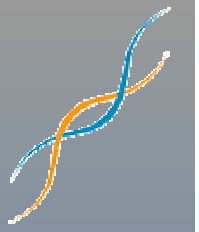


Computational Biology

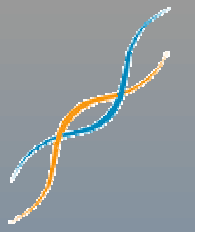
Molecular Biology
Basic Concepts

Prof. Dr. Antônio Augusto Fröhlich
Charles Ivan Wust

LISHA - UFSC

{guto|charles}@lisha.ufsc.br
http://www.lisha.ufsc.br/~{guto|charles}

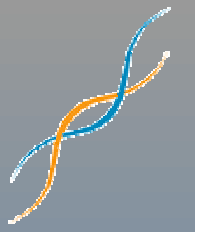
September 2003



Life

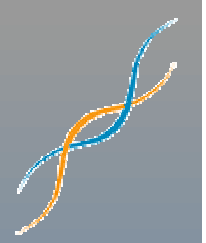
- Living things
  - Active participation;
  - Array of chemical reactions that never cease.
- Started about 3.5 million years ago.
- Composed of similar molecular chemestry (biochemestry)
  - Proteins
  - Nucleic acids

"We *are* our proteins." (Russel Doolittle)



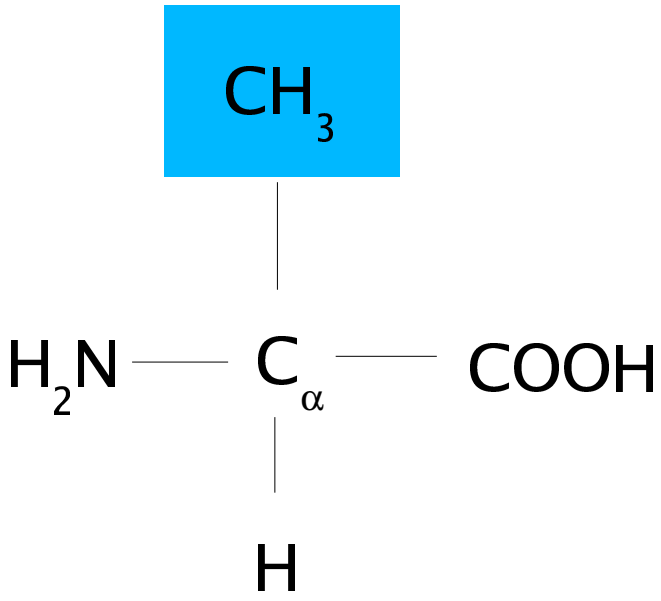
Proteins

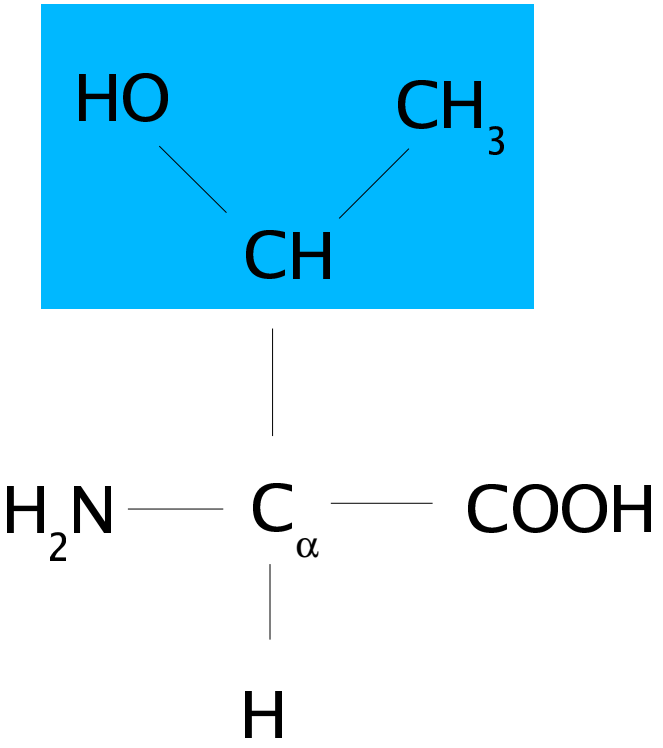
- Several functions
  - Structural: tissues building blocks
  - Enzymes: catalyst of chemical reactions
  - Other: oxygen transport, antibody defense, ...
- Chain of aminoacids



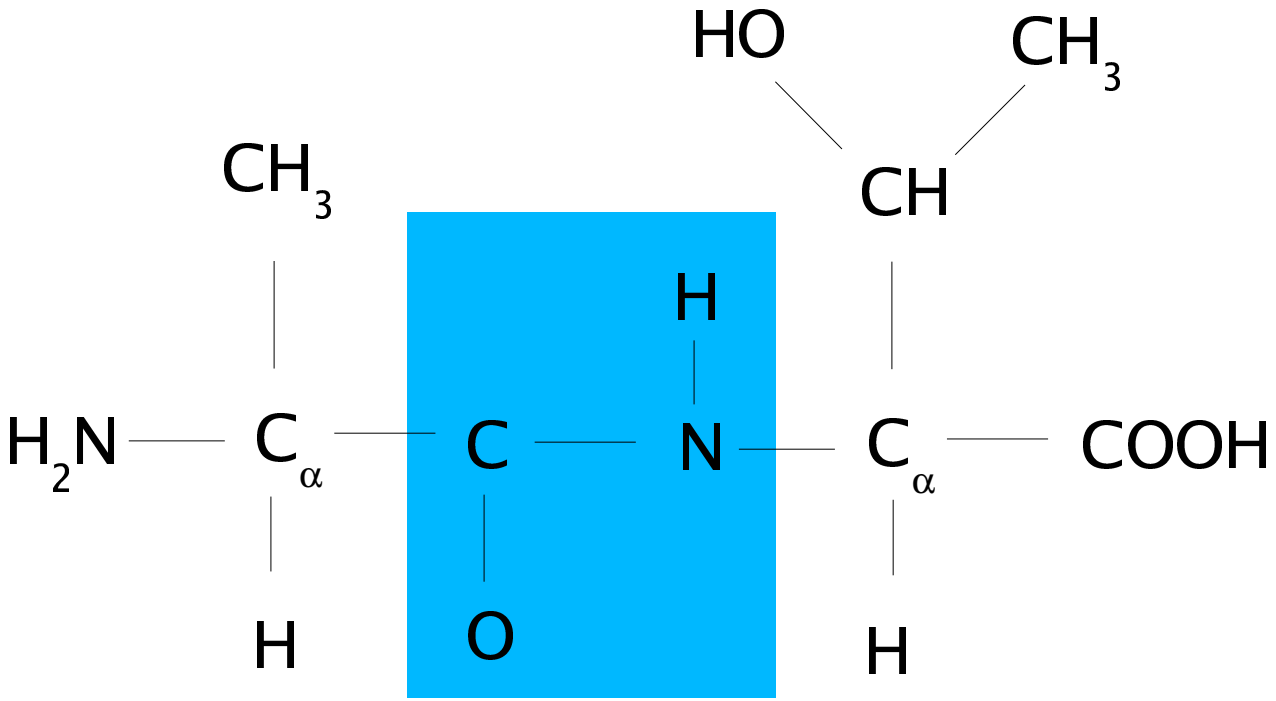
Computational Biology

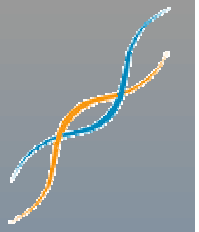
Amino Acid Examples





Peptide Bonds



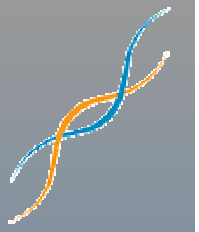


Computational Biology

The 20 amino acids

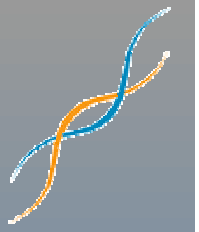
	1-letter code	*3-letter code*	*Name*
1	A	Ala	Alanine
2	C	Cys	Cysteine
3	D	Asp	Aspartic Acid
4	E	Glu	Glutamic Acid
5	F	Phe	Phenylalanine
6	G	Gly	Glycine
7	H	His	Histidine
8	I	Ile	Isoleucine
9	K	Lys	Lysine
10	L	Leu	Leucine

	1-letter code	*3-letter code*	*Name*
11	M	Met	Methionine
12	N	Asn	Asparagine
13	P	Pro	Proline
14	Q	Gln	Glutamine
15	R	Arg	Arginine
16	S	Ser	Serine
17	T	Thr	Threonine
18	V	Val	Valine
19	W	Trp	Tryptophan
20	Y	Tyr	Tyrosine



Protein Structure

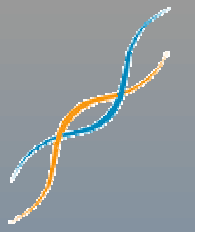
- Fold in 3 dimensions
  - *Primary structure*: Sequence of residues;
  - *Secondary structure*: “Local” structures (i.e. helices)
  - *Tertiary structure*: secondary structure on a global level, the overral 3-D structure of single proteins.
  - *Quartiary structure*: Product of the in interaction between different proteins.



Computational Biology

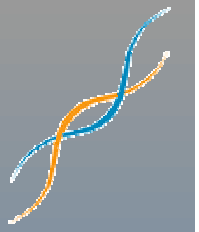
Protein 3-D Structure

http://www.the-scientist.com/images/yr2002/apr29/



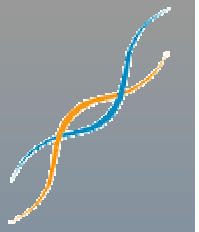
Nucleic Acids

- Two kinds
  - Ribonucleic acid (RNA)
  - Desoxyribonucleic acid (DNA)



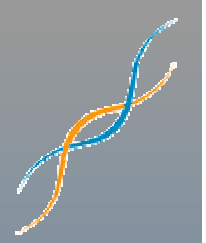
DNA

- Double chain of simpler molecules
  - Free-floating in lower organisms (prokariotes)
  - Inside the nucleus in higher organisms (euchariotes), and in mitochondria and chloroplasts (plants only)
- Single chain = strand
  - Repetition of nucleotides (sugar+phosphate+base)
  - Has an orientation (5' -> 3' is default)
  - 4 types of bases
    - (A)denin, (G)uanin, (C)ytosine and (T)imine



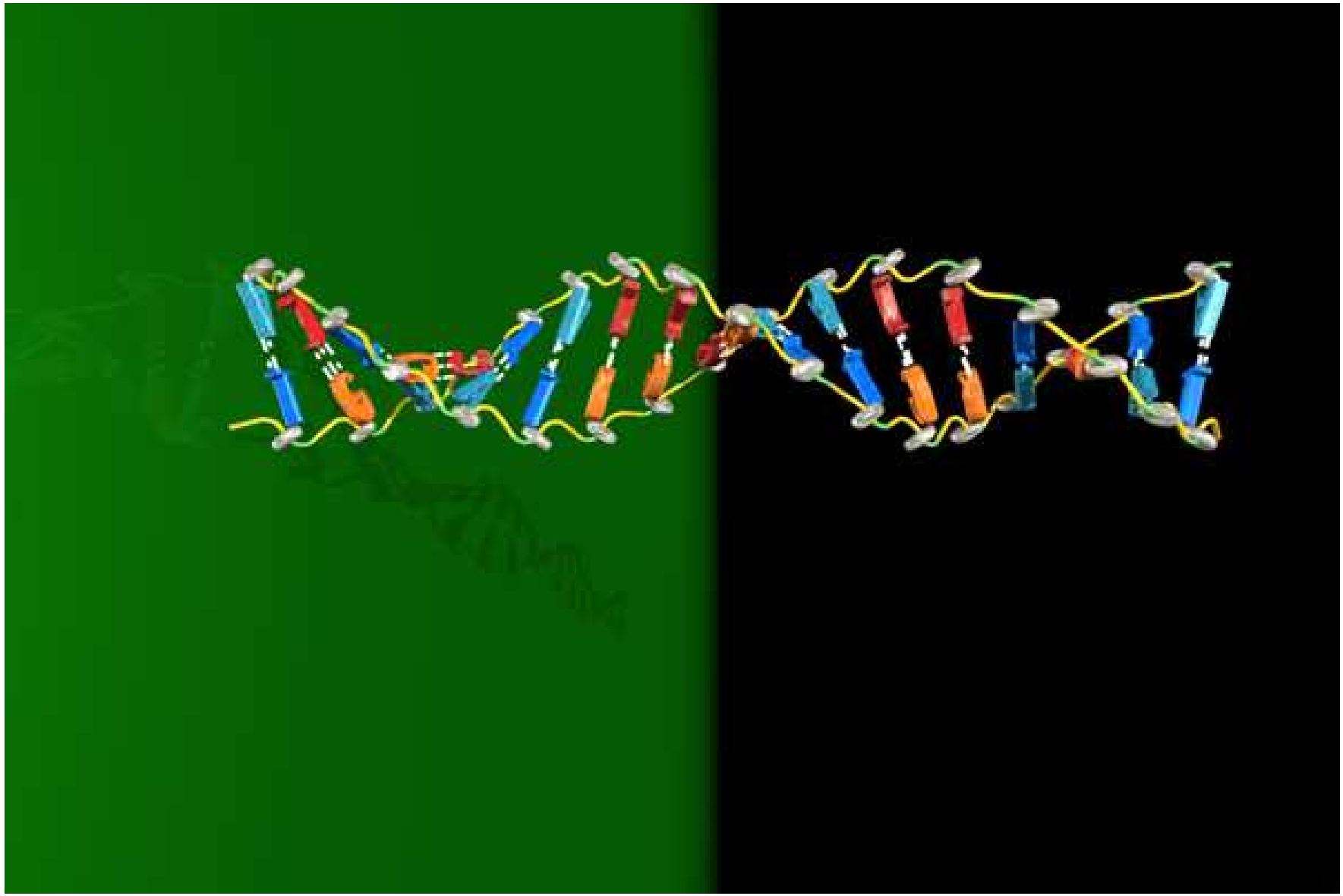
DNA

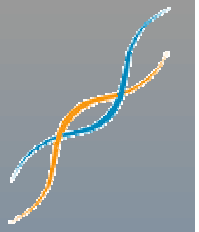
- Double strands
  - Duble helix structure discovered by Watson and Crick in 1953
- Each base in one strand is paired with a base in the other strand
  - Complementary bases (A + T, G + C)
- DNA lenth measured in *base pairs* (**bp**)
- Strands have reverse orientation (antiparallel)
- Each strand is the *reverse complement* of the oposite



Computational Biology

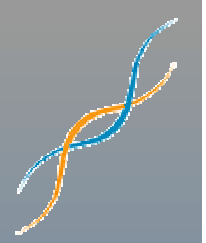
DNA



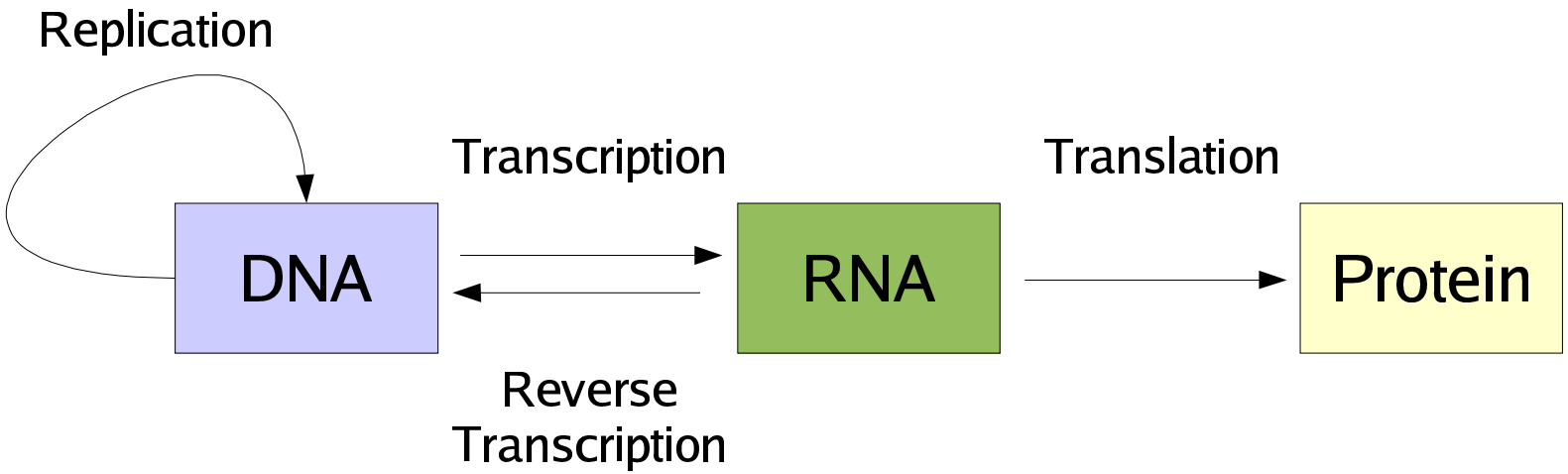


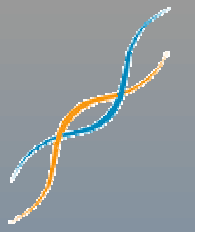
RNA

- Very similar to DNA
- Differences
  - Sugar is ribose, instead of 2'-deoxyribose
  - *(U)racil* instead of *(T)hymine*
  - Single stranded
  - Several different functions, instead of information encoding



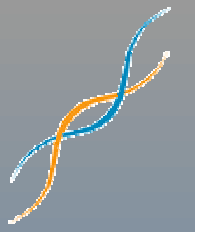
Central Dogma of Molecular Biology





Transcription

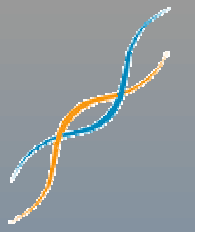
- Certain stretches of DNA encode information for building proteins, but not all of it.
- Each kind of protein usually correspond to one and only one stretch in the DNA: a ***gene***
- Sequence of amino acids in a protein are specified by triplets of nucleotides (codons), according to the genetic code.



Computational Biology

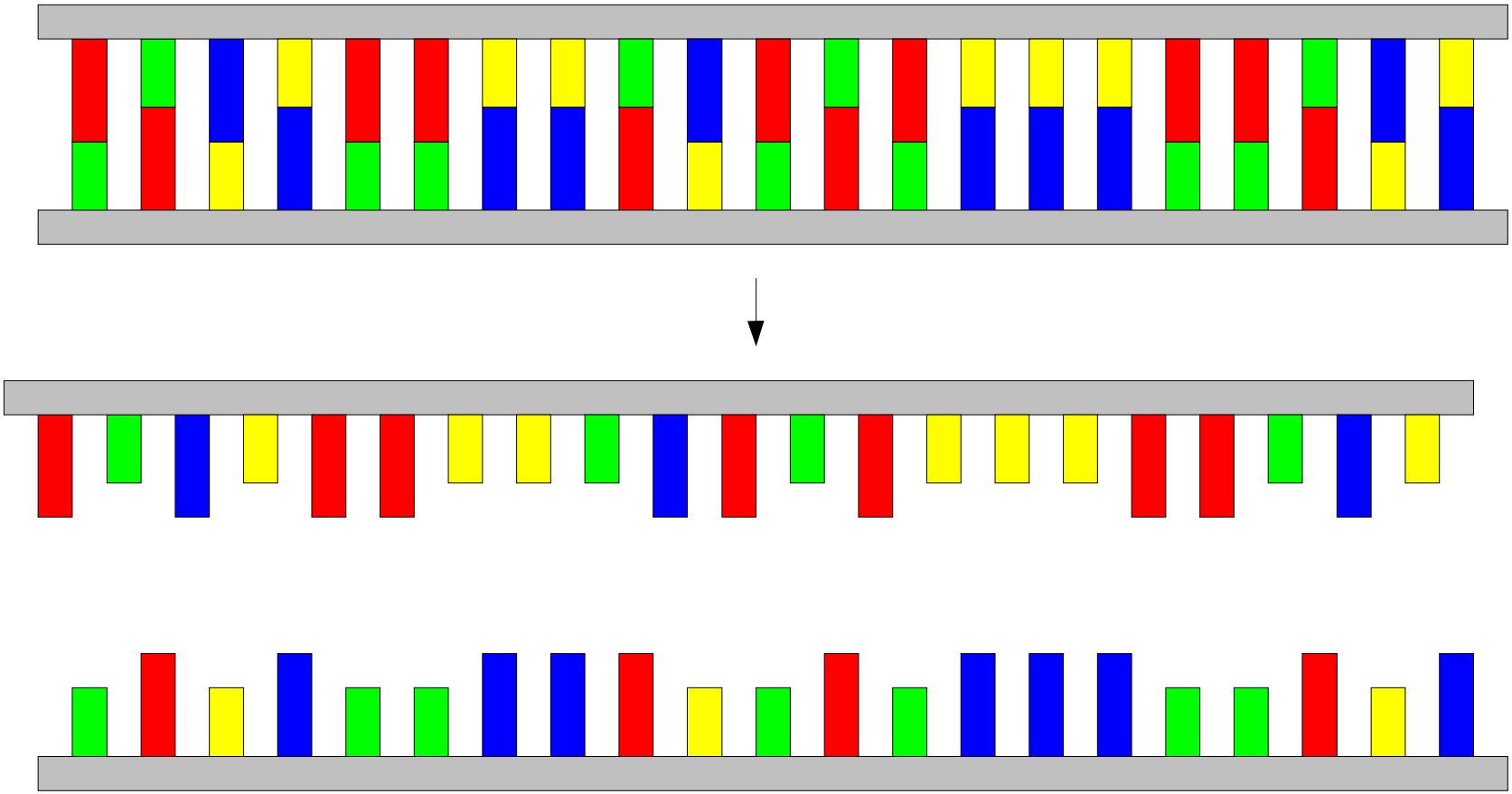
The genetic code

First pos.	*Second Pos.*	*Third Pos.*			
		G	**A**	**C**	**T**
G	**G**	Gly	Gly	Gly	Gly
	A	Glu	Glu	Asp	Asp
	C	Ala	Ala	Ala	Ala
	T	Val	Val	Val	Val
A	**G**	Arg	Arg	Ser	Ser
	A	Lys	Lys	Asn	Asn
	C	Thr	Thr	Thr	Thr
	T	Met	Ile	Ile	Ile
C	**G**	Arg	Arg	Arg	Arg
	A	Gln	Gln	His	His
	C	Pro	Pro	Pro	Pro
	T	Leu	Leu	Leu	Leu
T	**G**	Trp	STOP	Cys	Cys
	A	STOP	STOP	Tyr	Tyr
	C	Ser	Ser	Ser	Ser
	T	Leu	Leu	Phe	Phe



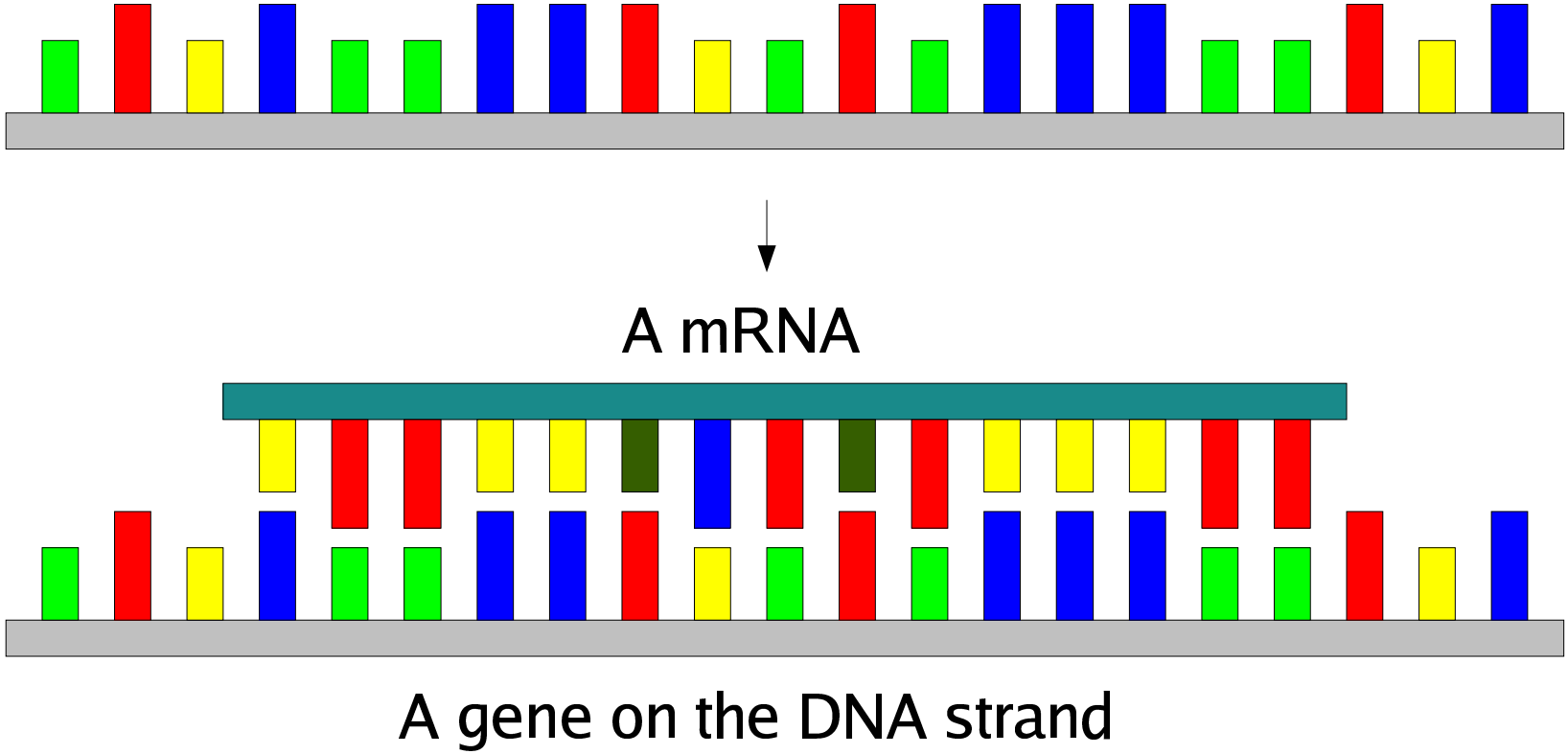
Transcription

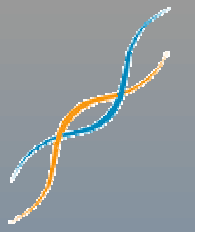
- The codon ATG signals the start of a gene
- DNA double strand is separated



Transcription

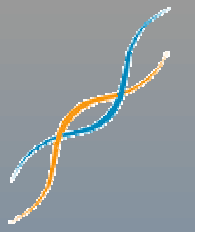
- The gene that is located in one of the strands is used as a tamplate to create an messanger RNA (mRNA) molecule, similar to DNA





Transcription

- In more evolved organisms, there is the splicing phenomenom: some parts of the gene is not copied to mRNA.
  - The parts that are not used in the mRNA are called *introns*, and what is used is called *exons*.
  - DNA corresponding to non-spliced mRNA is called genomic DNA.
  - DNA corresponding to spliced mRNA is called complementary DNA, or simply *cDNA*
- The cDNA can be obtained by reverse transcription



Translation

- Translation is the process of creating proteins from RNA.
  - This happens with the aid of a molecule called tRNA, in the cell ribossomes