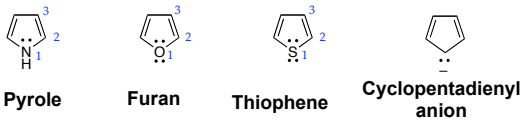
Chapter 28 Biomolecules: Heterocycles and Nucleic Acids

Heterocycles: cyclic organic compounds that contain rings atoms other than carbon (N,S,O are the most common).

28.1 Five-Membered Unsaturated Heterocycles (please read)
28.2 Structures of Pyrrole, Furan, and Thiophene (please read)



28.3 Electrophilic Substitution Reactions of Pyrrole, Furan, and Thiophene (please read)
28.4 Pyridine, a Six-Membered Heterocycle (please read)
28.5 Electrophilic Substitution of Pyridine (please read)
28.6 Nucleophilic Substitution of Pyridine (please read)
28.7 Fused-Ring Heterocycles (please read)

These sections contain some important concepts that were covered previously.

418

28.8 Nucleic Acids and Nucleotides

Nucleic acids are the third class of biopolymers (polysaccharides and proteins being the others)

Two major classes of nucleic acids
- deoxyribonucleic acid (DNA): carrier of genetic information
- ribonucleic acid (RNA): an intermediate in the expression of genetic information and other diverse roles

The Central Dogma (F. Crick):

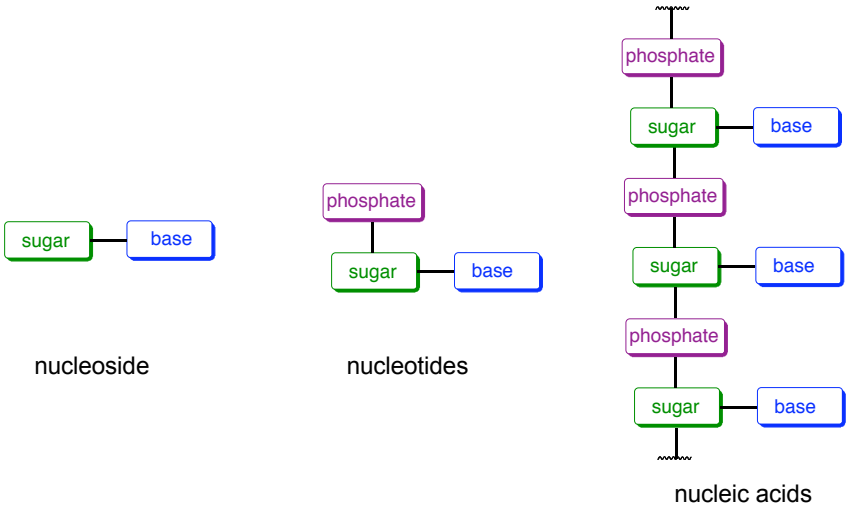
DNA (genome) ⟶ mRNA ⟶ Protein (proteome)

The monomeric units for nucleic acids are nucleotides

Nucleotides are made up of three structural subunits
1. Sugar: ribose in RNA, 2-deoxyribose in DNA
2. Heterocyclic base
3. Phosphate

419

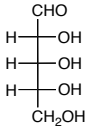
Nucleoside, nucleotides and nucleic acids

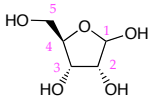


The chemical linkage between monomer units in nucleic acids is a phosphodiester

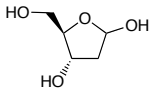
420

The sugar: based on the furanose form D-ribose
ribose for RNA; 2-deoxyribose for DNA



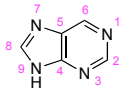


D-ribose

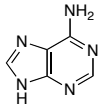


D-2-deoxyribose

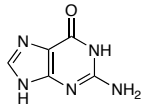
The heterocyclic base; there are five common bases for nucleic acids



purine

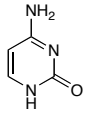


adenine (A)
DNA/RNA

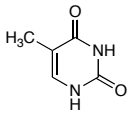


guanine (G)
DNA/RNA

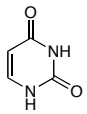
pyrimidine



cytosine (C)
DNA/RNA



thymine (T)
DNA

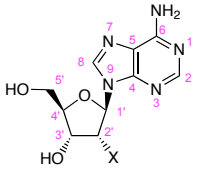


uracil (U)
RNA

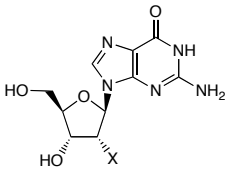
421

Nucleosides: sugar + base

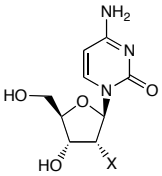
ribonucleosides or 2’-deoxyribonucleosides



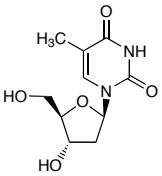
RNA: X= OH, adenosine (A)
DNA: X= H, 2'-deoxyadenosine (dA)



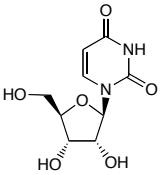
RNA: X= OH, guanosine (G)
DNA: X= H, 2'-deoxyguanosine (dG)



RNA: X= OH, cytidine (C)
DNA: X= H, 2'-deoxycytidine (dC)



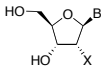
DNA: thymidine (T)



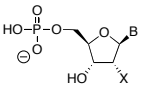
RNA: R= H, uridine (U)

422

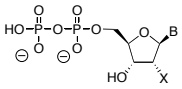
Nucleotides: nucleoside + phosphate



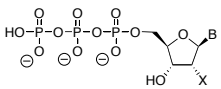
ribonucleoside (X=OH)
deoxyribonucleoside (X=H)



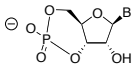
ribonucleotide 5'-monophosphate (X=OH)
deoxyribonucleotide 5'-monophosphate (X=H)



nucleotide 5'-diphosphate



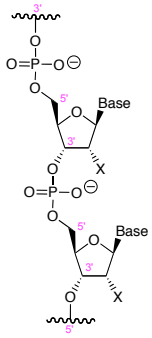
nucleotide 5'-triphosphate



ribonucleotide
3',5'-cyclic phosphophosphate

28.9 Structure of nucleic acids:
The chemical linkage between nucleotide units in nucleic acids is a phosphodiester, which connects the 5’-hydroxyl group of one nucleotide to the 3’-hydroxyl group of the next.

DNA is negatively charged



423

DNA sequences are written from left to right from the 5’ to 3’

5’-ATCGCAT-3’ or 5’-d(ATCGCAT)-3’

28.10 Base Pairing in DNA: The Watson–Crick Model

Chargaff’s Rule: the number of A = T and G = C in DNA

Two polynucleotide strands, running in opposite directions (*anti-parallel*) and coiled around each other in a *double helix*.

The strands are held together by complementary hydrogen-bonding between specific pairs of bases.

“Molecular Structure of Nucleic Acids” Watson J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738

"Molecular Structure of Deoxypentose Nucleic Acids" Wilkins, M. H. F.; Stokes A.R.; Wilson, H. R. *Nature* **1953**,*171*, 738-740.

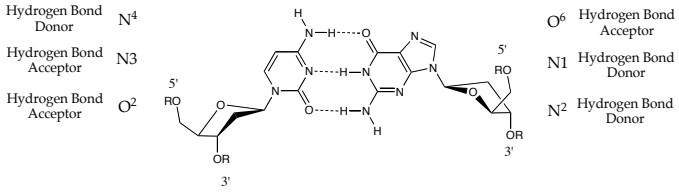
”Molecular Configuration in Sodium Thymonucleate,” Franklin, R.; Gosling, R. G. *Nature* **1953**, *171*, 740-741

1962 Nobel Prize in Medicine: F. H. C. Crick, J. D. Watson, Maurice F. H. Wilkins, *“for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material.”*

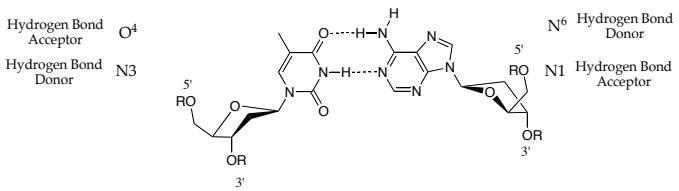
424

Anti-parallel, complementary hydrogen-bonding of DNA base pairs

Antiparallel C-G Pair

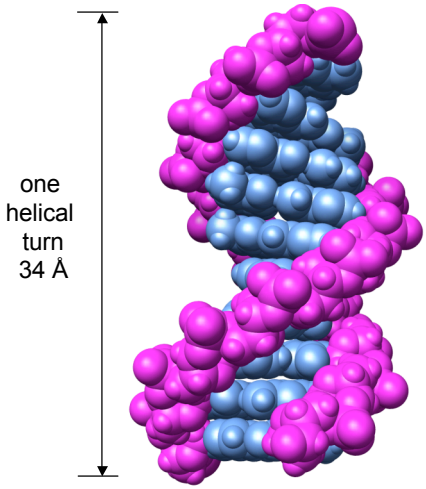


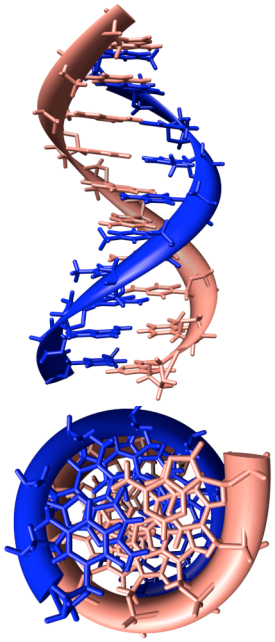
Antiparallel T-A Pair



425

DNA double helix

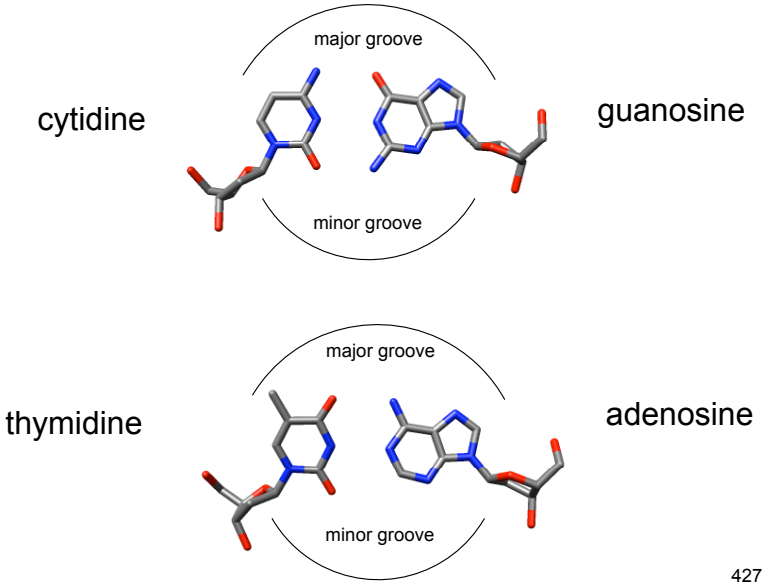




backbone: deoxyribose and phosphodiester linkage
bases

426

DNA Grooves



427

"It has not escaped our attention that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." Watson & Crick

28.11 Nucleic Acids and Heredity
The Central Dogma (F. Crick):

DNA $\xleftarrow{\text{replication}}$ DNA (genome) $\xrightarrow{\text{transcription}}$ mRNA $\xrightarrow{\text{translation}}$ Protein (proteome)

Expression and transfer of genetic information:

- **Replication**: process by which DNA is copied with very high fidelity.
- **Transcription**: process by which the DNA genetic code is read and transferred to messenger RNA (mRNA). This is an intermediate step in protein expression
- **Translation**: The process by which the genetic code is converted to a protein, the end product of gene expression.

The DNA sequence codes for the mRNA sequence, which codes for the protein sequence

428

28.12 Replication of DNA: DNA is replicated by the coordinated efforts of a number of proteins and enzymes.

Each cell contains about two meters of DNA. The DNA must be "packaged" into the cell nucleus by super-coiling and knotting.

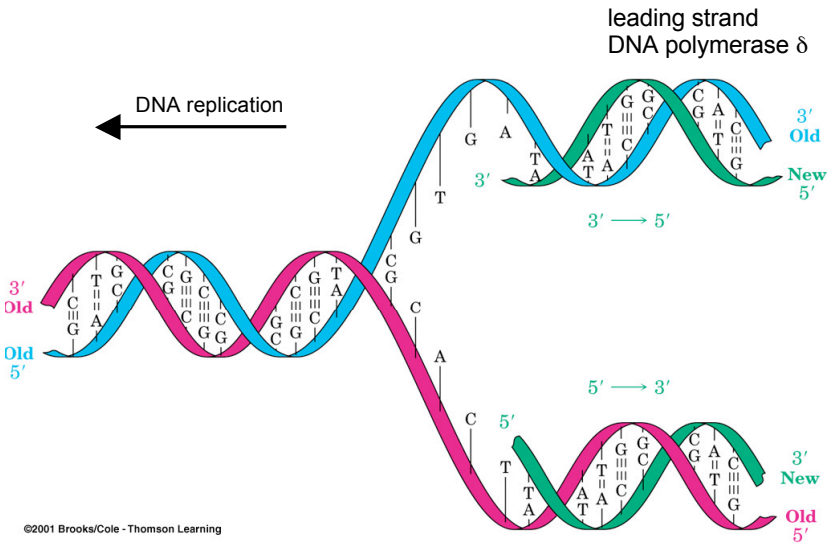
For replication, DNA must be unknotted, uncoiled and the double helix unwound.

Topoisomerase: Enzyme that unknots and uncoils DNA
Helicase: Protein that unwinds the DNA double helix.
DNA polymerase: Enzyme replicates DNA using each strand as a template for the newly synthesized strand.

DNA replication is *semi-conservative*: Each new strand of DNA contains one parental (old, template) strand and one daughter (newly synthesized) strand

429

Unwinding of DNA by helicases expose the DNA bases (replication fork) so that replication can take place

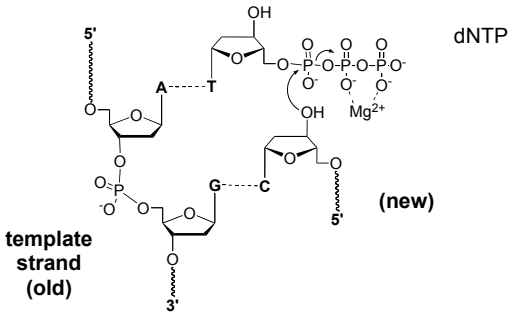


430

DNA Polymerase: the new strand is replicated from the 5' → 3' (start from the 3'-end of the template)

DNA polymerases are Mg^{2+} ion dependent

The deoxynucleotide 5'-triphosphate (dNTP) is the reagent for nucleotide incorporation

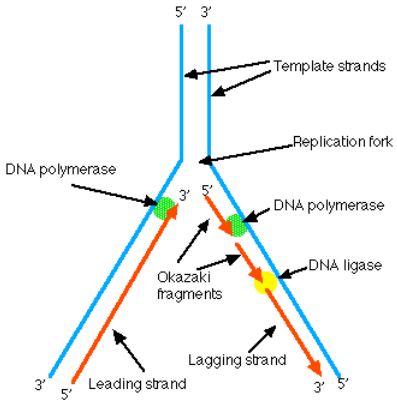


3'-hydroxyl group of the growing DNA strand acts as a nucleophile and attacks the α-phosphorus atom of the dNTP.

431

Replication of the *leading strand* occurs continuously in the 5' → 3' direction of the new strand.

Replication of the *lagging strand* occurs discontinuously. Short DNA fragments are initially synthesized and then ligated together. *DNA ligase* catalyzes the formation of the phosphodiester bond between pieces of DNA.



432

animations of DNA processing: http://www.wehi.edu.au/education/wehi-tv/dna/

DNA replication occurs with very high fidelity:

- Most DNA polymerases have high intrinsic fidelity
- Many DNA polymerases have "proof-reading" (exonuclease) activity
- Mismatch repair proteins seek out and repair base-pair mismatches due to unfaithful replication

28.13 Structure and Synthesis of RNA: Transcription

RNA contains ribose rather than 2-deoxyribose and uracil rather than thymine

There are three major kinds of RNA

- messenger RNA (mRNA):
- ribosomal RNA (rRNA)
- transfer RNA (tRNA)

DNA is found in the cell nucleus and mitochondria; RNA is more disperse in the cell.

433

Transcription: only one of the DNA strands is copied (coding or antisense strand). It sequence is converted to the complementary sequence in mRNA (template or sense strand), which codes for the amino acid sequence of a protein (or peptide)

28.14 RNA and Protein Biosynthesis: Translation

- proteins are synthesized in the cytoplasm on ribosomes.
- mRNA is the template for protein biosynthesis.
- a three base segment of mRNA (codon) codes for a an amino acid.

THE STANDARD GENETIC CODE

UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

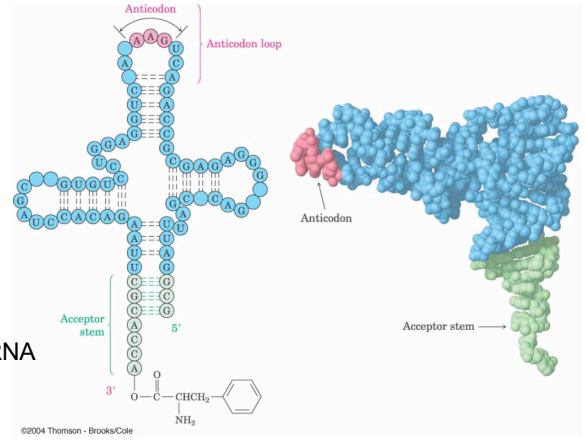
AUG is part of the initiation signal, as well as being the codon for internal methionine.

434

The "anticodon" region of of tRNA is complementary for a to the mRNA codon sequence.

The t-RNA carries an amino acid on the 3'-hydroxyl (aminoacyl t-RNA) and the ribosome catalyzes amide bond formation.

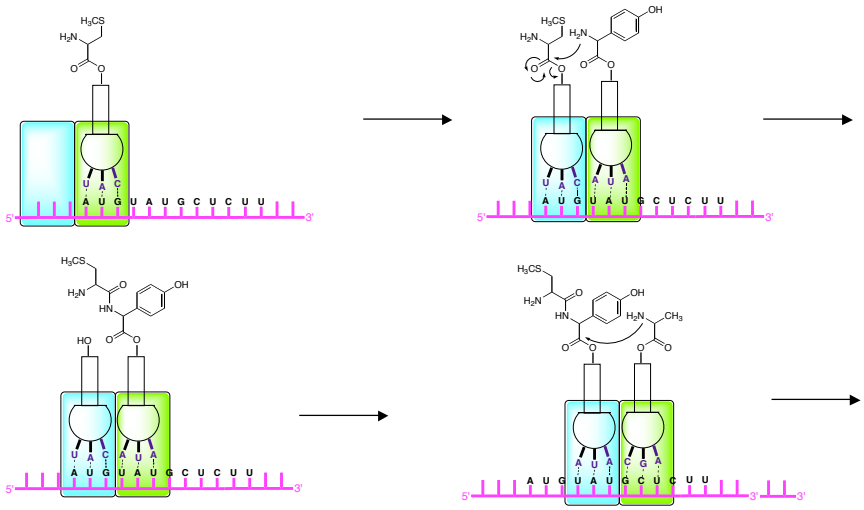
Although single-stranded, the are complementary sequences within tRNA that give it a defined conformation



aminoacyl t-RNA

435

Ribosomal protein synthesis



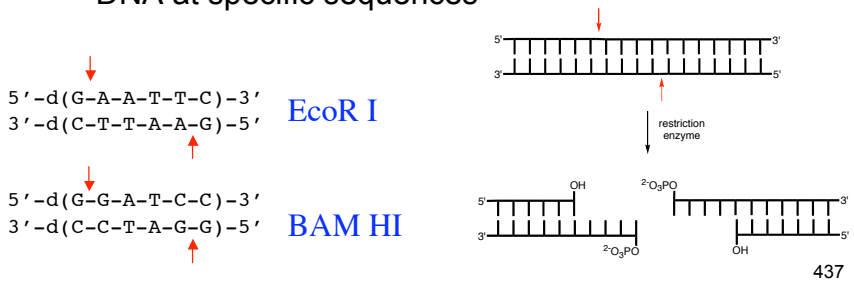
436

DNA sequencing

Maxam-Gilbert: relys on reagents that react with a specific DNA base that can subsequent give rise to a sequence specific cleavage of DNA

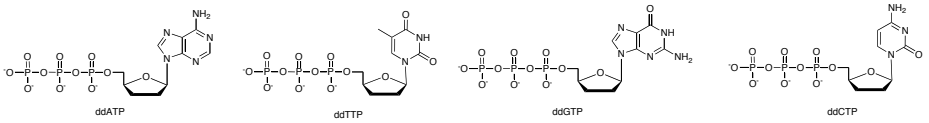
Sanger: Enzymatic replication of the DNA fragment to be sequenced with DNA polymerase, Mg^{+2}, and dideoxynucleotides triphosphate (ddNTP) that truncate DNA replication

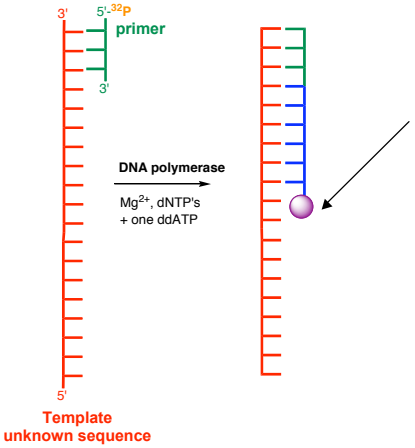
Restriction endonucleases: Bacterial enzymes that cleave DNA at specific sequences



437

Sanger Sequencing dideoxynucleotides triphosphate (ddNTP)

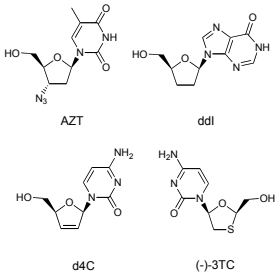




When a ddNTP is incorporated elongation of the primer is terminated

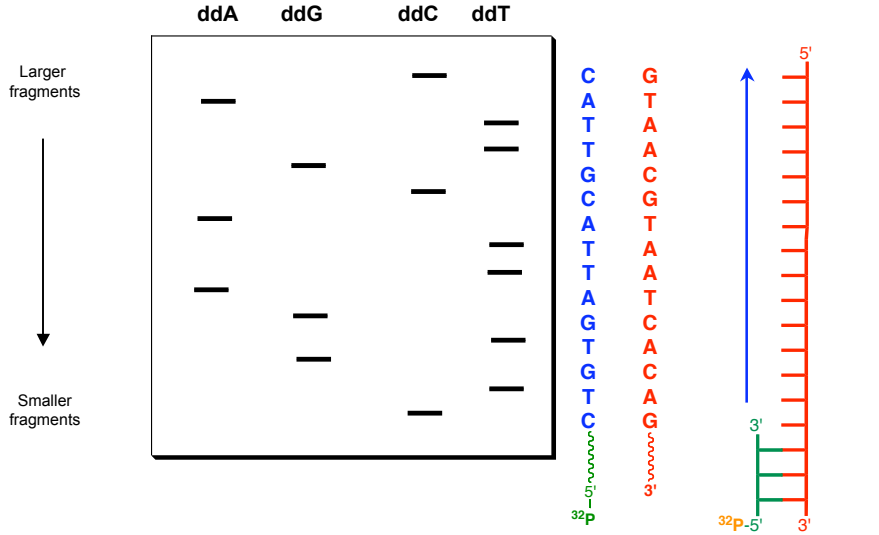
The ddNTP is specifically incorporated opposite its complementary nucleotide base

Anti-Viral Nucleosides



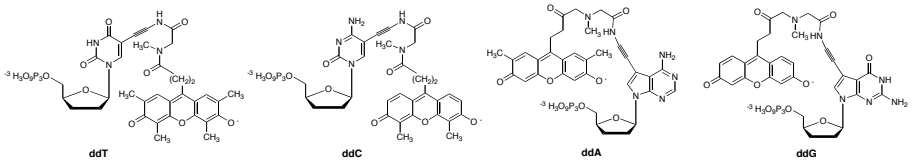
438

Sanger Sequencing



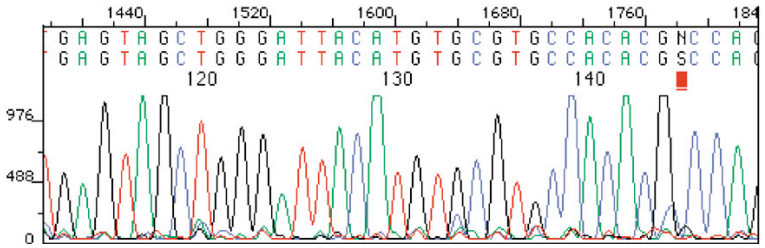
439

Automated Sanger Sequencing with flourescent ddNTP's



Excitation: ~ 490 nM, Emission: ddT= 526 nm. ddC= 519 nm, ddA= 512 nm, ddG= 505 nm

Sanger sequencing using flourescent ddNTP's: terminated DNA strands are separated by capillary electrophoresis, detected and identified by their fluorescence emission



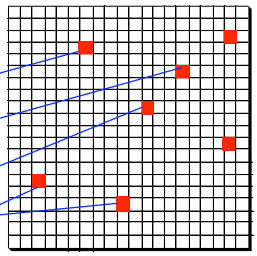
440

Sequencing on a "chip" (*Lagniappe*)

Spatially addressable: 8-mer chip: 65, 536 different sequences
12-mer chip: 1,677,216 different sequences

DNA fragment ⟶ DNA–fluorophore

Place DNA on the chip, then wash away non-specific hybridization after 1-10 hrs. Raise temperature and "melt" away partially hybridized sequences.



```
3'-ACGGTGCG
    CGGTGCGA
     GGTGCGAG
      GTGCGAGA
       TGCGAGAA
        GCGAGAAT    etc
```

3'-ACGGTGCGAGAAT---5' (from the chip)
5'-TGCCACGCTCTTA---3'

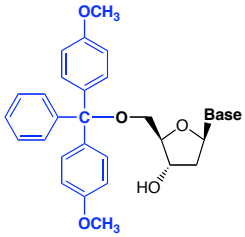
441

28.16 DNA Synthesis: short segments of DNA can be efficiently by automated, solid-phase methods.

Solid support: controlled pore glass (silica)
linked to the 3'-hydroxyl group of the first nucleotide
solid phase DNA synthesis is from 3'→5', which is the opposite direction from nature.

Protected nucleotide reagents:
5'-protecting group: 4,4'-dimethyoxytrityl (trityl is a triphenylmethyl group), abbreviated DMT

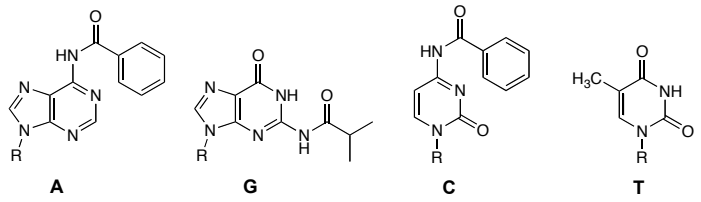
5'-DMT ether



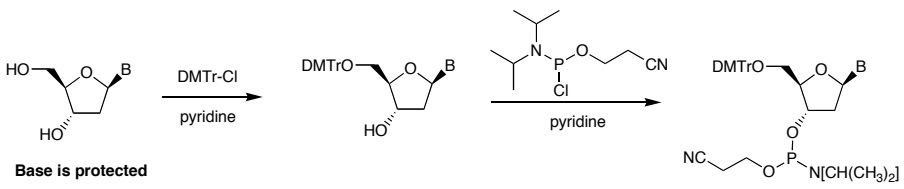
Removes with mild acid (Cl_3CCO_2H)

442

The base amino groups of dA, dG and dC must be protected, usually as an amides. The base of T does not require further protection.

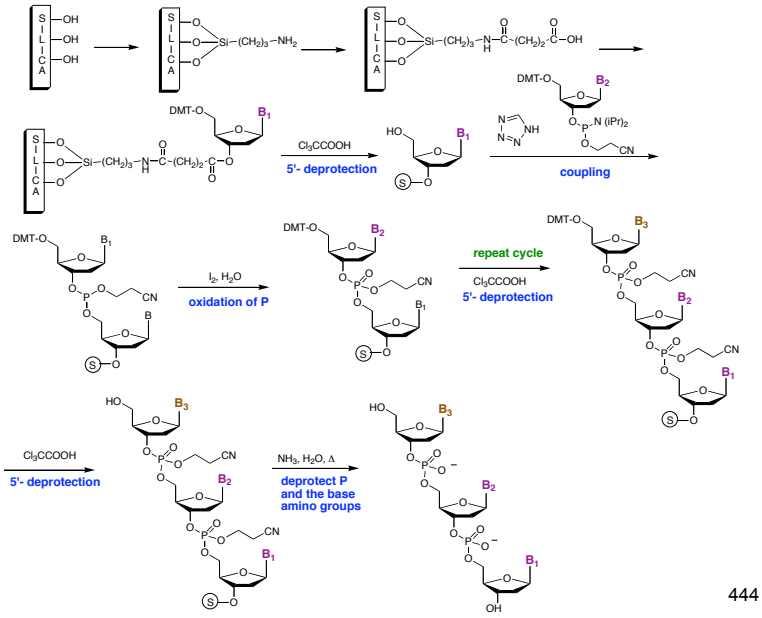


The 3'-phosphorous group: phosphoramidite



443

Automated, solid-phase DNA synthesis



444

28.17 Polymerase Chain Reaction (PCR): method for amplifying DNA using DNA polymerase and cycling temperature

Heat stable DNA Polymerases (from *archaea*):

Taq: thermophilic bacteria (hot springs)- no proof reading

Pfu: geothermic vent bacteria- proof reading

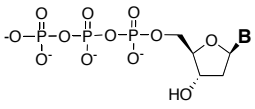
Mg^{2+}

two Primer DNA strands (synthetic, large excess)

one sense primer and one antisense primer

one Template DNA strand

dNTP's



KARY B. MULLIS, 1993 Nobel Prize in Chemistry for his invention of the polymerase chain reaction (PCR) method.

445

A typical PCR temperature cycle

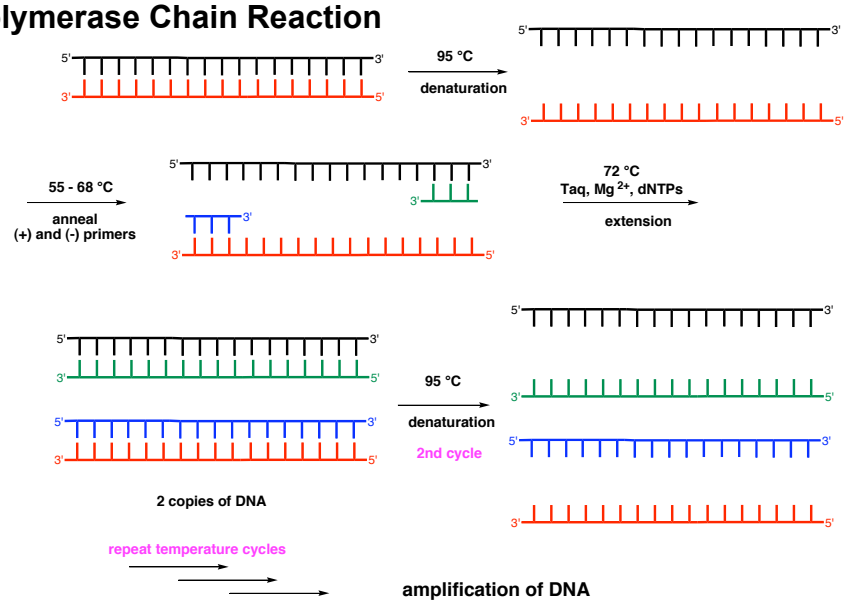
Denaturation:	94 °C	0.5 - 1 min	
Annealing:	55-68 °C	0.5 - 1 min	5 °C below the T_m of the primer
Extension: (replication)	72 °C	1 min	+ 1 min per Kb of DNA
# of cycles	25 - 35		
Final extension	72 °C	10 min	

1 x 2 = 2 x 2 = 4 x 2 = 8 x 2 = 16 x 2 = 32 x 2 = 64 x 2 = 128 x 2 = 256 x 2 = 512 x 2 = 1,024 x 2 = 2,048 x 2 = 4,096 x 2 = 8,192 x 2 = 16,384 x 2 = 32,768 x 2 = 65,536 x 2 = 131,072 x 2 = 262,144 x 2 = 524,288 x 2 = 1,048,576

In principle, over one million copies per original, can be obtained after just twenty cycles

446

Polymerase Chain Reaction



For a PCR animation go to: http://www.blc.arizona.edu/INTERACTIVE/recombinant3.dna/pcr.html
http://users.ugent.be/~avierstr/principles/pcrani.html

447