

MUTATION

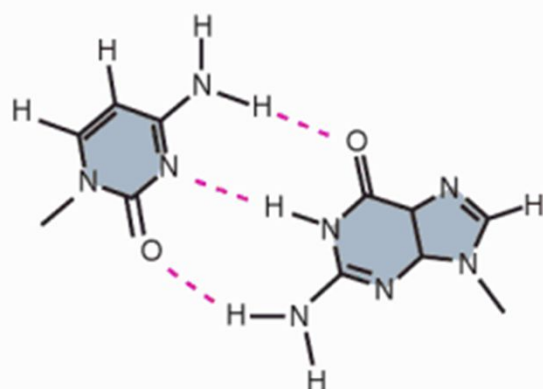
- Changes in the **nucleotide sequence** of DNA
- May occur in **somatic cells** (aren't passed to offspring)
- May occur in **gametes** (eggs & sperm) and be passed to offspring

MUTATIONS

- Point mutations
- Chromosome mutations

Point Mutations

- Change of a **single** nucleotide
- Includes the deletion, insertion, or substitution of **ONE** nucleotide in a gene



Mutation rate

Any base pair
1 in 10^9 – 10^{10}
generations



Any gene
1 in 10^5 – 10^6
generations



The genome
1 in 300
generations

FIGURE 1.25 A base pair is mutated at a rate of 10^{-9} – 10^{-10} per generation, a gene of 1000 bp is mutated at $\sim 10^{-6}$ per generation, and a bacterial genome is mutated at 3×10^{-3} per generation.

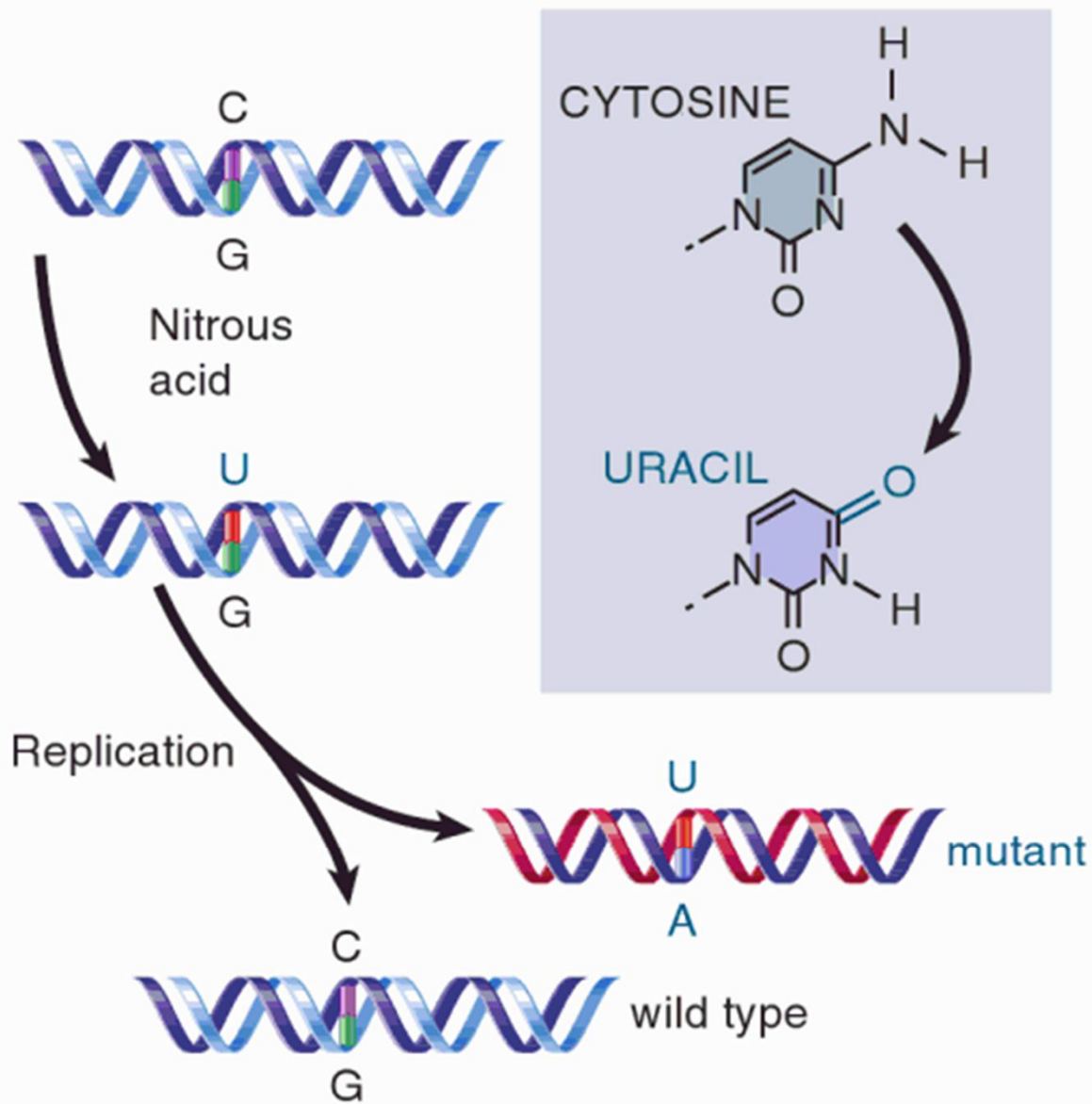


FIGURE 1.26 Mutations can be induced by chemical modification of a base.

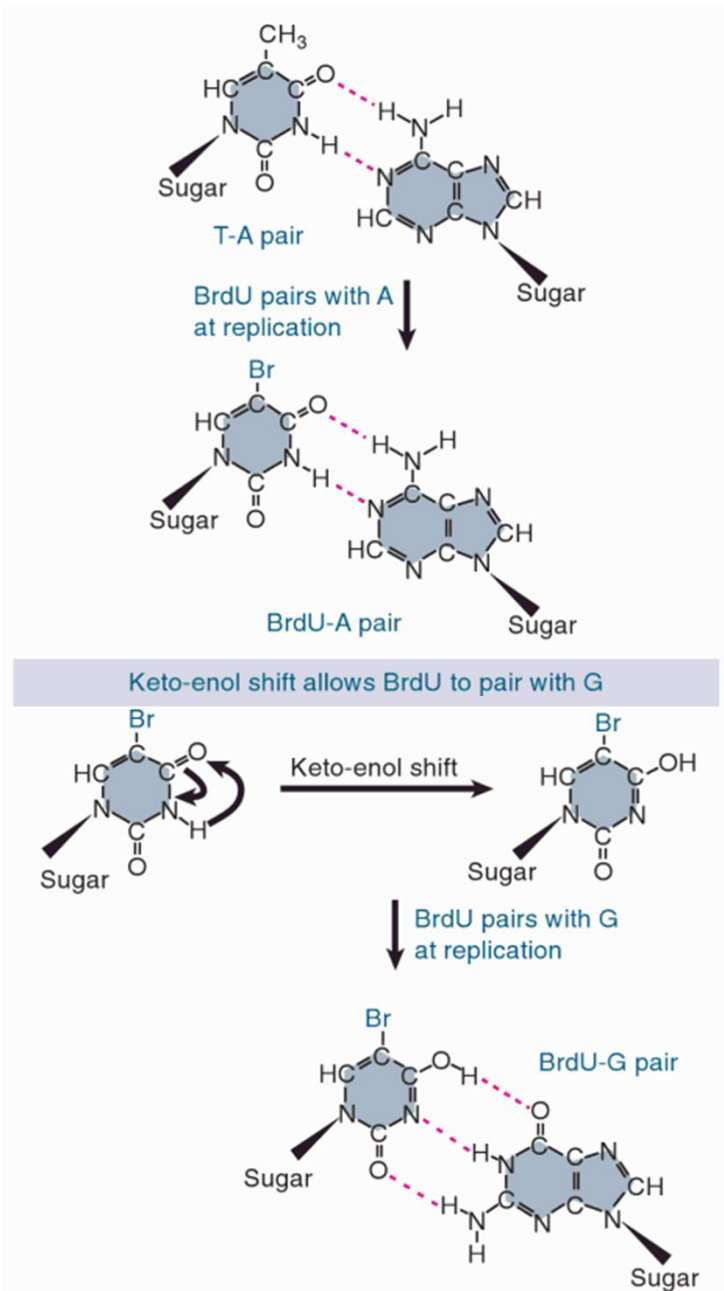


FIGURE 1.27 Mutations can be induced by the incorporation of base analogs into DNA.

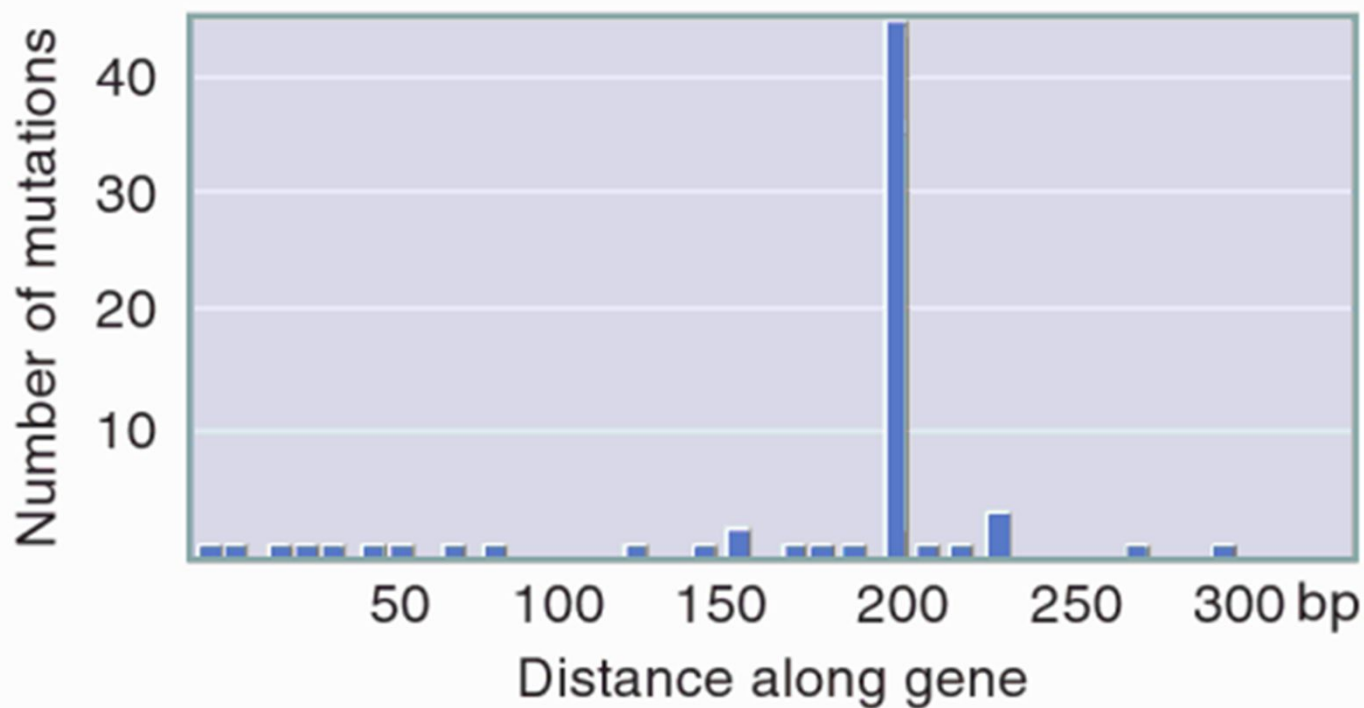


FIGURE 1.29 Spontaneous mutations occur throughout the *lacI* gene of *E. coli*, but are concentrated at a hotspot.

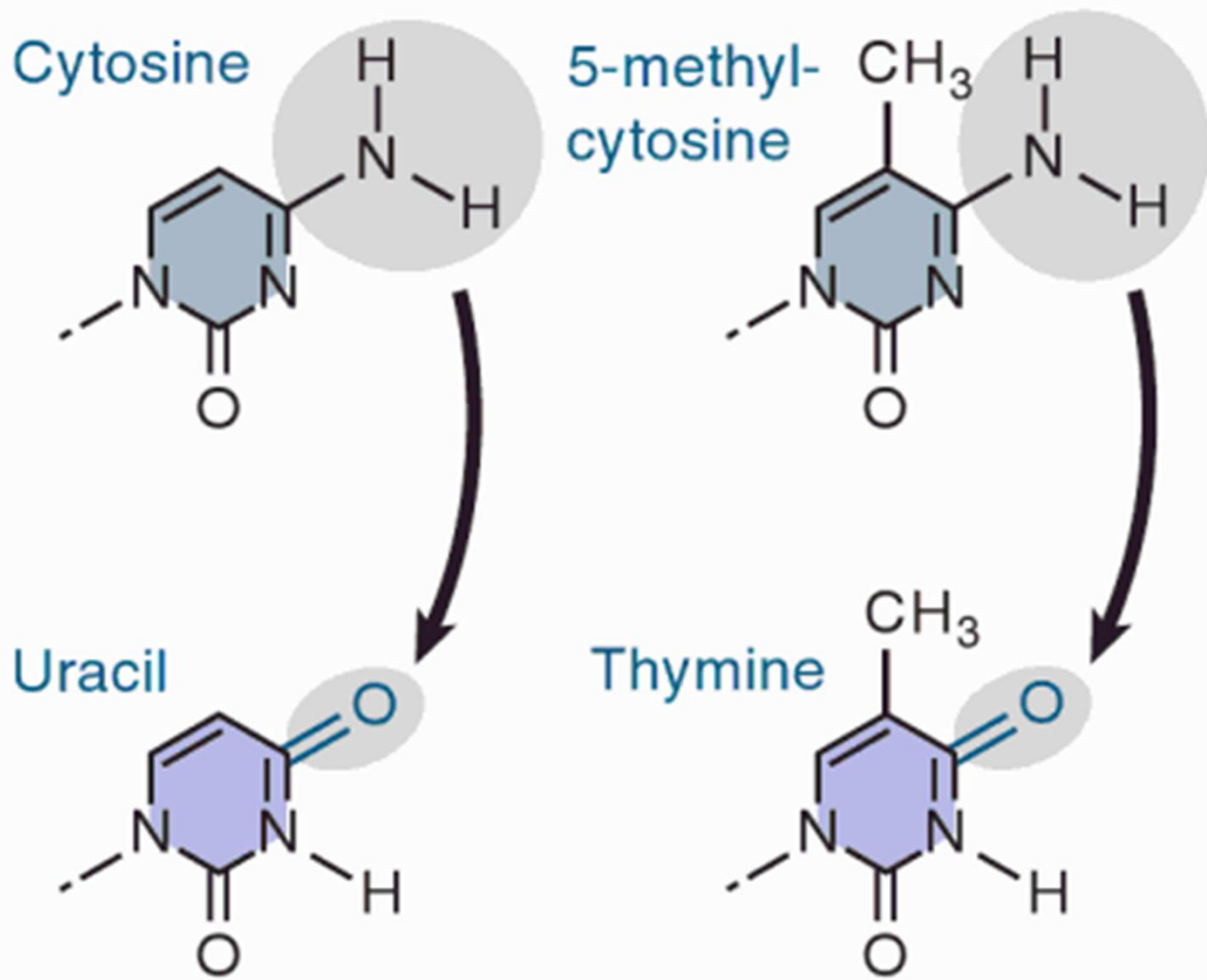


FIGURE 1.30 Deamination of cytosine produces uracil, whereas deamination of 5-methylcytosine produces thymine.

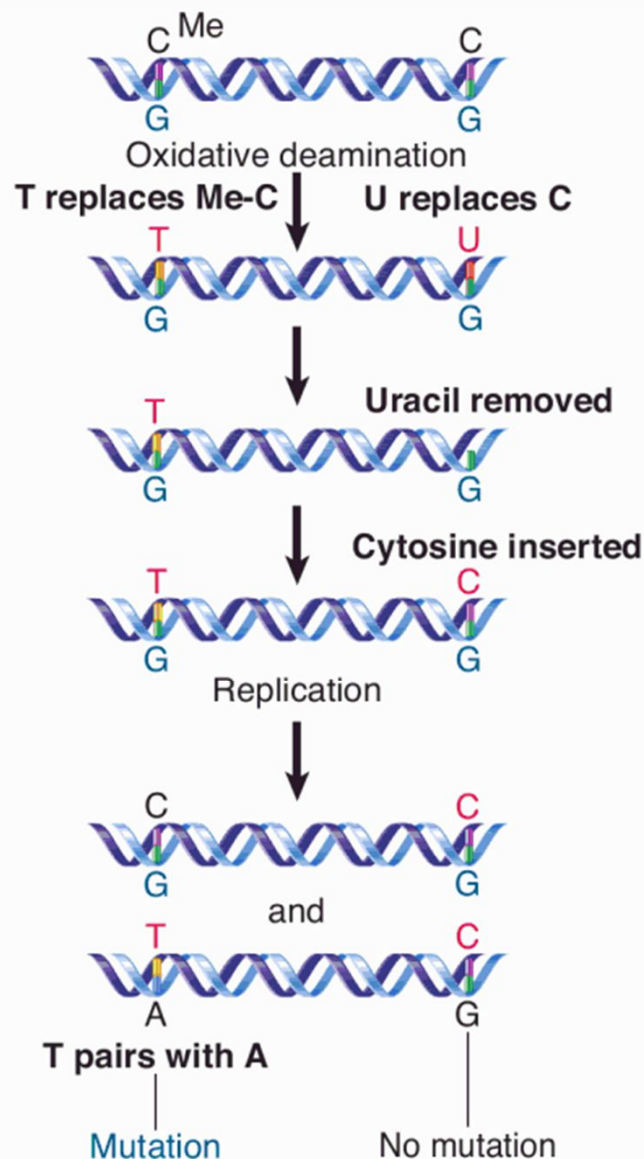


FIGURE 1.31 The deamination of 5-methylcytosine produces thymine (by C-G to T-A transitions), whereas the deamination of cytosine produces uracil (which usually is removed and then replaced by cytosine).

Chromosome mutations

- Inversions
- Insertions and deletions
- Translocations

What is a GENE?

A GENE is defined as the whole nucleotide sequence needed for the synthesis of a functional polypeptide

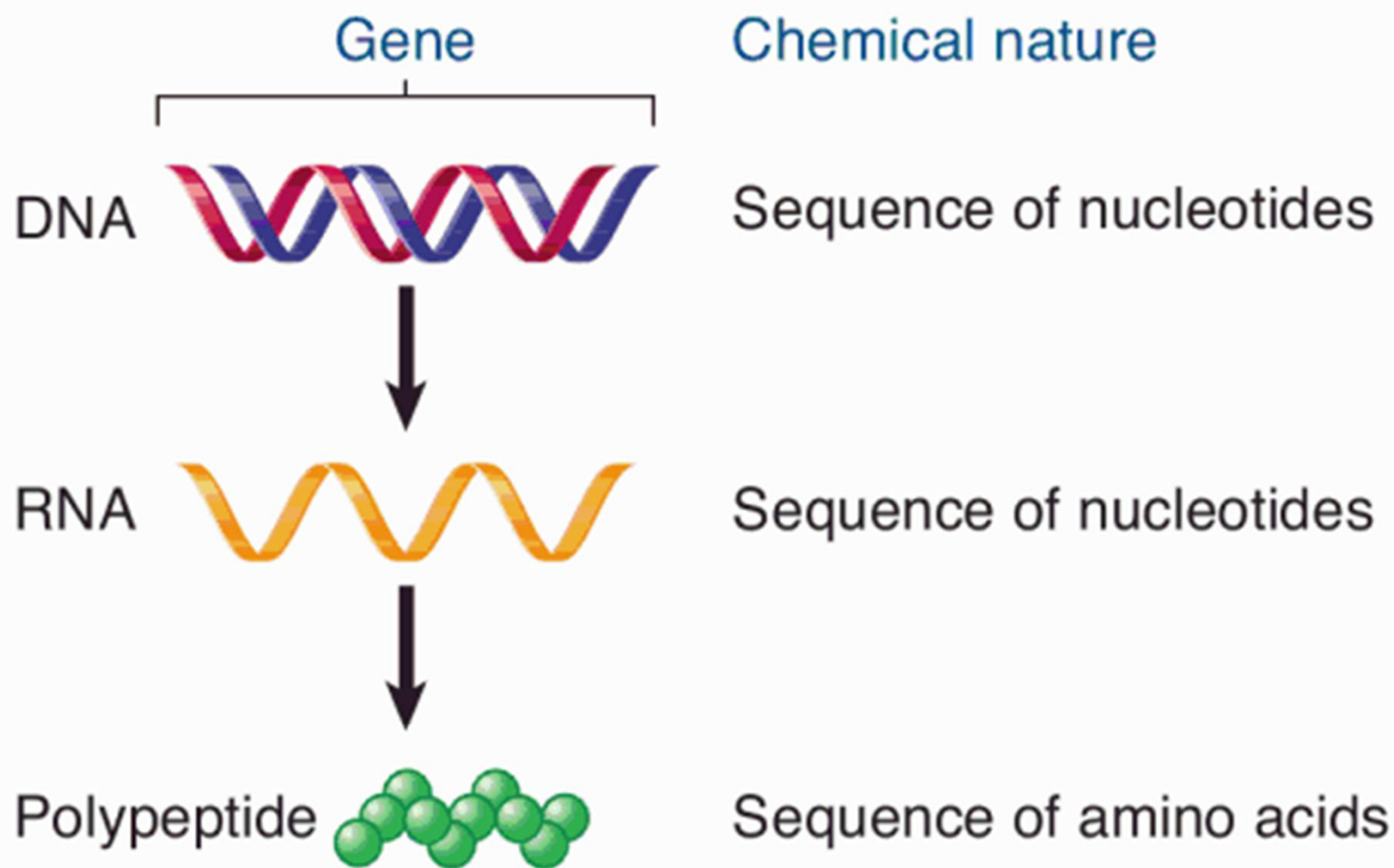
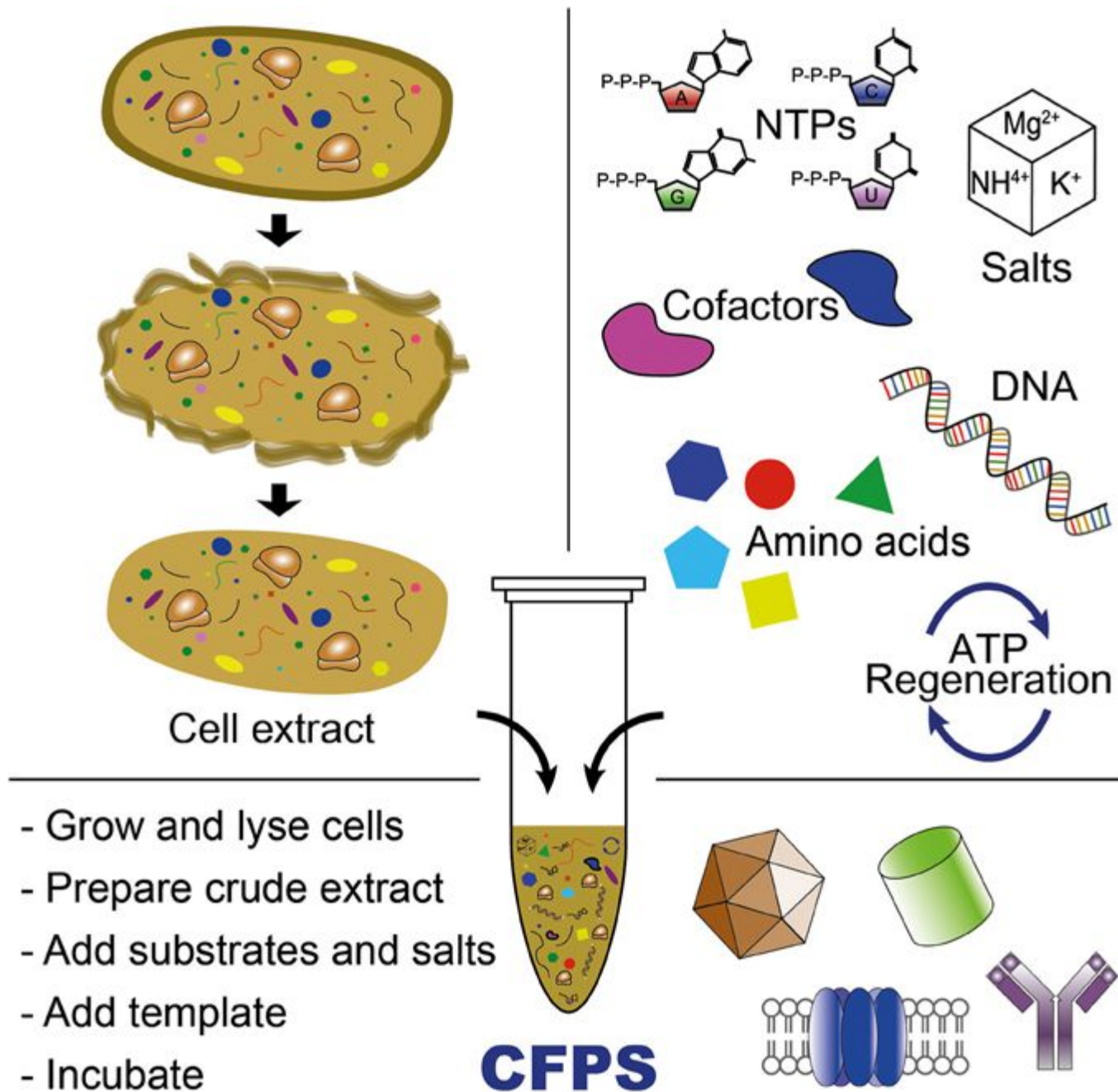


FIGURE 1.2 A gene codes for an RNA, which may code for protein.



The genetic code

What is the right way to come from the information coded in DNA to the amino acid sequence of the codified protein?

4 bases → 20 amino acids

		Second base			
		U	C	A	G
First base	U	UUU } Phe UUC } F UUA } Leu UUG } L	UCU } UCC } Ser UCA } S UCG }	UAU } Tyr UAC } Y UAA } STOP UAG } STOP	UGU } Cys UGC } C UGA } STOP UGG } Trp W
	C	CUU } CUC } Leu CUA } L CUG }	CCU } CCC } Pro CCA } P CCG }	CAU } His CAC } H CAA } Gln CAG } Q	CGU } CGC } Arg CGA } R CGG }
	A	AUU } AUC } Ile AUA } AUG } Met M	ACU } ACC } Thr ACA } T ACG }	AAU } Asn AAC } N AAA } Lys AAG } K	AGU } Ser AGC } S AGA } Arg AGG } R
	G	GUU } GUC } Val GUA } V GUG }	GCU } GCC } Ala GCA } A GCG }	GAU } Asp GAC } D GAA } Glu GAG } E	GGU } GGC } Gly GGA } G GGG }

FIGURE 25.1 All the triplet codons have meaning: 61 represent amino acids and 3 cause termination (stop codons).

Assegnazione dei codoni mediante uso di copolimeri ripetuti contenenti due o tre nucleotidi

Copolimero	Codoni presenti	Amminoacidi incorporati	Codoni assegnati
(CU) _n	CUC/UCU/CUC...	leucina serina	CUC UCU
(UG) _n	UGU/GUG/UGU...	cisteina valina	UGU GUG
(AC) _n	ACA/CAC/ACA...	treonina istidina	ACA CAC
(AG) _n	AGA/GAG/AGA...	arginina glutamina	AGA GAG
(AUC) _n	AUC/AUC/AUC... UCA/UCA/UCA... CAU/CAU/CAU...	poliisoleucina poliserina poliistidina	AUC UCA CAU

UUU UUC	UCU UCC	UAU UAC	UGU UGC
UUA UUG	UCA UCG	UAA UAG	UGA [UGG]
CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC CAA CAG	CGU CGC CGA CGG
AUU AUC AUA AUG	ACU ACC ACA ACG	AAU AAC AAA AAG	AGU AGC AGA AGG
GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC GAA GAG	GGU GGC GGA GGG

Third-base relationship

Third bases
with same
meaning

Codon
number

	} Third base irrelevant	U, C, A, G	32
		U, C, A	3
	} Purines differ from pyrimidines	A or G	14
		U or C	10
	Unique	G only	2

FIGURE 25.3 Third bases have the least influence on codon meanings. Boxes indicate groups of codons within which third-base degeneracy ensures that the meaning is the same.

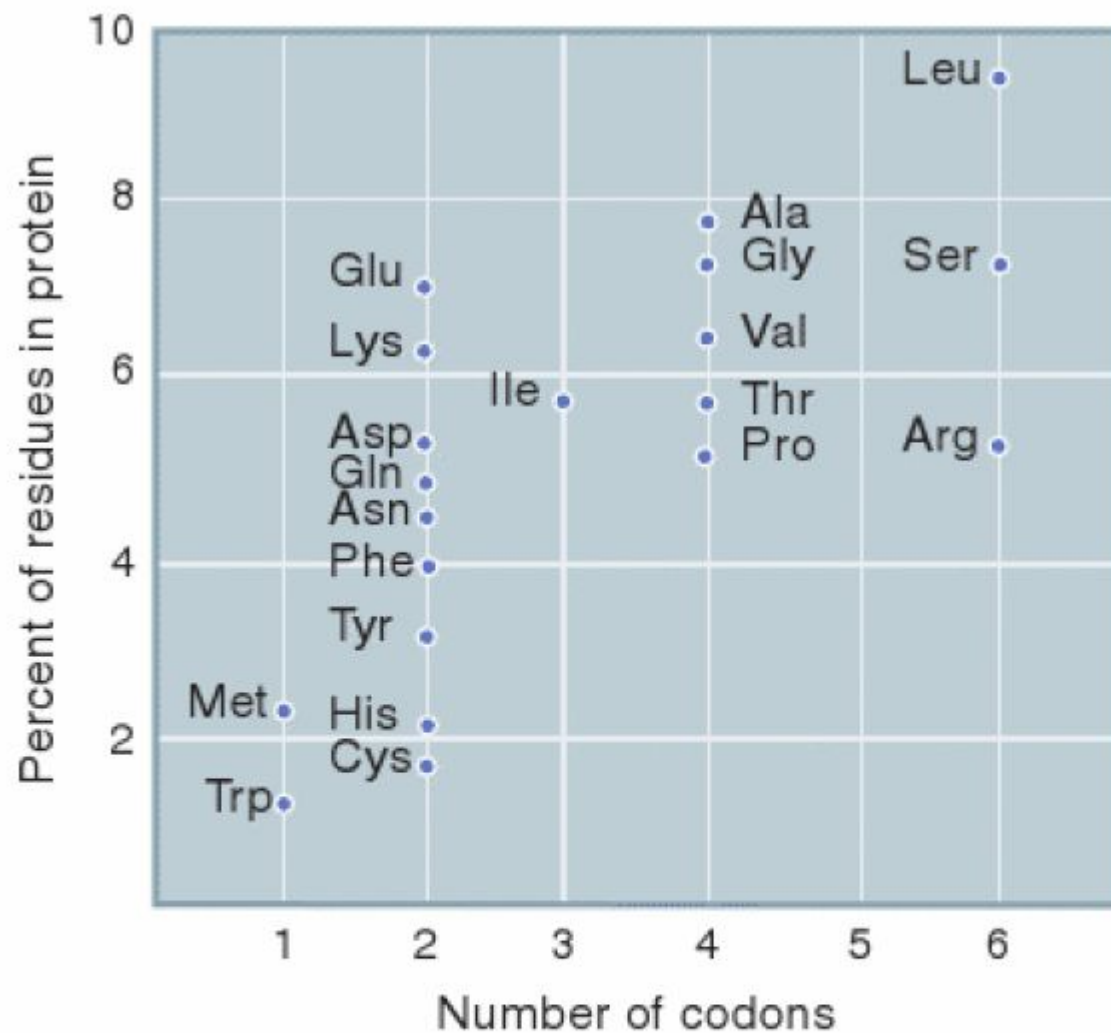


FIGURE 25.2 Some correlation of the frequency of amino acid use in proteins with the number of codons specifying the amino acid is observed. An exception is found for amino acids specified by two codons, which occur with a wide variety of frequencies.

Genetic code

- It is degenerated
- It is not ambiguous

Genetic code is universal

Exceptions:

UGA = Trp in Mycoplasma (bacteria)

UAA, UAG = Gln in Tetrahymena (protozoa)

mtDNA Genetic code in Mammals

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

mtDNA Genetic Code variation for mammals, fruit flies and yeasts

RNA Codon	Nuclear Genetic Code	mtDNA Genetic Code		
		Mammals	Drosophila	Yeasts
UGA	STOP	Tryptophan	Tryptophan	Tryptophan
AGA, AGG	Arginine	STOP	Serine	Arginine
AUA	Isoleucine	Methionine	Methionine	Methionine
AUU	Isoleucine	Methionine	Methionine	Methionine
CUU, CUC, CUA, CUG	Leucine	Leucine	Leucine	Threonine

Codon Usage for *E.coli* and *H.sapiens*

		Second Letter																
		U			C			A			G							
First Letter	U	Phe	UUU	16	14	Ser	UCU	46	7	Tyr	UAU	18	17	Cys	UGU	30	21	U
			UUC	84	86		UCC	32	36		UAC	82	83		UGC	70	79	C
	Leu	UUA	2	1		UCA	1	4	Stop	UAA	-	-	Stop	UGA	-	-	A	
		UUG	2	5		UCG	0	11	Stop	UAG	-	-	Trp	UGG	100	100	G	
C		Leu	CUU	2	3	Pro	CCU	4	15	His	CAU	22	15	Arg	CGU	74	6	U
			CUC	3	23		CCC	0	51		CAC	78	85		CGC	26	45	C
			CUA	1	3		CCA	11	9	Gln	CAA	6	10		CGA	0	5	A
			CUG	90	65		CCG	85	25		CAG	94	90		CGG	0	25	G
A	Ile	AUU	16	15	Thr	ACU	47	9	Asn	AAU	1	18	Ser	AGU	1	5	U	
		AUC	83	81		ACC	48	59		AAC	99	82		AGC	20	38	C	
		AUA	1	4		ACA	2	9	Lys	AAA	81	17	Arg	AGA	0	4	A	
	Met	AUG	100	100		ACG	3	23		AAG	19	83		AGG	0	15	G	
G	Val	GUU	60	3	Ala	GCU	51	11	Asp	GAU	26	18	Gly	GGU	57	9	U	
		GUC	2	26		GCC	4	60		GAC	74	82		GGC	42	58	C	
		GUA	28	2		GCA	27	9	Glu	GAA	82	13		GGA	0	7	A	
		GUG	10	69		GCG	18	20		GAG	18	87		GGG	1	26	G	

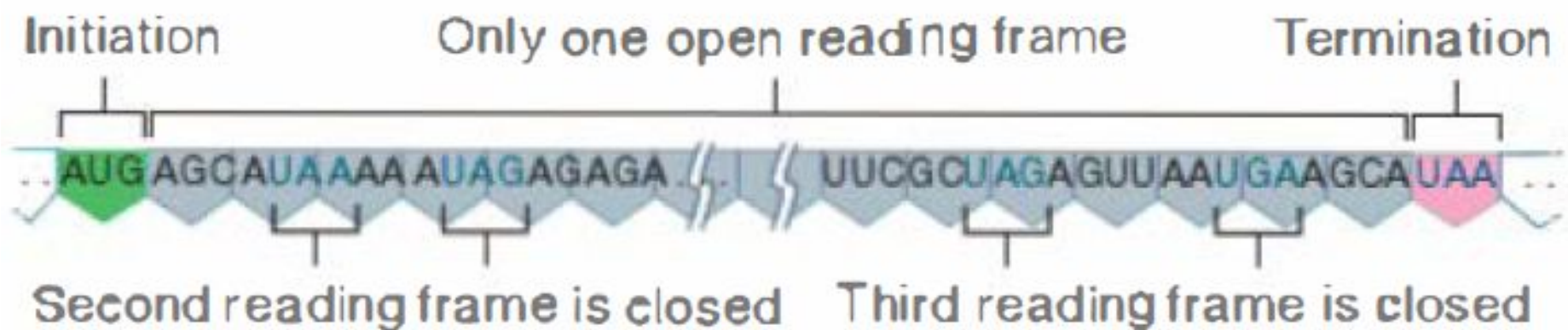
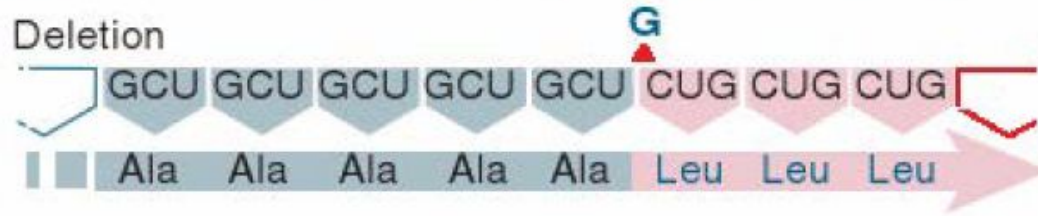
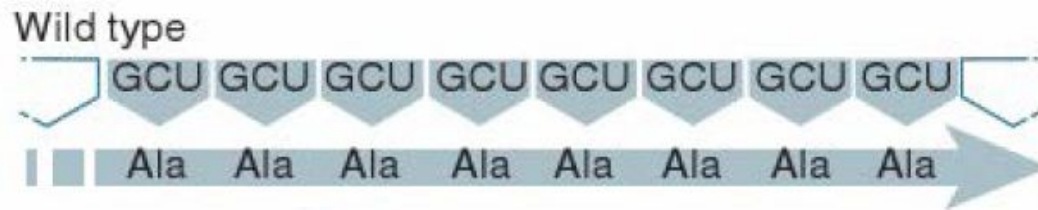


FIGURE 2.11 An open reading frame starts with AUG and continues in triplets to a termination codon. Closed reading frames may be interrupted frequently by termination codons.



■ Wild-type sequence ■ Mutant sequence

ACRIDINE treated DNA
(molecule causing
frameshift mutations)
shows that the genetic
code is read in
nonoverlapping triplets
from a fixed starting
point (Met).

FIGURE 2.10 Frameshift mutations show that the genetic code is read in triplets from a fixed starting point.