NEWS AND VIEWS

The longest synthetic gene. . .so far

LESS than two years ago, an attempt to synthesize a human interferon gene would have literally been impossible. For it is just two years since nucleotide sequences of interferon genes were first determined by using reverse transcriptase to make DNA from messenger-RNA molecules taken from cells induced to produce interferon. Since then, similar DNA copies of mRNA have made it possible to show that there is a whole family of interferon genes in the human genome — several corresponding to leukocyte interferon (the α variety) and fewer corresponding to fibroblast interferon (called β) and the more elusive immune interferon (γ). Inevitably, the amino acid sequence of each of these interferons has not yet been worked out directly, by chemical analysis; instead, the protein structure has mostly been inferred by applying the genetic code to the sequence of nucleotides in the gene. So it is, to say the least, a technical triumph that a group assembled by ICI should now have carried through the total synthesis of a human interferon gene (see Edge *et al.*, p.756).

This gene is the largest so far synthesized, no fewer than 514 nucleotides long. Given that it has been necessary to string together the nucleotides of both the complementary DNA strands, more than 1,000 nucleotides have been assembled in a predetermined way.

The chemistry of the synthesis is more easily described than carried out. The starting point is an initial nucleotide coupled to a polyacrylamide resin. Further nucleotides are added in pairs, by means of a coupling agent in anhydrous pyridine, and the authors say that the time taken for each cycle is a mere hour and a half. At that rate, ambitious people might calculate, it should be possible to assemble 5,000 nucleotides in a year, or a whole genome in . . ., well, a few centuries. Calculations such as these show that even the modifications of DNA-synthesizing techniques worked out by the ICI group fall a long way short of assembling whole genomes.

Indeed, as now described, short sections of each strand 15 or more nucleotides in length have to be made separately so as to keep the yield of the synthesis respectably high, and then short pieces of the double helix assembled before being joined together in the complete gene. That the gene is complete has been shown by its incorporation into a plasmid, its amplification and subsequent analysis. Edge and his colleagues do not say whether their plasmids have been able to direct the synthesis of proteins — and if so, what they are — but *something* is produced, for a further report on its biological activity is promised. It would be a great surprise if the product were not α-interferon.

The gene now synthesized is not however a naturally occurring gene but an analogue, designed so as to simplify the synthesis, but also to avoid combinations of neighbouring nucleotides that may impede expression of the gene in its unnatural environment of the plasmid. It should make one of the versions of α-interferon and, with luck, should do so more efficiently. But it would be sheer good fortune if the gene now synthesized turned out to be a superabundantly efficient interferon gene. The importance of this development is not the gene itself, but the demonstration of the technique. It is a little like climbing Everest, known to be easier the second time.

But these days, it is possible to buy machines that will automatically assemble nucleotides in some predetermined sequence. What can be so special about the synthesis of this artificial gene? For one thing, the simple assembly of a score or so of nucleotides is only half the battle. Purification is a tedious and even uncertain business. Prudent people make sure that the nucleotide sequence they have made is indeed the one they set out to make by subsequent analysis. But the most important need in the construction of a largish gene is an intelligent strategy for deciding what pieces of the gene to make, and in what order to assemble them. The use of computers to spot sequences in the target gene that might cause trouble during its assembly is also obligatory.

This does not mean that the work now reported is simply a technical *tour de force*. In the past few years, synthetic pieces of genes have been widely used in several different ways — as probes for finding within some intact genome a piece of genetic material containing some particular sequence, for example. Those whose business is to insert foreign genes into plasmid DNA increasingly find it safest to resynthesize the ends of the genes back to the nearest convenient restriction site so as to ensure that nucleotides are not lost when the tails needed for ligation to the plasmid DNA are added chemically. Fashioning pieces of DNA expected to have important functions such as those acting as sites for the promotion of replication has also become fashionable and valuable. But for a group with its roots in the pharmaceutical industry, it must be especially pleasing that it is now possible to plan systematically to vary the constitution of a gene for the sake of the variations of protein structure that result. Is there just a trace of nostalgia in the statement that "classical medicinal structure–activity analysis should [now] be possible with large peptides"?

The prospect is exciting. There is already ample evidence, derived from the experience of traditional pharmacology, that natural products can usually be modified to make more efficient drugs. The changes rung on the structure of the naturally occurring antibiotics in the past several decades are a sign of that. The interferons promise to be a nightmare for pharmacologists — whatever their biological effects, unspecificity seems to be their common characteristic. So it is natural that in the past few months, there has been great interest in hybrid molecules of interferon — one end from one and the other from another. Now it should be possible to set about making more subtle modifications, not those that depend on where some restriction site within the gene happens to fall. The result should help throw light on the functions of different parts of the interferon molecule. Yet improving on naturally occurring interferon is bound to be a daunting task. While reliable biological assays of the effectiveness of molecules related to interferon are few and far between, and while the efficiency with which interferon genes embodied in plasmids yield protein product is as low as at present, it would be hoping for too much that the uncertainties that abound in the genetic engineering of medicinal interferon will quickly be cleared up.

0028-0836/81/340667-02$01.00

©1981 Macmillan Journals Ltd