The Artificial Gene: It's Synthesized and It Works in Cells

The first wholly artificial gene that functions in a living cell has been synthesized by Nobel laureate Har Gobind Khorana and his associates at the Massachusetts Institute of Technology. The synthetic gene has been shown to correct a mutational defect in bacteriophage lambda, a virus that infects bacteria, by replacing a defective gene. The synthesis and demonstration that the gene works have been hailed by other scientists as a major accomplishment in genetics.

The complete synthesis, announced at the recent Centennial Meeting of the American Chemical Society by Hans-Joachim Fritz and Ramamoorthy Belagaje of MIT, was the fruit of 9 years of work by 24 postdoctoral fellows in Khorana's laboratory. Their approach to the project has been substantially different from that of most other investigators who have synthesized large polynucleotides. Others generally use a naturally occurring polynucleotide as a template to guide the activity of bacterial enzymes that form a complementary copy of the template. Khorana's group, however, chemically synthesized 39 short fragments—each containing 10 to 15 nucleotides—that, together, comprise both strands of the double-stranded gene.

These segments were chosen so that a short piece of single-stranded DNA extends from each end when two complementary fragments are allowed to form a complex. The extended segment then serves as a splint to which an adjoining fragment can be complexed for joining by a DNA ligase. In this way, they constructed a complete double-stranded unit with an overall length of 207 nucleotides.

The biologically active product for which the gene codes is tyrosine transfer RNA (a form of RNA that attaches to the amino acid tyrosine and carries it to the ribosome, where it is incorporated into a growing protein chain), which contains 86 nucleotides. The gene proper, whose sequence was determined by John Smith and Sidney Altman at the Medical Research Council in Cambridge, England, contains 126 nucleotides. The initial product of the gene also contains 126 nucleotides, but for some as yet unknown reason, the extra 41-nucleotide segment is cleaved away by an enzyme within the cell. The synthesis of this portion of the gene was completed 3 years ago (*Science*, 28 September 1973, p. 1235).

The rest of the synthetic gene is composed of a 56-nucleotide promoter or start signal on one end and a 25-nucleotide terminator or stop signal on the other end. Khorana's group has spent the last 3 years determining the sequence of these segments and synthesizing them. It was the addition of these segments that made the gene functional.

That functionality was tested both in test tube experiments and by incorporating the gene into a mutant strain of bacteriophage lambda. This strain, when allowed to infect the bacterium *Escherichia coli*, produces short, nonfunctional proteins, apparently as the result of a defect in the terminator signal for the tyrosine transfer RNA gene. As a result, the mutant normally does not proliferate in *E. coli*. But when the synthetic gene was incorporated into the mutant bacteriophage, the mutant proliferated as well as normal bacteriophage lambda in the bacteria, indicating both that functional proteins were being produced and that the gene was functioning.

Correcting genetic defects in a bacteriophage is not, of course, the goal of the experiment. Rather, that goal is to learn more about the gene itself and how it is regulated. The chemical synthesis of a gene, Khorana argues, is a much more versatile technique than the use of a template because it allows changes in the gene sequence to be made at will. Individual nucleotides or long sequences can be altered in a known fashion, and the effects of these alterations observed in the test tube or in the cell. It should be particularly interesting to study such changes in the regulatory segments of the gene; it might be possible in this manner to learn much more about the mechanisms by which gene expression is controlled and, eventually, about why certain genes in cells such as cancer cells appear to be out of control. It might also make possible new methods of obtaining biological products. One suggestion is to incorporate a gene for insulin in a bacterium such as *E. coli* so that the valuable protein could be harvested from bacterial fermentation instead of from animals.

A particularly intriguing aspect of Khorana's work was the discovery of large regions of symmetry in the regulatory sequences. That is, a sequence of bases in one part of the single-stranded gene is complementary to another sequence of bases, read backwards, further down the strand (Fig. 1). A very high degree of symmetry was found in the promoter, a lesser degree in the terminator. Khorana suggests that the regions of symmetry might permit the regulatory sequence of a single-stranded gene to fold back on itself and thereby form a three-dimensional shape that could be recognized by the enzymes that use the gene as a template for the synthesis of RNA.

Khorana has been very careful to draw a distinction between his studies and those on recombinant DNA. In the latter case, naturally occurring genes from different organisms are joined together to study the expression of genes, to alter the characteristics of the organism in which the gene resides, or to produce large amounts of DNA. Khorana's work, at least so far, involves the synthesis of a gene that is already present in, and absolutely necessary to, all living cells. The technique could, however, be used to synthesize and introduce into an organism a gene that does not occur naturally in that organism. This technique is more tedious and time-consuming than that used in recombining DNA, but the results—and the potential for hazard—should be about the same. It may thus be that Khorana's techniques will be closely examined by some of the scientists who urge a limitation on work with recombinant DNA.

—THOMAS H. MAUGH II

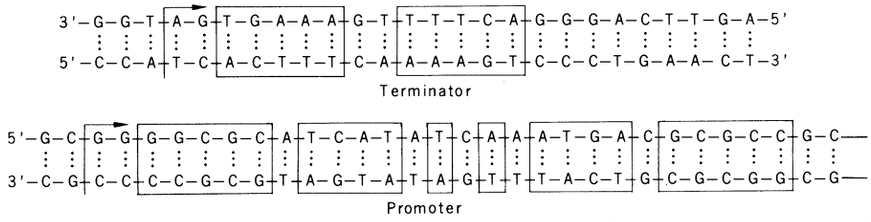

Fig. 1. Nucleotide sequences of the terminator region and part of the promoter region of tyrosine transfer RNA. The regions of symmetry are enclosed by the boxes. (A, adenine; T, thymine; G, guanine; C, cytosine.)