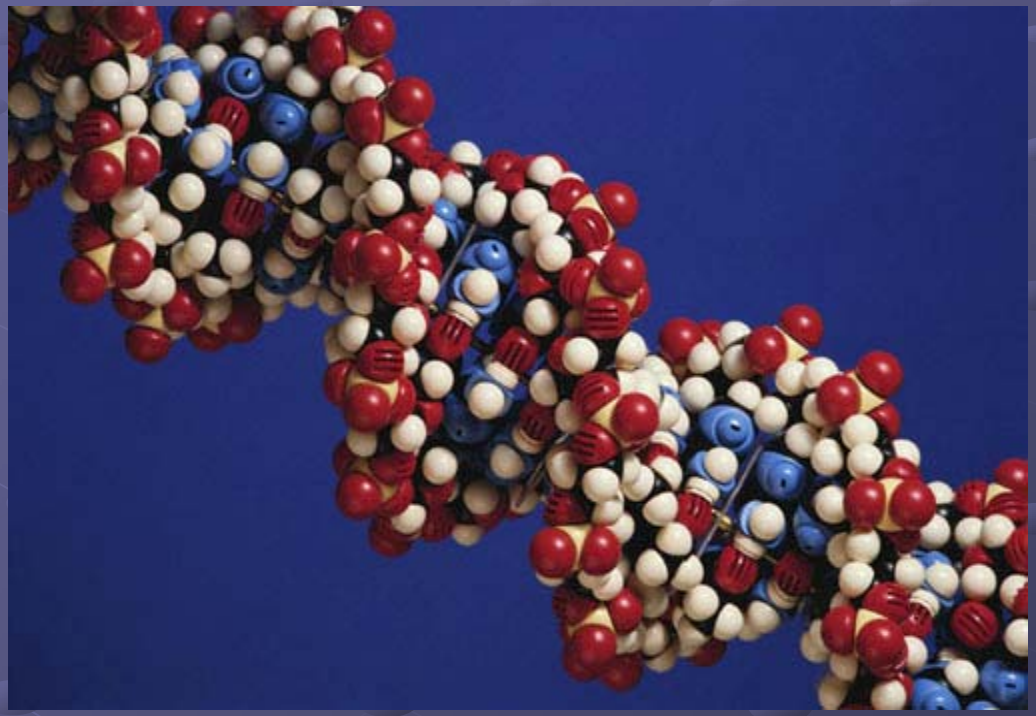
Basic Principles of Forensic Molecular Biology and Genetics

Nucleic Acids: Structure and Reactivity



DNA Structure and Reactivity

"We are still far from the time when people will understand the curious relationship between one fragment of nature and another, which all the same explain each other and enhance each other." --Vincent van Gogh

DNA Structure and Reactivity Cont.

- **Central Dogma of Molecular Biology**

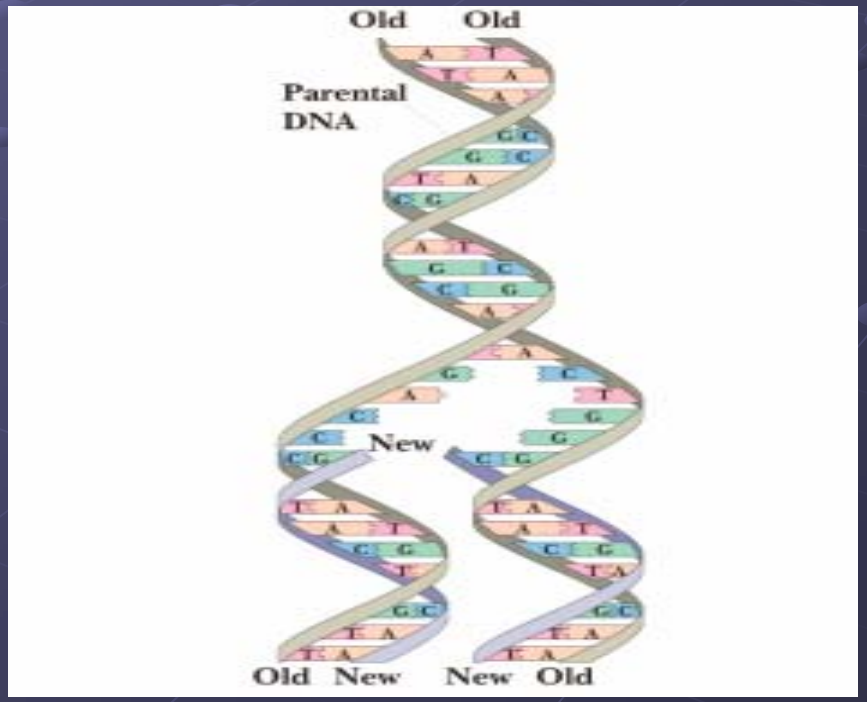
DNA → RNA → protein

However:	DNA dependent DNA polymerase	(DNA →DNA)
	RNA dependent RNA polymerase	(RNA → RNA)
	Reverse transcriptase	(RNA →cDNA)

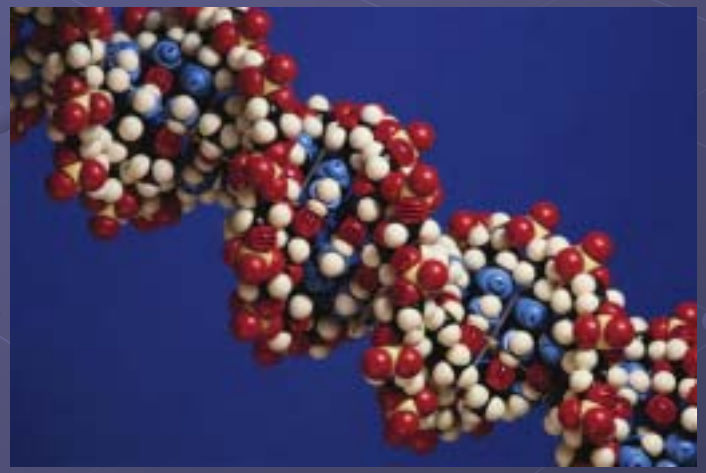
DNA Structure and Reactivity Cont.

- **Structure-Function**
  - “It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.”

    Watson and Crick 1953

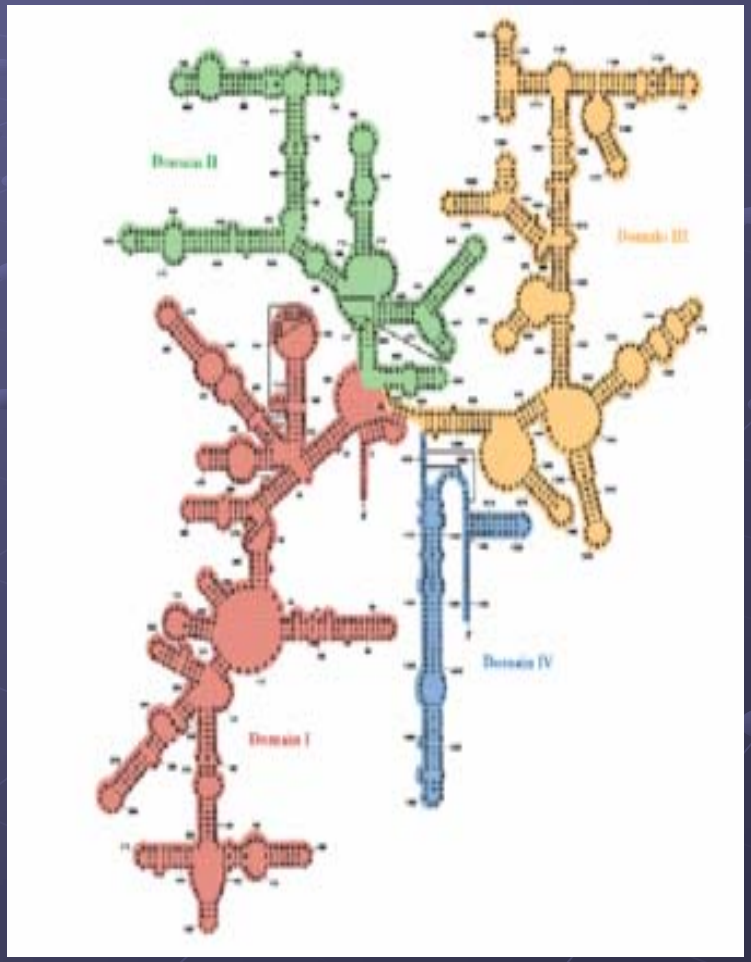


- **B-Form DNA**



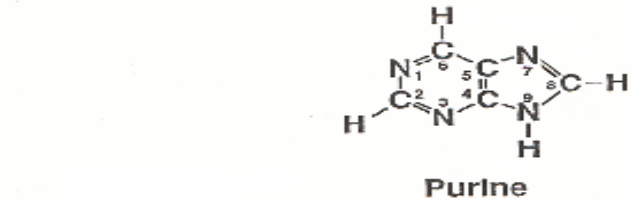
DNA Structure and Reactivity Cont.

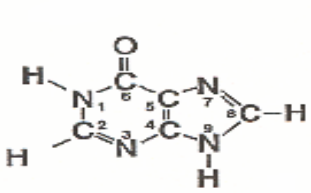
- **RNA Structure/Function**
  - Structure and function relationship not as obvious as DNA

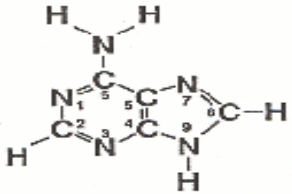


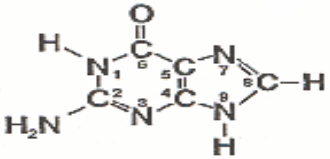
DNA Structure and Reactivity Cont.

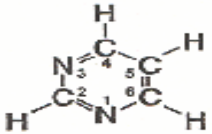
- **DNA Components**
  - Bases
    - Purines: A, G, H
    - Pyrimidines: T, C, U
    - 5-Methyl C (at CpG and CpXpG)

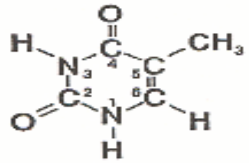

Purine

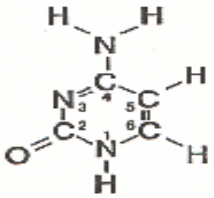

Hypoxanthine (H)

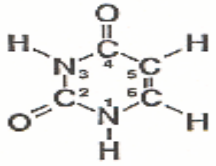

Adenine (A)

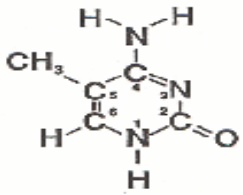

Guanine (G)


Pyrimidine


Thymine (T)

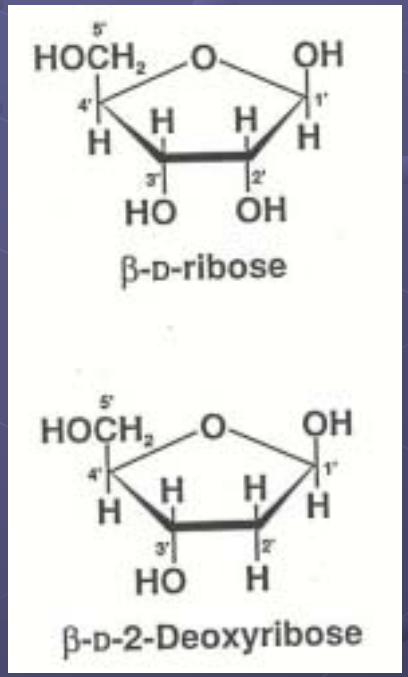

Cytosine (C)

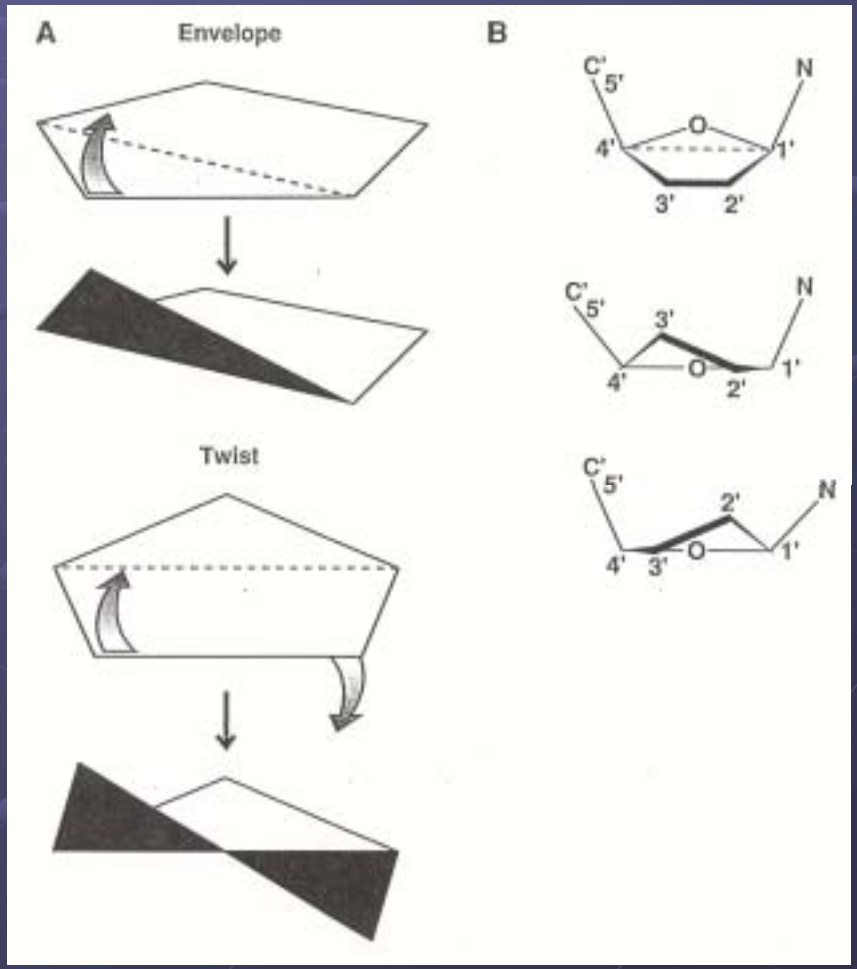

Uracil (U)


5-Methylcytosine

DNA Structure and Reactivity Cont.

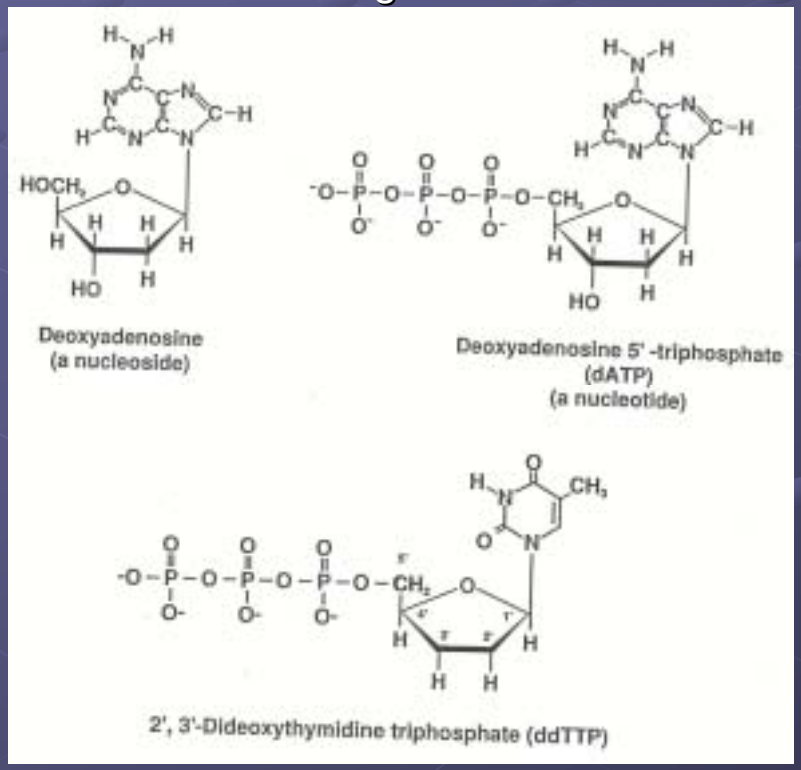
- **DNA Components Cont.**
  - Sugars
    - β-D-ribose
    - β-D-2 deoxyribose
    - C2’ endo versus C3’ endo form
    - Envelope v. twist forms

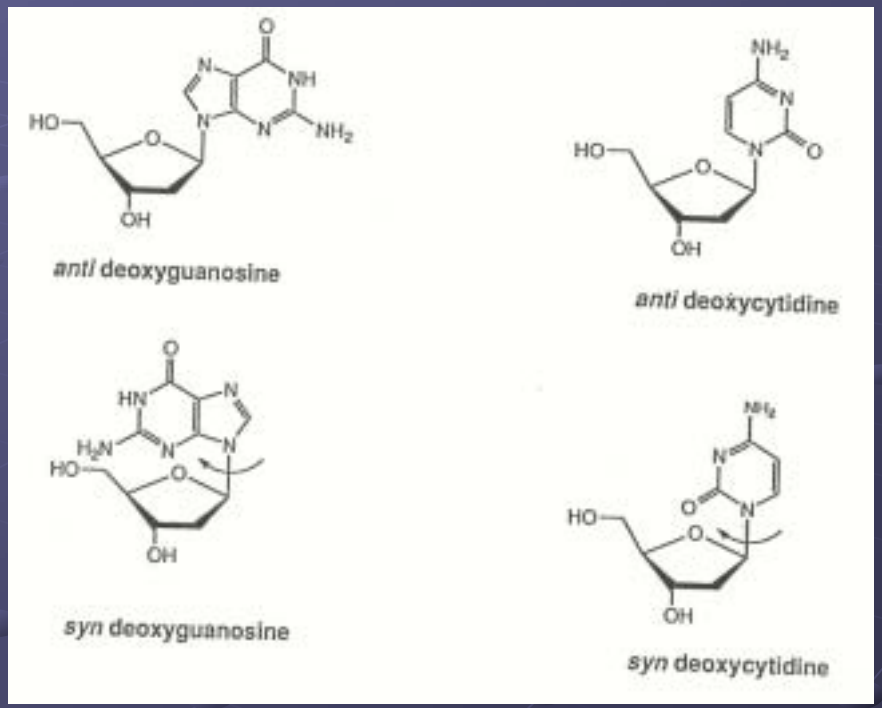




DNA Structure and Reactivity Cont.

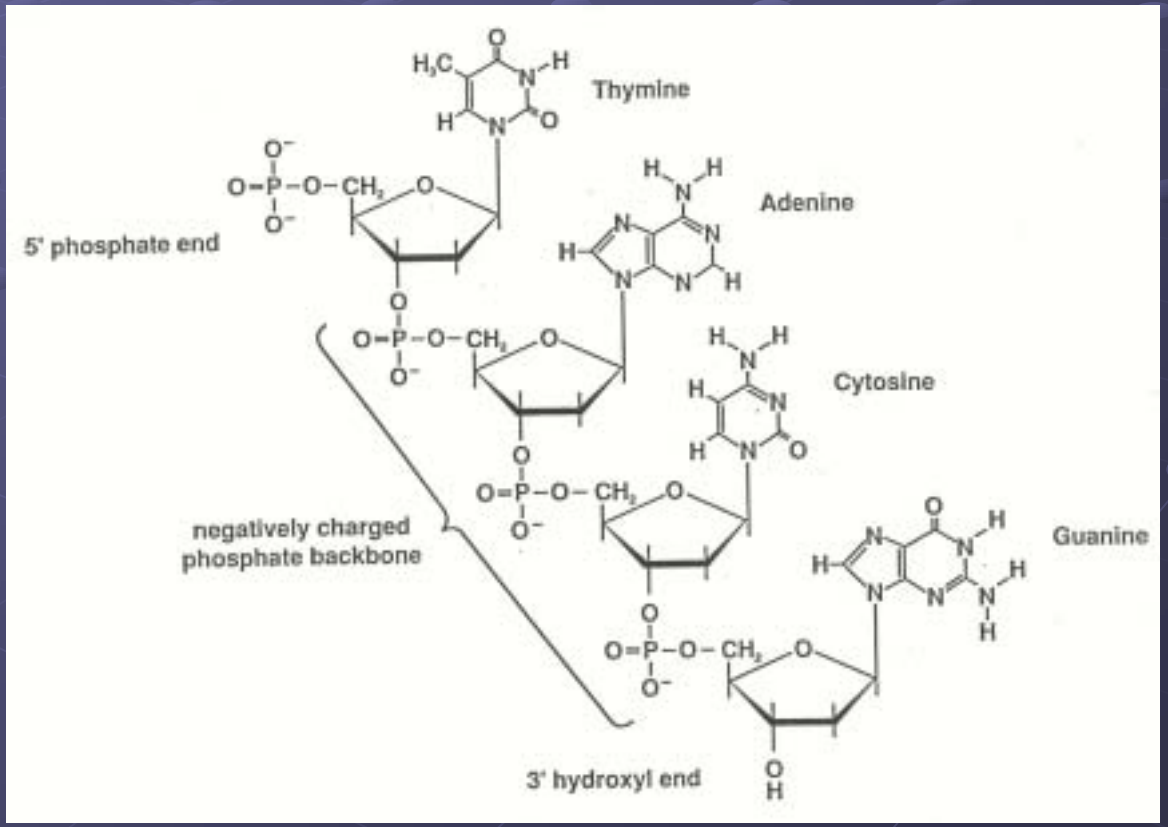
- **DNA Components Cont.**
  - Nucleosides/nucleotides
    - dA, dATP, ddTTP
    - Anti- v. Syn- N glycosidic bond angles





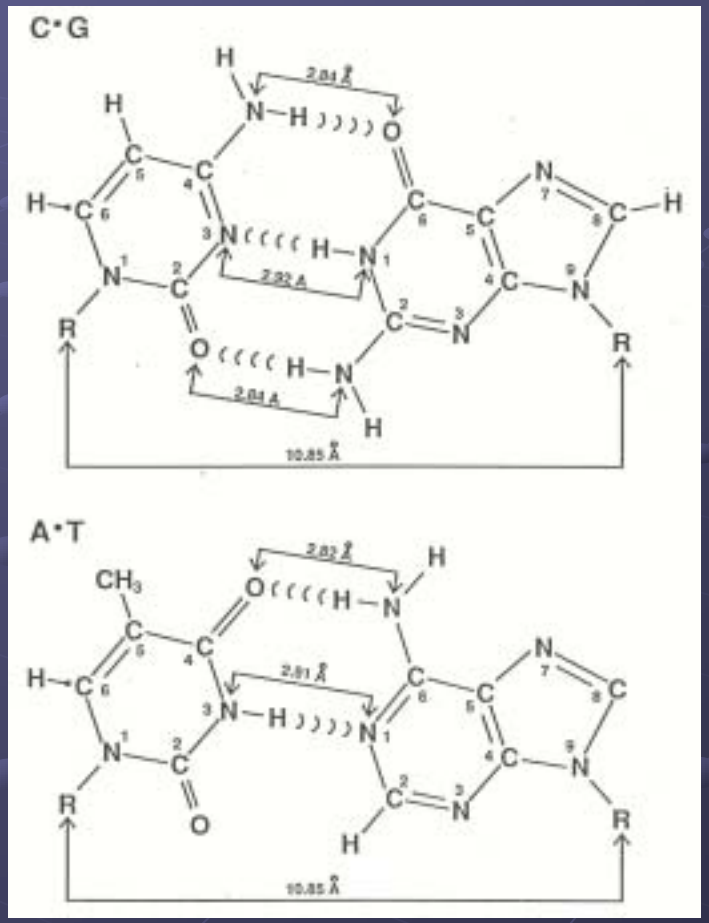
DNA Structure and Reactivity Cont.

- DNA Components Cont.
  - Single polynucleotide chain
    - 5’ and 3’ ends



DNA Structure and Reactivity Cont.

- Base Pair Interactions
  - Watson -Crick (W-C) base pairs
    - C·G, A·T
    - H-bonds (~2-3 kcal/mole) (normally ~3-7 kcal/mole; c.f. covalent bonds, ~80-100 kcal/mole)
  - Stacking energies
  - Orientations in 3D space

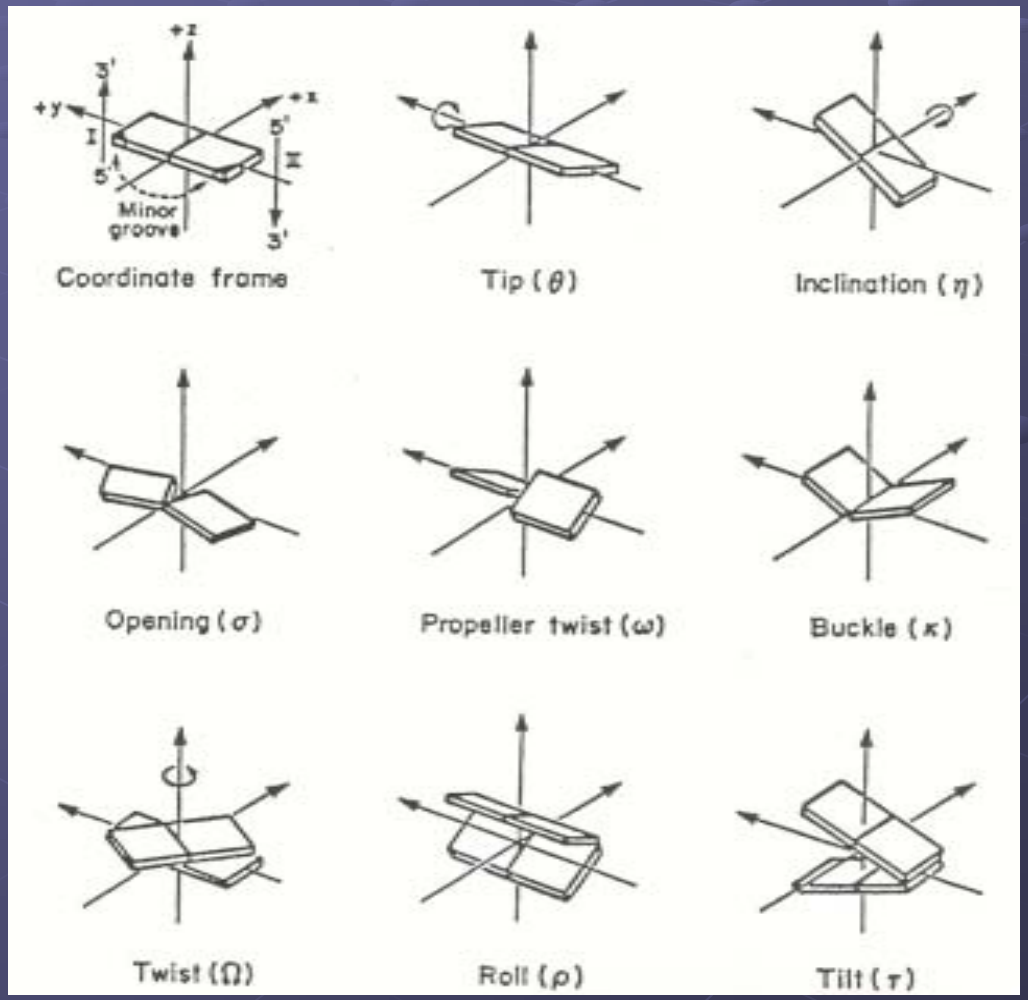


Base Pair Stacking Energies

Dinucleotide base pairs	Stacking energies (kcal/mol/stacked pair)[a]
(GC) - (GC)	−14.59
(AC) - (GT)	−10.51
(TC) - (GA)	− 9.81
(CG) - (CG)	− 9.69
(GG) - (CC)	− 8.26
(AT) - (AT)	− 6.57
(TG) - (CA)	− 6.57
(AG) - (CT)	− 6.78
(AA) - (TT)	− 5.37
(TA) - (TA)	− 3.82

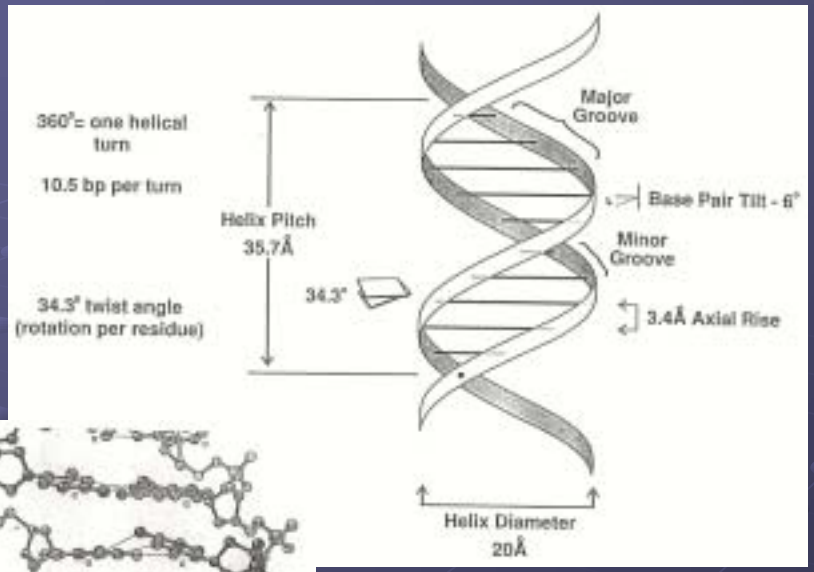
DNA Structure and Reactivity Cont.

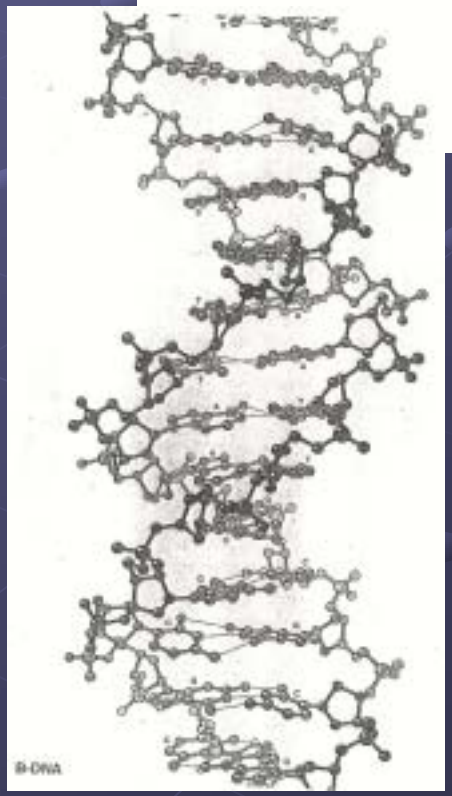
- Base Pair Interactions Cont.



DNA Structure and Reactivity Cont.

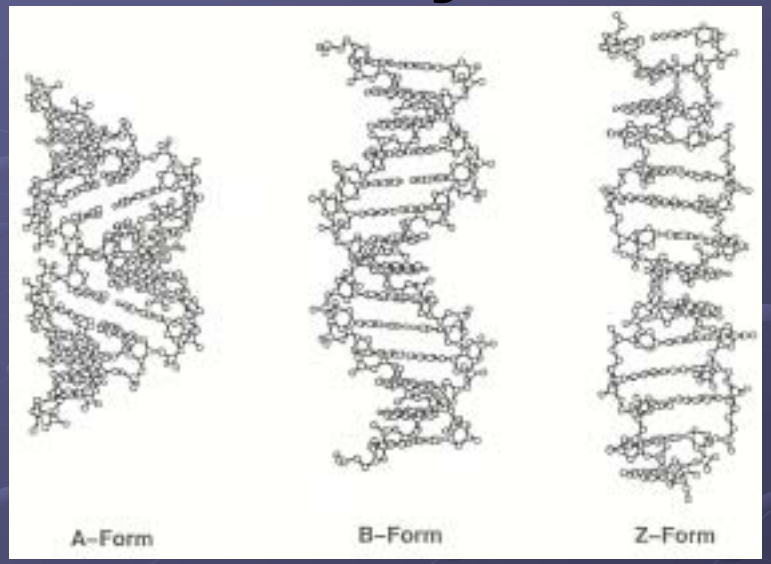
- **B Form DNA: Structural Parameters**
  - **Helix pitch (35.7 Å)**
  - **Helix diameter (20 Å)**
  - **Helical turn (= 360º = 10.5 bp per turn)**
  - **Twist angle (rotation per residue) (34.4º)**
  - **Base pair tilt (minus 6º)**
  - **Axial rise (3.4 Å)**
- **B Form DNA: Double Helix**
  - Molecular model





DNA Structure and Reactivity Cont.

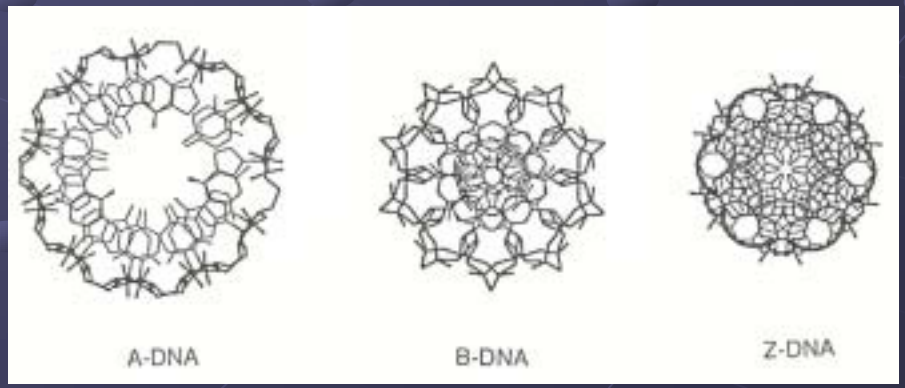
- Non -B Form DNA: A and Z Helices
  - Table
  - Molecular models



Helix Parameters

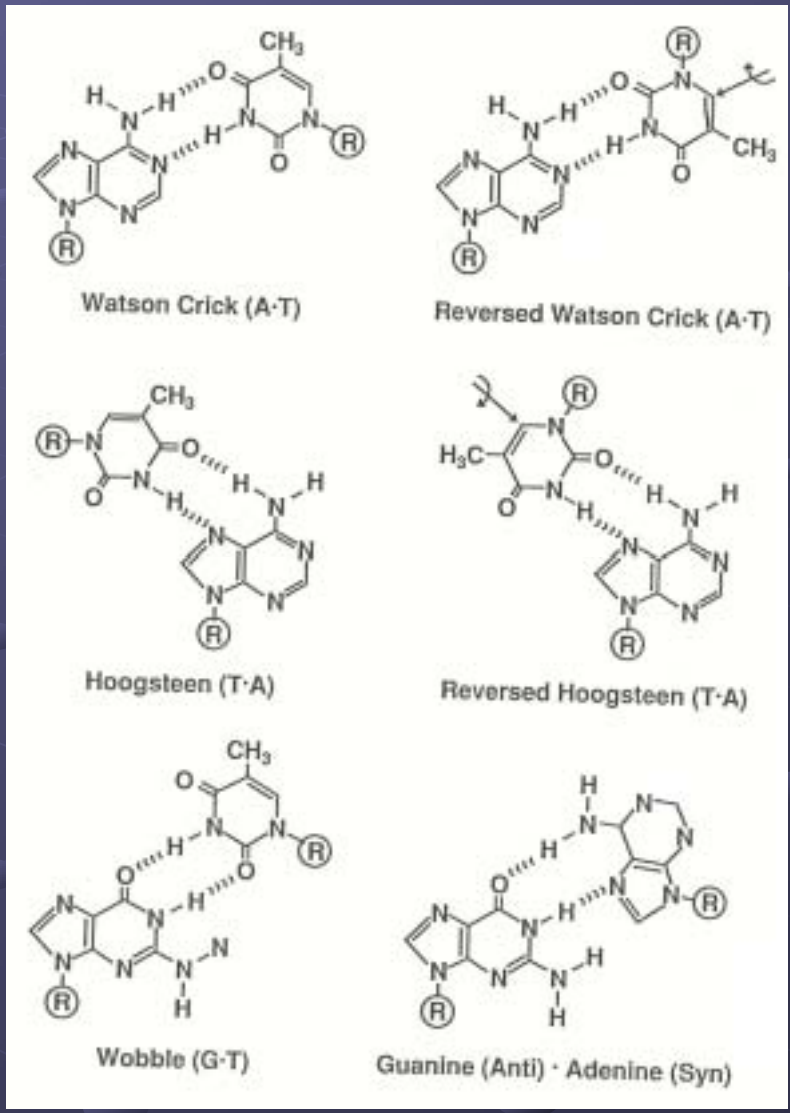
Parameter	A-DNA	B-DNA	Z-DNA
Helix sense	Right	Right	Left
Residue per turn	11	10 (10.5)[a]	12
Axial rise (Å)	2.55	3.4	3.7
Helix pitch (°)	28	34	45
Base pair tilt (°)	20	−6	7
Rotation per residue (°)	33	36 (34.3)[a]	−30
Diameter of helix (Å)	23	20	18
Glycosidic bond configuration			
dA, dT, dC	anti	anti	anti
dG	anti	anti	syn
Sugar pucker			
dA, dT, dC	C3′ endo	C2′ endo	C2′ endo
dG	C3′ endo	C2′ endo	C3′ endo
Intrastrand phosphate–phosphate distance (Å)			
dA, dT, dC	5.9	7.0	7.0
dG	5.9	7.0	5.9

[a]Values in parentheses are the residues per turn and rotation per residue for B-form DNA as it exists in solution of physiological ionic strength. Other values are taken from X-ray diffraction data.



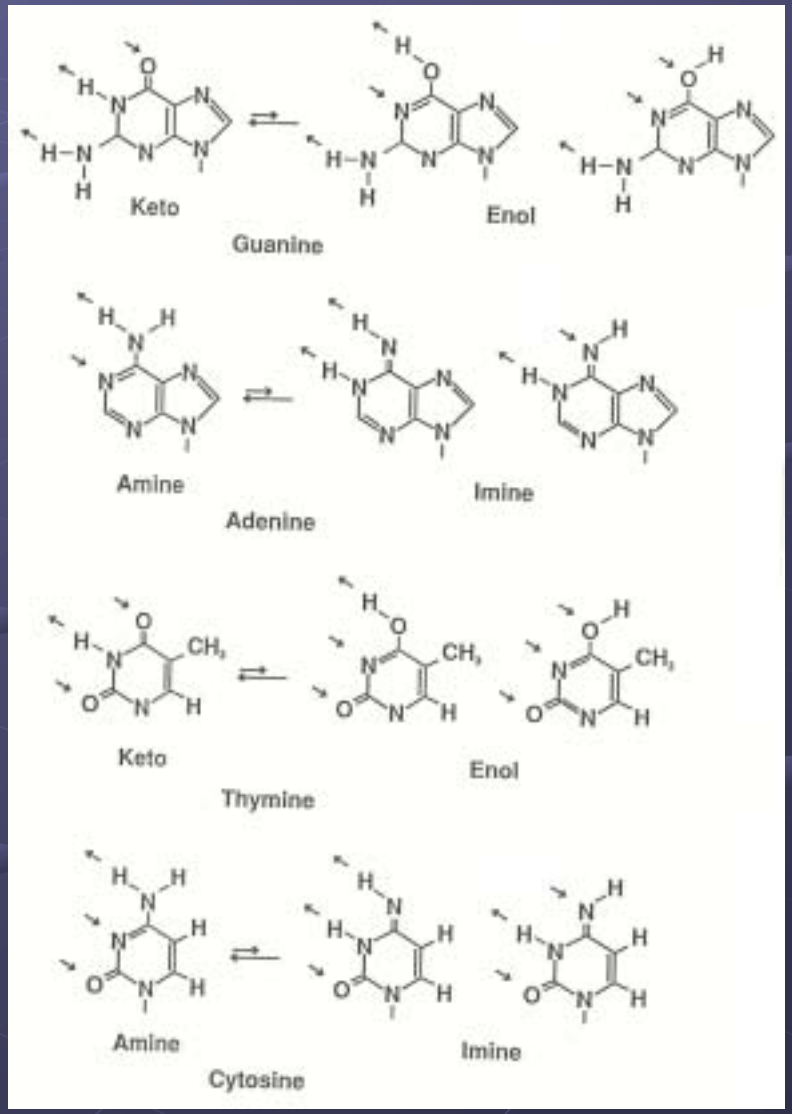
DNA Structure and Reactivity Cont.

- **Non-canonical Base Pairing Schemes**
  - Reversed W-C
  - Hoogsteen, Reversed Hoogsteen [8]
  - Wobble, Guanine (*anti*)-Adenine (*syn*) [8]



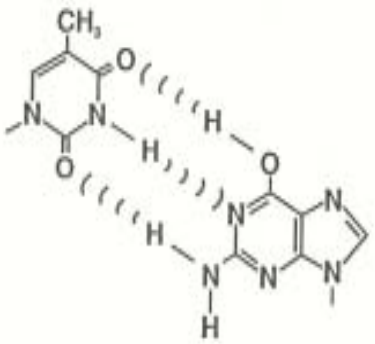
DNA Structure and Reactivity Cont.

- **Non-canonical Base Pairing Schemes Cont.**
  - **Tautomerization**
    - keto-enol, amine-imine

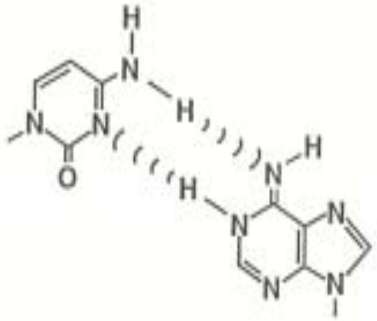


DNA Structure and Reactivity Cont.

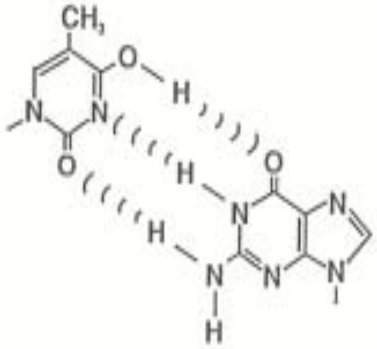
base pairing of tautomers



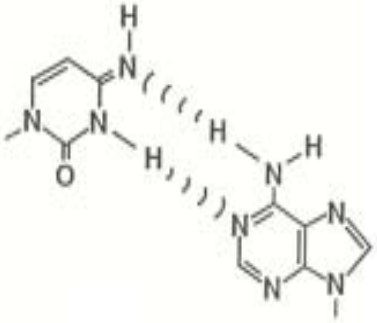
Thymine•Guanine (enol)



Cytosine•Adenine (imino)

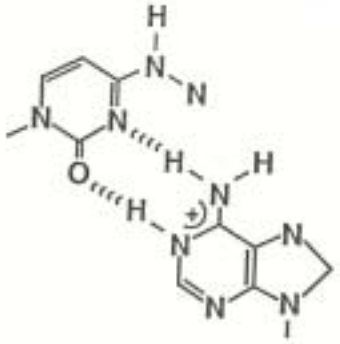


Thymine (enol)•Guanine

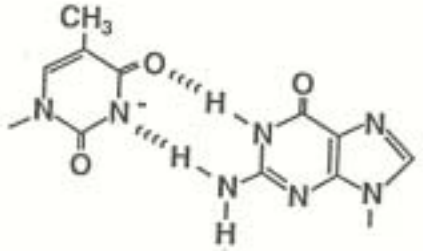


Cytosine (imino)•Adenine

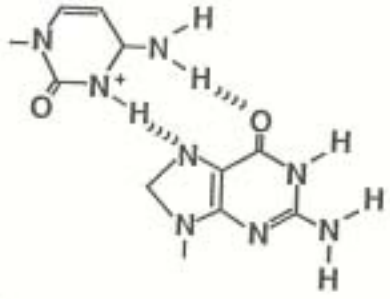
base pairing of the ionized forms



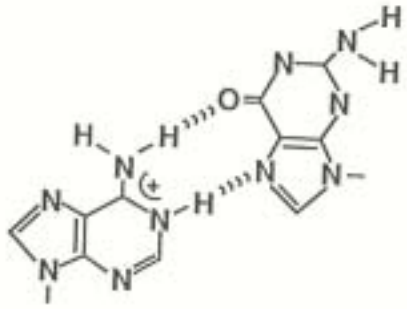
Cytosine•Adenine (ionized)
Wobble



Thymine (ionized)•Guanine



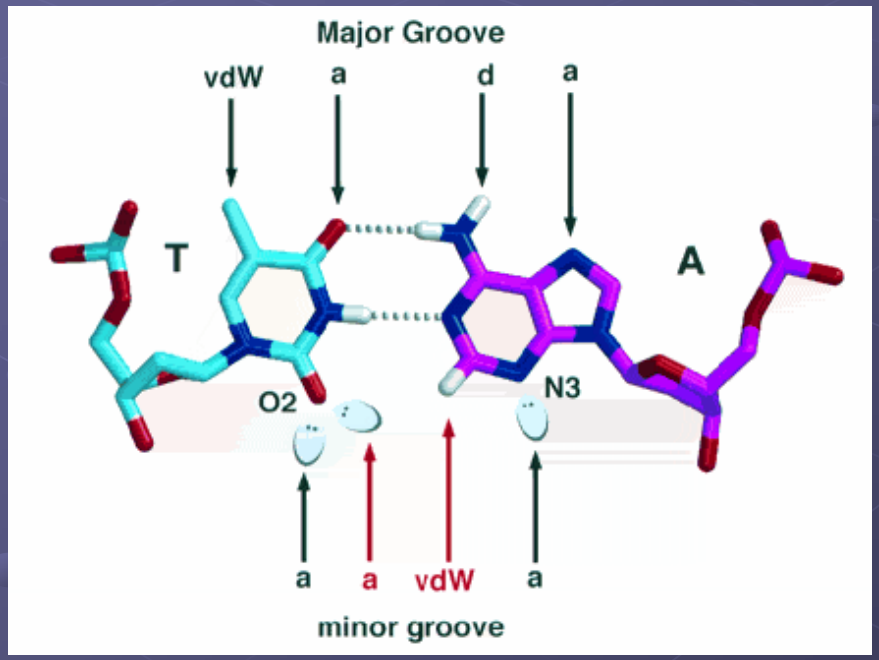
Cytosine (ionized)•Guanine
Hoogsteen

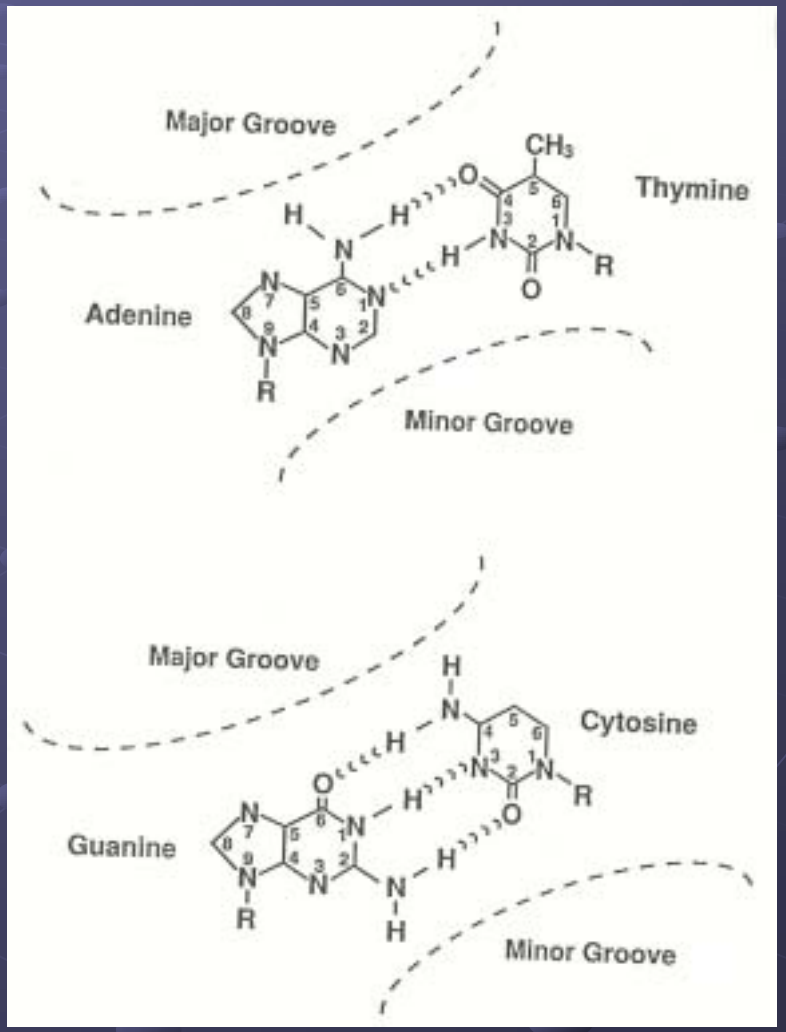


Adenine (ionized)•Guanine (syn)

DNA Structure and Reactivity Cont.

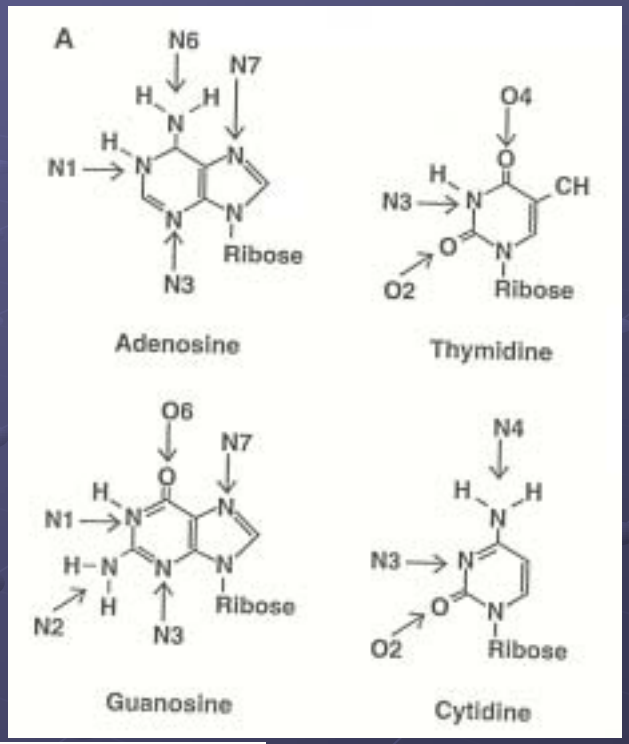
- Accessibility of Functional Groups in DNA
  - Major and minor groove accessibility



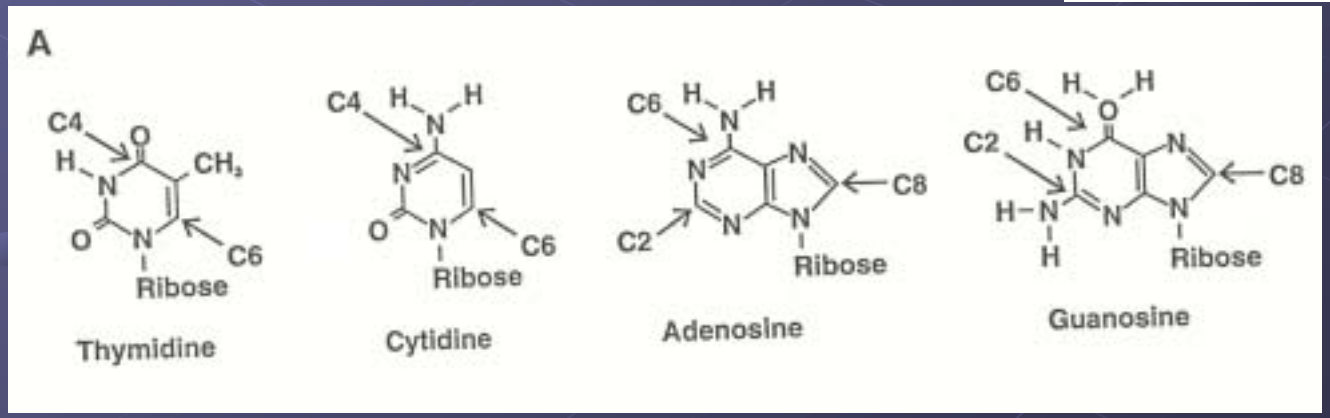


DNA Structure and Reactivity Cont.

sites of electrophilic attack

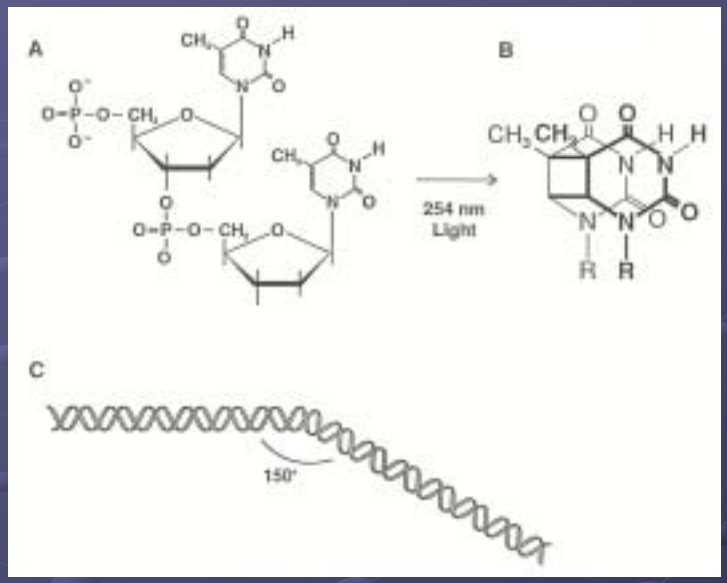


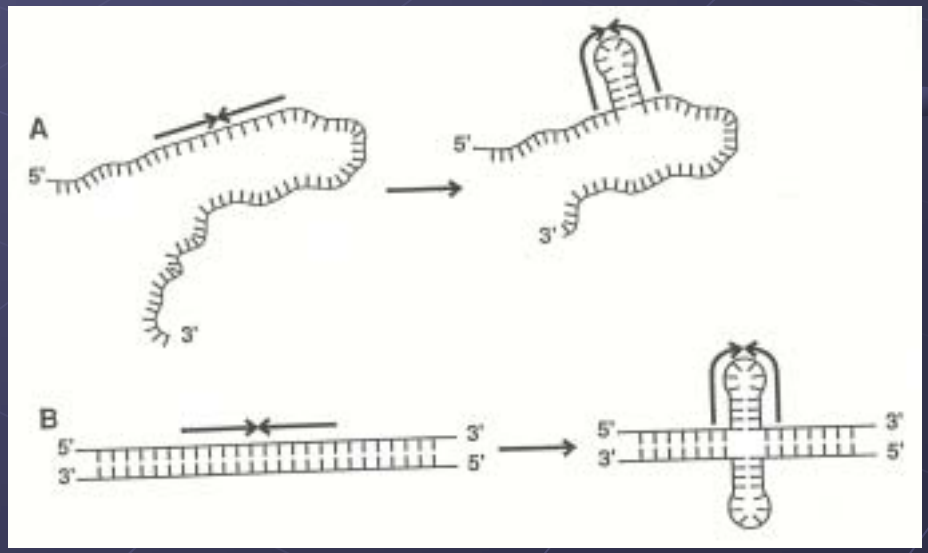
sites of nucleophilic attack



DNA Structure and Reactivity Cont.

- Non-Helical Structures
  - Bent DNA
    - pyrimidine dimers
    - multiple tracts of A phased by 10 bp
  - Cruciform structures
  - Supercoiling (L= T + W)
  - Triplex DNA
  - Four stranded DNA





DNA Structure and Reactivity Cont.

- Methylation of DNA

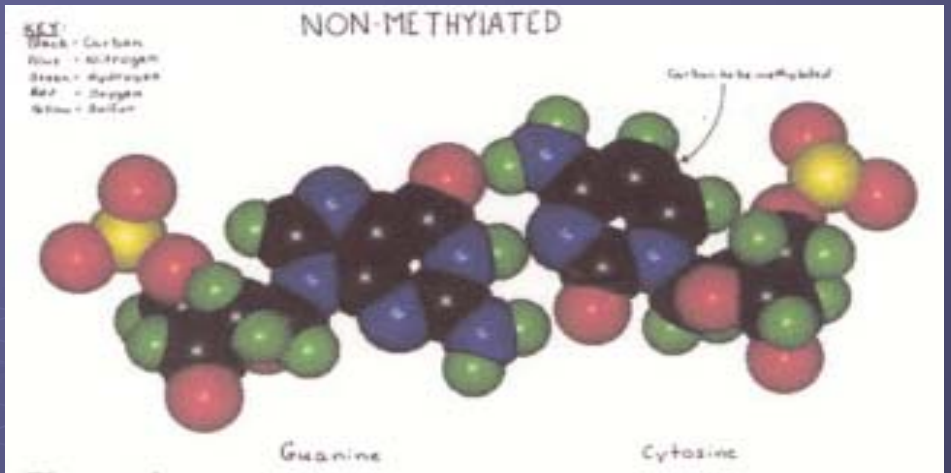


Figure 1

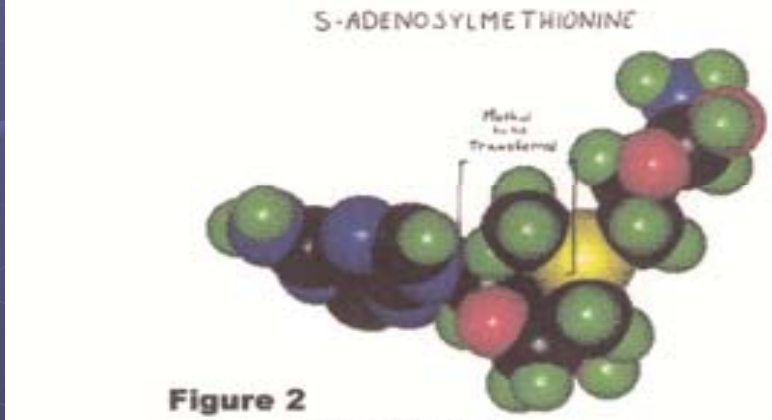


Figure 2

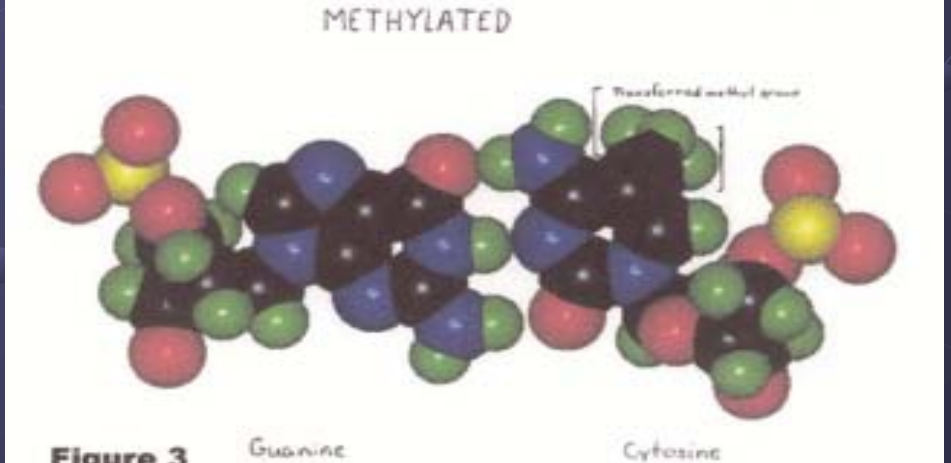


Figure 3

DNA Structure and Reactivity Cont.

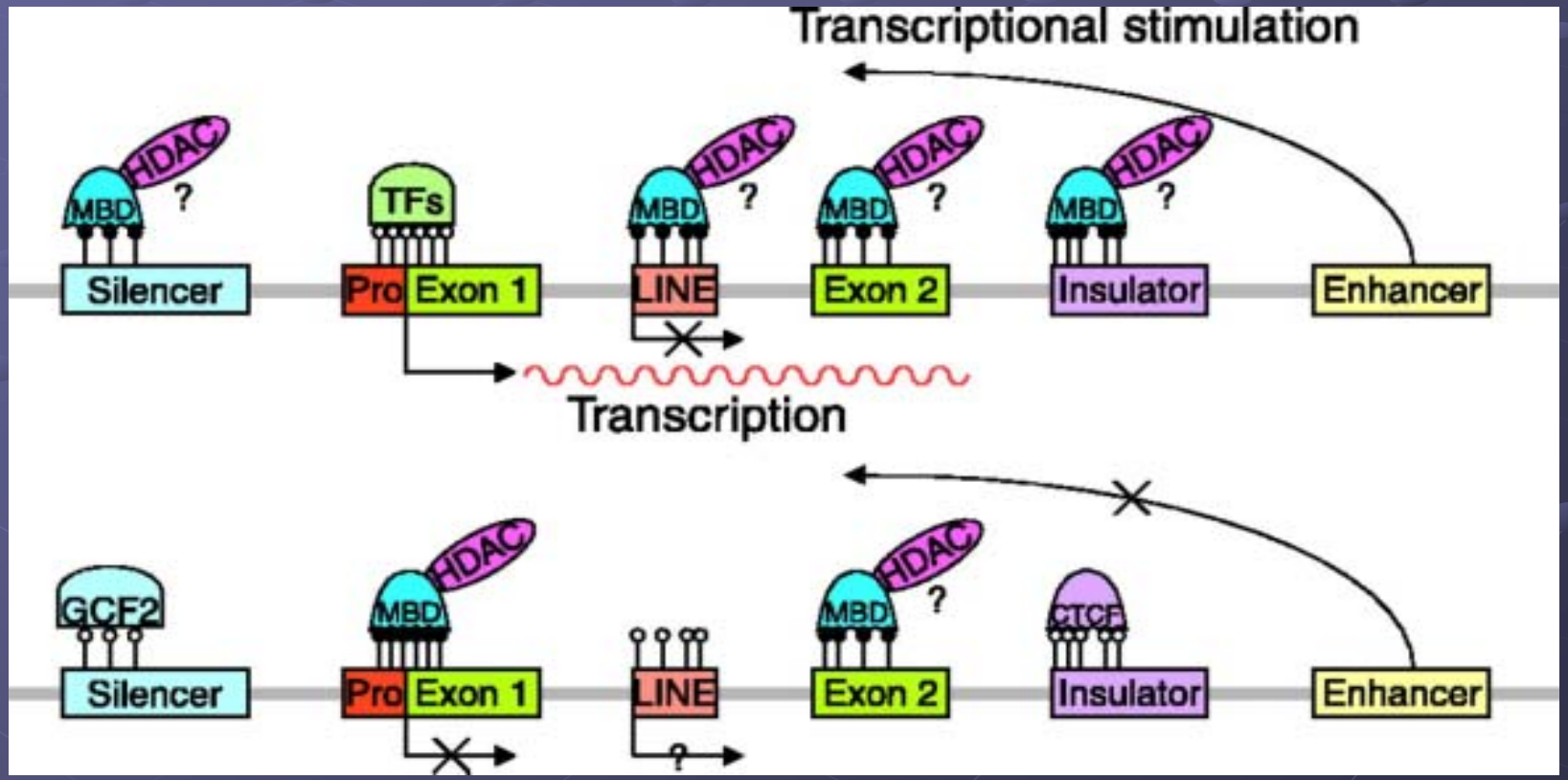
- Methylation of DNA
  - Example of epigenetics
    - any gene regulatory activity that doesn't involve changes to the DNA sequence and that can persist through one or more generations
  - Essential for development of mammals (why is unknown)
  - 3 DNA cytosine methyltransferases cloned: DNMT1, -3a, -3b
  - Alters major groove
  - CpG promoter methylation
    - prevents transcription initiation
    - silencing of genes
      - on inactive X
      - parasitic DNA
      - can 'spread' (methylation attracts more methylation)
      - imprinted genes

DNA Structure and Reactivity Cont.

- **Methylation of DNA Cont.**
  - CpG islands
    - CpG dinucleotides are underrepresented in mammalian DNA (likely due to mutagenic potential of C deamination)
    - Despite above, are found in genes (promoters, but also first exons and 3' end)
    - >200 bp, high GC content, observed/expected ratio >0.6 (needs revising)
  - Why are there varying methylation patterns in eukaryotes? (c.f prokaryotes where all sites are methylated in the presence of the appropriate methytransferase)?
    - Exclusion of access to methylation sites by DNA bound proteins
      - Removal of SP1 binding sites flanking CpG island ----> de novo methylation during development
    - Methylation targeting mechanism steered by sequence specific binding proteins
      - DNMT1, DNMT3a associate with Rb, E2F1, histone deacetylase (HDACs), RP58 (transcriptional repressor)

DNA Structure and Reactivity Cont.

- **Methylation of DNA Cont.**
  - Mechanisms of Transcriptional Silencing
    - Inhibits initiation, not elongation in mammals



DNA Structure and Reactivity Cont.

- **Methylation of DNA Cont.**
  - Methylation and Tissue Specific Gene Expression (TSGE)
    - DNA modifications are tissue specific
    - Link between methylation and TSGE not clear and in some doubt
  - Methylation and Human Disease
    - C methylation is major contributor to disease- causing germline mutations and somatic mutations causing cancer
    - Abnormal methylation of promoters of regulatory genes causes silencing (e.g. p16 tumor suppressor gene)

DNA Structure and Reactivity Cont.

- **Methylation of DNA Cont.**
  - Methylation and Human Disease Cont.
    - Methylation is also important for development after birth
      - ICF (immunodeficiency, centromeric instability, facial anomalies syndrome)
        - Mutation in DNMT3b ---> undermethylation of satellite DNA and specific chromosome decondensation
      - Rett Syndrome
        - Common kind of mental retardation in young girls
        - Mutation in MBD (methylated CpG binding protein)
        - May not be able to interpret methylation signal
  - Major Function of Methylation?
    - Long term silencing of noncoding DNA in the genome, which contains a substantial proportion of repetitive elements

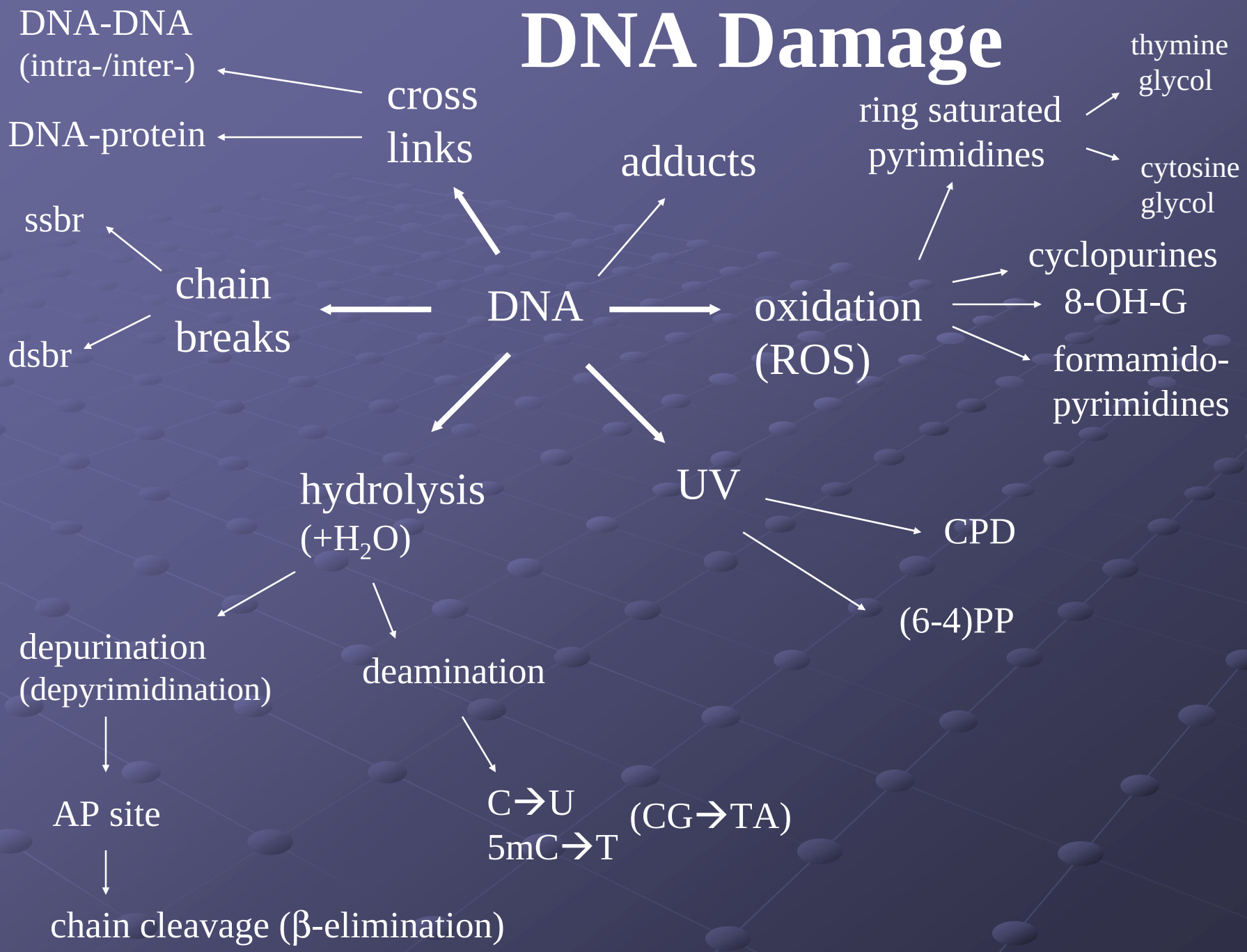
DNA Damage and Repair

Primary structure of DNA is dynamic and subject to constant change such as:

transposition, changes in chemistry or sequence of nucleotides due to:

- spontaneous damage
  - replication, recombination, repair
  - inherent instability of specific chemical bonds
  - physical or chemical agents from the environment
- environmental damage
  - physical agents
  - chemical agents

DNA Damage


DNA-DNA
(intra-/inter-)
DNA-protein
cross
links
adducts
ring saturated
pyrimidines
thymine
glycol
cytosine
glycol
ssbr
dsbr
chain
breaks
DNA
oxidation
(ROS)
cyclopurines
8-OH-G
formamido-
pyrimidines
hydrolysis
(+H2O)
UV
CPD
(6-4)PP
depurination
(depyrimidination)
deamination
AP site
C→U
5mC→T
(CG→TA)
chain cleavage (β-elimination)

DNA Damage and Repair

- **Damage**

  Damage refers to the situation whereby the primary structure is covalently altered other than as result of epigenetic gene regulation. Damage can be spontaneous or environmental in nature.

- **Mismatches**

  Difference in free energy between complementary and non-complementary bases is only ~2-3 kcal/mol (a single hydrogen bond) -→ translates to potential replication error frequency of 1-10 % per base. However, error frequency is 6-9 orders of magnitude less!

	cumulative error frequency
Base pairing	~ 10^{-1}-10^{-2}
DNA polymerase actions (base selection/proofreading)	~ 10^{-5}-10^{-6}
Accessory proteins (e.g. SSB protein)	~ 10^{-7}
Post-replicative mismatch correction	~ 10^{-10}

DNA Damage and Repair Cont.

- **Tautomeric Shifts**
  - If a base in a template strand exists in its rare tautomeric form misincorporation in the daughter strand can result.

    Imino:

    C (imino) will pair with A

    A (imino) will pair with C

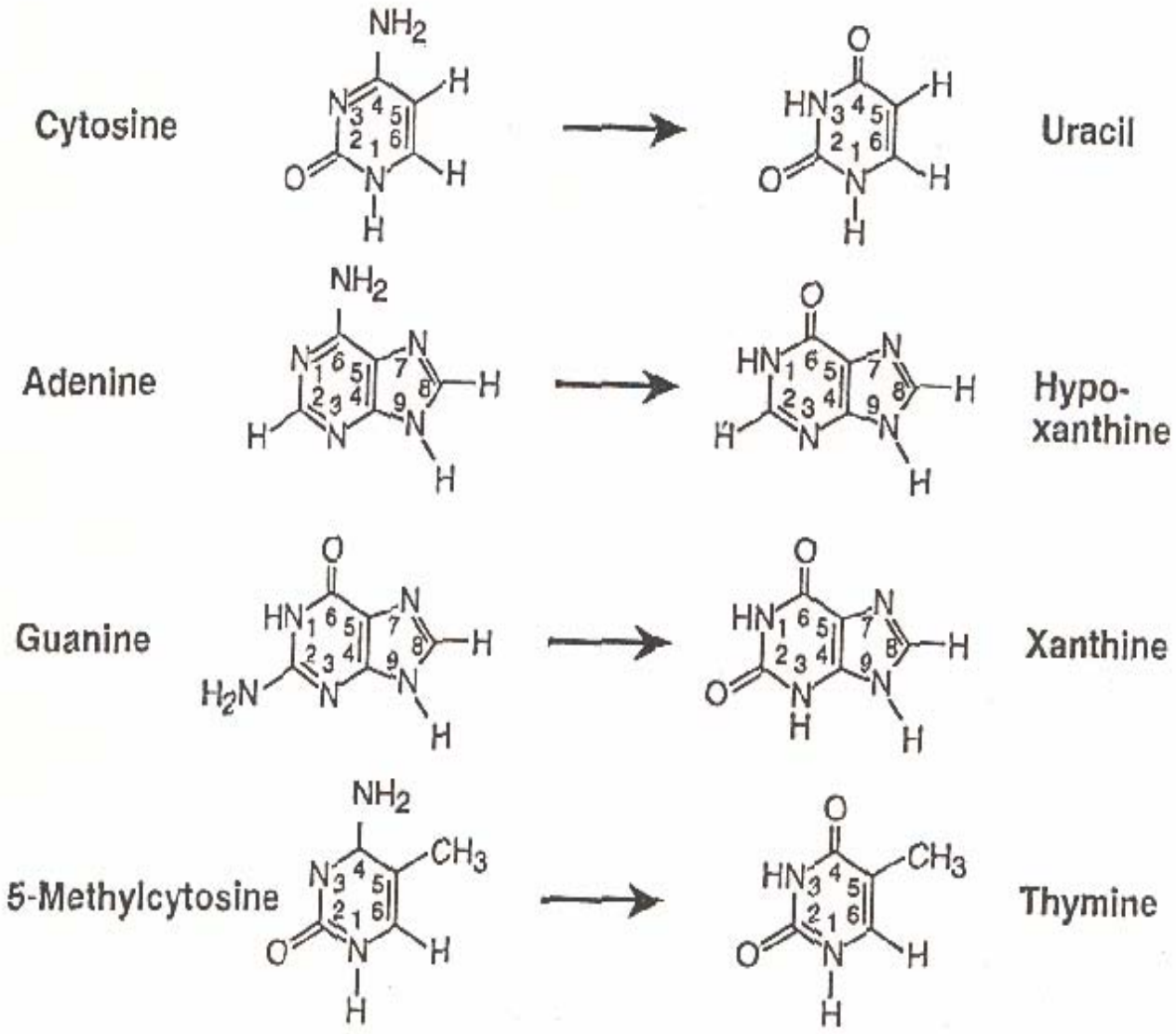
    T (enol) will pair with G

    G (enol) will pair with T

- **Deamination**

  Bases containing exocyclic amino groups can undergo deamination.

Deamination



Deamination (cont'd)

C →U

Mutagenic lesion: If not repaired, can result in a G.C →A.T transition.

It is speculated that the reason for the use of T (methylated U) in DNA instead of U (as in RNA) allows for the facile detection of the deamination product of C → Uracil excised rapidly in DNA by a uracil-DNA glycosylase.

Rate of C deamination in ssDNA > 100 fold increased over that in dsDNA.

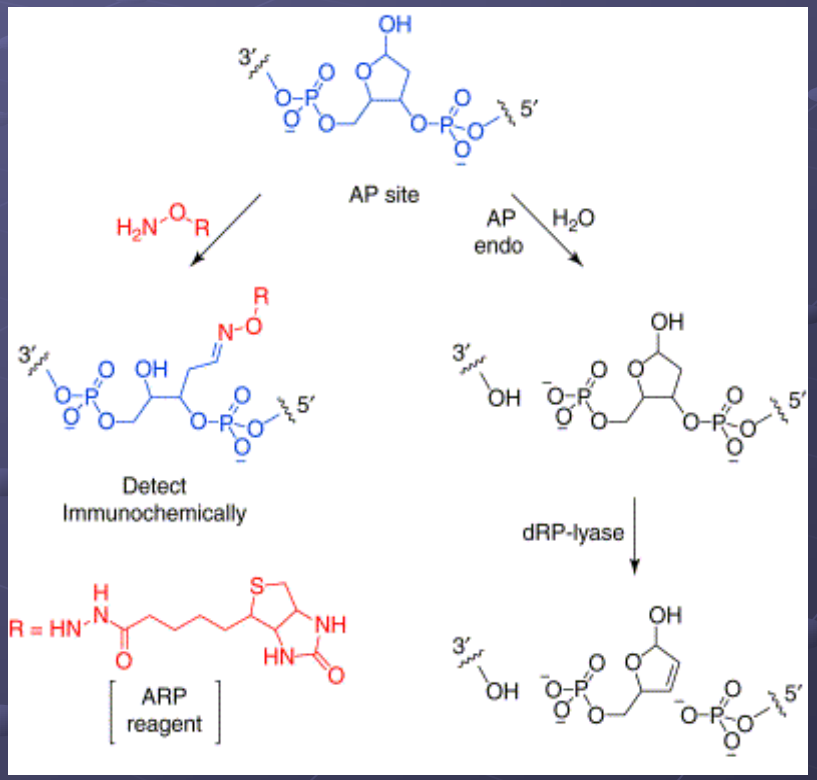
5-MeC →T

Highly mutagenic due to degree of inefficiency of MMR system: results in G.C →A.T transition.

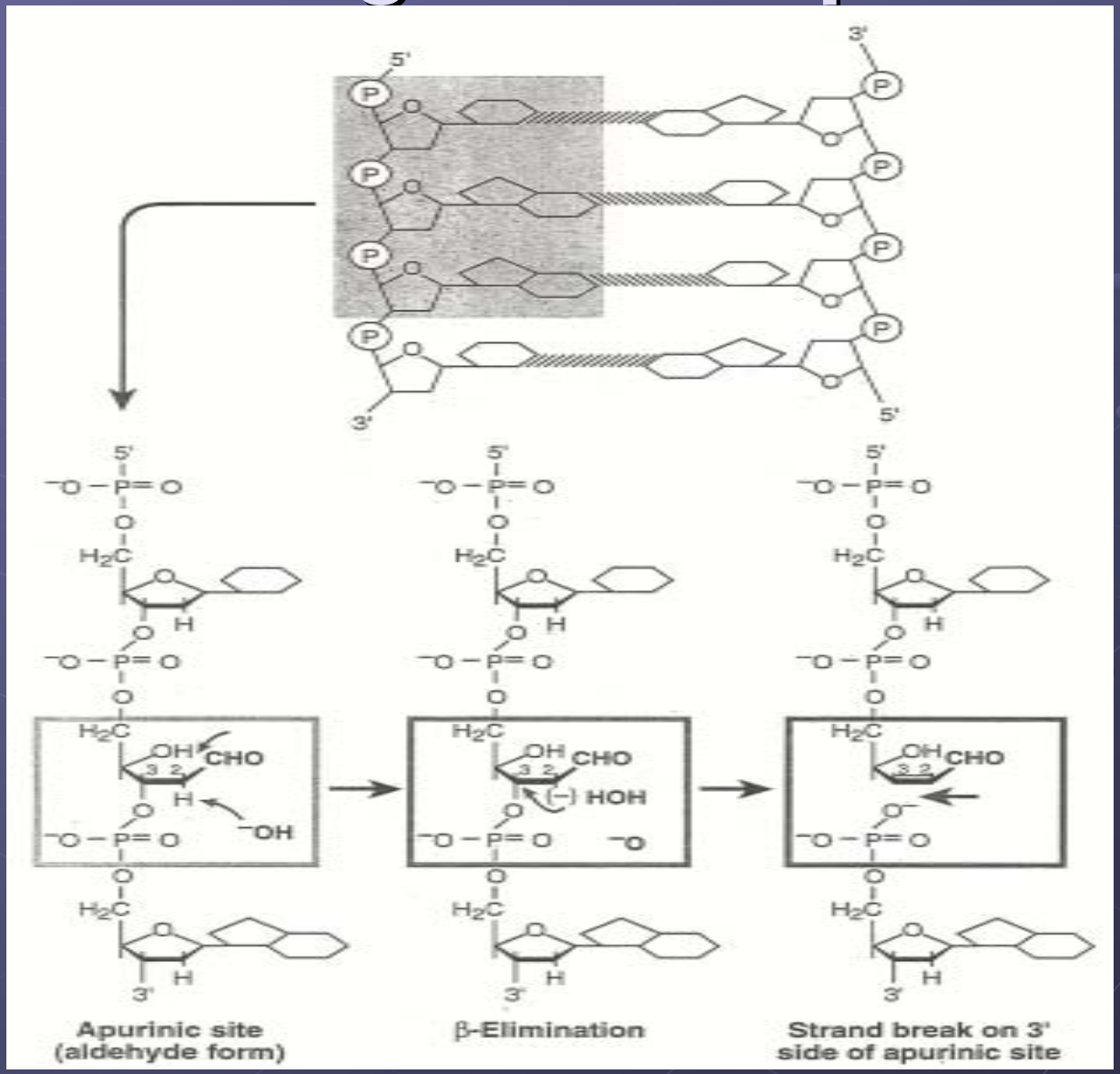
A, G deamination occurs at a much reduced rate compared to C (< 2%). Rate is 10-4 that of depurination (see below).

DNA Damage and Repair Cont.

- **Depurination /depyrimidination**
  Apurinic/apyrimidinic (AP) sites can exist at a level of 50,000-200,000 AP sites per genome in human tissues.
  - AP site deoxyribose exists in an equilibrium between the closed furanose and open aldehyde form, the 3'phosphodiester bonds of which are labile. These can be hydrolysed by a b-elimination reaction in which the pentose carbon b to the aldehyde is activated at alkaline pH and elevated temperature.
  - AP sites are repaired by 5' phosphodiester hydrolysis (by AP endonuclease) followed by 3' phosphate elimination (by dRp-lyase activity of DNA polymerase.



DNA Damage and Repair Cont.



DNA Damage and Repair Cont.

- **Oxidative Damage**

  Reactive oxygen species (ROS) can cause oxidative damage to DNA. ROS include hydroxyl radicals, hydrogen peroxide and singlet oxygen with the hydroxyl radical being the most important.

  $RH_2 + \cdot OH \rightarrow RH + H_2O$

  Superoxide radicals (produced as a side product of mitochondrial respiration) can be dismutated into hydrogen peroxide which is then converted to the hydroxyl radical through a Fenton type reaction.

DNA Damage and Repair Cont.

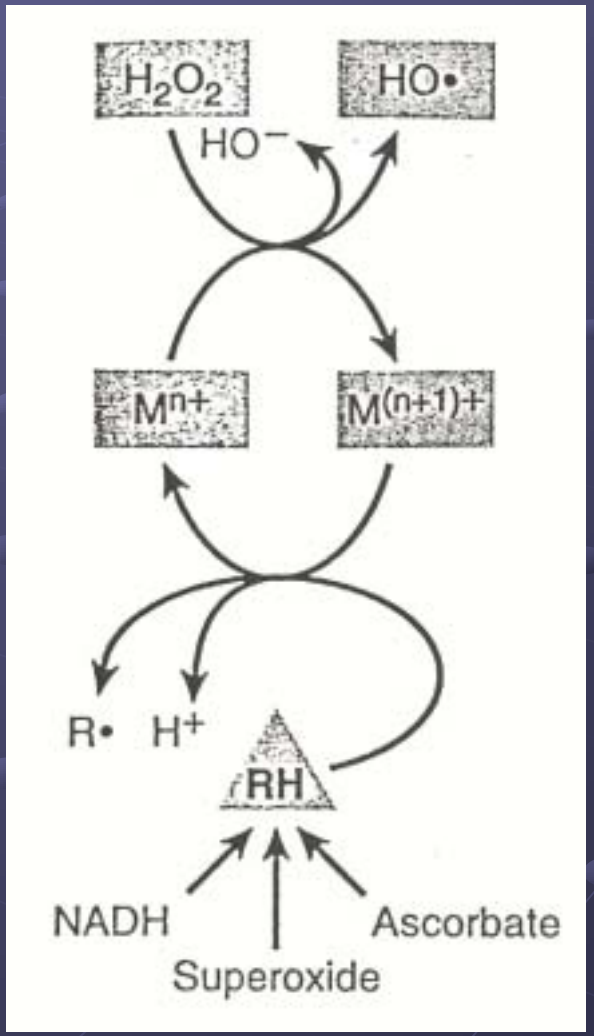
- **Oxidative Damage Cont.**

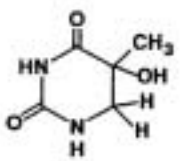
  OH. -Mediated oxidative damage can be of the following types:

  - **Oxidized bases**
  - **Abasic sites**
  - **DNA-DNA intra-strand adducts**
  - **DNA strand breaks**
  - **DNA-protein crosslinks**

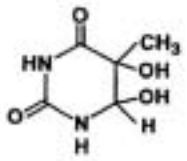
  We will concentrate on oxidized bases, particularly:

  - **Guanine →8-oxoguanosine**
  - **Thymine →thymidine glycol, hydantoins**

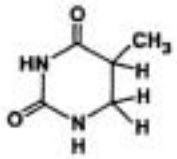




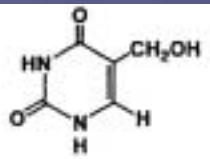
5-hydroxy-6-hydro-thymine



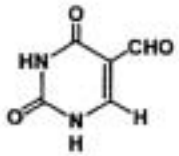
thymine glycol



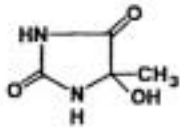
5,6-dihydrothymine



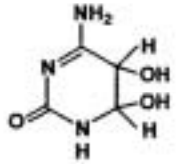
5-hydroxymethyl-uracil



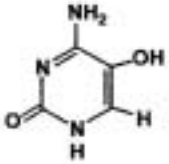
5-formyluracil



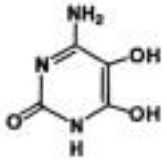
5-hydroxy-5-methyl-hydantoin



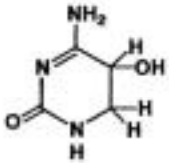
cytosine glycol



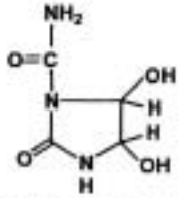
5-hydroxycytosine



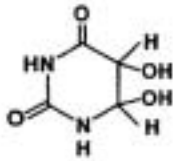
5,6-dihydroxycytosine



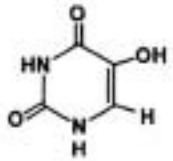
5-hydroxy-6-hydro-cytosine



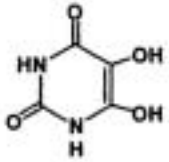
trans-1-carbamoyl-2-oxo-4,5-dihydroxyimidazolidine



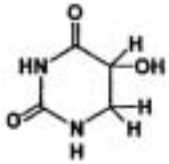
uracil glycol



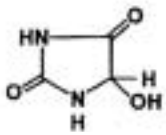
5-hydroxyuracil



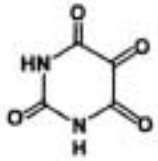
5,6-dihydroxyuracil



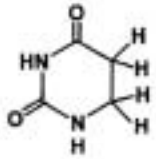
5-hydroxy-6-hydro-uracil



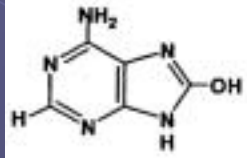
5-hydroxyhydantoin



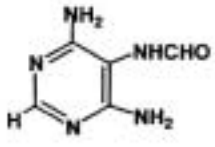
alloxan



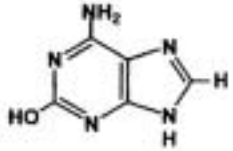
5,6-dihydrouracil



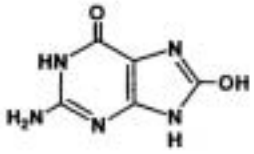
8-hydroxyadenine



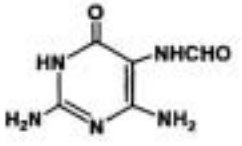
4,6-diamino-5-form-amidopyrimidine



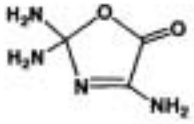
2-hydroxyadenine



8-hydroxyguanine



2,6-diamino-4-hydroxy-5-formamidopyrimidine



oxazolone

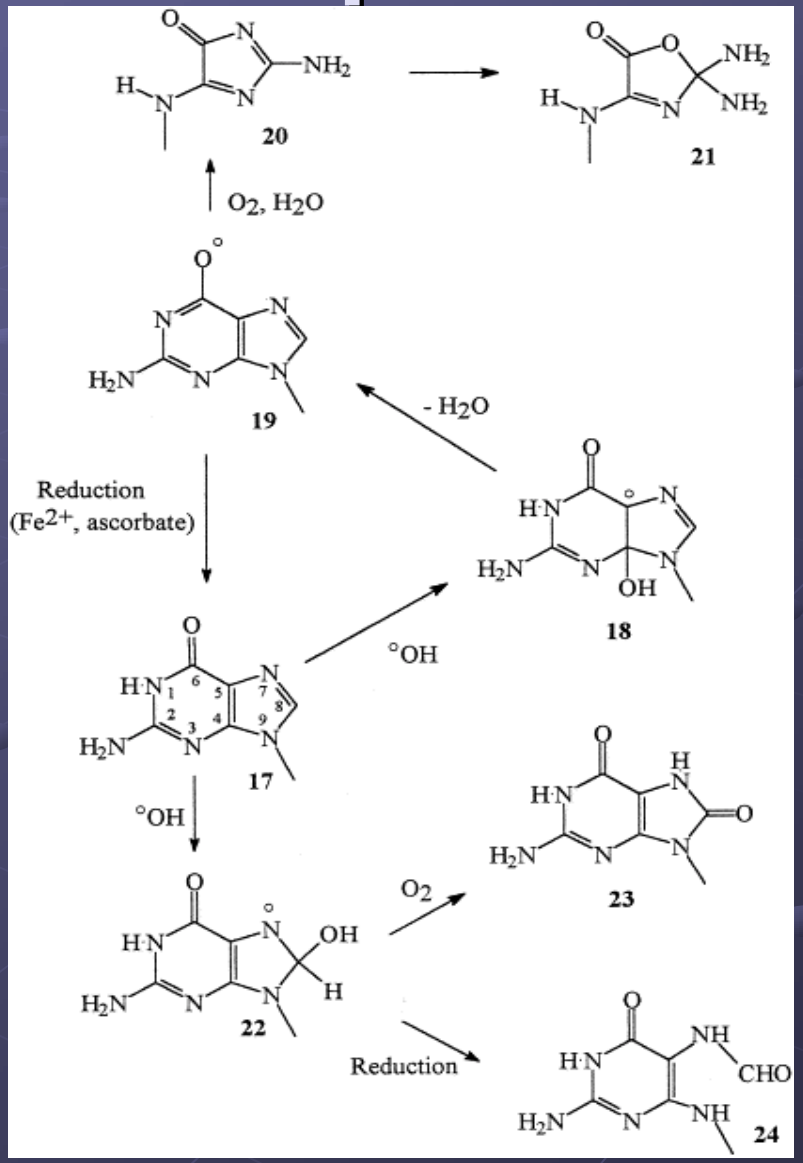
DNA Damage and Repair Cont.

- **Oxidative Damage Cont.**

Guanine →8-oxoguanosine [18]

Guanine is the most susceptible DNA target to OH. mediated oxidation.

8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodGuo) (**23**) is a mispairing lesion during replication.

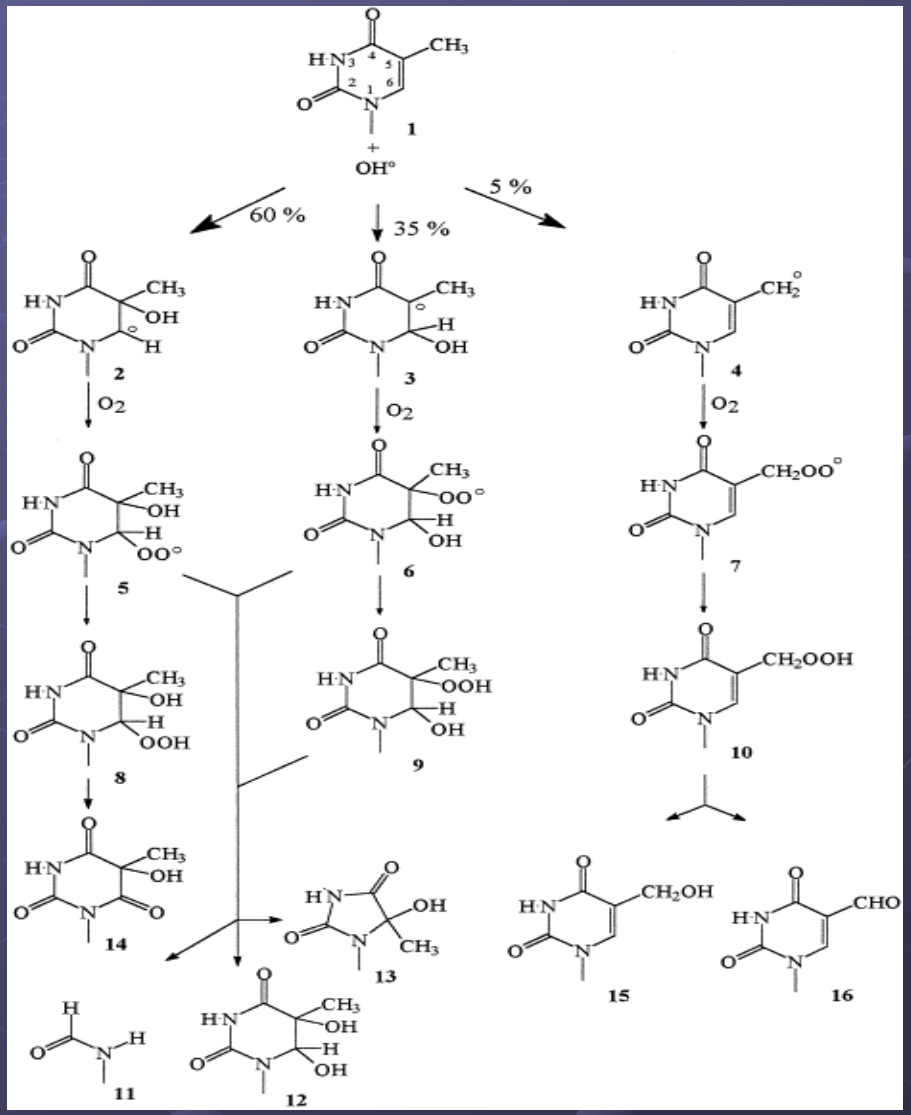


DNA Damage and Repair Cont.

- **Oxidative Damage Cont.**

<u>Thymine →thymidine glycol, hydantoins</u>

The predominant products are 5,6-dihydroxy-5,6-dihydro-thymidine (thymidine glycol) (**12**) and 5-hydroxy-5-methylhydantoin (**13**). These lesions block replication.



DNA Damage and Repair Cont.

- **Environmental Damage**
  - **Ionizing radiation**

    Radiolysis of water

    $H_2O \rightarrow H_2O^+ + e^-$

    Formation of other ROS

    $H_2O^+ + H_2O \rightarrow OH\cdot + H_3O^+$

    $OH\cdot + OH\cdot \rightarrow H_2O_2$

    $e_{aq} + O_2 \rightarrow O_2^-$

    $2\ O_2^- + 2H^+ \rightarrow O_2 + H_2O_2$

    Base damage

    Damage is $OH\cdot$ –mediated, similar to that previously described for oxidative damage. Phenomenon of 'local multiply damaged sites' → single energy deposition event can result in several radical reactions in the immediate vicinity.

    - **Sugar damage and strand breaks**

      Damage to sugar residues less frequent than damage to bases (by ~3 fold). Due to modified structures at the 3' end (e.g. phosphoglycolate O_3POCH_2COO-) precludes repair by a simple ligation step.

      Single and strand breaks are also formed.

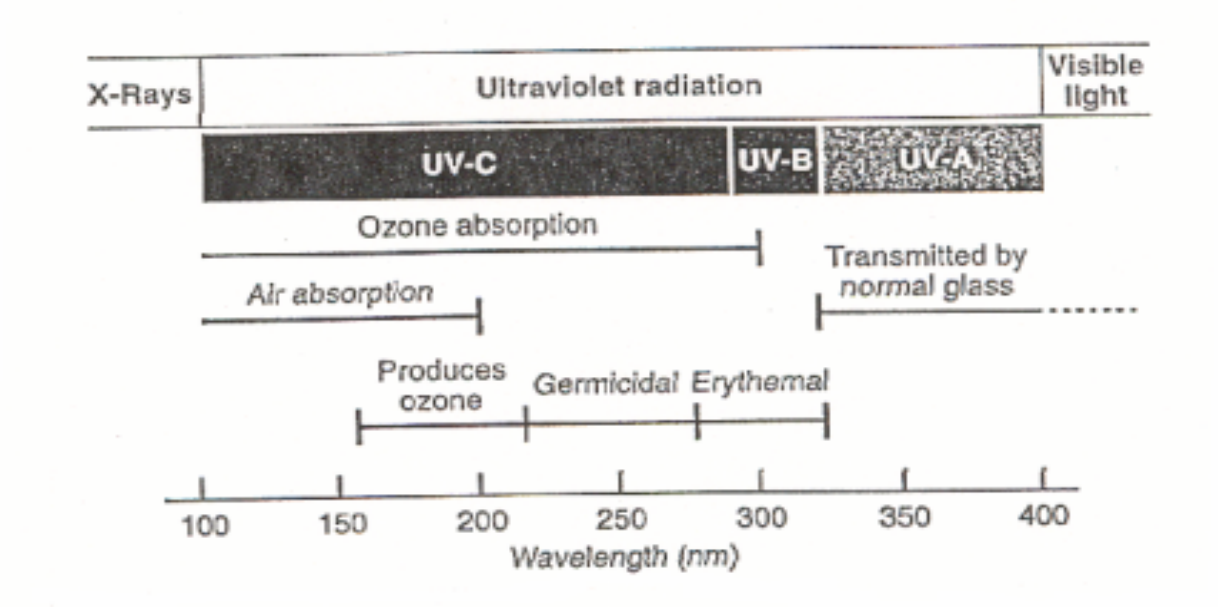
      Controversy exists due to the formation of these lesions at non-physiological doses.

DNA Damage and Repair Cont.

- **UV Radiation**

UV Radiation Spectrum

UV spectrum divided into UV-A (400-320 nm), UV-B (320-290 nm) and UV-C (290-100 nm). Solar radiation consists of mainly UV-A and UV-B since ozone absorbs at 300 nm.



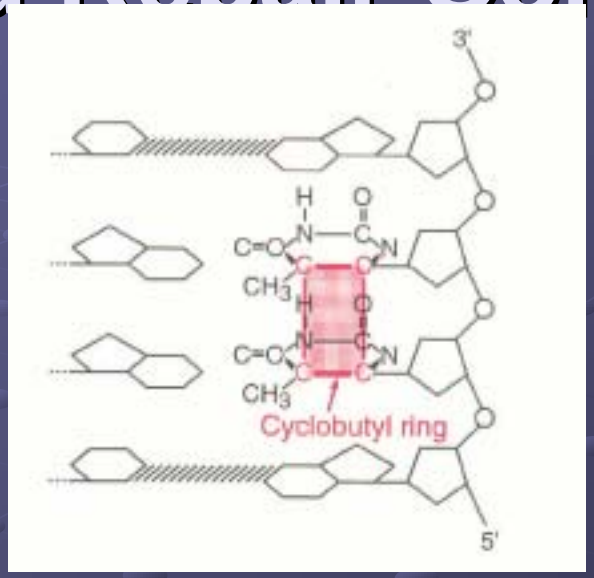
DNA Damage and Repair Cont.

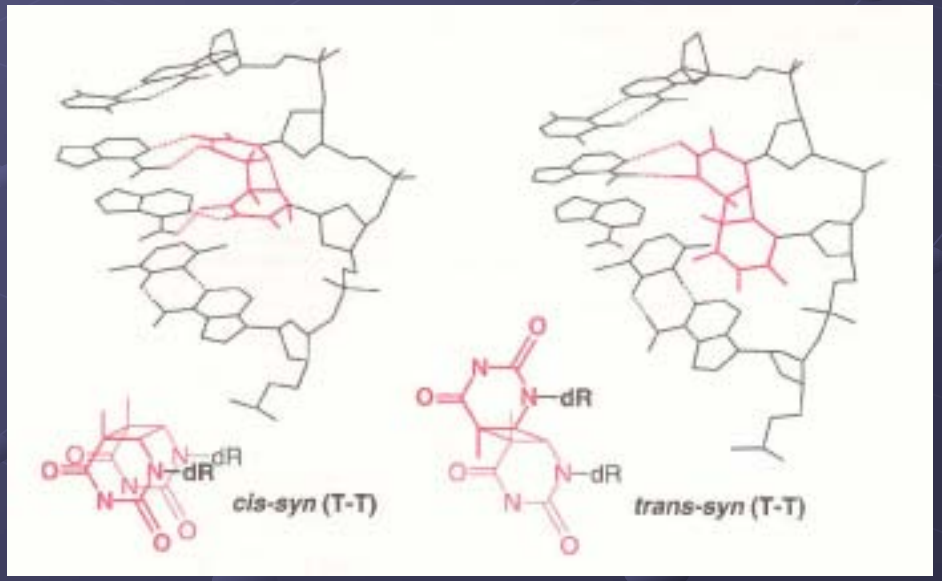
- UV Radiation Cont.

CPD

When DNA is exposed to wavelengths approaching its absorption maximum (260 nm), adjacent pyrimidines become covalently linked by the formation of a cyclobutane ring structure resulting from saturation of their respective 5,6 double bonds → called a cyclobutane pyrimidine dimer (CPD)

- 12 isomeric forms (only 4 found in any quantity)
  - *cis-syn* in B form DNA
  - *trans-syn* in denatured DNA, ssDNA, B-Z DNA junctions
  - extremely stable to pH and temperature extremes





DNA Damage and Repair Cont.

- **UV Radiation Cont.**

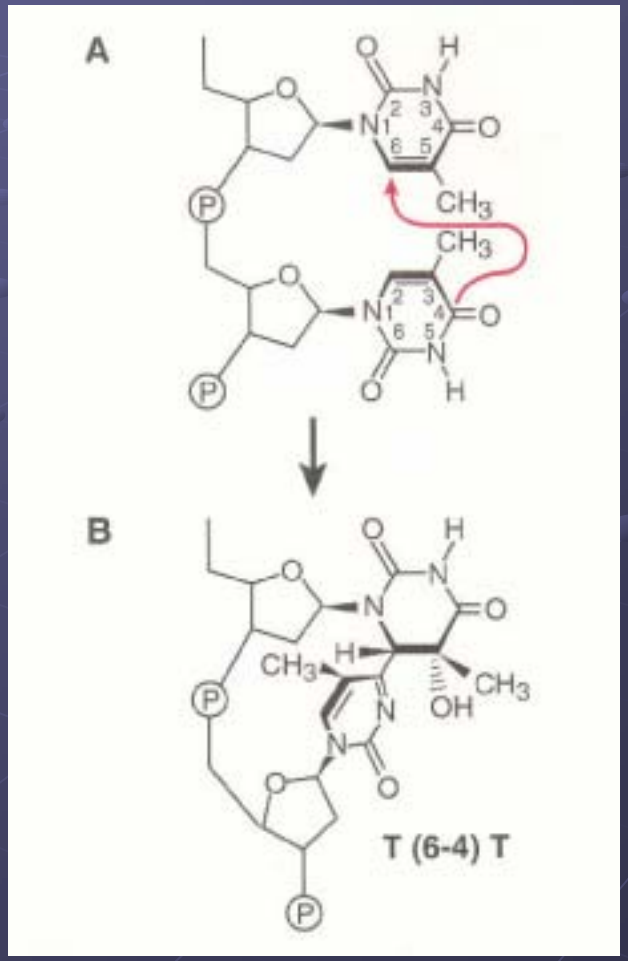
CPD Cont.

- blocks DNA polymerase read through
- bulky , helix distorting lesions
- T-T >> C-T, T-C >> C-C (68:13:16:3)
- CPD yield dependent upon sequence context outside of dimer
- process: Py + Py $\xleftarrow{UV}$ Py <>Py
  - equilibrium exists such that the amount of dimers 7% of total thymine

(6-4) PP

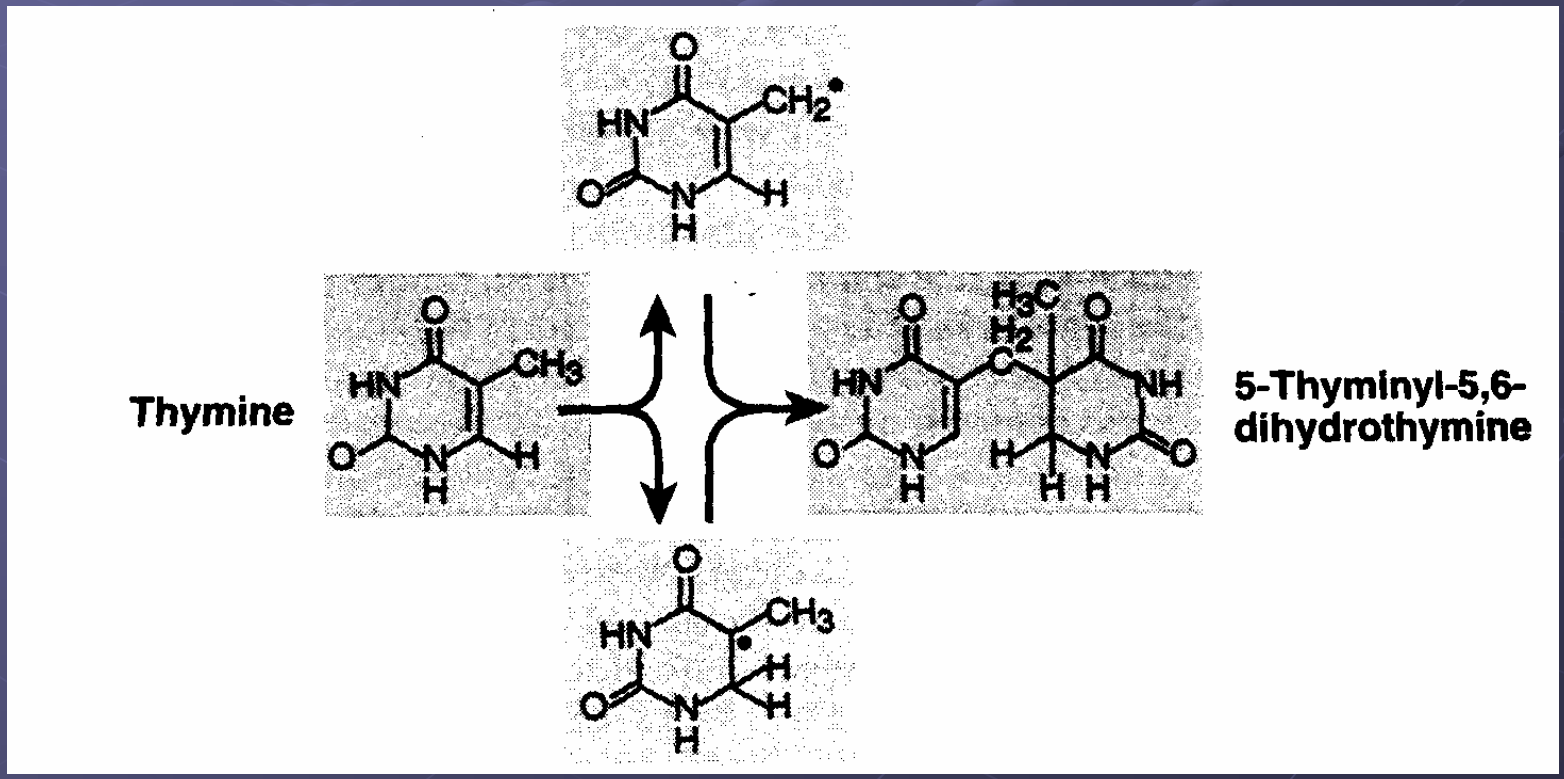
Alkali labile lesions also formed in UV treated DNA, called pyrimidine-pyrimidone (6-4) lesions or photoproducts, (6-4)PPs.

- causes major helical distortions
- blocks DNA polymerase read through
- TC, CC more common than TT (CT not seen)
- (6-4)PP — $\xrightarrow{313nm}$ — Dewar isomer



Spore photoproduct

- uv irradiated spores of B. subtilis
- related to state of hydration (A form in dehydrated DNA)



DNA Damage and Repair Cont.

- Chemical Agents
  - Alkylating agents
    Electrophilic compounds that react with nucleophilic centers in DNA.
    - O6-methylguanine
      - generated endogenously
      - pairs with C or T
      - MGMT enzyme
    - N7-methylguanine
    - N3-methyladenine (MPG glycosylase)
  - Cross Linking Agents
    Bifunctional alkylating agents can cause intra- or inter-strand crosslinks
    - Nitrogen and sulfur mustard, mitomycin and *cis*-platinum
    - Also caused by UV and ionizing radiation

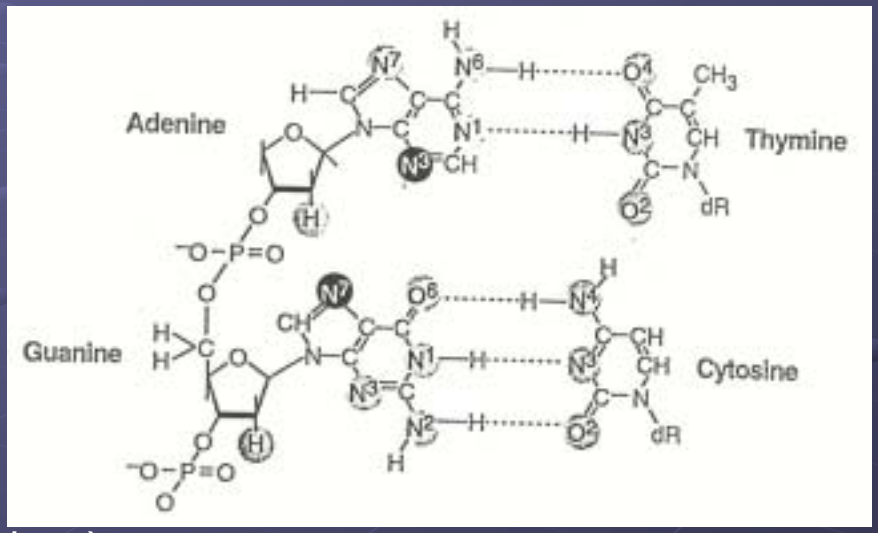


Table 3–1 *Cellular responses to DNA damage*

Response	Mechanism
Reversal of DNA damage	*Enzymatic photoreactivation* Repair of spore photoproduct Repair of O^6-alkylguanine, O^4-alkylthymine, and alklyphosphotriesters *Ligation of DNA strand breaks*
Excision of DNA damage	*Base excision repair* Nucleotide excision repair *Mismatch repair*
Tolerance of DNA damage	Replicative bypass of template damage with gap formation and recombination Translesion *DNA synthesis*

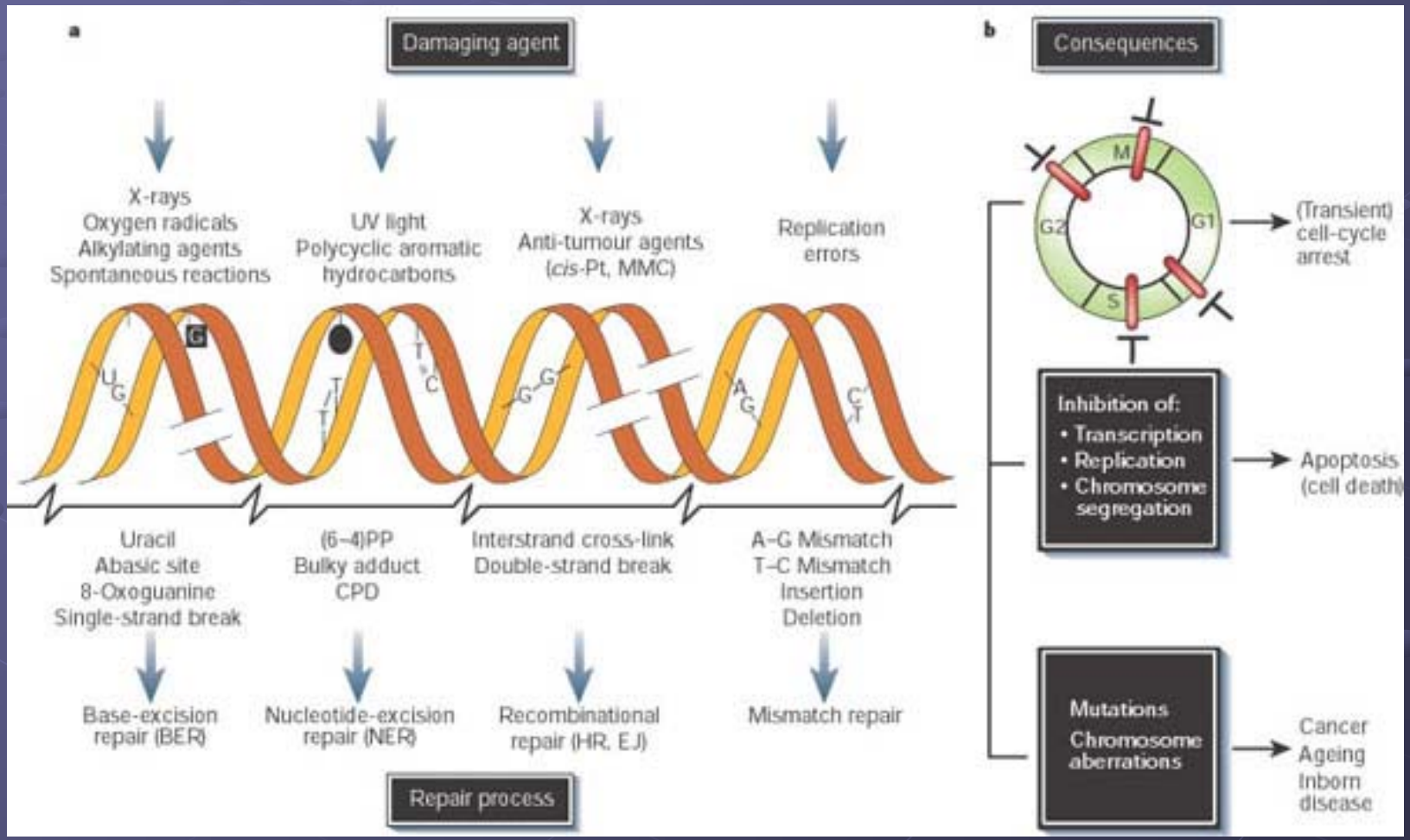
DNA Damage and Repair Cont.

- **DNA Repair**

  Cancer is a genetic disease in the sense that mutations can activate proto-oncogenes or inactivate tumor suppressor genes. However, in addition to these spontaneous mutations, cancer risk is increased due to the 'mutator' phenotype caused by inherited or acquired faulty DNA repair systems. The main mammalian DNA repair systems include:

  - Direct Reversal
  - Replicational Bypass (Translesion Synthesis)
  - Nucleotide excision repair (NER)
  - Base excision repair (BER)
  - Homologous recombination
  - End joining
  - Mismatch repair

DNA Damage and Repair Cont.



DNA Damage and Repair Cont.

Table 1 Human syndromes with defective genome maintenance

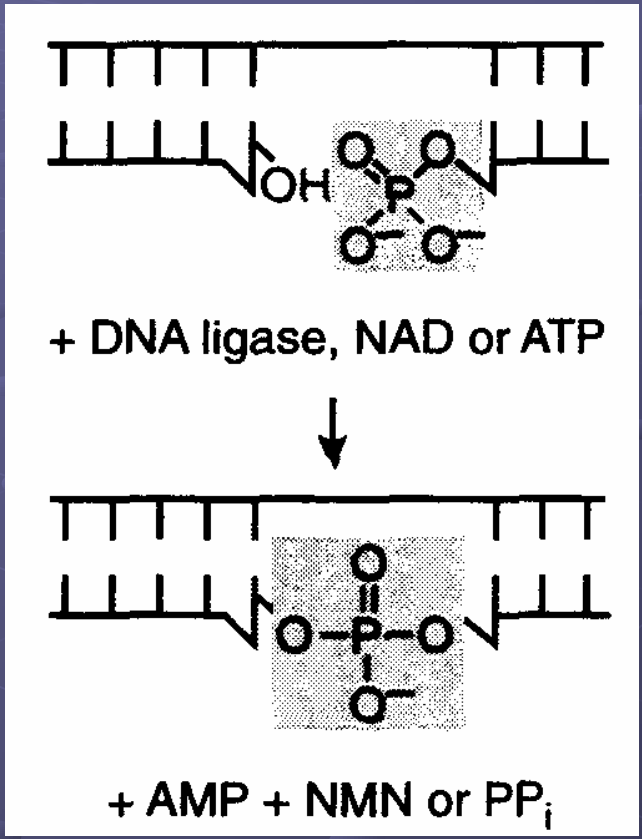
Syndrome	Affected maintenance mechanism	Main type of genome instability	Major cancer predisposition
Xeroderma pigmentosum	NER (± TCR)	Point mutations	UV-induced skin cancer
Cockayne syndrome	TCR	Point mutations	None*
Trichothiodystrophy	NER / TCR	Point mutations	None*
Ataxia telangiectasia (AT)	DSB response/repair	Chromosome aberrations	Lymphomas
AT-like disorder	DSB response/repair	Chromosome aberrations	Lymphomas
Nijmegen breakage syndrome	DSB response/repair	Chromosome aberrations	Lymphomas
BRCA 1/BRCA2	HR	Chromosome aberrations	Breast (ovarian) cancer
Werner syndrome	HR?/TLS?	Chromosome aberrations	Various cancers
Bloom syndrome	HR?	Chromosome aberrations (SCE↑)	Leukaemia, lymphoma, others
Rothmund–Thomson syndrome	HR?	Chromosome aberrations	Osteosarcoma
Ligase IV deficiency†	EJ	Recombination fidelity	Leukaemia(?)
HNPCC	MMR	Point mutations	Colorectal cancer
Xeroderma pigmentosum variant	TLS‡	Point mutations	UV-induced skin cancer

*Defect in transcription-coupled repair triggers apoptosis, which may protect against UV-induced cancer.
†One patient with leukaemia and radiosensitivity described with active-site mutation in ligase IV.
‡Specific defect in relatively error-free bypass replication of UV-induced cyclobutane pyrimidine dimers.
Abbreviations: BER, base-excision repair; DSB, double-strand break; HNPCC, hereditary non-polyposis colorectal cancer; HR, homologous recombination; MMR, mismatch repair; NER, nucleotide-excision repair; SCE, sister-chromatid exchange; TCR, transcription-coupled repair; TLS, translesion synthesis.

DNA Repair

Direct Reversal

- Enzymatic photoreactivation
- Repair of spore photoproduct
- Repair of O^6 alkyl guanine, O^4-alkyl thymine and alkylphosphotriesters
- Ligation of single strand breaks

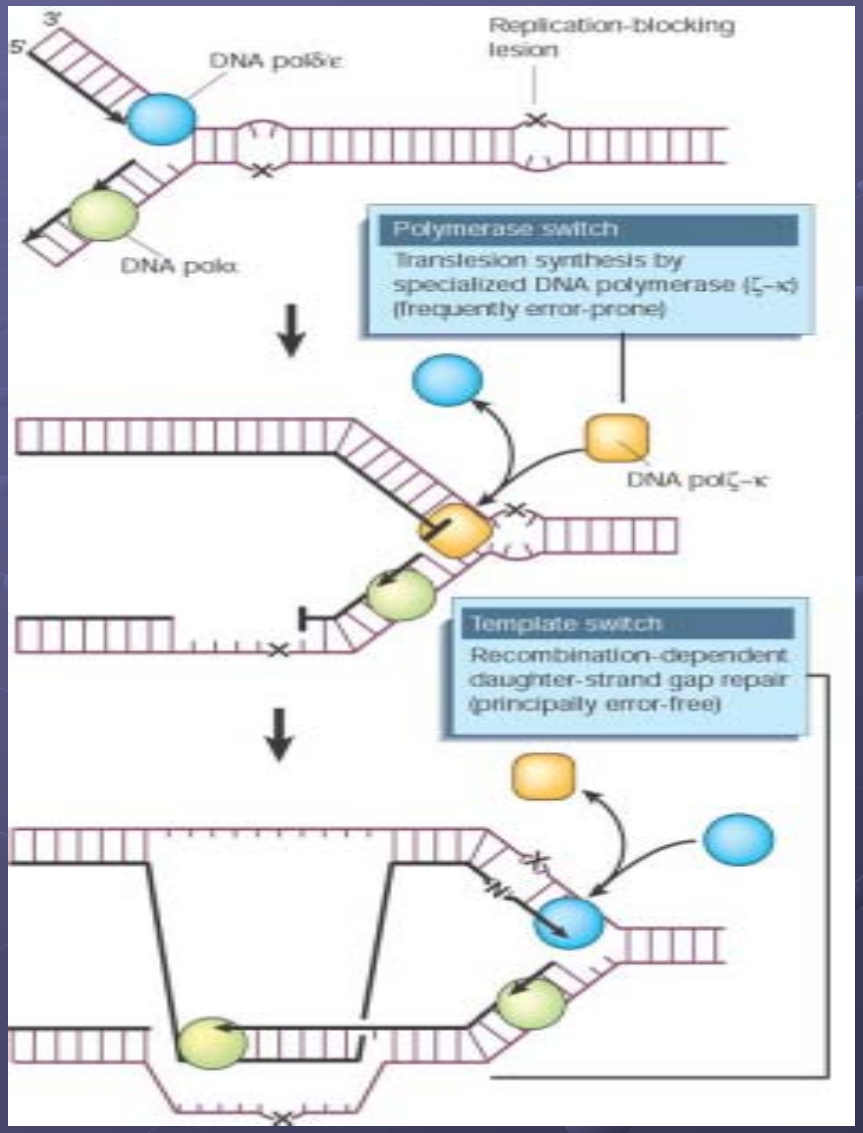


Ligation of single stranded breaks

Replicational Bypass

- To minimize cell death from replication blockage a process of translesion synthesis (TLS) has evolved
  - TLS-, SOS-, mutagenic-, error-prone-, Y-family- DNA polymerases
- Bacterial mutatants non-mutable by UV radiation in 1976
  - Mutations are not accidental but result from active chemical process
- Hypothesized that error prone DNA polymerases exist that could insert random nucleotides in a template independent fashionand that DNA damage induced mutagenesis might be a specific cellular response to damage
- In E. coli, SOS response involves >30 inducible genes
  - Initially *recA and lexA* mutants
  - Then, *umuC* and *umuD* (pol V) and *dinB* (pol IV)
  - Genome screens for homologs resulted in the identification of an extended superfamily (Y family)

DNA Damage and Repair Cont.



Nucleotide Excision Repair (NER)

NER deals with a wide class of helix distorting lesions. Over 25 proteins participate in the NER pathway. Main function is to remove UV photoproducts, crosslinks and other bulky lesions. Two distinct pathways:

- Global genome NER
  - **surveys entire genome for distortions**
- Transcription coupled repair (TCR)
  - **acts on damage that blocks RNA polymerase**

Xeroderma Pigmentosa (XP) is a disorder of the NER system which increases a patients risk for developing skin cancer >2000 fold.

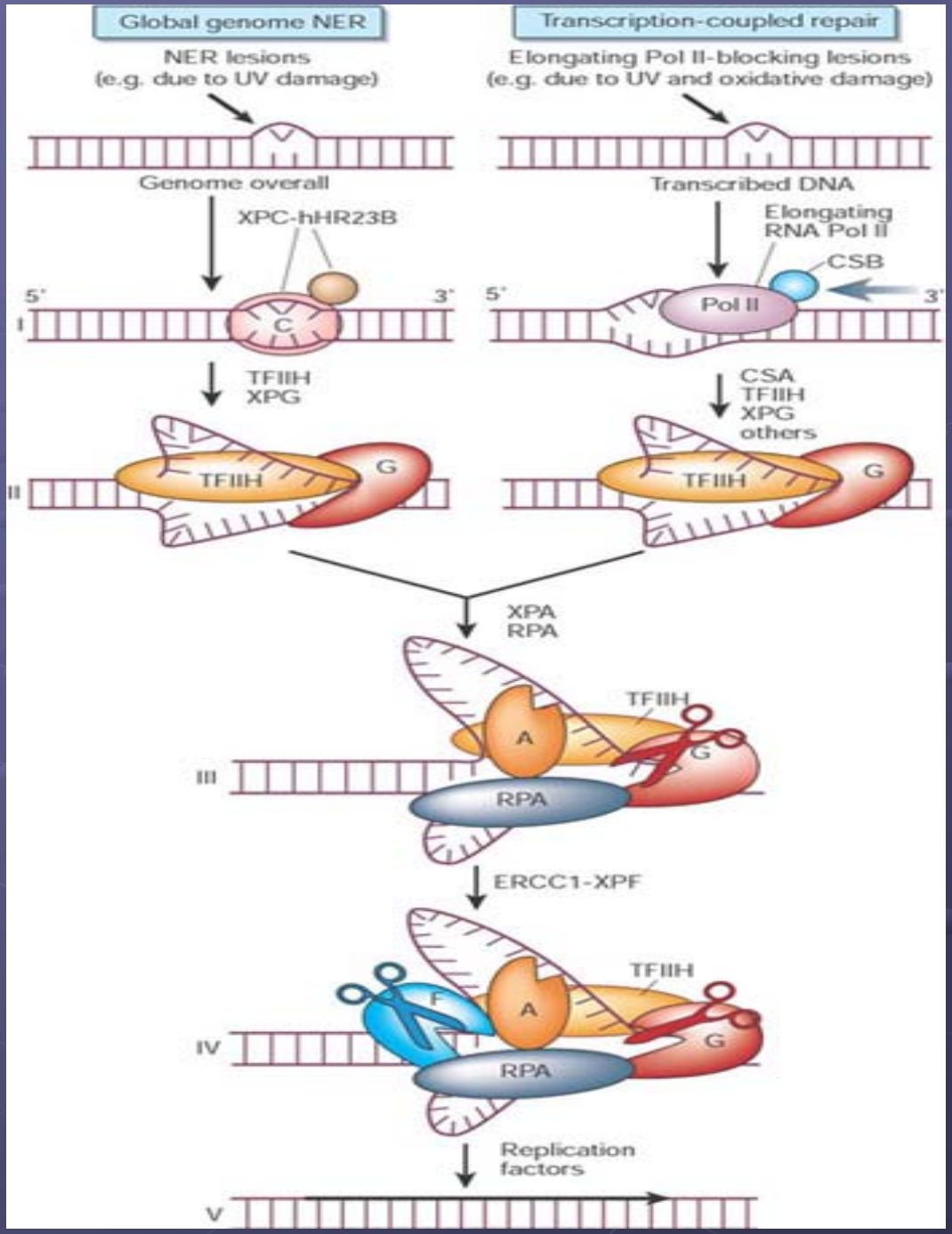
- caused by inability to excise UV photoproducts (e.g. CPDs)
- mutations in one of seven genes (XPA-XPG)

Cockayne's syndrome (CS) is a TCR specific disorder caused by mutations in the CSA or CSB genes.

- no predisposition to cancer
- CS cells particularly susceptible to apoptosis
  - **protects against tumorigenesis**

NER

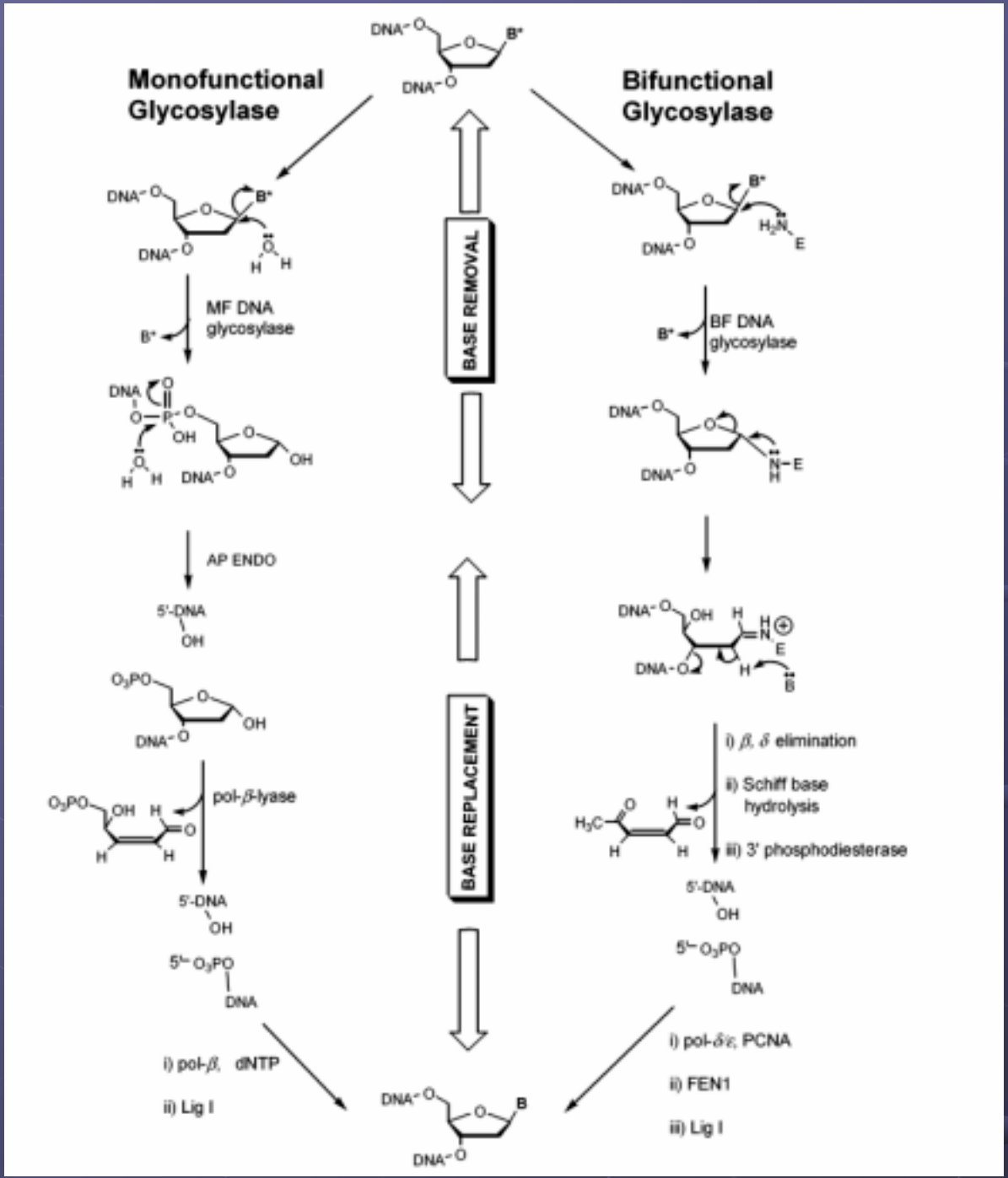
- Damage recognition
- DNA unwinding (discrimination of damaged from non-damaged strand)
- Incision on both sides of the lesion
  - Removal of ~25-30 nucleotides
- Repair synthesis
- Ligation


Global genome NER
NER lesions
(e.g. due to UV damage)
Genome overall
XPC-hHR23B
5'
3'
I
C
TFIIH
XPG
II
TFIIH
G
Transcription-coupled repair
Elongating Pol II-blocking lesions
(e.g. due to UV and oxidative damage)
Transcribed DNA
Elongating
RNA Pol II
CSB
5'
Pol II
3'
CSA
TFIIH
XPG
others
TFIIH
G
XPA
RPA
TFIIH
A
G
III
RPA
ERCC1-XPF
TFIIH
F
A
G
IV
RPA
Replication
factors
V

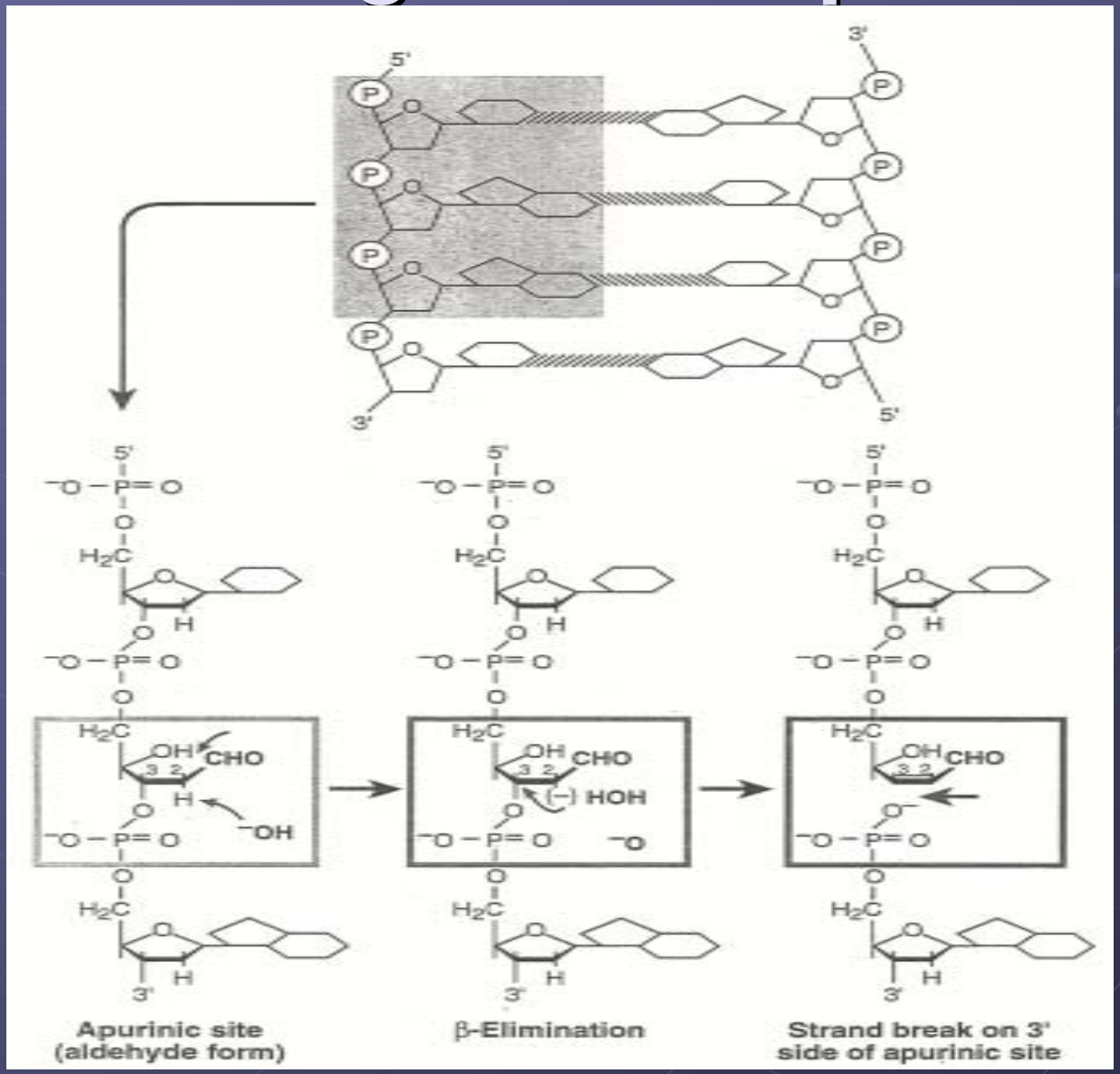
DNA Damage and Repair Cont.

- **DNA Repair Cont.**
  - Base Excision Repair (BER)

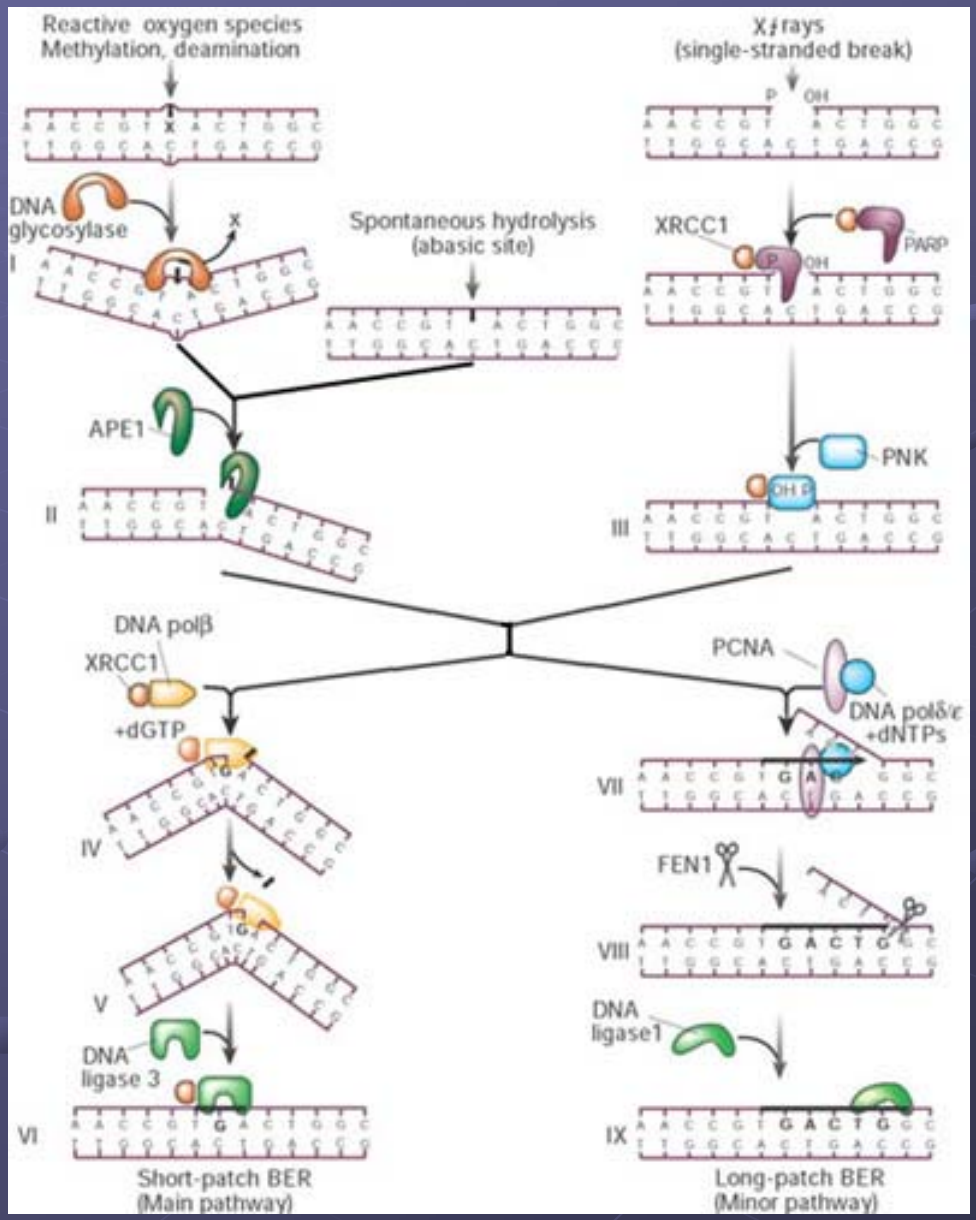
    BER is the cell's main protection against ROS, methylation and deamination.
    - no disorders caused by inherited deficiencies in BER have been identified
      - partial redundancies between different glycosylases
    - however, inactivation of BER core proteins causes embryonic lethality


Monofunctional Glycosylase
Bifunctional Glycosylase
MF DNA glycosylase
BF DNA glycosylase
BASE REMOVAL
AP ENDO
5'-DNA
OH
pol-β-lyase
BASE REPLACEMENT
i) β, δ elimination
ii) Schiff base hydrolysis
iii) 3' phosphodiesterase
i) pol-β, dNTP
ii) Lig I
i) pol-δε, PCNA
ii) FEN1
iii) Lig I

DNA Damage and Repair Cont.



DNA Damage and Repair Cont.



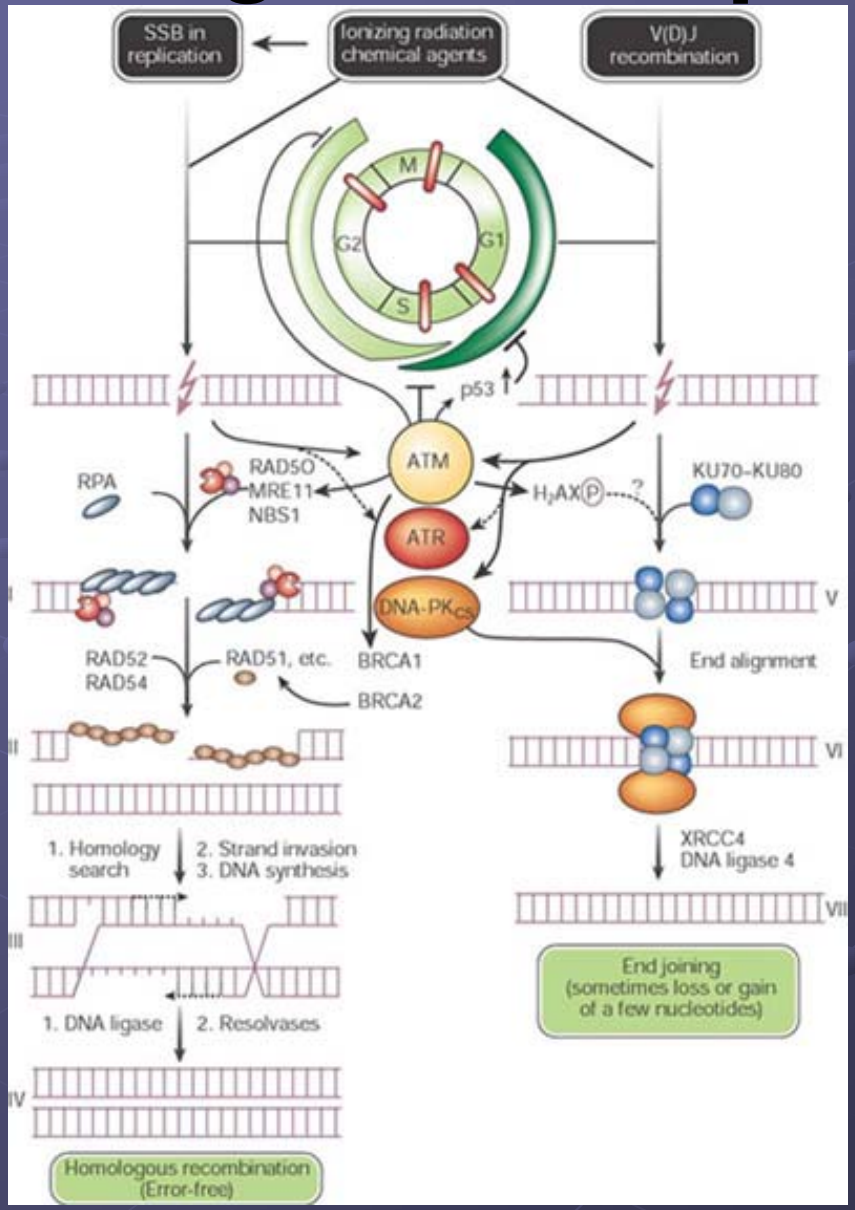
DNA Damage and Repair Cont.

- **DNA Repair Cont.**
  - Homologous Recombination (HR) and End Joining (EJ)

    Double strand breaks (DSBs) arise due to:
    - Ionizing radiation/X rays
    - Free radicals

    Once DSB is detected, a complex machinery is mobilized to halt cell cycle progression and to recruit repair factorsIf second undamaged copy of template is available --> HRIf not, ER --> more error prone

DNA Damage and Repair Cont.



DNA Damage and Repair Cont.

- **DNA Repair Cont.**
  - **Mismatch Repair**

    MMR removes:
    - **Nucleotides mispaired during replication**
    - **Insertion/deletion loops due to slipped strand mispairing of repetitive sequences (microsatellites) during replication or recombination**

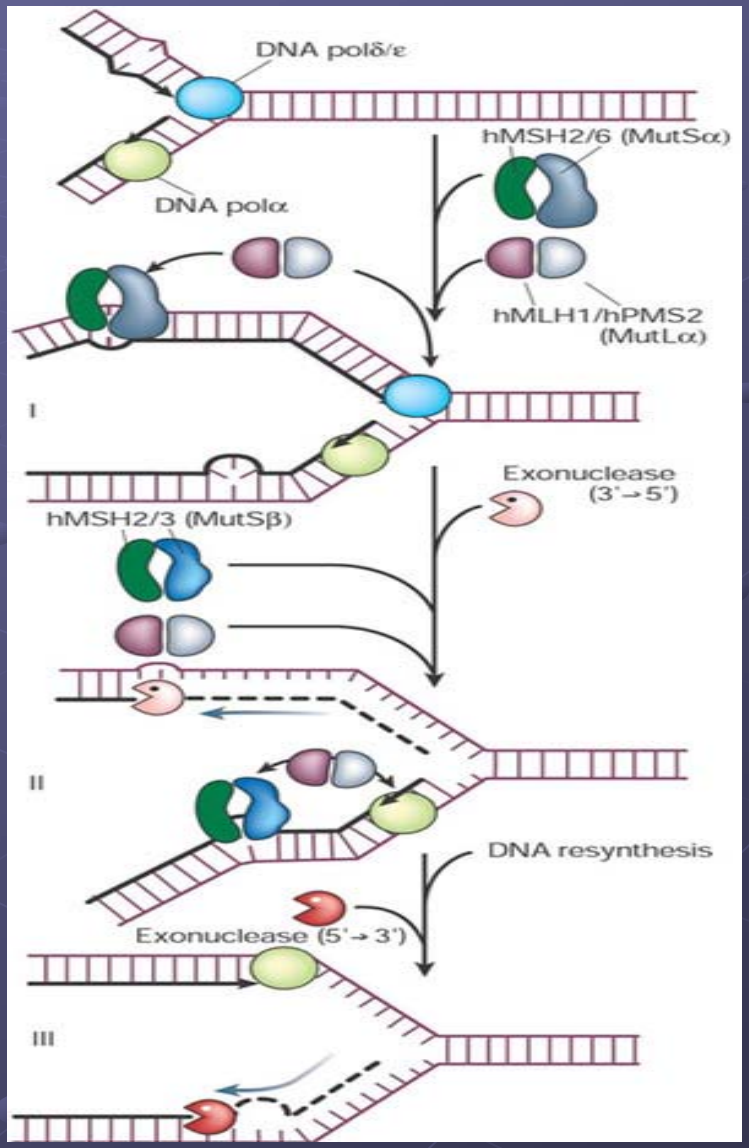
    The microsatellite instability (MSI) phenotype of hereditary non-polyposis colorectal cancer (HNPCC) caused mainly by mutations in MMR genes
    - **especially hMLH1 and hMSH2**

DNA Damage and Repair Cont.

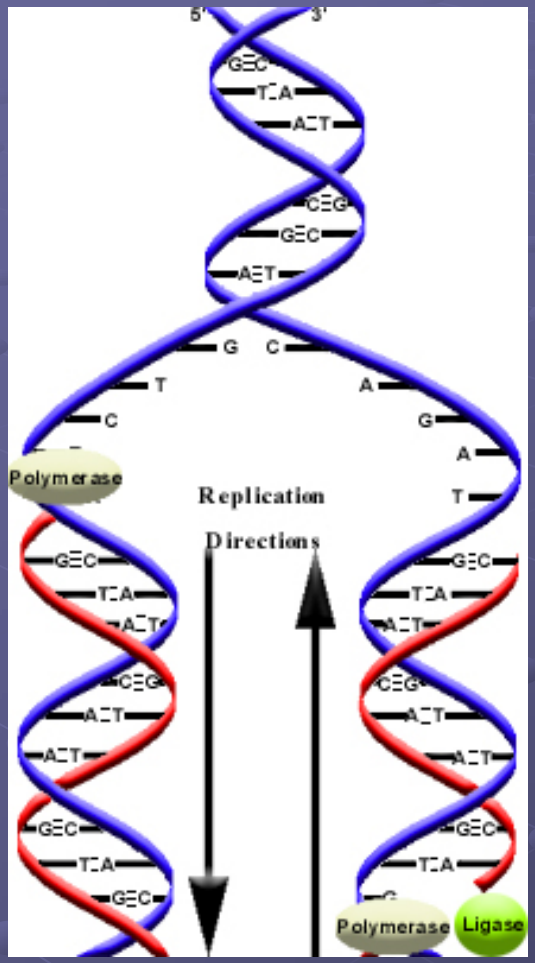
E. coli	*S. cerevisiae*	*H. sapiens*	Function
MutS	*Msh2*	*MSH2*[a]	Recognizes the mismatch
	Msh6	*MSH6*[a]	Forms a complex with *Msh2* (*MSH2*)
	Msh3	*MSH3*	Forms a complex with *Msh2* (*MSH2*)
	Msh1	Not identified	Functions in mitochondrial MMR
	Msh4	*MSH4*	Required for meiotic recombination
	Msh5	*MSH5*	Required for meiotic recombination
MutL	*Mlh1*	*MLH1*[a]	Couples mismatch recognition and subsequent repair
	Pms1	*PMS2*[a]	Forms a complex with *Mlh1* (*MLH1*)
	Mlh2	*PMS1*[a]	Forms a complex with *Mlh1* (*MLH1*)
	Mlh3	*MLH3*	Forms a complex with *Mlh1* (*MLH1*)

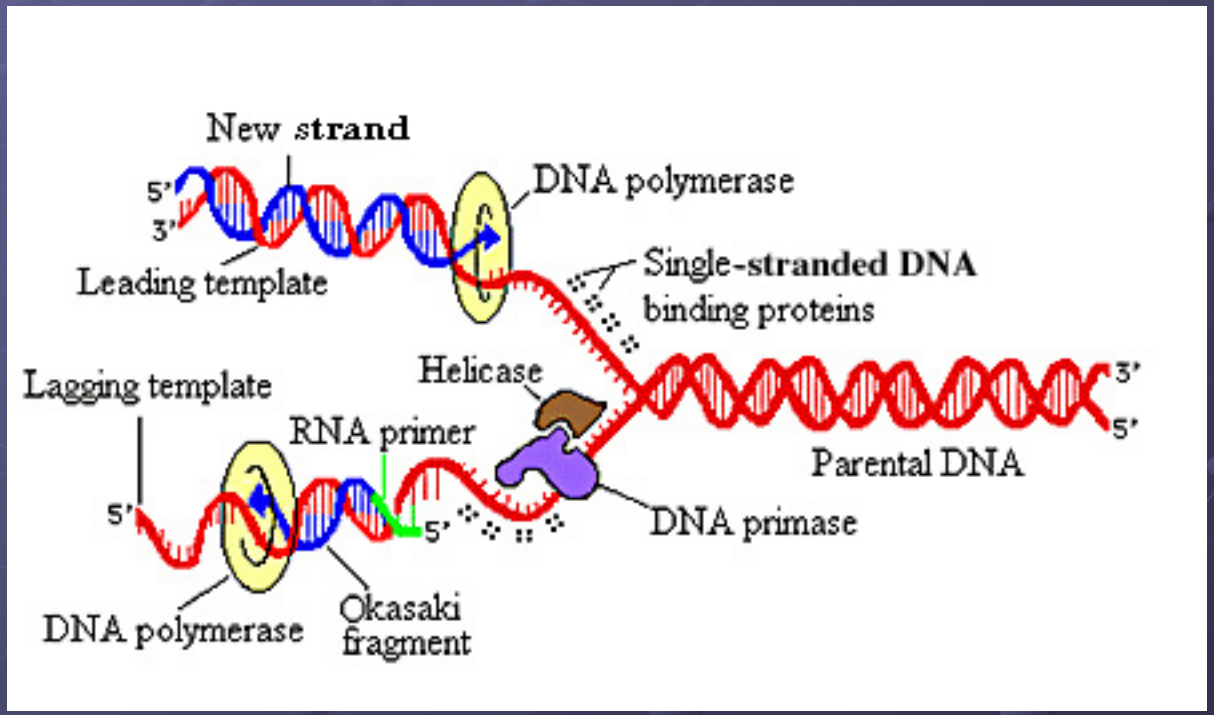
MMR gene	Mechanism of biallelic inactivation		Microsatellite instability phenotype
	Sporadic	Hereditary	
MSH2	Somatic mutation + LOH[a] or two somatic mutations	Germline mutation + LOH or germline + somatic mutation	High degree of instability of di-, mononucleotide, and other short tandem repeats
MLH1	Promoter hypermethylation (biallelic) (or somatic mutation + LOH, or two somatic mutations)	Germline mutation + LOH	High degree of instability of di-, mononucleotide, and other short tandem repeats
MSH6	Frequent frameshift mutations in coding C_8 repeat[b]	Germline mutation + somatic (frameshift) mutation or germline mutation + LOH	Low degree of instability with preferential involvement of mononucleotide repeats
MSH3	Frequent frameshift mutations in coding A_8 repeat[b]	None identified	Data not available[c]
PMS2	Two somatic mutations	Germline + somatic mutation	High degree of instability of dinucleotide and other repeats
PMS1	Data not available	Data not available	Data not available
MLH3	Data not available	Data not available	Data not available

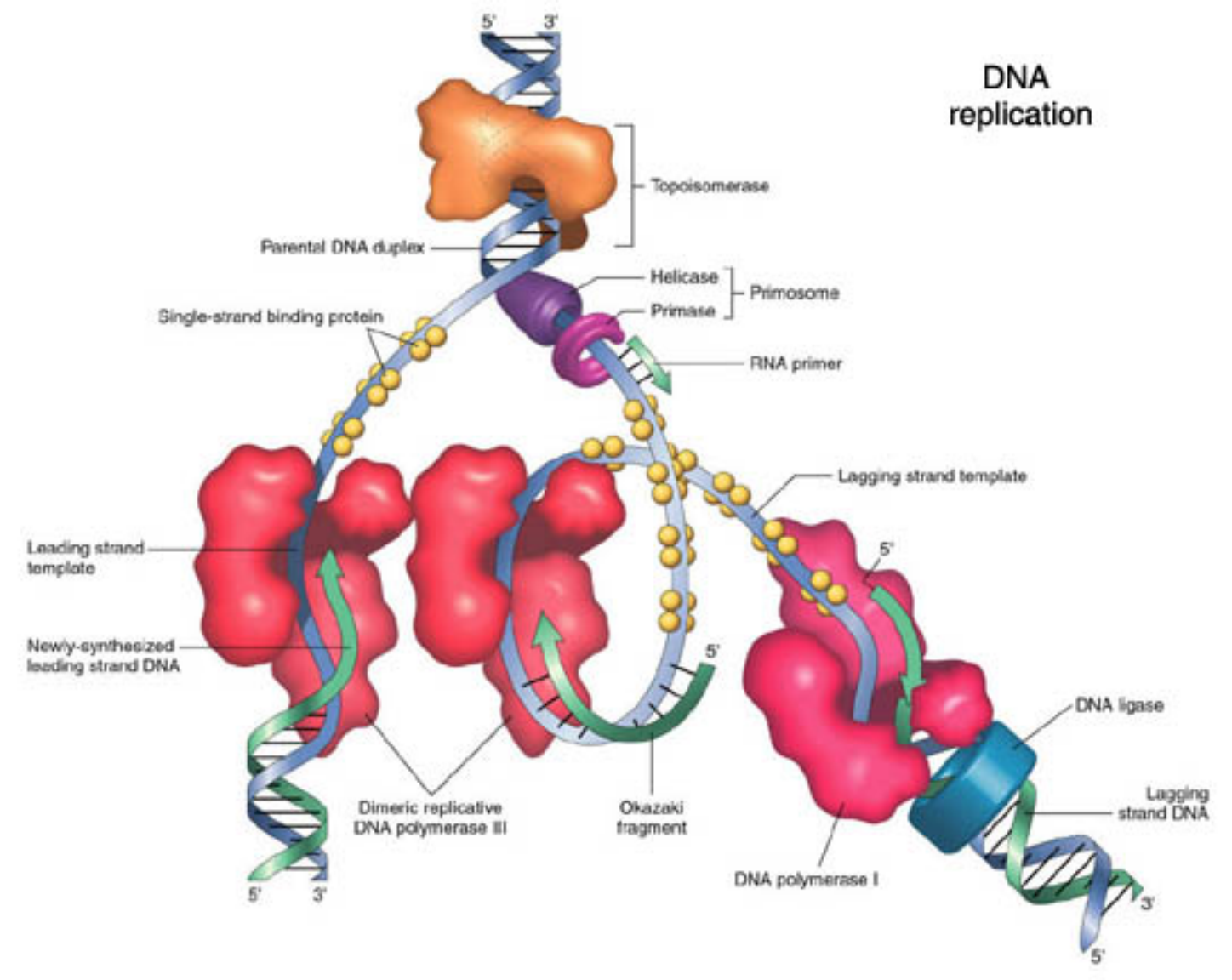
DNA Damage and Repair Cont.



DNA Replication






DNA replication
5'
3'
Topoisomerase
Parental DNA duplex
Helicase
Primase
Primosome
Single-strand binding protein
RNA primer
Lagging strand template
Leading strand template
Newly-synthesized leading strand DNA
5'
5'
DNA ligase
Lagging strand DNA
Dimeric replicative DNA polymerase III
Okazaki fragment
DNA polymerase I
5'
3'
3'
5'

Functional RNA Molecules

- rRNA
- mRNA
- tRNA
- catalytic RNA
- small nuclear RNA (snRNA)(splicing)
- guide RNA (silencing)
- telomerase RNA (chromosome end replication)
- signal recognition particle (SRP) RNA (protein translocation)
- small nucleolar RNA (snoRNA) (rRNA modification)
- micro-RNA

RNA Structure

- The difference in the biology of RNA compared with DNA stems from the presence in RNA of the 2’ OH residue in the sugar ring. This creates a nucleophilic center that changes the reactivity of the molecule and sterically hinders the formation of a ‘B form ‘ double helical’ structure.

RNA Structure

- RNA synthesized as single strands
- folds into stable structure, e.g. tRNA
  - short regions of W-C base pairing interrupted by:
    - non-canonical pairs
    - loops, including hairpins
    - bulged nucleotides
    - more complex structures e.g. pseudoknots, junctions
  - 2D structure predicted by use of thermodynamic secondary structure prediction algorithms
    - includes covariation between two nucleotides that conserve W-C --> evidence of a base pair between these nucleotides
- modified nucleotides and sugars
  - pseudouridine, 5' methyl cytosine, ribothymidine (tRNA greatest number and percentage of these)
  - 2'-O-methyl substituents
- dsRNA --> A-form
  - not > 10 bp
  - stabilized by H-bonding of the 2'-OH group with neighboring phosphates or other riboses
  - Non W-C pairs common e.g. G-U and A-C

mRNA

Life cycle

Nucleus: transcription
capping
methylation
polyadenylation
splicing
editing

Translocation to cytoplasm

Cytoplasm: translation
degradation
poly(A) remodeling/localisation

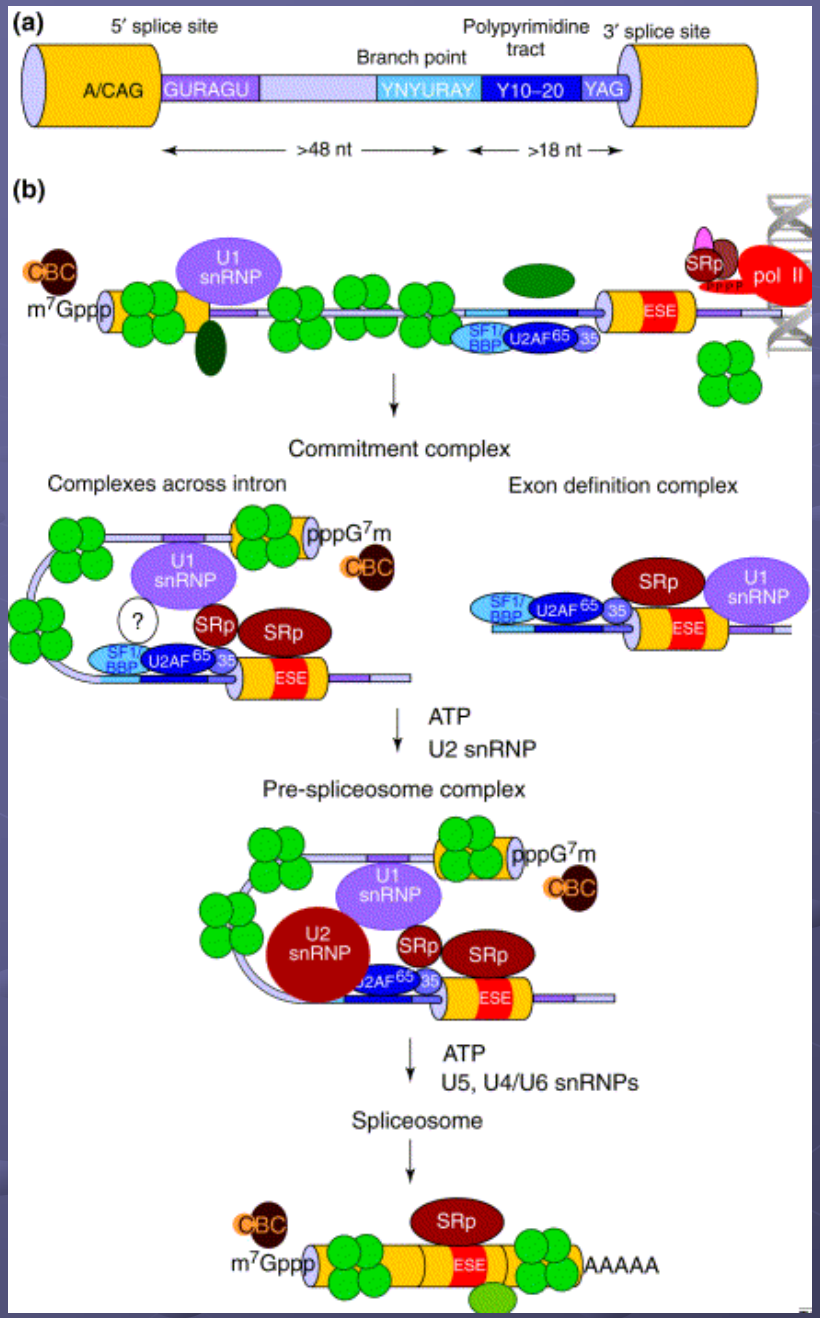
Structural elements

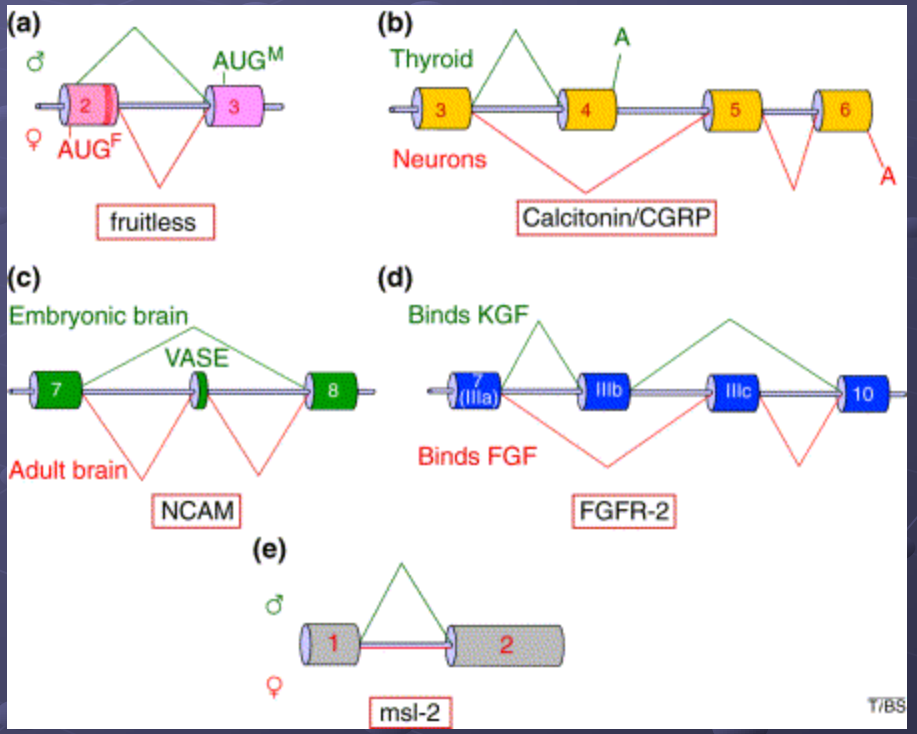
- 5’ cap
- 5’ UTR
- coding region
- 3’ UTR
- poly(A)

RNA Splicing

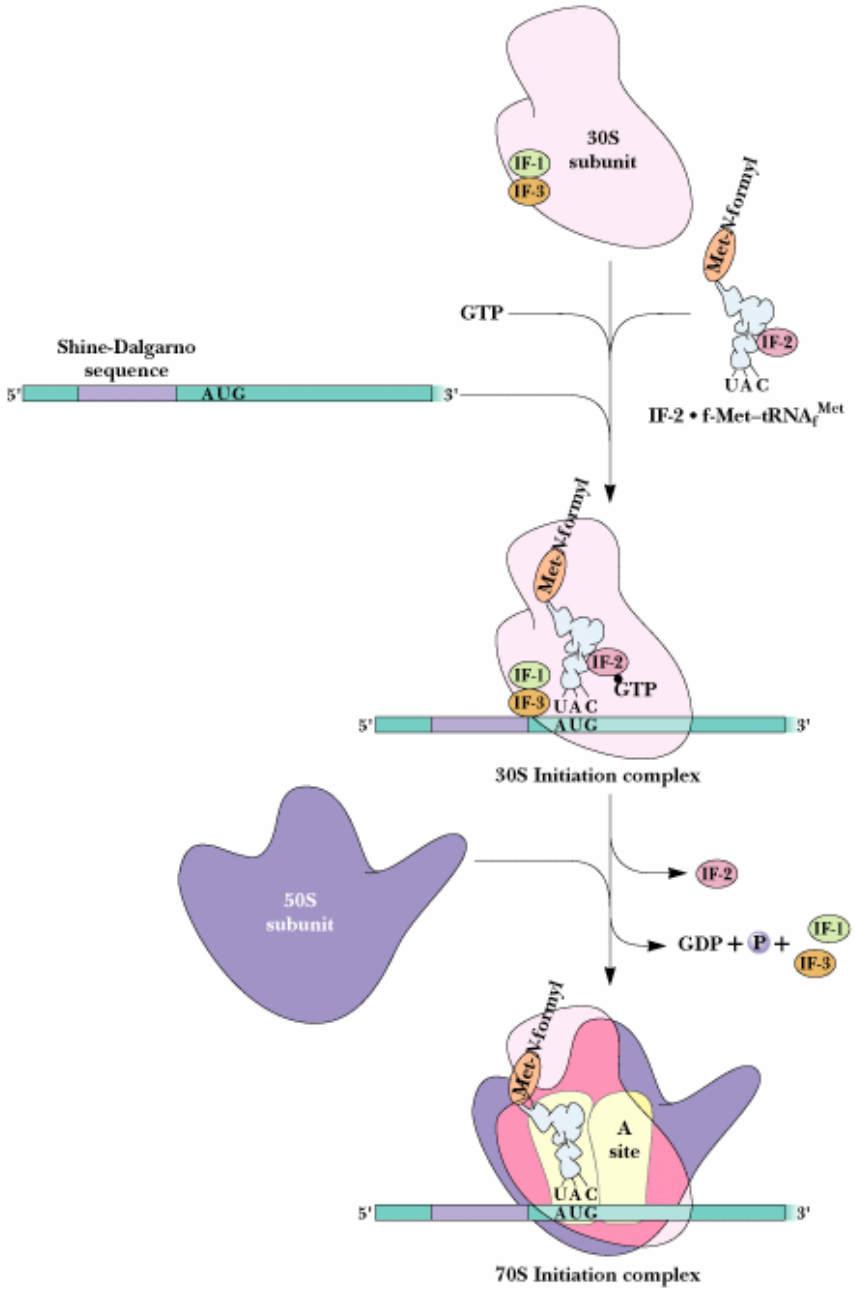
Eukaryotic mRNAs are transcribed as precursors (pre-mRNA) containing intervening sequences (introns). These sequences are subsequently removed and the flanking regions containing the exons are spliced together to form mature mRNA.

Splicing

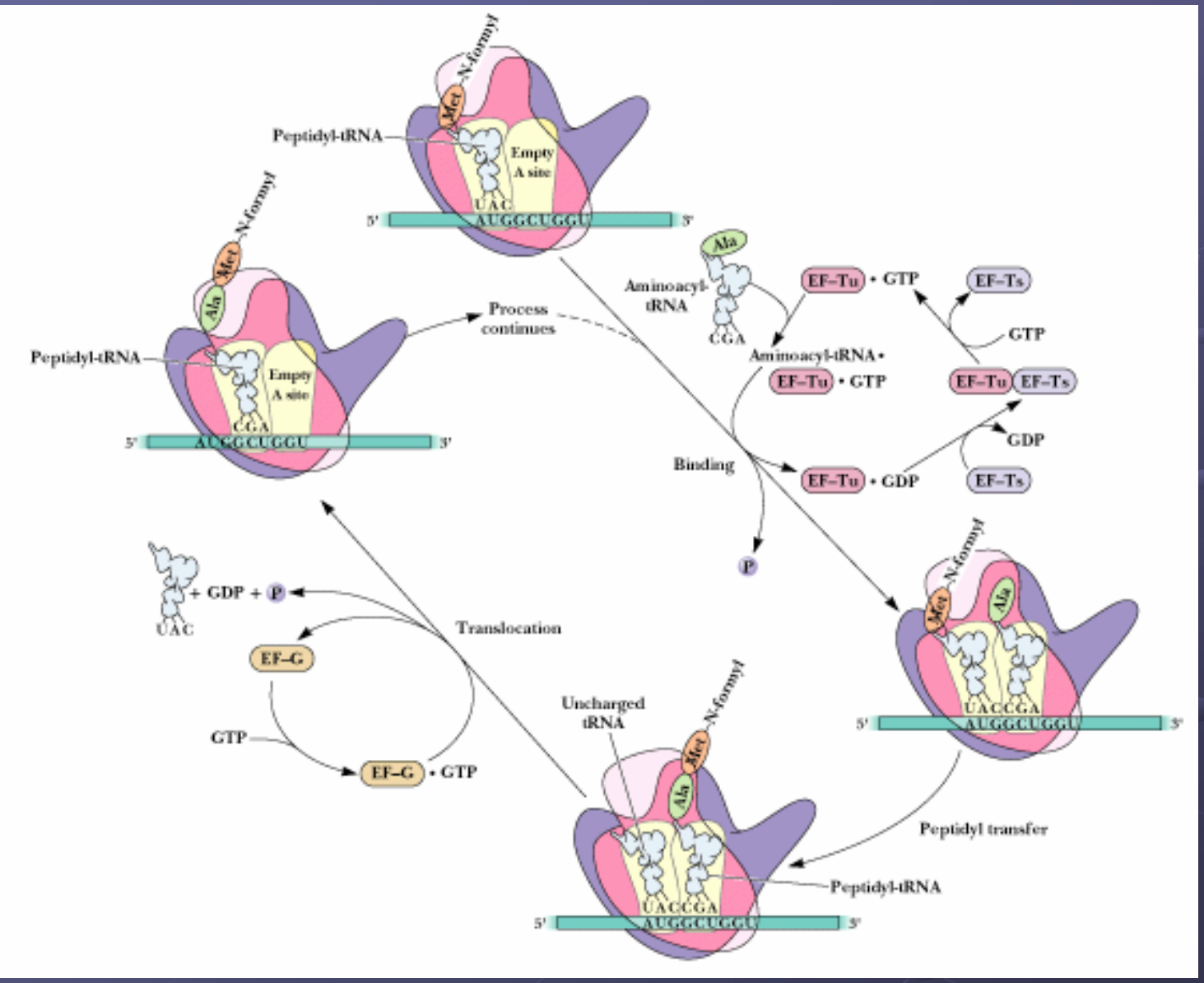

(a)
5′ splice site
Polypyrimidine tract
3′ splice site
Branch point
A/CAG
GURAGU
YNYURAY
Y10–20
YAG
>48 nt
>18 nt
(b)
CBC
m7Gppp
U1 snRNP
SRp
pol II
SF1/BBP
U2AF65
35
ESE
Commitment complex
Complexes across intron
Exon definition complex
pppG7m
?
ATP
U2 snRNP
Pre-spliceosome complex
U2 snRNP
Spliceosome
AAAAA

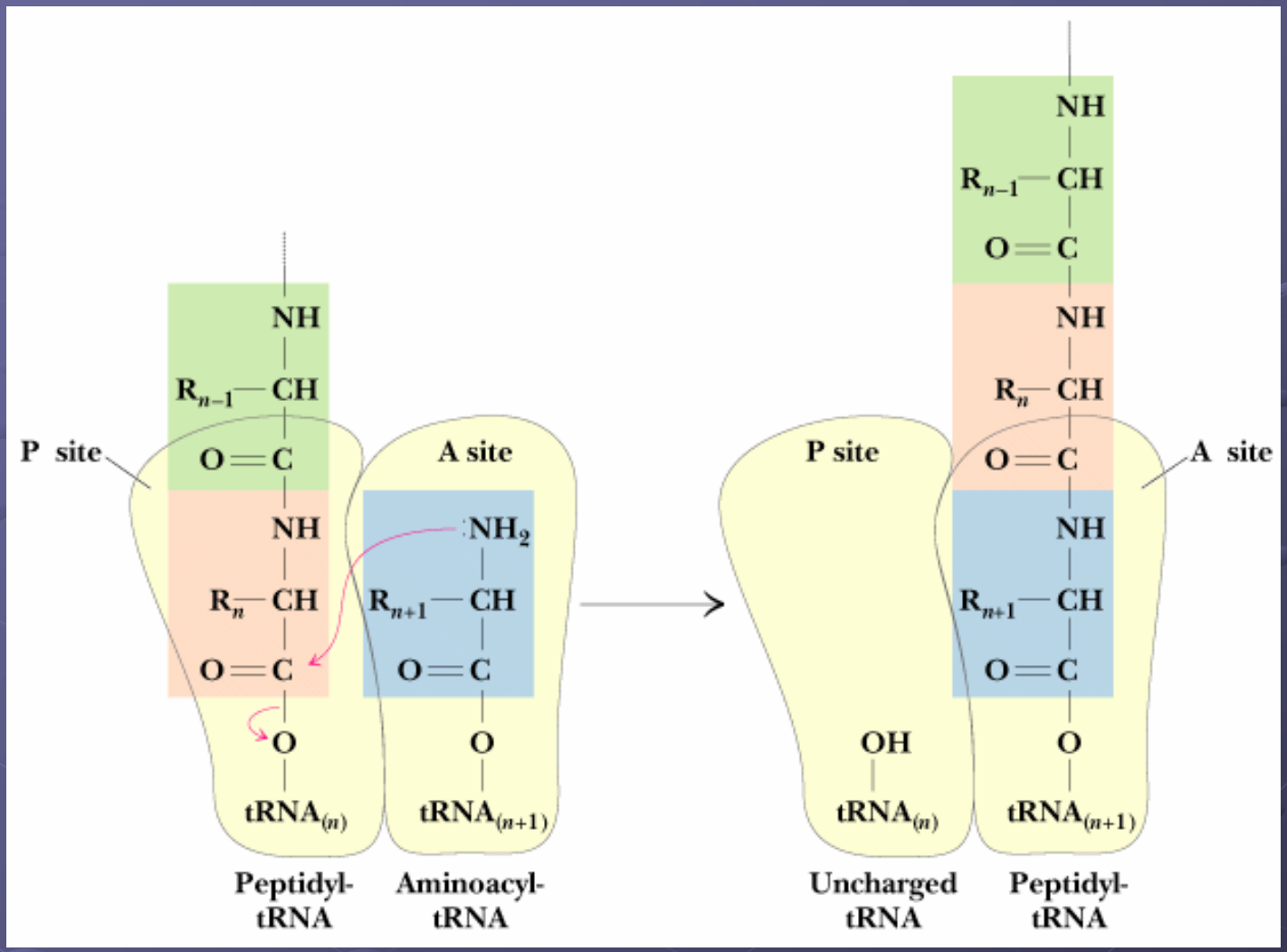

(a)
♂
AUGM
2
3
♀
AUGF
fruitless
(b)
Thyroid
A
3
4
5
6
Neurons
A
Calcitonin/CGRP
(c)
Embryonic brain
VASE
7
8
Adult brain
NCAM
(d)
Binds KGF
7 (IIIa)
IIIb
IIIc
10
Binds FGF
FGFR-2
(e)
♂
1
2
♀
msl-2
T/BS

Translation Initiation

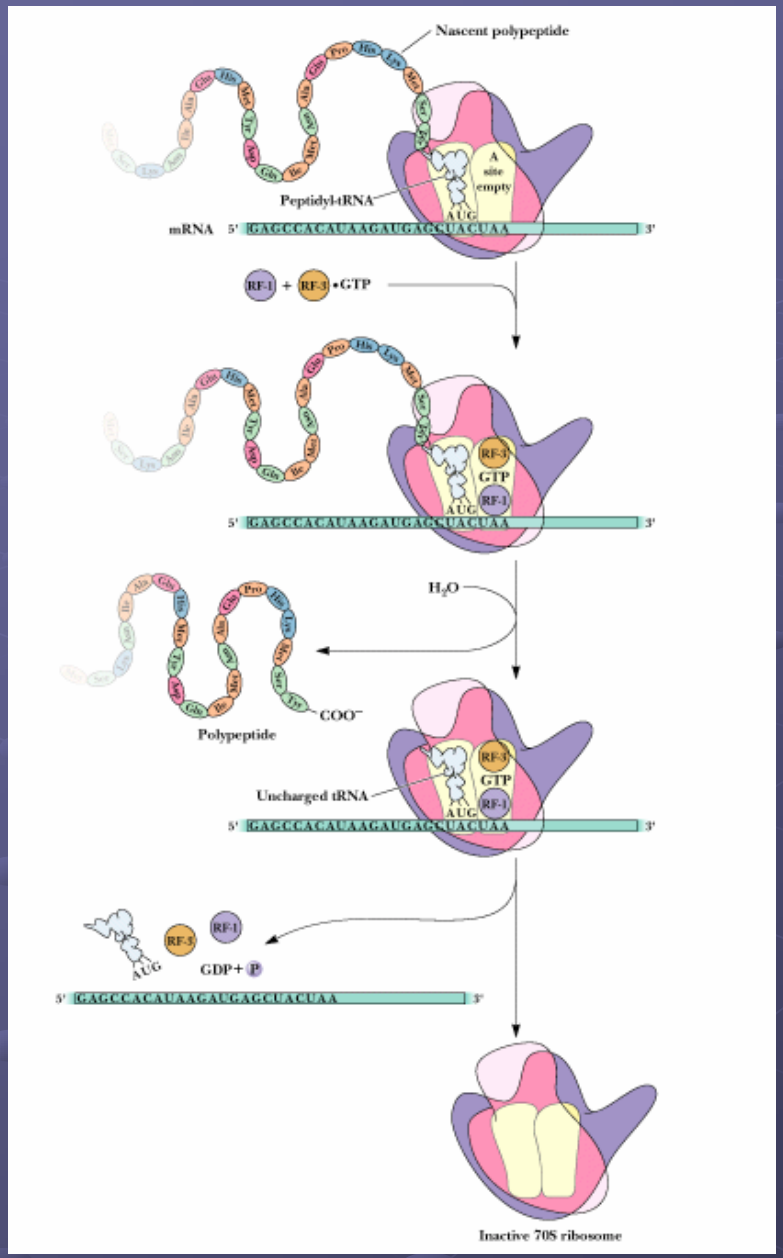


Translation Elongation

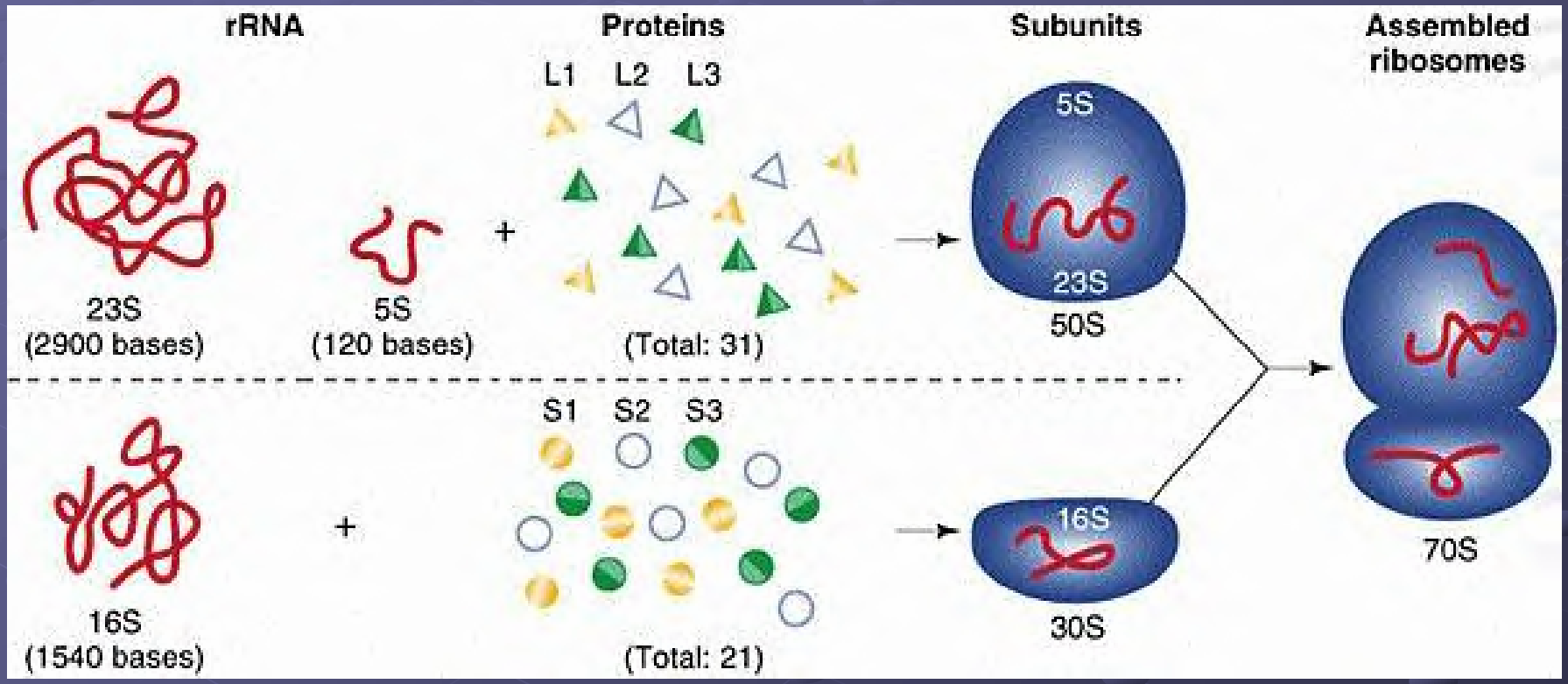


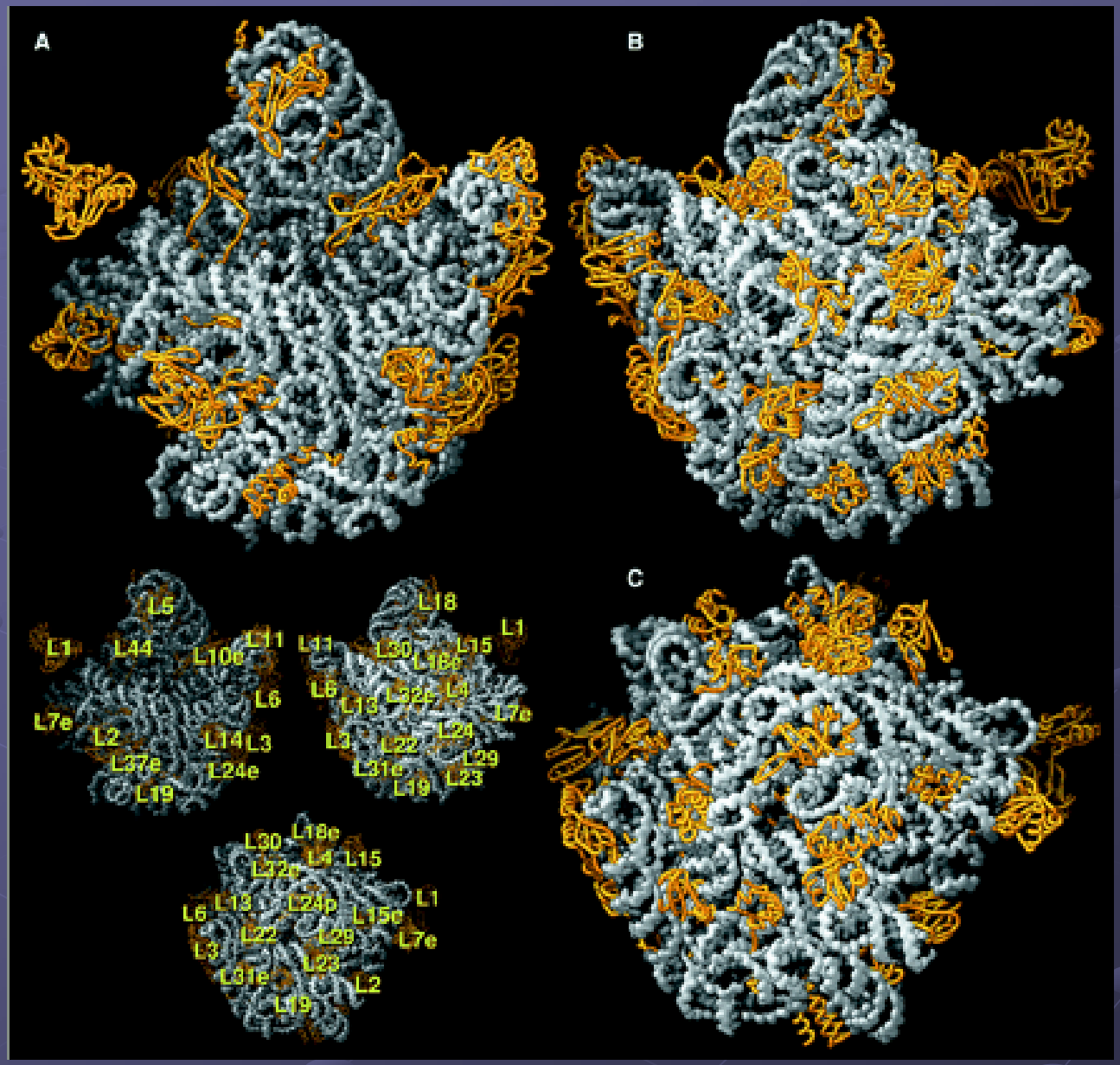


Peptide Bond Formation

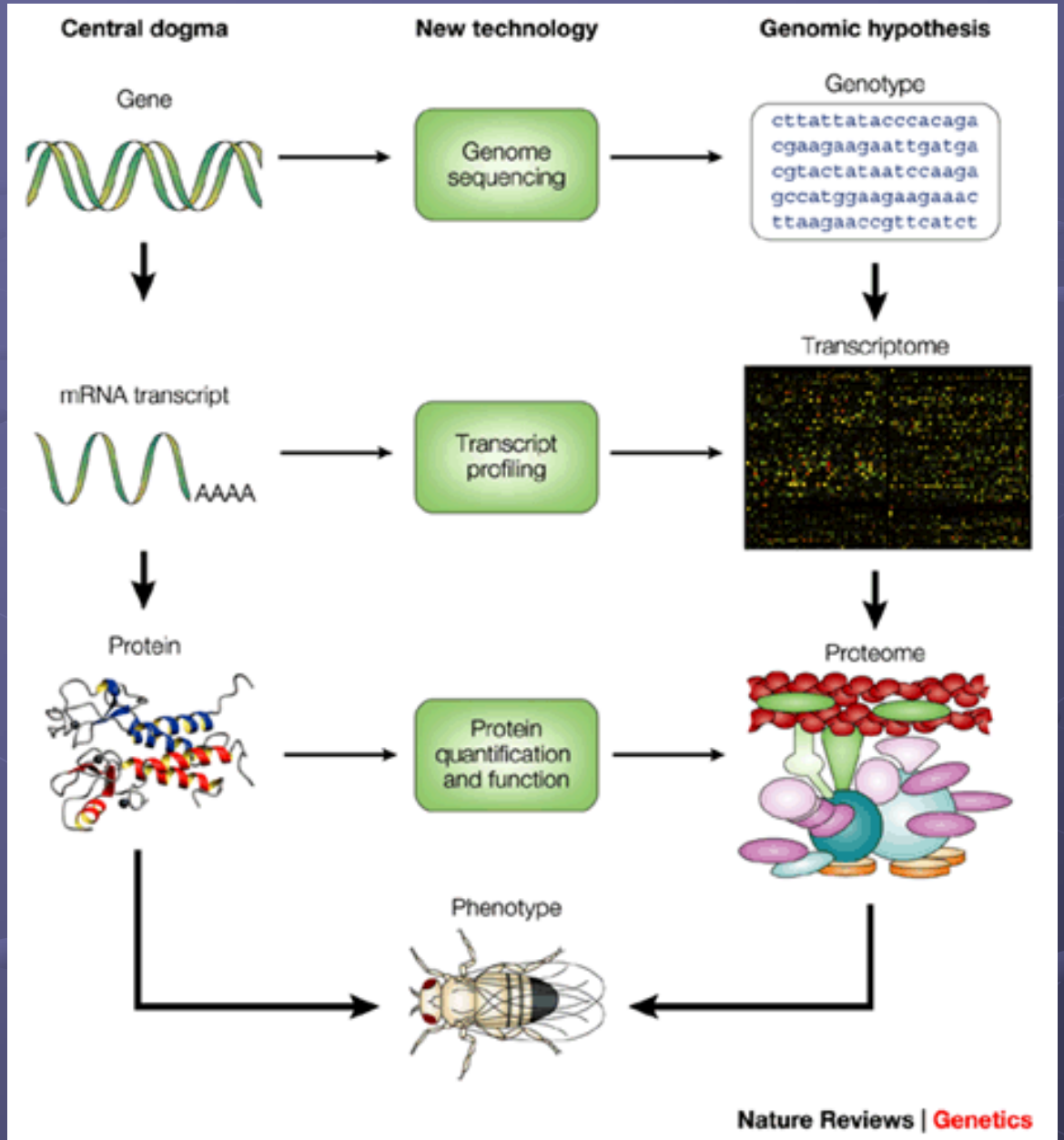


Translation: Peptide Release


rRNA
Proteins
Subunits
Assembled ribosomes
L1 L2 L3
5S
23S
50S
23S (2900 bases)
5S (120 bases)
(Total: 31)
S1 S2 S3
16S
30S
70S
16S (1540 bases)
(Total: 21)


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

Gene expression profiling



- Terminally differentiated cells
  - have a pattern of gene expression that is unique to each cell type
  - manifested by presence and relative abundance of specific mRNAs
- Transcriptome
  - The unique pattern of gene expression within each cell type
- Microarrays
  - Used to monitor gene expression in different tissues

"This is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."

Speech given by Winston Churchill at the Lord Mayor's Luncheon, Mansion House, London, November 10, 1942.