

The genetic code

The genetic code links the nucleotide base sequence in mRNA to amino acids in a protein. Every three nucleotides in mRNA is called a **codon**. A codon is the unit of the genetic code and each codon codes for one amino acid (or stop signal) in the synthesis of a protein molecule.

There are 64 codons (4^3) as there are four (4) bases (A, U, G and C) and three (3) positions in a codon. Protein molecules are synthesised using a combination of twenty (20) amino acids. The genetic code is described as **degenerate** as some amino acids have several codons. The genetic code is also described as **non-overlapping** as nucleotide bases are not shared between adjacent codons.

The genetic code on mRNA is shown in the table below.

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

The names and codes of the twenty amino acids are shown in the table below.

ala	alanine	gly	glycine	pro	proline
arg	arginine	his	histidine	ser	serine
asn	asparagine	ile	isoleucine	thr	threonine
asp	aspartic acid	leu	leucine	trp	tryptophan
cys	cysteine	lys	lysine	tyr	tyrosine
gln	glutamine	met	methionine	val	valine
glu	glutamic acid	phe	phenylalanine	stop	stop codon