INTERNATIONAL PEACE FOUNDATION

(Feb. 5 -12, 2004) Bangkok

THE DISCOVERY OF SPLIT GENES

Prof. Richard J. Roberts

In 1976 life was very simple. Many molecular biologists thought that the basic secrets of life were known and that all that remained for molecular biologists was to mop up the details. In 1977 all of that was to change, when Richard Gelinas, Louise Chow, Tom Broker and myself made the startling discovery that genes in higher organisms had a structure that was fundamentally different from that found in bacteria. In this lecture I will briefly recount the background to the discovery, explain how the discovery was made and speak a little about the further ramifications that have since followed.

By 1976 we had a very good idea of what a typical bacterial gene looked like and had glimpsed how it was controlled. It was known that the information encoded in the gene was read by first transcribing it into a messenger RNA (mRNA), which was then passed to a special machine in the cell called the ribosome. This would then enable the translation into protein of the message encoded in the RNA. This decoding of the mRNA required that the individual bases in the RNA were read three at a time, with each triplet coding for a single amino acid that was to be added into a growing protein chain. The complete segment of DNA that was transcribed into mRNA is called a gene and the sequences at its 5'-end contained signals to the ribosome, telling it where to begin translation and signals at the 3'-end of the mRNA that told the ribosome where to stop translation. The gene itself was surrounded by sequences in the DNA that told an enzyme called RNA polymerase where to being transcribing the RNA and also signals telling it where to stop. The signal telling it where to start making RNA was called a promoter and it was this signal that we were especially interested in.

It was known in bacteria that the promoter sequences were fairly short and contained very specific well-conserved sequences. Richard Gelinas and I wanted to know if the promoter sequences in higher organisms were the same as the ones in bacteria. We set out to explore this question by using a virus that grows in human cells called Adenovirus-2. We began to work out methods whereby first we could map the exact start of the mRNA sequences made from Ad-2 mRNA. We knew that there should be some 20 different mRNAs made by Ad-2 and we focused our attention on the late mRNAs that were made in great abundance. The idea was that if we could work out the sequence of one of these mRNAs right at its 5' end, then we would merely need to locate the corresponding DNA sequence, work out the DNA sequences that preceded it, and we would have the promoter. To do this we developed a new technique that would enable us to catch short sequences from the extreme 5' end of all of the Ad-2 mRNAs, anticipating that there would be some 15-20 different such sequences corresponding to each of the individual mRNAs that we knew should be there. Using techniques developed in Fred Sangers' lab, we planned to separate these different 5'-ends in two dimensions and then to choose the most abundant for our studies.

Imagine our surprise when we found not 20 different 5' ends, but only one! At first we thought we had made a mistake, but much repetition and many different approaches convinced us that in fact, all Ad-2 mRNAs, no matter where they were made, all had the same 5'-end. This meant that for most mRNAs, the left-hand end and the main part of the mRNA were encoded a long way apart and were not adjacent to each other as was known to be the case for bacteria. This meant that the mechanism for making mRNAs in Ad-2 must be fundamentally different from the mechanism used in bacteria. We tried many biochemical experiments to try to prove this and to uncover the nature of the mechanism. However, it was not until March, 1977, that the key experiment occurred to me one Saturday morning. It involved electron microscopy and we consulted with our colleagues Louis Chow and Tom Broker, both excellent electron microscopists to see if they could perform the experiment in collaboration with us. They agreed. Richard Gelinas made the

appropriate substrates and on Tuesday, the first experiment was done. Louise Chow reported that sure enough, the genes in Ad-2 were split into pieces as we had suspected. Split genes had been discovered and fairly soon it became clear that in eukaryotes there was an extra step between DNA and protein that required the splicing of the mRNA following the original transcription event.

This work had led to the discovery of complicated machinery dedicated to RNA splicing and this split nature of genes has considerably complicated their analysis in higher organisms. Because the rules that govern this RNA splicing are still not fully worked out, at the present time we cannot just look at the human DNA sequence that was reported three years ago and know where the genes are without additional biochemical work. We also know that because of RNA splicing, a single gene can often make many different mRNAs. Thus, the one-to-one correspondence between a gene and its protein product that is found in bacteria, does not apply to humans. This is one reason why higher organisms, especially man, hae been able to develop the complicated lifestyle that we now see. In 1977 we learnt that molecular biology was not dead! Since then, there have been many exciting discoveries that have made biology the vibrant science that it is today.