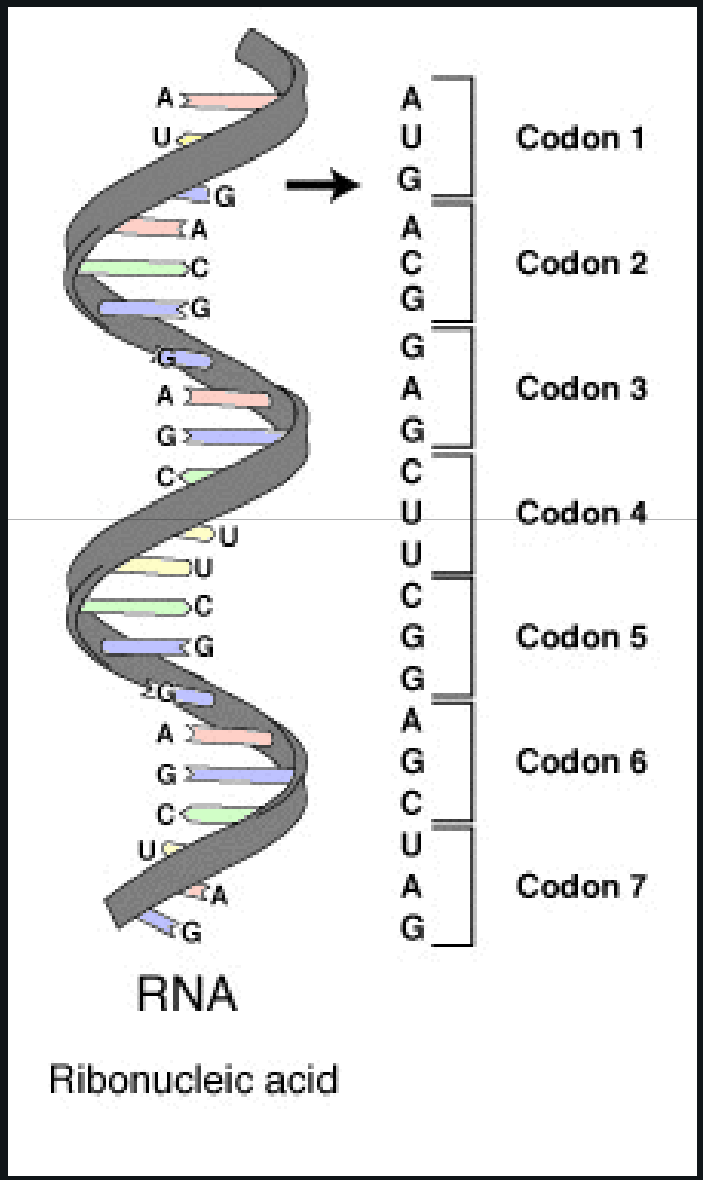
Genetic Code

By

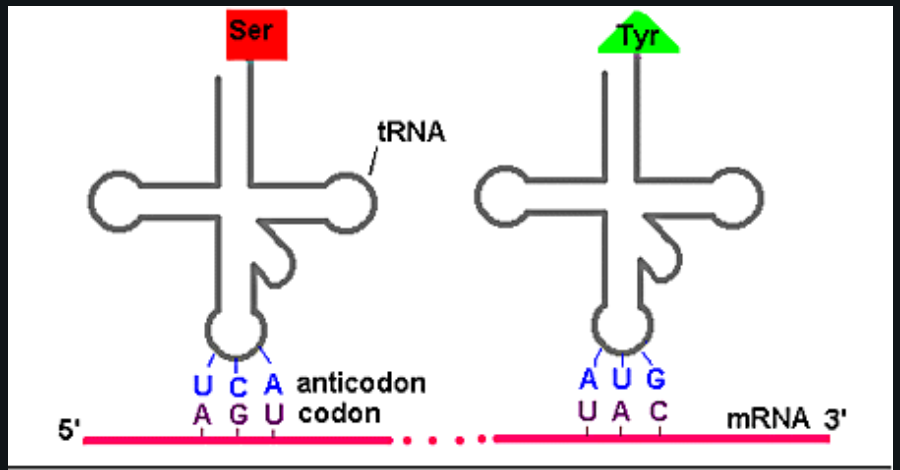
PRAMOD KUMAR MAHISH
Asst. Professor (Biotechnology)
Govt. Digvijay P. G. College Rajnandgaon (C.G.)

Definition

- The genetic code is the set
- in which information encoded within genetic material (DNA or mRNA sequences)
- is translated into proteins (amino acid sequences) by living cells.


A
U
G
A
C
G
G
G
A
G
C
U
U
C
G
G
A
G
C
U
A
G
Codon 1
Codon 2
Codon 3
Codon 4
Codon 5
Codon 6
Codon 7
RNA
Ribonucleic acid

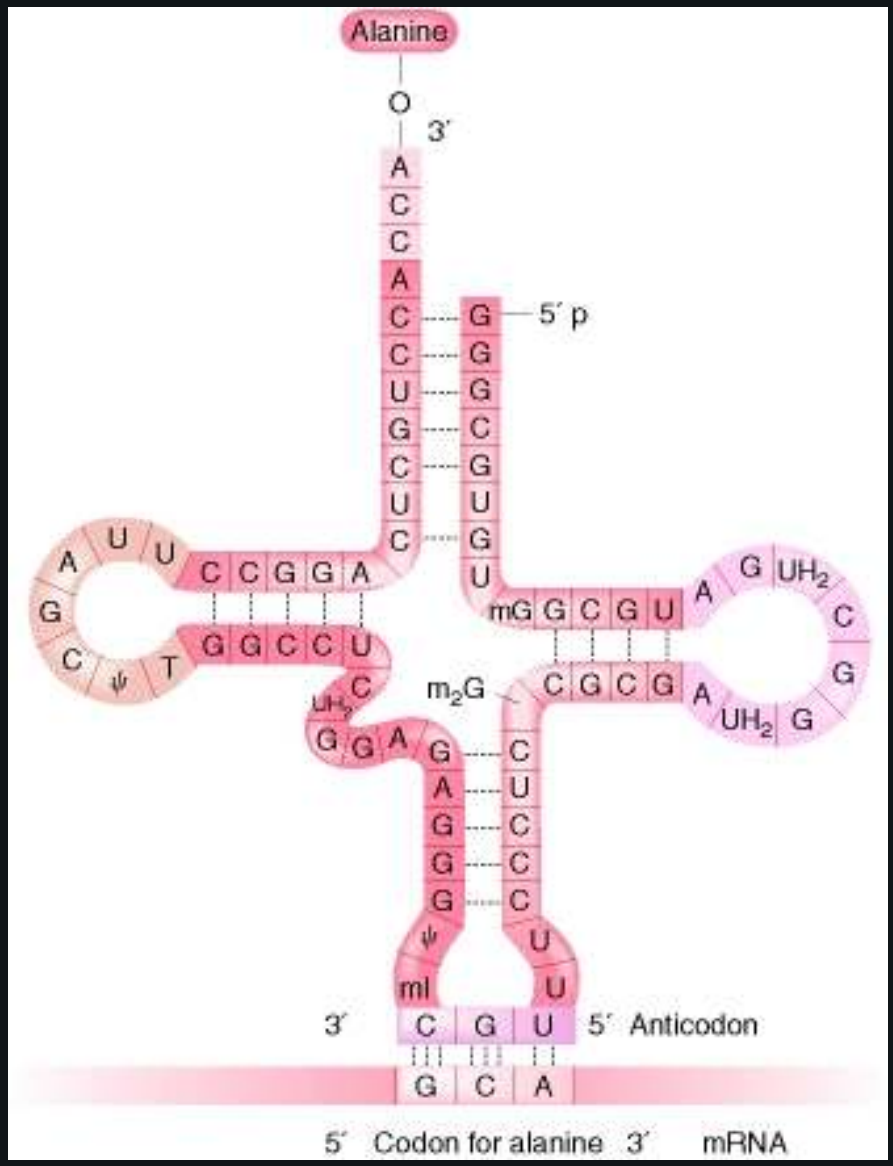
- Biological decoding is accomplished by the ribosome, which links amino acids in an order specified by mRNA, using transfer RNA (tRNA) molecules to carry amino acids and to read the mRNA three nucleotides at a time.

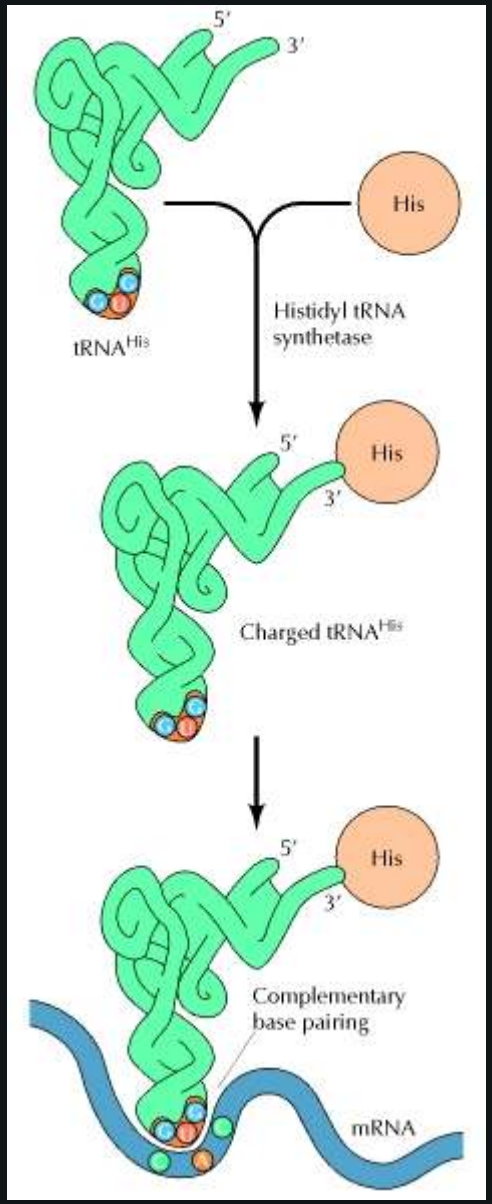


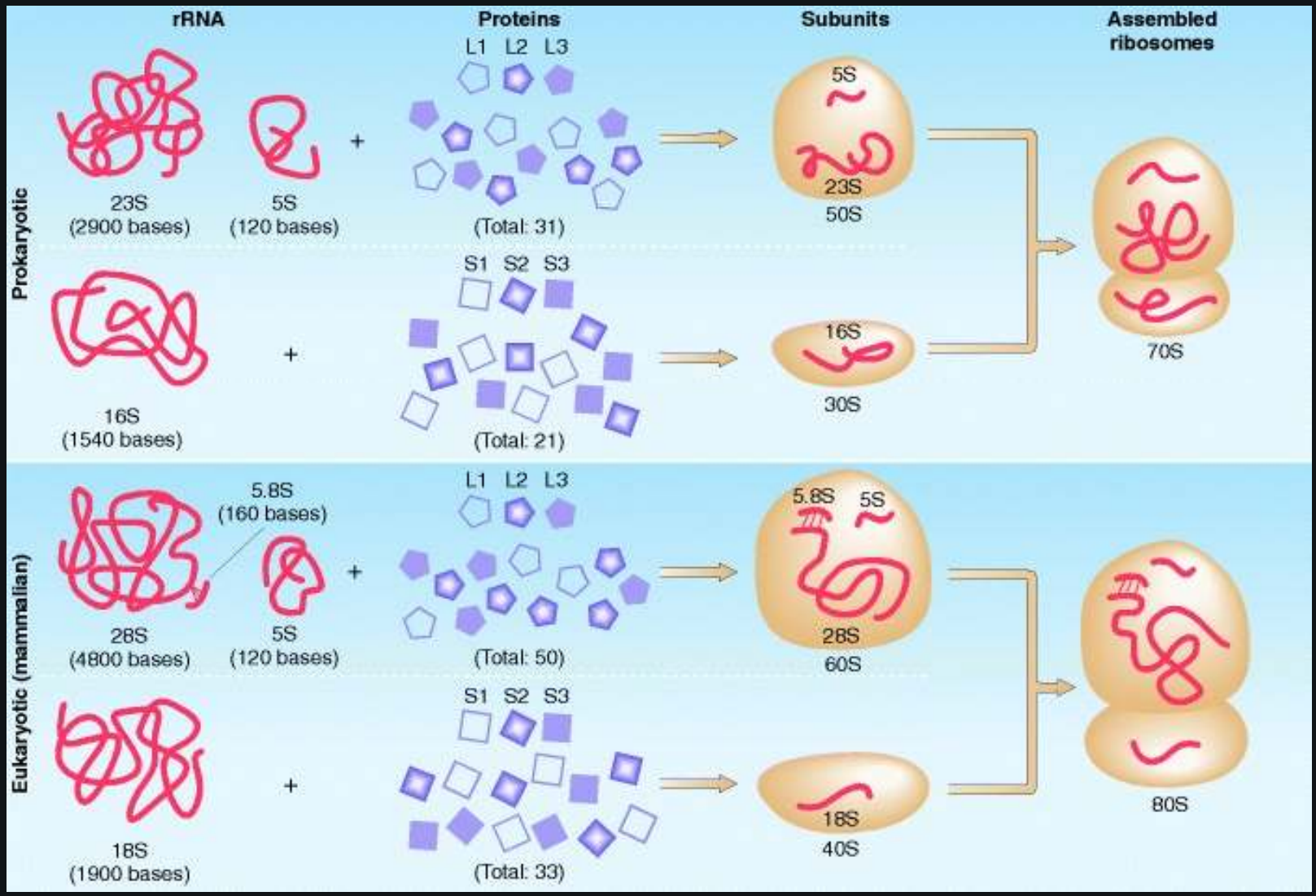
2nd base in codon

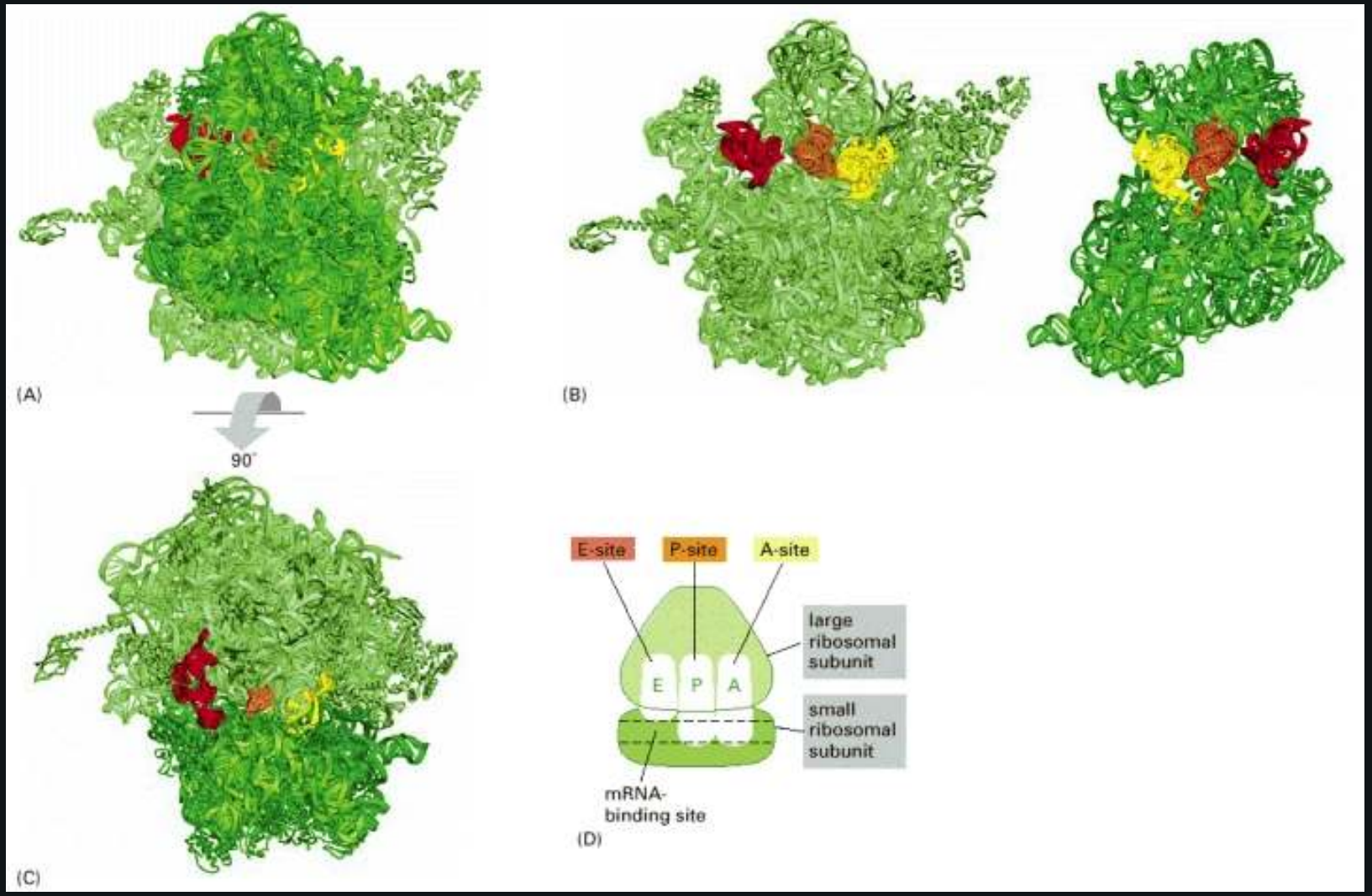
1st base in codon	U	C	A	G	3rd base in codon
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

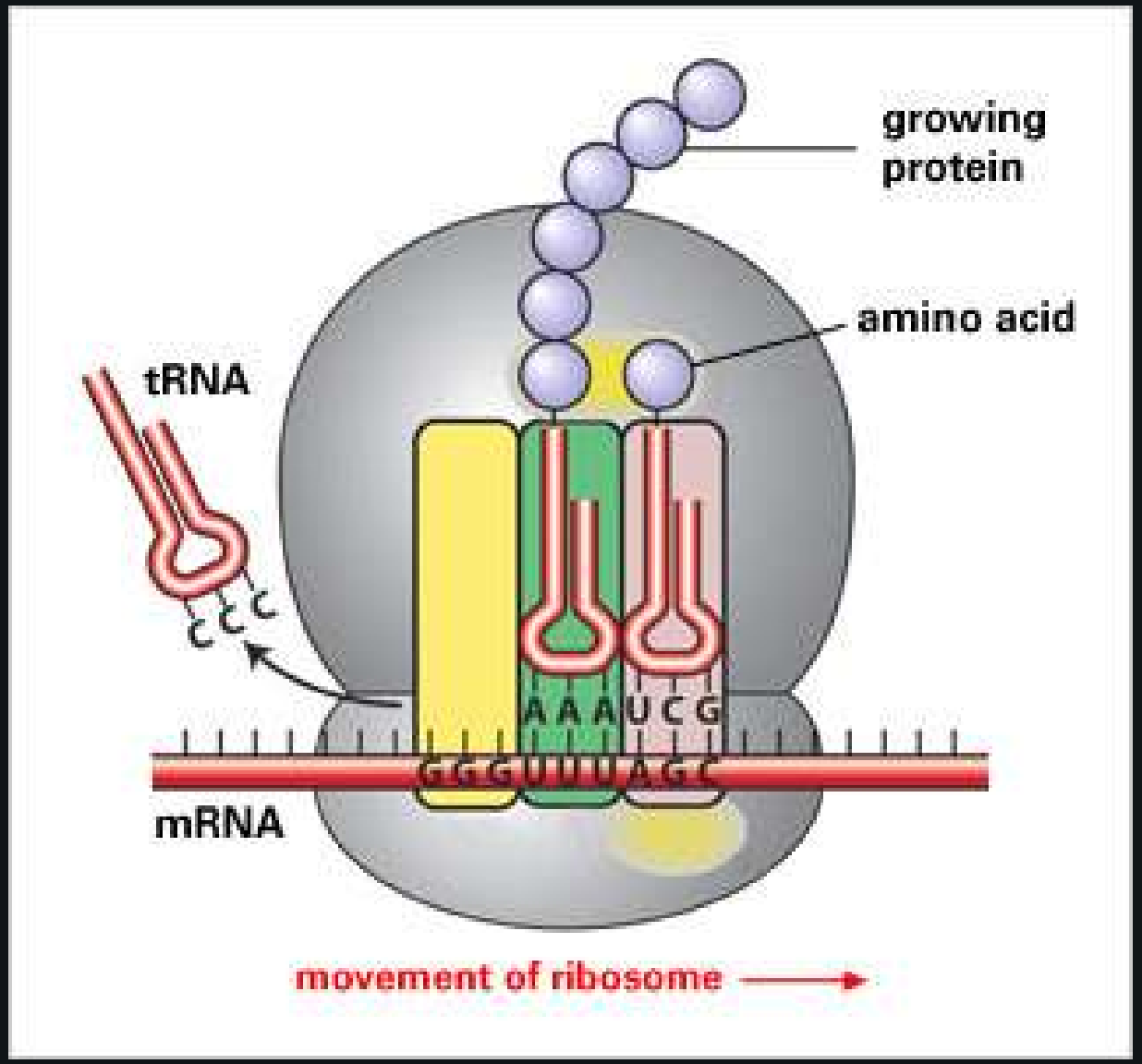
The Genetic Code

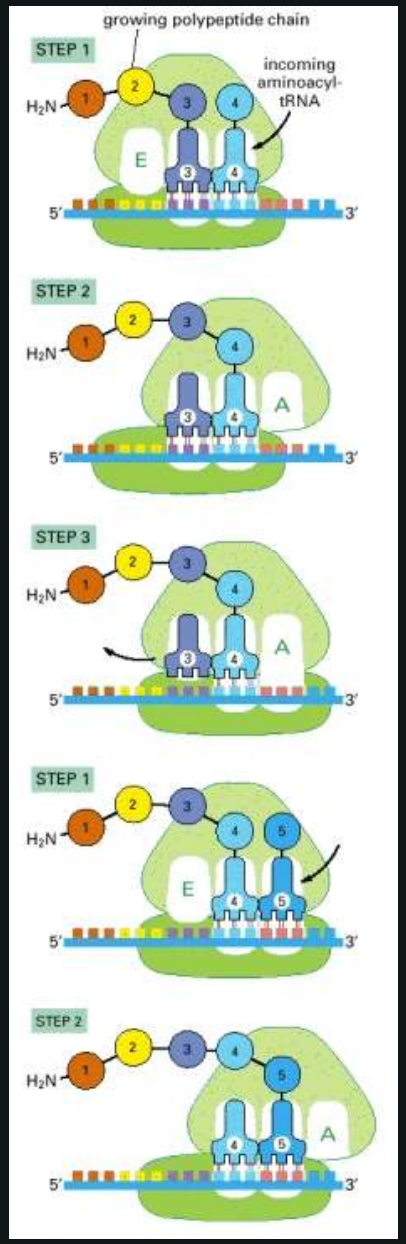

Alanine
O
3′
A
C
C
A
C
C
C
U
G
C
C
U
C
G
G
G
G
G
C
G
U
U
G
U
5′ p
A
U
U
C C G G A
G
C
ψ
T
G G C C U
UH2 C
G G A
mG G C G U
A
G UH2
C
G
C G C G
A
UH2 G
m2G
G
A
G
G
G
ψ
ml
C
U
C
C
C
U
U
3′
C G U
5′ Anticodon
G C A
5′ Codon for alanine 3′ mRNA

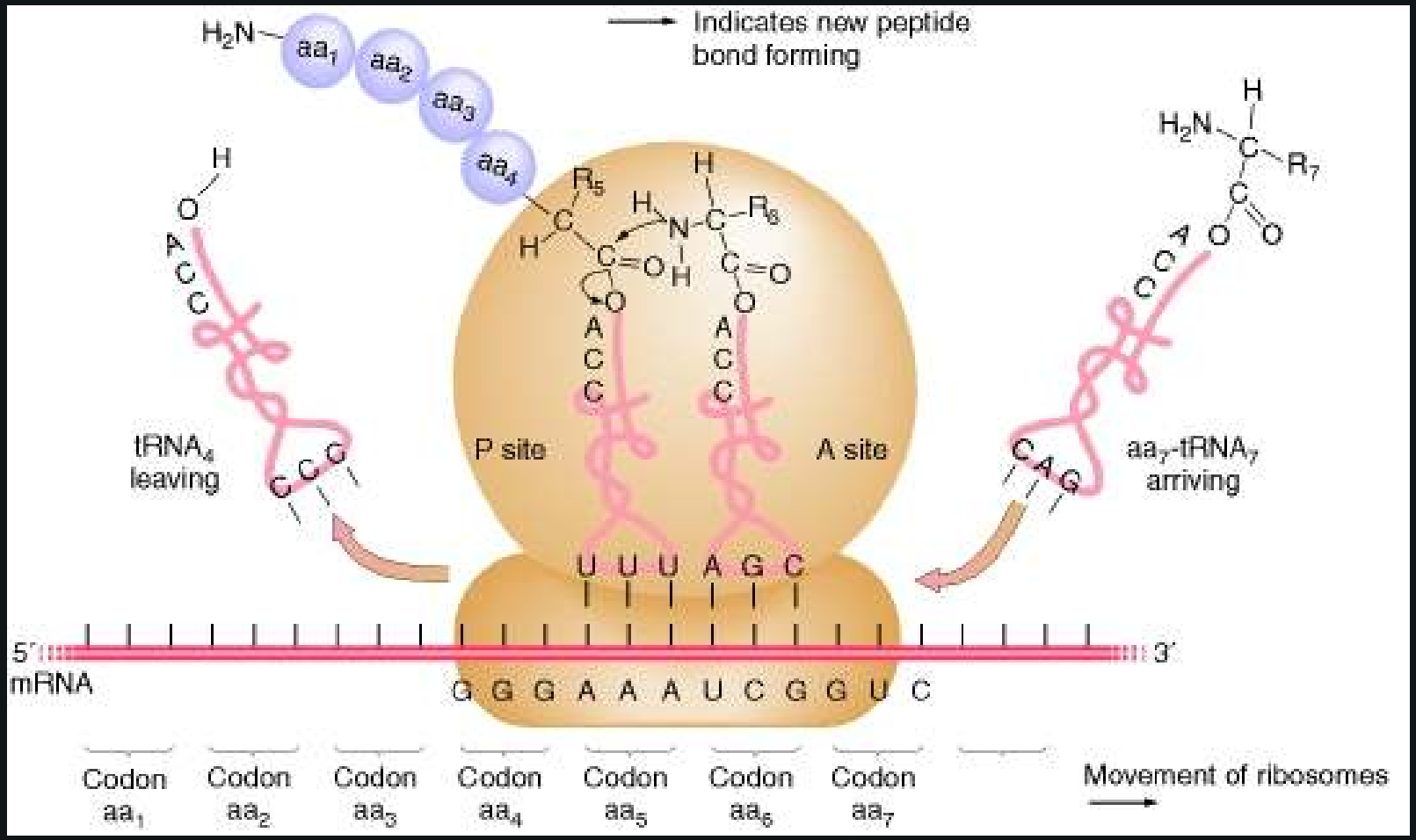

5′
3′
His
Histidyl tRNA synthetase
tRNA^{His}
5′
His
3′
Charged tRNA^{His}
5′
His
3′
Complementary base pairing
mRNA


rRNA
Proteins
Subunits
Assembled ribosomes
Prokaryotic
23S
(2900 bases)
5S
(120 bases)
L1 L2 L3
(Total: 31)
5S
23S
50S
16S
(1540 bases)
S1 S2 S3
(Total: 21)
16S
30S
70S
Eukaryotic (mammalian)
5.8S
(160 bases)
28S
(4800 bases)
5S
(120 bases)
L1 L2 L3
(Total: 50)
5.8S
5S
28S
60S
18S
(1900 bases)
S1 S2 S3
(Total: 33)
18S
40S
80S


(A)
90°
(B)
(C)
E-site
P-site
A-site
large ribosomal subunit
E
P
A
small ribosomal subunit
mRNA-binding site
(D)


growing protein
amino acid
tRNA
A A A U C G
G G G U U U A G C
mRNA
movement of ribosome

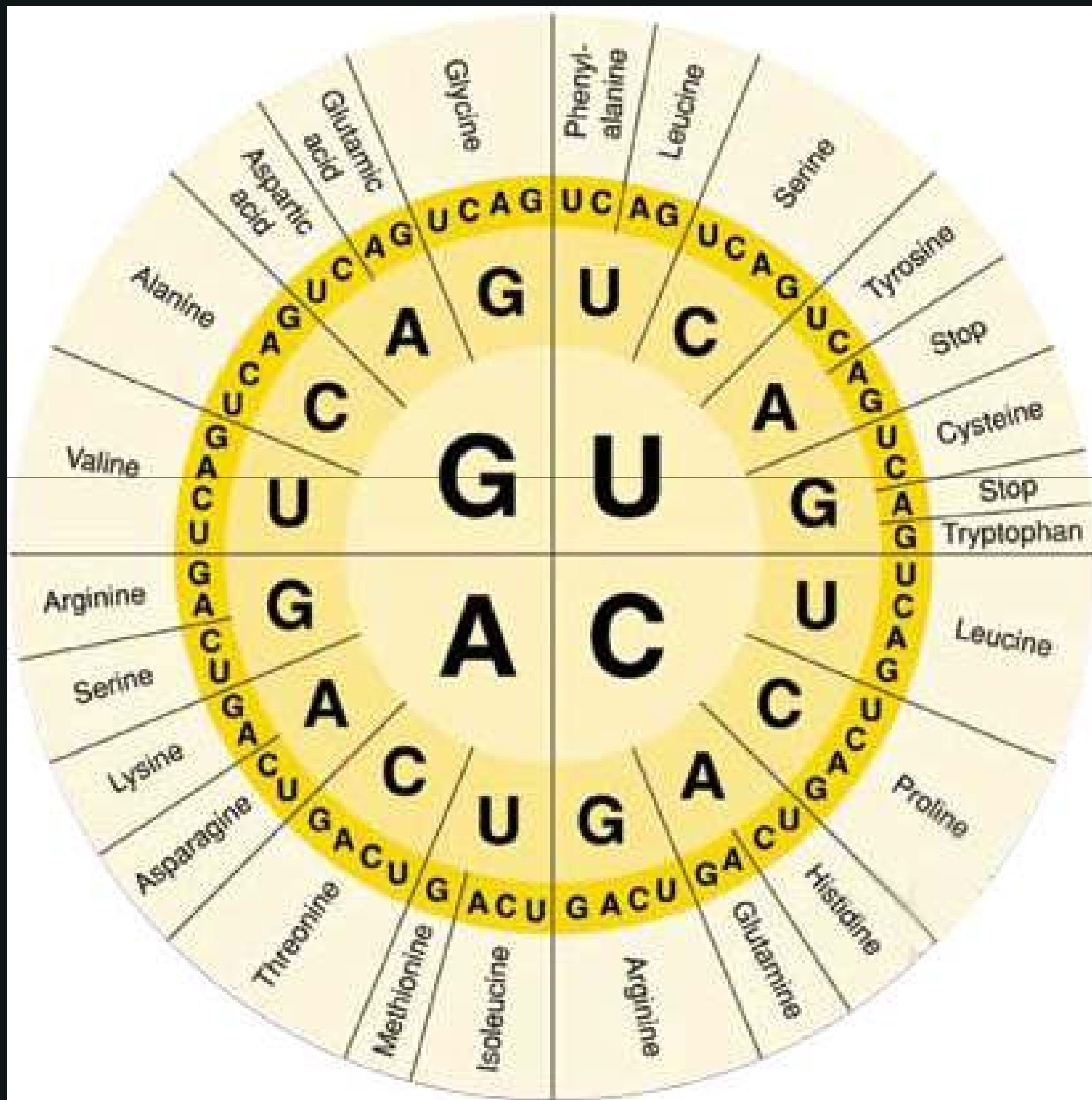

growing polypeptide chain
STEP 1
incoming aminoacyl-tRNA
H_2N
E
5′
3′
STEP 2
A
STEP 3
STEP 1
STEP 2


Indicates new peptide bond forming
tRNA4 leaving
P site
A site
aa7-tRNA7 arriving
5′
mRNA
3′
U U U A G C
G G G A A A U C G G U C
Codon aa1
Codon aa2
Codon aa3
Codon aa4
Codon aa5
Codon aa6
Codon aa7
Movement of ribosomes

nonpolar | polar | basic | acidic | (stop codon)

Standard genetic code

1st base	2nd base								3rd base
	U		C		A		G		
U	UUU	(Phe/F) Phenylalanine	UCU	(Ser/S) Serine	UAU	(Tyr/Y) Tyrosine	UGU	(Cys/C) Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA		UCA		UAA	Stop (*Ochre*)	UGA	Stop (*Opal*)	A
	UUG		UCG		UAG	Stop (*Amber*)	UGG	(Trp/W) Tryptophan	G
C	CUU	(Leu/L) Leucine	CCU	(Pro/P) Proline	CAU	(His/H) Histidine	CGU	(Arg/R) Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	(Gln/Q) Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	(Ile/I) Isoleucine	ACU	(Thr/T) Threonine	AAU	(Asn/N) Asparagine	AGU	(Ser/S) Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine	A
	AUG[A]	(Met/M) Methionine	ACG		AAG		AGG		G
G	GUU	(Val/V) Valine	GCU	(Ala/A) Alanine	GAU	(Asp/D) Aspartic acid	GGU	(Gly/G) Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	(Glu/E) Glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G


G U A C
Phenyl-alanine
Leucine
Serine
Tyrosine
Stop
Cysteine
Stop
Tryptophan
Leucine
Proline
Histidine
Glutamine
Arginine
Isoleucine
Methionine
Threonine
Asparagine
Lysine
Serine
Arginine
Valine
Alanine
Aspartic acid
Glutamic acid
Glycine

Characteristics of Genetic Code

1. *Composed of 4 nucleotides on mRNA (A,U,G,C)*

2. *Read in triplets (3 nucleotides / codon)*

3. *64 possible codons*
   Start: AUG (Also Met)
   Stop: UAG (Amber); UGA (Opal); UAA (Ochre)
   Amino Acids: 60 Codons + AUG

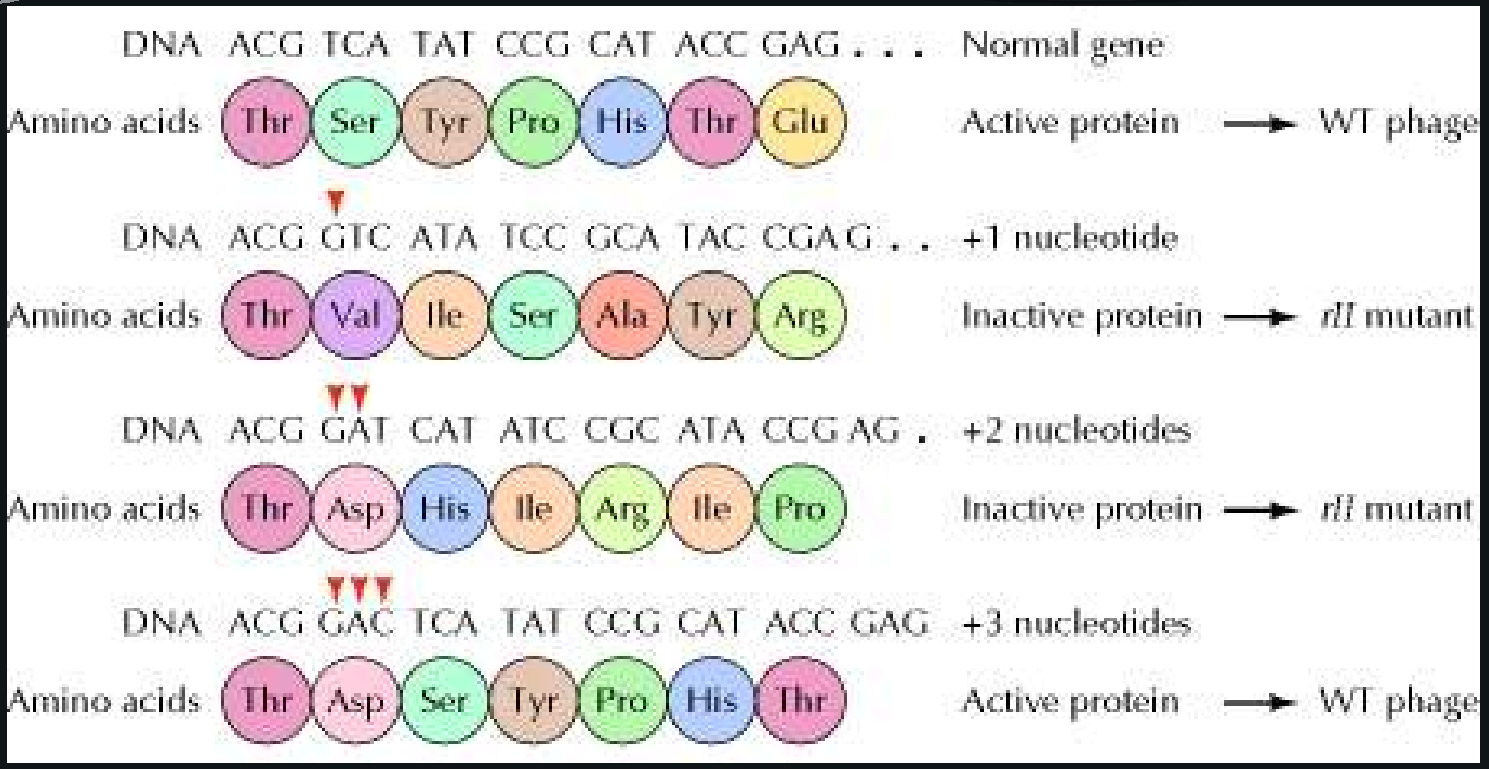
4. *Code is Degenerate: >1 codon / amino acid*

5. *Code is non-overlapping*
   Each nucleotide part of only 1 codon

6. *AUG (Start) Defines Reading Frame*

   5' ------- AUG – $(XXX)^n$ – STOP ---------- 3'

- The genetic code is unambiguous: each codon specifies one amino acid only
- The code is degenerate. One amino acid may be specified by more than one codon. For instance serine is encoded by six codons, glycine by four and lysine by two.
- Note that in most cases sufficient coding is performed by the first two bases, the third (or wobble) base playing a minor role. For instance the four codons that specify glycine (GGU, GGC, GGA and GGG) all start with GG.
-



Genetic evidence for a triplet code

A series of mutations consisting of additions of one, two, or three nucleotides were studied in the *rII* gene of bacteriophage T4. Additions of one or two nucleotides alter the reading frame of the remainder of the gene. Therefore, all the subsequent amino acids are abnormal, and an inactive protein is produced, giving rise to mutant phage. Additions of three nucleotides, however, alter only a single amino acid. The reading frame of the remainder of the gene is normal, and an active protein giving rise to wild-type (WT) phage is produced.

- **The genetic code is composed of nucleotide triplets.** In other words, three nucleotides in mRNA (a codon) specify one amino acid in a protein.
- **The code is non-overlapping.** This means that successive triplets are read in order. Each nucleotide is part of only one triplet codon.
- **The genetic code is unambiguous.** Each codon specifies a particular amino acid, and only one amino acid. In other words, the codon ACG codes for the amino acid threonine, and only threonine.
- **The genetic code is degenerate.** In contrast, each amino acid can be specified by more than one codon.
- **The code is nearly universal.** Almost all organisms in nature (from bacteria to humans) use exactly the same genetic code. The rare exceptions include some changes in the code in mitochondria, and in a few protozoan species.

The Nobel Prize in Physiology or Medicine 1968

Robert W. Holley, H. Gobind Khorana, Marshall W. Nirenberg

The Nobel Prize in Physiology or Medicine 1968

Robert W. Holley

H. Gobind Khorana

Marshall W. Nirenberg

Robert W. Holley

Har Gobind Khorana

Marshall W. Nirenberg

The Nobel Prize in Physiology or Medicine 1968 was awarded jointly to Robert W. Holley, Har Gobind Khorana and Marshall W. Nirenberg *"for their interpretation of the genetic code and its function in protein synthesis"*.

The Genetic Code table proposed in 1965 by Khorana

The Genetic Code

1st letter	2nd letter				3rd letter
	U	C	A	G	
U	PHE	SER	TYR	CYS	U
	PHE	SER	TYR	CYS	C
	LEU	SER	C.T.	NONS	A
	LEU	SER	C.T.	TRY	G
C	LEU	PRO	HIS	ARG	U
	LEU	PRO	HIS	ARG	C
	LEU	PRO	GLN	ARG	A
	LEU	PRO	GLN	ARG	G
A	ILEU	THR	ASN	SER	U
	ILEU	THR	ASN	SER	C
	ILEU	THR	LYS	ARG	A
	MET (C.I.)	THR	LYS	ARG	G
G	VAL	ALA	ASP	GLY	U
	VAL	ALA	ASP	GLY	C
	VAL	ALA	GLU	GLY	A
	VAL (C.I.)	ALA	GLU	GLY	G

Thank you