

Chapter 28 – Biomolecules: Nucleic Acids

Chapter Outline

I. Nucleic acids (Sections 24.1 – 24.2).

A. Nucleotides (Section 24.1).

1. Nucleotides are composed of a heterocyclic purine or pyrimidine base, an aldopentose, and a phosphate group.
 - a. In RNA, the purines are adenine and guanine, the pyrimidines are uracil and cytosine, and the sugar is ribose.
 - b. In DNA, thymine replaces uracil, and the sugar is 2'-deoxyribose.
2. Positions on the base receive non-prime superscripts, and positions on the sugar receive prime superscripts.
3. The heterocyclic base is bonded to C1' of the sugar.
4. DNA is vastly larger than RNA and is found in the cell nucleus.

B. Nucleic acids.

1. Nucleic acids are composed of nucleotides connected by a phosphate ester bond between the 5' ester of one nucleotide and the 3' hydroxyl group of another.
 - a. One end of the nucleic acid polymer has a free hydroxyl group and is called the 3' end.
 - b. The other end has a free phosphate group and is called the 5' end.
2. The structure of a nucleic acid depends on the order of bases.
3. The sequence of bases is described by starting at the 5' end and listing the bases by their one-letter abbreviations.

C. Base-pairing in DNA (Section 24.2).

1. DNA consists of two polynucleotide strands coiled in a double helix.

Adenine and thymine hydrogen-bond with each other, and cytosine and guanine hydrogen-bond with each other.
2. Because the two DNA strands are complementary, the amount of A equals the amount of T, and the amount of C equals the amount of G.
3. The double helix is 2.0 Å wide, there are 10 bases in each turn, and each turn is 3.4 Å in height.
4. The double helix has a major groove and a minor groove into which polycyclic aromatic molecules can intercalate.

D. The "central dogma" of molecular genetics.

1. The function of DNA is to store genetic information and to pass it on to RNA, which uses it to make proteins.
2. Replication, transcription and translation are the three processes that are responsible for carrying out the central dogma.

II. The transfer of genetic information (Sections 28.3 – 28.5).

A. Replication of DNA (Section 28.3).

1. Replication is the enzyme-catalyzed process whereby DNA makes a copy of itself.
2. Replication is semiconservative: each new strand of DNA consists of one old strand and one newly synthesized strand.
3. How replication occurs:
 - a. The DNA helix partially unwinds.
 - b. New nucleotides form base pairs with their complementary partners.

- c. Formation of new bonds is catalyzed by DNA polymerase and takes place in the 5' → 3' direction.
Bond formation occurs by attack of the 3' hydroxyl group on the 5' triphosphate, with loss of a diphosphate leaving group.
- d. Both new chains are synthesized in the 5' → 3' direction.
 - i. One chain is synthesized continuously (the leading strand).
 - ii. The other strand is synthesized in small pieces, which are later joined by DNA ligase enzymes (the lagging strand).

B. Transcription (Section 28.4).

- 1. There are 3 types of RNA:
 - a. Messenger RNA (mRNA) carries genetic information to ribosomes when protein synthesis takes place.
 - b. Ribosomal RNA (rRNA), complexed with protein, comprises the physical makeup of the ribosomes.
 - c. Transfer RNA (tRNA) brings amino acids to the ribosomes, where they are joined to make proteins.
- 2. DNA contains "promoter sites", which indicate where mRNA synthesis is to begin, and base sequences that indicate where mRNA synthesis stops.
RNA polymerase binds to the promoter sequence.
- 3. mRNA is synthesized in the nucleus by transcription of DNA.
 - a. The DNA partially unwinds, forming a "bubble".
 - b. Ribonucleotides form base pairs with their complementary DNA bases.
 - c. Bond formation occurs in the 5' → 3' direction.
 - d. Only one strand of DNA (the template strand) is transcribed.
 - e. Thus, the synthesized mRNA is a copy of the coding strand with U replacing T.
- 4. Synthesis of mRNA is not necessarily continuous.
 - a. Often, synthesis begins in a region of DNA called an exon and is interrupted by a seemingly nonsensical region of DNA called an intron.
 - b. In the final mRNA, the nonsense sections have been removed and the remaining pieces have been spliced together.

C. Translation (Section 28.5).

- 1. Translation is the process in which proteins are synthesized at the ribosomes by using mRNA as a template.
- 2. The message delivered by mRNA is contained in "codons" – 3-base groupings that are specific for an amino acid.
 - a. Amino acids are coded by 61 of the possible 64 codons.
 - b. The other 3 codons are "stop" codons.
- 3. Each tRNA is responsible for bringing an amino acid to the growing protein chain.
 - a. A tRNA has a cloverleaf-shaped secondary structure and consists of 70–100 ribonucleotides.
 - b. Each tRNA contains an anticodon complementary to the mRNA codon.
- 4. The protein chain is synthesized by enzyme-catalyzed peptide bond formation.
- 5. A 3-base "stop" codon on mRNA signals when synthesis is complete.

IV. DNA technology (Sections 28.6 – 28.8).

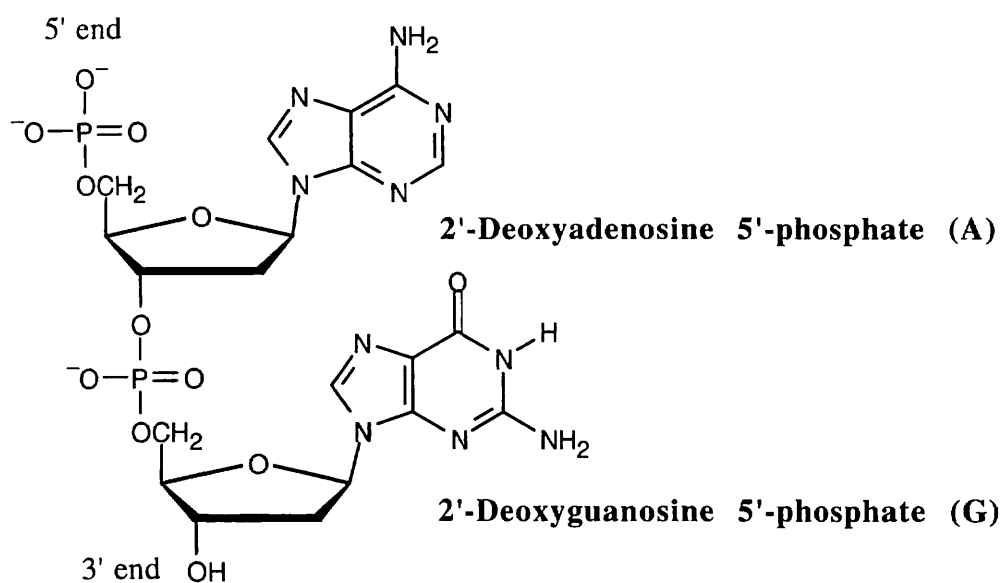
A. DNA sequencing (Section 28.6).

- 1. Before sequencing, the DNA chain is cleaved at specific sites by restriction endonucleases.
 - a. The restriction endonuclease recognizes both a sequence on the coding strand and its complement on the template strand.
 - b. The DNA strand is cleaved by several different restriction endonucleases, to produce fragments that overlap those from a different cleavage.

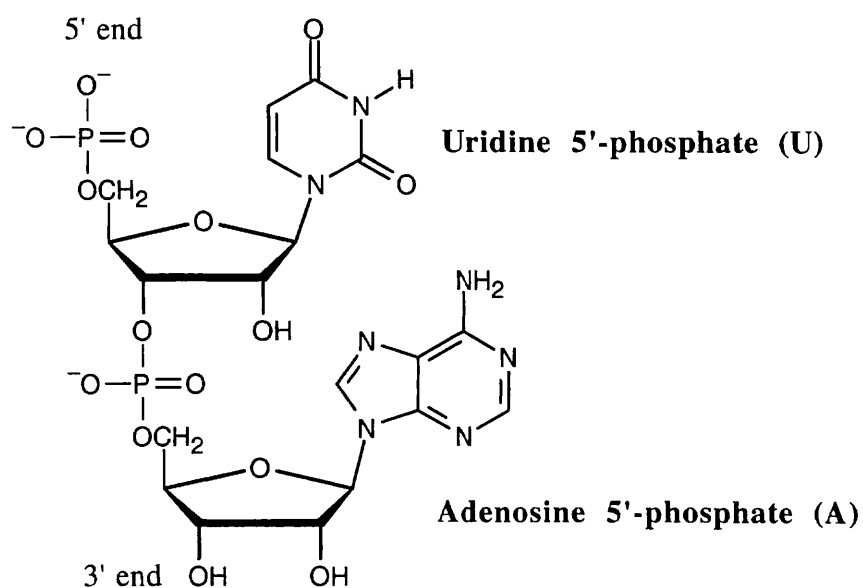
2. Maxam–Gilbert DNA sequencing.
This method uses chemical techniques.
 3. Sanger dideoxy DNA sequencing.
 - a. The following mixture is assembled:
 - i. The restriction fragment to be sequenced.
 - ii. A primer (a small piece of DNA whose sequence is complementary to that on the 3' end of the fragment).
 - iii. The 4 DNA nucleotide triphosphates.
 - iv. Small amounts of the four dideoxynucleotide triphosphates, each of which is labeled with a different fluorescent dye.
 - b. DNA polymerase is added to the mixture, and a strand begins to grow from the end of the primer.
 - c. Whenever a dideoxynucleotide is incorporated, chain growth stops.
 - d. When reaction is complete, the fragments are separated by gel electrophoresis.
 - e. Because fragments of all possible lengths are represented, the sequence can be read by noting the color of fluorescence of each fragment.
- B. DNA synthesis (Section 28.7).
1. DNA synthesis is based on principles similar to those for protein synthesis.
 2. The following steps are needed:
 - a. The nucleosides are protected and bound to a silica support.
 - i. Adenine and cytosine bases are protected by benzoyl groups.
 - ii. Guanine is protected by an isobutyryl group.
 - iii. Thymine isn't protected.
 - iv. The 5' –OH group is protected as a DMT ether.
 - b. The DMT group is removed.
 - c. The polymer-bound nucleoside is coupled with a protected nucleoside containing a phosphoramidite group.
 - i. One of the phosphoramidite oxygens is protected as a β -cyano ether.
 - ii. Tetrazole catalyzes the coupling.
 - d. The phosphite is oxidized to a phosphate with I_2 .
 - e. Steps b – d are repeated until the desired chain is synthesized.
 - f. All protecting groups are removed and the bond to the support is cleaved by treatment with aqueous ammonia.
- C. The polymerase chain reaction (Section 28.8).
1. The polymerase chain reaction (PCR) can produce vast quantities of a DNA fragment.
 2. The key to PCR is *Taq* DNA polymerase, a heat-stable enzyme.
Newer heat-stable DNA polymerase enzymes have become available.
 3. Steps in PCR:
 - a. The following mixture is heated to 95°C (a temperature at which DNA becomes single-stranded);
 - i. *Taq* polymerase.
 - ii. Mg^{2+} ion.
 - iii. The 4 deoxynucleotide triphosphates.
 - iv. A large excess of two oligonucleotide primers, each of which is complementary to the ends of the fragment to be synthesized.
 - b. The temperature is lowered to 37°C–50°C, causing the primers to hydrogen-bond to the single-stranded DNA.
 - c. After raising the temperature to 72°C, *Taq* catalyzes the addition of further nucleotides, yielding two copies of the original DNA.
 - d. The process is repeated until the desired quantity of DNA is produced.

Solutions to Problems

28.1



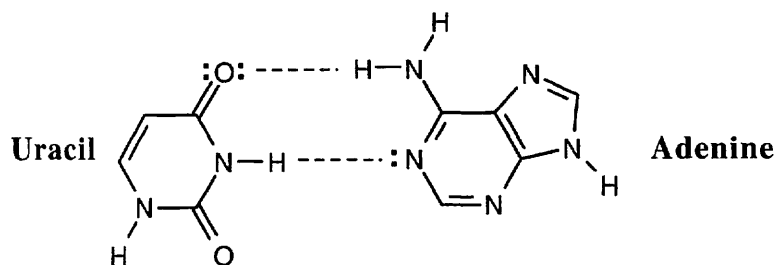
28.2



28.3 DNA (5' end) GGCTAATCCGT (3' end) is complementary to
DNA (3' end) CCGATTAGGCA (5' end)

Remember that the complementary strand has the 3' end on the left and the 5' end on the right.

28.4



28.5 DNA (5' end) GATTACCGTA (3' end) is complementary to
 RNA (3' end) CUAAUGGCAU (5' end)

28.6 RNA (5' end) UUCGCAGAGU (3' end)
 DNA (3' end) AAGCGTCTCA (5' end) template strand

28.7–28.8 Several different codons can code for the same amino acid. The corresponding anticodon follows the slash mark after each codon. The mRNA codons are written with the 5' end on the left and the 3' end on the right, and the tRNA anticodons have the 3' end on the left and the 5' end on the right.

Amino acid:	Ala	Phe	Leu	Tyr
Codon sequence/ tRNA anticodon:	GCU/CGA GCC/CGG GCA/CGU GCG/CGC	UUU/AAA UUC/AAG	UUA/AAU UUG/AAC CUU/GAA CUC/GAG CUA/GAU CUG/GAC	UAU/AUA UAC/AUG

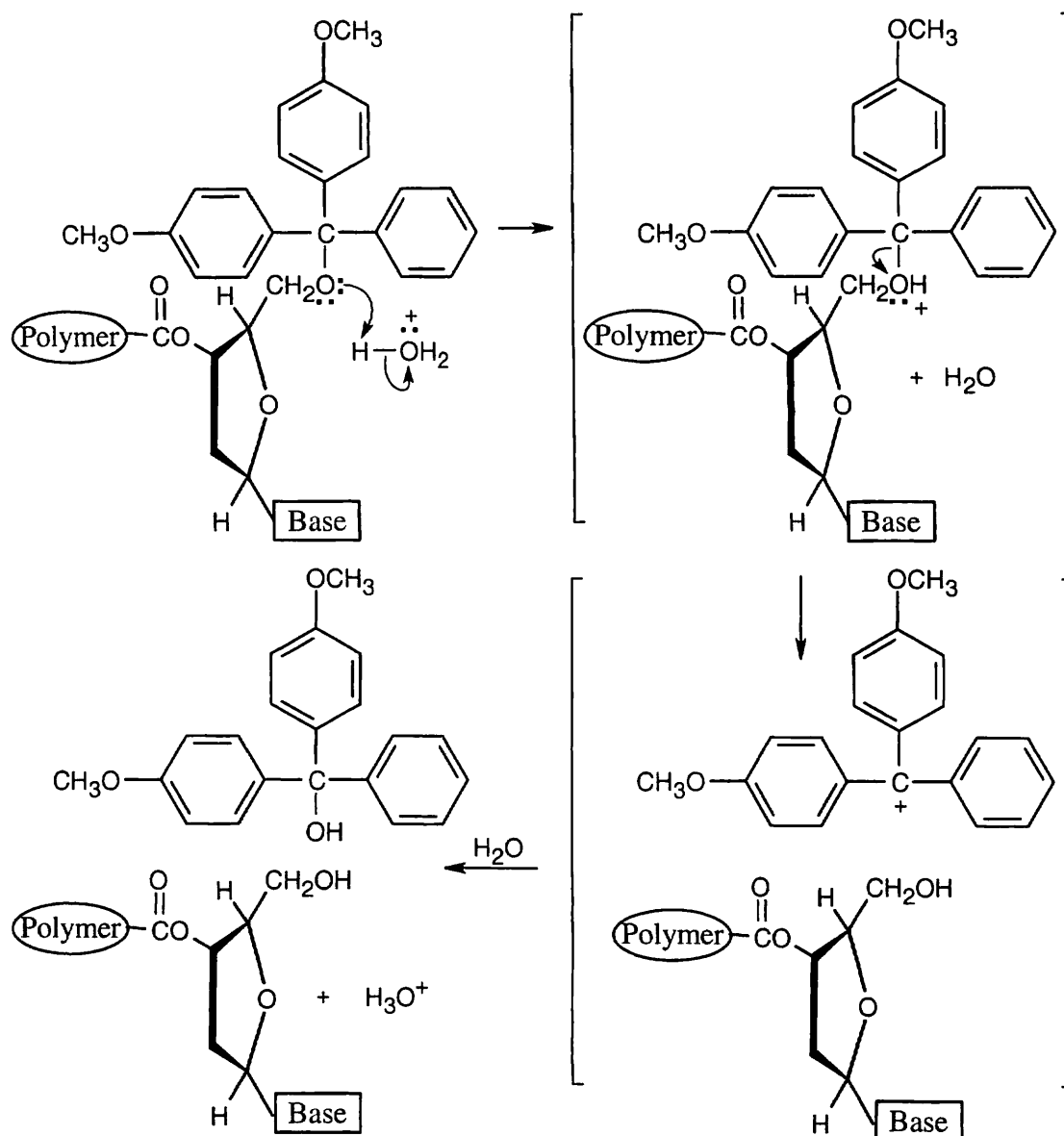
28.9–28.10

The mRNA base sequence: (5' end) CUU–AUG–GCU–UGG–CCC–UAA (3' end)

The amino acid sequence: Leu—Met—Ala—Trp—Pro—(stop)

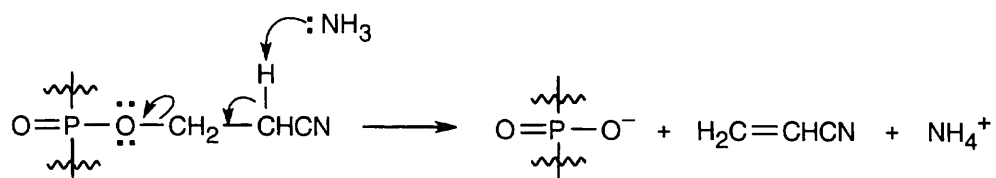
The DNA sequence: (3' end) GAA–TAC–CGA–ACC–GGG–ATT (5' end)
 (template strand)

28.11



Cleavage of DMT ethers proceeds by an S_N1 mechanism and is rapid because the DMT cation is unusually stable.

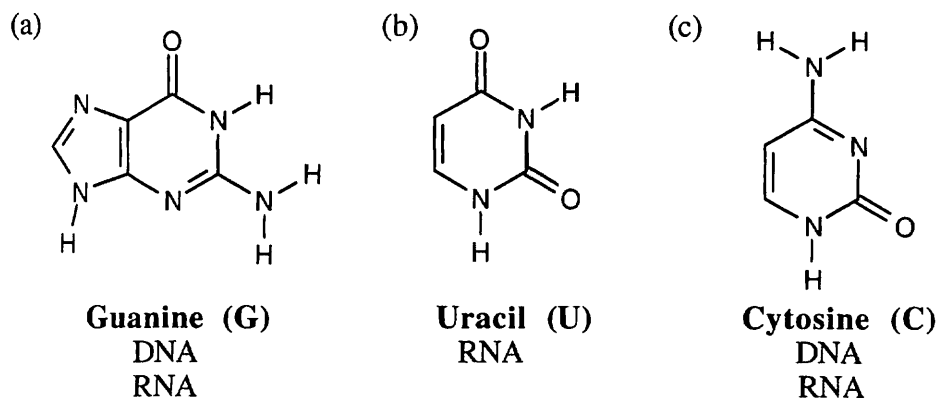
28.12



This is an $E2$ elimination reaction, which proceeds easily because the hydrogen α to the nitrile group is acidic.

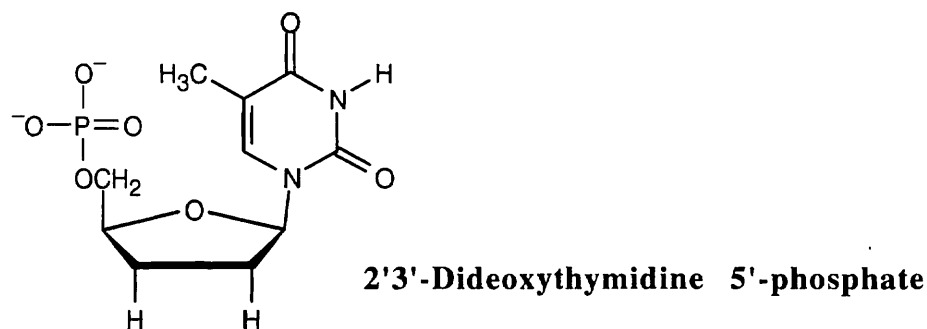
Visualizing Chemistry

28.13



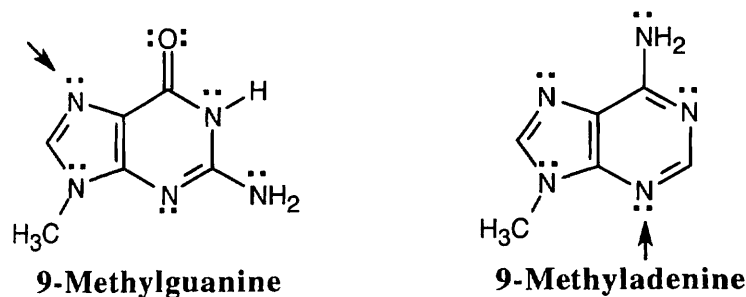
All three bases are found in RNA, but only guanine and cytosine are found in DNA.

28.14



The triphosphate made from 2'-deoxythymidine 5' phosphate is labeled with a fluorescent dye and used in the Sanger method of DNA sequencing. Along with the restriction fragment to be sequenced, a DNA primer, and a mixture of the four dNTPs, small quantities of the four labeled dideoxynucleotide triphosphates are mixed together. DNA polymerase is added, and a strand of DNA complementary to the restriction fragment is synthesized. Whenever a dideoxynucleotide is incorporated into the DNA chain, chain growth stops. The fragments are separated by electrophoresis, and each terminal dideoxynucleotide can be identified by the color of its fluorescence. By identifying these terminal dideoxynucleotides, the sequence of the restriction fragment can be read.

28.15 According to the electrostatic potential map, the nitrogen at the 7 position of 9-methylguanine is more electron-rich and should be more nucleophilic. Thus 9-methylguanine should be the better nucleophile.



Additional Problems

28.16 The DNA that codes for natriuretic peptide (32 amino acids) consists of 99 bases; 3 bases code for each of the 32 amino acids in the chain (96 bases), and a 3-base "stop" codon is also needed.

28.17 *Position 9:*

Horse amino acid = Gly	Human amino acid = Ser
mRNA codons (5' → 3'):	
GGU GGC GGA GGG	UCU UCC UCA UCG AGU AGC
DNA bases (template strand 3' → 5'):	
<u>CCA</u> <u>CCG</u> CCT CCC	AGA AGG AGT AGC <u>TCA</u> <u>TCG</u>

The underlined horse DNA base triplets differ from their human counterparts (also underlined) by only one base.

Position 30:

Horse amino acid = Ala	Human amino acid = Thr
mRNA codons (5' → 3'):	
GCU GCC GCA GCG	ACU ACC ACA ACG
DNA bases (template strand 3' → 5'):	
CGA CGG CGT CGC	TGA TGG TGT TGC

Each of the above groups of DNA bases from horse insulin has a counterpart in human insulin that differs from it by only one base. It is possible that horse insulin DNA differs from human insulin DNA by only two bases out of 159!

28.18 The percent of A always equals the percent of T, since A and T are complementary. The percent G equals the percent C for the same reason. Thus, sea urchin DNA contains about 32% each of A and T, and about 18% each of G and C.

28.19 Even though the stretch of DNA shown contains UAA in sequence, protein synthesis doesn't stop. The codons are read as 3-base individual units from start to end, and, in this mRNA sequence, the unit UAA is read as part of two codons, not as a single codon.

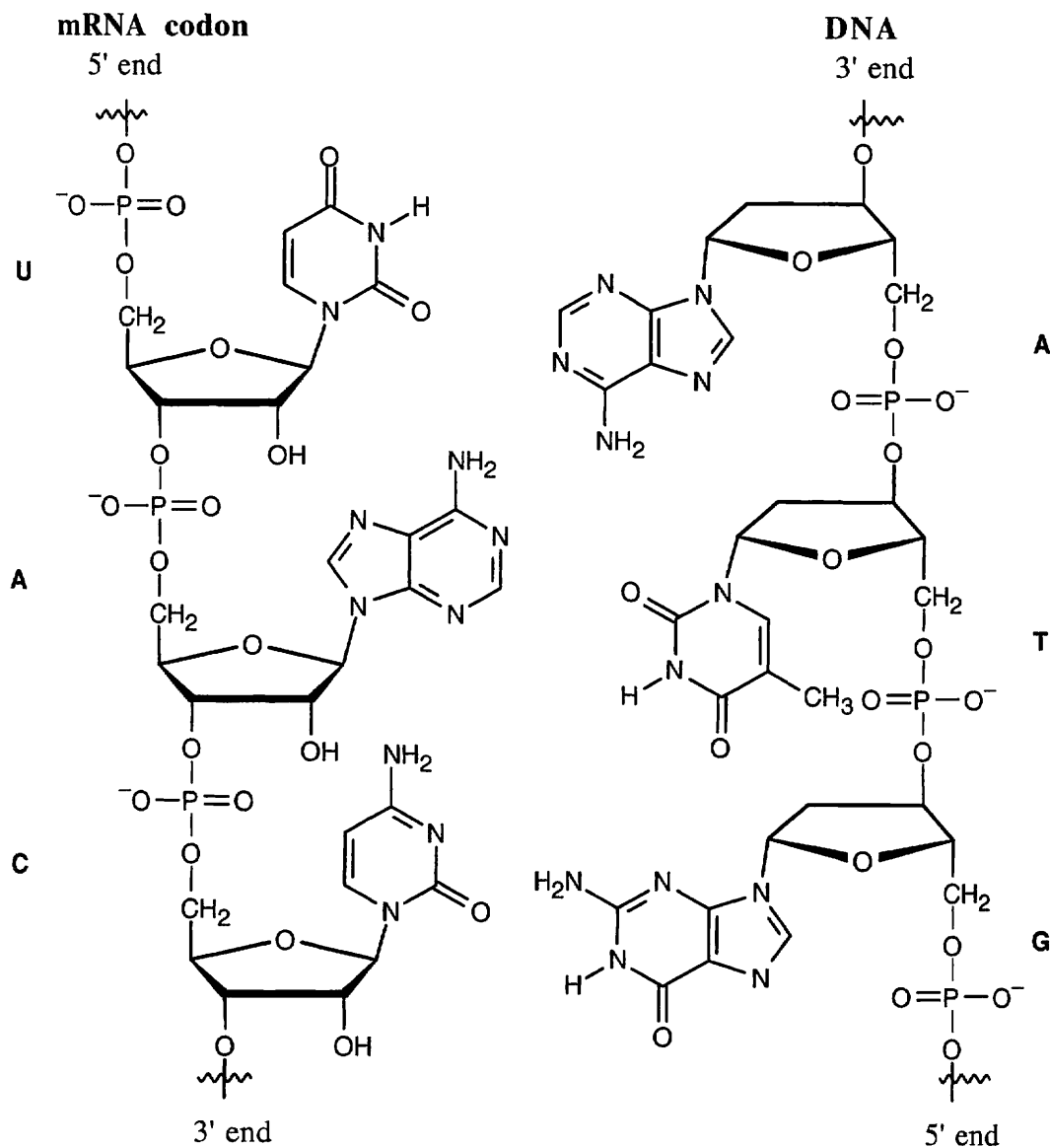
28.20 Restriction endonucleases cleave DNA base sequences that are palindromes, meaning that the sequence reads the same as the complement when both are read in the (5') to (3') direction. Thus, the sequence in (c), CTCGAG is recognized. The sequence in (a), GAATTC, is also a palindrome and is recognized by a restriction endonuclease. The sequence in (b) is not recognized.

28.21–28.23

mRNA codon :	(5'→3')	(a) AAU	(b) GAG	(c) UCC	(d) CAU
Amino acid:		Asn	Glu	Ser	His
DNA sequence:	(3'→5')	TTA	CTC	AGG	GTA
tRNA anticodon:	(3'→5')	UUA	CUC	AGG	GUA

The DNA sequence of the template strand is shown.

28.24–28.25 UAC is a codon for tyrosine. It was transcribed from ATG of the template strand of a DNA chain.



28.26 Tyr—Gly—Gly—Phe—Met (stop) is coded by

UAC	GGU	GGU	UUU	AUG	UAA
UAU	GGC	GGC	UUC		UAG
	GGA	GGA			UGA
	GGG	GGG			

A total of $2 \times 4 \times 4 \times 2 \times 1 \times 3 = 194$ different mRNA sequences can code for metenkephalin!

28.27 Angiotensin II: Asp—Arg—Val—Tyr—Ile—His—Pro—Phe (stop)

mRNA sequence: GAU CGU GUU UAU AUU CAU CCU UUU UAA
 (5'→3') GAC CGC GUC UAC AUC CAC CCC UUC UAG
 CGA GUA AUA CCA UGA
 CGG GUG CCG
 AGA
 AGG

As in the previous problem, many mRNA sequences (13,824) can code for angiotensin II.

28.28 DNA coding strand (5'→3'): CTT—CGA—CCA—GAC—AGC—TTT

mRNA (5'→3'): CUU—CGA—CCA—GAC—AGC—UUU

Amino acid sequence: Leu—Arg—Pro—Asp—Ser—Phe

The mRNA sequence is the complement of the DNA template strand, which is the complement of the DNA coding strand. Thus, the mRNA sequence is a copy of the DNA coding strand, with T replaced by U.

28.29 mRNA sequence (5'→3'): CUA—GAC—CGU—UCC—AAG—UGA

Amino Acid: Leu—Asp—Arg—Ser—Lys (stop)

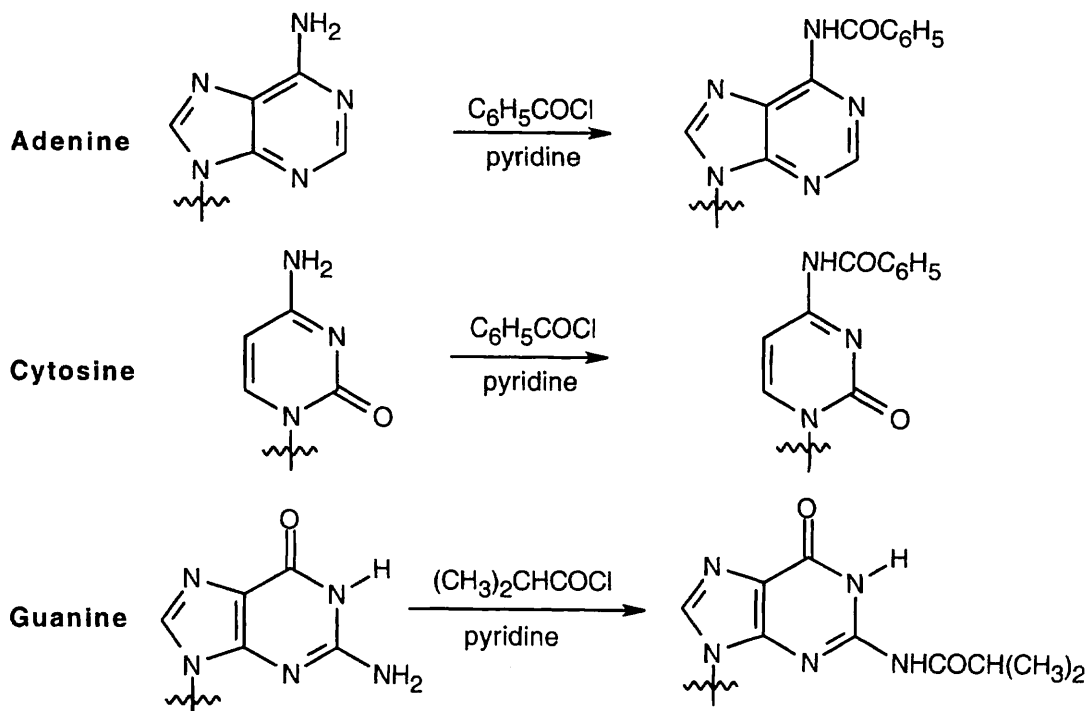
28.30

	<i>Original Sequence</i>	<i>Miscopied Sequence</i>
DNA coding strand (5'→3'):	-CAA-CCG-GAT-	-CGA-CCG-GAT-
mRNA sequence (5'→3'):	-CAA-CCG-GAU-	-CGA-CCG-GAU-
Amino acid sequence:	-Gln—Pro—Asp-	-Arg—Pro—Asp-

If this gene sequence were miscopied in the indicated way, a glutamine in the original protein would be replaced by an arginine in the mutated protein.

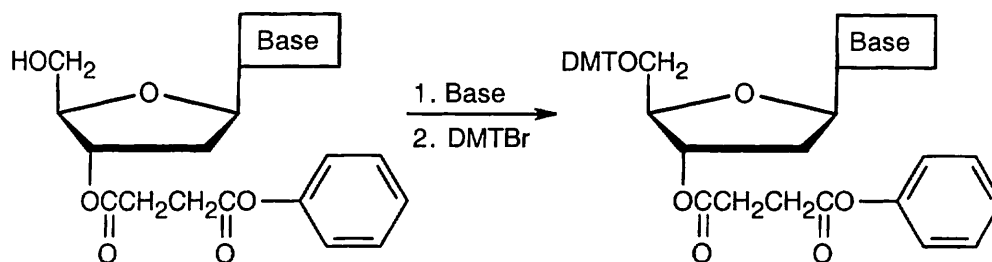
28.31 1. First, protect the nucleotides.

(a) Bases are protected by amide formation.

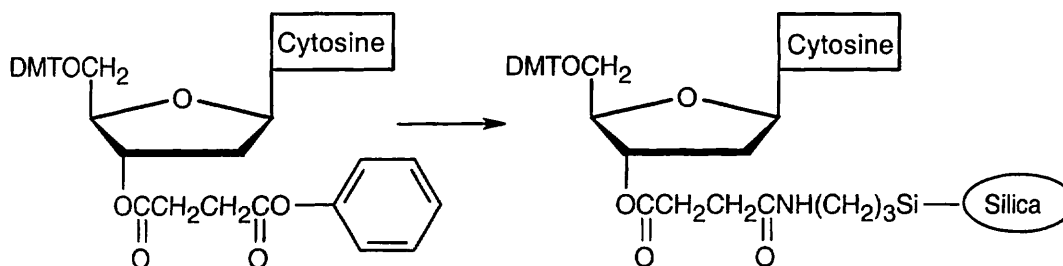


Thymine does not need to be protected.

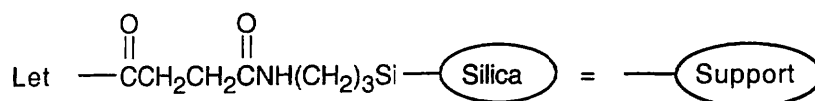
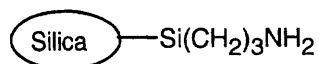
(b) The 5' hydroxyl group is protected as its *p*-dimethoxytrityl (DMT) ether.



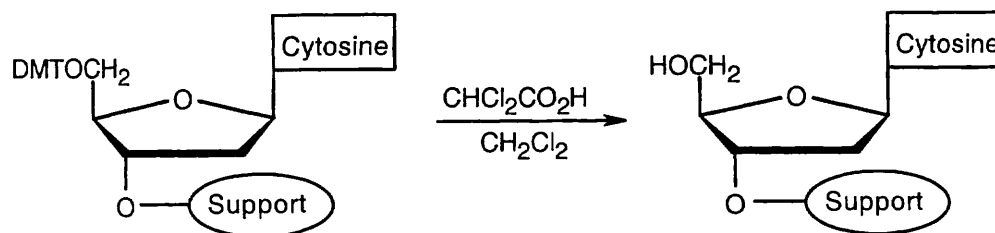
2. Attach a protected 2-deoxycytidine nucleoside to the polymer support.



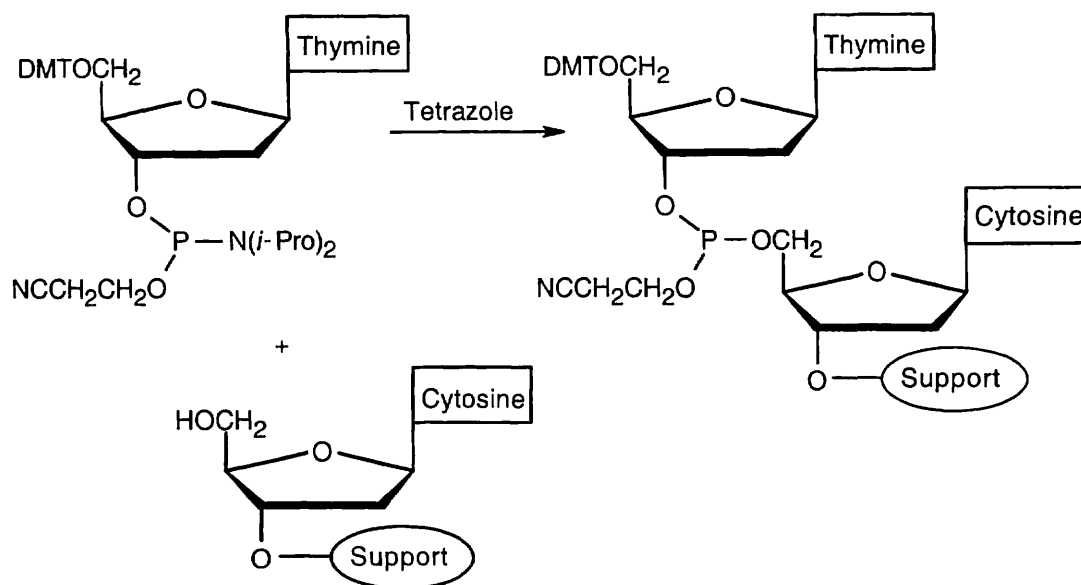
+



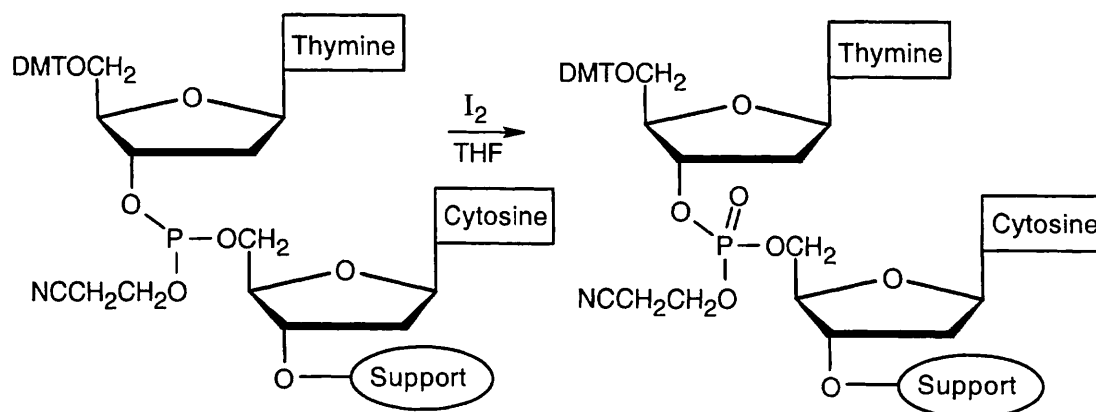
3. Cleave the DMT ether.



4. Couple protected 2'-deoxythymidine to the polymer-2'-deoxycytidine. (The nucleosides have a phosphoramidite group at the 3' position.)

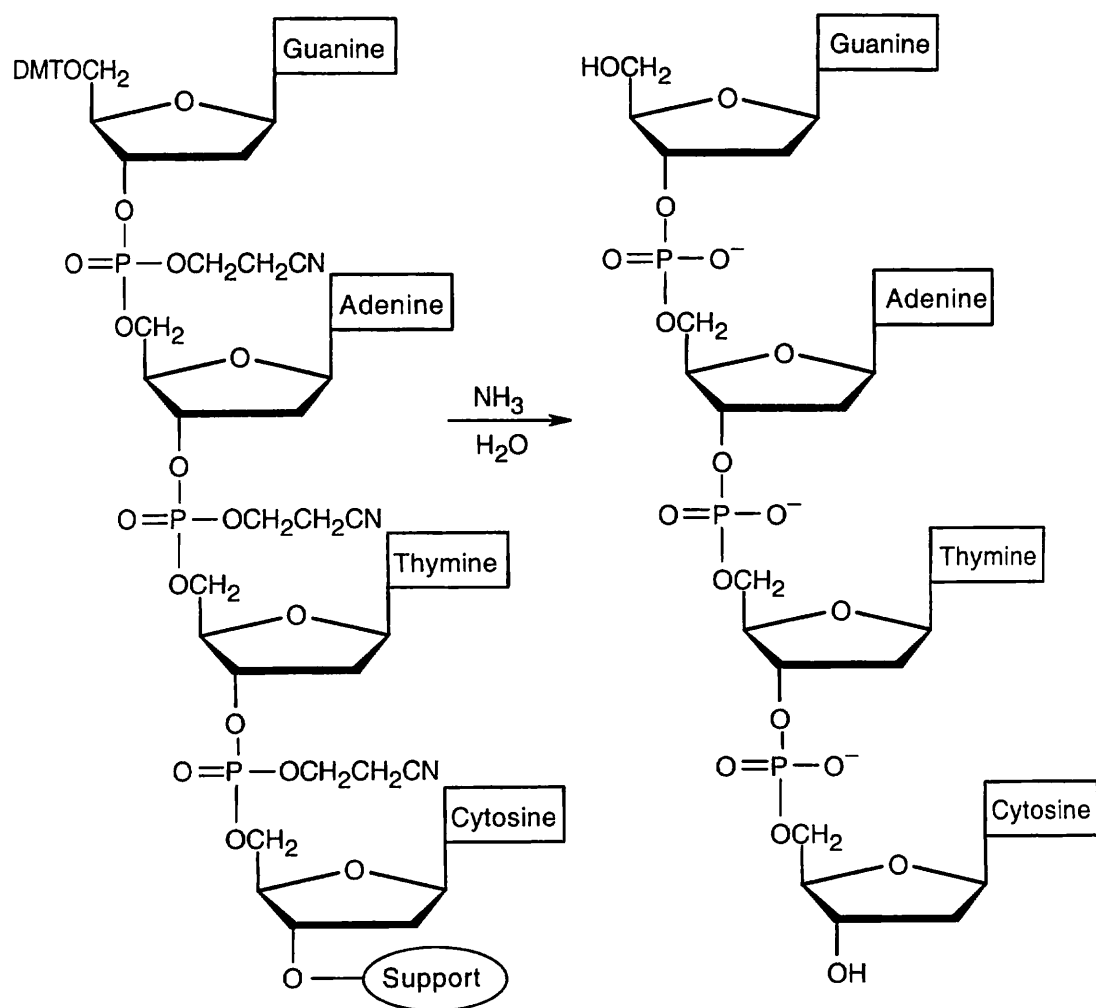


5. Oxidize the phosphite product to a phosphate triester, using iodine.



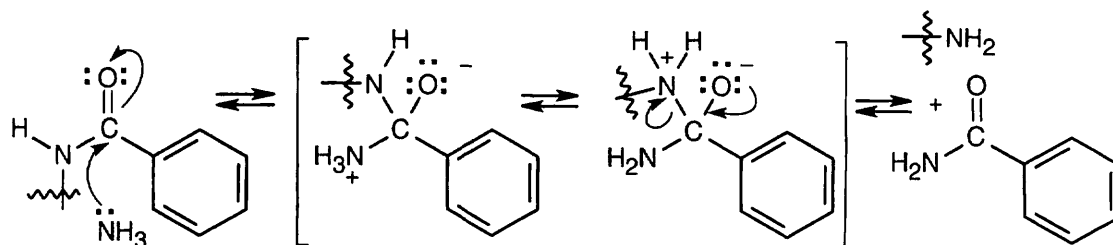
6. Repeat steps 3–5 with protected 2'-deoxyadenosine and protected 2'-deoxyguanosine.

7. Cleave all protecting groups with aqueous ammonia to yield the desired sequence.

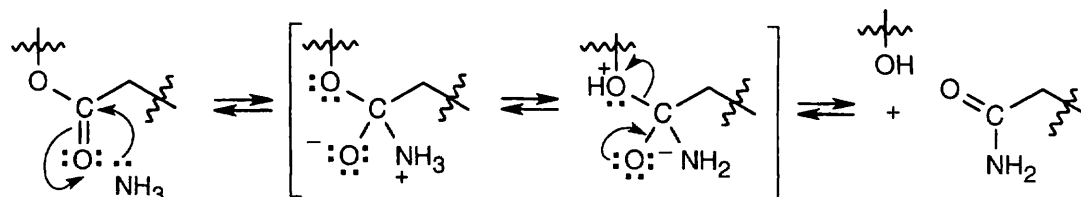


28.32 Both of these cleavages occur by the now-familiar nucleophilic acyl substitution route. A nucleophile adds to the carbonyl group, a proton shifts location, and a second group is eliminated.

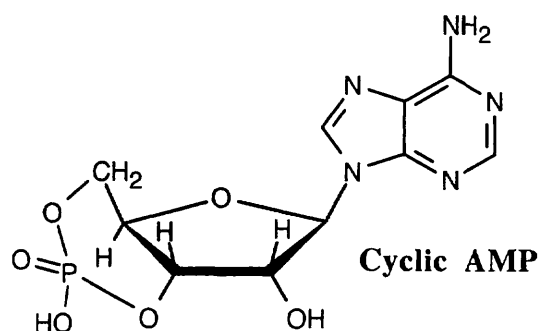
Deprotection at 1:



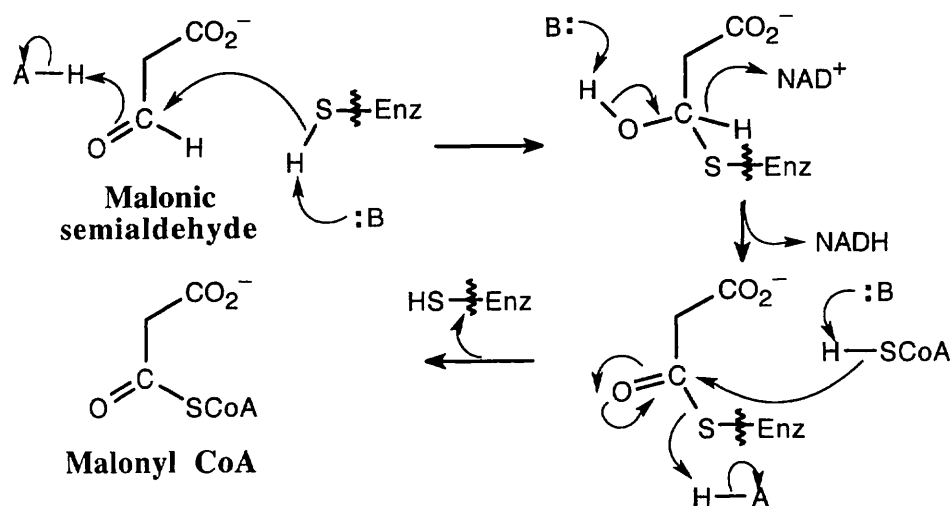
Deprotection at 2:



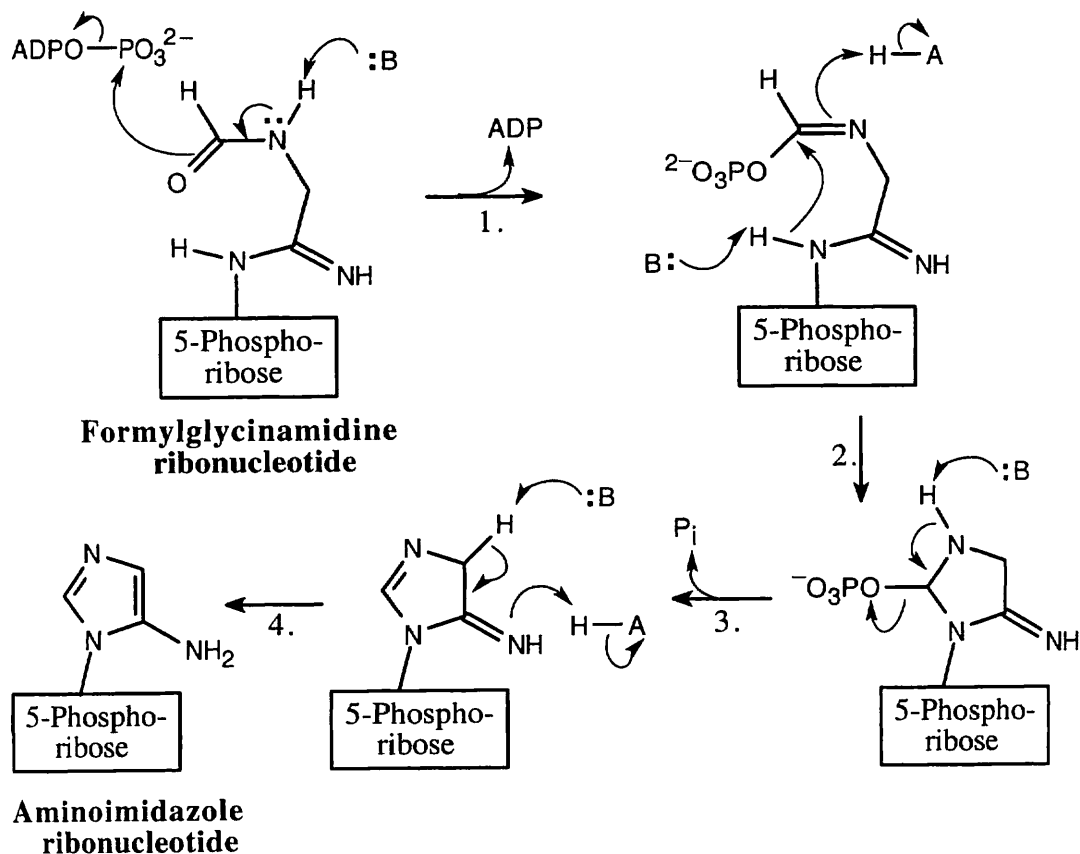
28.33



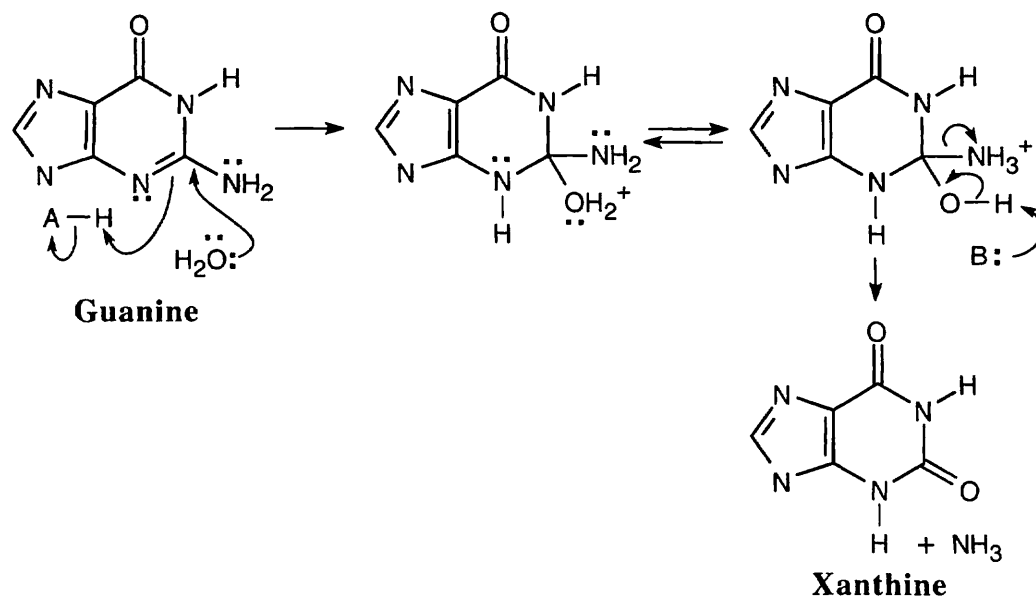
28.34 This reaction involves addition of a thiol residue of the enzyme to malonic semialdehyde, yielding a hemithioacetal. Oxidation by NAD^+ , followed by nucleophilic acyl substitution by CoA, gives malonyl CoA.



28.35 The steps: (1) phosphorylation; (2) cyclization; (3) loss of phosphate; (4) tautomerization.



28.36



28.37 Both steps are nucleophilic acyl substitutions.

