GENETIC CODE

Dinesh D. Khedkar

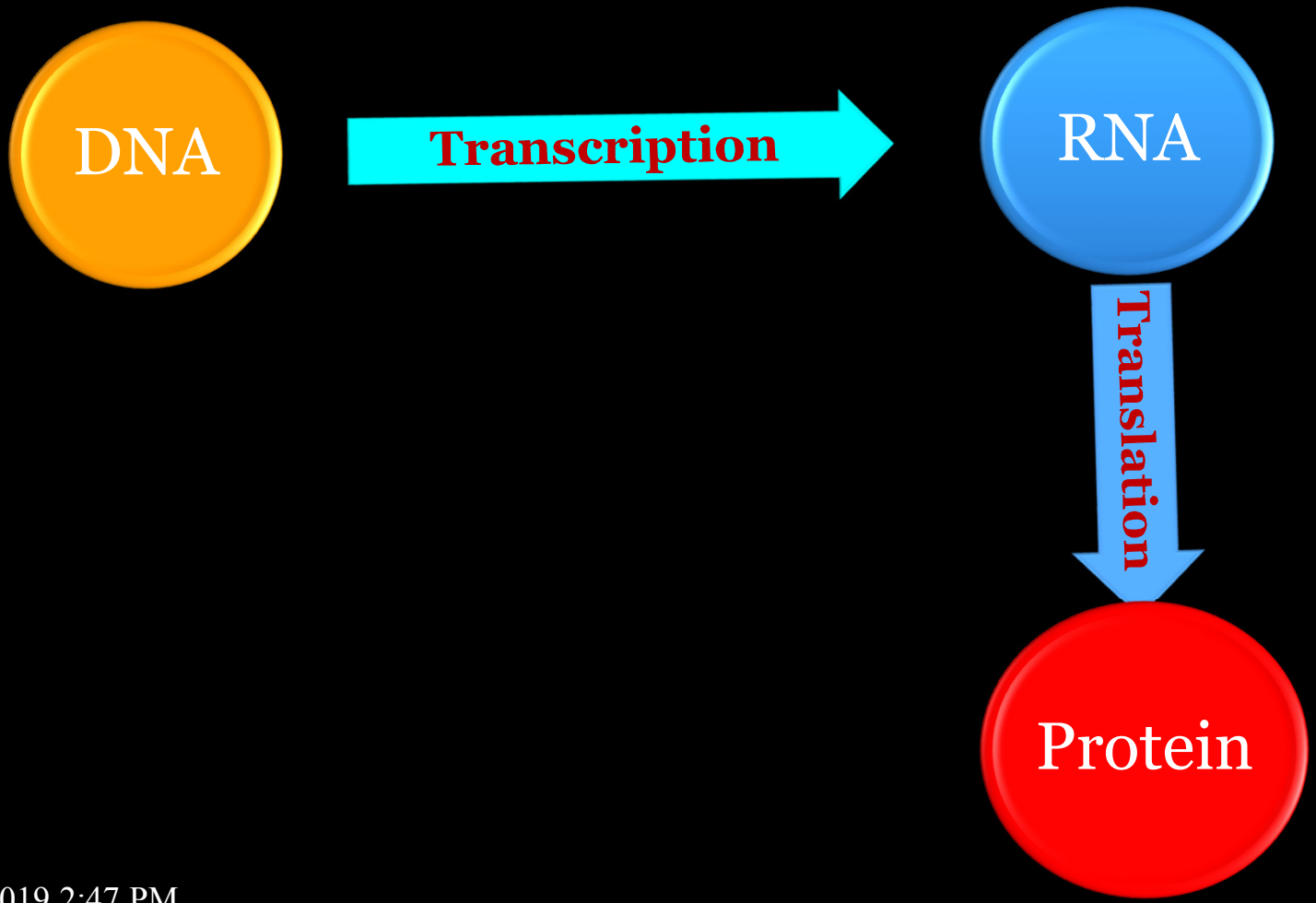
Department of Botany

A UGC Awarded College with Potential for Excellence
NAAC Reaccredited "A" - "Very Good" Grade

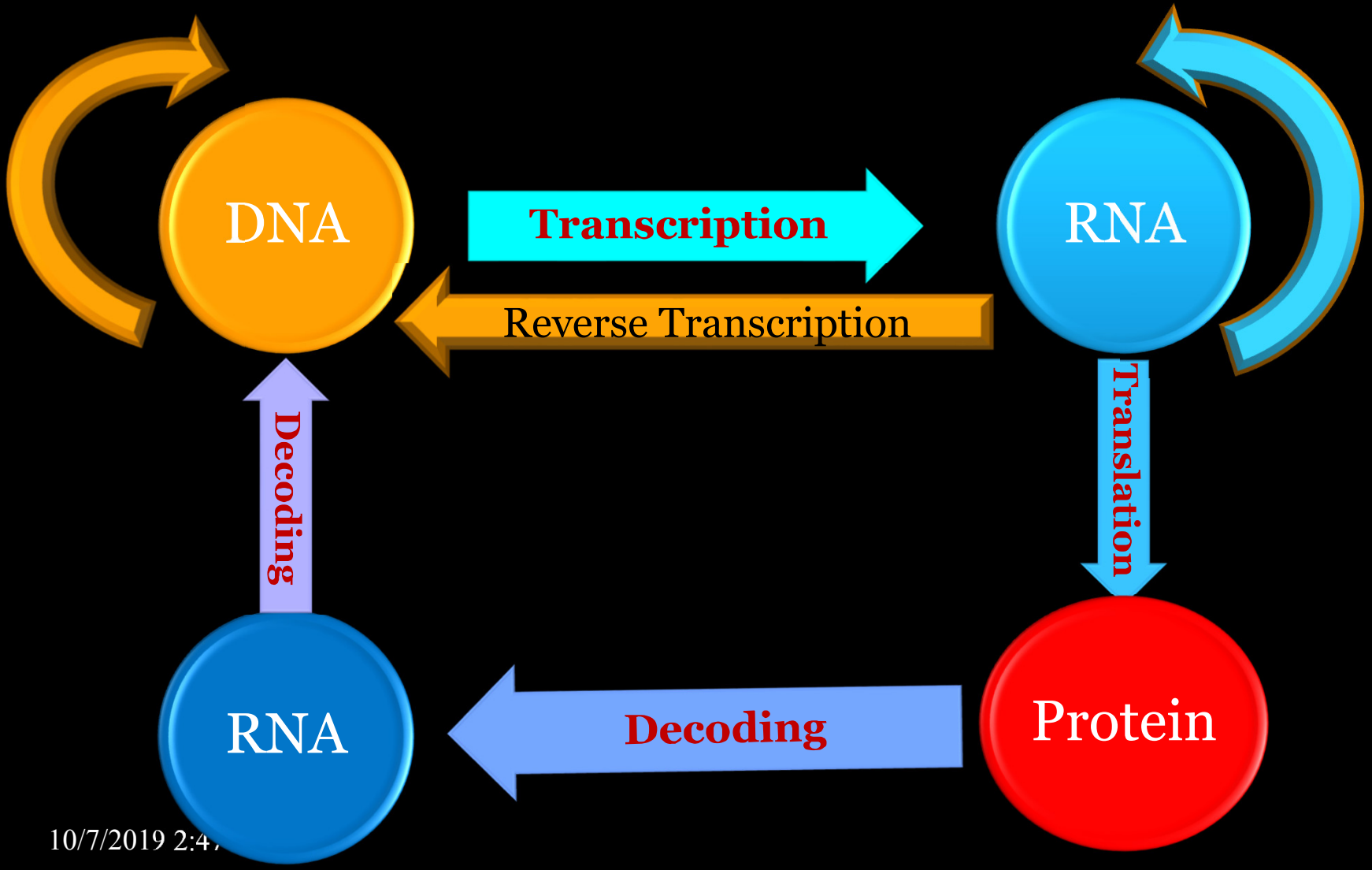
Shri Shivaji Science College

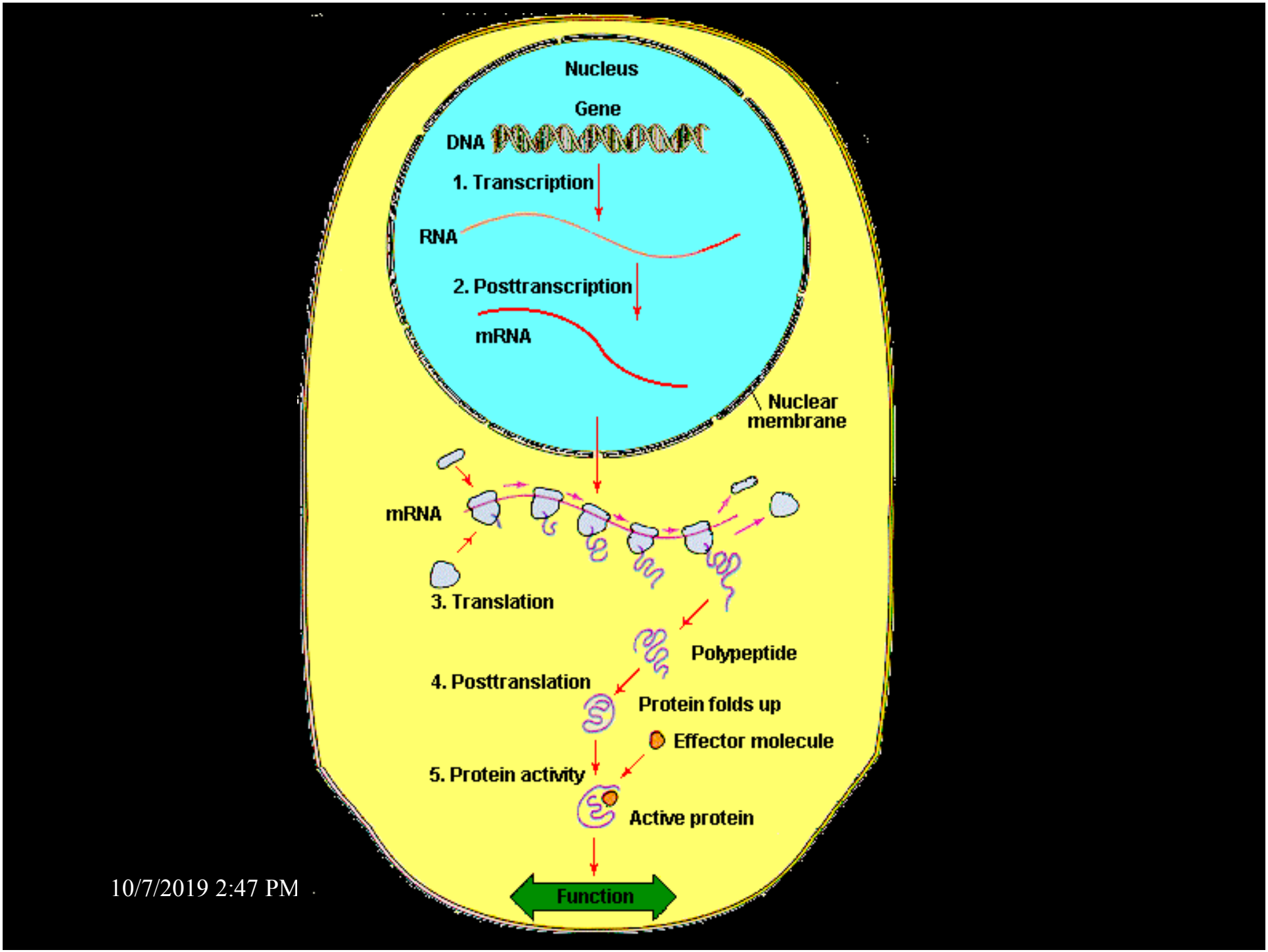
Amravati, Maharashtra - 444 603
Phone No.: 0721-2660855 Fax: 0721-2665485
e-mail : shivajiscamt@gmail.com
Web site : www.shivajiscamt.org

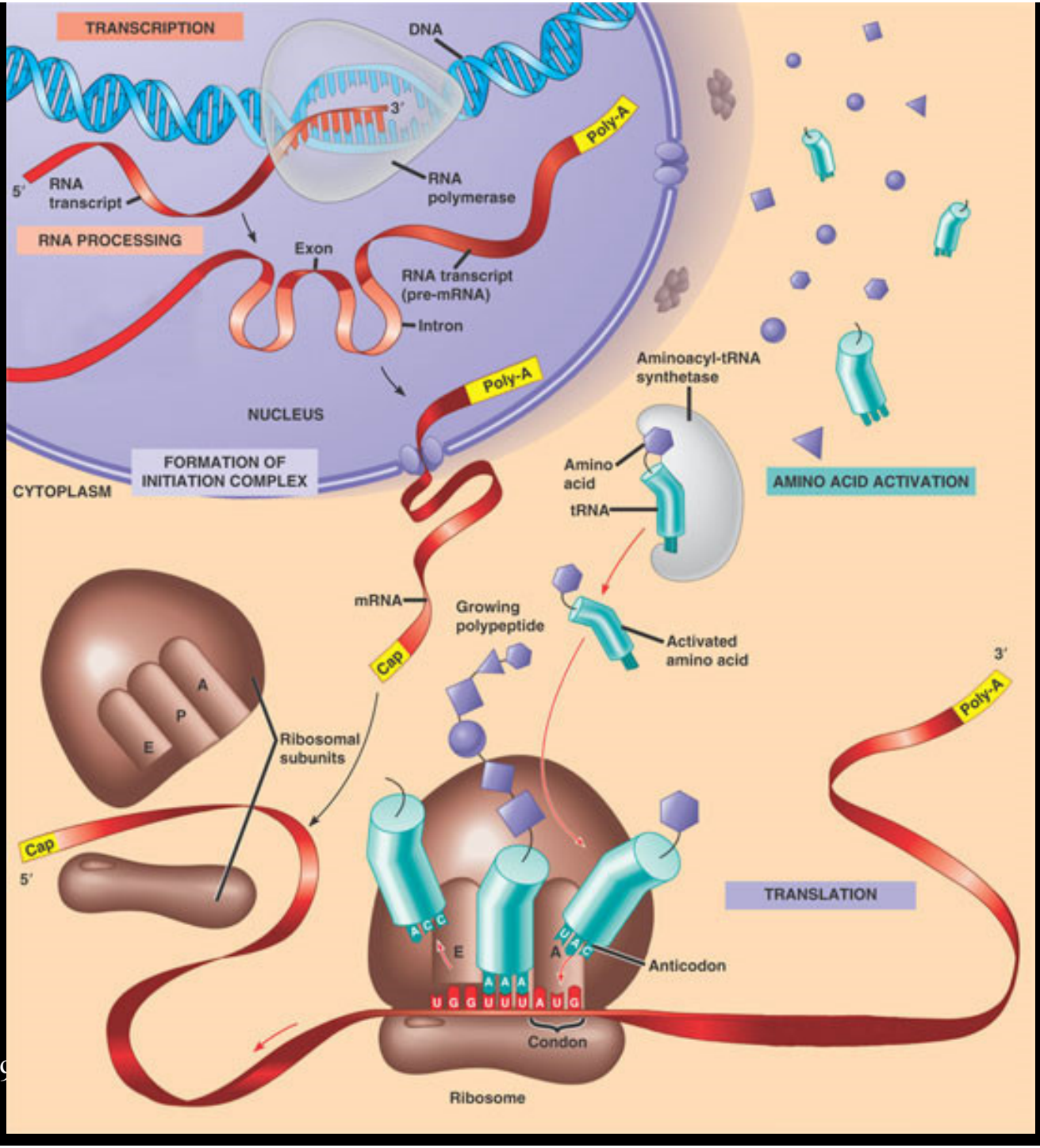
CENTRAL DOGMA



MODERN CENTRAL DOGMA

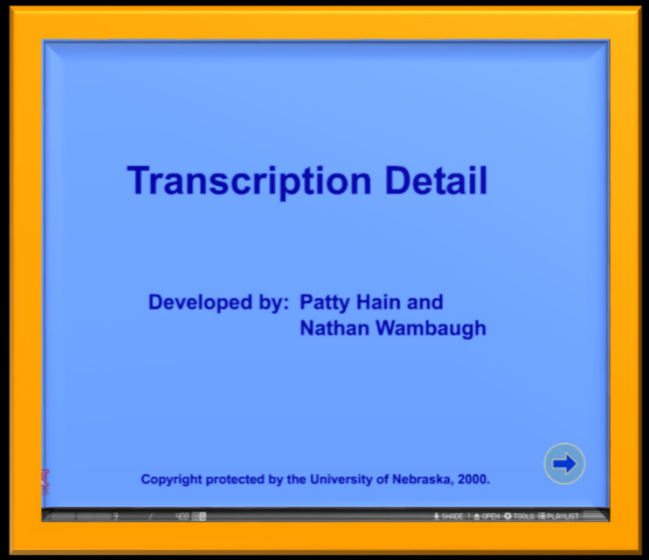



Nucleus
Gene
DNA
1. Transcription
RNA
2. Posttranscription
mRNA
Nuclear membrane
mRNA
3. Translation
Polypeptide
4. Posttranslation
Protein folds up
Effector molecule
5. Protein activity
Active protein
Function


TRANSCRIPTION
DNA
3′
Poly-A
5′
RNA transcript
RNA polymerase
RNA PROCESSING
Exon
RNA transcript (pre-mRNA)
Intron
Poly-A
NUCLEUS
CYTOPLASM
FORMATION OF INITIATION COMPLEX
Aminoacyl-tRNA synthetase
Amino acid
tRNA
AMINO ACID ACTIVATION
mRNA
Growing polypeptide
Activated amino acid
Cap
A
P
E
Ribosomal subunits
3′
Poly-A
Cap
5′
TRANSLATION
A C C
E
A
U A G
Anticodon
A A A
U G G U U U A U G
Condon
Ribosome

PROTEIN SYNTHESIS

Central dogma



10/7/2019 2:47 PM

GENETIC CODE

In the flow of genetic information in living organisms the language of the all molecules should be read by the enzymes or any concerning reader.

- DNA could be read by Bases
- Proteins could be read by Amino Acids
- mRNA could be read by ?

–It will be read by codon

GENETIC CODE

Genetic code is established to have an understanding of molecular language.

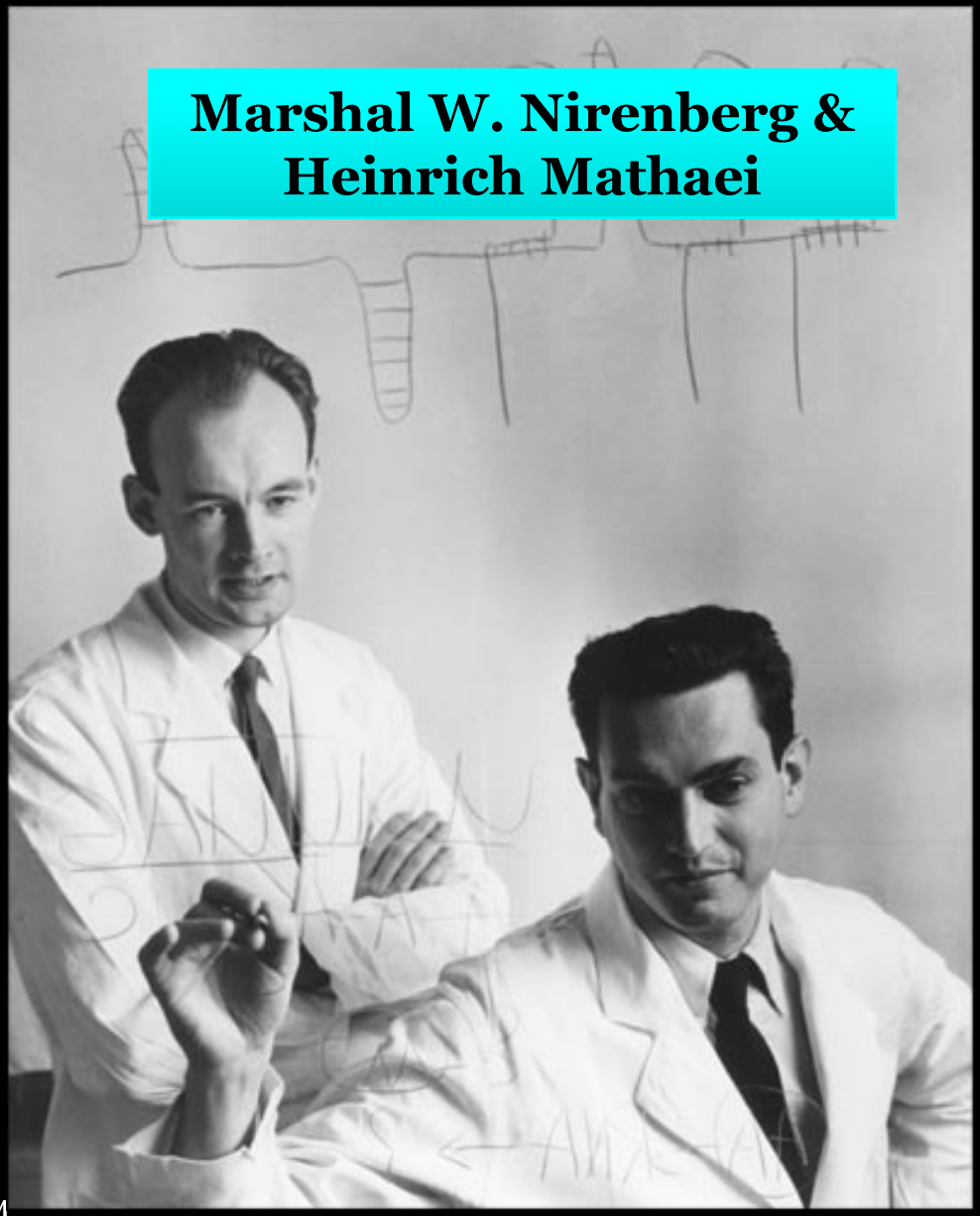
In the years after 1953, scientists scrambled to be the first to decipher the genetic code. In an attempt to make the race interesting, theoretical physicist and astronomer **George Gamow** came up with a plan. He organized an exclusive club, the **"RNA Tie Club,"**

M. W. Nirenberg, H. G. Khoorana and R. W. Holey (1964) gave a first clue to prepare dictionary of codons

Codon assignment was studied through the synthesis of the artificial mRNA by ***Polynucleotide Phoshorylase*** enzyme

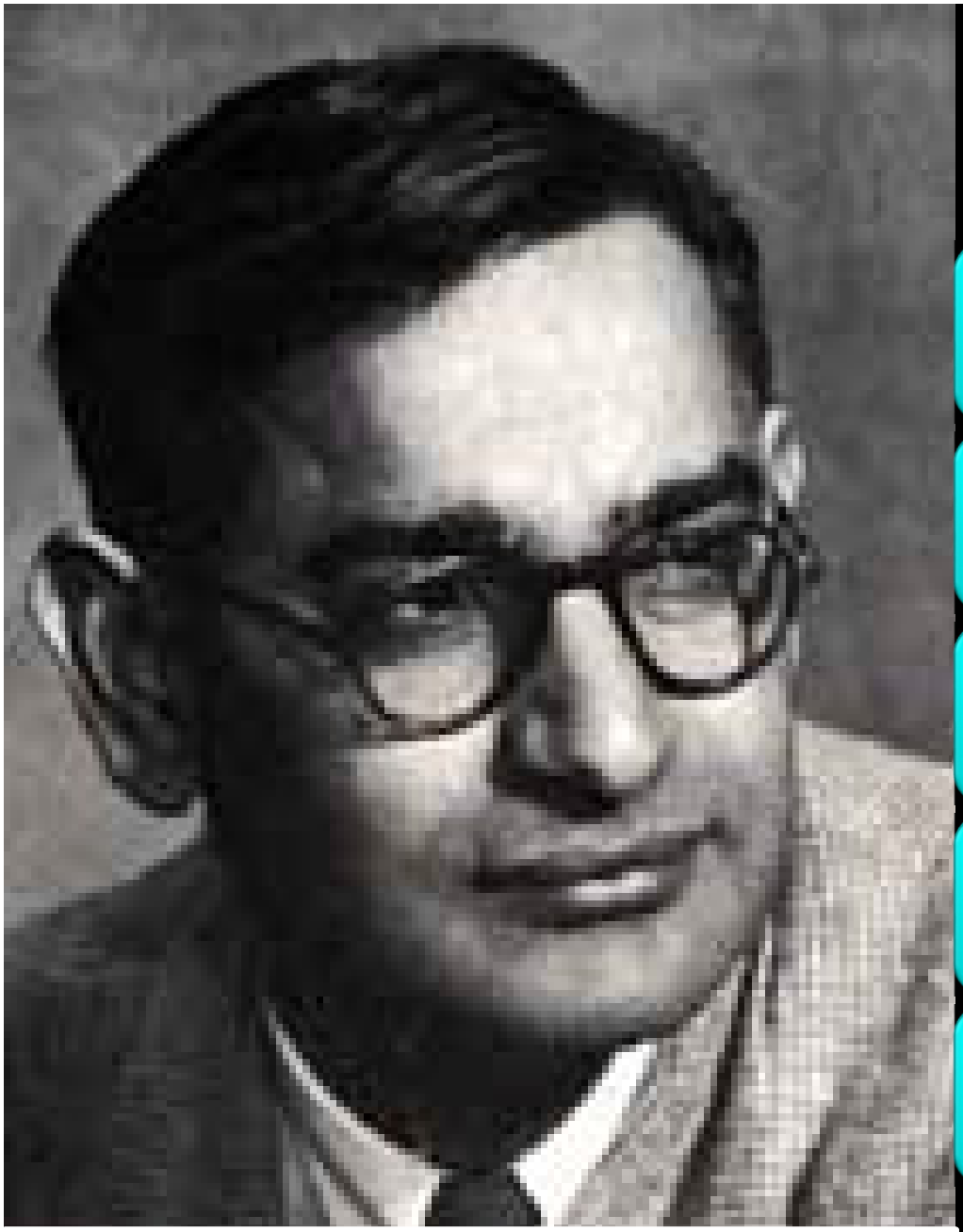
M. W. Nirenberg and Matthaei (1961) synthesized polyphenyl alanine amino acid by poly U mRNA

Marshal W. Nirenberg & Heinrich Mathaei



NOBEL PRIZE, 1968

M. W. Nirenberg **R. W. Holley** **H.G. Khorana**

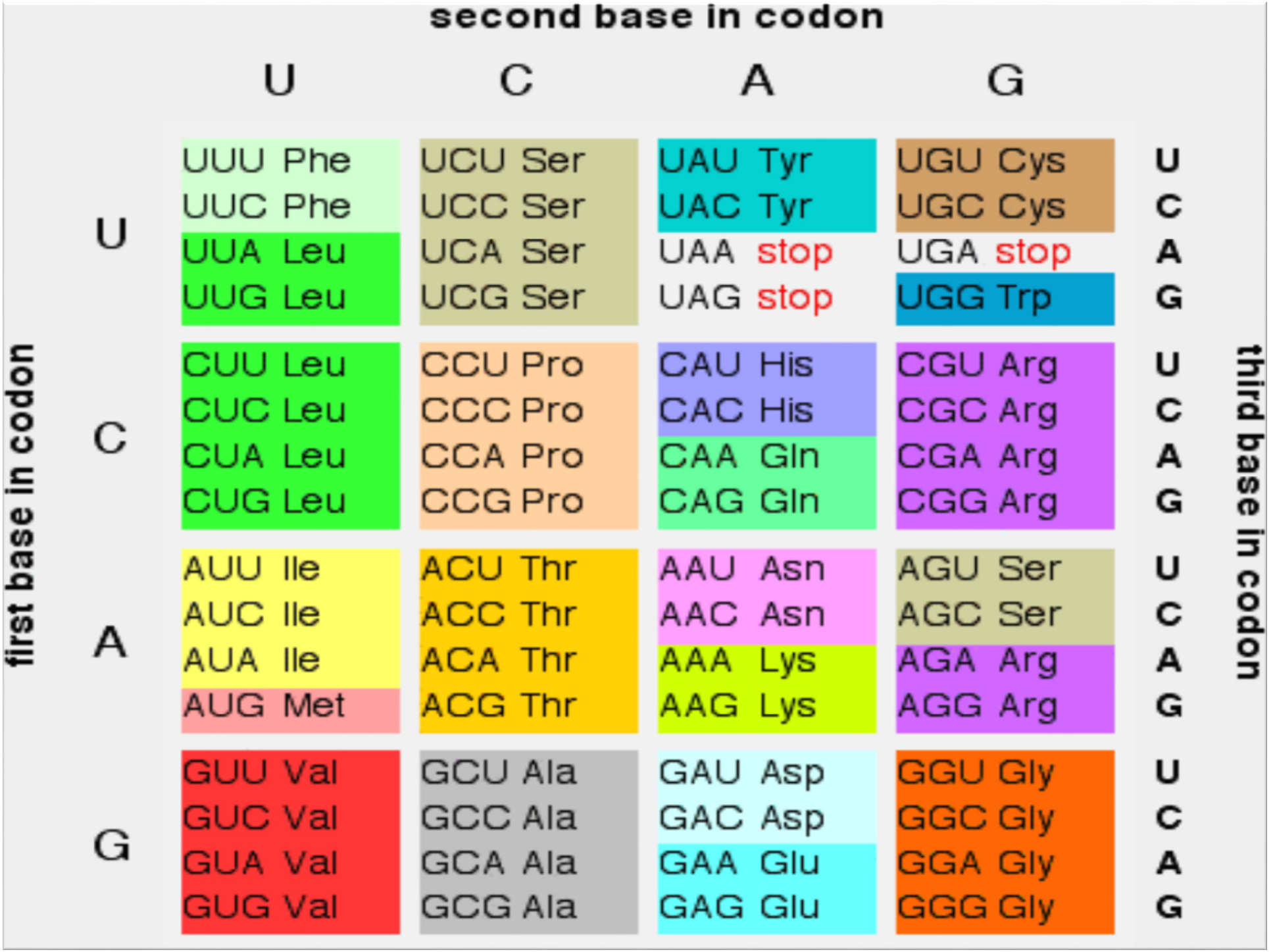


Har Gobind Khorana

- Born in 1922, India
- B. Sc. & M. Sc.: Punjab Uni.
- Ph. D. : Uni. of Liverpool
- Teacher at Uni. of Wisconsin
- Nobel, 1968 – synthetic gene & cracking of genetic code

GENETIC CODE

- Poly A, U, C and G mRNA provided information of amino acid coding.
- UUU - Polyphenyl Alanine
- AAA - Lysine
- CCC - Proline
- GGG - Glycine
- UUG - Leucine
- AUG- Metheonine
- CCU - Proline
- This leads to further study of codons and their codes


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

The table shows the 64 codons and the amino acid for each. The direction of the mRNA is 5' to 3'.					
1st Base	2nd base				3rd Base
	U	C	A	G	
U	UUU (Phe/F)Phenylalanine UUC (Phe/F)Phenylalanine UUA (Leu/L)Leucine UUG (Leu/L)Leucine	UCU (Ser/S)Serine UCC (Ser/S)Serine UCA (Ser/S)Serine UCG (Ser/S)Serine	UAU (Tyr/Y)Tyrosine UAC (Tyr/Y)Tyrosine UAA Ochre (Stop) UAG Amber (Stop)	UGU (Cys/C)Cysteine UGC (Cys/C)Cysteine UGA Opal (Stop) UGG (Trp/W)Tryptophan	U C A G
C	CUU (Leu/L)Leucine CUC (Leu/L)Leucine CUA (Leu/L)Leucine CUG (Leu/L)Leucine	CCU (Pro/P)Proline CCC (Pro/P)Proline CCA (Pro/P)Proline CCG (Pro/P)Proline	CAU (His/H)Histidine CAC (His/H)Histidine CAA (Gln/Q)Glutamine CAG (Gln/Q)Glutamine	CGU (Arg/R)Arginine CGC (Arg/R)Arginine CGA (Arg/R)Arginine CGG (Arg/R)Arginine	U C A G
A	AUU (Ile/I)Isoleucine AUC (Ile/I)Isoleucine AUA (Ile/I)Isoleucine AUG (Met/M)Methionine	ACU (Thr/T)Threonine ACC (Thr/T)Threonine ACA (Thr/T)Threonine ACG (Thr/T)Threonine	AAU(Asn/N)Asparagine AAC (Asn/N)Asparagine AAA (Lys/K)Lysine AAG (Lys/K)Lysine	AGU (Ser/S)Serine AGC (Ser/S)Serine AGA (Arg/R)Arginine AGG (Arg/R)Arginine	U C A G
G	GUU (Val/V)Valine GUC (Val/V)Valine GUA (Val/V)Valine GUG (Val/V)Valine	GCU (Ala/A)Alanine GCC (Ala/A)Alanine GCA (Ala/A)Alanine GCG (Ala/A)Alanine	GAU (Asp/D)Aspartic acid GAC (Asp/D)Aspartic acid GAA (Glu/E)Glutamic acid GAG (Glu/E)Glutamic acid	GGU (Gly/G)Glycine GGC (Gly/G)Glycine GGA (Gly/G)Glycine GGG (Gly/G)Glycine	U C A G

The table shows the 64 codons and the amino acid for each. The direction of the mRNA is 5' to 3'.

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

Wobbling

GENETIC CODE

Degeneracy

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Genetic Code properties

1. Triplet Nature
2. Non Overlapping
3. Non Punctuating
4. Degeneracy (Synonyms)
5. Non Ambiguous
6. Universal
7. Wobbling

TRIPLET NATURE

- Three bases of m RNA coding one amino acid
- Total standard amino acids are 20
- If Singlet – Specifying 4 amino acids
- If Doublet - Specifying 16 amino acids
- If Triplet - Specifying 64 amino acids
- Evidences-
- Effects of addition or deletion of one nucleotide
- Frame shift mutation
- Point mutation, etc.

NON OVERLAPPING

- Codons are purely non overlapping
- Change in a base can affect more than one codon

OVERLAPPING

CAG CAG CAG CAG 4 Codons

Gln Gln Gln Gln

One CAG GCA AGC CAG GCA AG 6 Codons

Gln Ala Ser Gln Ala Ser

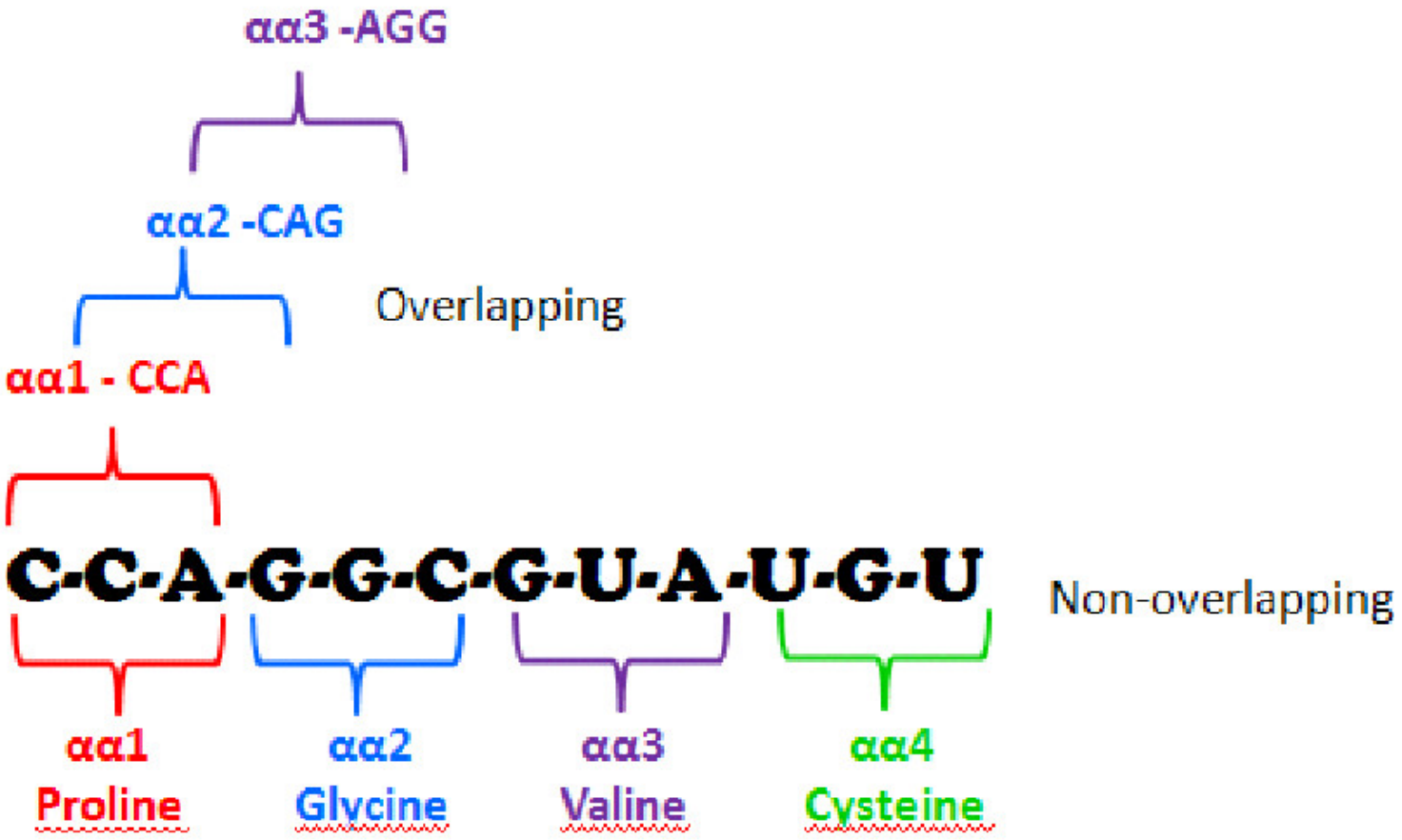
Two CAG AGC GCA CAG AGC GCA CAG

Gln Ala Ser Gln Ala Ser Gln

AGC GCA CAG

Ala Ser Gln 10 Codons

The genetic code is non-overlapping



NON PUNCTUATING

A women without her man is nothing

A women without her man, is nothing.

A women: without her, man is nothing.

NON PUNCTUATING

ललीत घोडयावर बस.

ललीत, घोडया, वर बस.

- I like cooking dogs and animals

NON PUNCTUATING

- Genetic code is non punctuating
- The reading frame of m RNA could not have any break during translation
- Punctuation in codon may be lethal
- It is called as comma less language

DEGENERACY (SYNONYMS)

Lacking some usual or expected property, Corrupted

GENETIC CODE

One amino acid is coded by many (determined) number of codons

Codons are degenerate, they do not posses independent coding by them only

These codon are called as Synonymous Codons

Eight groups of dictionary coding just one amino acid

Unmixed families – Needs reading only first two bases (8 Mixed Families)

Mixed families – Group is coding two different amino acid or stop codon (8 Mixed Families)

DEGENERACY (SYNONYMS)

GENETIC CODE

Synonymous Codons

No. of Codons		Amino Acid
Six	–	Ser, Arg, Leu
Four	–	Val, Pro, Thr, Ala, Gly
Three	–	Ile
Two	–	Phe, Tyr, Asn, His, Gln, Asp, Lys, Glu, Cys,
One	–	Met, Trp

NON AMBIGUOUS

Where do you place your bones when you eat?

डोरकव सर एकदम हाडाचे शिक्षक आहेत

NON AMBIGUOUS

- A particular codon will always code for a specific amino acid
- **Exceptions**
  - –AUG – N – Formyl Metheonine
  - –GUG – Valine, Methionine

- UGA can code for selenocysteine and UAG can code for pyrrolysine Selenocysteine is now viewed as the 21st amino acid, and pyrrolysine is viewed as the 22nd.

UNIVERSAL

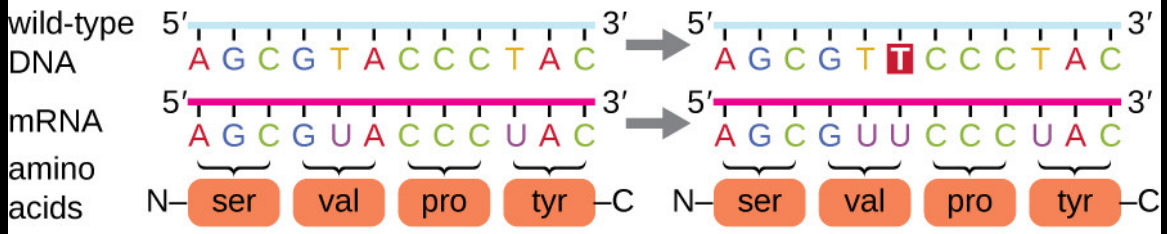
- All the living organisms are having same meaning of the code
- In 1980, discrepancies in the code were thought
- Organeller genomes are having different meaning
- Bonitz (1980) Proposed new genetic code for mitochondrial DNA, Ciliated Protozoans, Mycoplasms, etc.

UNIVERSAL

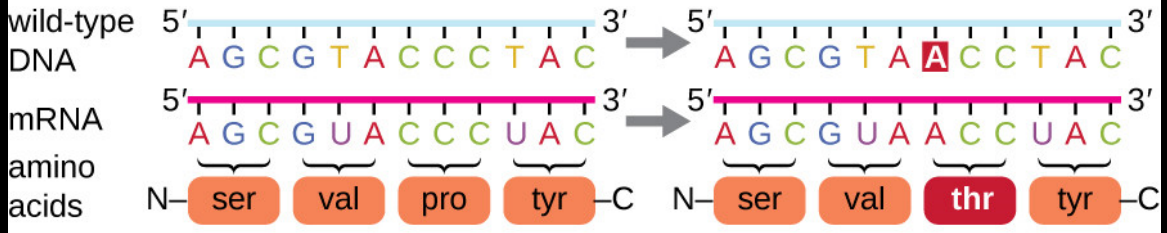
- 22 anticodons in place of 55
- 4 stop codon (AGG) in stead of three
- Non Universal amongst mitochondria also

point mutation: substitution of a single base

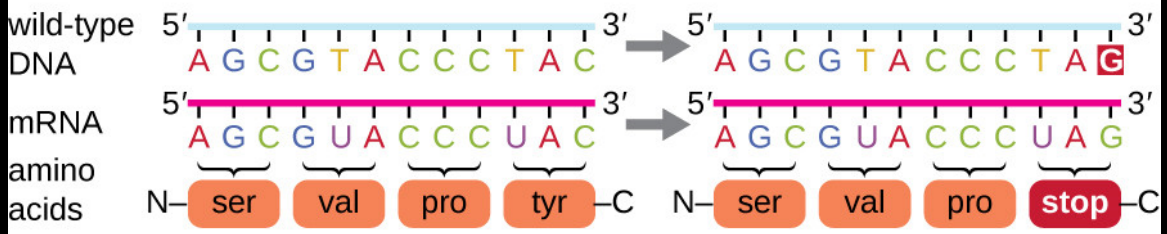
silent: has no effect on the protein sequence



missense: results in an amino acid substitution

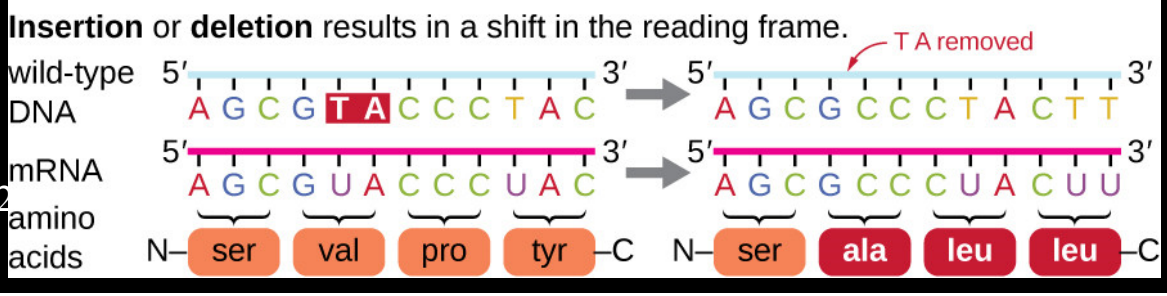


nonsense: substitutes a stop codon for an amino acid



frameshift mutation: insertion or deletion of one or more bases

Insertion or **deletion** results in a shift in the reading frame.



BONITZ'S DICTIONARY

CODONS	GENERAL	ALTERNATIVE
CUX	Leu	**Thr (Yeast)**
AUA	Ile	**Met (Yeast, Dro.)**
UGA	Stop	**Trp (Myco. Higher plants)**
AGA/AGC	Arg	**Stop (Yeast, Vertibrates)**
CGG	Arg	**Trp (Higher plants)**
UAA/UAG	Stop	**Glu (Protozons)**
UAG	**Stop**	**Ala / Leu (Higher Plants)**

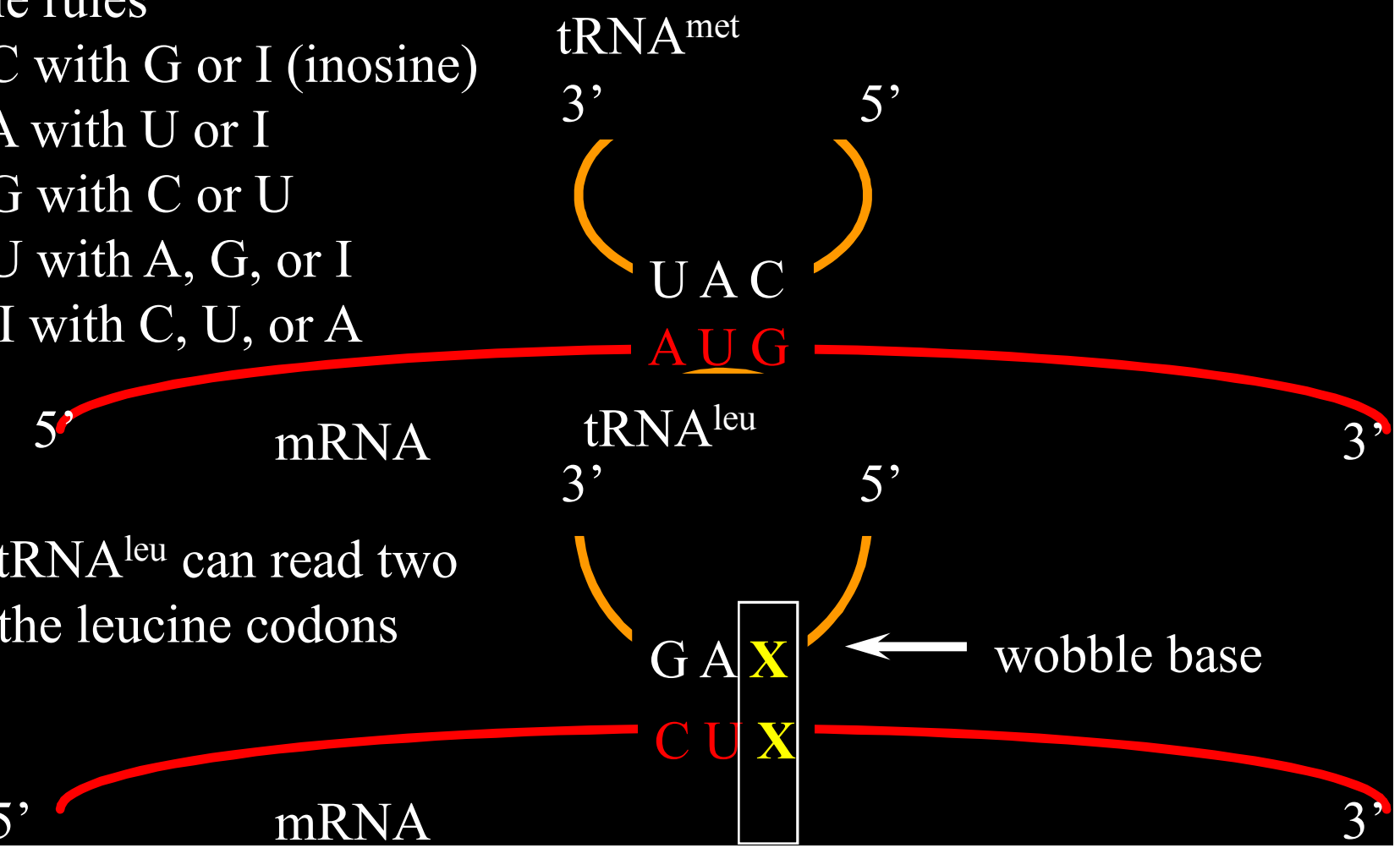
WOBBLING

GENETIC CODE

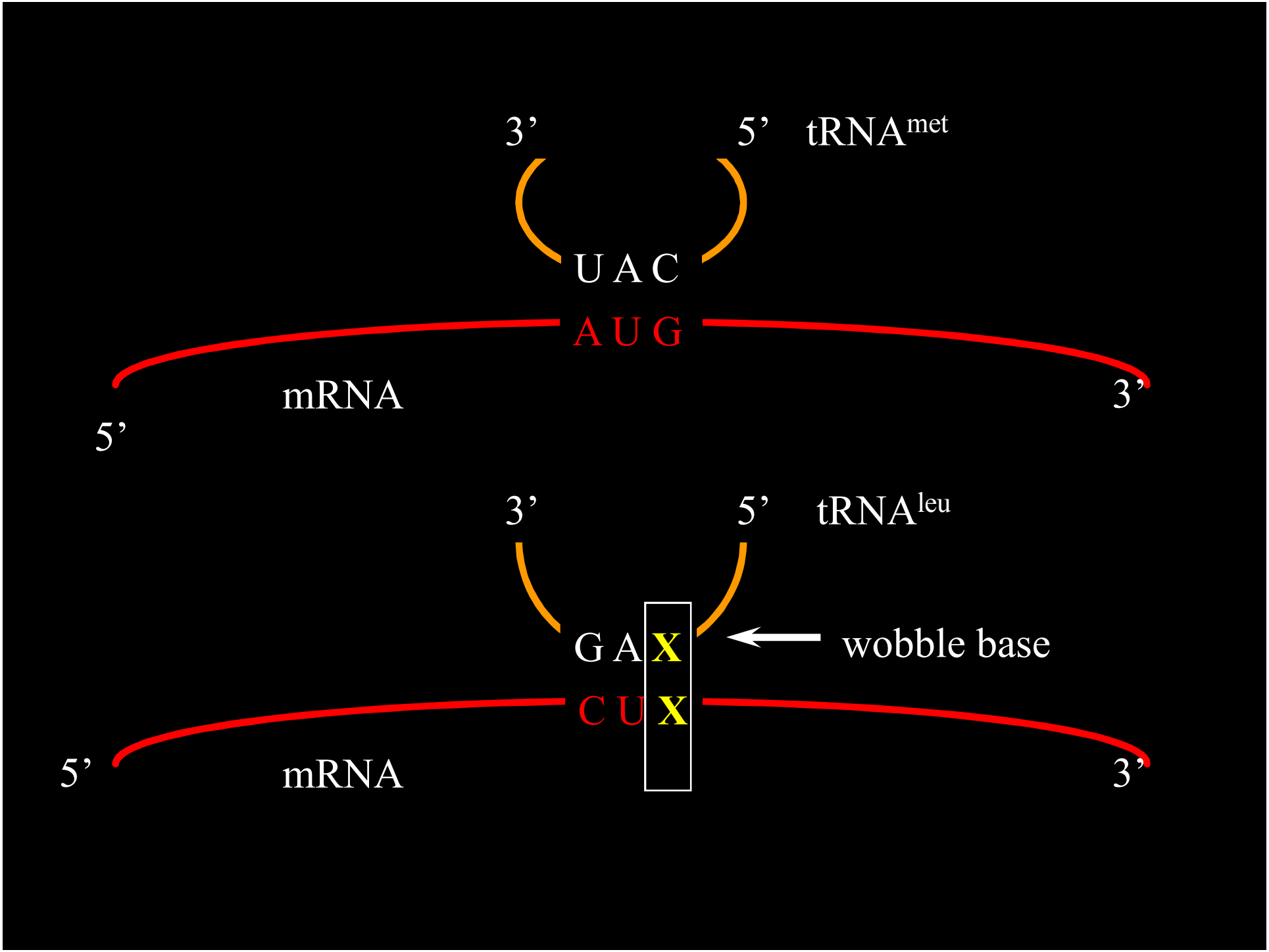
- Wobble hypothesis proposed by Dr. F. H. C. Crick (1965)
- Third base of the codon is not important
- The specificity of the codon is determined by first two bases
- Same t RNA can recognize more than one codon CUX, UUA, UUG (Leu) By Same t RNA
- This phenomenon is responsible for evolution of genetic code
- Recently, 24 t RNA enough in general and 22 in mitochondria

Codon-anticodon interactions

- codon-anticodon base-pairing is antiparallel
- the third position in the codon is frequently degenerate
- one tRNA can interact with more than one codon (therefore 50 tRNAs)
- wobble rules
  - C with G or I (inosine)
  - A with U or I
  - G with C or U
  - U with A, G, or I
  - I with C, U, or A



- one $tRNA^{leu}$ can read two of the leucine codons


3’
5’
tRNA met
U A C
A U G
mRNA
5’
3’
3’
5’
tRNA leu
G A X
wobble base
C U X
5’
mRNA
3’

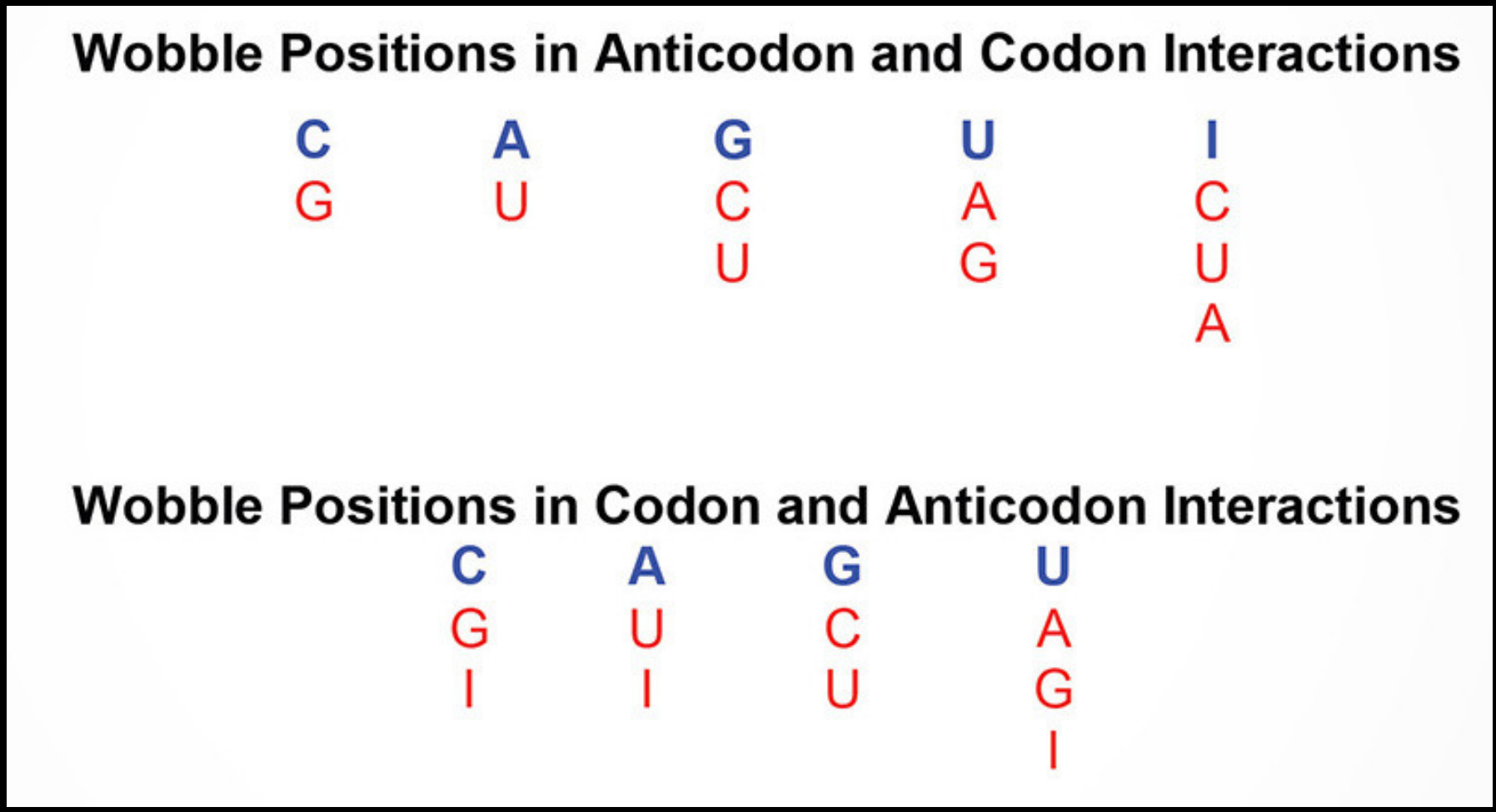
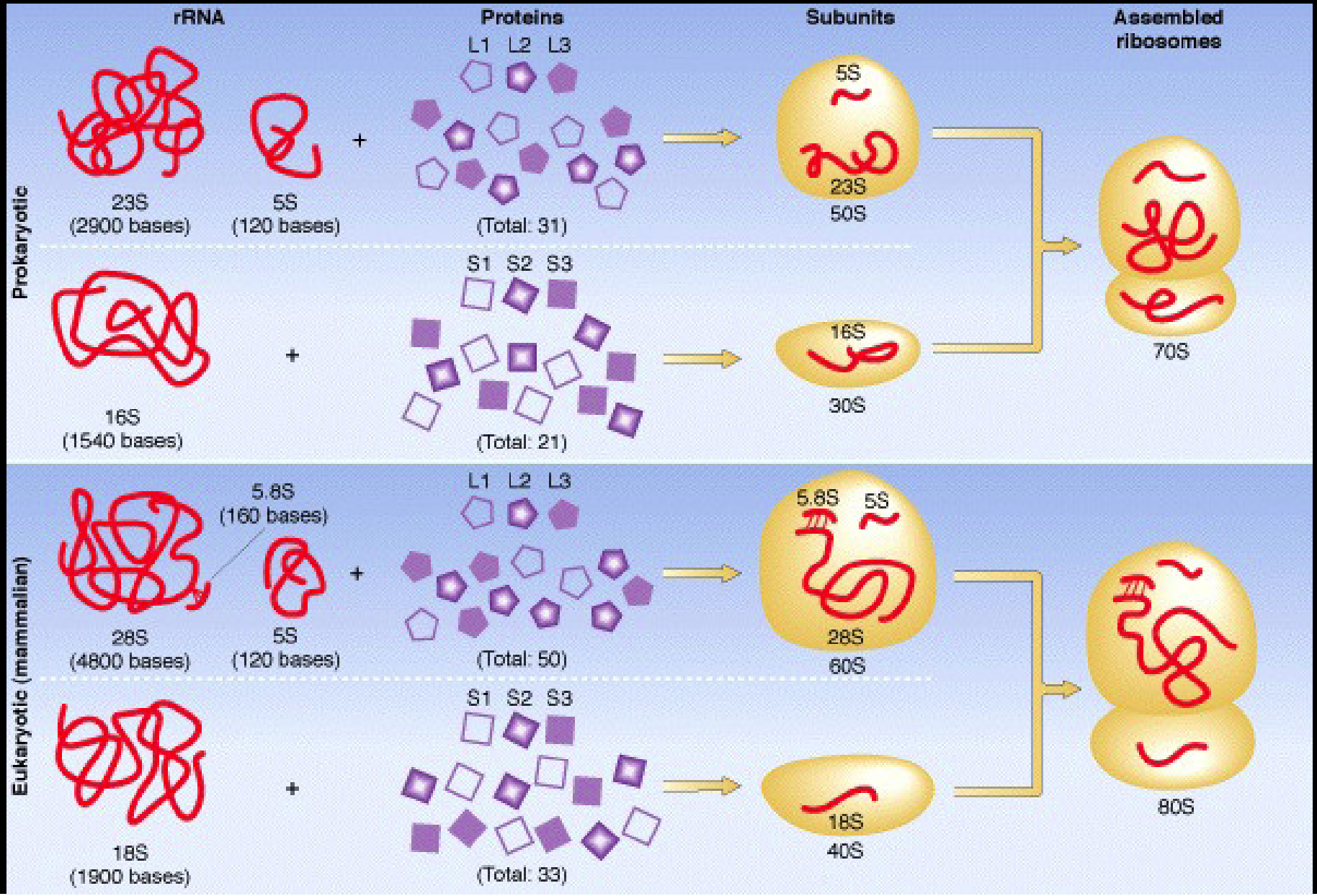
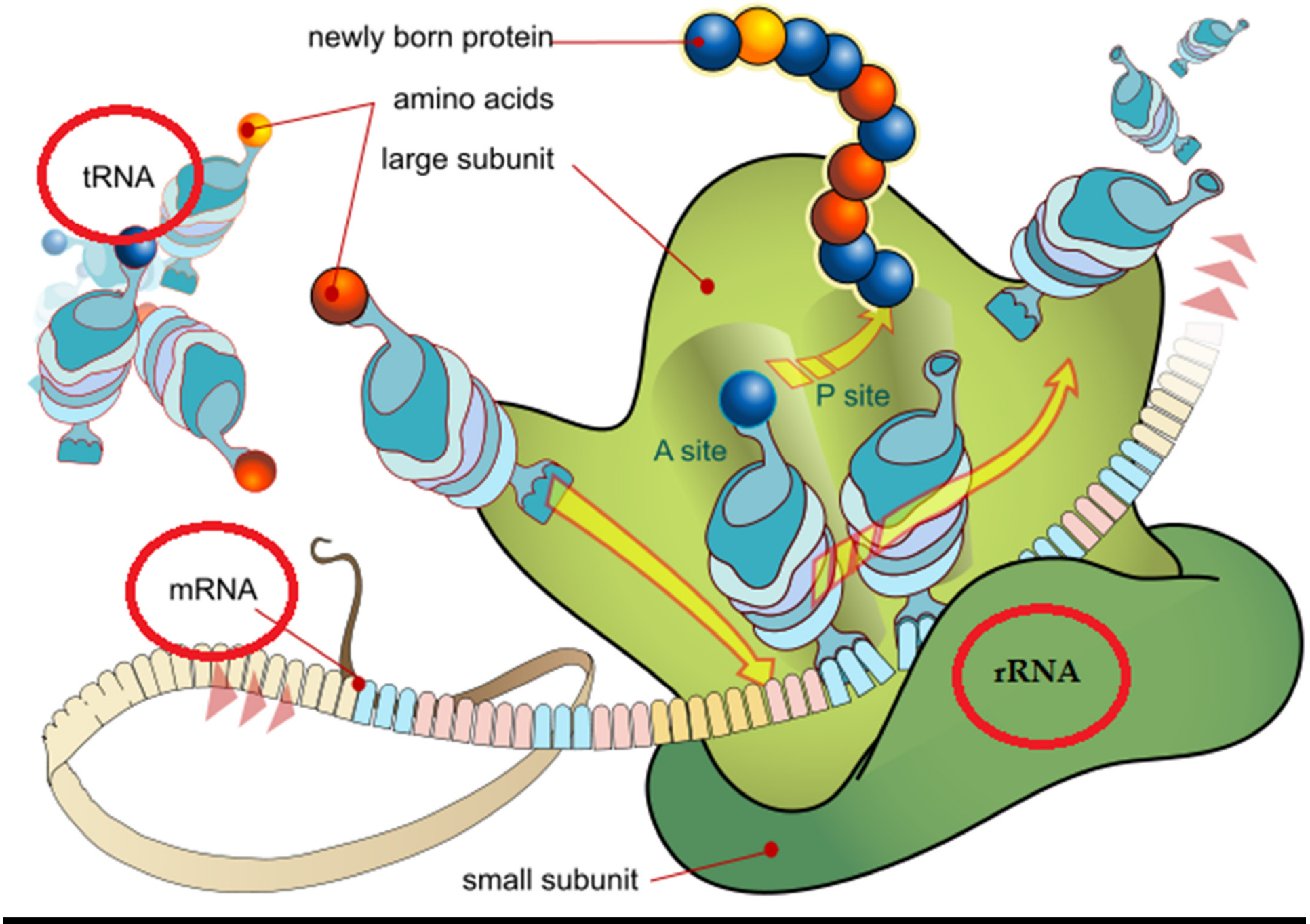
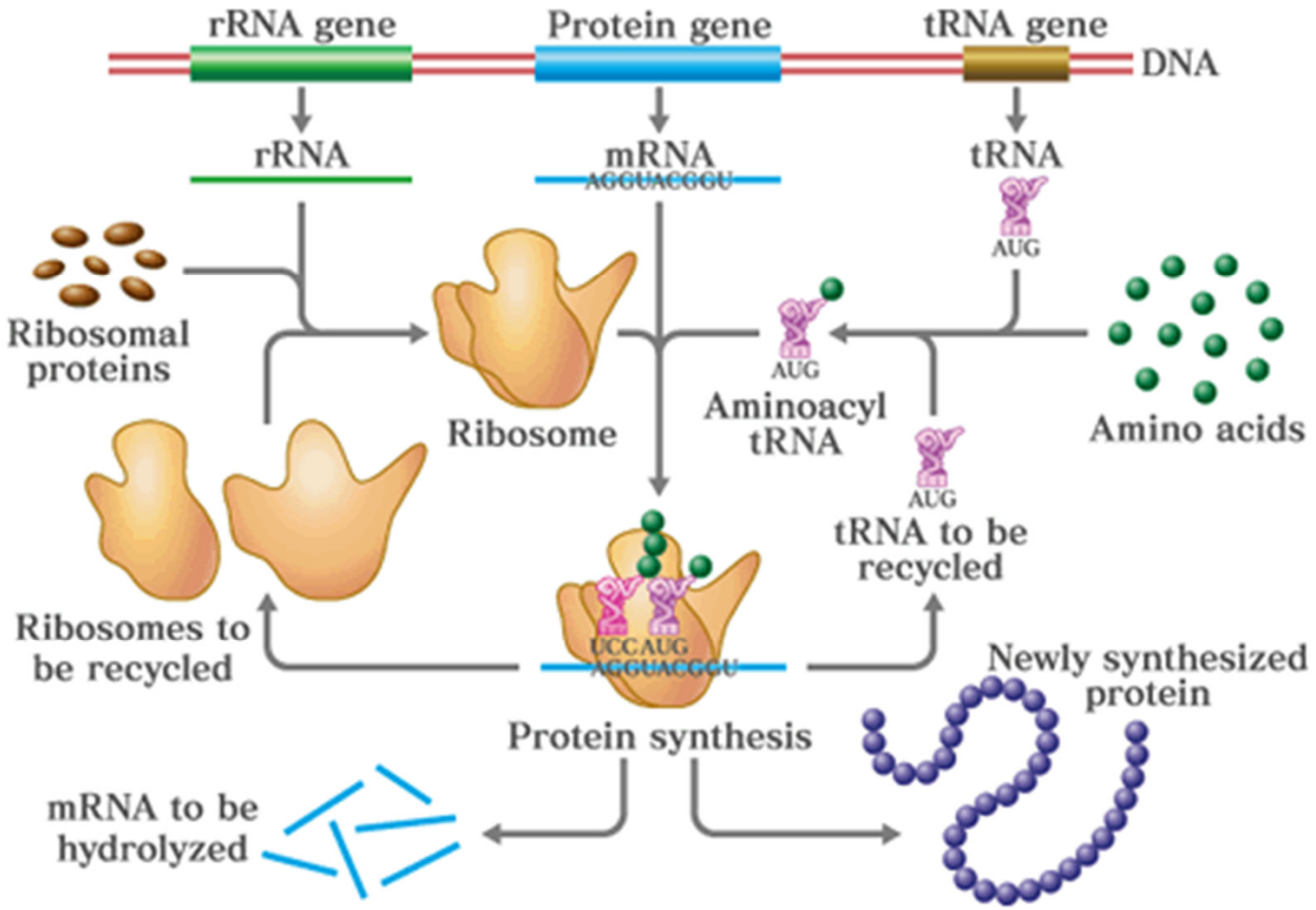


Diagram showing the various modified nucleotides of tRNAs that are found in the wobble position in the anticodon. The top half shows the wobble nucleotides of the anticodon in blue and the various nucleotides (in red) of the wobble position of the codon that can be found in non-Watson-Crick base-pairs. The lower panel illustrates the opposite showing the wobble nucleotides of the codon in blue and the associated wobble nucleotides of the anticodon in red.

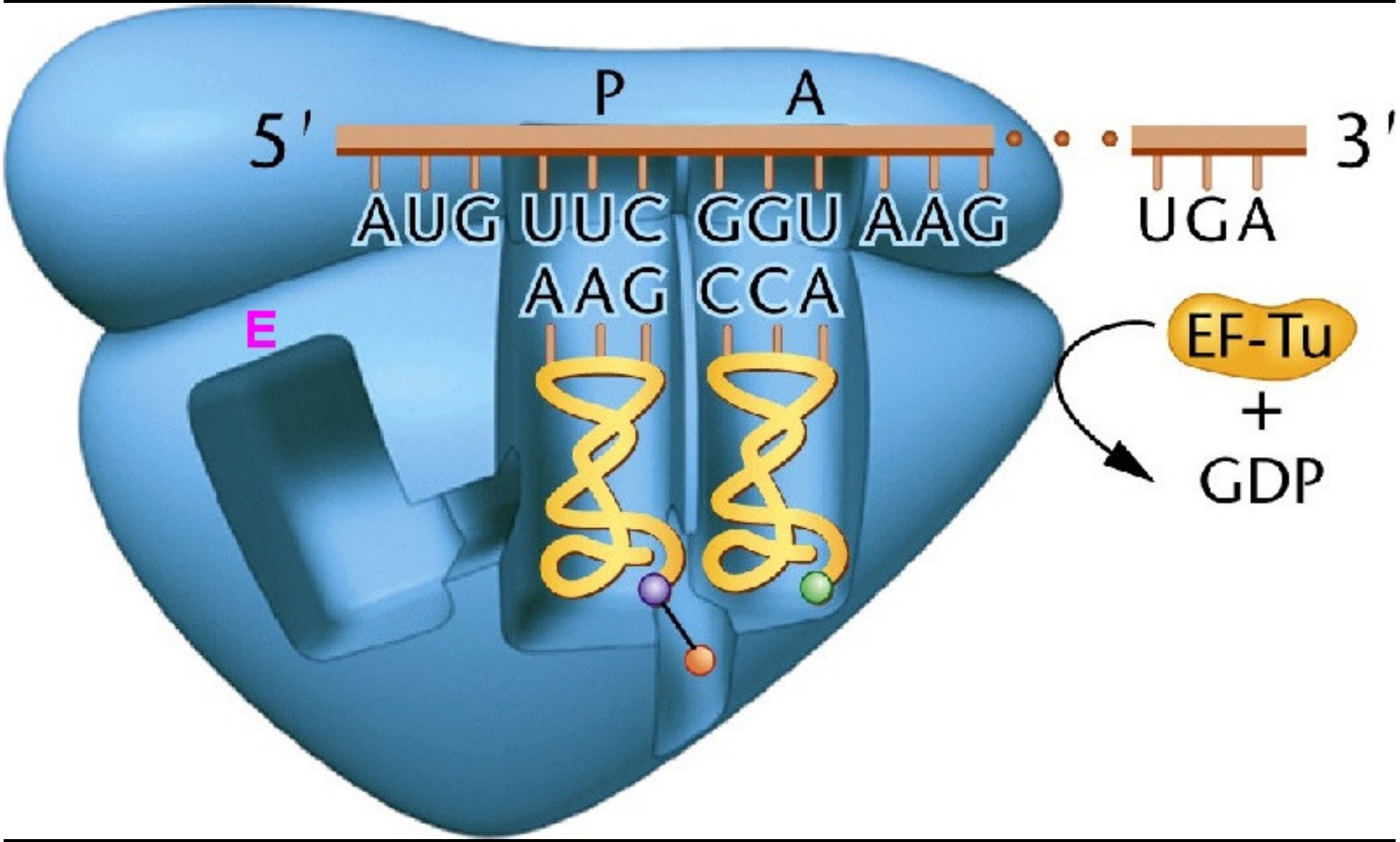
RIBOSOME – AS TRANSLATION MACHINERY

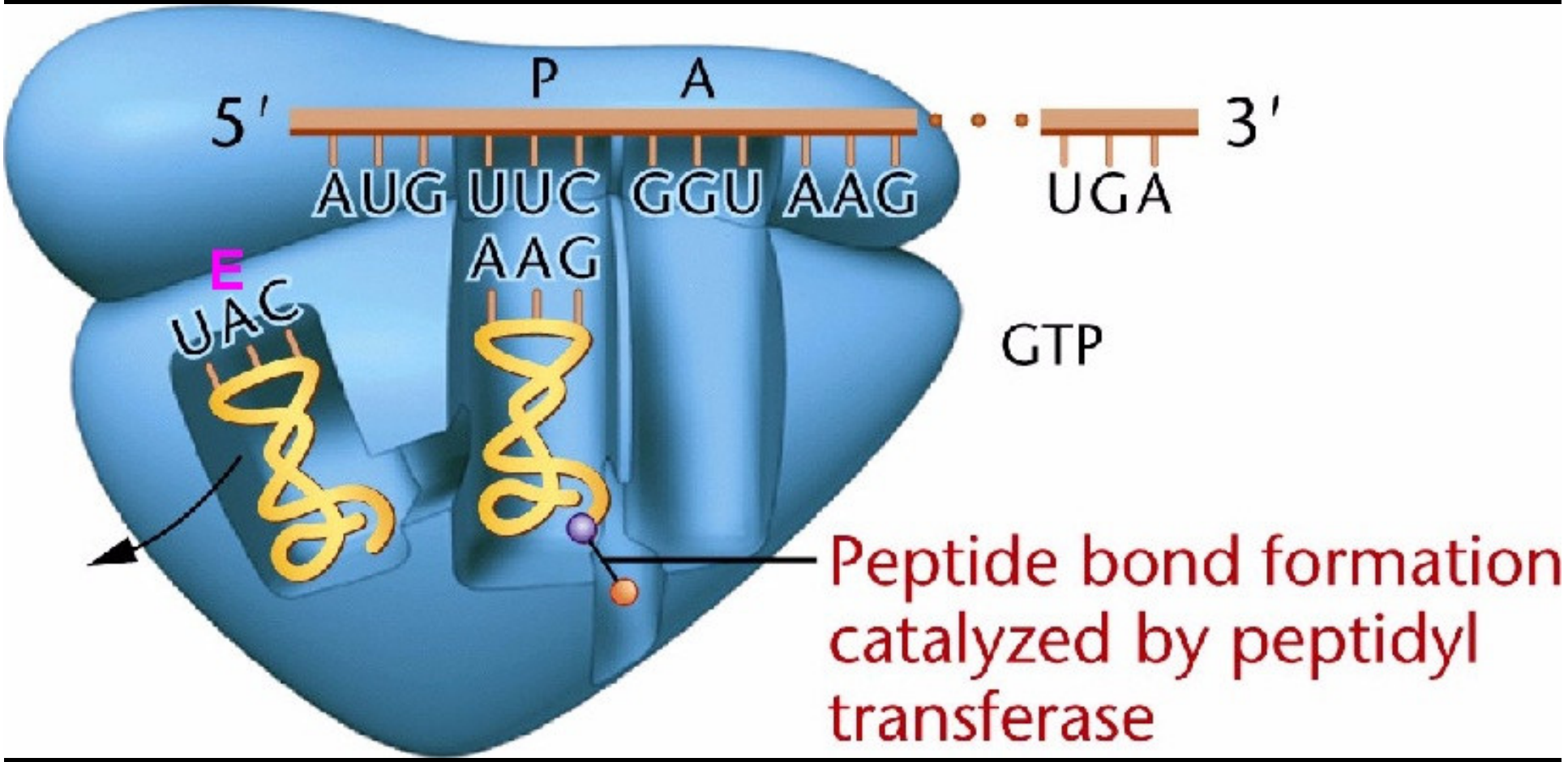


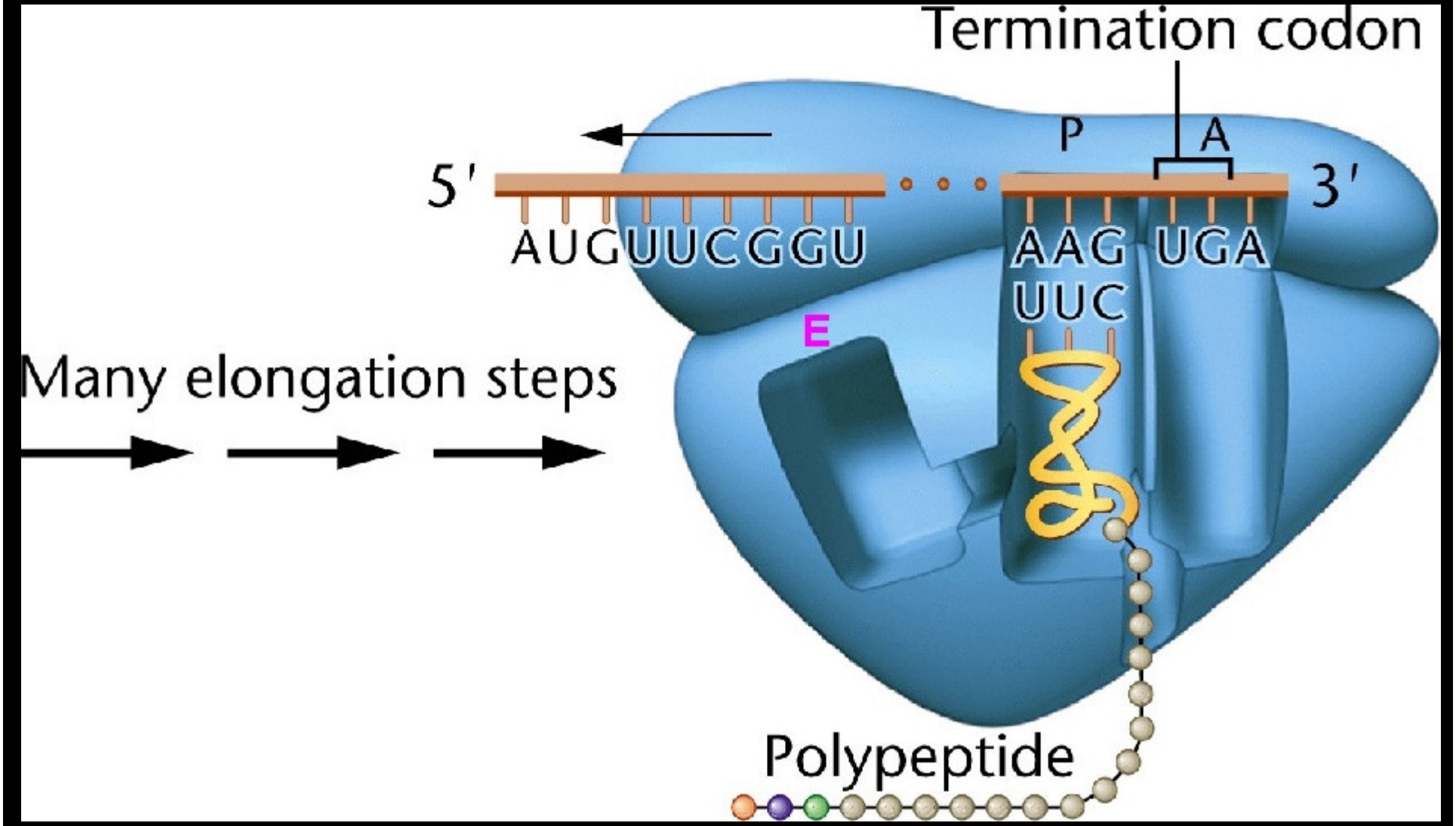

newly born protein
amino acids
large subunit
tRNA
P site
A site
mRNA
rRNA
small subunit


rRNA gene
Protein gene
tRNA gene
DNA
rRNA
mRNA
AGGUACGGU
tRNA
AUG
Ribosomal proteins
Ribosome
AUG
Aminoacyl tRNA
Amino acids
AUG
tRNA to be recycled
Ribosomes to be recycled
UCC AUG
AGGUACGGU
Newly synthesized protein
Protein synthesis
mRNA to be hydrolyzed

ZOOM


P
A
5′
3′
AUG UUC GGU AAG
UGA
AAG CCA
E
EF-Tu
+
GDP


P
A
5′
3′
AUG UUC GGU AAG
UGA
E
AAG
UAC
GTP
Peptide bond formation catalyzed by peptidyl transferase


Termination codon
P
A
5′
3′
AUGUUCGGU
AAG
UGA
UUC
E
Many elongation steps
Polypeptide

Archetypal Code

•In 1966, Jukes presented a concept of premitive code, in which one anticodon will pair with family of codons due to Wobbling at first base of anticodon and third base of codon

SECOND PART

Mechanism involved in the interaction of specific t RNA with corresponding aminoacyl synthatase will provides second genetic code

An understanding of simply codons will not give reliable and justifiable data about molecular language

The protein formation (Three Dimensional Structure) is very essential to study

BIOCHEMICAL ELUCIDATION OF GENETIC CODE

Breaking of code is serendipitous.

Experimental Approaches

Assignment of codons with unknown sequences-

1. Polyuridylic Acid Method
2. Copolymer Method

Assignment of codons with known sequences-

3. Binding Technique
4. Repetitive Sequencing Technique

ASSIGNMENT OF CODONS WITH UNKNOWN SEQUENCES-

- Under influence of Amino Acid sequence in Protein
- mRNA sequencing is not well establish technique
- Indirect method to crack code
- Requirements –
  - Cell free Amino Acid incorporating system
  - Polymerizing Enzyme for ribonucleoside tripho.

POLYURIDYLIC ACID METHOD:

Marshal Nirenberg & Heinrich Mathei, 1961

UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU UUU
Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe

AAA – Lysine

CCC – Proline

GGG – Not successful as it is forming secondary structure

Severo Ochoa was also deeply involved –

Discovered Polynucleotide Phosphorylase Enzyme

Received Nobel Prize 1959

2. COPOLYMER TECHNIQUE

Nirenberg used mixture of two or more ribonucleoside diphosphate

UDP + CDP → (Polynucleotide phosphorylase) → Polynucleotides

3 : 1

UUU
UUC
UCU
CUU
CCU

Obtained phenyl alanine and serine in 3:1 Ratio

Serine contains 2 Us and 1 C

Exact sequence could not achieved

Codon assignment due to A:C, 5:1

Amino Acid	Codon composition
Lysine	3 A
Asparagine, Glutamine, Threonine	2 A 1 C
Histidine, Proline, Threonine	1 A 2 C
Proline	3 C

Assignment of codons with known sequences-

3. Binding Technique

4. Repetitive Sequencing Technique

BINDING TECHNIQUE

Marshal Nirenberg & Philip Leder, 1964

- ✓ Synthetic m RNA,
- ✓ Ribosome &
- ✓ Particular aminoacyl–t RNA

Nirenberg

Leder

Codon1+ Ribosome +AA 1+tRNA

Ribosome -Codon1-AA 1+tRNA1

RADIOACTIVITY TEST ON NITROCELLULOSE PAPER

Codon1+ Ribosome +AA 1+AA 2+AA 3+AA 4+AA 5 + AA6 + AA 7 + AA 8 + AA 9 + AA 10 + AA 11 + tRNA

Ribosome -Codon1-AA 11+tRNA

- Only 45 codons could be worked out.

Assignment of codons with known sequences-

3. Binding Technique

4. Repetitive Sequencing Technique

Har Gobind Khorana

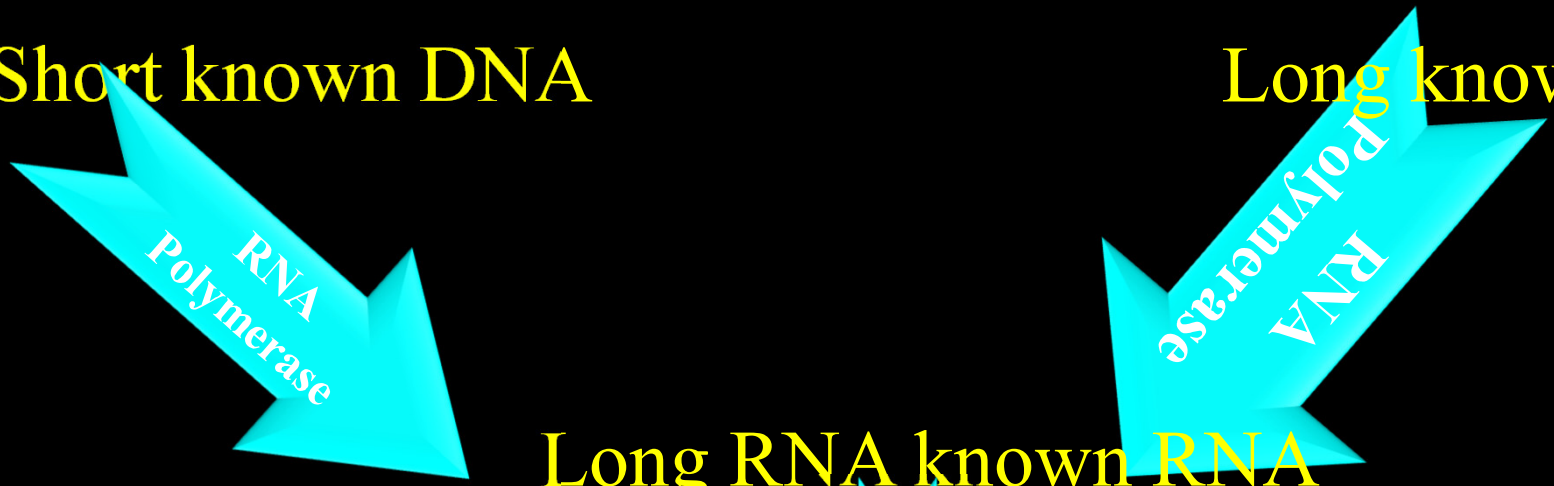
- Born in 1922, India
- B. Sc. & M. Sc.:Punjab Uni.
- Ph. D. : Uni. of Liverpool
- Teacher at Uni. of Wisconsin
- Nobel, 1968 – synthetic gene & cracking of genetic code

REPETITIVE SEQUENCING TECHNIQUE

In - vitro chemical synthesis of DNA

Short known DNA

Long known DNA



Long RNA known RNA

In – vitro Translation

Peptide of known sequence

Homopolymers & Heteropolymers were formed

Conclusions

Codon specificity for Amino Acid

Information is conveyed through RNA

Triplet & Non overlapping Genetic code

Polarity of codons

Frame importance was studied

EVOLUTION OF GENETIC CODE

- **Crick (1968)** - The genetic code evolved from a simpler form that encoded fewer amino acids.
- **Wong (1975)** - The genetic code coevolved with the invention of biosynthetic pathways for new amino acids.
- As soon as there were amino acids and nucleic acids available (produced abioticaly), both began to bind to each other.
- Knight and Landweber (2000) - It now seems clear that "the code probably underwent a process of expansion from relatively few amino acids to the modern complement of 20"

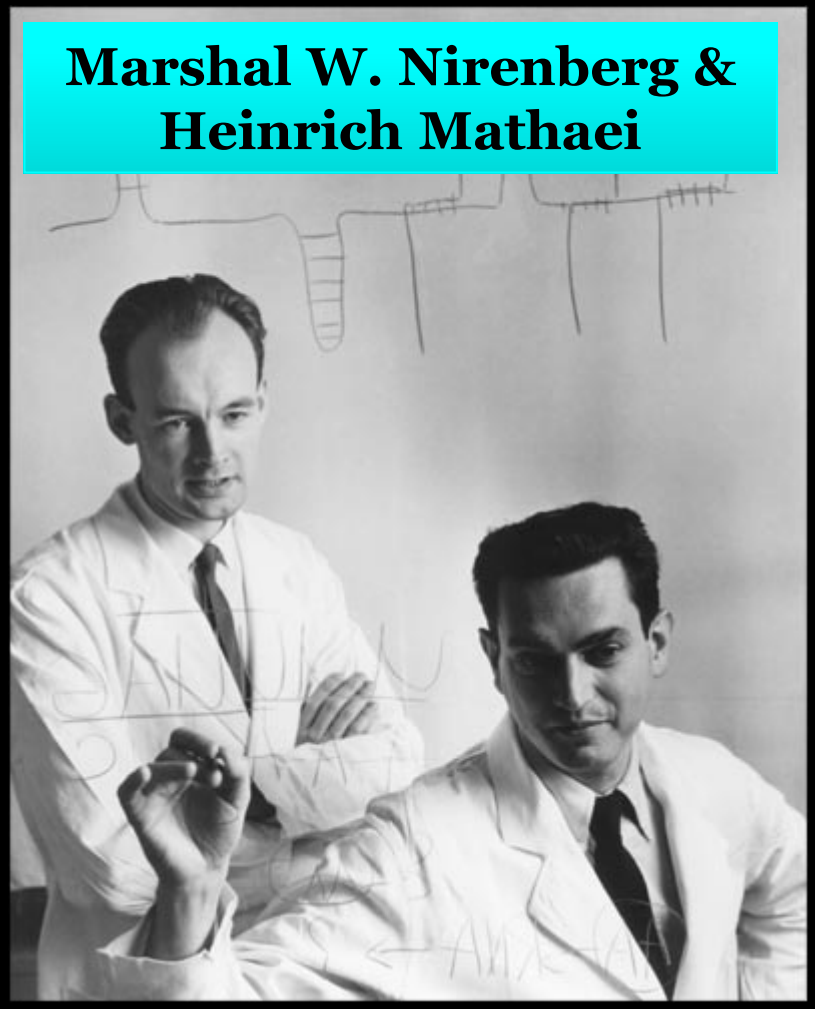
SIMPLER TO COMPLEX

Code of homopolymers

Code of 2 bases in triplet

Code of 3 bases in triplet

Marshal W. Nirenberg & Heinrich Mathaei



Leder

Robert W. Holley

Nirenberg receiving Nobel, 1968