N Bases / Nucleosides / Nucleotides / Nucleic Acid Structures (Review)

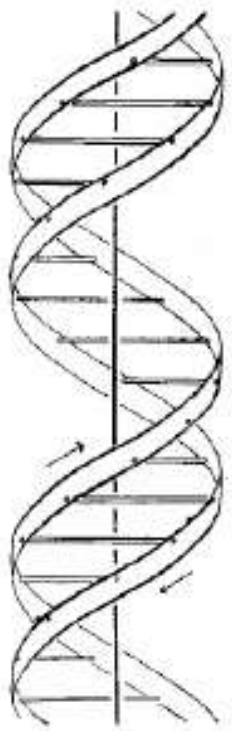
Goals for this review unit:

1. **Recognize the common building blocks of nucleic acids: names / 1-letter abbrev.**
2. **Nomenclature for nucleosides and nucleotides (structure ofATP)**
3. **Primary structures of RNA and DNA**
4. **Conformations in DNAs**
5. **Characteristics of B-DNA, A-DNA and Z-DNA**
6. **Denaturation of DNA**
7. **Features of RNA / Functions of RNA**
8. **DNA Sequencing (Maxam – Gilbert vs. Sanger Dideoxy)**

The Birth of Molecular Biology: DNA Structure

inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

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. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There



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

Nature – 1953

Nature – 2001

The Nobel Prize in Physiology or Medicine 1962

"for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material"

Francis Harry Compton Crick

1/3 of the prize

United Kingdom

MRC Laboratory of Molecular Biology
Cambridge, United Kingdom

b. 1916
d. 2004

James Dewey Watson

1/3 of the prize

USA

Harvard University
Cambridge, MA, USA

b. 1928

Maurice Hugh Frederick Wilkins

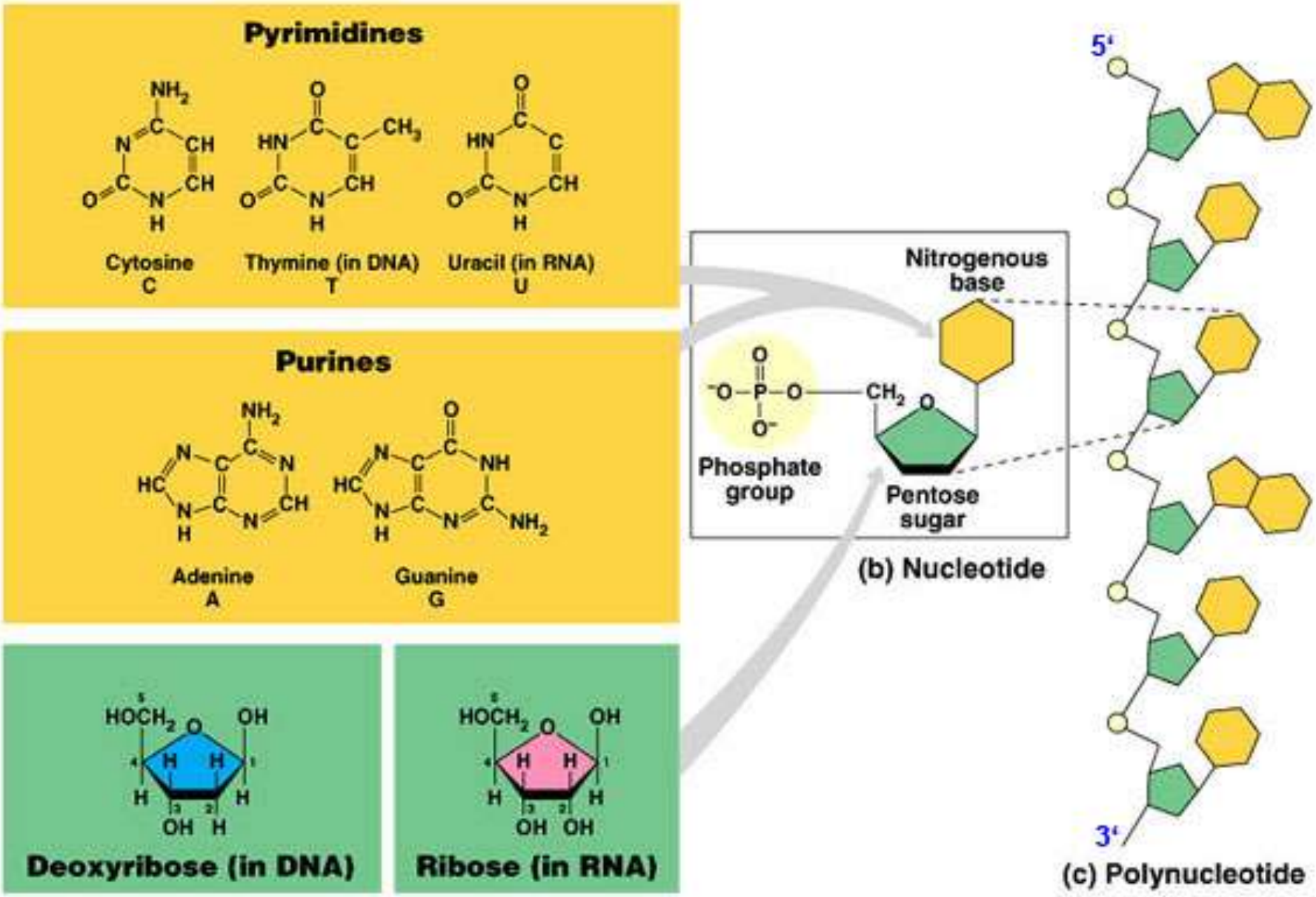
1/3 of the prize

United Kingdom and New Zealand

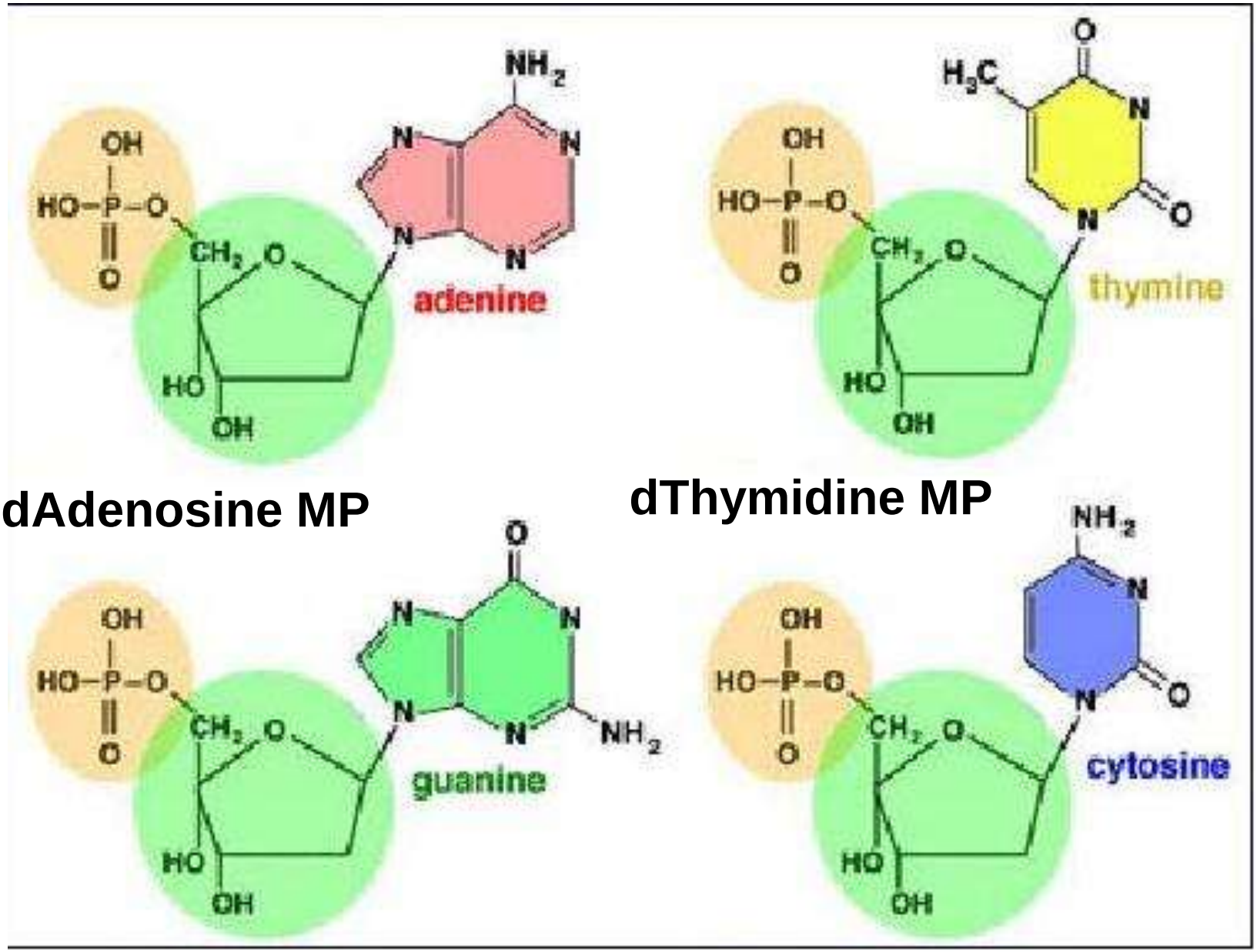
London University
London, United Kingdom

b. 1916
(in Pongaroa, New Zealand)
d. 2004

Left to right: Maurice Wilkins, John Steinbeck, John Kendrew, Max Perutz, Francis Crick and Jim Watson after the Nobel Ceremony in Stockholm in December 1962.



Copyright © Pearson Education, Inc., publishing as Benjamin Cummings



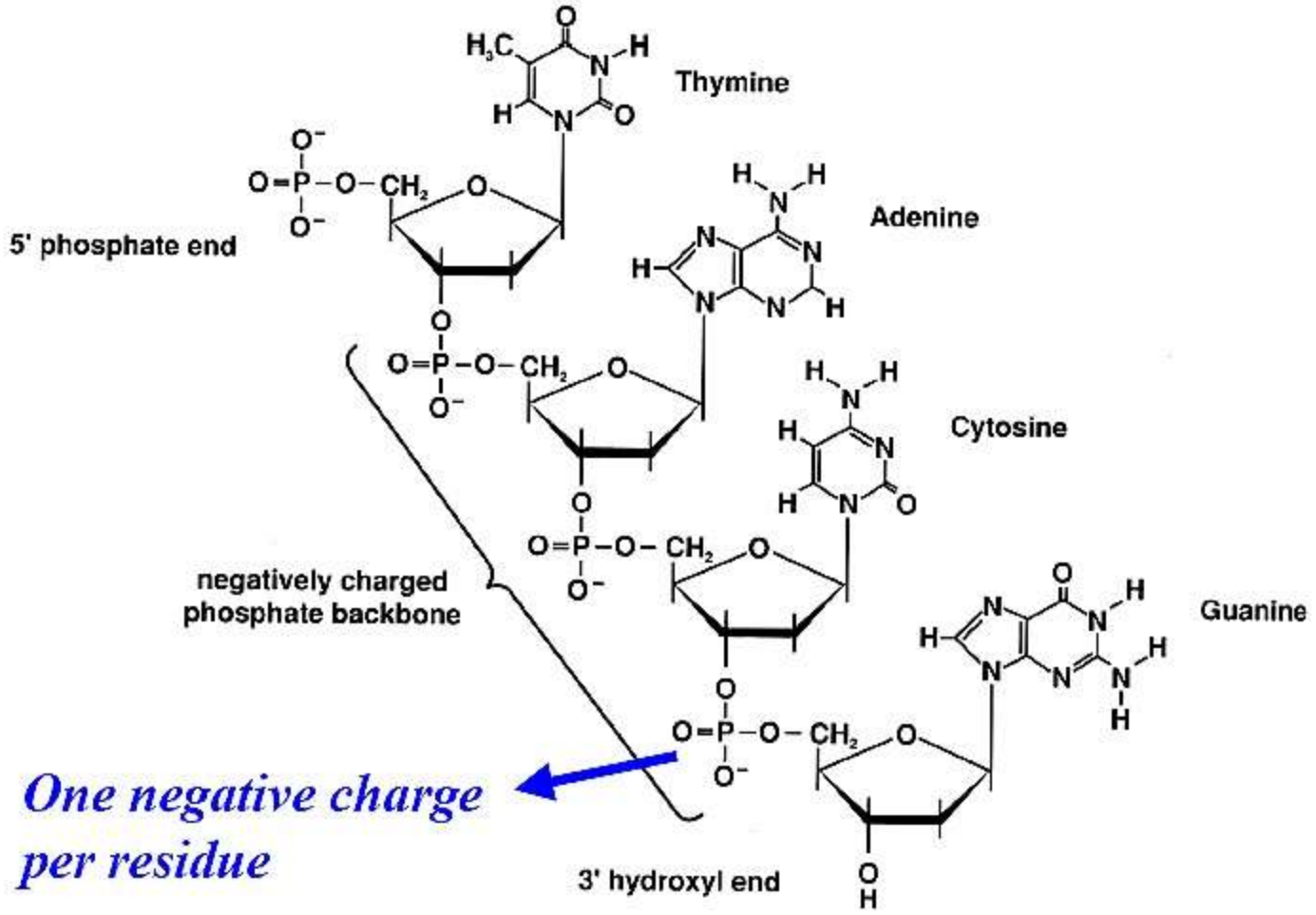
dAdenosine MP

dThymidine MP

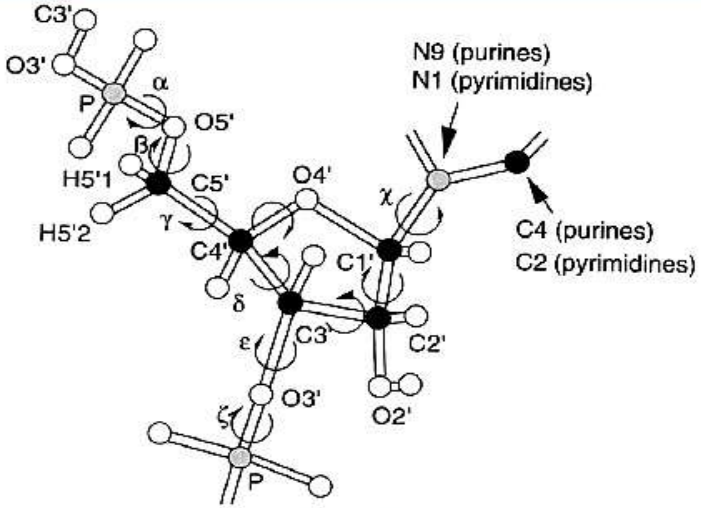
dGuanosine MP

dCytidine MP

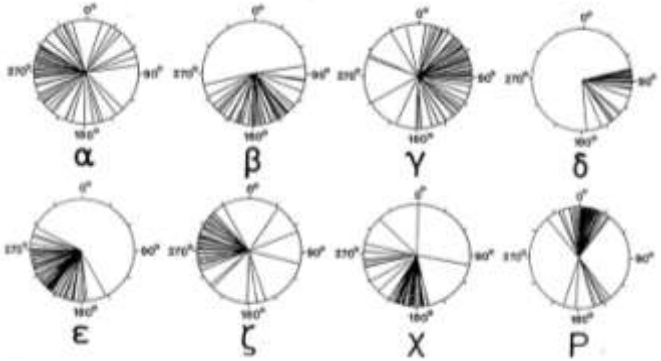
DNA primary structure



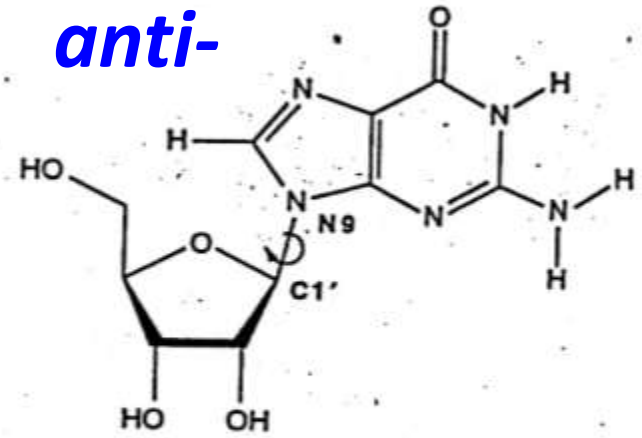
Rotational angles of phosphodiester chain



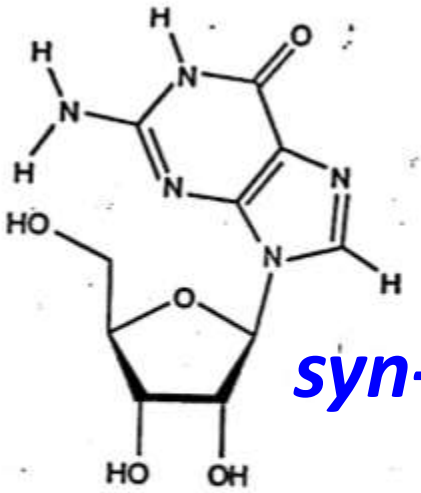
	B-DNA	A-DNA	Z-DNA C	Z-DNA G
$\alpha(\omega)$	−41	−90	138	100
$\beta(\phi)$	136	211	−94	−108
$\gamma(\psi)$	38	47	80	−70
$\delta(\psi')$	139	83	48	−130
$\varepsilon(\phi')$	−133	−185	180	−140
$\zeta(\omega')$	−57	−45	−170	56
χ	78	27	20	−100



anti-

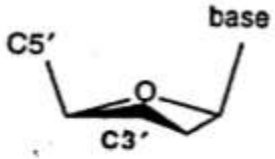


guanosine-anti



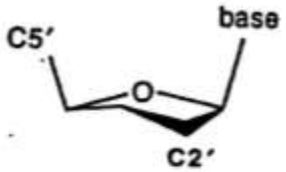
guanosine-syn

syn-



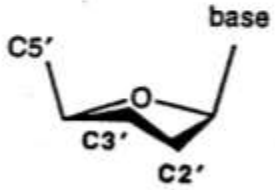
C3′ endo

C3’ endo



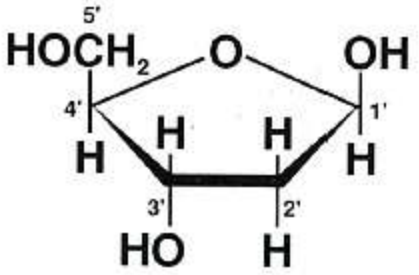
C2′ exo

C2’ exo

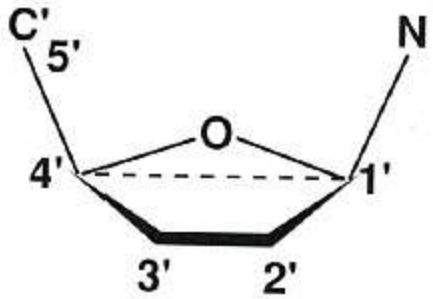


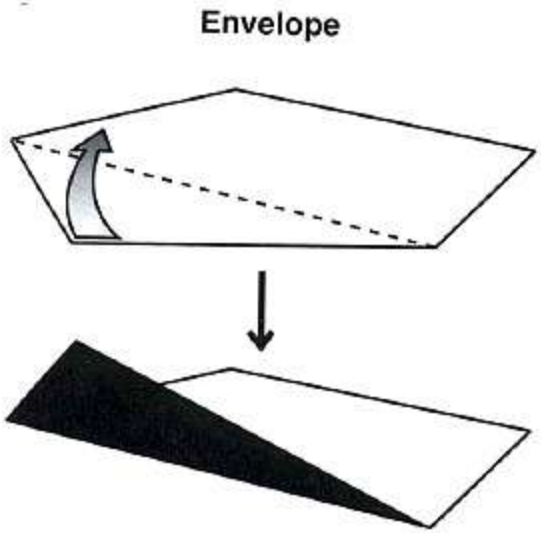
C3′ endo–C2′ exo

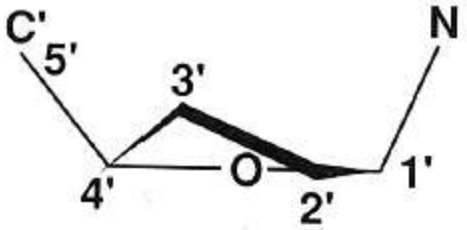
Sugar pucker in DNA



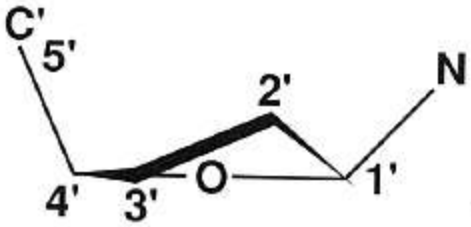
β-D-2-Deoxyribose





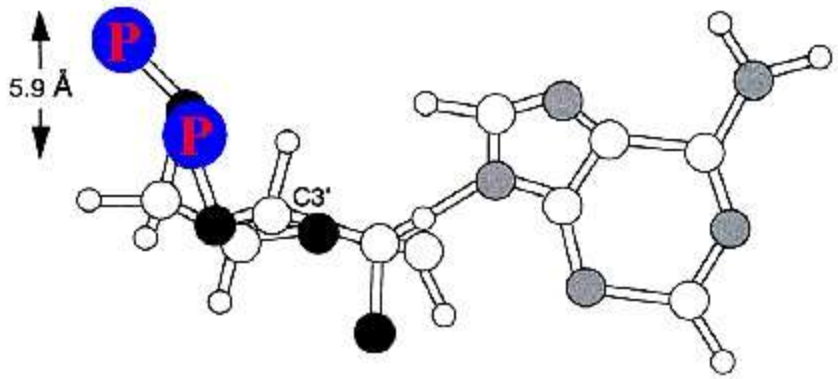


C3' endo

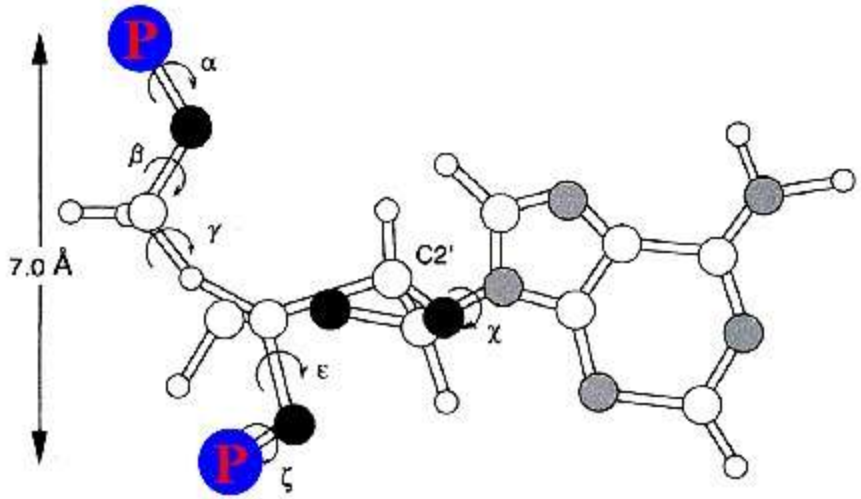


C2' endo

Sugar pucker in DNA



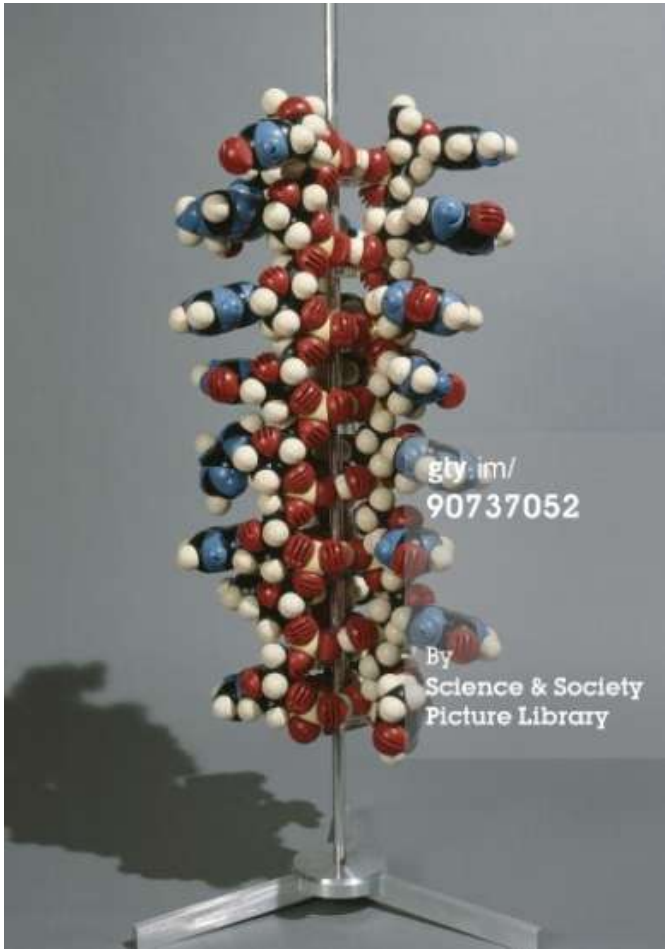
C3' endo

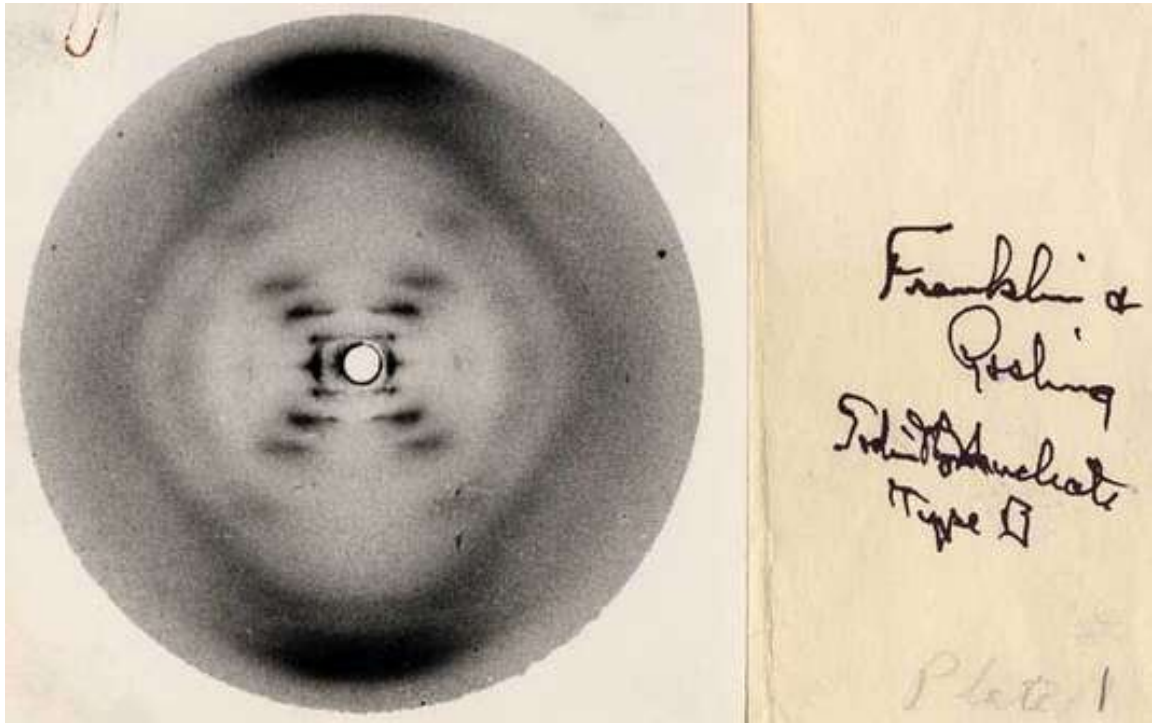


C2' endo

Pauling triple helix model

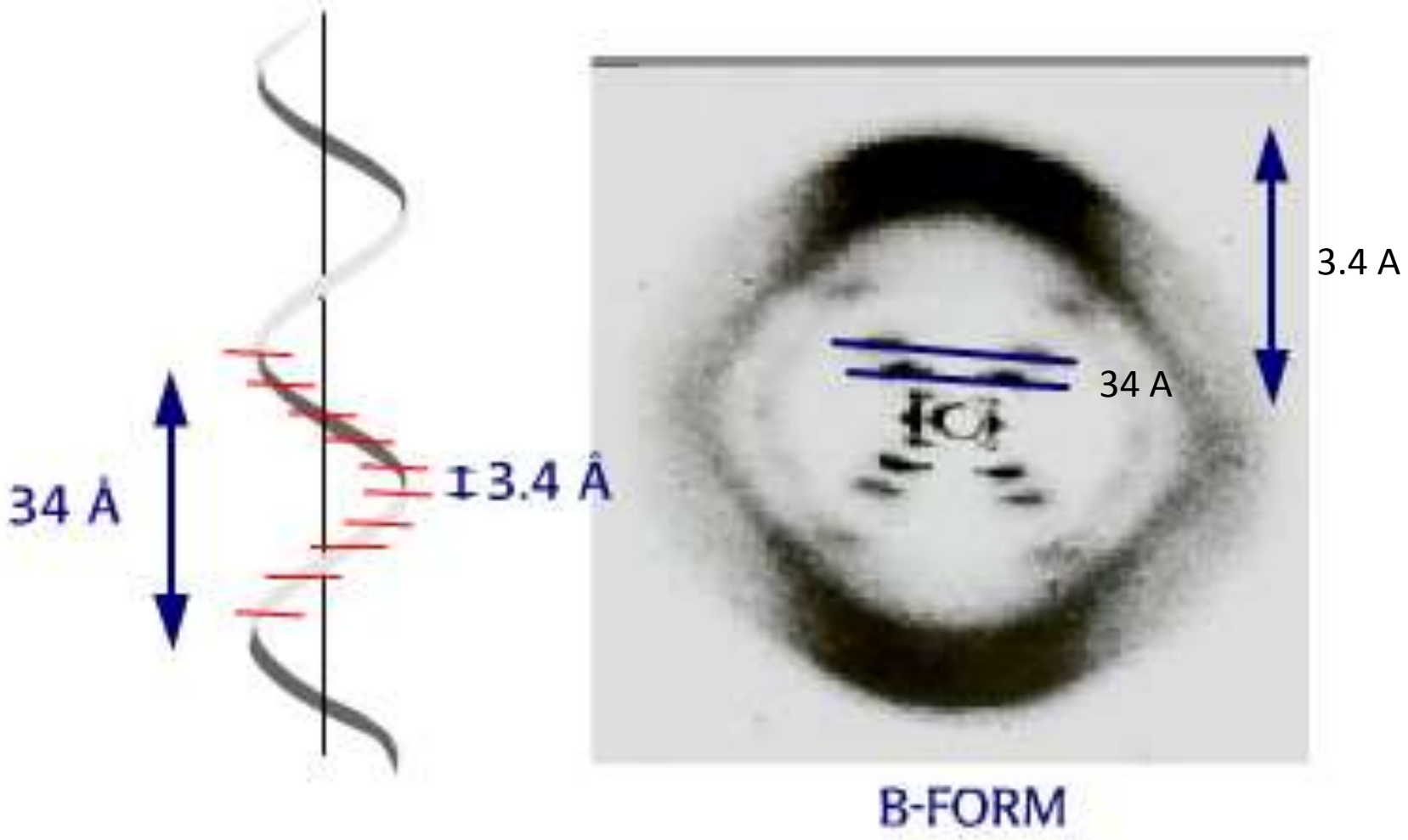
One of the failed hypothetical models of DNA is Linus Pauling's triple helix model. This structure would be unstable under normal cellular conditions.





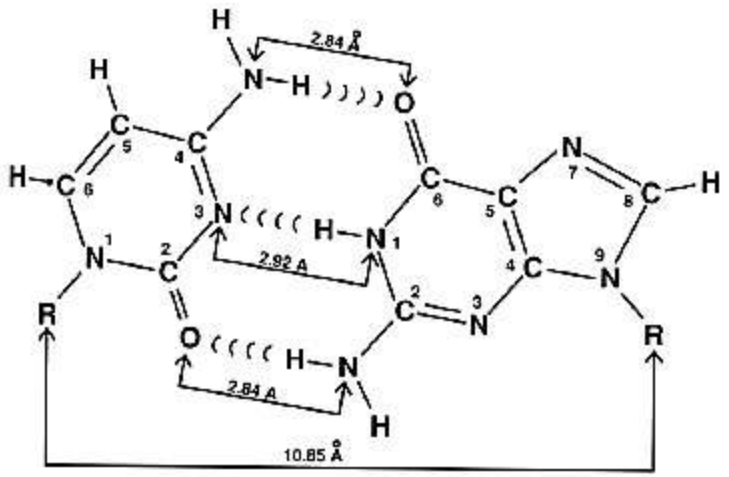
http://www.hhmi.org/biointeractive/dna/DNAi_pauling_triple_helix.html

Rosalind Franklin and Raymond Gosling's B-DNA X-ray Image (interpretation)

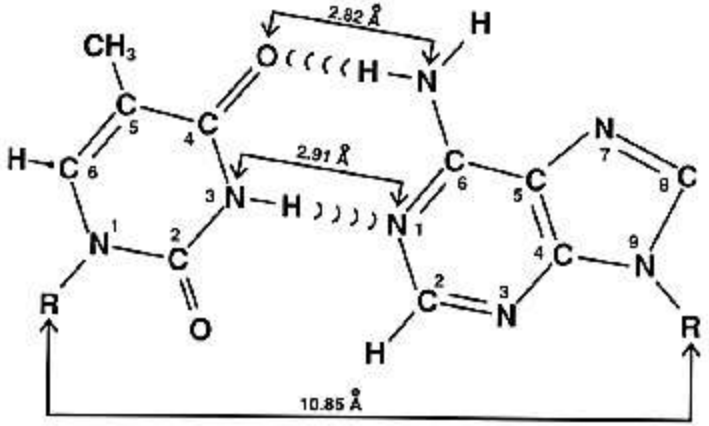


http://www.dnalc.org/view/15014-Franklin-s-X-ray-diffraction-explanation-of-X-ray-pattern-.html

Watson-Crick base pairs



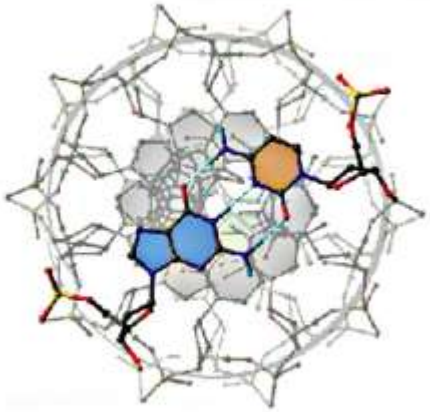
G-C base pair



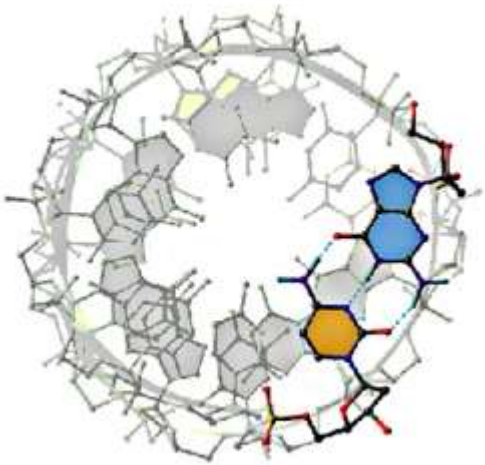
A-T base pair

HHMI: http://www.hhmi.org/biointeractive/dna/DNAi_watson_basepairing_video.html

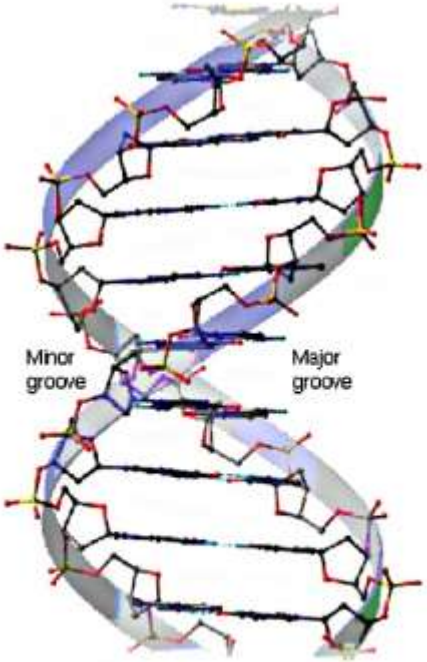
A and B Double Helices



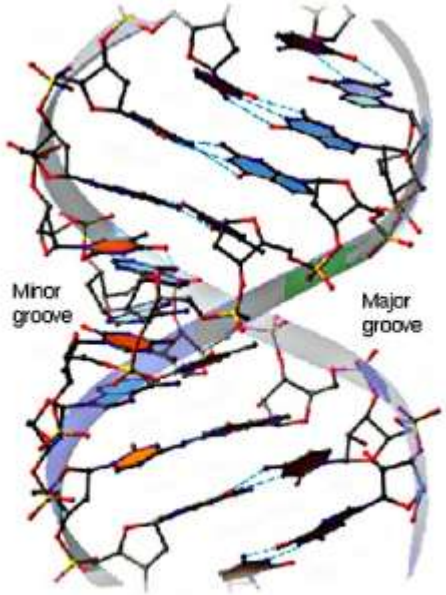
(a) B-DNA, end-on view



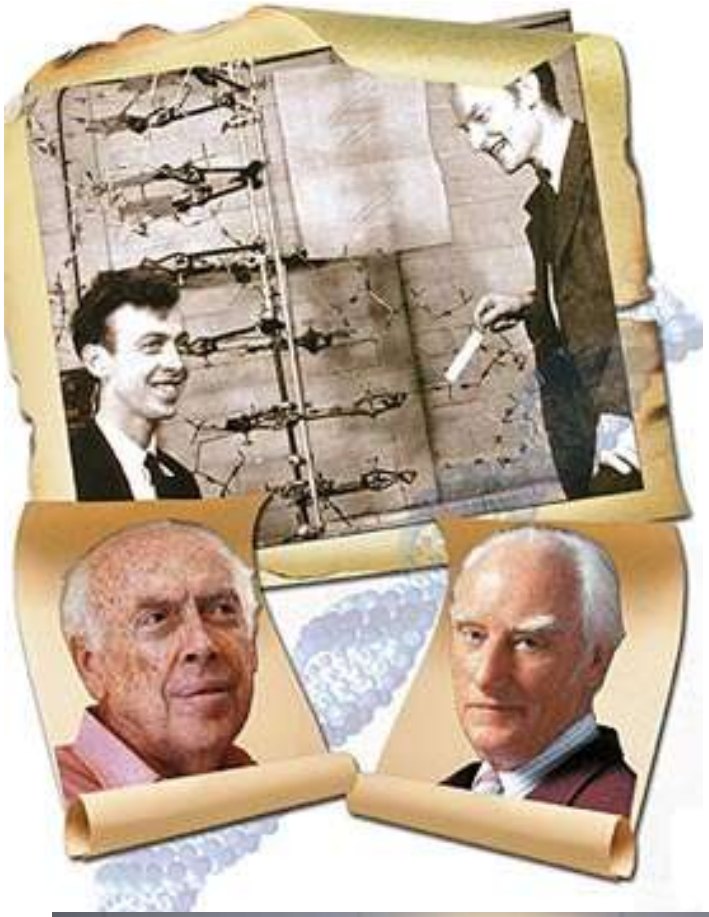
(c) A-DNA, end-on view



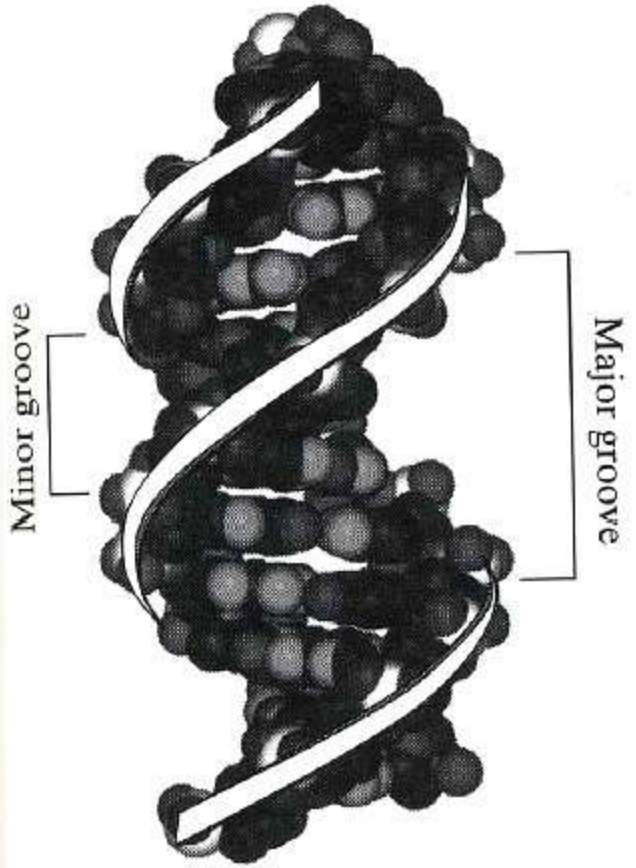
(b) B-DNA, side view



(d) A-DNA, side view



Structure of double stranded DNA (B-DNA)



B-DNA

Right handed helix

10.5 residue per turn

Helix pitch = 34Å

Base pair tilt-helical axis = -6°

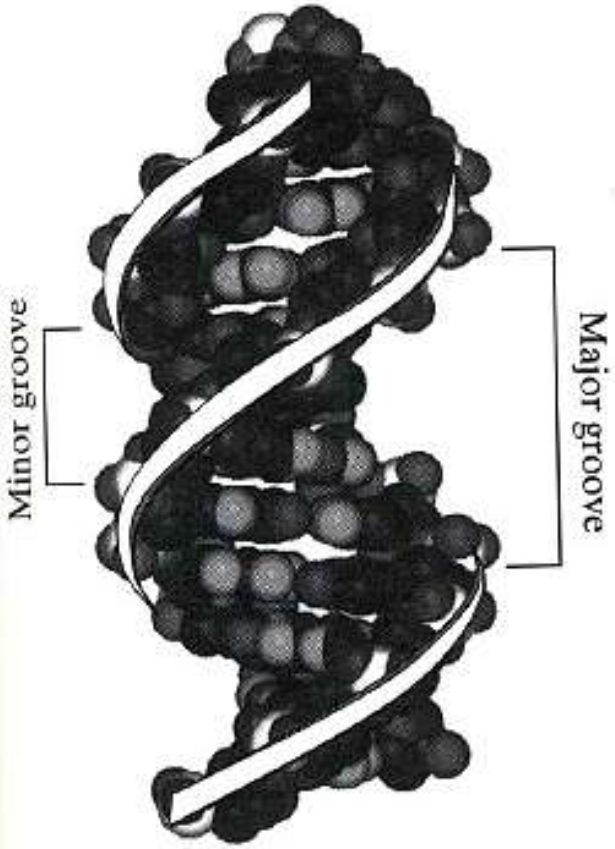
Diameter = 20Å

Sugar pucker
dA, dT, dC, dA: C2' endo

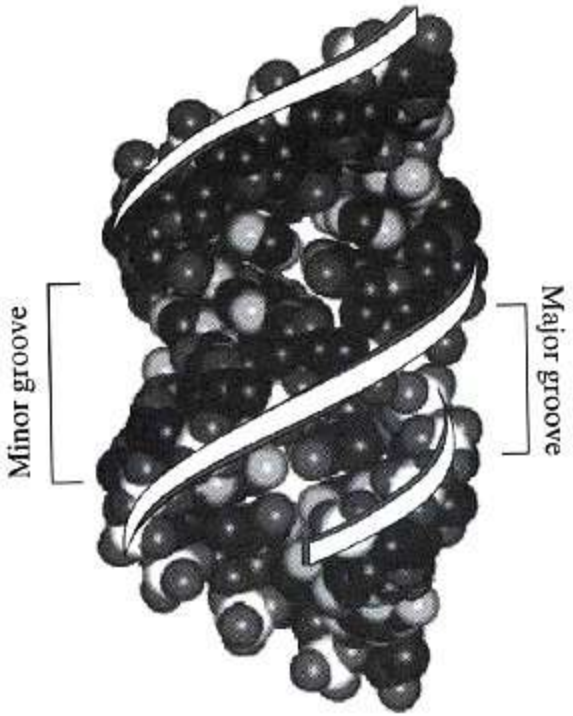
Glycosidic bond
dA, dT, dC, dA: anti

Minor grove show base diversity

A-DNA vs. B. DNA



B-DNA



A-DNA

A - DNA

11 residue per turn

Helix pitch = 28Å

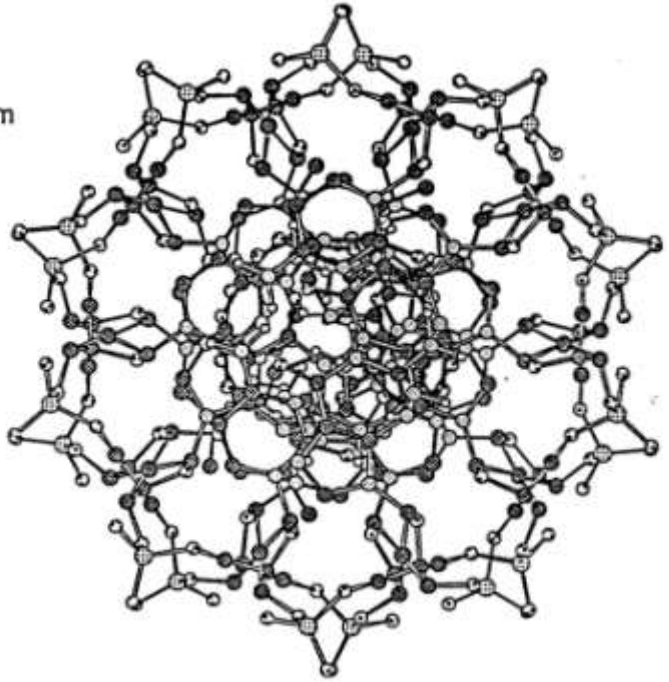
Base pair tilt = 20°

Diameter = 23Å

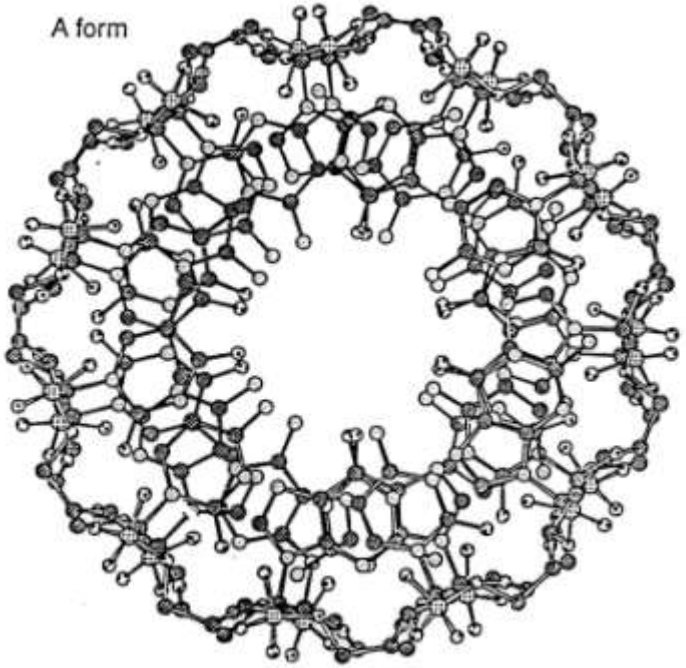
Sugar pucker
C3' endo

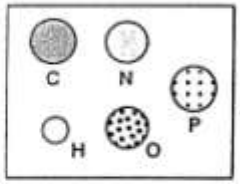
Glycosidic bond
anti

B form

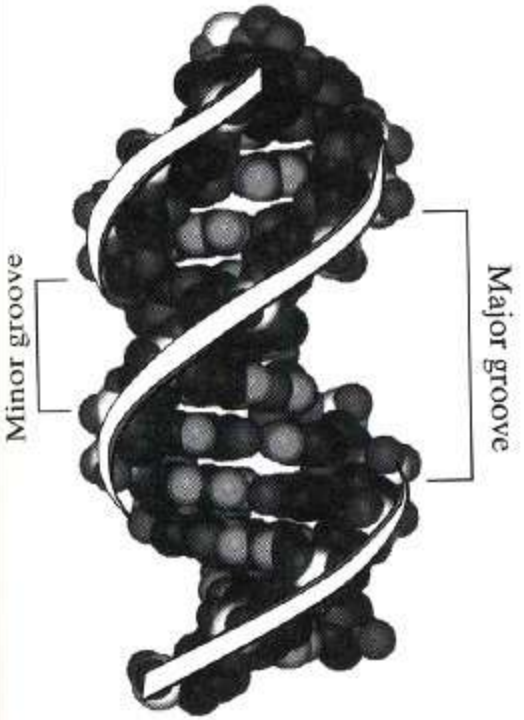


A form

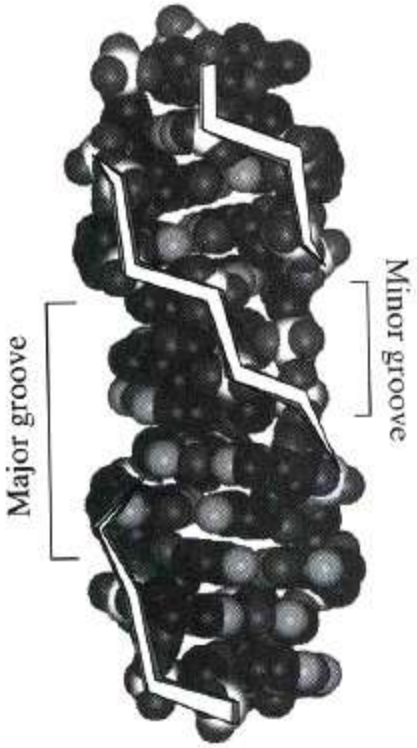




Z-DNA vs. B-DNA



B-DNA



Z-DNA

Z - DNA

left handed helix

12 residue per turn

Helix pitch = 45Å

Base pair tilt = 7°

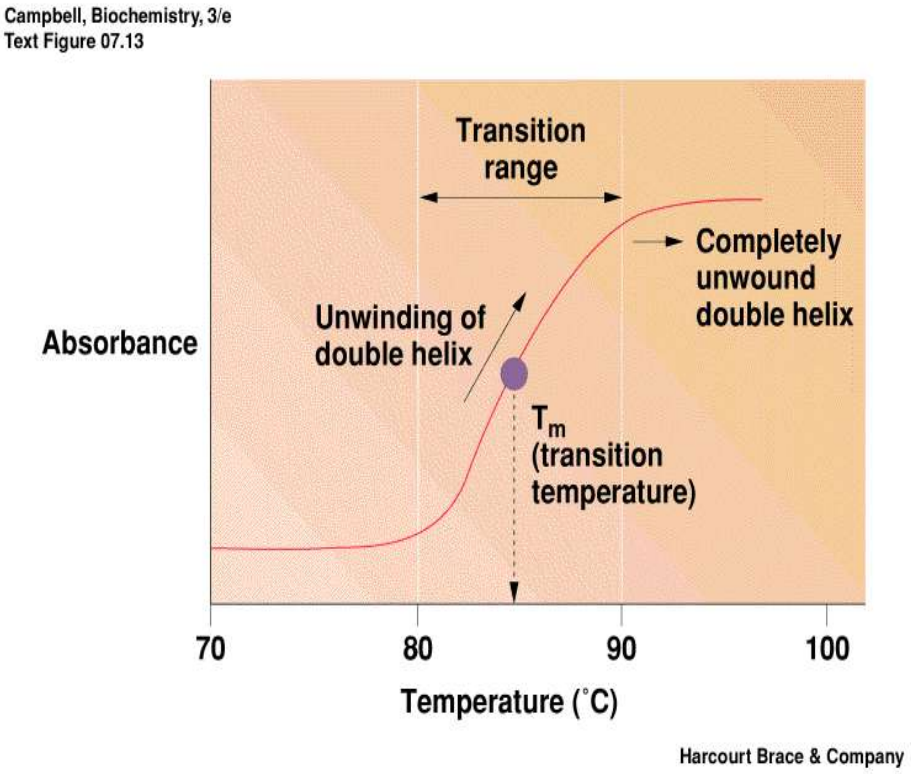
Diameter = 18Å

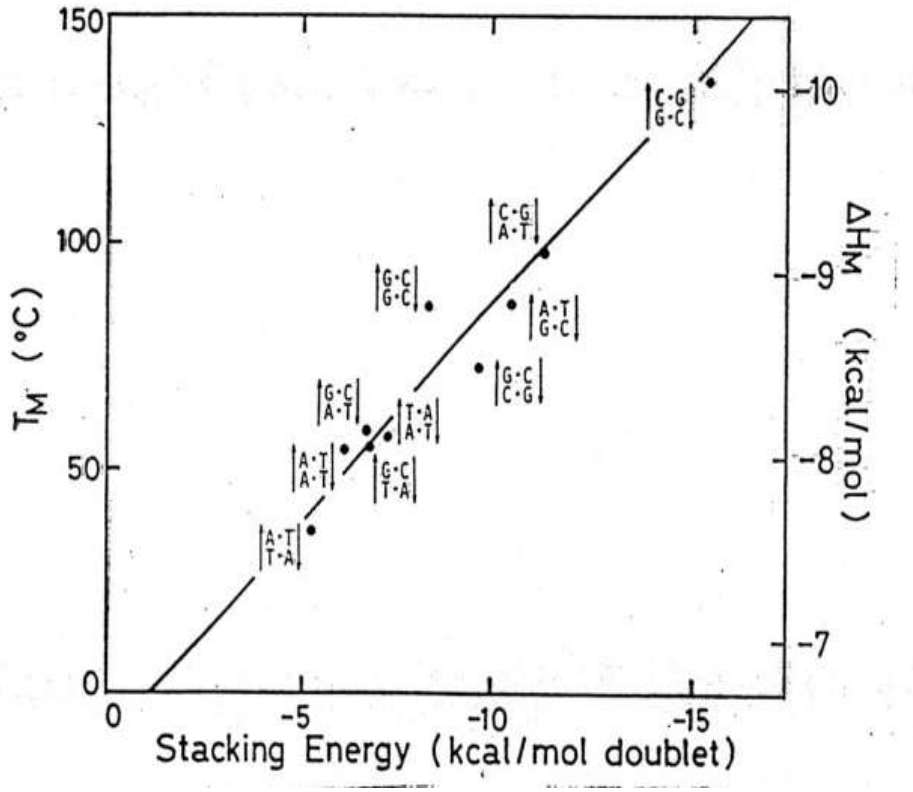
Sugar pucker
dA, dT, dC: C2' endo
dG: C3' endo

Glycosidic bond
dA, dT, dC: anti
dG: syn

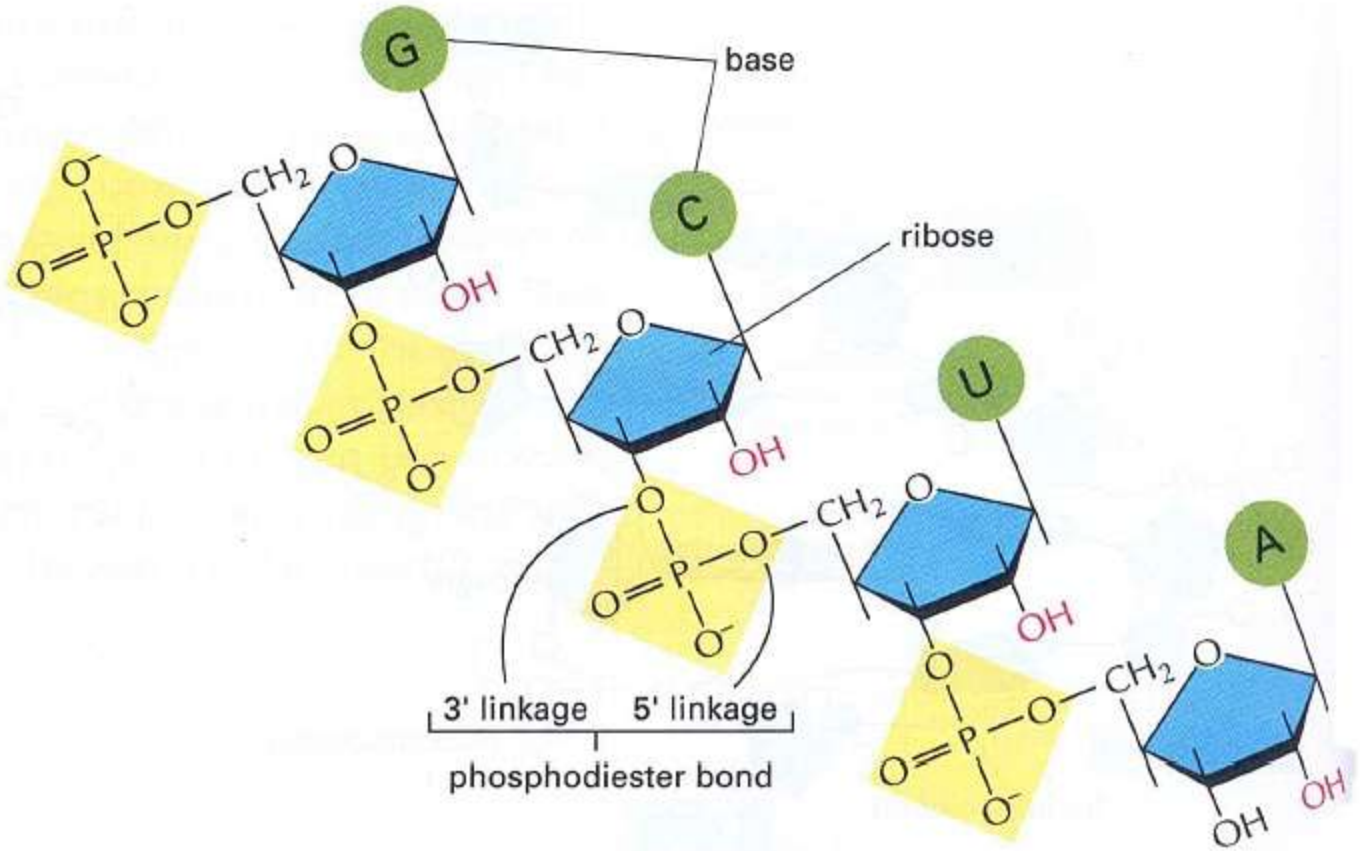
Denatuted DNA:

Heat denaturation of DNA is called "melting," The purine and pyrimidine bases exhibit very strong p-p* transitions around 260 nm. *E. coli* DNA absorption is only about 60% of that predicted from the weighted average spectrum based on its composition, this loss of intensity is called ***hypochromism***. Since the absorpance goes up as DNA "unwinds", it can be used to monitor the unstacking of DNA.

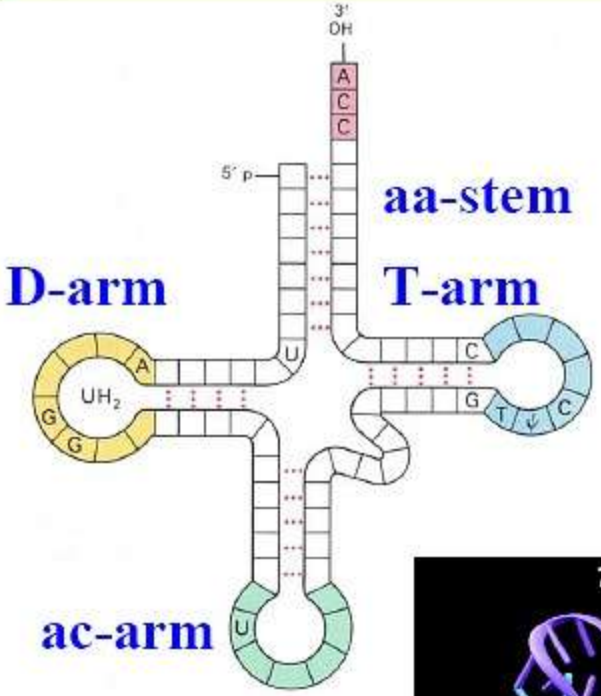


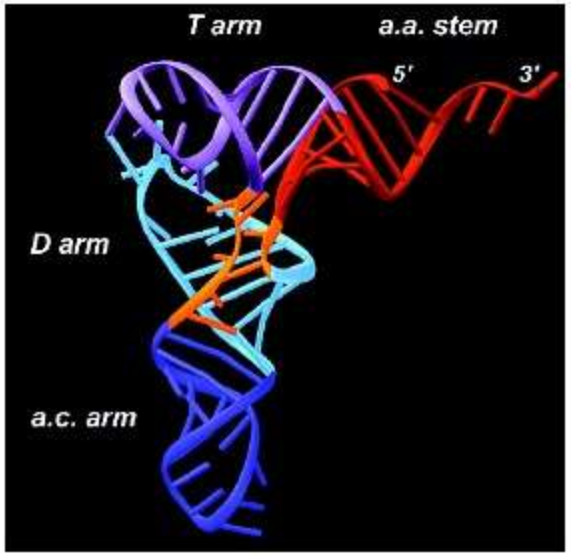


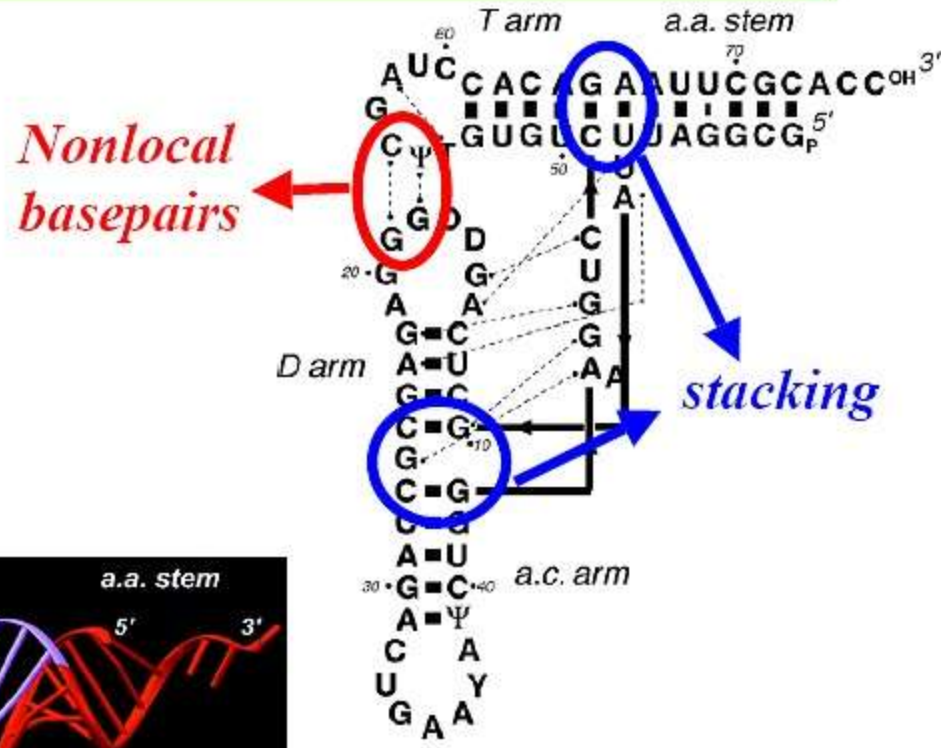
RNA primary structure



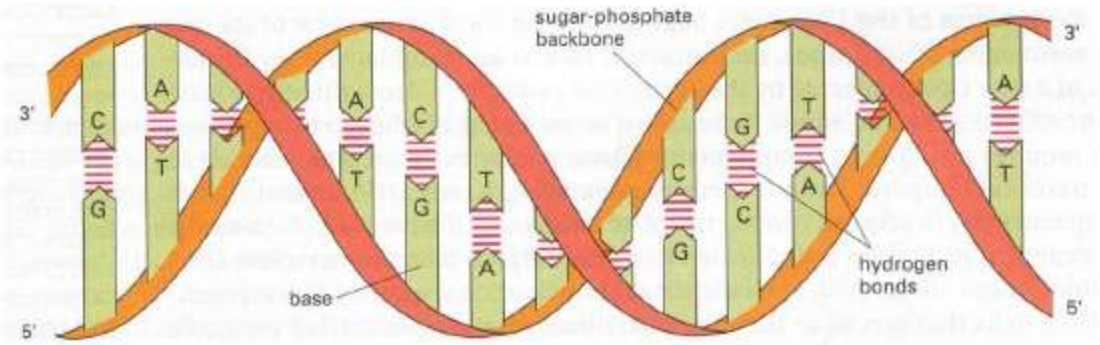
Transfer-RNA







Genetic information



... G T A C T G A A C G C A G G T...

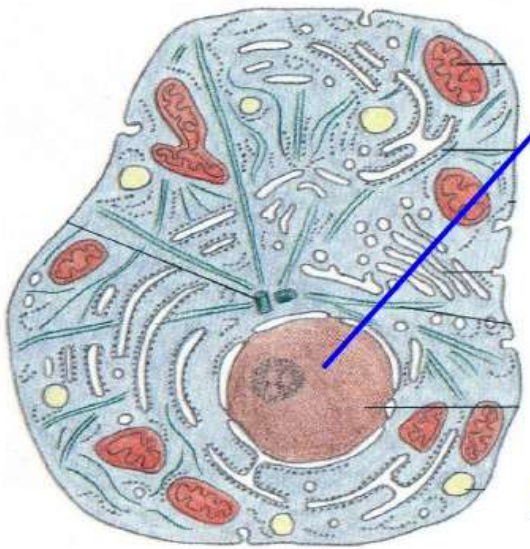
Genetic code

Human being: ~ 3,000,000,000 base-pairs
~ 30,000 – 40,000 Genes *

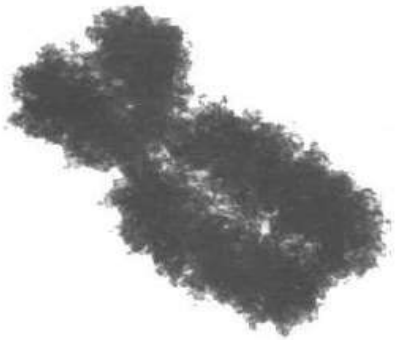
(Public Human Genome Project and Celera Genomics)

* ***Recent estimates put the number of human genes closer to 19,000!***

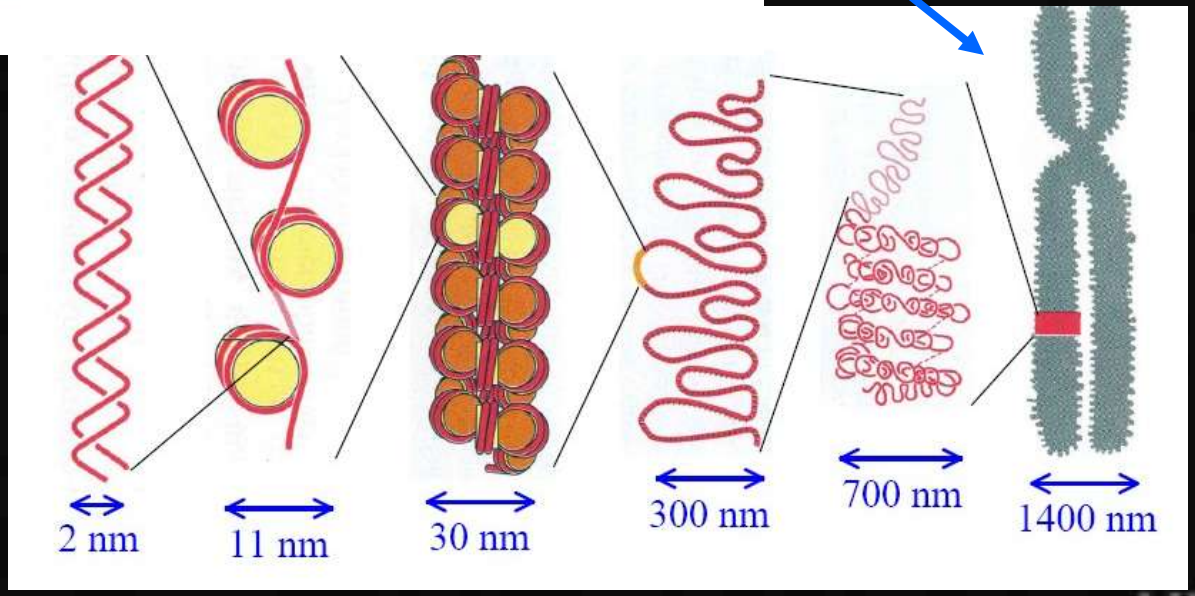
Chromosome



Chromosomes



Human being has 24 chromosomes.



HHMI: http://www.hhmi.org/biointeractive/dna/DNAi_packaging_vo2.html

Sequencing DNA

Prior to the mid-1970's no method existed by which DNA could be directly sequenced. Knowledge about gene and genome organization was based upon studies of prokaryotic organisms and the primary means of obtaining DNA sequence was so-called reverse genetics in which the amino acid sequence of the gene product of interest is back-translated into a nucleotide sequence based upon the appropriate codons.

- **Maxam-Gilbert DNA Sequencing**
- **Sanger (didexoy) DNA Sequencing**

The Nobel Prize in Chemistry 1958

"for his work on the structure of proteins, especially that of insulin"

Frederick Sanger

United Kingdom

University of Cambridge
Cambridge, United Kingdom

b. 1918

The Nobel Prize in Chemistry 1980

"for his fundamental studies of the biochemistry of nucleic acids, with particular regard to recombinant-DNA"

"for their contributions concerning the determination of base sequences in nucleic acids"

Paul Berg

1/2 of the prize

USA

Stanford University
Stanford, CA, USA

b. 1926

Walter Gilbert

1/4 of the prize

USA

Harvard University,
Biological Laboratories
Cambridge, MA, USA

b. 1932

Frederick Sanger

1/4 of the prize

United Kingdom

MRC Laboratory of
Molecular Biology
Cambridge, United
Kingdom

b. 1918

Maxam-Gilbert DNA Sequencing

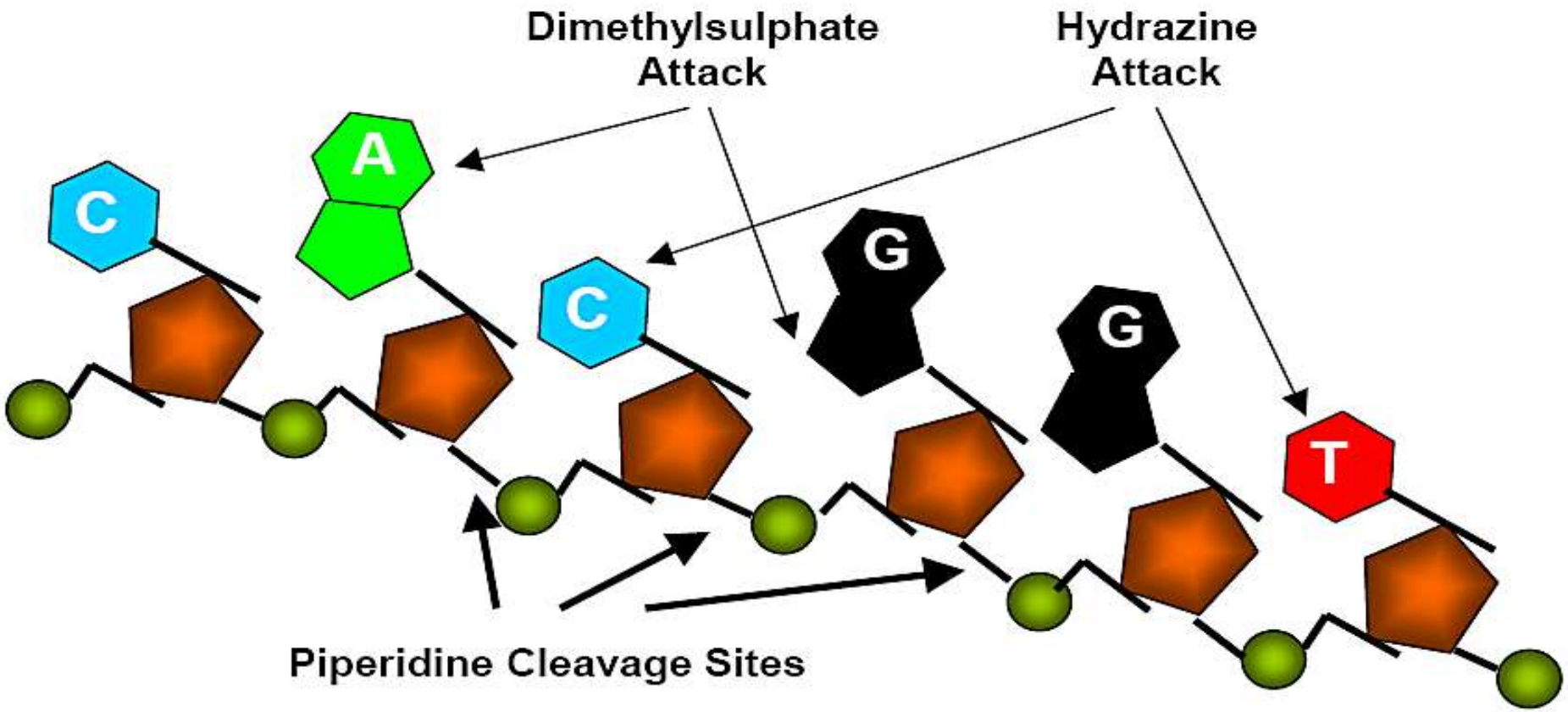


Figure 1. Chemical targets in the Maxam-Gilbert DNA sequencing strategy. Dimethylsulphate or hydrazine will attack the purine or pyrimidine rings respectively and piperidine will cleave the phosphate bond at the 3' carbon.

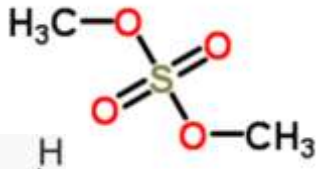
http://www.idtdna.com/support/technical/TechnicalBulletinPDF/DNA_Sequencing.pdf

Allan Maxam / Walter Gilbert DNA Sequencing

Sequencing single-stranded DNA

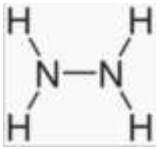
Two-step catalytic process:

1) Break glycoside bond between the ribose sugar and the base / displace base

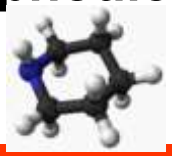


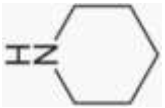
Purines react with dimethyl sulfate

Pyrimidines react with hydrazine (toxic)



2) Piperidine catalyzes phosphodiester bond cleavage where base displaced





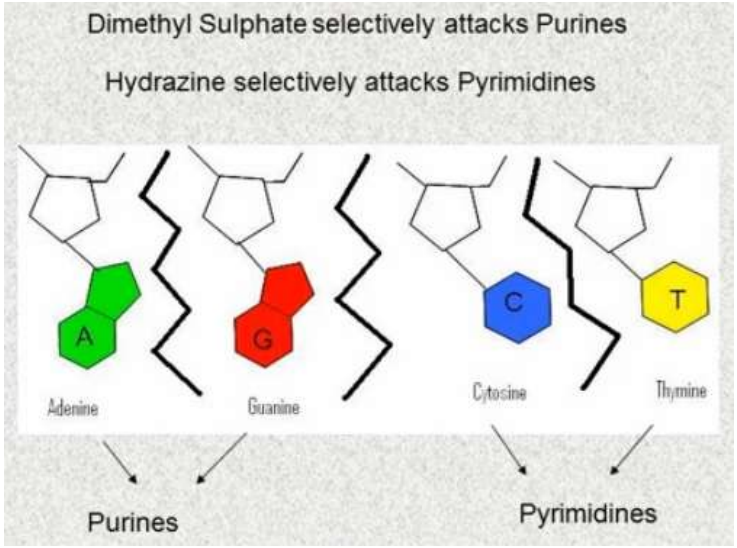
"G" - dimethyl sulfate and piperidine

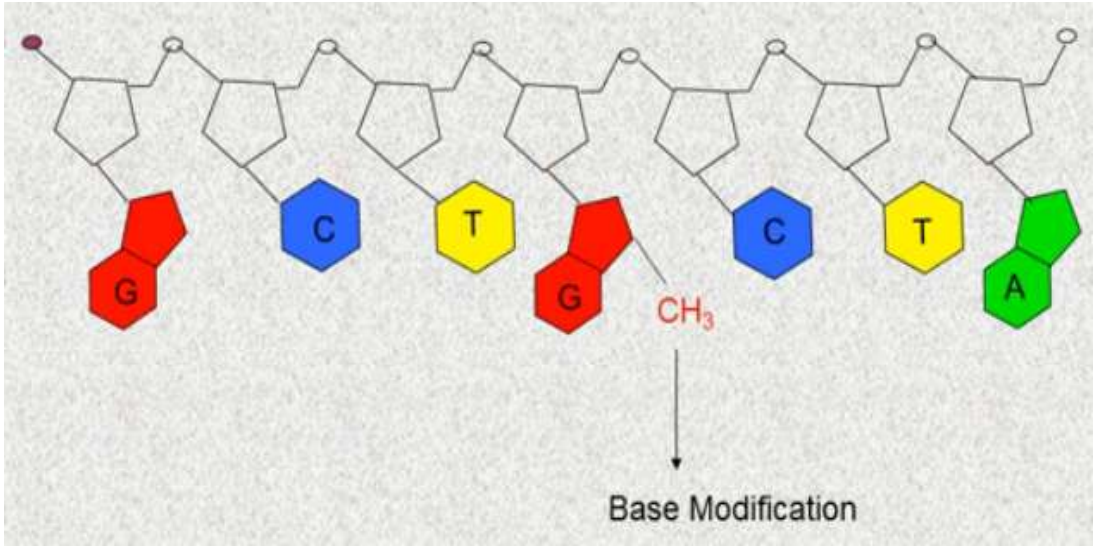
"A + G" - dimethyl sulfate and piperidine in formic acid

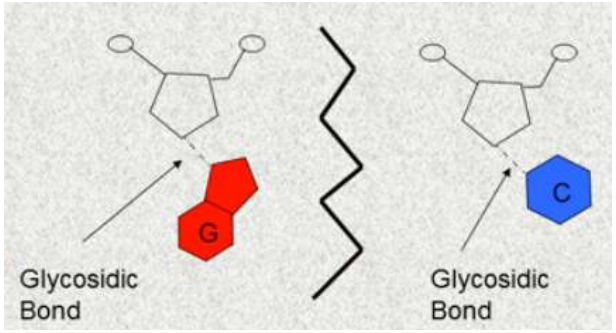
"C" - hydrazine and piperidine in 1.5M NaCl

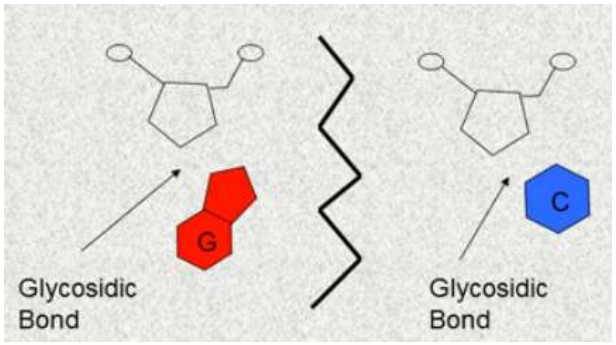
"C + T" - hydrazine and piperidine

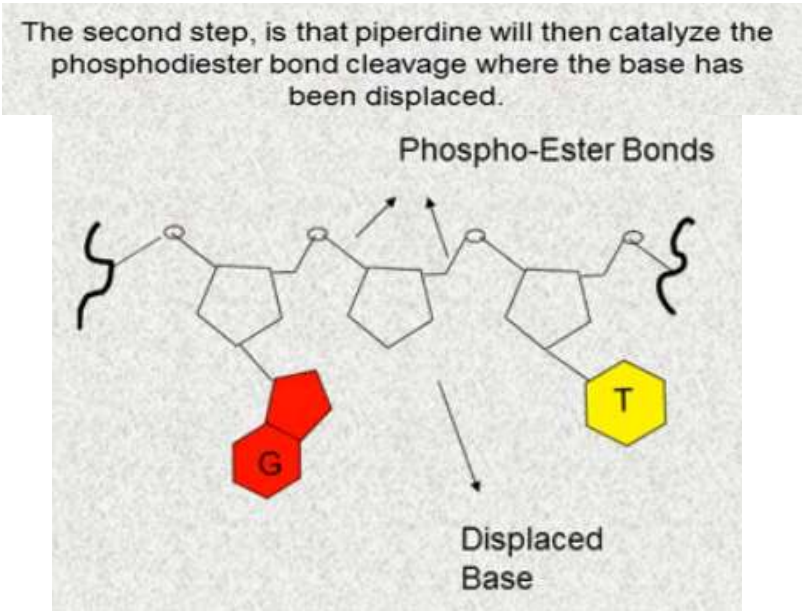
Maxam Gilbert Sequencing











Maxam Gilbert Sequencing by ChurchStreet105
http://www.youtube.com/watch?v=lqWZ-duHfu8&feature=related

Chemical Reagents and Conditions Employed For Maxam-Gilbert Sequencing.

Guanine – Dimethyl Sulphate followed by Piperdine

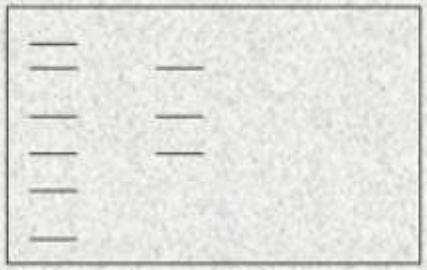
Guanine & Adenine – Dimethyl Sulphate in Formic Acid followed by Piperdine

G A +G

Chemical Reagents and Conditions Employed For Maxam-Gilbert Sequencing.

Cytosine & Thymine – Hydrazine followed by perperdine

Cytosine – Hydrazine in 2M NaCl followed by perperdine



C+T C

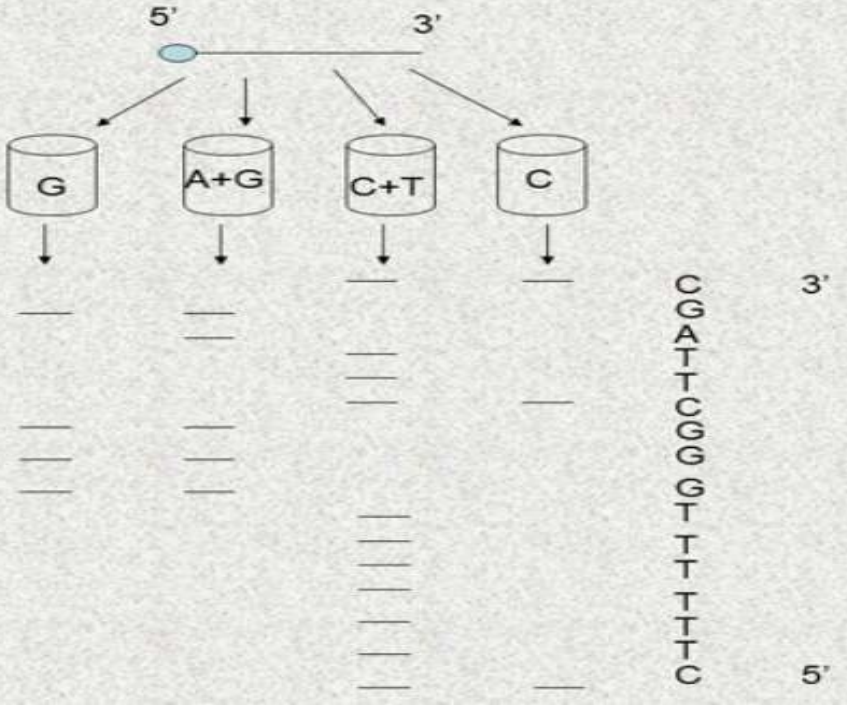
Sequenced Chain

^{32}PGpCpTpGpCpTpApGpGpTpGpCpCpGpApGpC

G G G G G G G

Cleaved Fragments

^{32}P
^{32}PGpCpTp
^{32}PGpCpTpGpCpTpAp
^{32}PGpCpTpGpCpTpApGp
^{32}PGpCpTpGpCpTpApGpGpTp
^{32}PGpCpTpGpCpTpApGpGpTpGpCpCp
^{32}PGpCpTpGpCpTpApGpGpTpGpCpCpGpAp
^{32}PGpCpTpGpCpTpApGpGpTpGpCpCpGpApGpC



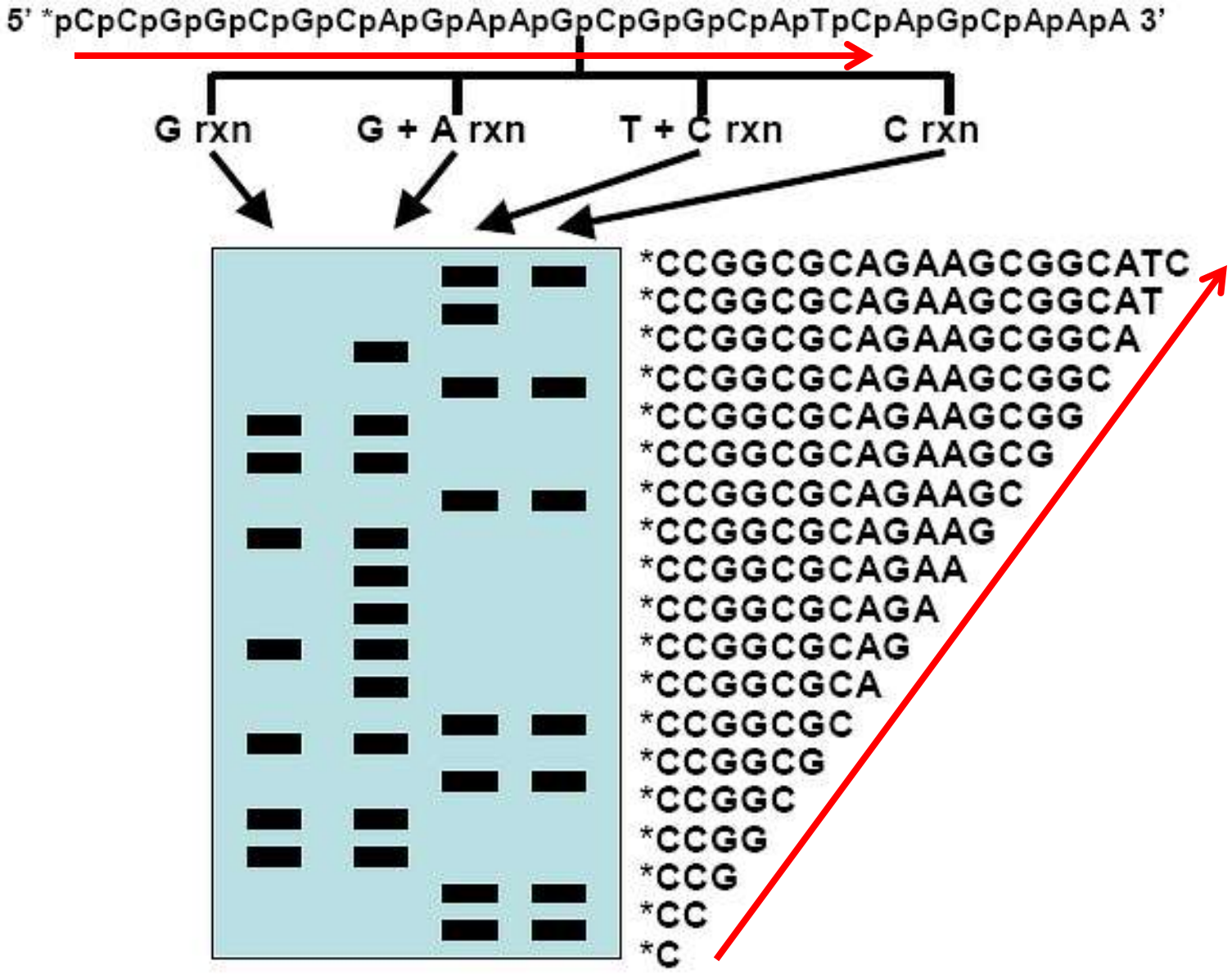


Figure 2. The Maxam-Gilbert manual sequencing scheme. The target DNA is radiolabeled and then split into the four chemical cleavage reactions. Each reaction is loaded onto a polyacrylamide gel and run. Finally, the gel is autoradiographed and base calling proceeds from bottom to top.

Proc. Natl. Acad. Sci. USA
Vol. 74, No. 2, pp. 560–564, February 1977
Biochemistry

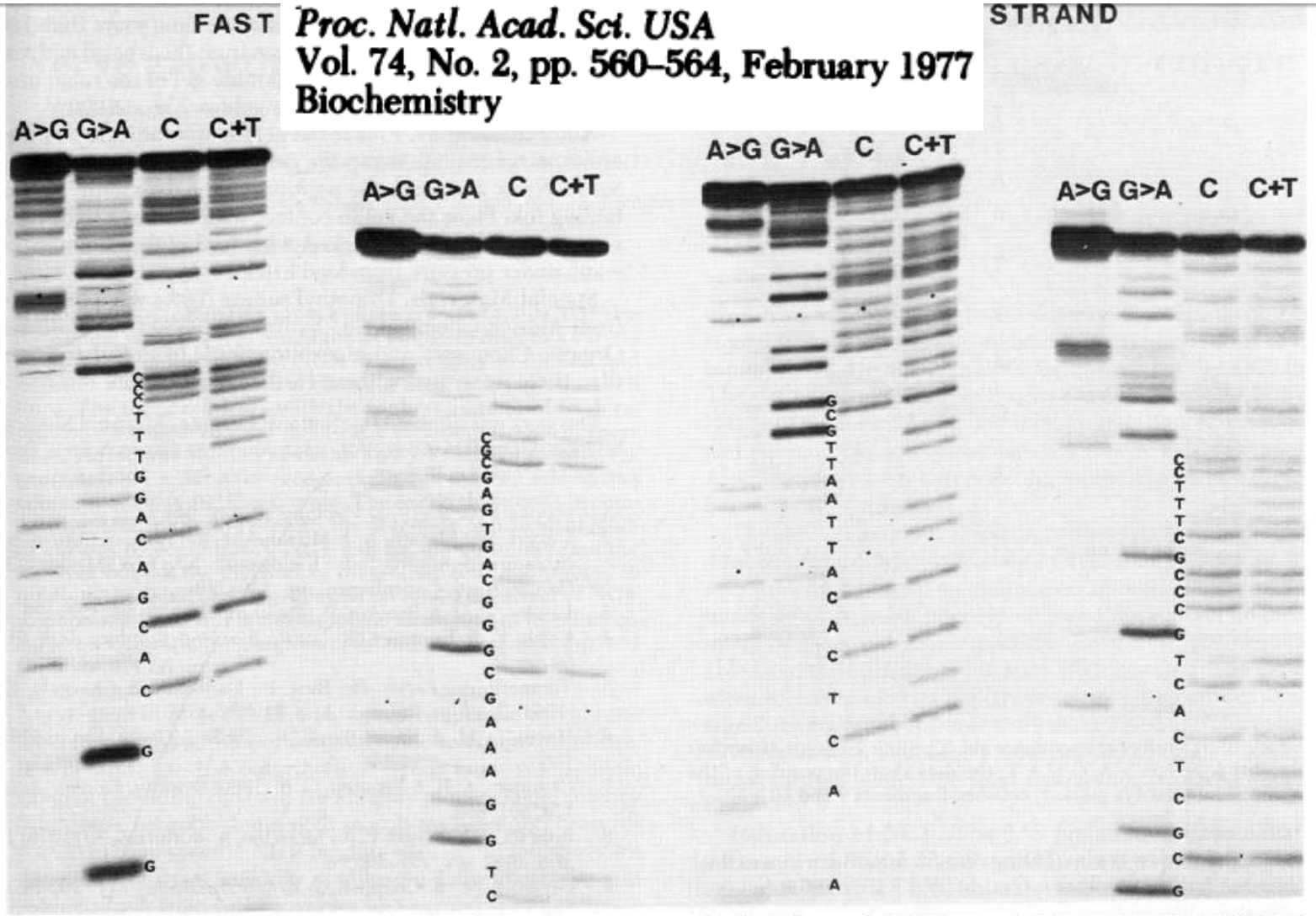


FIG. 2. Autoradiograph of a sequencing gel of the complementary strands of a 64-base-pair DNA fragment. Two panels, each with four reactions, are shown for each strand; cleavages proximal to the 5′ end are at the bottom on the left. A strong band in the first column with a weaker band in the second arises from an A; a strong band in the second column with a weaker band in the first is a G; a band appearing in both the third and fourth columns is a C; and a band only in the fourth column is a T. To derive the sequence of each strand, begin at the bottom of the left panel and read upward until the bands are not resolved; then, pick up the pattern at the bottom of the right panel and continue upward. One-tenth of each strand, isolated from the gel of Fig. 1, was used for each of the base-modification reactions. The dimethyl sulfate treatment was 50 mM for 30 min to react with A and G; hydrazine treatment was 18 M for 30 min to react with C and T and 18 M with 2 M NaCl for 40 min to cleave C. After strand breakage, half of the products from the four reactions were layered on a 1.5 × 330 × 400 mm denaturing 20% polyacrylamide slab gel, pre-electrophoresed at 1000 V for 2 hr. Electrophoresis at 20 W (constant power), 800 V (average), and 25 mA (average) proceeded until the xylene cyanol dye had migrated halfway down the gel. Then the rest of the samples were layered and electrophoresis was continued until the new bromphenol blue dye moved halfway down. Autoradiography of the gel for 8 hr produced the pattern shown.

Maxam-Gilbert DNA Sequencing

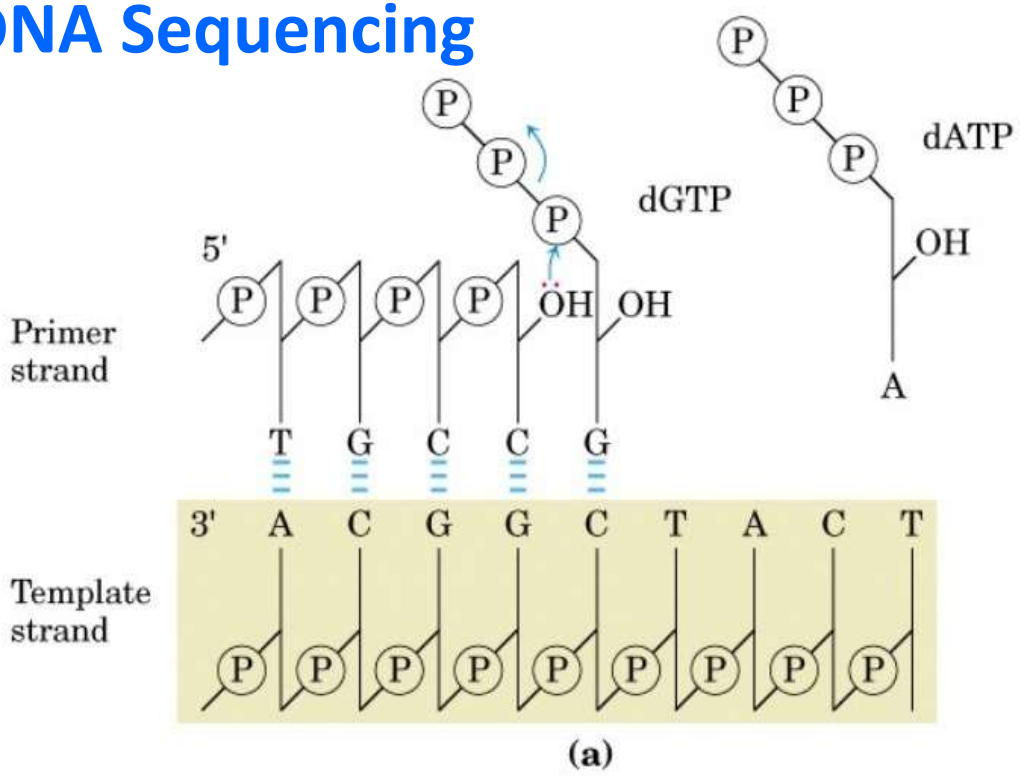
- **200-300 bases of DNA sequence every few days**
- **Use large amounts of radioactive material, ^{35}S or ^{32}P**
- **Constantly pouring large, paper thin acrylamide gels**
- **Hydrazine is a neurotoxin**

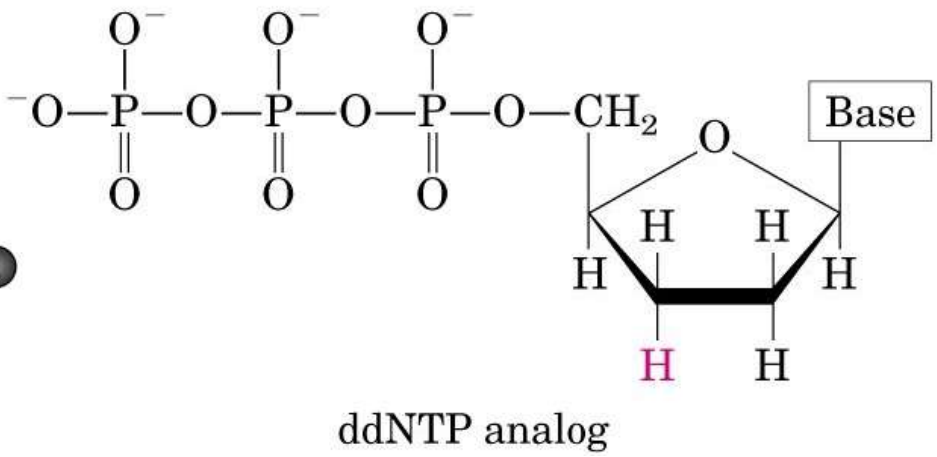
Early Benefits -

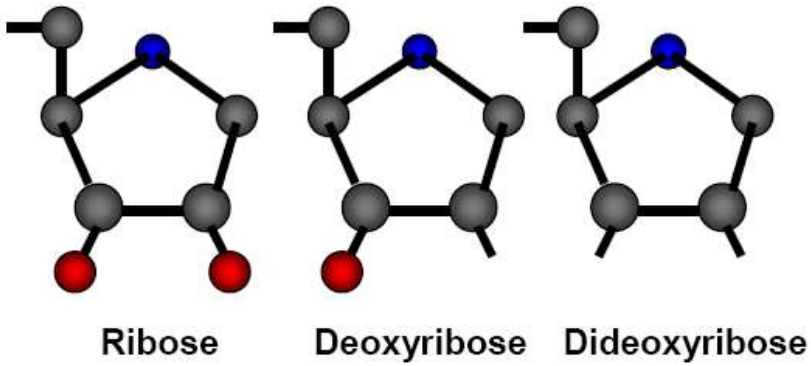
Discovery that the gene for ovalbumin in chicken and the gene encoding β-globin in rabbit contained non-coding gaps in the coding regions. These gaps were flanked by the same dinucleotides in the two genes; GT on the 5' end of the gaps and AG on the 3' end of the gaps. Soon, the terms intron and exon were added to the genetic lexicon to describe the coding and non-coding regions of eukaryotic genes (1977).

Fred Sanger (dideoxy) DNA Sequencing

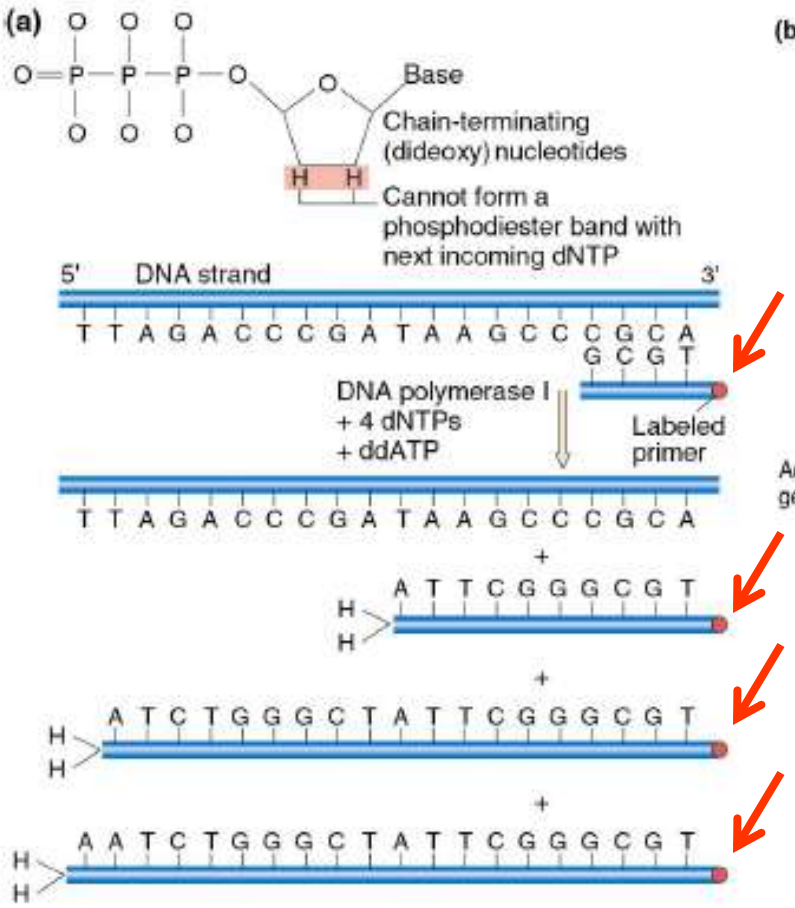
Sanger knew that, whenever a dideoxynucleotide was incorporated into a polynucleotide, the chain would irreversibly stop, or terminate. Thus, the incorporation of specific dideoxynucleotides in vitro would result in selective chain termination.

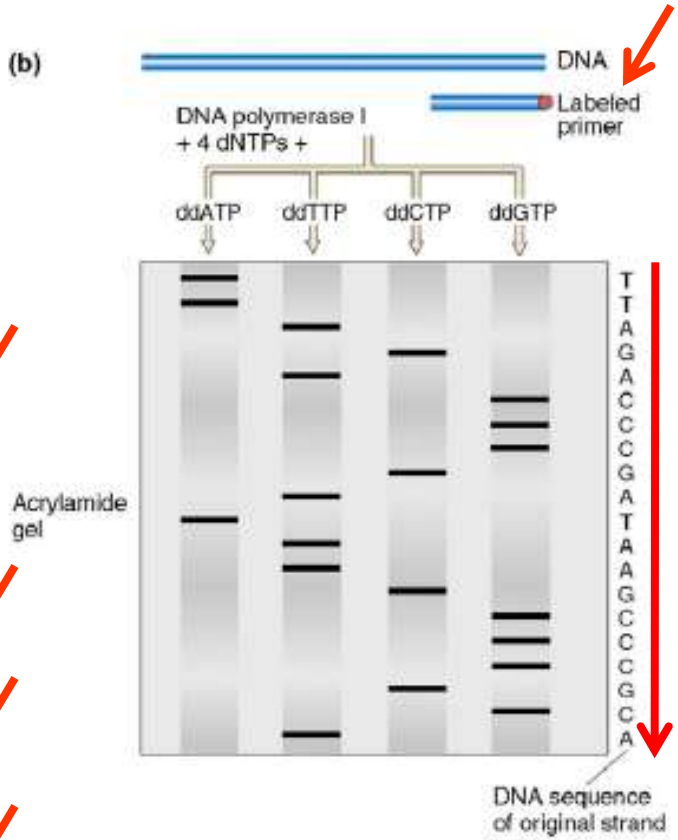


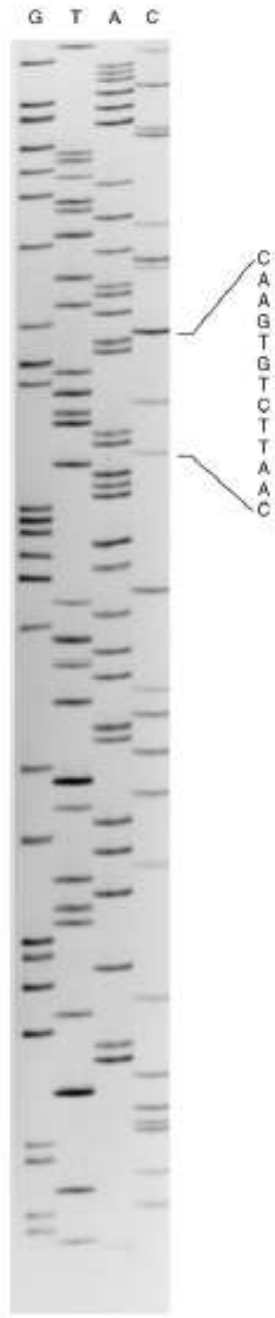




Sanger (dideoxy) DNA Sequencing







Advantages of dideoxy DNA Sequencing

- **Elimination of dangerous chemicals (hydrazine)**
- **Greater efficiency**

Chemical synthesis of a 25-mer oligonucleotide takes more than two hours.

Taq polymerase makes DNA strands off of a template at rate of about 500 bases per minute (20 x 120 = 2400)

→ **High Throughput Methods (Human Genome Project)**

Automated Fluorescence Sequencing

In 1986, Leroy Hood and colleagues reported on a DNA sequencing method in which the radioactive labels, autoradiography, and manual base calling were all replaced by fluorescent labels, laser induced fluorescence detection, and computerized base calling.

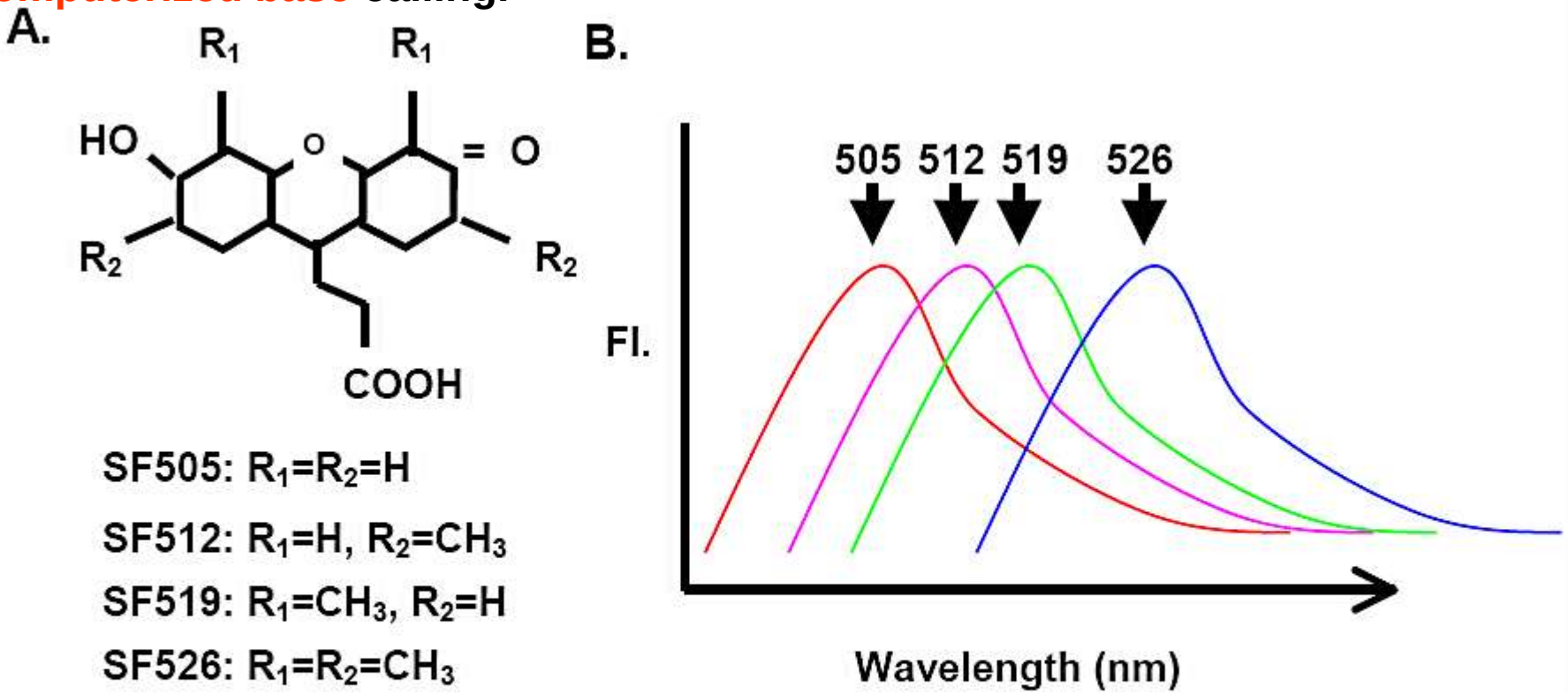
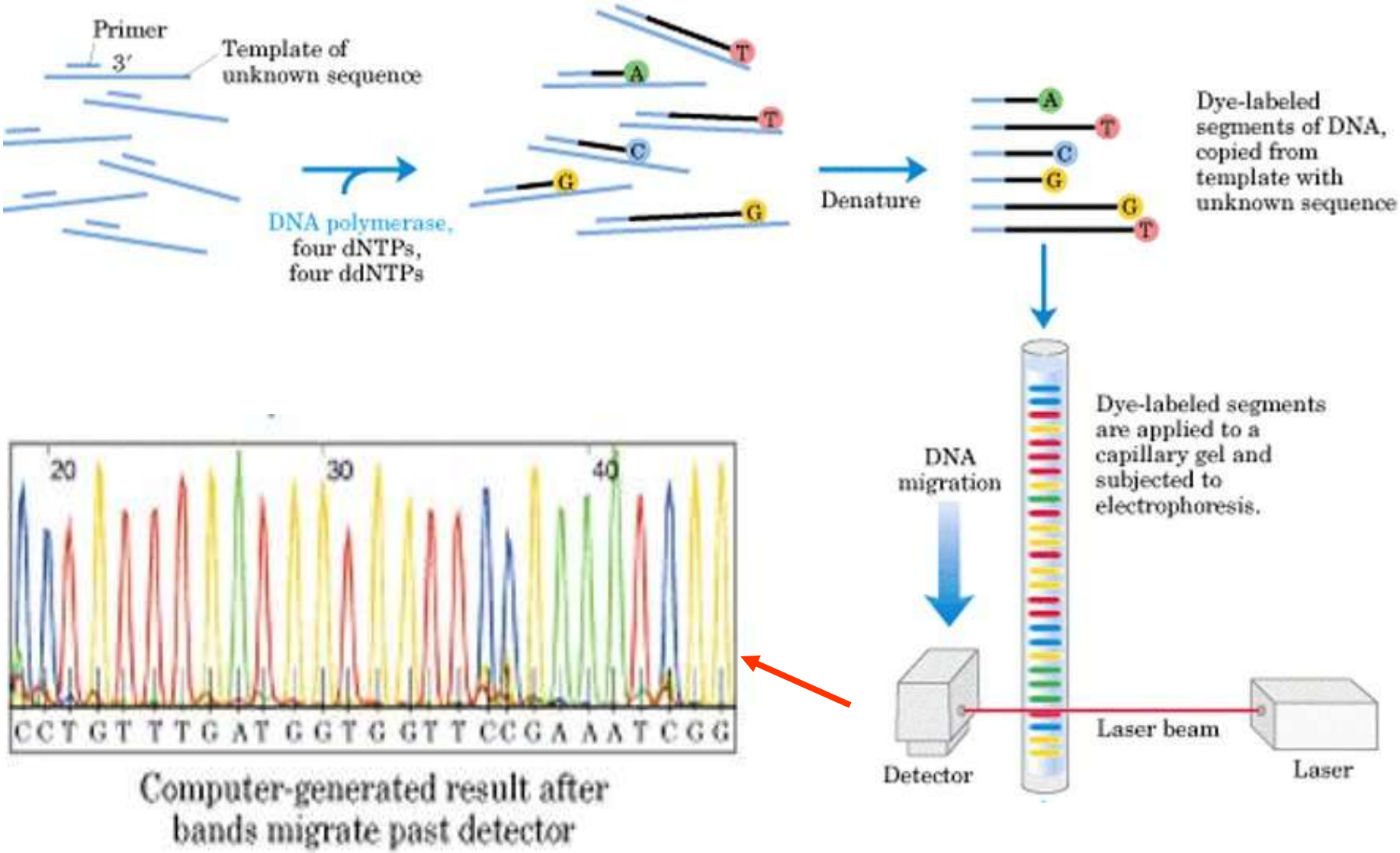


Figure 5. A. Chemical structure of the four succinylfluorescein dyes developed at DuPont. B. Normalized fluorescence emission spectra for each of the four dyes following excitation at 488nm. Shifts in the spectra were achieved by changing the side groups R_1 and R_2.

Automated DNA sequencing



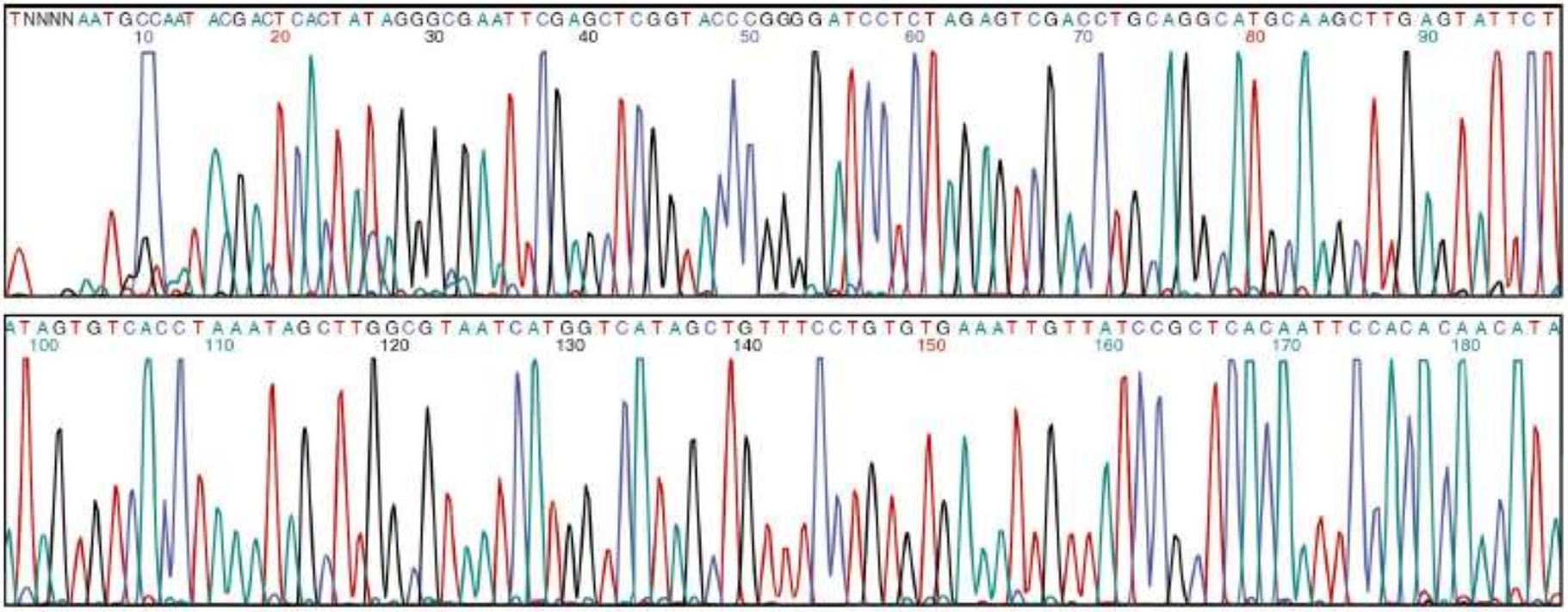
Automated dye-terminator sequencing

4-fluorescently labelled dideoxy dye terminators

ddATP
ddGTP
ddCTP
ddTTP

pool and load in a single well or capillary

- scan with laser + detector specific for each dye
- automated base calling
- very long reads (~ 1000 bases)/run



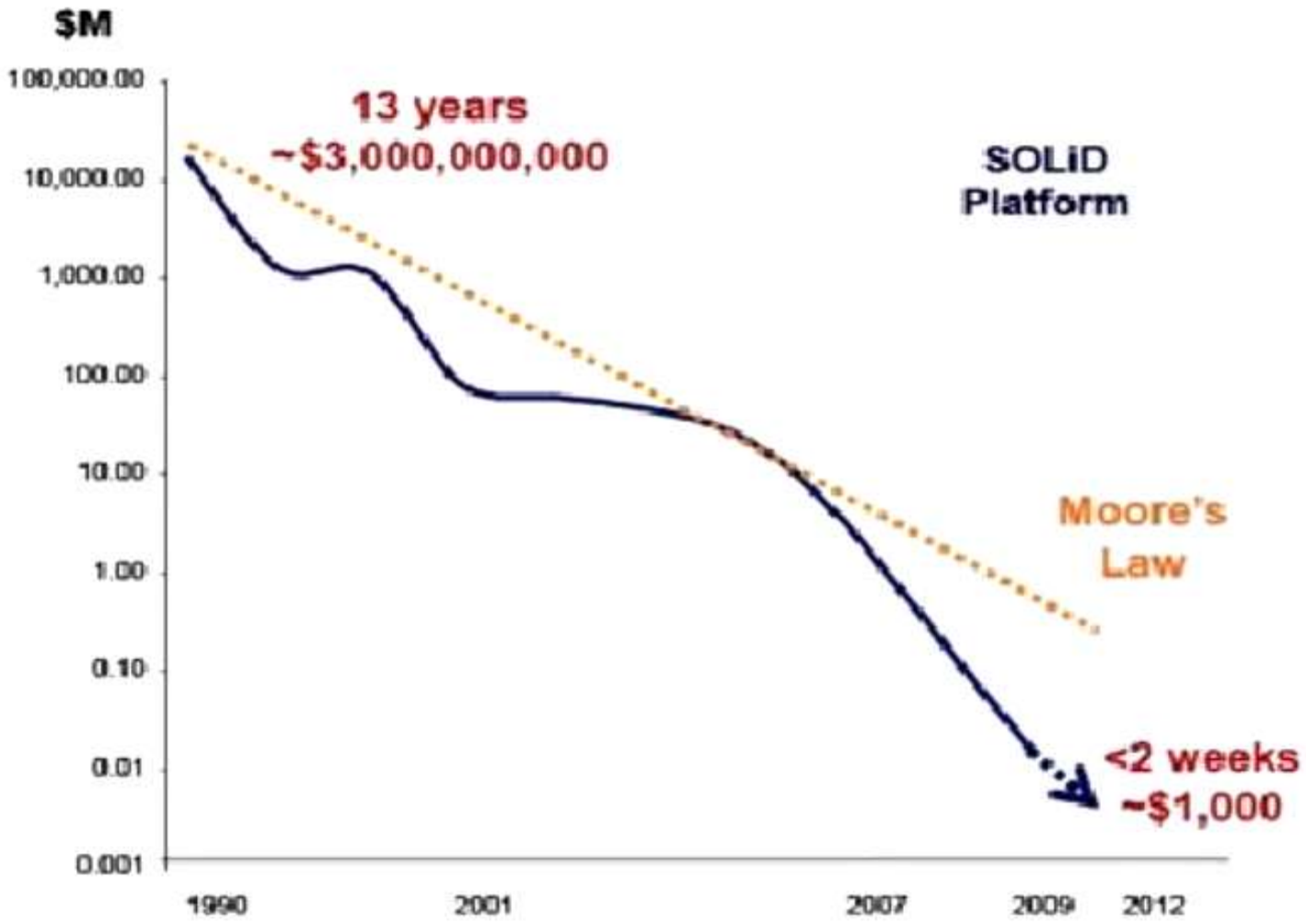
[DNA dideoxysequencing animation]

Sequencing DNA

Prior to the mid-1970's no method existed by which DNA could be directly sequenced. Knowledge about gene and genome organization was based upon studies of prokaryotic organisms and the primary means of obtaining DNA sequence was so-called reverse genetics in which the amino acid sequence of the gene product of interest is back-translated into a nucleotide sequence based upon the appropriate codons.

- **Maxam-Gilbert DNA Sequencing** – chemical sequencing
- **Sanger (didexoy) DNA Sequencing** – dideoxy sequencing
- **Next Generation DNA Sequencing (Applications)**
  - *Illumina* – bridging PCR / reversible dye terminator
  - *454 sequencing* – emulsion PCR / pyrosequencing

Cost per Human Genome



http://blogs.forbes.com/sciencebiz/2010/06/03/your-genome-is-coming/

Sanger sequencing

Robert Koch

Clonal population

Frederick Sanger

Chemistry

Leroy Hood

Automation

NextGen sequencing

- 'Bridging' PCR
  - Dye terminator
    - illumina
- Emulsion PCR
  - Ligation
    - Applied Biosystems
  - Pyro-sequencing
    - 454 SEQUENCING

Next-Generation Sequencers

		Read Length	Gb/run	Technology
illumina	GA_{IIx}	2 x 100+ bp	20+ Gb	• Bridge amplification • Reversible terminators
454 SEQUENCING	GS FLX Titanium	1 x 400-600bp 2 x 140-200bp	0.4-0.6 Gb	• Emulsion PCR amplification • Homopolymers detected by an increase in signal proportional to length
applied biosystems	SOLiD 3	2 x 50bp	20+ Gb	• Emulsion PCR amplification • Ligation-based sequencing • Alignment in color space
Helicos BioSciences Corporation	Single Molecule Sequencer	2 x 25-55bp	21-28 Gb	• No amplification • Single molecule sequencing

Monday: Next-Gen Sequencing (Dr. Hunicke-Smith / Wedneday: Expression (Dr. Zhang)

The University of Texas at Austin, Genomic Sequencing and Analysis Facility

or

for short

The Good, Bad, and Ugly of Next-Gen Sequencing

Scott Hunicke-Smith

THE SPECTACULAR ARCHITECTURE OF THE RIBOSOME AND CLUES ABOUT ITS ORIGIN

The following slides on ribosome structure are taken from material presented at the IUCr Congress in Madrid 2011.

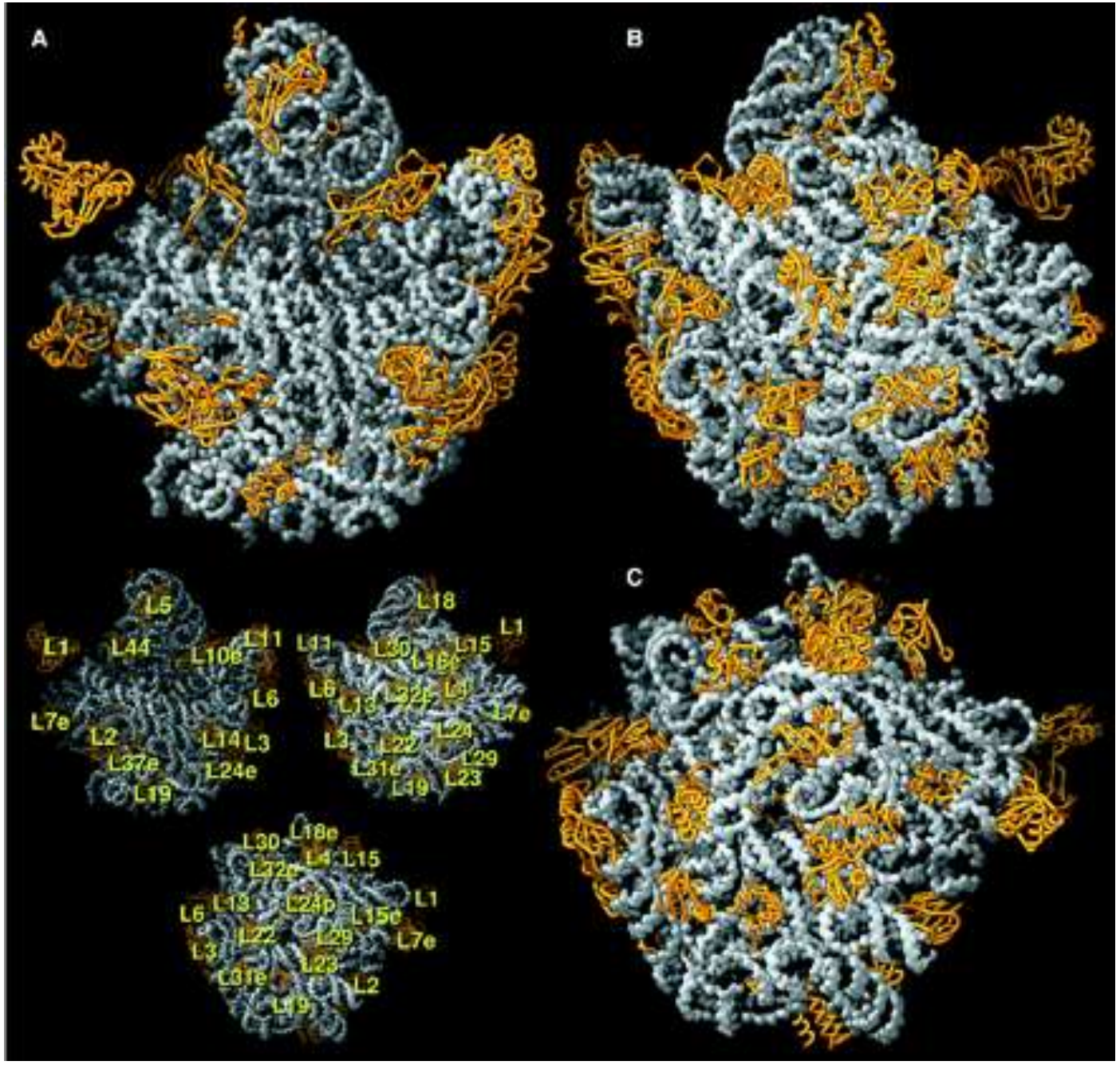
ADA YONATH

WEIZMANN INSTITUTE OF SCIENCE

Ribosomes have been considered non-crystallizable

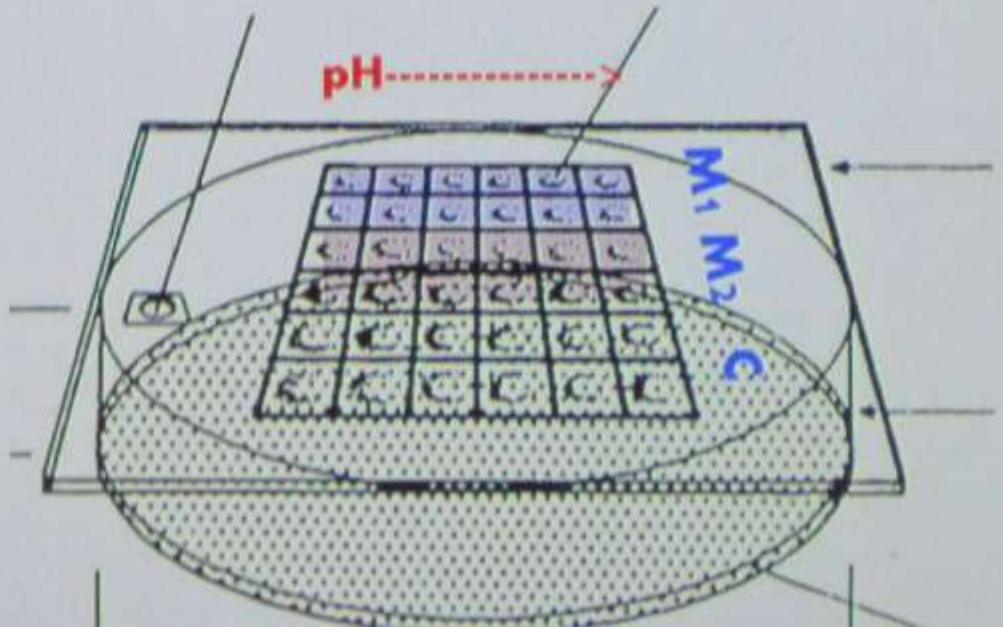
owing to:

- their high degree of internal mobility
- considerable flexibility
- functional heterogeneity
- marked tendency to deteriorate
- chemical complexity
- large size and asymmetric nature

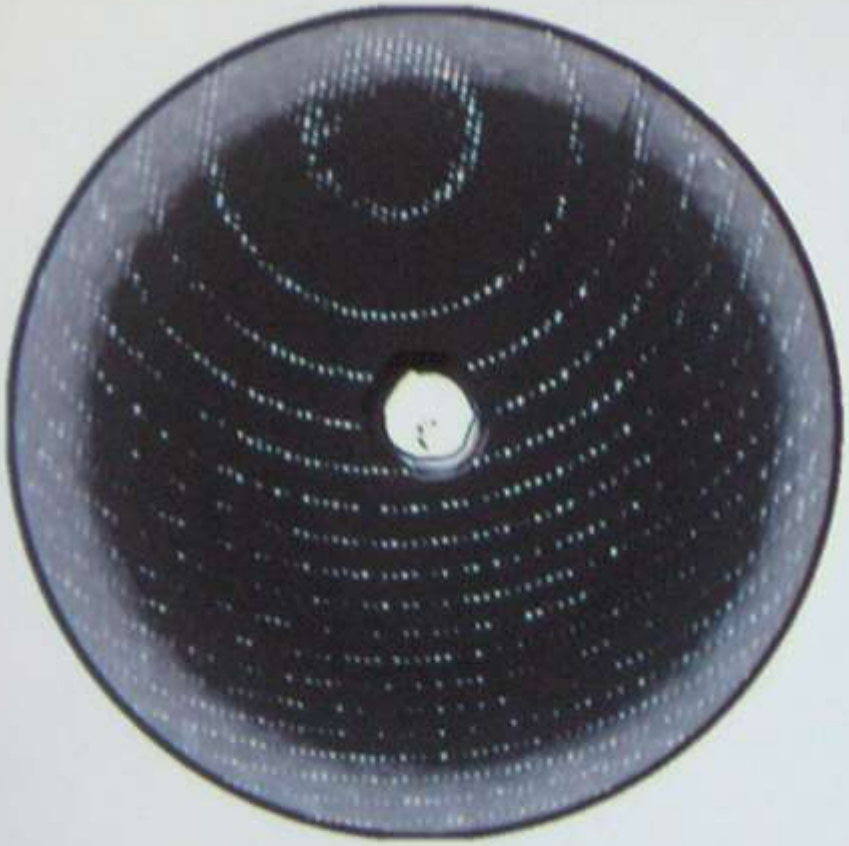

A
B
C
L5
L1
L44
L10e
L11
L41
L18
L1
L30
L15
L16e
L6
L13
L32e
L9
L7e
L7e
L2
L14
L3
L3
L22
L24
L37e
L24e
L31e
L29
L19
L49
L23
L18e
L30
L4
L15
L32e
L1
L6
L13
L24p
L15e
L22
L29
L7e
L3
L23
L31e
L2
L19

Head Full of Ribosomes

Micro crystals, not useful for crystallographic analysis, but indicating high potential, were obtained after screening of 25,000 conditions (within 6 months)

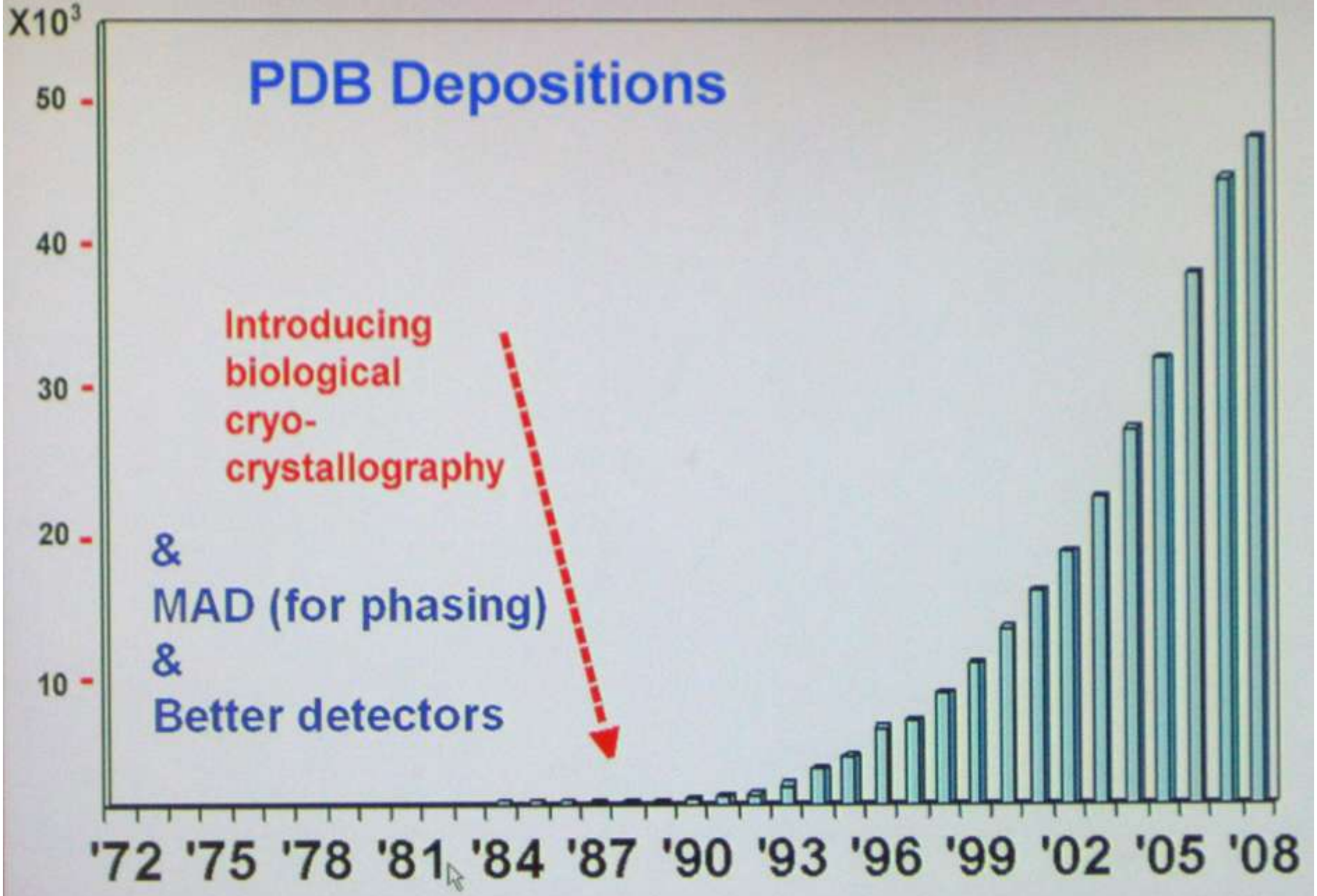


H50S diffraction, 1986



Time = 0

After Exposure of 0.1 second


X10^3
PDB Depositions
50
40
30
20
10
Introducing
biological
cryo-
crystallography
&
MAD (for phasing)
&
Better detectors
'72 '75 '78 '81 '84 '87 '90 '93 '96 '99 '02 '05 '08