Cell Biology, Genetics and Evolution

Genetic Code

By: Shozab Seemab Khan (PhD Scholar)

Genetic Code

- The genetic code is a set of rules used by cells to translate the information in a gene (DNA or RNA) into proteins. Proteins are made of amino acids, and the genetic code tells the cell which amino acids to use and in what order.
- **A. Triplet Code (Codons)**
- **Codons:** The genetic code is read in groups of three nucleotides, called codons. Each codon corresponds to a specific amino acid or a stop signal during protein synthesis.

Genetic Code

- 1. A triplet code comprised of three nucleotide bases in a sequence.
- 2. How many triplet codes?
- 20 common amino acids in a protein
- 4 different bases on DNA   A,T,C, & G
- 4 different bases on RNA   U,A,G, & C
- Total 64 Codons

Why Triplet Code

- The logic is that the nucleotide code must be able to specify the placement of 20 amino acids. Since there are only four nucleotides, a code of single nucleotides would only represent four amino acids, such that A, C, G and U could be translated to encode amino acids.
- A doublet code would involve two nucleotides per codon, where each nucleotide could be one of the four bases: These doublet codons would be:
- AA, AC, AG, AU
- CA, CC, CG, CU
- GA, GC, GG, GU
- UA, UC, UG, UU

Why Triplet Code

- If the genetic code had evolved as a doublet system, it would have been limited to encoding only 16 amino acids or required a different mechanism for encoding more amino acids.
- A triplet code could make a genetic code for 64 different combinations (4 X 4 X 4) genetic code and provide plenty of information in the DNA molecule to specify the placement of all 20 amino acids.

Why Triplet Code

Type of code	Number of permutations
Singlet	$4^1 = 4$
Doublet	$4^2 = 16$
Triplet	$4^3 = 64$
Quadruplet	$4^4 = 256$
Pentuplet	$4^5 = 1024$

Only the triplet code really looks feasible

Genetic Code

- For example, the RNA codon AUG (ATG in DNA) codes for the amino acid methionine, which is often the start signal for protein synthesis.
- The codon UAA is a stop signal, meaning it tells the cell to stop building the protein.
- **B. Redundancy**
- The genetic code is redundant, meaning that multiple codons can code for the same amino acid. For example, both GAA and GAG code for the amino acid glutamic acid.
- This redundancy helps protect against mutations—if one nucleotide changes, it might not always affect the amino acid being produced.

Genetic Code

- **C. Translation Process**
- The mRNA (messenger RNA) transcribes the genetic information from DNA and takes it to the ribosome, where the codons are read and translated into a chain of amino acids (a protein).
- tRNA (transfer RNA) brings the correct amino acids to the ribosome by matching its anticodon with the mRNA codon.
- The ribosome assembles the amino acids in the correct sequence to form the protein.

Genetic Code

- ❖ **D. Start and Stop Codons**
- ❖ **Start codon:**
- ❖ The codon AUG signals the start of translation and also codes for methionine.
- ❖ **Stop codons:**
- ❖ Codons like UAA, UGA, and UAG do not code for any amino acid and signal the end of translation, stopping the protein-building process. They are also called non-sense codons.

Second letter

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

Amino Acids with Most (6) Codons

- The amino acid Leucine (Leu), Serine (Ser) and Arginine (Arg) has the most codons in the standard genetic code, with six codons. These codons are:
- Leucine (Leu): CUU, CUC, CUA, CUG, UUA, UUG
- Serine (Ser): UCU, UCC, UCA, UCG, AGU, AGC
- Arginine (Arg): CGU, CGC, CGA, CGG, AGA, AGG
- This redundancy is part of the genetic code's "degeneracy," which provides help against mutations.

Amino Acids with 4 Codons

- The amino acids that are encoded by four codons in the standard genetic code are:
- Alanine (Ala): GCU, GCC, GCA, GCG
- Glycine (Gly): GGU, GGC, GGA, GGG
- Proline (Pro): CCU, CCC, CCA, CCG
- Threonine (Thr): ACU, ACC, ACA, ACG
- Valine (Val): GUU, GUC, GUA, GUG

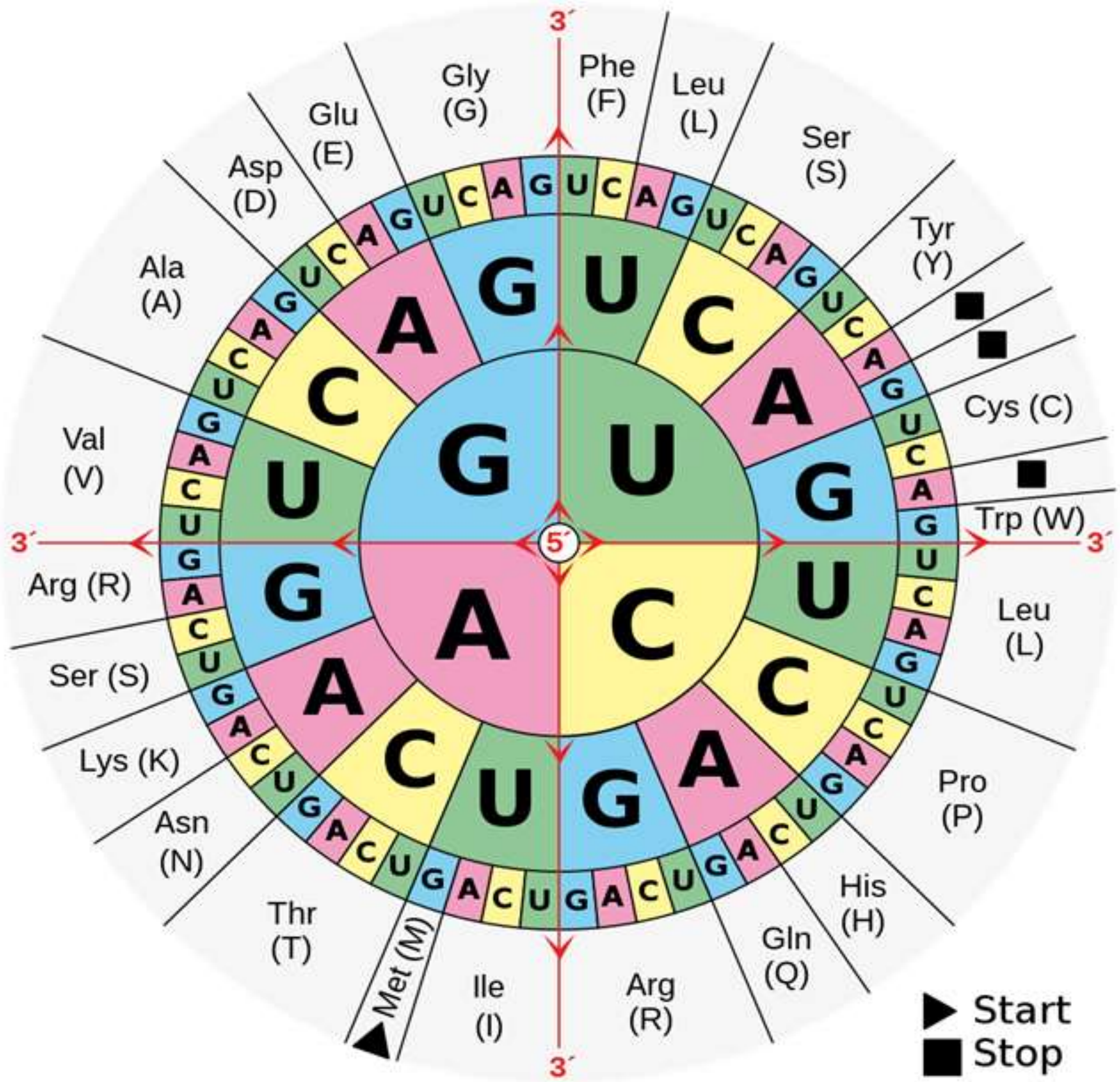
Amino Acids with 2 Codons

- The amino acids that are encoded by only two codons in the standard genetic code are:

Asparagine (Asn):	AAU, AAC
Aspartic acid (Asp):	GAU, GAC
Glutamine (Gln):	CAA, CAG
Glutamic acid (Glu):	GAA, GAG
Histidine (His):	CAU, CAC
Lysine (Lys):	AAA, AAG
Phenylalanine (Phe):	UUU, UUC
Tyrosine (Tyr):	UAU, UAC

Amino Acids with Only 1 Codon

- The amino acid that has only one codon in the standard genetic code is Methionine (Met), encoded by:
- AUG
- Additionally, Tryptophan (Trp) also has only one codon:
- UGG
- Methionine is particularly significant because AUG also serves as the start codon, marking the beginning of translation in most proteins.

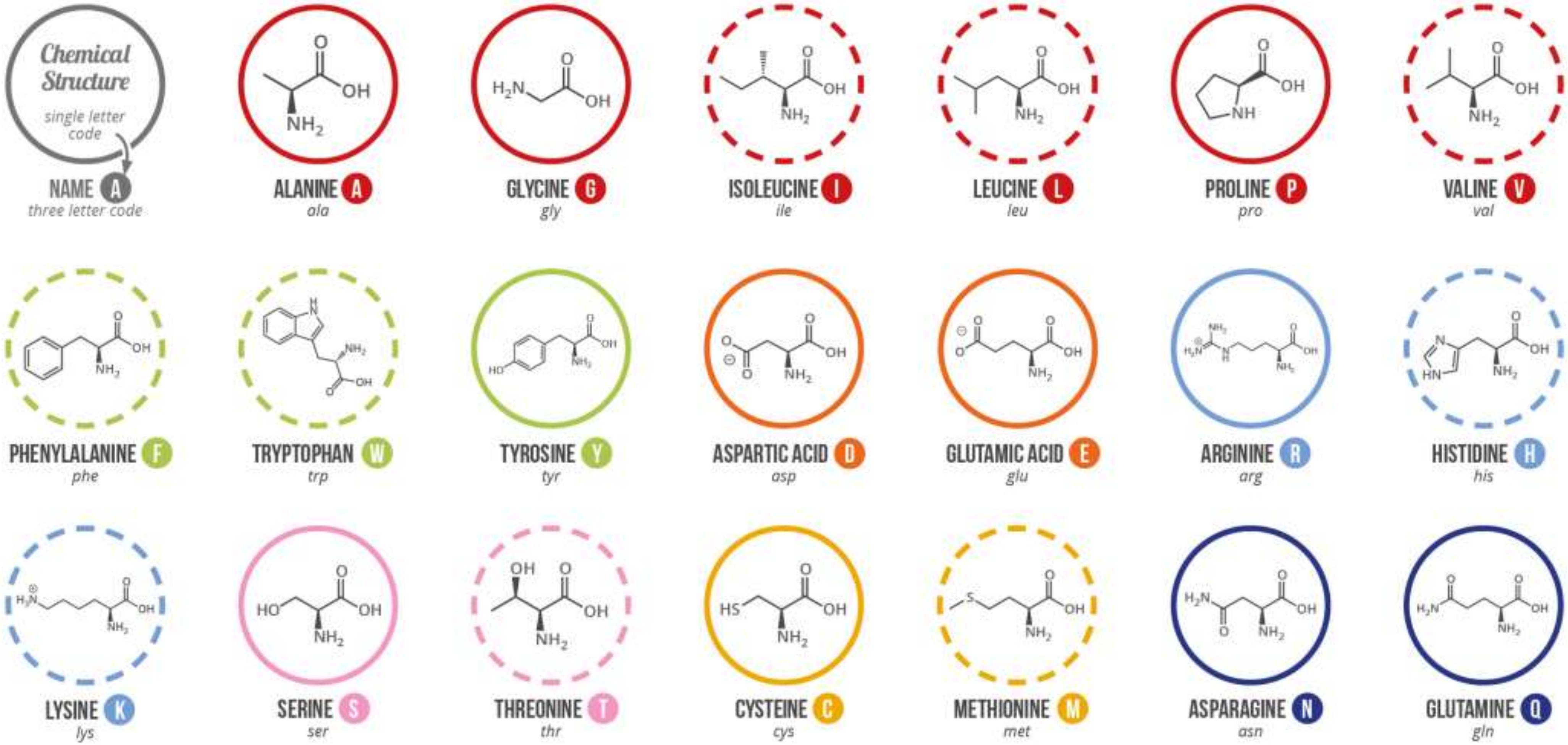


3´
5´
Phe (F)
Leu (L)
Ser (S)
Tyr (Y)
Cys (C)
Trp (W)
Leu (L)
Pro (P)
His (H)
Gln (Q)
Arg (R)
Ile (I)
Met (M)
Thr (T)
Asn (N)
Lys (K)
Ser (S)
Arg (R)
Val (V)
Ala (A)
Asp (D)
Glu (E)
Gly (G)
Start
Stop

A GUIDE TO THE TWENTY COMMON AMINO ACIDS

AMINO ACIDS ARE THE BUILDING BLOCKS OF PROTEINS IN LIVING ORGANISMS. THERE ARE OVER 500 AMINO ACIDS FOUND IN NATURE - HOWEVER, THE HUMAN GENETIC CODE ONLY DIRECTLY ENCODES 20. 'ESSENTIAL' AMINO ACIDS MUST BE OBTAINED FROM THE DIET, WHILST NON-ESSENTIAL AMINO ACIDS CAN BE SYNTHESISED IN THE BODY.

Chart Key: ALIPHATIC AROMATIC ACIDIC BASIC HYDROXYLIC SULFUR-CONTAINING AMIDIC NON-ESSENTIAL ESSENTIAL



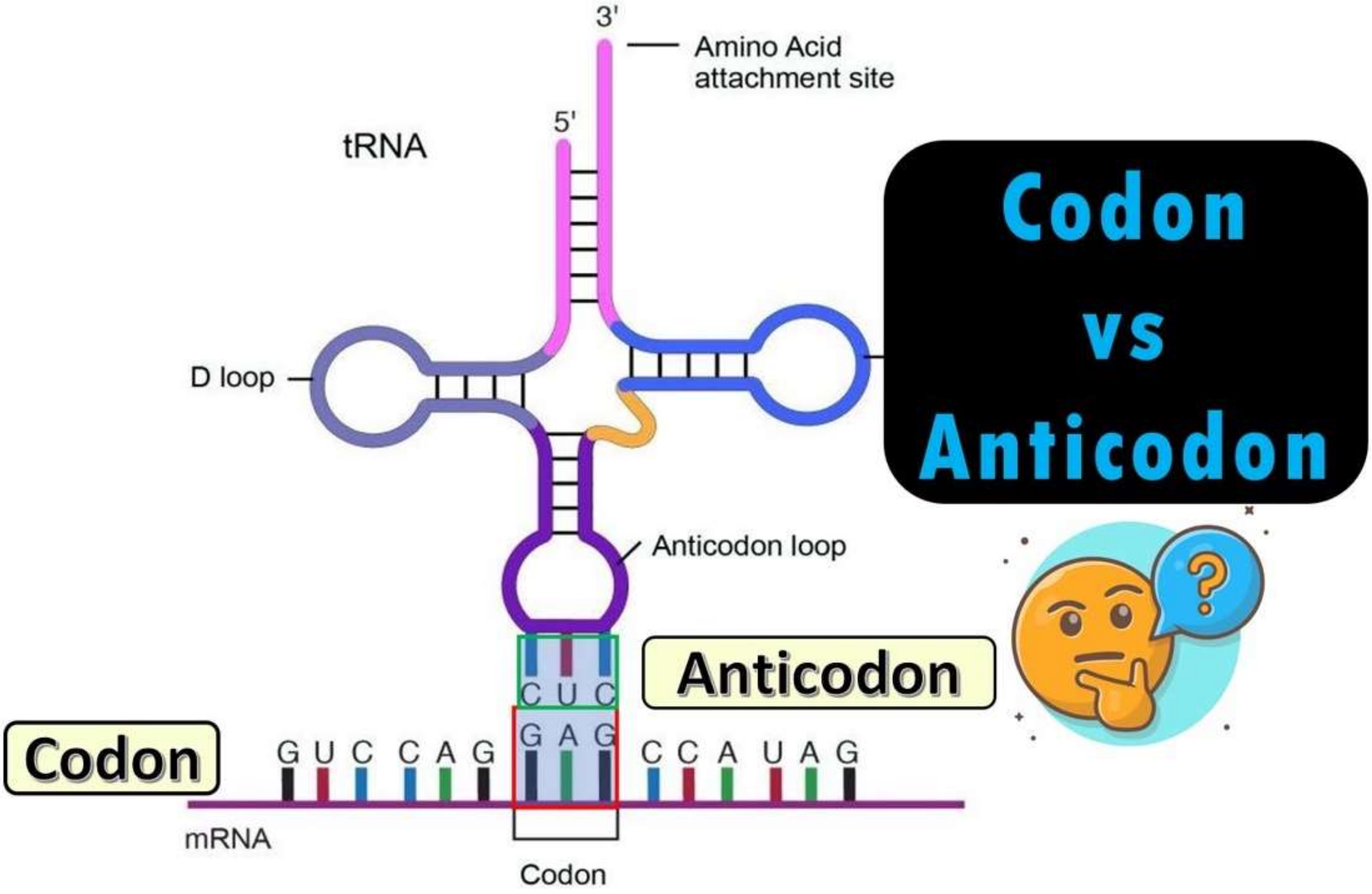
Note: This chart only shows those amino acids for which the human genetic code directly codes for. Selenocysteine is often referred to as the 21st amino acid, but is encoded in a special manner. In some cases, distinguishing between asparagine/aspartic acid and glutamine/glutamic acid is difficult. In these cases, the codes asx (B) and glx (Z) are respectively used.

Anticodon

- An anticodon is a sequence of three nucleotides found in a molecule of transfer RNA (tRNA). Its primary function is to pair with a complementary codon on a messenger RNA (mRNA) during the process of translation in protein synthesis.
- **Codon:** A triplet of bases on mRNA that specifies an amino acid.
- **Anticodon:** A triplet of bases on tRNA that is complementary (opposite) to the mRNA codon.

Anticodon

- ❖Each anticodon on tRNA matches a specific codon on mRNA.
- ❖The tRNA carries an amino acid that corresponds to the codon it pairs with.
- ❖This pairing helps ensure that the correct amino acid is added to the growing protein chain.
- ❖For example, if the codon on mRNA is AUG (which codes for methionine), the anticodon on tRNA will be UAC, and the tRNA will carry methionine.


3'
Amino Acid attachment site
5'
tRNA
D loop
Anticodon loop
Codon vs Anticodon
Anticodon
C U C
G A G
Codon
G U C C A G
C C A U A G
mRNA
Codon

Genetic code in mitochondrial DNA

- The genetic code in mitochondrial DNA (mtDNA) differs slightly from the standard (nuclear) genetic code.
- While the majority of organisms use the "universal" genetic code for translating messenger RNA into proteins, mitochondria have evolved unique variations to accommodate their specialized roles in energy production.
- These variations occur in the codon-to-amino acid assignments.

Key Features of the Mitochondrial Genetic Code

- **Different Stop Codon Usage:**
- UGA, which is a stop codon in the universal code, codes for tryptophan (Trp) in mitochondrial DNA.
- AGA and AGG, which typically code for arginine in the universal code, are stop codons in mammalian mitochondrial DNA.
- **Start Codons:**
- In addition to the universal start codon AUG (methionine), mitochondria may use AUA (which typically codes for isoleucine) as a start codon for methionine.

Key Features of the Mitochondrial Genetic Code

- Changes in Codon Assignments:
- **AUA:** Codes for methionine (instead of isoleucine in the standard code).
- **UGA:** Codes for tryptophan (instead of a stop codon).
- **AGA and AGG:** Code for stop codons (in mammals mitochndria, rather than arginine).
- **CUU, CUC, CUA, and CUG:** These codons still code for leucine, but there are some differences in usage among different species' mitochondrial codes.

Variation Among Species

- The mitochondrial genetic code varies among organisms. For example:
- In humans, the specific modifications above apply.
- In yeast mitochondria, there are different variations (e.g., CUN codons code for threonine rather than leucine).
- In protozoans and some invertebrates, additional codon reassignments exist.

Universality of the Genetic Code

- The genetic code is nearly universal but not entirely.
- In most organisms, the genetic code is consistent: the same codons specify the same amino acids across the vast majority of species. This universality is one of the strongest pieces of evidence for the shared evolutionary origin of life.
- **For example:**
- AUG codes for methionine (start codon in most organisms).
- UUU codes for phenylalanine in virtually all species.

Exceptions to Universality

- Some organisms and organelles use variations of the standard genetic code. These exceptions often occur in specialized environments or organelles and include:
- **Mitochondria:**
- Mitochondria have their own genetic code, which differs slightly from the universal code.
- **Example:** UGA codes for tryptophan in human mitochondria, whereas it is a stop codon in the universal code.
- **Ciliates:**
- Certain ciliates (single-celled eukaryotes) reassign stop codons:
- UAA and UAG may code for glutamine instead of serving as stop codons.

Exceptions to Universality

- **Yeast Mitochondria:**
- CUN codons code for threonine instead of leucine.
- **Mycoplasma and Some Bacteria:**
- UGA may code for tryptophan instead of being a stop codon.
- **Nuclear Code Variants in Some Organisms:**
- Specific organisms, like some protists, have slight variations in their nuclear genetic codes.

Reasons for Variations

- **Evolutionary Adaptation:**
- Variations likely arose due to the evolutionary pressure to optimize translation for specific environments or organelles.
- **Isolation:**
- Organelles like mitochondria originated from endosymbiotic bacteria, which allowed the development of distinct genetic codes.
- **Reduced Complexity:**
- Mitochondria and other organelles often have smaller genomes and simplified translation machinery, which led to reassignment of codons.

Thank You!