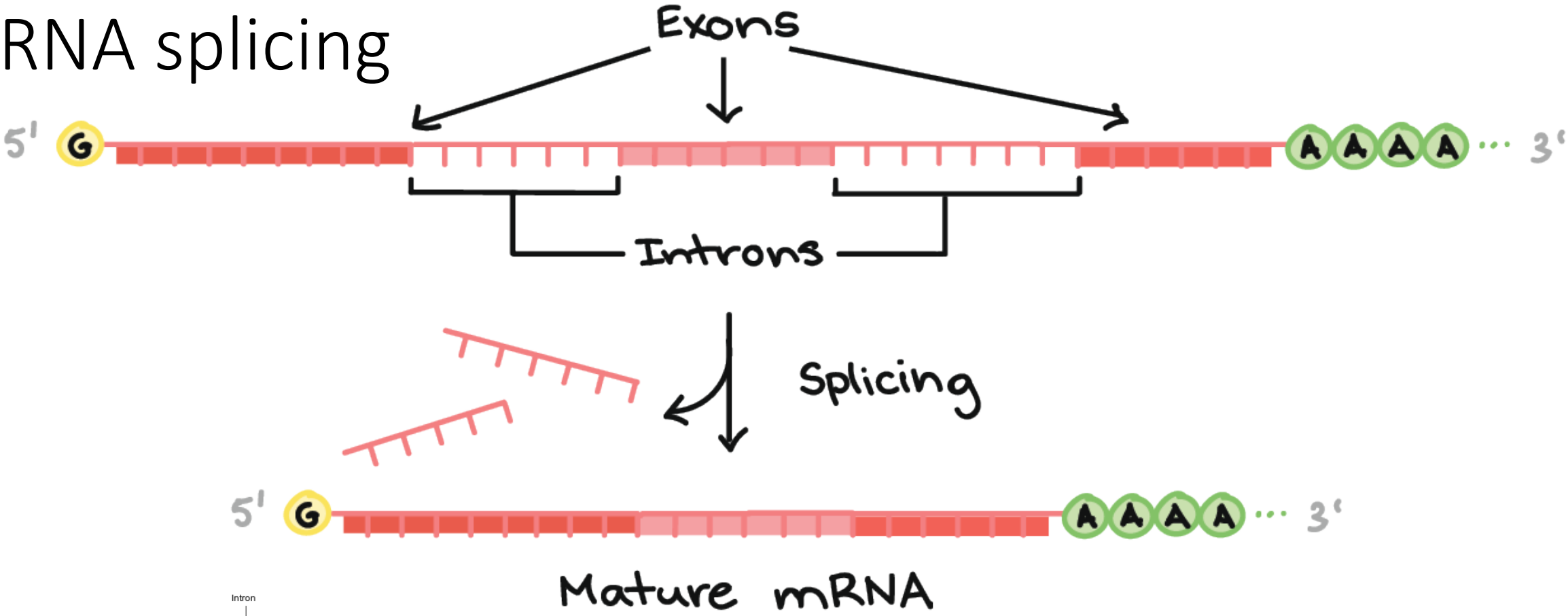
1. One of the terminal phosphate groups is removed by RNA triphosphatase, leaving a bisphosphate group (i.e. 5′(ppN)[pN]$_n$);
2. GTP is added to the terminal bisphosphate by mRNA guanylyltransferase, losing a pyrophosphate from the GTP substrate in the process. This results in the 5′–5′ triphosphate linkage, producing 5′(Gp)(ppN)[pN]$_n$;
3. The 7-nitrogen of guanine is methylated by mRNA (guanine-*N*7-)-methyltransferase, with *S*-adenosyl-L-methionine being demethylated to produce *S*-adenosyl-L-homocysteine, resulting in 5′(m7Gp)(ppN)[pN]$_n$ (cap-0);
4. Cap-adjacent modifications can occur, normally to the first and second nucleotides, producing up to 5′(m7Gp)(ppN*)(pN*)[pN]$_n$ (cap-1 and cap-2);[7]
5. If the nearest cap-adjacent nucleotide is 2′-*O*-ribose methyl-adenosine (i.e. 5′(m7Gp)(ppAm)[pN]$_n$), it can be further methylated at the N6 methyl position to form N^6-methyladenosine, resulting in 5′(m7Gp)(ppm6Am)[pN]$_n$.[3]

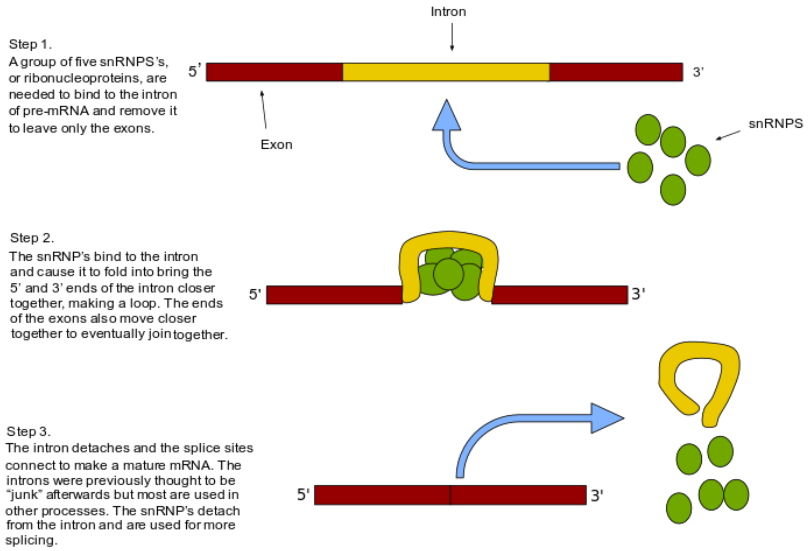
poly-A tail

Poly-A Tail:

- **Enzymatic (post-transcriptional)** additon of 20-300 adenine nucleotides
- helps inhibit degradation of the mRNA
- important for nuclear export

mRNA splicing





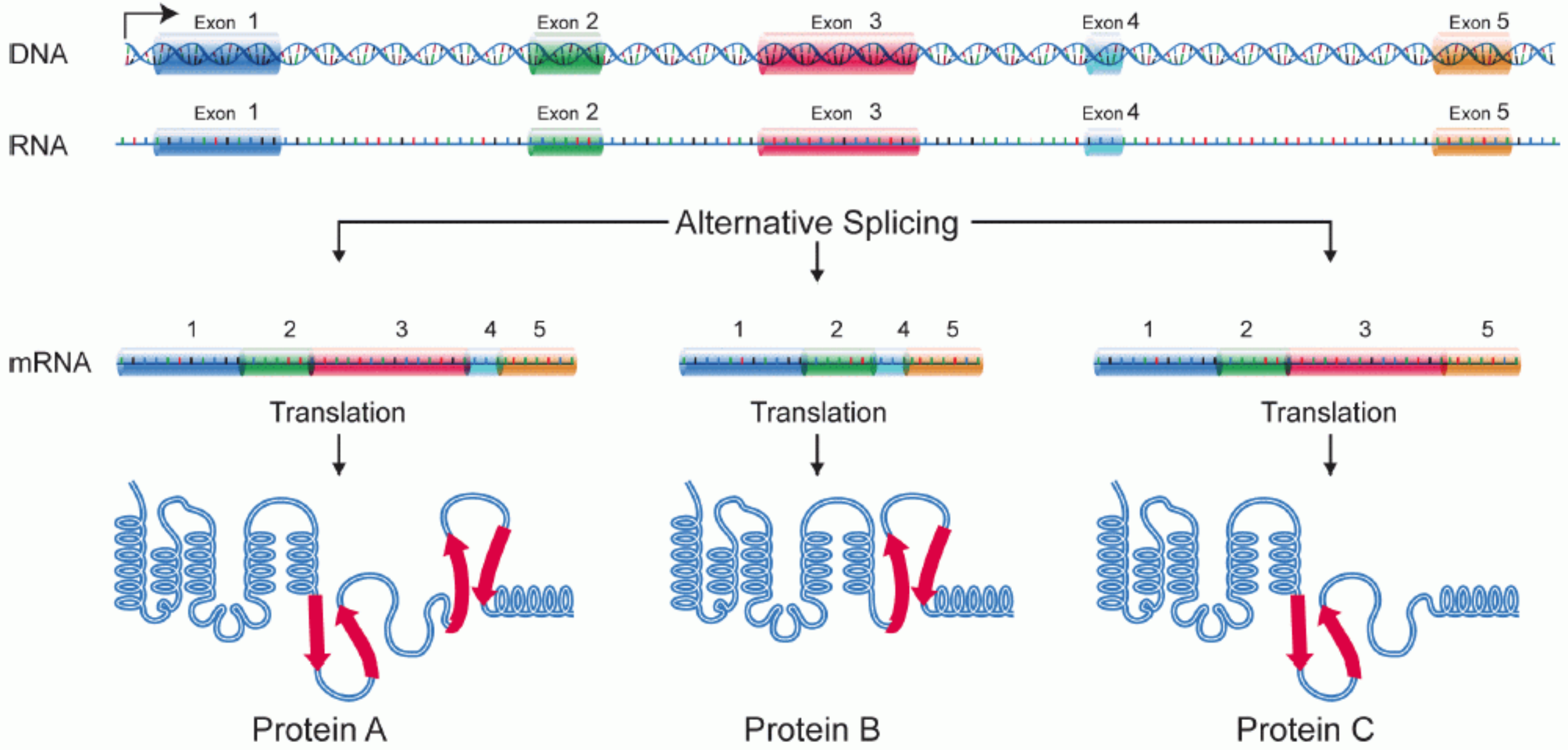
Spliceosome:

5 small nuclear RNAs (snRNAs)

+ proteins

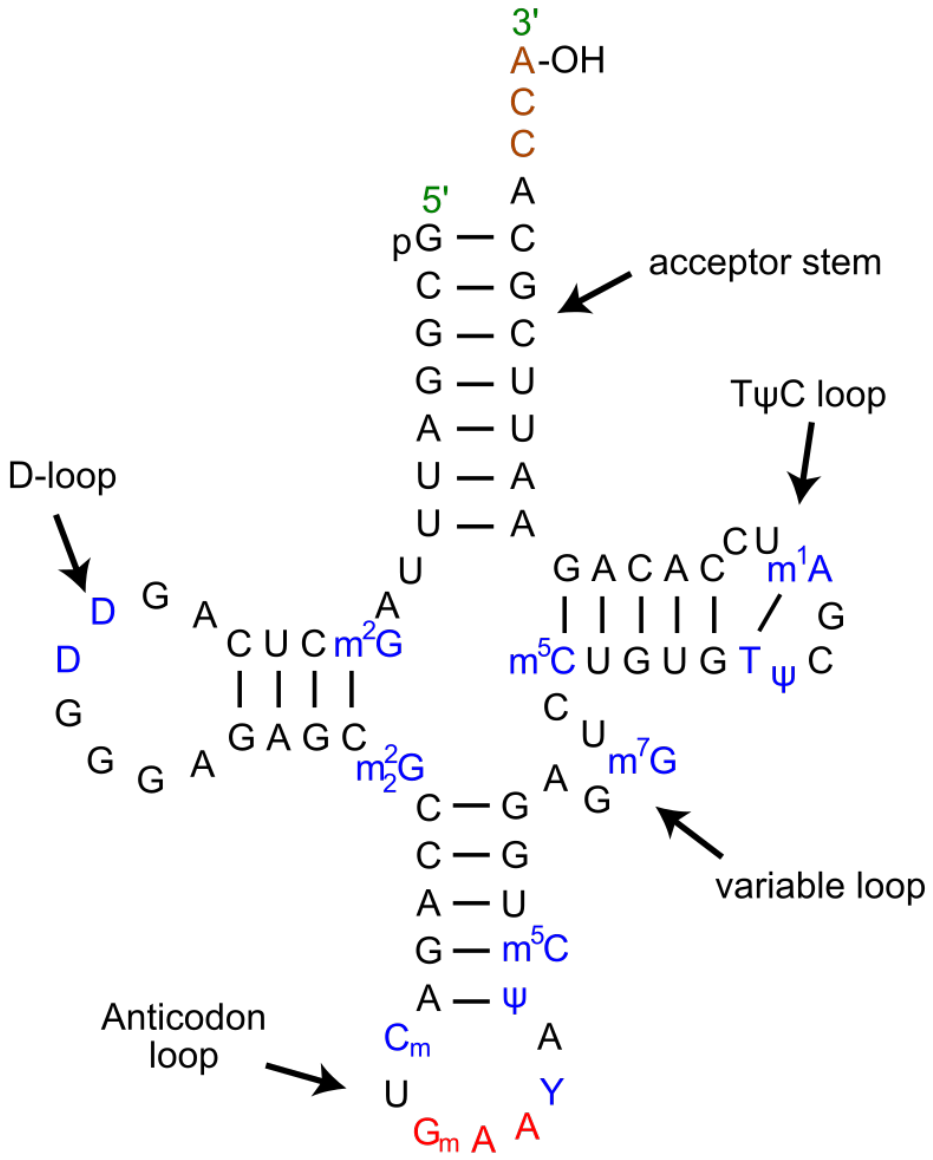
= small nuclear ribonucleoprotein complex (snRNP)

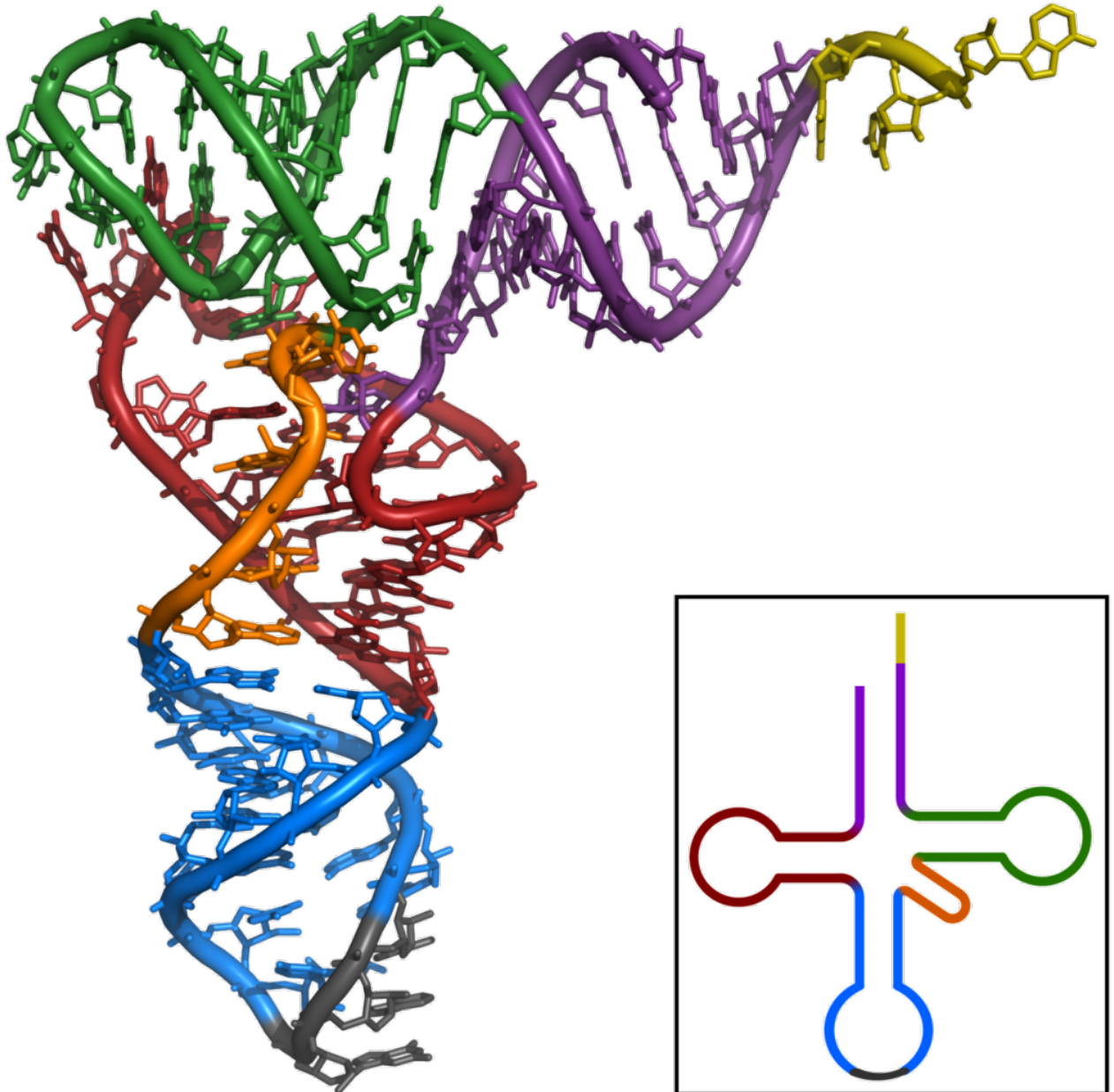
mRNA alternative splicing



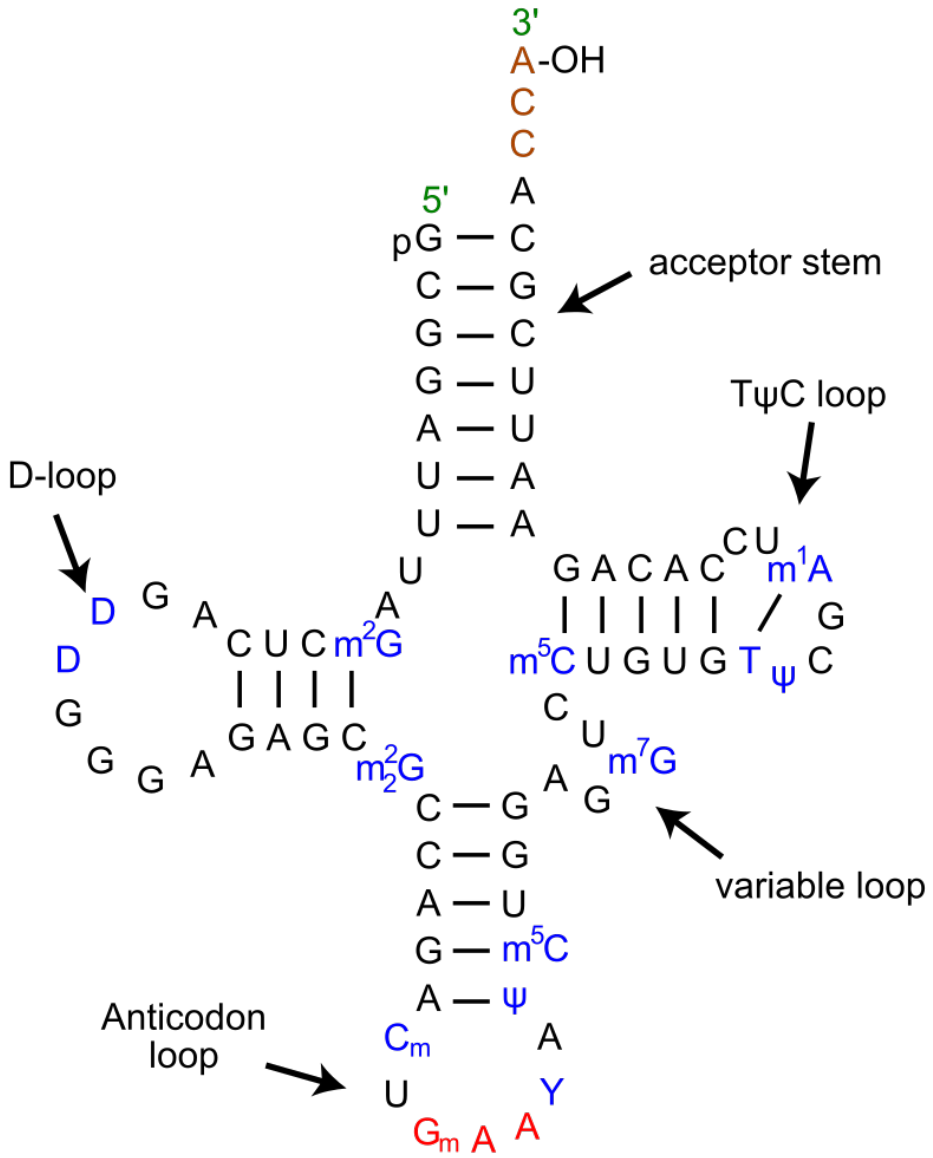
tRNA

tRNA (transfer RNA)



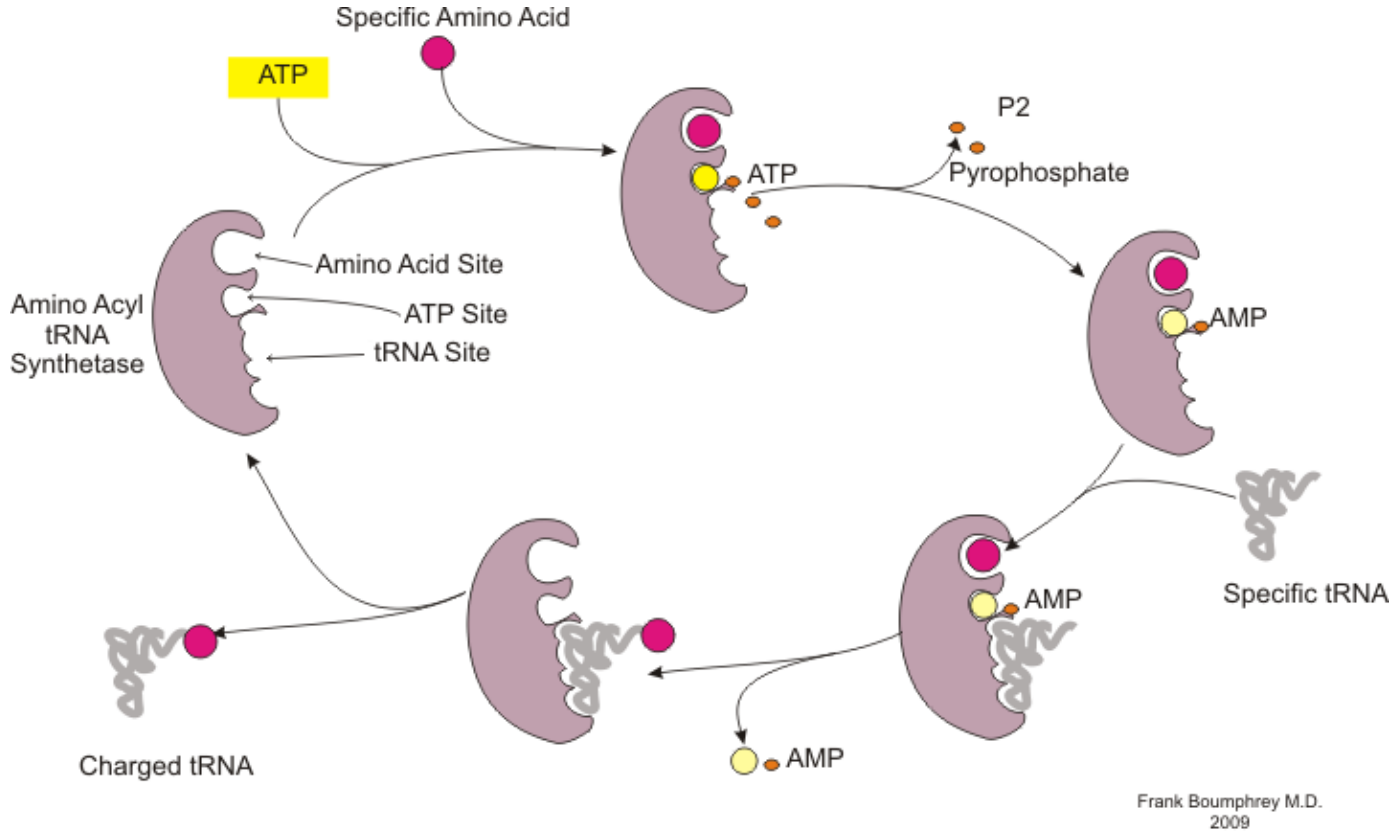


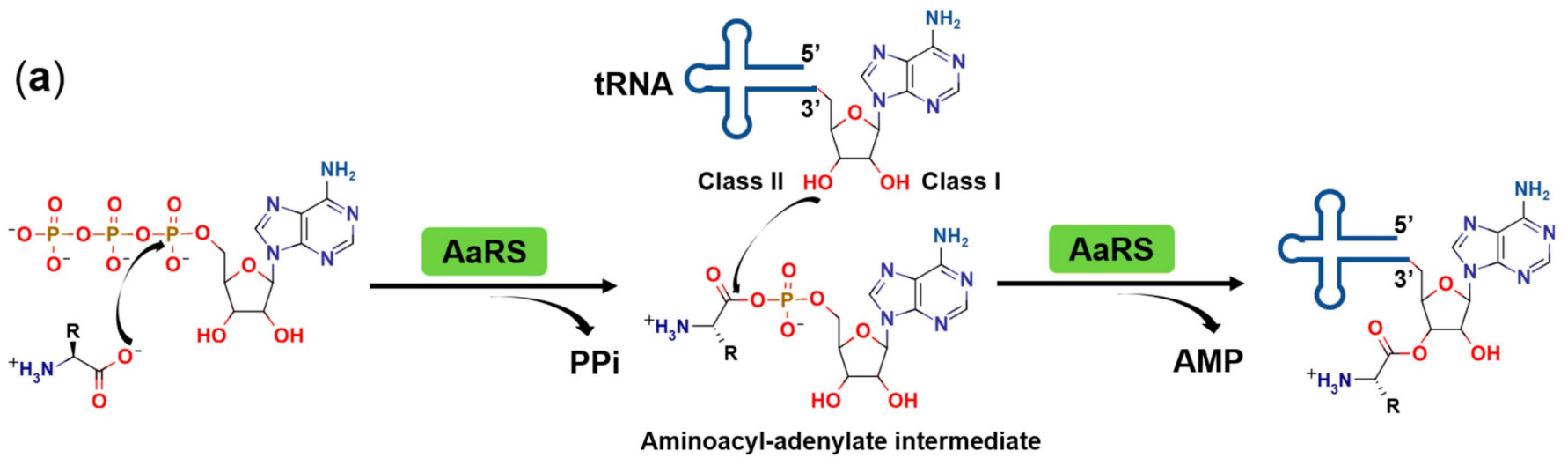
tRNA (transfer RNA)



- The **acceptor stem** is a 7- to 9-base pair (bp) stem made by the base pairing of the 5′-terminal nucleotide with the 3′-terminal nucleotide (which contains the CCA 3′-terminal group used to attach the amino acid). In general, such 3′-terminal tRNA-like structures are referred to as 'genomic tags'. The acceptor stem may contain non-Watson-Crick base pairs.[7][9]
- The **CCA tail** is a cytosine-cytosine-adenine sequence at the 3′ end of the tRNA molecule. The amino acid loaded onto the tRNA by aminoacyl tRNA synthetases, to form aminoacyl-tRNA, is covalently bonded to the 3′-hydroxyl group on the CCA tail.[10] This sequence is important for the recognition of tRNA by enzymes and critical in translation.[11][12] In prokaryotes, the CCA sequence is transcribed in some tRNA sequences. In most prokaryotic tRNAs and eukaryotic tRNAs, the CCA sequence is added during processing and therefore does not appear in the tRNA gene.[13]
- The **D loop** is a 4- to 6-bp stem ending in a loop that often contains dihydrouridine.[7]
- The **anticodon loop** is a 5-bp stem whose loop contains the anticodon.[7] The tRNA 5′-to-3′ primary structure contains the anticodon but in reverse order, since 3′-to-5′ directionality is required to read the mRNA from 5′-to-3′.
- The **TΨC loop** is named so because of the characteristic presence of the unusual base Ψ in the loop, where Ψ is pseudouridine, a modified uridine. The modified base is often found within the sequence 5'-TΨCGA-3', with the T (ribothymidine, m5U) and A forming a base pair.[14]
- The **variable loop** sits between the anticodon loop and the ΨU loop and, as its name implies, varies in size from 3 to 21 bases.[15]

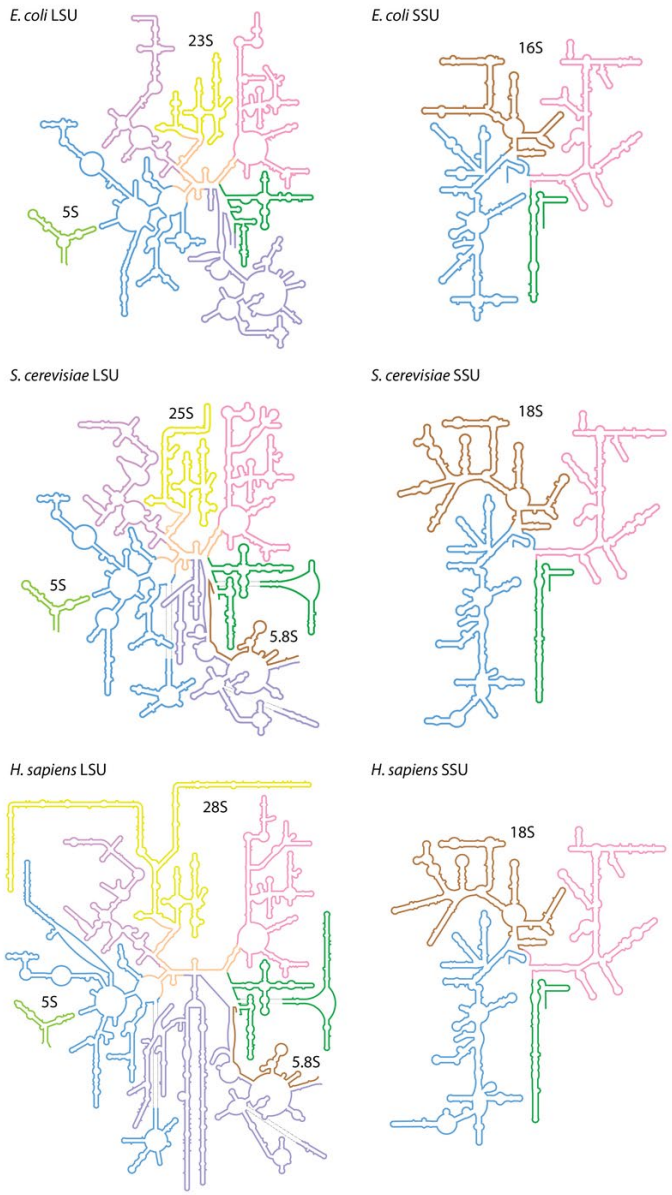
tRNA (transfer RNA)



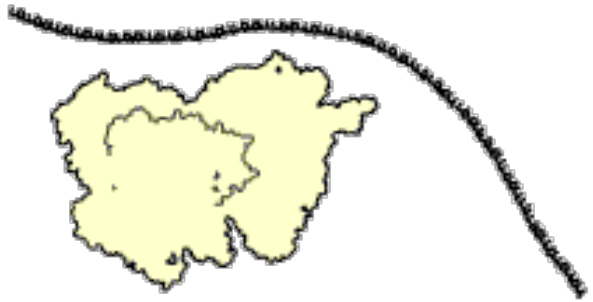


rRNA

rRNA (ribosomal RNA)



Ribosomal ribonucleic acid (**rRNA**) is a type of non-coding RNA which is the primary component of ribosomes, essential to all cells. rRNA is a ribozyme which carries out protein synthesis in ribosomes. Ribosomal RNA is transcribed from ribosomal DNA (rDNA) and then bound to ribosomal proteins to form small and large ribosome subunits. rRNA is the physical and mechanical factor of the ribosome that forces transfer RNA (tRNA) and messenger RNA (mRNA) to process and translate the latter into proteins.[1] Ribosomal RNA is the predominant form of RNA found in most cells; it makes up about 80% of cellular RNA despite never being translated into proteins itself. Ribosomes are composed of approximately 60% rRNA and 40% ribosomal proteins by mass.



miRNA