NUCLEIC ACIDS
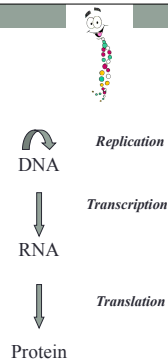
ECS129
Instructor: Patrice Koehl
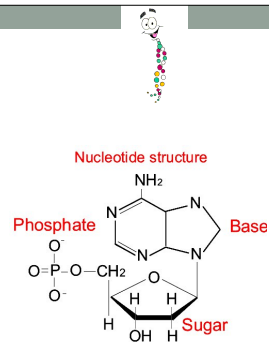

Nucleic Acids

· Nucleotides
· DNA Structure
· RNA
  · Synthesis
  · Function
  · Secondary structure
  · Tertiary interactions
  · Wobble hypothesis

DNA → *Replication*

↓ *Transcription*

RNA

↓ *Translation*

Protein

Nucleic Acids

· Nucleotides
· DNA Structure
· RNA
  · Synthesis
  · Function
  · Secondary structure
  · Tertiary interactions
  · Wobble hypothesis

Nucleotide structure

NH_2

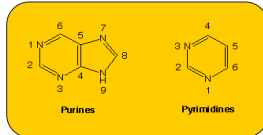Phosphate — Base

$O=P-O-CH_2$

Sugar

OH

Nucleotides

- Nucleotides are found primarily as the monomeric units comprising the major nucleic acids of the cell, RNA and DNA
- Other functions of nucleotides:
  - serving as energy stores (mainly ATP)
  - controlling numerous enzymatic reactions through allosteric effects on enzyme activity
 - mediators of numerous important cellular processes such as second messengers in signal transduction events

Nucleotides

- The nucleotides found in cells are derivatives of the heterocyclic highly basic, compounds, purine and pyrimidine:

Purines   Pyrimidines

Nucleotides

These "bases" are attached to sugar rings: ribose (RNA), deoxyribose (DNA):

ribose   2-deoxyribose
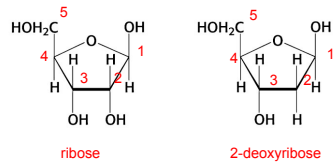
Base Formula	Base (X=H)	Nucleoside X=ribose or deoxyribose	Nucleotide X=ribose phosphate
	Cytosine, C	Cytidine, A	Cytidine monophosphate CMP
	Uracil, U	Uridine, U	Uridine monophosphate UMP
	Thymine, T	Thymidine, T	Thymidine monophosphate TMP
	Adenine, A	Adenosine, A	Adenosine monophosphate AMP
	Guanine, G	Guanosine, A	Guanosine monophosphate GMP

Base conformation

The base can exist in 2 distinct orientations about the N-glycosidic bond.
These conformations are identified as, *syn* and *anti.*
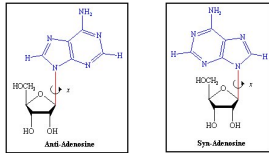
The *anti* conformation predominates.

Sugar conformation
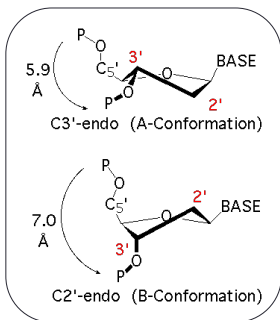
The ribose is a flexible ring that has two preferred conformations in polynucleotides:

-C3'-endo, found mostly in RNA and in DNA "single strand"

-C2'-endo, found mostly in DNA
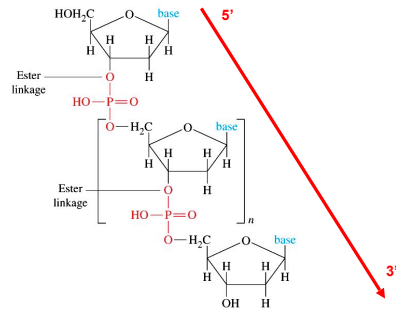
5.9 Å

C3'-endo (A-Conformation)

7.0 Å

C2'-endo (B-Conformation)

Polynucleotides

HOH₂C — O — base   **5'**

Ester linkage — O
HO — P = O
O — H₂C — O — base

Ester linkage — O
HO — P = O
O — H₂C — O — base   **3'**

Nucleic Acids

- Nucleotides
- DNA Structure
- RNA
  - Synthesis
  - Function
  - Secondary structure
  - Tertiary interactions
  - Wobble hypothesis

DNA: double stranded helix

As you can see from the model, the structure is basically a--...long,...twisty,...uh, rope-laddery-type...thingy.

© 1995 GABE MARTIN
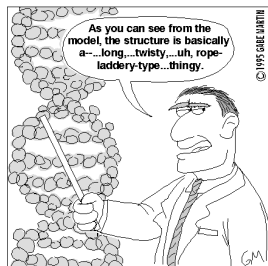
1953: The structure of the DNA molecule is first described.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

DNA: base pairs

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine ; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

Excerpt from Watson and Crick, Nature, 4356, 737-728 (1953)

DNA: base pairs

DNA: double stranded helix

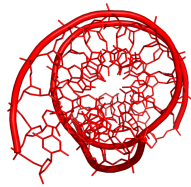We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate di-ester groups joining β-D-deoxyribofuranose residues with 3′.5′ linkages.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Excerpts from Watson and Crick, Nature, 4356, 737-728 (1953)
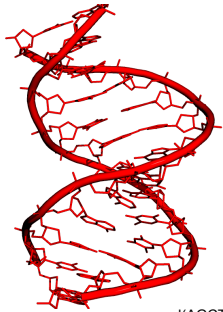
DNA: a style?

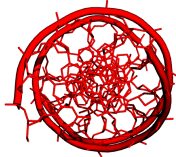Perth, Australia

Life Science, UC Davis
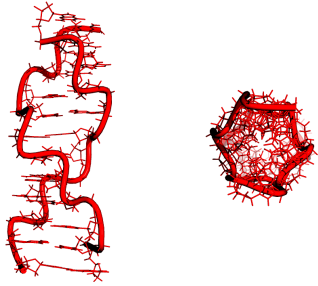

Chambord, France

DNA Structures: A-DNA

d(AGCTTGCCTTGAG)•d(CTCAAGGCAAGCT)
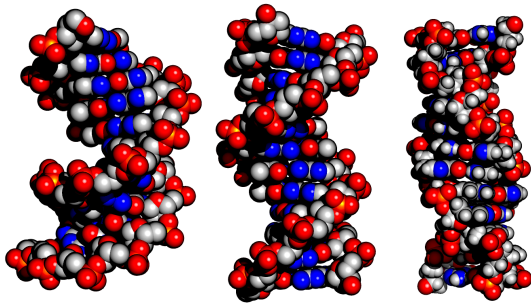

DNA Structures: B-DNA

GCGAATTCGCG)•d(CGCGAATTCGCG)

DNA Structures: Z-DNA

d(CGCGCGCGCGCGCG)•d(CGCGCGCGCGCGCG)

DNA Structures: A, B and Z

DNA Structures: A, B and Z

Property	A-DNA	B-DNA	Z-DNA
Helix	Right-handed	Right-handed	Left-handed
Sugar	C3'-endo	C2'-endo	C2' endo (C) C3' endo (G)
Base pairs /turn	11	10	12
Pitch	28 Å	34 Å	44.6 Å
Tilt	20 deg	0	-7 deg
Rise /bp	2.3 Å	3.4 Å	3.7 Å
Diameter	23 Å	20 Å	17 Å

DNA: organization

(a) Supercoiled
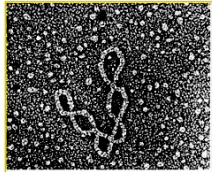
(b) Relaxed circle
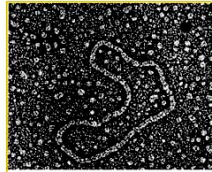
Semi-conservative DNA replication

- Helicase: separates the two DNA strands, starting at replication origins (rich in A-T base pairs)
- RNA primase: inserts a starter of RNA nucleotides at the initiation point
- DNA polymerase binds a complementary leading strand of DNA nucleotides starting at the 3'end of the RNA primer
- Exonuclease removes RNA primer, which are replace with DNA nucleotides by DNA polymerase

Tutorial : http://www.pbs.org/wgbh/aso/tryit/dna/

Nucleic Acids

- Nucleotides
- DNA Structure
- RNA
  - Synthesis
  - Function
  - Secondary structure
  - Tertiary interactions
  - Wobble hypothesis

DNA

Replication

Transcription

RNA

Translation

Protein

Transcription and 5' capping

DNA

5'
Exon   Intron
5' Cap

RNA polymerase

Completion of primary transcript

Primary transcript
5'   Noncoding end sequence   3'

Cleavage, polyadenylation, and splicing

Mature mRNA   5'   AAA(A)$_n$ 3'

---

mRNA synthesis

The mRNA is formed by adding nucleotide that are complementary to the template strand

DNA coding strand   DNA

5'
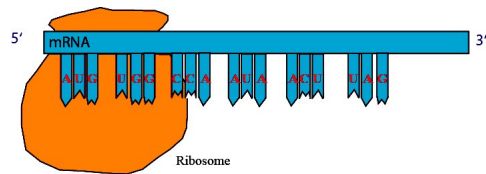3'   3'   3'
5'
DNA template strand   5'
RNA

Nucleic Acids

- Nucleotides
- DNA Structure
- RNA
  - Synthesis
  - Function
  - Secondary structure
  - Tertiary interactions
  - Wobble hypothesis

DNA   Replication

Transcription

RNA

Translation

Protein

rRNA molecules are a major component of ribosomes.

tRNA molecules bring amino acids (aa) to the ribosome.

Large subunit

(Small subunit)

mRNA molecules feed through the ribosome.

Three types of RNA play a role in protein synthesis

TRANSLATION

- The process of reading the mRNA sequence and creating the protein is called translation
- Protein are made of amino acids (20 different, 9 "essentials")
- 3 bases or nucleotides make one codon
- Each codon specifies one amino acid : genetic code

mRNA

5′   3′

AUG   UGG   CCA   AUA   ACU   UAG

MET / codon start   TRP   PRO   ILE   THR   codon stop

GENETIC CODE

2nd base in codon

| 1st base in codon		U	C	A	G	3rd base in codon
	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	STOP	STOP	A
		Leu	Ser	STOP	Trp	G
	C	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
		Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
	A	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
		Ile	Thr	Lys	Arg	A
		Met/Start	Thr	Lys	Arg	G
	G	Val	Ala	Asp	Gly	U
		Val	Ala	Asp	Gly	C
		Val	Ala	Glu	Gly	A
		Val	Ala	Glu	Gly	G

GENETIC CODE

- *The genetic code is universal:*
  - Virtually all organisms use the same genetic code
  - Virtually all organisms use the same 20 amino acids

- *The genetic code is degenerate:*
  6 codons code for Leu

- *Some codons have special functions:*
  AUG encodes for Met, and is codon start
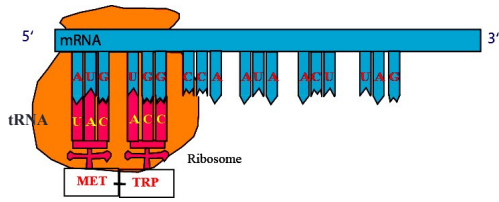 UAA, UAG and UGA are termination codons
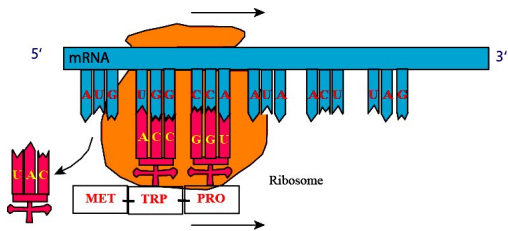

Translation : initiation

Translation : initiation
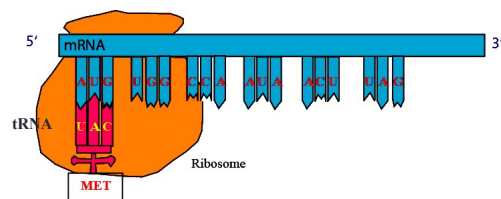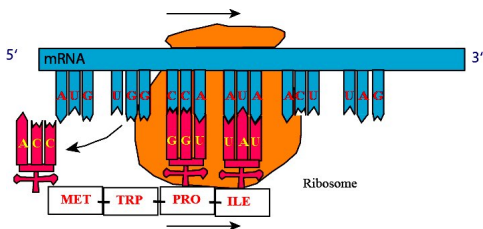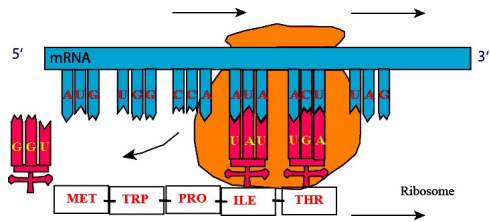
Translation : elongation
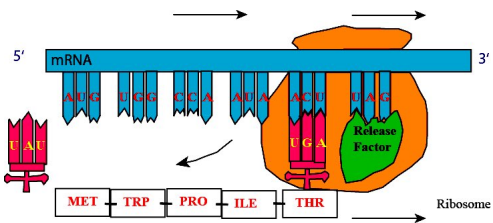

Translation : elongation
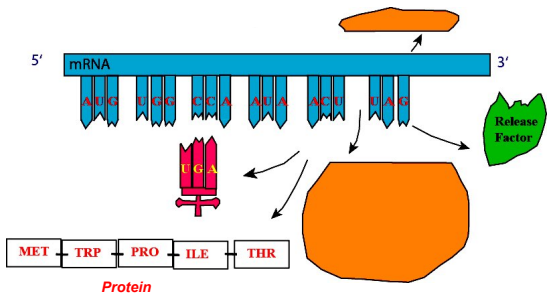

Translation : elongation

Translation : elongation
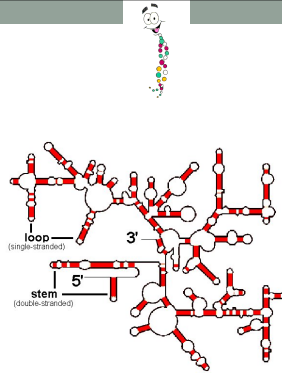

Translation : termination

Translation : termination

Nucleic Acids
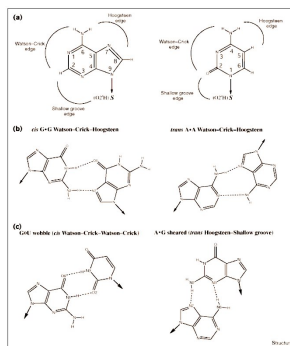
- Nucleotides
- DNA Structure
- RNA
  - Synthesis
  - Function
  - Secondary structure
  - Tertiary interactions
  - Wobble hypothesis

loop
(single-stranded)

3'

stem
(double-stranded)

5'

RNA

- RNA versus DNA:
  - ribose instead of deoxyribose
  - Uracyl instead of Thymine
  - mostly single stranded

- Three major RNAs:
  - mRNA (messenger RNA)   : DNA transcript
  - tRNA (transfer RNA)   : transfer amino acid during
                              protein synthesis
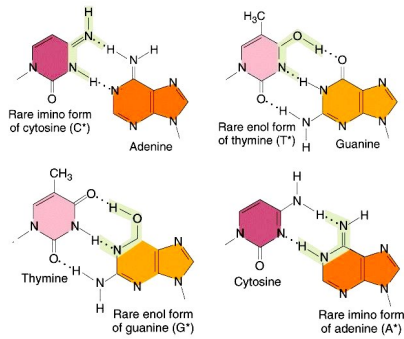 - rRNA   : ribosomal RNA

- Can be active

RNA: non standard base pairs

(a)

(b) G•G Watson-Crick-Hoogsteen   trans A•A Watson-Crick-Hoogsteen

(c) G•U wobble (cis Watson-Crick-Watson-Crick)   A•G sheared (trans Hoogsteen-Shallow groove)

*(Westhof and Fritsch,
RNA folding: beyond Watson Crick pairing,
Structure, 8:55-65 (2000)*

*Database of non-canonical base pairs:
http://prion.bchs.uh.edu/bp_type/*

RNA: non standard base pairs

Rare imino form of cytosine (C*) — Adenine

Rare enol form of thymine (T*) — Guanine

Thymine — Rare enol form of guanine (G*)

Cytosine — Rare imino form of adenine (A*)

RNA secondary structures

· G-C and A-U form hydrogen bonded base pairs and are said to be complementary

· Base pairs are approximately coplanar and are almost always stacked onto other base pairs in an RNA structure. Contiguous base pairs are called stems.

· Unlike DNA, RNA is typically produced as a single stranded molecule which then folds intra-molecularly to form a number of short base-paired stems. This base-paired structure is called RNA secondary structure.

RNA: Hairpins

Single stranded subsequences bounded by base pairs are called loops. A loop at the end of a stem is called a hairpin loop. Simple substructures consisting of a single stem and loop are called stem loops, or hairpins.


TGCGATACTCATCGCA

5'

Hairpin
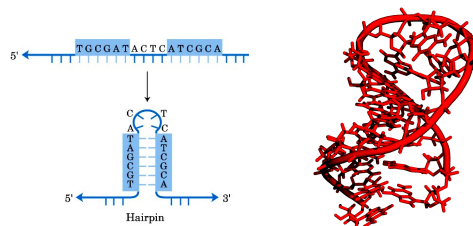
RNA secondary structures

Single stranded bases within a stem form a bulge or bulge loop if the single stranded bases are on only one side of the stem.

If single stranded bases interrupt both sides of a stem, they form an internal (interior) loop.
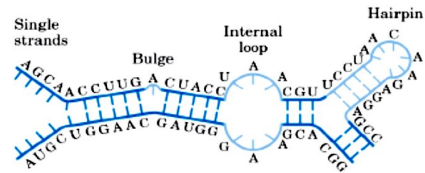
Nucleic Acids

- Nucleotides
- DNA Structure
- RNA
  - Synthesis
  - Function
  - Secondary structure
  - Tertiary interactions
  - Wobble hypothesis
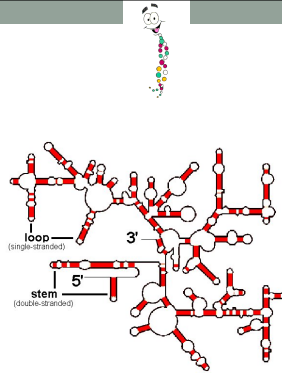
RNA "tertiary interactions"

In addition to secondary structural interactions in RNA, there are also tertiary interactions, including: (A) pseudoknots, (B) kissing hairpins and (C) hairpin-bulge contacts.

Pseudoknot   **Kissing hairpins**   **Hairpin-bulge**

Pseudoknots

Kissing hairpins

tRNA structure

tRNA: small molecules (73 to 93 nucleotides) with cloverleaf secondary structure

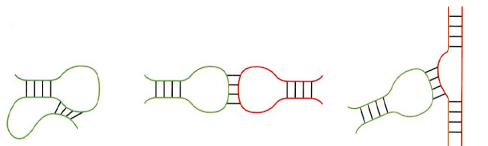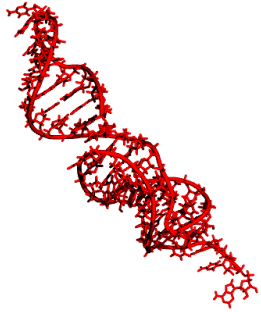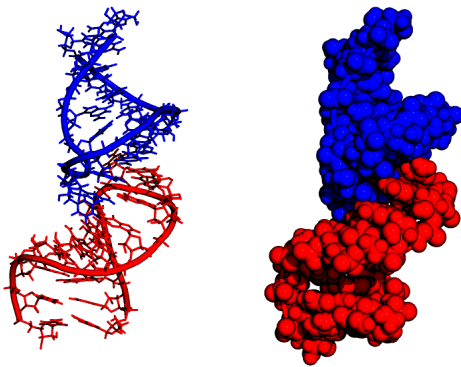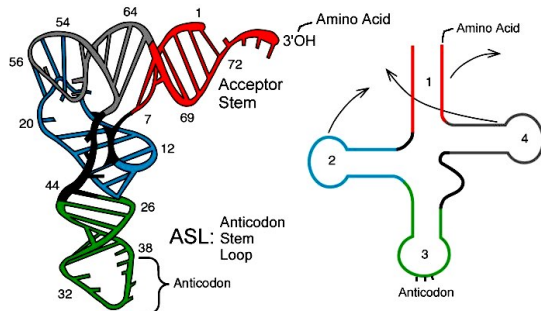# Nucleic Acids

- Nucleotides
- DNA Structure
- RNA
  - Synthesis
  - Function
  - Secondary structure
  - Tertiary interactions
  - Wobble hypothesis

loop
(single-stranded)

3'

stem
(double-stranded)

5'

The wobble hypothesis (Crick, 1966)

Ester bond

tRNA
molecule

CHR — Amino acid

NH₃⁺

Intramolecular
base-pairing

Anticodon

mRNA 5'———G-C-C———3'
Codon

In 1965, Holley determined the sequence
of yeast tRNA(ala): he found the
nucleotide Inosine at the 5' end in the
anticodon.

The wobble hypothesis (Crick, 1966)

Crick proposed the Wobble hypothesis to generalize Holley's observation:
*- Interaction between codon and anticodon need to be exact at first
two positions.*
*- The third interaction can be less restrictive, and can include non standard
base-pairing.*

This hypothesis accounts for the degeneracy
In the Genetic Code.

5' anticodon base	3' codon base
A	U
C	G
G	C or U
U	A or G
I	A or C or U

Wobble base pairs

Inosine   **Uracil**

Summary (1)

- Nucleotides have three parts: sugar (ribose in RNA, deoxyribose in DNA), base (purine,A, G, and pyrimidine, C, T or U), and phosphate group.
- Nucleotide can polymerise to form polynucleotides, or "strands".
- DNA (deoxyribo nucleic acid) is a double stranded helix, where the two strands run in opposite directions and are maintained together by hydrogen bonds. Base pairs include one purine and one pyrimidine (A-T and G-C).
- There are three main forms of DNA helices: A, B and Z.
- DNA molecules have topological constraints, such as supercoiling.

Summary (2)

- Only one DNA strand is used for RNA synthesis: the "template" strand, which is complementary to the coding strand. The sequence of the mRNA is the sequence of the coding strand, where T are replace by U.

- Three types of RNA are involved in protein synthesis: messenger RNA (mRNA, carries the information), transfer RNA (tRNA, brings the correct amino acid during synthesis), and ribosomal RNA (rRNA, major consituent of the ribosome, where protein synthesis occurs.

- The message carried by the mRNA is read as a collection of "words" of 3 letters, or codons. There are 64 codons, that code for 20 amino acids. AUG is the initiation codon, which codes for Methionine. UAA, UAG and UGA are stop codons. There is redundancy in the genetic code, related to the third base in the codon.

Summary (3)

- RNA bases can be free, involved in base pairs, or base triplets.

- RNA contains single stranded regions, hairpin loops, bulges, and internal loops (secondary structures)

- RNA secondary structures can interact to form pseudoknots, kissing hairpins, or hairpin-bulge complexes.

- The wobble hypothesis is based on the presence in some tRNA of Inosine at the 5' end of the anticodon. It is one possible explanation of the degeneracy of the genetic code.