

GENETIC CODE

The number and the sequences of bases in mRNA, specifying an amino acid, is known as codon.

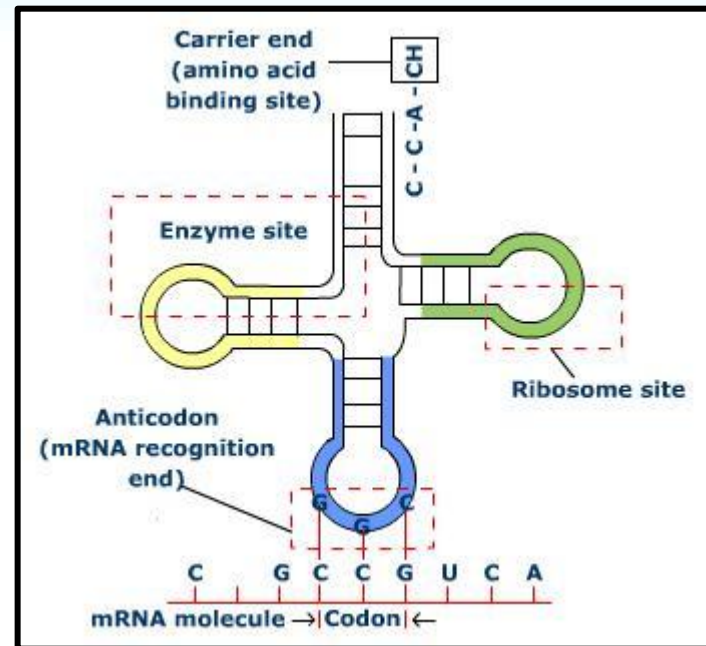
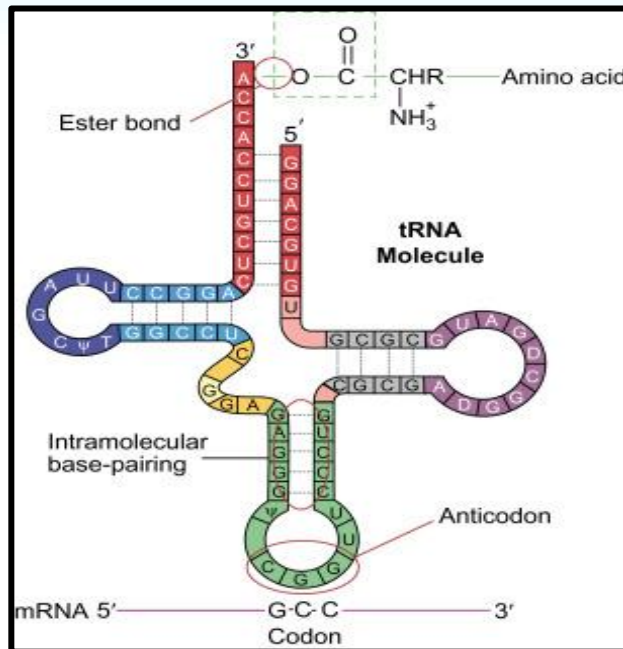
m-RNA

AUGCGGAAAUUCCACUUGGCU



Polypeptide met-arg-lys-phe-his-leu-ala

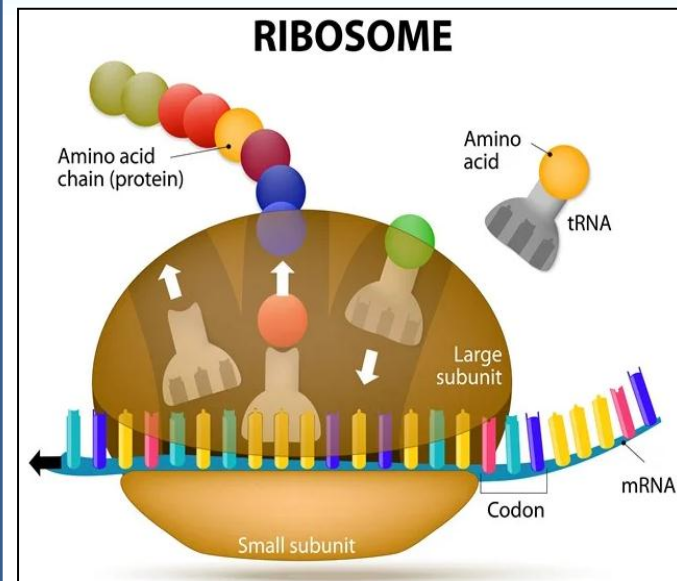
The set of bases in a tRNA, that base pair with a codon of an mRNA is known as anticodon.



For example the codon 5' GCC 3' has the anticodon 3' CGG 5'.

Essential	Conditionally Non-Essential	Non-Essential
Histidine	Arginine	Alanine
Isoleucine	Asparagine	Asparatate
Leucine	Glutamine	Cysteine
Methionine	Glycine	Glutamate
Phenylalanine	Proline	
Threonine	Serine	
Tryptophan	Tyrosine	
Valine		
Lysine		

The set of all the codons, that specify the 20 amino acids, is termed as the genetic code, genetic language or coding dictionary.



		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 Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	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 } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg 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

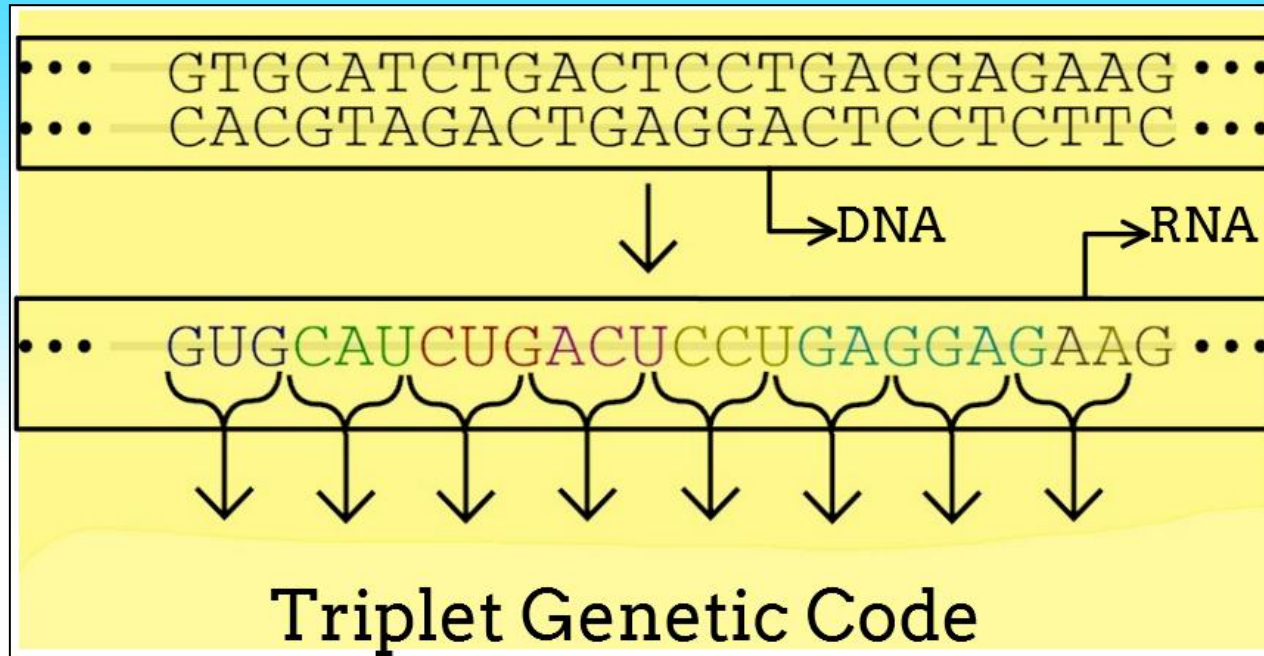
Four bases – A, T, G & C

20 different amino acids

Two bases per codon would yield only $4^2 = 16$ possible codon – **clearly not enough**

Three bases per codon yield $4^3 = 64$ possible codon – **an apparent excess**

The code is a triplet code – each mRNA codon that specifies an amino acid in a polypeptide chain consists of three nucleotides.



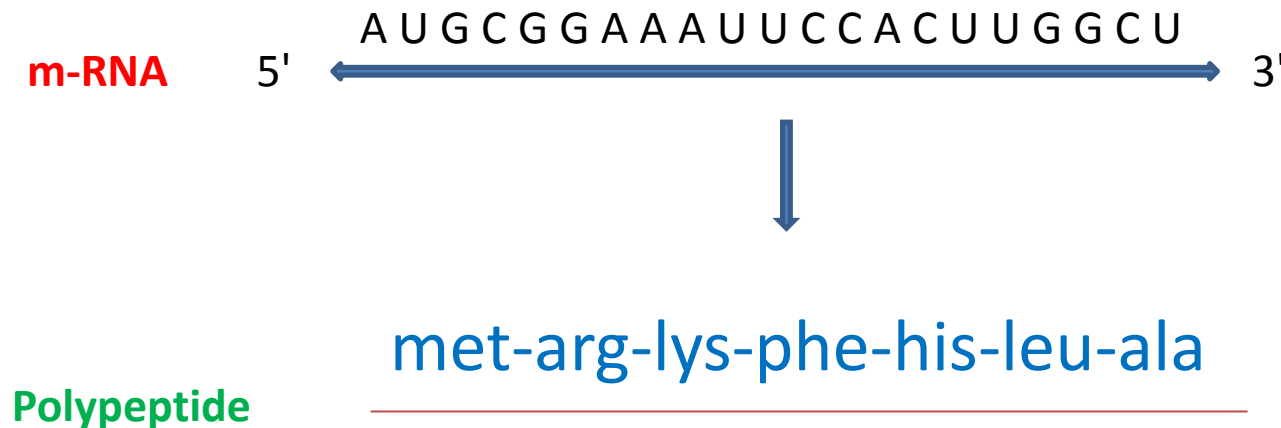
m-RNA 5' AUGCGGAUAUCCACUUGGCU 3'



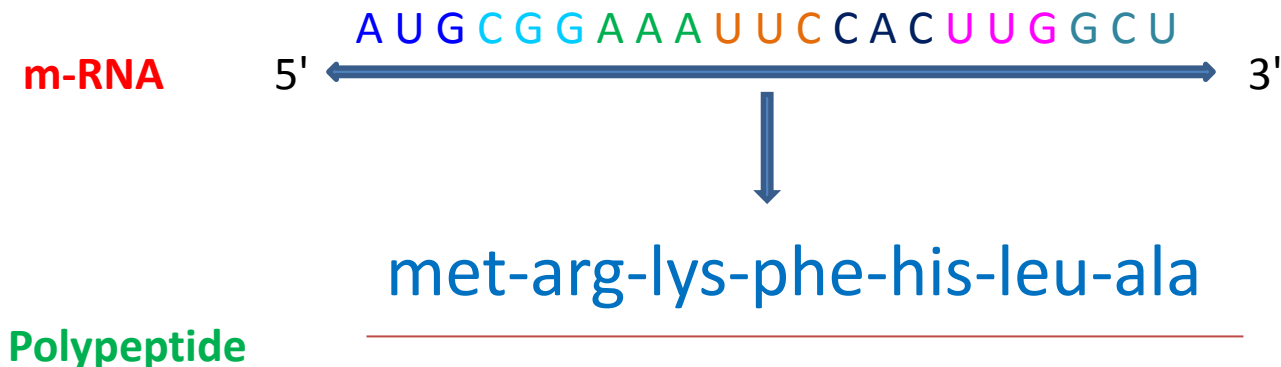
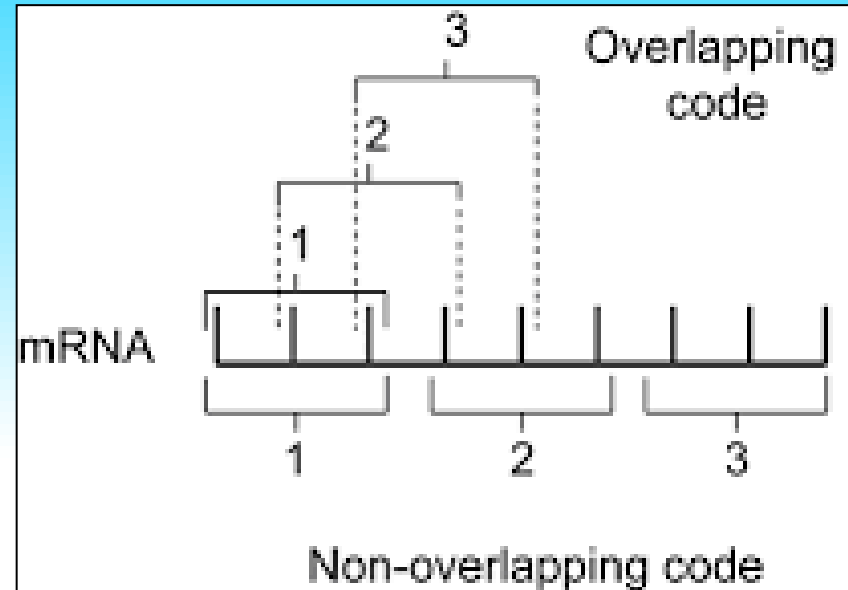
Polypeptide

met-arg-lys-phe-his-leu-ala

The code is comma free – the mRNA is read continuously. Three nucleotides (one codon) at a time without skipping any nucleotides of the message.



The code is non-overlapping – the mRNA is read in successive groups of three nucleotides



The code is almost universal – all organisms share the same genetic language.

**Thus lysine is coded by AAA or AAG
arginine by CGU, CGC, CGA, CGG, AGA
& AGG & so on.**

**The code, however is not completely universal.
For example the mitochondria of some organisms
such as mammals have minor changes in the code.**

The code is degenerate –

Two exceptions – AUG (methionine) & UGG (tryptophan).

More than one codon occurs for each amino acid. This multiple coding is called the degeneracy of genetic code.

UUU UUC	Phe	CAU CAC	His	UUA UUG	Leu	CAA CAG	Glu
CUU CUC CUA CUG	Leu	CGU CGC CGA CGG	Arg	GCU GCC GCA GCG	Ala	GGU GGC GGA GGG	Gly

All codons are not used equally. For example, UUC is more frequently used codon for phenyl alanine followed by UUU.

The code has start and stop signals –

AUG - start codon- specifying Methionine

**GUG – (Start codon in rare cases) – then specify Methionine;
But code valine when used as intermediate codon.**

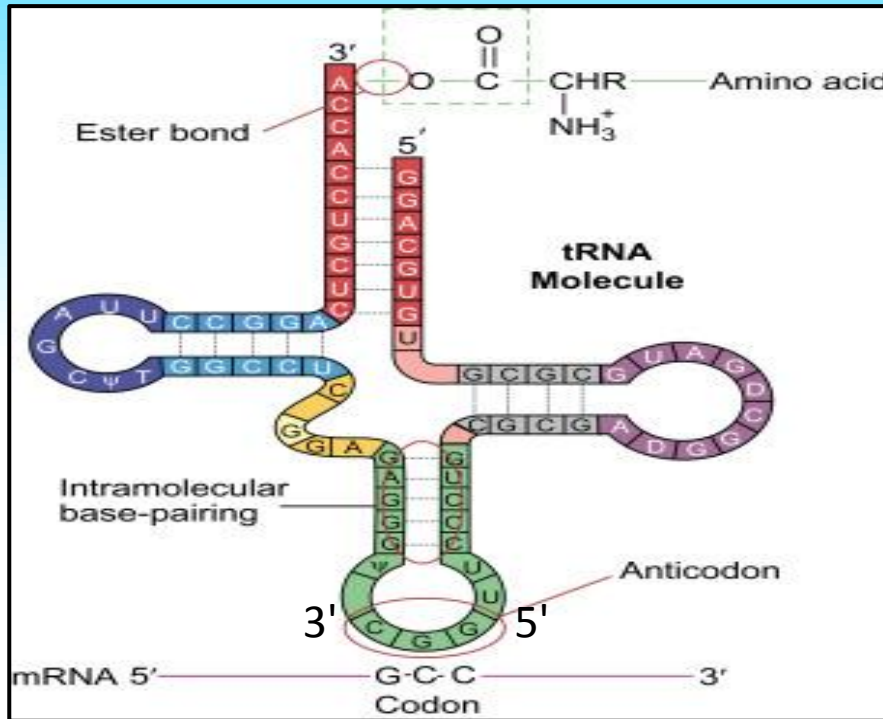
Only 61 of the 64 codons specify amino acids, these codons are called **sense codons.**

The three other codons – **UAG (Amber), **UAA** (Ochre) & **UGA** (Opal) do not specify an amino acid and no tRNA in normal cell carry the appropriate anticodon. These three codons are the **stop or non-sense codons or chain terminating codons**.**

Wobble occurs in the anticodon –

61 sense codons can be read by fewer than 61 distinct tRNA

The wobble hypothesis proposed by F. Crick



Base at the 5' end of the anticodon (complementary to the base of the 3' end of the codon) is not as constrained as the other two bases. This feature allows for less exact base pairing, so that the base at the 5' end of anticodon can potentially pair with one or three different bases at the 3' end of the codon.