Chapter 23
Nucleic Acids

from
Organic Chemistry

by
Robert C. Neuman, Jr.
Professor of Chemistry, emeritus
University of California, Riverside

orgchembyneuman@yahoo.com
<http://web.chem.ucsb.edu/~neuman/orgchembyneuman/>

Chapter Outline of the Book

****

I. Foundations

1. Organic Molecules and Chemical Bonding
2. Alkanes and Cycloalkanes
3. Haloalkanes, Alcohols, Ethers, and Amines
4. Stereochemistry
5. Organic Spectrometry

II. Reactions, Mechanisms, Multiple Bonds

6. Organic Reactions *(*Not yet Posted*)
7. Reactions of Haloalkanes, Alcohols, and Amines. Nucleophilic Substitution
8. Alkenes and Alkynes
9. Formation of Alkenes and Alkynes. Elimination Reactions
10. Alkenes and Alkynes. Addition Reactions
11. Free Radical Addition and Substitution Reactions

III. Conjugation, Electronic Effects, Carbonyl Groups

12. Conjugated and Aromatic Molecules
13. Carbonyl Compounds. Ketones, Aldehydes, and Carboxylic Acids
14. Substituent Effects
15. Carbonyl Compounds. Esters, Amides, and Related Molecules

IV. Carbonyl and Pericyclic Reactions and Mechanisms

16. Carbonyl Compounds. Addition and Substitution Reactions
17. Oxidation and Reduction Reactions
18. Reactions of Enolate Ions and Enols
19. Cyclization and Pericyclic Reactions *(*Not yet Posted*)

V. Bioorganic Compounds

20. Carbohydrates
21. Lipids
22. Peptides, Proteins, and α–Amino Acids
23. Nucleic Acids

****

**Note: Chapters marked with an (*) are not yet posted.*

23: Nucleic Acids

23: Nucleic Acids

- *Structures of Nucleic Acids*
- *Replication, Transcription, and Translation*
- *Nucleotide Biosynthesis and Degradation*

Preview

Nucleic acids (DNA and RNA) perform a variety of crucial functions in organisms. DNA stores and transfers genetic information, it serves as the template for the synthesis of new DNA and RNAs, while RNAs carry out protein synthesis. Nucleic acids contain only a few different components, but they have great structural diversity. This diversity results from the many possible combinations of those few components due to the large sizes of DNA and RNA. We will see that our study of nucleic acids brings together information from our earlier studies of *carbohydrates* (Chapter 20) as well as *amino acids* and *proteins* (Chapter 22).

23.1 Structures of Nucleic Acids

The two classes of nucleic acids are **DNA** (**deoxyribonucleic acid**) and **RNA** (**ribonucleic acid**). While they have significantly different structures, we can describe both DNA and RNA as **polynucleotides** (polymers of **nucleotides**).

Nucleotides and Nucleosides (23.1A)

Each *nucleotide* subunit of a *nucleic acid* contains a *phosphate* group, a *sugar* component, and a *heterocyclic* ring system (**heterocyclic base**) (Figure 23.01). The portion of the *nucleotide* containing just the sugar and heterocyclic base is called a **nucleoside**.

Figure 23.01

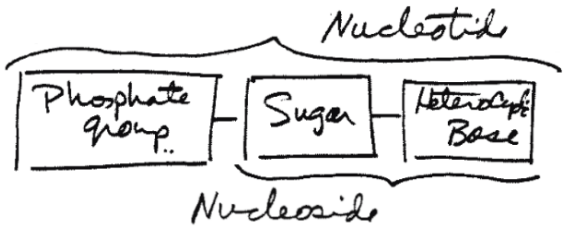
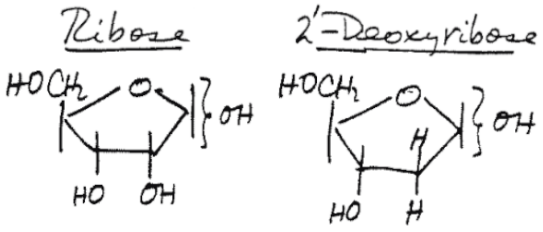


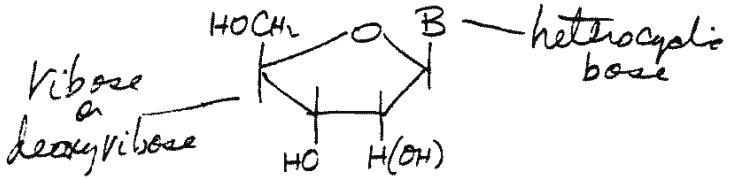
Figure 23.02



The Sugar. The *sugar* component of RNA nucleotides (or nucleosides) is *ribose*, while that of DNA nucleotides (or nucleosides) is *2'-deoxyribose* (no OH on C2') (Chapter 20) (Figure 23.02). The ribose and 2'-deoxyribose units exist as *furanose* forms (Chapter 20) in both RNA and DNA.

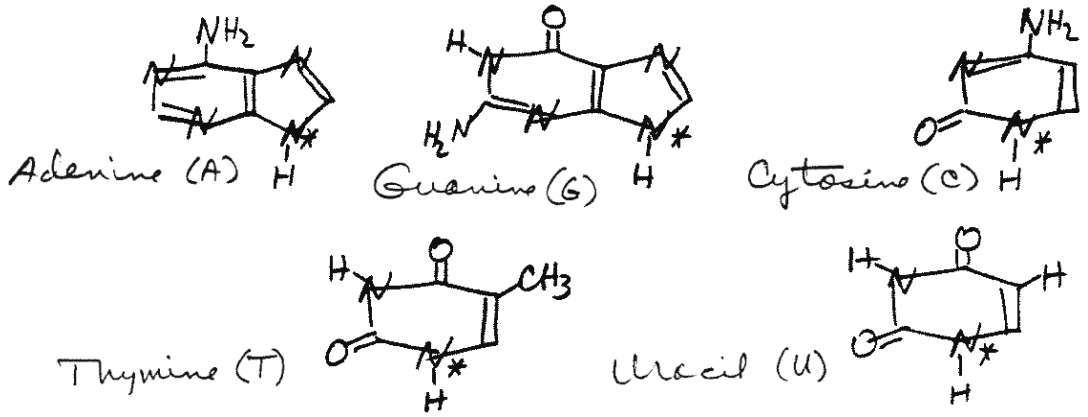
The Heterocyclic Bases. Each *heterocyclic base* (abbreviated **B**) bonds to the anomeric carbon (C1') of the *ribose* or *deoxyribose* ring with a β-C-N-glycosidic bond (Chapter 20).

Figure 23.03



The four heterocyclic bases in *DNA* nucleotides (or nucleosides) are **adenine** (**A**), **guanine** (**G**), **cytosine** (**C**), and **thymine** (**T**).

Figure 23.04



Each bonds to the C1' of *deoxyribose* at N* as shown below for *adenine* (Figure 23.05). The heterocyclic bases in *RNA* nucleotides (or nucleosides) similarly bond to *ribose*. They include *A*, *G*, and *C*, but **uracil** (**U**) replaces *thymine* (*T*). *U* is structurally similar to *T* except that the C5-CH_3 group of *T* is absent in *U*. RNA molecules can have other heterocyclic bases in addition to *A*, *G*, *C*, and *U*. *Adenine* (*A*) and *guanine* (*G*) are **purines** because they have the same ring skeleton as *purine*. *Cytosine* (*C*), *thymine* (*T*), and *uracil* (*U*), are **pyrimidines** because they have the ring skeleton of *pyrimidine*.

Figure 23.05

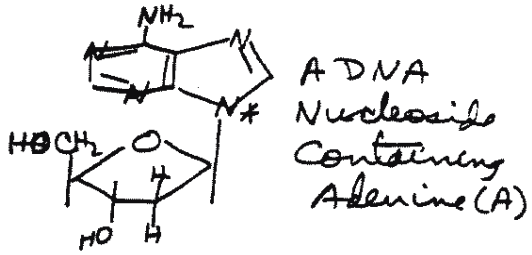
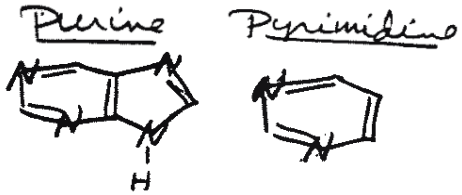
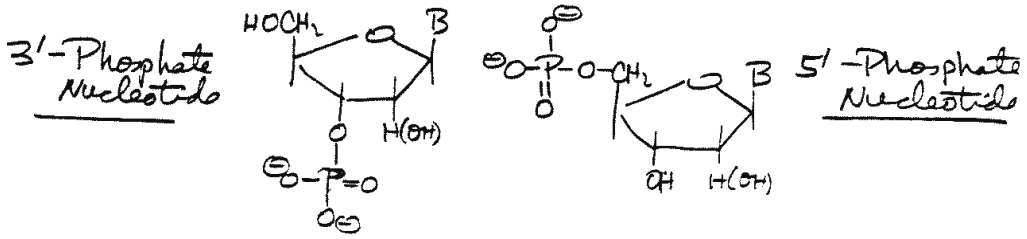


Figure 23.06



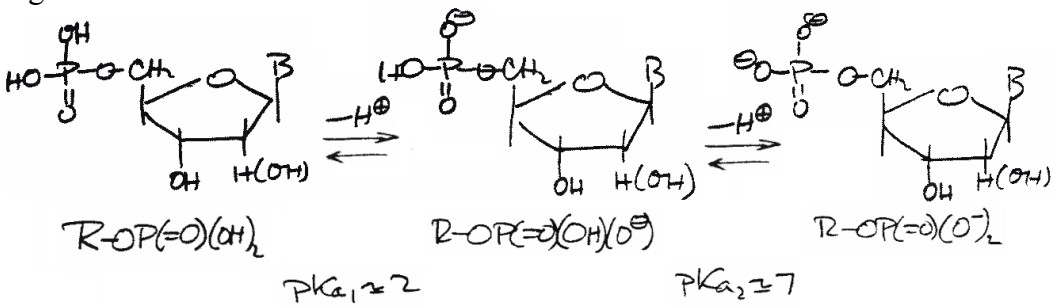
The Phosphate Groups. The phosphate groups of nucleotides bond to C3' or C5' of the ribose or deoxyribose rings (Figure 23.07) [next page]. We will see that this is a consequence of the way nucleotides join as polynucleotides in DNA or RNA.

Figure 23.07



We can represent a nucleotide as R-OP(=O)$(O^-)_2$ (or R-OPO_3^{-2}) where R is a *nucleoside* (*sugar-base*). The phosphate groups are anions at physiological pH because their fully protonated forms are diprotic acids (R-OP(=O)$(OH)_2$ or R-OPO_3H_2) with $pK_{a1} \approx 2$ and $pK_{a2} \approx 7$.

Figure 23.08



Nucleotide and Nucleoside Nomenclature. Each *nucleoside* has a single name, while each *nucleotide* has two names (Table 23.1). The prefix *deoxy* indicates that *deoxyribose* replaces *ribose*, and the numbers *3′* or *5′* show where the phosphate attaches to the sugar ring.

Table 23.1. Names of Nucleosides and Nucleotides.

Base	Nucleoside	Nucleotide
Purines		
Adenine (A)	Adenosine	Adenosine 3'(or 5')-phosphate 3'(or 5')-Adenylic acid
	Deoxyadenosine	Deoxyadenosine 3'(or 5')-phosphate 3'(or 5')-Deoxyadenylic acid
Guanine (G)	Guanosine	Guanosine 3'(or 5')-phosphate 3'(or 5')-Guanylic acid
	Deoxyguanosine	Deoxyguanosine 3'(or 5')-phosphate 3'(or 5')-Deoxyguanylic acid
Pyrimidines		
Cytosine (C)	Cytidine	Cytidine 3'(or 5')-phosphate 3'(or 5')-Cytidylic acid
	Deoxycytidine	Deoxycytidine 3'(or 5')-phosphate 3'(or 5')-Deoxyytidylic acid
Uracil (U)	Uridine	Uridine 3'(or 5')-phosphate 3'(or 5')-Uridylic acid
Thymine (T)	Deoxythymidine	Deoxythymidine 3'(or 5')-phosphate 3'(or 5')-Deoxythymidylic acid

Polynucleotide Structure (23.1B)

DNA and RNA have **sugar-phosphate backbones** (Figures 23.09 and 23.10).

The Sugar-Phosphate Backbone. We can view the polynucleotide strands of DNA or RNA as many *nucleosides* linked by phosphate groups (P) at the 3' and 5' carbons of the sugar furanoside rings (S) (Figure 23.09). As a result, RNA and DNA have *sugar-phosphate backbones* with heterocyclic bases (*B*) attached to the anomeric C (C1') of each sugar ring. Each polynucleotide strand has a 5' end (the terminal phosphate attached to C5' of a terminal nucleoside) and a 3' end (a terminal phosphate attached to C3' of the other terminal nucleoside) (Figure 23.10). Phosphate groups in the sugar-phosphate backbone of nucleic acids fully ionize at physiological pH ($(RO)_2P(=O)O^-$) since their protonated forms are monoprotic acids $(RO)_2P(=O)OH$ with a $pK_a \approx 2$ (Figure 23.11) [next page]. This is why DNA and RNA are called nucleic *acids*.

Figure 23.10 (and 23.09)

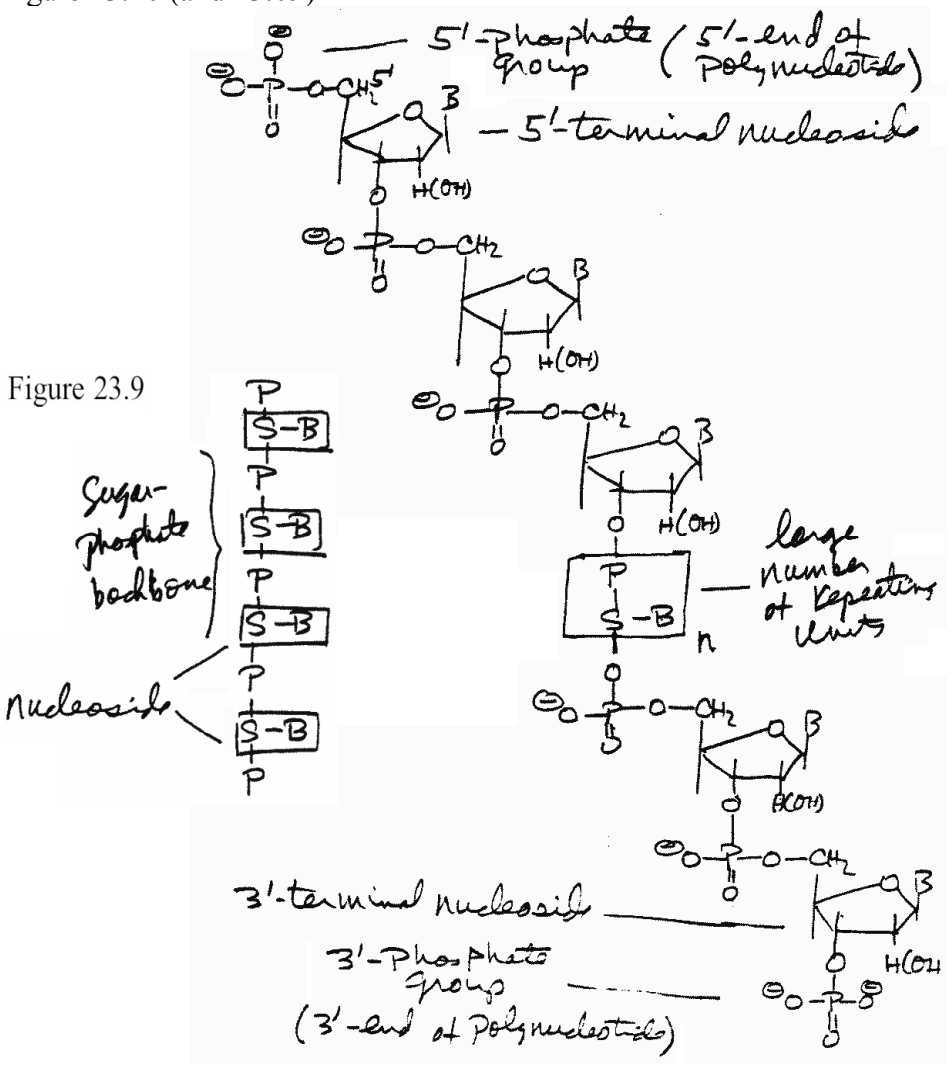
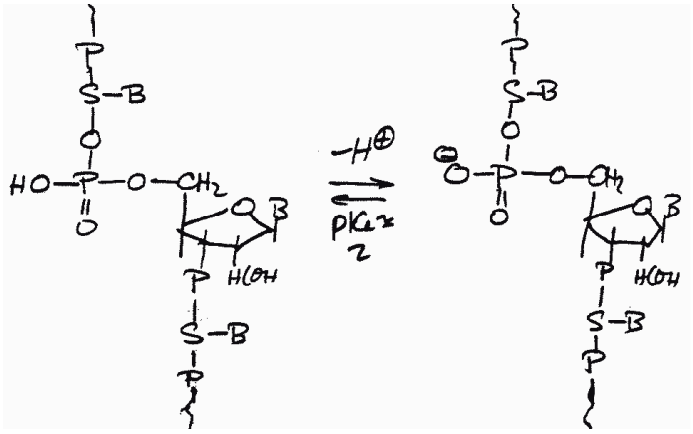


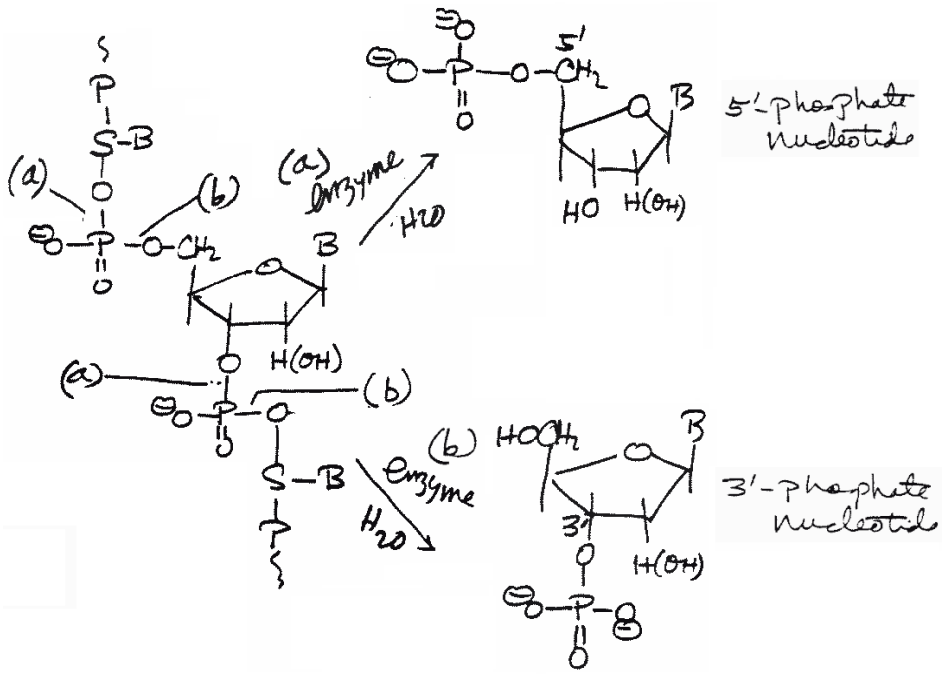
Figure 23.9

Figure 23.11



Hydrolysis of Polynucleotides. Polynucleotides cleave into individual nucleotides during enzymatic hydrolysis (Figure 23.12). Enzymes cleave 3' or 5' CO-P bonds resulting in the formation of 5' or 3'-phosphate nucleotides, respectively.

Figure 23.12

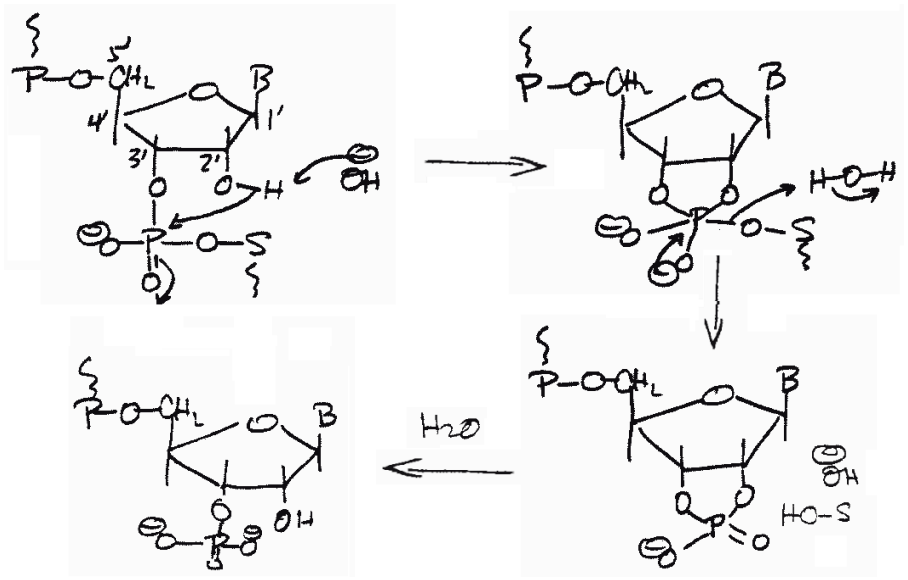


Polynucleotides of DNA are more stable to basic hydrolysis than those of RNA. DNA nucleotides have deoxyribose (no OH on C2') in their sugar-phosphate backbone precluding intramolecular participation of the C2'-OH that occurs during basic hydrolysis of RNA (Figure 23.13) [next page].

Comparative Structures of DNA and RNA (23.1C)

With the exception of nucleic acids in some viruses, DNAs contain two intertwining polynucleotide strands while RNAs have only a single polynucleotide strand.

Figure 23.13

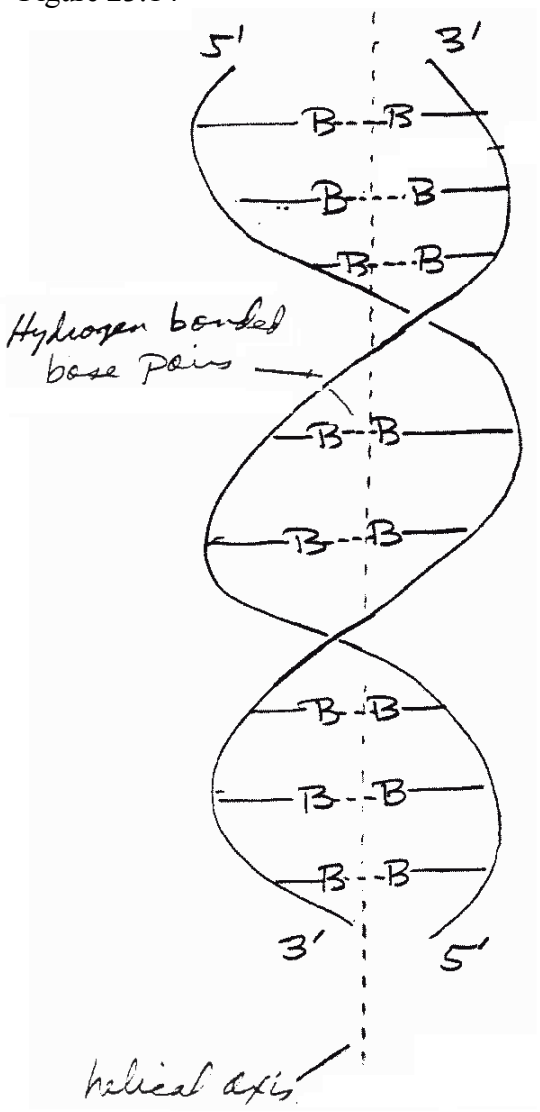


The DNA Double Helix. DNA consists of two α-helical polynucleotide strands intertwined to form a **double helix** (Figure 23.14) [next page]. The two strands run in opposite directions (3'→5' and 5'→3') so we describe them as **antiparallel**. We will see that heterocyclic bases (B) on one strand form hydrogen bonds with those on the other strand. The resultant hydrogen bonded **base pairs** stack above and below each other like the steps of a ladder. The most common form of DNA is **B-DNA** where the strands are right-handed α-helices and the helical axis passes directly through the center of the base pairs. DNA molecules are not straight rods, but bend, loop, and coil.

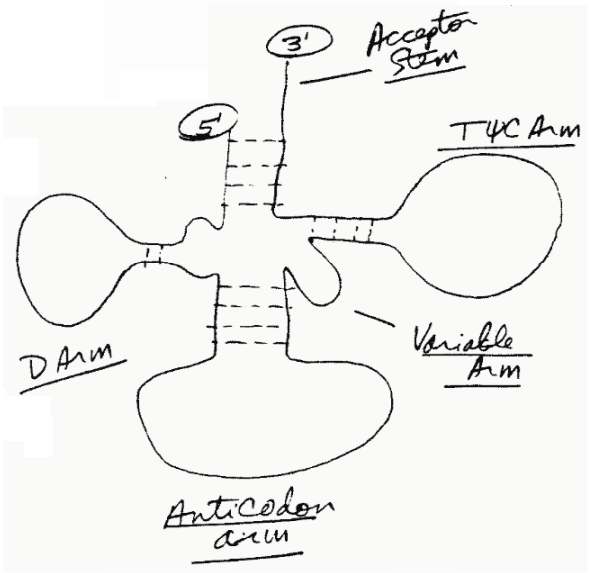
> **Other Forms of DNA**. **A-DNA** is a form of DNA where the base pairs tip with respect to the helical axis and the axis does not pass through the base pairs. *B-DNA* reversibly transforms into *A-DNA* by a change in the moisture content of the atmosphere around the DNA. **Z-DNA** has a left-handed double helix. *B-DNA* is the most prevalent of the three forms.

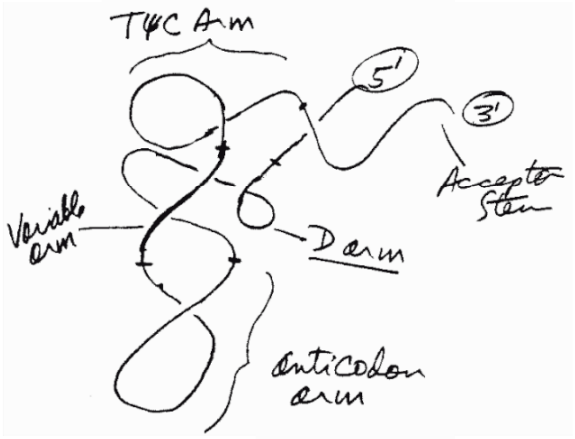
RNA Polynucleotides. RNA molecules are single stranded and there are several different types including **ribosomal RNA** (**rRNA**), **messenger RNA** (**mRNA**), and **transfer RNA** (**tRNA**). *tRNA* molecules are relatively small and structurally well characterized. They have three **arms**, and a **stem** that includes both the 3' and 5' ends of the polynucleotide strand (Figure 23.15) [next page]. Regions in each *arm* have hydrogen bonds between heterocyclic bases on the same polynucleotide strand. The shape (*3° structure*) of tRNAs is like an "*L*" (Figure 23.16) [next page]. We show regions corresponding to the *stem* and the various *arms* or "**loops**" for comparison with the representation in Figure 23.15. We will consider these RNAs again when we discuss protein synthesis later in the chapter.

Figure 23.14



Figures 23.15 and 23.16





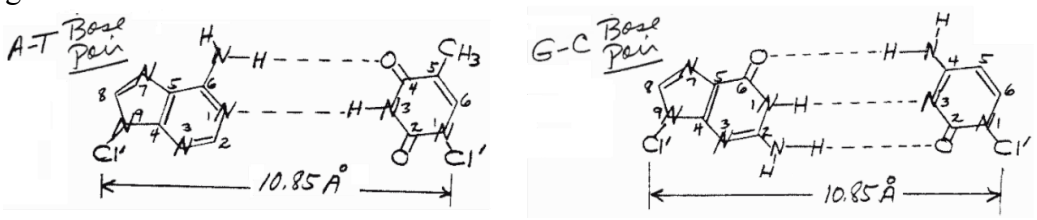
Sizes of DNA and RNA. *DNA* molecules are very large ranging from polynucleotide strands of 5,000 to 300,000 nucleotides in viruses, more than 4,500,000 nucleotides in some bacteria, and 2,900,000,000 nucleotides in humans. The extended length of those strands can be as much as 0.2 mm in viruses, more than 1.5 mm in bacteria, and almost 1 m (100 cm) in humans. The size of *RNA* molecules depends on the type. *tRNA* molecules range from 60 to 95 nucleotides, some *rRNA* molecules in *E. coli* have polynucleotide strands of ~ 120, ~1500, and ~2900 nucleotide units, while *mRNA* ranges from hundreds to thousands of nucleotides.

Base Pairing (23.1D)

Hydrogen bonding (*base pairing*) between heterocyclic bases is very selective in DNA, but less so in RNA.

DNA. Base pairs in DNA are either *A-T* (*adenine-thymine*), or *G-C* (*guanine-cytosine*). These A-T and G-C pairs are called "**Watson-Crick**" base pairs after the British chemists James Watson and Francis Crick who described the structure of DNA in 1953 and subsequently received the 1960 Nobel Prize in Medicine for this achievement.

Figure 23.17



In A-T or G-C, one base is a *purine* while the other is a *pyrimidine*. As a result, the theoretical distance between anomeric C's (C1') of their sugars is the same for both pairs and polynucleotide strands of DNA can have uniform separation.

Hydrogen bond donors (N-H groups) and acceptors (O= or -N= atoms) line up well in A-T and G-C, but you can also draw hydrogen bonds between *A* and *C*, and between *G* and *T*. However, neither of these alternate purine-pyrimidine pairs fits together as well as *A-T* and *G-C* and neither has 3 hydrogen bonds to *G* or *C* that exist in a *G-C* base pair. (We sometimes show the number of hydrogen bonds in these base pairs by writing them as *A=T* or *G≡C*.)

Heterocyclic Base Association Constants. Experimentally measured association constants (K) for formation of hydrogen bonded pairs of free heterocyclic bases suggest that *A-T* and *G-C* are the most thermodynamically favorable base pairs in DNA. Approximate values of K for A, G, C, and U (Table 23.2) are greatest for A-U (or U-A) (analogous to A-T), and G-C (or C-G). K's were determined for U rather than T, but values for *T* should be comparable to those for *U* in spite of the C5-CH_3 group in *T*.

Table 23.2. Approximate Association Constants (K, M^{-1}) for Hydrogen Bonded Dimer Formation from Free Heterocyclic Bases.

Base	A	G	C	U
A	3	-	-	**100**
G	-	5,000	**50,000**	-
C	-	**50,000**	28	-
U	**100**	-	-	6

RNA. RNAs have only one polynucleotide strand, but folds in the strand allow hydrogen bonding between some of the bases. These base pairs are usually *G-C* or *A-U* (equivalent to *A-T*), but other pairs are present since RNAs contain a number of modified heterocyclic bases (Figure 23.18)[next page]. *A* and *U* also form an energetically favorable **Hoogsteen** base pair in some RNAs (Figure 23.19)[next page].

Figure 23.18

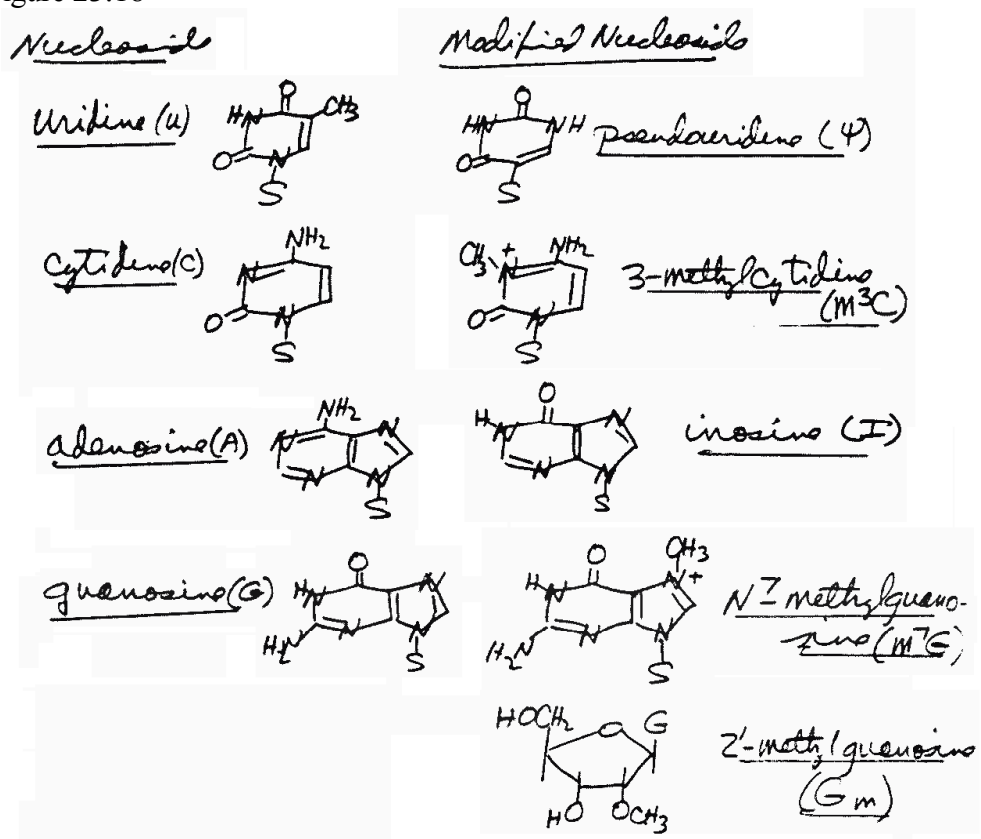


Figure 23.19

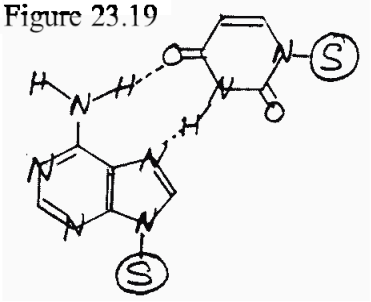


Figure 23.21

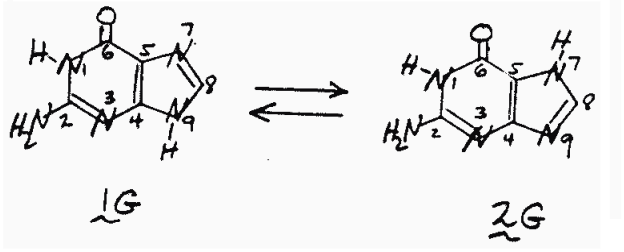
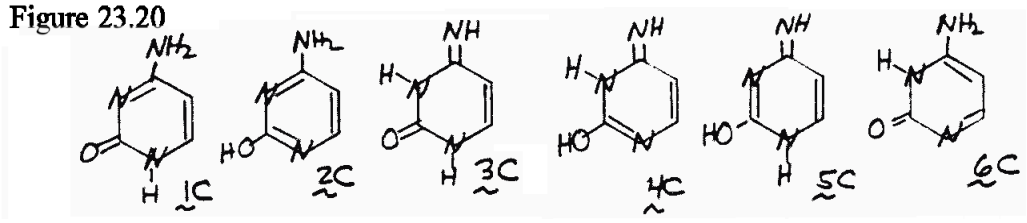


Figure 23.20



Tautomers of Heterocyclic Bases. We can write a number of different tautomeric forms for each heterocyclic base such as the 6 tautomeric structures shown here for *cytosine* (*C*) (Figure 23.20) [above]. Structure **1C** that we have shown throughout the chapter is also the most stable form for free *C* in aqueous solution. The tautomeric forms that we have shown

for *A*, and for *T* (or *U*) in nucleosides are also the most favorable tautomers of the free bases in aqueous solution, but *G* has two relatively favorable tautomers (Figure 23.21) [previous page]. Structure **1G** is the form found in nucleic acids since *G* bonds to the anomeric carbon of ribose or deoxyribose at its N9 nitrogen.

Forces that Influence Nucleic Acid Structures (23.1E)
The same forces that determine protein structure (Chapter 22) influence nucleic acid structures. They include *hydrogen bonding*, *hydrophobic bonding*, and *ionic interactions*.

Hydrogen Bonding. The order of bases on each strand of DNA must be complementary so that each base pair is *A*-*T* or *G*-*C*. However, the energy of the hydrogen bonds in these base pairs is no greater than that which we expect for hydrogen bonding of these bases to water. For this reason, base pairing does not appear to be the primary force stabilizing the DNA double helix.

Hydrophobic Bonding. As with proteins (Chapter 22), hydrophobic interactions provide the major stabilizing force for nucleic acids. These hydrophobic interactions occur between bases stacked above and below each other in the double helix. The facts that heterocyclic bases stack with each other in single strands of RNA, and when they are free in aqueous solution, demonstrate the energetic preference for base stacking.

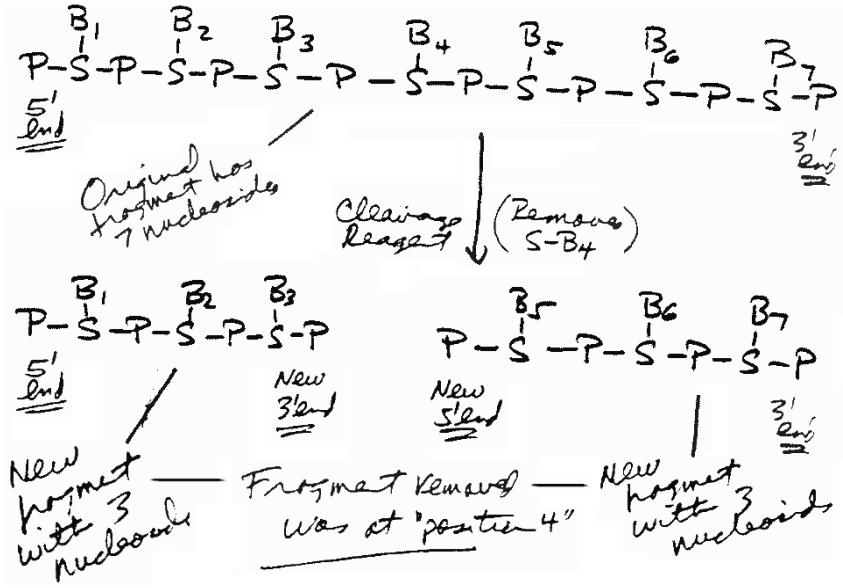
Ionic Interactions. Electrostatic repulsion between negatively charged phosphates in the sugar-phosphate backbone destabilizes all structures of nucleic acids with strands in close proximity. Association of the phosphate groups with cations such as Mg^{+2} diminishes these repulsive forces.

Sequencing Nucleic Acids (23.1F)
A knowledge of the *sequence* of nucleotides in nucleic acids is crucial to understanding their function in organisms.

Sequencing Strategy. Biochemists use the same general strategy for sequencing nucleic acids that they use for proteins (Chapter 22). Fragment sequences provide the information that permits assembly of the sequence of the full polynucleotide. There are a number of different polynucleotide sequencing methods including **chemical sequencing** that we describe here. While biochemists now primarily use other methods, *chemical sequencing* is historically important and its organic reactions are particularly relevant to our studies of organic chemistry.

Chemical Sequencing. A **cleavage reagent** removes a nucleoside and cleaves a polynucleotide into two new fragment strands (Figure 23.22). One fragment has a new 3'-phosphate end while the other has a new 5'-phosphate end. We can identify the position in the original strand of the nucleoside that the reagent removed by determining the number of nucleosides in either of these new fragments. *Cleavage reagents* selectively remove nucleosides with specific heterocyclic bases so they also identify the heterocyclic base on the nucleoside at that position.

Figure 23.22

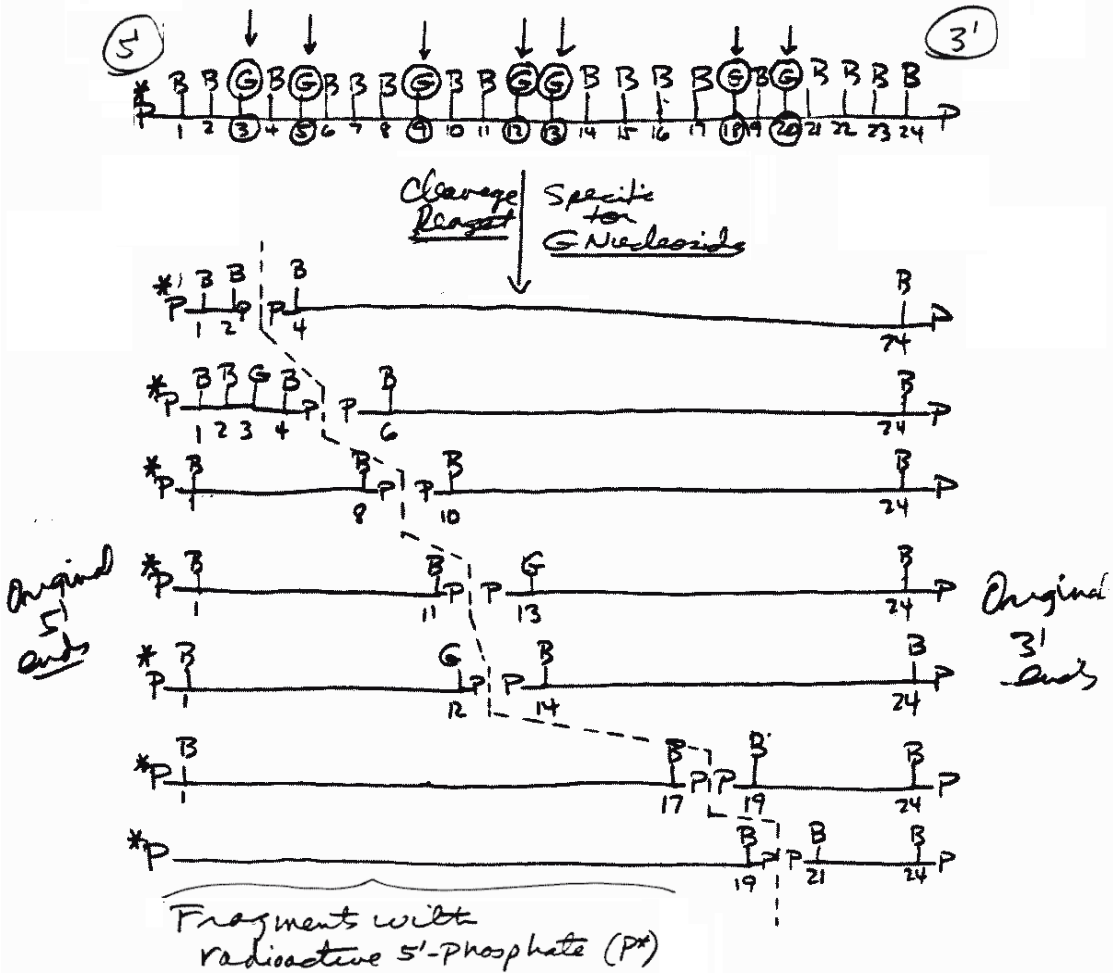


The four available cleavage reagents remove *G* nucleosides, *A* and *G* nucleosides, *C* nucleosides, or *C* and *T* nucleosides. If we use the reagent that randomly removes one *sugar-G* from each strand in the sample, we obtain a mixture of fragments that originally had a *sugar-G* at their new 3' and new 5' ends (Figure 23.23) [next page]. Biochemists label the original strands with radioactive phosphate at their 5' ends before chemical cleavage. As a result, cleavage fragments with new *3'-phosphate ends* have the 5' radioactive phosphate label, while fragments with new *5'-phosphate ends* have no radioactive label. By determining the number of nucleosides in fragments with radioactive phosphate, we establish positions of *sugar-G* in the original polynucleotide strand with respect to its original 5'-phosphate end.

Analysis of Cleavage Fragments. Reaction mixtures arising fromuse of the four cleavage reagents are simultaneously separated using **gel electrophoresis**. This technique is described in biochemistry text books. Cleavage fragments migrate toward the positive (+) electrode according to their size (number of nucleotides) with smallest fragments migrating fastest. An **autoradiograph** images the radioactive fragments and their relative positions reflect their sizes. A comparison of the data from all four cleavage

reactions on the same *autoradiograph*, we establish the positions of each *A*, *T*, *G*, and *C* nucleoside with respect to the 5'-phosphate end of the original polynucleotide strand.

Figure 23.23

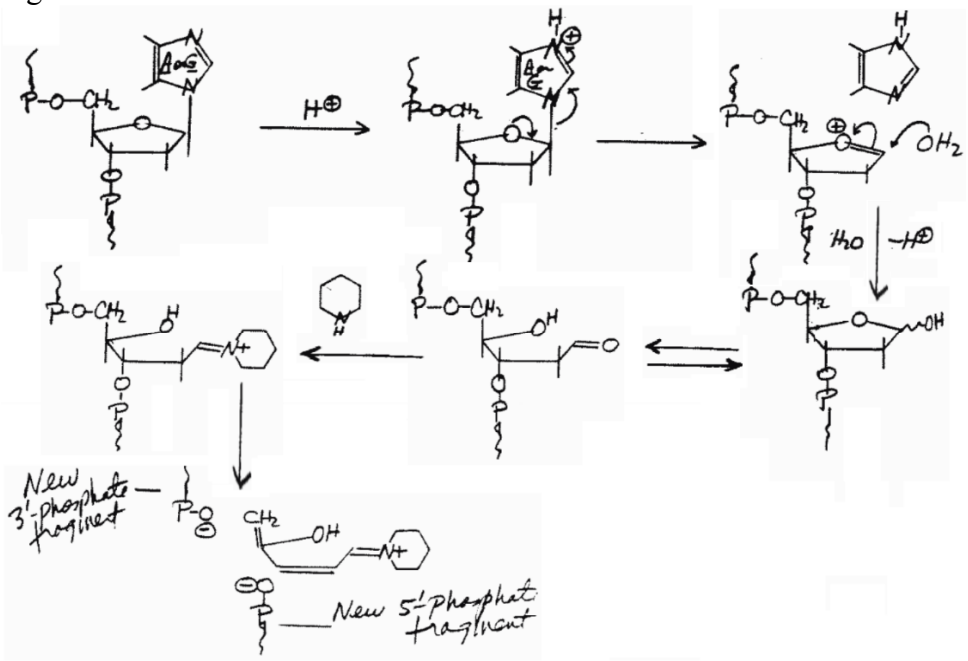


Chemical Cleavage Reagents and their Reactions (23.1G)

The cleavage reagents delete specific nucleosides from a polynucleotide by first reacting with the heterocyclic base and then its sugar component.

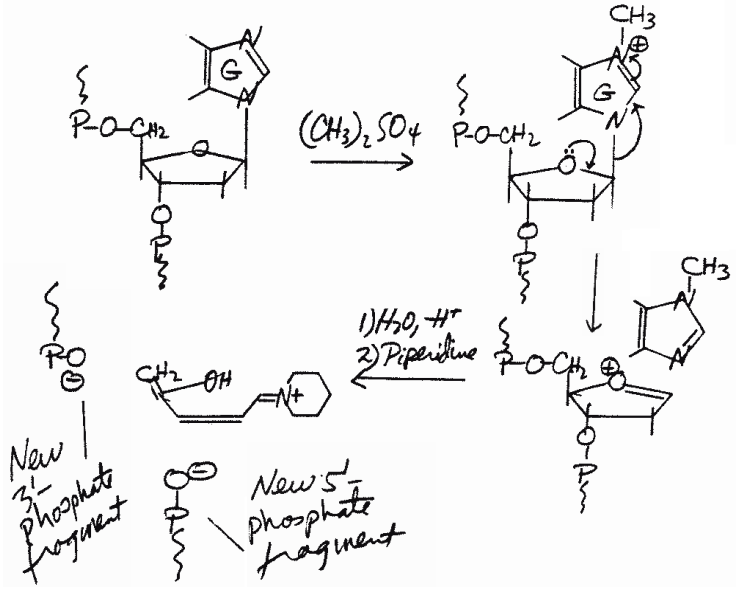
A and G Nucleosides. We can delete *A* and *G* nucleosides from the polynucleotide by treating it with acid and then piperidine. Acid protonates the *purines A* and *G* on N7 making them good leaving groups from the anomeric C of their sugar rings (Figure 23.25) [next page]. (*Note - there is no Figure 23.24*) Water adds to the resulting cyclic oxonium ions (Chapter 20) giving furanose units still bonded to the *sugar-phosphate* backbone. Piperidine reacts with their aldose forms cleaving their phosphate bonds and releasing the two new polynucleotide fragments.

Figure 23.25



We identify *G* nucleosides by treating the polynucleotide with *dimethylsulfate* that methylates N7 of *G* (Figure 23.26). The resulting positively charged heterocyclic ring is a good leaving group, and hydrolysis followed by treatment with piperidine leads to loss of the methylated *sugar-G* nucleoside and cleavage as described above.

Figure 23.26



Dimethylsulfate also methylates *A*, but at N3 rather than N7. N3 methylated *A* is a relatively poor leaving group, so fragments from loss of *G* are more intense in the autoradiograph and we can distinguish them from those due to loss of *A* .

C and T Nucleosides. Hydrazine reacts with *C* and *T* nucleosides releasing a 5-membered heterocycle and forming an imine of their sugar component (Figures 23.27 and 23.28). Subsequent treatment with piperidine gives the two polynucleotide fragments. Treatment of the polynucleotide with hydrazine in 1 to 2 M NaCl removes only *C* nucleosides.

Figure 23.27

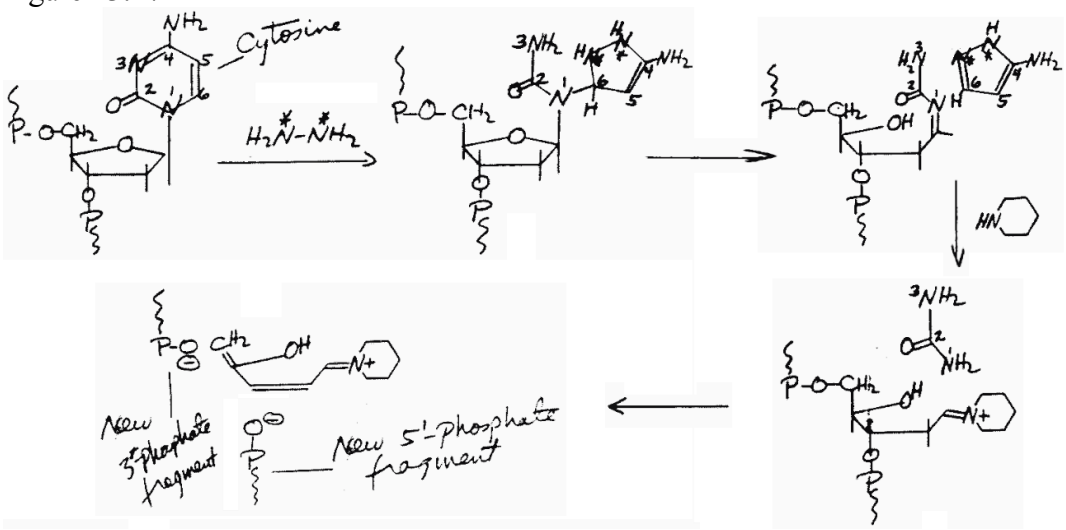
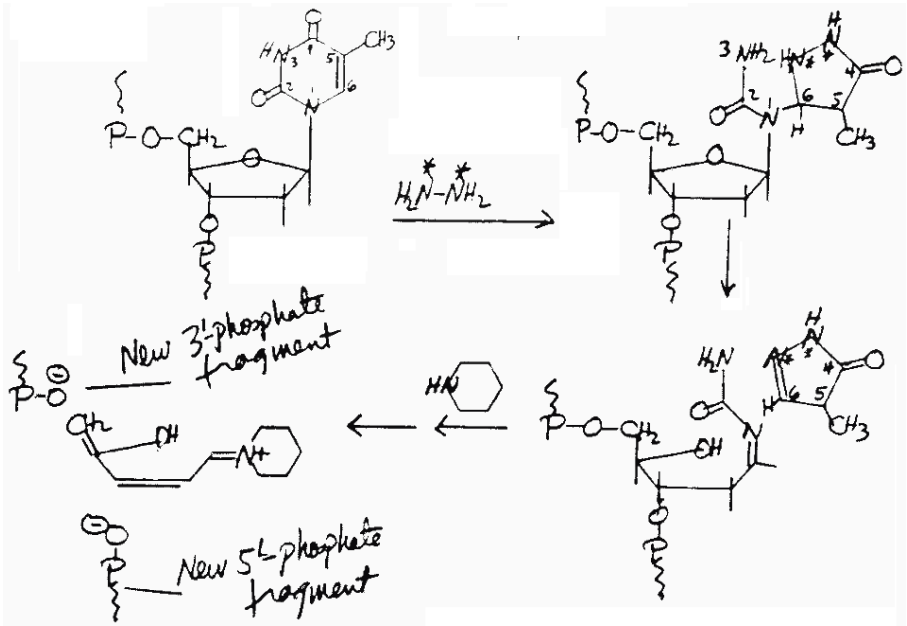


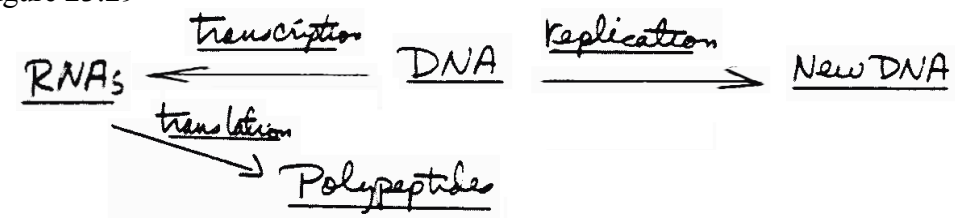
Figure 23.28



23.2 Replication, Transcription, and Translation

New DNA forms by **replication**. DNA is the template for the synthesis of RNA by **transcription**. RNA participates in synthesis of proteins from amino acids during **translation** (Figure 23.29) [next page].

Figure 23.29

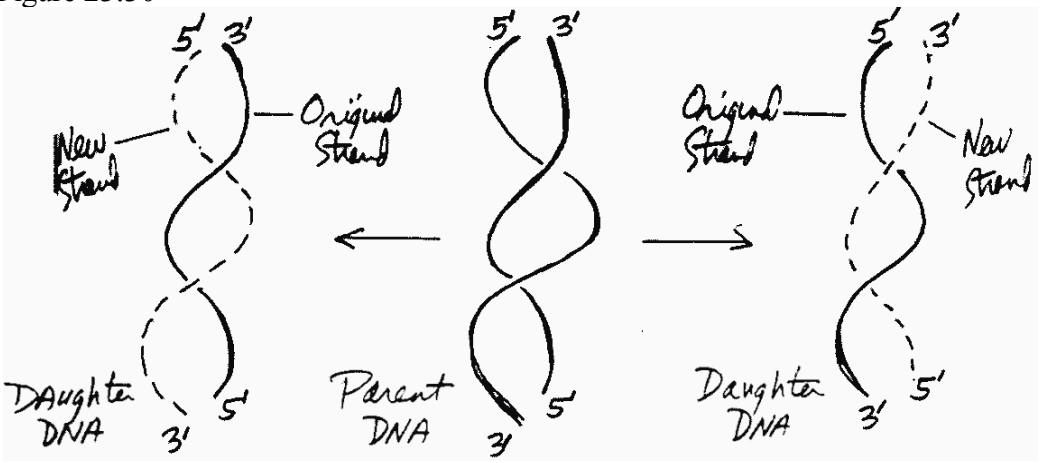


Replication (23.2A)

A double stranded DNA molecule becomes two identical double stranded DNA molecules during *replication*.

Replication is Semiconservative. We describe replication as **semiconservative** because each of the two new DNA molecules contains one strand of the original DNA molecule and one new strand.

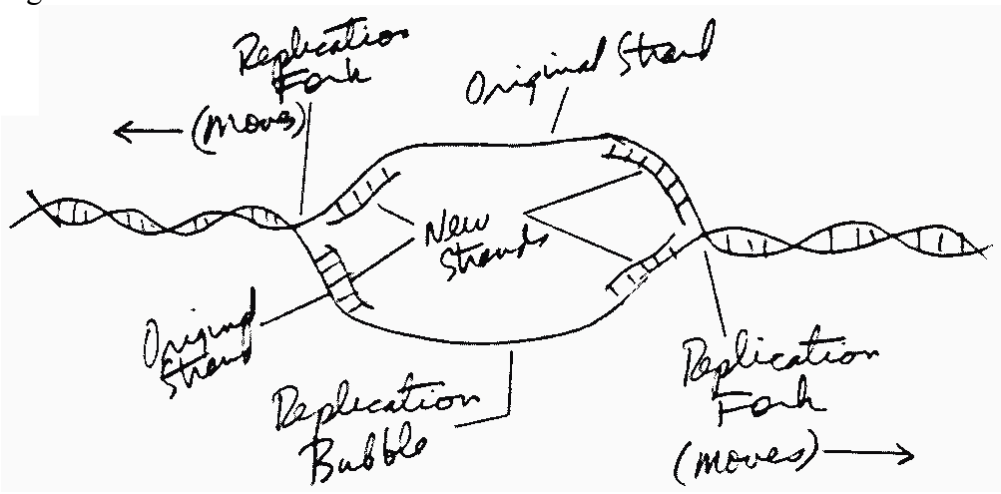
Figure 23.30



The original DNA molecule contains a 5'→3' strand and a complementary 3'→5' strand. One of the new DNA molecules contains the 5'→3' strand of its parent and a new complementary 3'→5' strand assembled from nucleotides during replication, while the other new DNA molecule contains the 3'→5' strand of its parent and a new 5'→3' strand. The new DNA strands of each **daughter** DNA develop within a **replication bubble** on the parent DNA molecule that disrupts hydrogen bonding between base pairs (Figure 23.31) [next page]. The two ends of the bubble are **forks** and new complementary strands assemble on both parent strands at both forks.

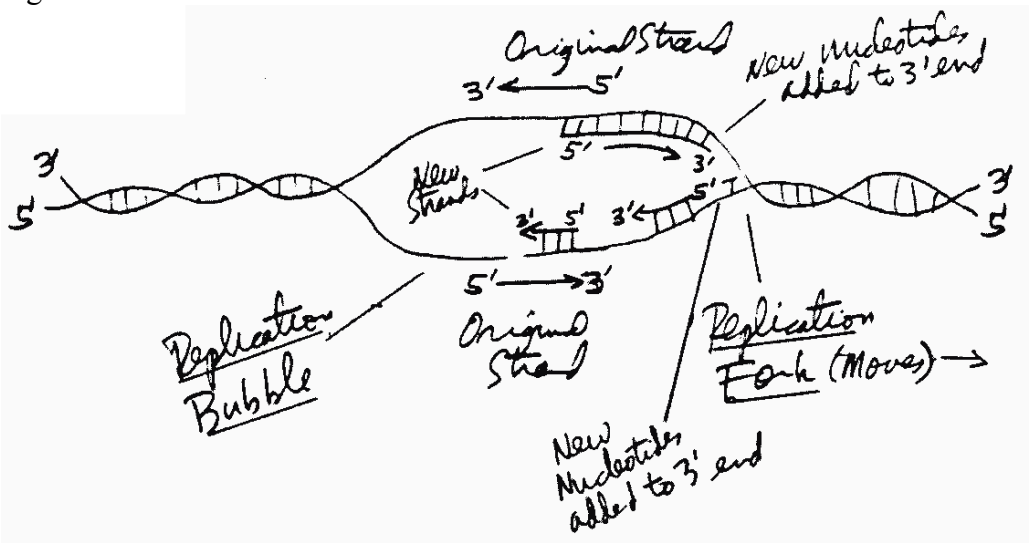
Replication Occurs 5'→3'. During replication, nucleotides add only to 3'-OH groups of new polynucleotide strands. This leads to a fundamental difference in the way the two new daughter strands grow. At the fork on your right in Figure 23.32 [next page], the new 5'→3' strand (the complement of the parent 3'→5' strand) grows by **continuous** addition of

Figure 23.31



nucleotides to its 3' end as the fork moves along the original DNA molecule. In contrast, the new 3'→5' strand (the complement to the parent 5'→3' strand) grows **discontinuously**. Short polynucleotide segments form in a 5'→3' direction and join to form the complete 3'→5' strand later in the overall assembly process. At the fork on your left, new *continuous* and *discontinuous* strands grow on opposite sides of the bubble from those at the fork on your right because the two forks move in different directions.

Figure 23.32

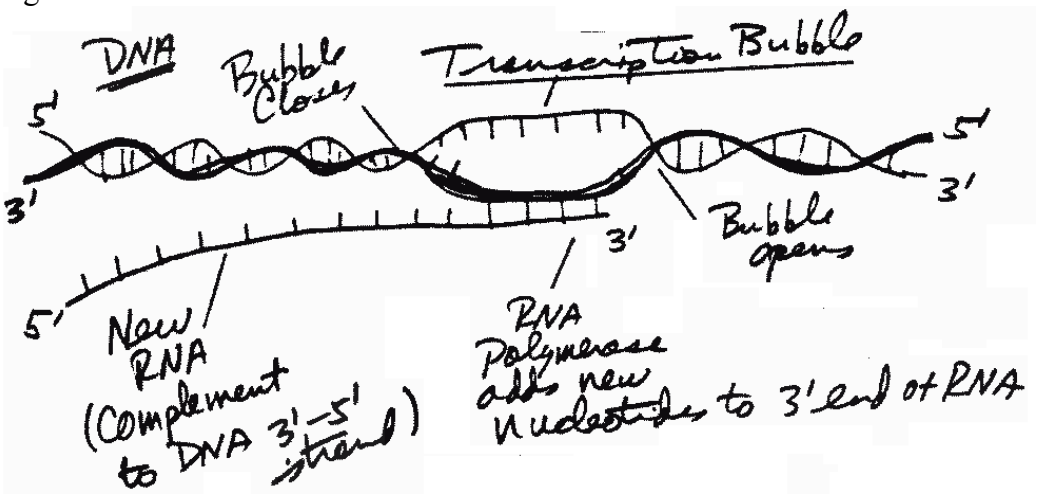


Transcription (23.2B)

Specific regions of the 3'→5' strand of DNA serve as templates for synthesis (*transcription*) of RNAs. DNA *transcribes* RNAs in 5'→3' directions (nucleotides add to the 3' end of the growing RNA strands) beginning at the 3' end of the DNA templates (Figure 23.33) [next page]. The resulting RNA strands are complementary to the template segments of the 3'→5'

DNA strands. Transcription occurs at a **transcription bubble** that has analogies to the *replication bubble* described above.

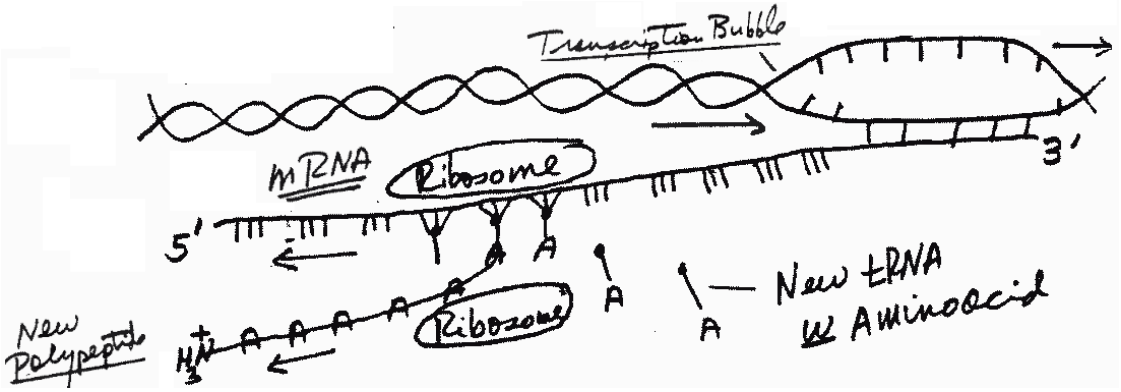
Figure 23.33



Translation (23.C)

Protein synthesis (*translation*) takes place in **ribosomes** containing *ribosomal RNA* (*rRNA*). Amino acids individually arrive at the *ribosome* brought by *transfer RNA* (*tRNA*) molecules that bind to *messenger RNA* (*mRNA*) just transcribed from DNA. The amino acids couple in a stepwise manner to yield the protein.

Figure 23.34



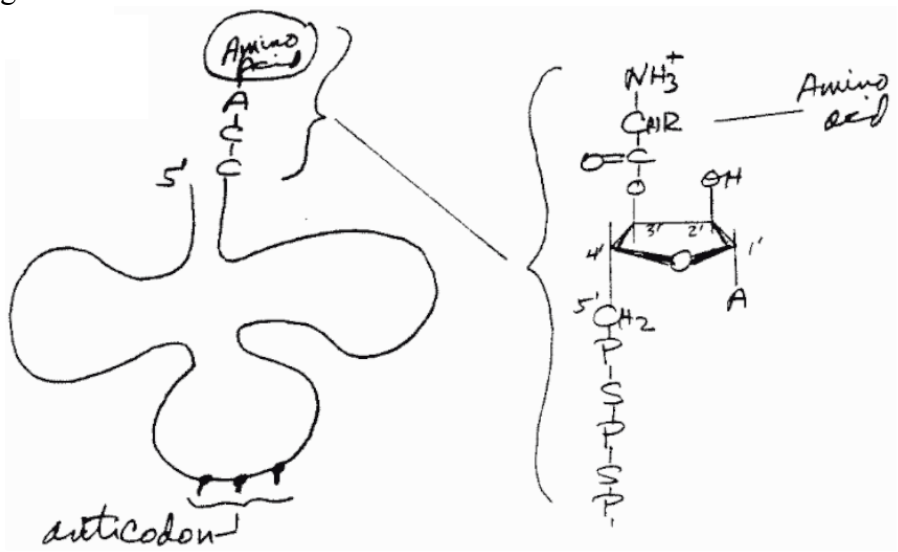
mRNA. The amino acid sequence in the protein results from the sequence of nucleotides in the *mRNA*. Three adjacent nucleosides in *mRNA* called a **codon** specify each amino acid (Table 23.3) [next page]. Since the *codon* GCU specifies alanine (Ala), the nucleoside sequence -GCU-GCU- in a mRNA specifies the amino acid sequence -Ala-Ala- in the protein. These *codons* are the **standard genetic code**. You can see in Table 23.3 [next page] that more than one *codon* specifies a particular amino acid, but in most organisms each *codon* specifies only one of the 20 standard amino acids.

Table 23.3. The Standard Genetic Code

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

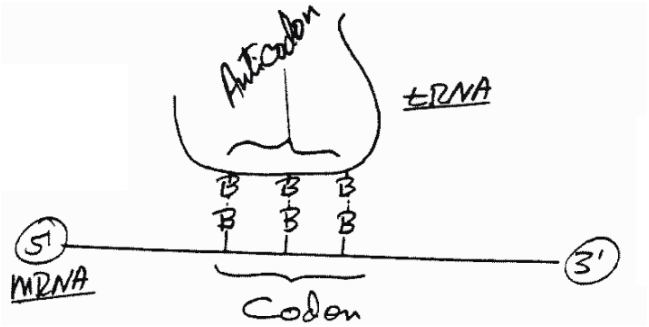
Amino Acid-tRNA Molecules. Amino acids covalently bind to **acceptor stems** of tRNAs that always terminate with the nucleoside sequence *CCA*. The carboxylate of the amino acid forms an ester with the 3'- or 2'-OH of ribose in the terminal *A* nucleoside.

Figure 23.35



The resulting *tRNA-amino acid molecules* hydrogen bond to a *codon* on *mRNA* specific to the amino acid.

Figure 23.36



This hydrogen bonding between the *codon* and a three-base sequence on tRNA called the **anticodon** depends only on the nucleoside sequence of *tRNA anticodon* and not the structure of the attached amino acid. The *tRNA* must first bind the correct amino acid or an incorrect amino acid will become part of the protein. The selectivity of a *tRNA* for a particular amino acid depends on the nucleoside content and sequence in both its *acceptor stem* and *anticodon loop*.

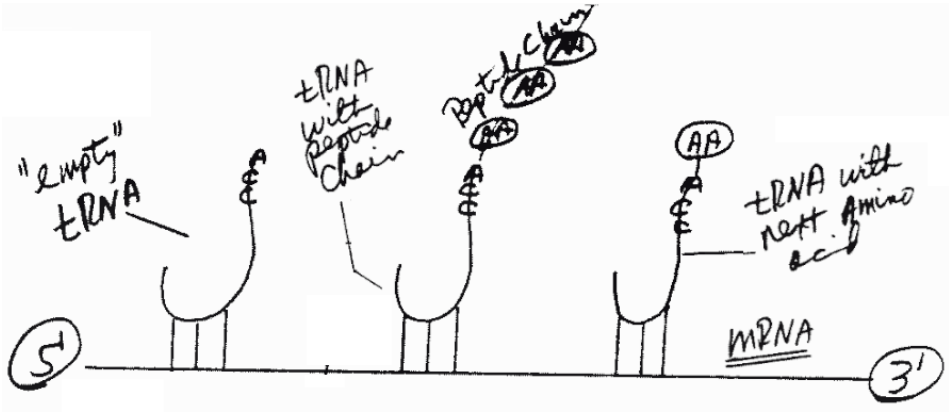
Codon-Anticodon Hydrogen Bonding. Different *codons* can hydrogen bond to the same *anticodon*. For example, the mRNA *codons* UUC and UUU for *Phe* bind the same tRNA. While the first two U's in UUC and UUU pair with A's in the tRNA *anticodon* (Watson-Crick base pairing), C (in UUC) and the third U (in UUU) each forms a base pair with the *modified base Gm* (see Figure 23.18) in the anticodon *GmAA*. In general, the first two bases in a *codon* must hydrogen bond to complementary bases in the *anticodon* (*G-C* or *A-U*), but the *codon's* third base has the apparent flexibility to form a *non-Watson-Crick base pair* with the remaining *anticodon* base.

> **Codon Sequence Order**. Nucleosides in mRNA codons (Table 23.3) are shown in their 5'→3' order. This order also specifies the nucleosides in the **first**, **second**, and **third positions** of the codon. In the UUC codon for *Phe*, U is in the *first position* at the 5' end of the sequence, while C is in the *third position* at the 3' end. *Anticodons* are also written in their 5'→3' order so the anticodon of *Phe* tRNA is written *GmAA* even though the codon and anticodon sequences hydrogen bond in opposite strand directions. The ability of codon bases in the *third position* to bond with a non-complementary base in the *first position* of the anticodon permits the variability of the third base in codons (see Table 23.3).

Steps in Protein Synthesis. Protein synthesis occurs in ribosomes through which the mRNA molecule moves. We can imagine a moment in time when three adjacent tRNAs are

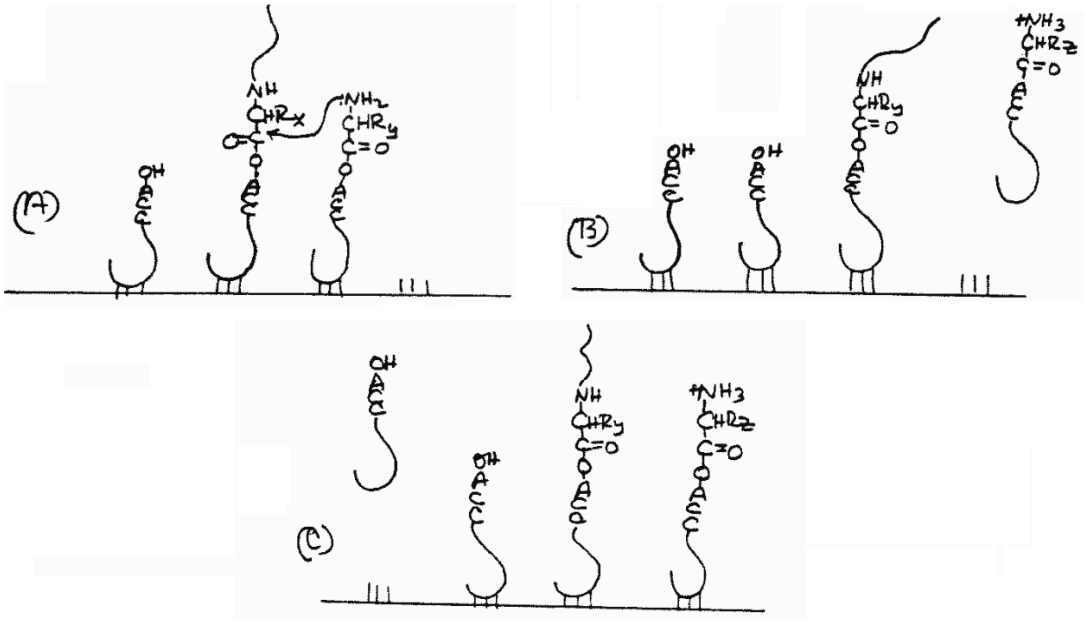
hydrogen bonded to mRNA in the ribosome (Figure 23.37). The middle tRNA carries the growing peptide (protein) chain, the tRNA on the 3' side (of mRNA) carries the next new amino acid that adds to the peptide, and the tRNA on the 5' side (of mRNA) is "empty" (it has no attached amino acid or peptide).

Figure 23.37



In a process called **transpeptidation**, the amino group of the amino acid-tRNA on the 3' side attacks C=O of the ester group binding the peptide chain to the middle tRNA (Figure 23.38A). This elongates the peptide chain by one amino acid and transfers it to the tRNA on the 3' side leaving a second "empty" tRNA on its 5' side (Figure 23.38B). The original "empty" tRNA on the 5' side leaves its site on mRNA and a new amino acid-tRNA binds on the 3' side (Figure 23.38C). The result is a new group of three tRNAs shifted (**translocated**) by one *codon* toward the 3' end of mRNA. These **chain elongation** steps repeat many times until all amino acids add to the peptide.

Figure 23.38



23.3 Nucleotide Biosynthesis and Degradation

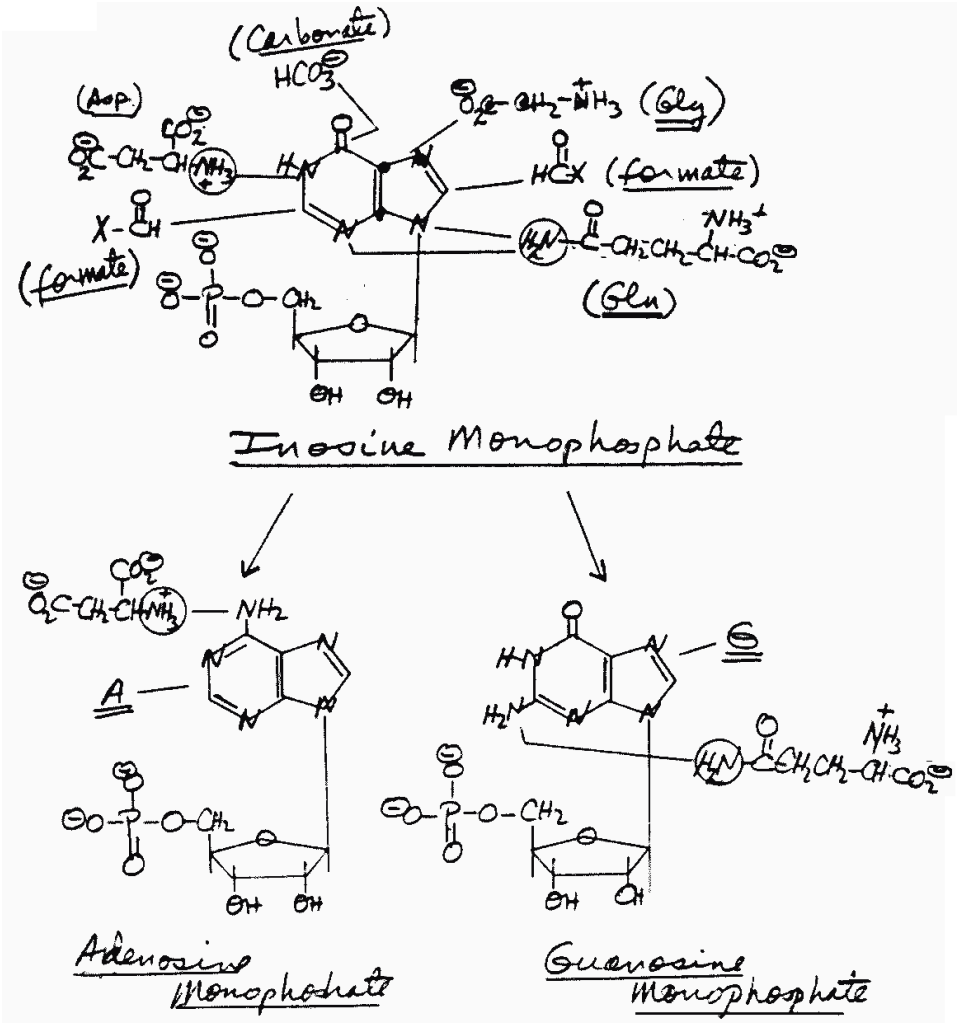
This section summarizes the biosynthetic origins and metabolic fates of the nucleotides of *A*, *G*, *C*, *U*, and *T*.

Biosynthesis (23.3A)

Purine and *pyrimidine* heterocyclic bases arise in different metabolic pathways.

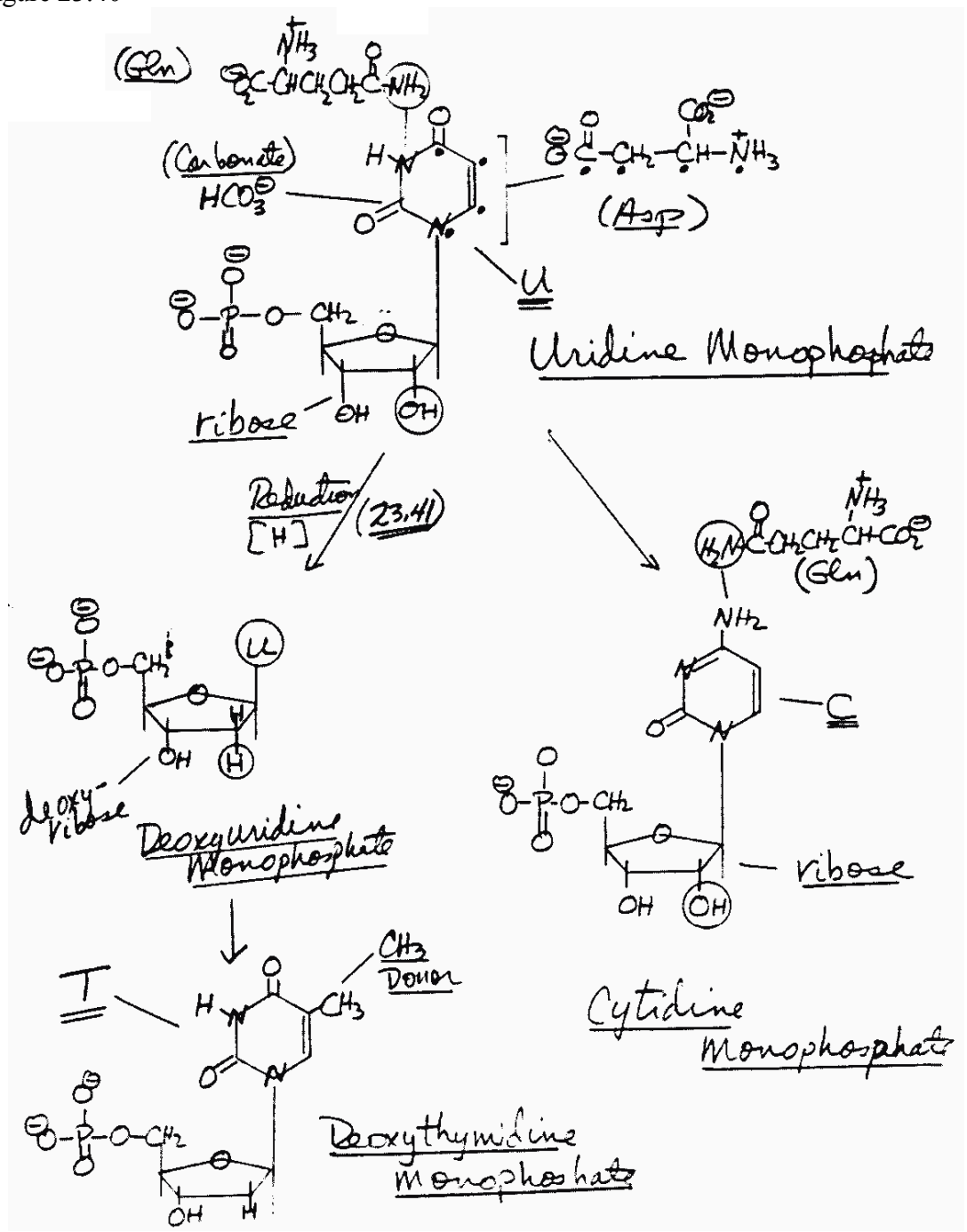
Purines. The purines *adenine* and *guanine* originate as ribose-5'-phosphate nucleotides from the common nucleotide intermediate *inosine monophosphate*. You can see that the individual atoms in *A* and *G* come from a variety of different sources.

Figure 23.39



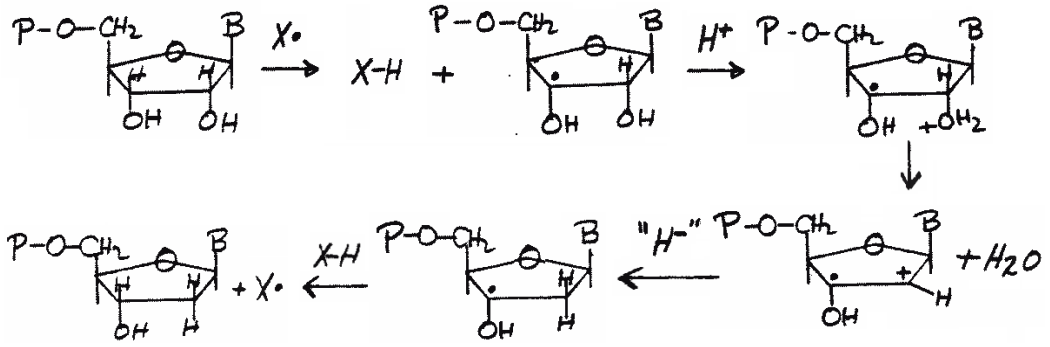
Pyrimidines. *Uracil* also comes from several different sources (Figure 23.40) [next page]. Its ribose-5'-phosphate nucleotide serves as the biosynthetic precursor of *cytosine* and *thymine* nucleotides. *Thymine* nucleotides contain *deoxyribose* and arise by enzymatic methylation of deoxyribose nucleotides of *uracil*.

Figure 23.40



Deoxyribose Nucleotides. While *deoxyribose* nucleotides of *thymine* come directly from *deoxyribose* nucleotides of *uracil*, *deoxyribose* nucleotides of *A*, *G*, *C*, and *U* come from their corresponding *ribose* nucleotides (Figure 23.41) [next page]. The multistep enzyme-catalyzed reduction reaction, where H replaces the 2'-OH group, involves radical and cation-radical intermediates.

Figure 23.41

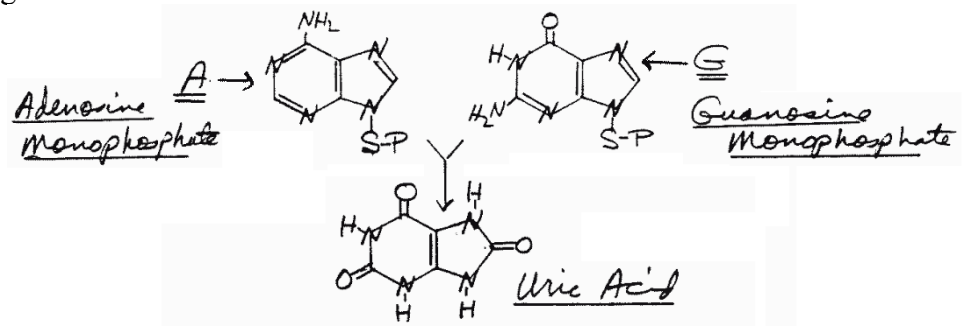


Degradation of Heterocyclic Bases (23.3B)

Just as they have different biosynthetic origins, the *purines* and *pyrimidines* have different metabolic fates.

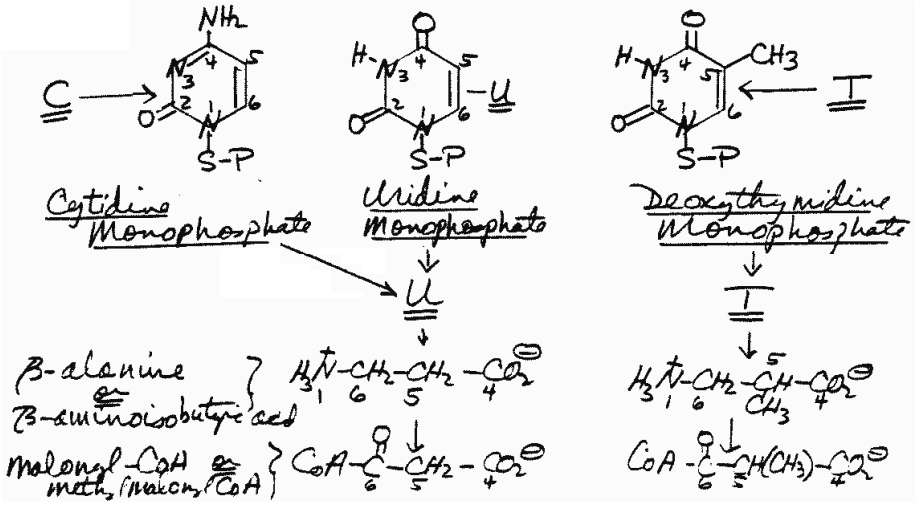
Purines. Organisms biosynthetically transform *adenine* or *guanine* into *uric acid*.

Figure 23.42



Pyrimidines. Nucleotides of *uracil* and *cytosine* metabolize to *β-alanine* (2-amino-propanoic acid) *via* the intermediate *uracil* free base.

Figure 23.43



In a chemically equivalent process, *thymine* nucleotides transform *via thymine* into *β-aminoisobutyric acid* (2-amino-1-methylpropanoic acid). β-Alanine and β-aminoisobutyric acid respectively become *malonyl-CoA* and *methylmalonyl-CoA* (Chapter 21).

Chapter Review

Structures of Nucleic Acids

(1) Nucleic acids are polynucleotides with alternating phosphate groups and ribofuranosides (or deoxyribofuranosides), with glycosidically bonded heterocyclic bases that are primarily adenine (A), guanine (G), cytosine (C), thymine (T), or uracil (U). (2) A furanoside and its heterocyclic base are a nucleoside, while a nucleotide is the 3' or 5' phosphate of a nucleoside. (3) Deoxyribonucleic acids (DNAs) have two intertwining α-helical polynucleotide strands with deoxyribose and A, G, C, and T. (4) Ribonucleic acids (RNAs) are single stranded polynucleotides with ribose, A, G, C, and U, as well as modified heterocyclic bases. (5) The base sequences of each DNA strand are complementary so that base pairs are A-T and G-C. (6) Base pairing in RNAs occurs between bases on the same folded strand. (7) Hydrophobic bonding primarily stabilizes nucleic acids. (8) Chemical sequencing removes specific nucleosides whose location is identified by the size of the resulting fragments.

Replication, Transcription, and Translation

(1) DNA replication is semiconservative and occurs by addition of nucleotides to the 3' ends of both the new discontinuous and continuous strands. (2) RNA forms as a continuous 5'→3' strand by transcription of a segment of the 3'→5' DNA strand. (3) mRNA serves as a template for protein synthesis (translation) as it is transcribed from a DNA segment. (4) tRNA molecules transport specific amino acids to the ribosome translation site and they bind to mRNA via anticodon-codon hydrogen bonding. (5) Codons are triplets of heterocyclic bases on mRNA that uniquely specify a specific amino acid. (6) An amino acid covalently binds to the acceptor stem of tRNA *via* an ester linkage between the carboxylic acid group of the amino acid and the 3' (or 2') OH of the ribose of the terminal *A* nucleoside. (7) Elongation of the peptide chain occurs by amide bond formation between the amino group of an amino acid and the C=O group of the peptide chain.

Nucleotide Biosynthesis and Degradation

(1) Purines and pyrimidines have different biosynthetic pathways from a variety of precursors. (2) Deoxythymidine nucleotides are formed by methylation of deoxyuridine nucleotides, but all other deoxyribose nucleotides are formed from their ribose nucleotides by replacement of the 2'-OH by H. (3) A and G nucleotides degrade to give uric acid, while C, U, and T nucleotides degrade to amino acids that are subsequently converted into malonyl Co-A or into methylmalonyl Co-A.