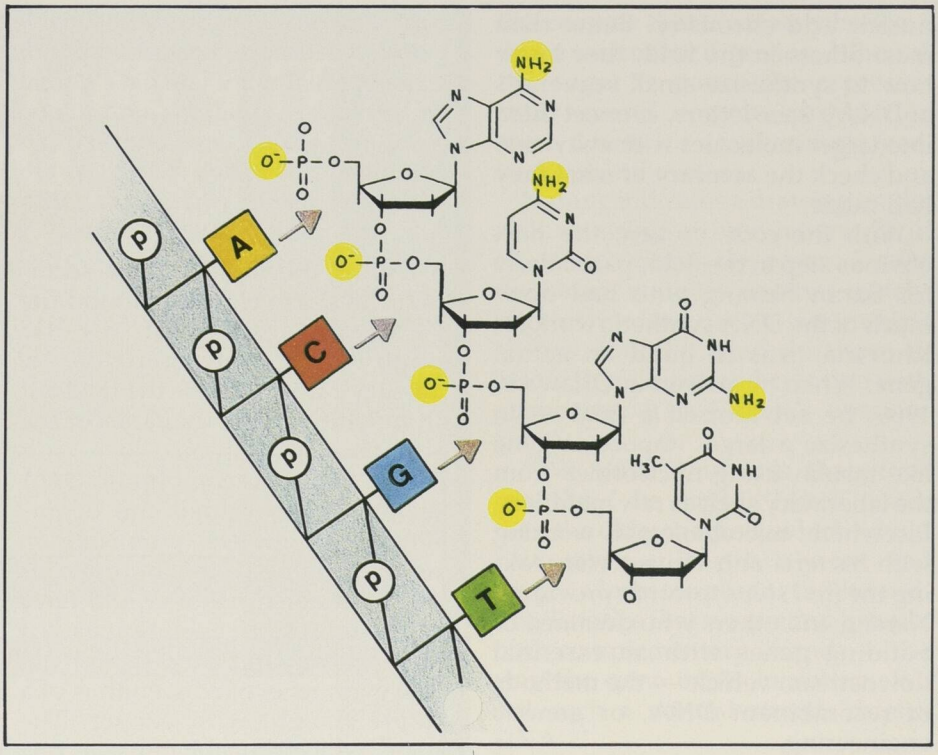
fer. They knew that the code in the DNA letter banks is first transcribed into another kind of linear molecular message — appropriately called "messenger" ribonucleic acid or messenger RNA — and transported out of the nucleus where it is translated into proteins. Like DNA, proteins are linear molecules too (at least in their primary structure), made up of 20-odd different kinds of smaller molecules called amino acids. From proteins, all else in life derives. They make up the struts and beams of all living structures, they are the enzymes that make the chemistry of life go, and they form a vast array of cellular messengers. Insulin is one of these. What Khorana and his team wanted to know was how the A, C, T and G bases code for the 20 amino acids. This relation is life's most ancient lexicon, the genetic code, created so early in evolution that it holds for every living species on our planet. "They were exciting days," recalls Narang. "Especially as we were in a tight race with other laboratories. Eventually, we confirmed what biologists had inferred from purely mathematical considerations — that each amino acid is coded for by a "triplet" of three bases on the DNA. For example, the amino acid methionine is coded for by the sequence A-T-G. If you have a protein like insulin, with 51 amino acids, its genetic code is 153 bases long on the DNA chain, or a string of 51 3-base codons, one after another in sequence." Equally important, the group showed it didn't really matter what type of cell was tested — bacterium, human, sunflower — the code was always the same. All of the variety of life issued from the same information processing system. Only the data were different.

The cracking of the code and proof of its universality marked a watershed in molecular biology comparable in importance to the discovery of DNA's double helix structure. In 1968, Gobind Khorana shared the Nobel prize in physiology and medicine for his work along with Marshall Nirenberg of the National Institutes of Health in Bethesda, Maryland, and Robert Holley of Cornell University in Ithaca, New York.

The skill that allowed Khorana's people to move forward in this field so quickly was in their ability to do

Building a strand of DNA is basically a masking job. The sylized DNA at left is much more complicated when you look at its elemental structure, shown on the right. The code letters, A, C, G, and T are actually nitrogen-containing ring compounds (A and G have two rings, C and T have only one) linked to the sugar ring deoxyribose which in turn is connected to a phosphate group. These three member units, called nucleotides, are the shelf materials that chemists use to build a strand of DNA; they are linked one after another to the lengthening chain by joining phosphate to deoxyribose in regular sequence. But, before making these nucleotide connections, chemists must mask certain 'active' groups (yellow) that would otherwise interfere with phosphate-sugar bonding.



John Bianchi

The A-chain of human insulin, in the form required to splice it into a plasmid, the small loops of bacterial DNA used to clone genes. The AATTC sequence on the left end is an enzyme 'recognition site' used to insert the gene into the plasmid, as is the right end CCTAG sequence (see plasmid graphic, page 25). The ATG immediately following is the universal signal to START transcribing the gene message, while the TGA at the other end is the signal to STOP. The 'triplet' sequence of bases in between specify the amino acids that make up the A-chain of insulin (the first 10 amino acids are listed along the top). Bioscientists clone the three proinsulin subunit chains in this way (A, B, and C) to obtain enough material to link them in the proper sequence of the complete gene, B-C-A).