# Deciphering the Genetic Code

Marshall Nirenberg*, Laboratory of Biochemical Genetics, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, 20892, USA.

The first question I asked as an independent investigator in 1958 using a bacterial cell-free protein synthesizing system was, "Does DNA or RNA, transcribed from DNA, code for protein synthesis?" We found that a fraction of *E. coli* ribosomal RNA, but not DNA, stimulated amino acid incorporation into protein, and that poly U, tested with 20 different amino acid solutions each with 19 cold and 1 radioactive amino acids, stimulated the incorporation of only phenylalanine into protein. These *in vitro* results demonstrated the existence of mRNA, showed that a series of Us corresponds to an RNA codon for phenylalanine, and suggested that other polynucleotides could be used to decipher other codons. We then showed that aminoacyl-tRNA is required for the synthesis of protein and that single-stranded poly U functions as mRNA, but not double- or triple-stranded poly U-poly A helices. This was the first antisense RNA experiment. Between 1961 and 1964 Severo Ochoa's laboratory and my laboratory determined the base compositions of codons for the other amino acids. We then showed that trinucleotides stimulate the binding of appropriate species of aminoacyl-tRNA to ribosomes and used this assay to determine the nucleotide sequences of RNA codons. We synthesized the 64 trinucleotides using 2 enzymatic methods and found that the third bases of synonym codons vary systematically and we identified 4 patterns of degeneracy. We determined the nucleotide sequences of 54 of the 64 codons; then Gobind Khorana and his colleagues reported the nucleotide sequences of RNA codons using 64 chemically synthesized trinucleotides. Khorana and his colleagues also synthesized RNA with 2, 3, or 4 base repeats and used the RNAs to direct protein synthesis. Clark and Marcker showed that N-formyl-methionine-tRNA initiates protein synthesis by recognizing AUG. UAA, UAG, and UGA were shown to correspond to termination of protein synthesis by Brenner and by Garen and their colleagues. The 21$^{st}$ amino acid, selenocysteine, recognizes UGA only if there is a downstream stem-loop in the mRNA. Pyrrolysine, the 22$^{nd}$ amino acid found only in a few species of bacteria recognizes UAG. We showed that the genetic codes of *E. coli, Xenopus,* and the hamster are identical. Subsequently, small variations were found in the genetic codes of some organisms and in mitochondria. Nevertheless, the genetic codes used by all forms of life studied are very similar. These results strongly suggest that the genetic code appeared very early during evolution, that all forms of life on Earth descended from a common ancestor, and therefore that all living things on Earth are related to one another. The molecular language solves the problem of biological time, for it is easier to construct a new organism using the information stored in DNA than it is to repair an aging, malfunctioning one.