= 16 amino acids . A code of 3 nucleotides could code for a maximum of 43 =
64 amino acids .

The Crick , Brenner et al. experiment first demonstrated that codons consist of three DNA bases ; Marshall Nirenberg and Heinrich J. Matthaei were the first to elucidate the nature of a codon in 1961 at the National Institutes of Health . They used a cell @-@ free system to translate a poly @-@ uracil RNA sequence (i.e. , UUUUU ...) and discovered that the polypeptide that they had synthesized consisted of only the amino acid phenylalanine . They thereby deduced that the codon UUU specified the amino acid phenylalanine . This was followed by experiments in Severo Ochoa 's laboratory that demonstrated that the poly @-@ adenine RNA sequence (AAAAA ...) coded for the polypeptide poly @-@ lysine and that the poly @-@ cytosine RNA sequence (CCCCC ...) coded for the polypeptide poly @-@ proline . Therefore , the codon AAA specified the amino acid lysine , and the codon CCC specified the amino acid proline . Using different copolymers most of the remaining codons were then determined . Subsequent work by Har Gobind Khorana identified the rest of the genetic code . Shortly thereafter , Robert W. Holley determined the structure of transfer RNA (tRNA) , the adapter molecule that facilitates the process of translating RNA into protein . This work was based upon earlier studies by Severo Ochoa , who received the Nobel Prize in Physiology or Medicine in 1959 for his work on the enzymology of RNA synthesis .

Extending this work , Nirenberg and Philip Leder revealed the triplet nature of the genetic code and deciphered the codons of the standard genetic code . In these experiments , various combinations of mRNA were passed through a filter that contained ribosomes , the components of cells that translate RNA into protein . Unique triplets promoted the binding of specific tRNAs to the ribosome . Leder and Nirenberg were able to determine the sequences of 54 out of 64 codons in their experiments . In 1968 , Khorana , Holley and Nirenberg received the Nobel Prize in Physiology or Medicine for their work .

= = Features = =

= = = Reading frame = = =

A codon is defined by the initial nucleotide from which translation starts and sets the frame for a run of uninterrupted triplets , which is known as an " open reading frame " (ORF) . For example , the string GGGAAACCC , if read from the first position , contains the codons GGG , AAA , and CCC ; and , if read from the second position , it contains the codons GGA and AAC ; if read starting from the third position , GAA and ACC . Every sequence can , thus , be read in its 5 ' ? 3 ' direction in three reading frames , each of which will produce a different amino acid sequence (in the given example , Gly @-@ Lys @-@ Pro , Gly @-@ Asn , or Glu @-@ Thr , respectively) . With double @-@ stranded DNA , there are six possible reading frames , three in the forward orientation on one strand and three reverse on the opposite strand . The actual frame from which a protein sequence is translated is defined by a start codon , usually the first AUG codon in the mRNA sequence .

In eukaryotes , ORFs in exons are often interrupted by introns .

= = = Start / stop codons = = =

Translation starts with a chain initiation codon or start codon . Unlike stop codons , the codon alone is not sufficient to begin the process . Nearby sequences such as the Shine @-@ Dalgarno sequence in E. coli and initiation factors are also required to start translation . The most common start codon is AUG , which is read as methionine or , in bacteria , as formylmethionine . Alternative start codons depending on the organism include " GUG " or " UUG " ; these codons normally represent valine and leucine , respectively , but as start codons they are translated as methionine or formylmethionine .

The three stop codons have been given names : UAG is amber , UGA is opal (sometimes also called umber) , and UAA is ochre . " Amber " was named by discoverers Richard Epstein and

Charles Steinberg after their friend Harris Bernstein , whose last name means " amber " in German . The other two stop codons were named " ochre " and " opal " in order to keep the " color names " theme . Stop codons are also called " termination " or " nonsense " codons . They signal release of the nascent polypeptide from the ribosome because there is no cognate tRNA that has anticodons complementary to these stop signals , and so a release factor binds to the ribosome instead .

= = = Effect of mutations = = =

During the process of DNA replication , errors occasionally occur in the polymerization of the second strand . These errors , called mutations , can affect the phenotype of an organism , especially if they occur within the protein coding sequence of a gene . Error rates are usually very low ? 1 error in every 10 ? 100 million bases ? due to the " proofreading " ability of DNA polymerases .

Missense mutations and nonsense mutations are examples of point mutations , which can cause genetic diseases such as sickle @-@ cell disease and thalassemia respectively . Clinically important missense mutations generally change the properties of the coded amino acid residue between being basic , acidic , polar or non @-@ polar , whereas nonsense mutations result in a stop codon .

Mutations that disrupt the reading frame sequence by indels (insertions or deletions) of a non @-@ multiple of 3 nucleotide bases are known as frameshift mutations . These mutations usually result in a completely different translation from the original , and are also very likely to cause a stop codon to be read , which truncates the creation of the protein . These mutations may impair the function of the resulting protein , and are thus rare in in vivo protein @-@ coding sequences . One reason inheritance of frameshift mutations is rare is that , if the protein being translated is essential for growth under the selective pressures the organism faces , absence of a functional protein may cause death before the organism is viable . Frameshift mutations may result in severe genetic diseases such as Tay @-@ Sachs disease .

Although most mutations that change protein sequences are harmful or neutral , some mutations have a beneficial effect on an organism . These mutations may enable the mutant organism to withstand particular environmental stresses better than wild type organisms , or reproduce more quickly . In these cases a mutation will tend to become more common in a population through natural selection . Viruses that use RNA as their genetic material have rapid mutation rates , which can be an advantage , since these viruses will evolve constantly and rapidly , and thus evade the defensive responses of e.g. the human immune system . In large populations of asexually reproducing organisms , for example , E. coli , multiple beneficial mutations may co @-@ occur . This phenomenon is called clonal interference and causes competition among the mutations .

= = = Degeneracy = = =