F. Sanger: Pioneer of primary sequences

K. BURTON

Department of Biochemistry, The University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K.

F. Sanger: An Honorary Advisory Editor of Bioscience Reports

A second Nobel Prize award for Chemistry to Dr. Frederick Sanger is fitting recognition to a great biochemical pioneer. After showing us how to sequence proteins for his first Nobel, he has developed powerful and simplified techniques for sequencing first RNA and then DNA.

When Sanger started to work on the structure of insulin in the mid-forties, it was not universally evident that the primary structures of biological macromolecules could be fully determined. Indeed, many biochemists suspected that proteins were so heterogeneous that sequences could never be elucidated.

Sanger’s approach was characteristically direct and practical. He used simple organic reactions of organic chemistry to develop his dinitrobenzene end-group method, a feature being the production of yellow peptides that were easily detected in paper or column chromatography. After the structure of insulin had been determined, Sanger continued to explore better separation methods for peptides and later started to apply them to oligonucleotides. With the purification of individual species of tRNA, and its sequence analysis by non-radioactive methods in the hands of Holley, Zachau, and others, Sanger saw that, as for proteins, easy detection and good resolution on chromatography paper would speed up RNA sequencing and make it more practicable. He used 32P labelling and autoradiography for detection, and a number of separation methods including ionophoresis on cellulose acetate and on ion-exchange papers to develop two-dimensional fingerprinting methods where the position of the
nucleotides often gave a provisional indication of their structures. As in Holley's work, the strategy of sequencing depended upon base-specific cleavage by ribonucleases, but it could still be relatively time-consuming to find the unambiguous overlaps needed to derive the complete sequences. Sanger's procedures were soon used by several groups all over the world to sequence many small RNA molecules. His own group also took up the challenge of larger molecules and determined the sequences of several interesting segments of R17 bacteriophage RNA. Comparison of corresponding nucleotide and amino acid sequences gave direct confirmation of the genetic code and demonstrated that degeneracy in the code could allow coding sequences of RNA to fold into hairpin loops with good Watson-Crick base-pairing.

Sanger now turned his attention to the single-stranded DNA bacteriophage, φX174, which is particularly suited to sequencing methods involving primed synthesis by DNA polymerases. He developed what is, in retrospect, a very simple idea, namely, to run a series of parallel syntheses each from the same initiation site determined by the primer but with preferential termination at different types of DNA nucleotide. Each incubation mix would give products of random length, but one reaction mixture would end preferentially at adenine nucleotides, the second at guanine nucleotides, and so on. The products are examined by a one-dimensional high-resolution ionophoresis system on thin polyacrylamide gels and the sequence can be read directly from observation of the pattern of bands. The first successful procedure, the 'plus and minus' method, was used to obtain the complete provisional sequence of φX174 DNA containing over 5000 nucleotides. An improved method was then developed using chain termination by 2'3'-dideoxynucleotide triphosphates and used to sequence those regions of the provisional sequence that had not been satisfactorily confirmed. This sequence demonstrated many interesting features, including the occurrence of genes which overlap each other in different translational reading frames.

The practical demonstration that such rapid sequencing was possible has assisted the development of other methods that can be applied directly to RNA as well as the well-known Maxam-Gilbert procedure for DNA sequencing. This method uses the same system of thin gels and parallel ladder patterns but is based on random base-specific cleavages of terminally-labelled DNA. The different methods are complementary and have many experimental manipulations in common, so that they can be used easily in the same laboratory, as convenience and their appropriate merits dictate.

The DNA sequencing methods can be applied to very long stretches of DNA, taking advantage of the specific cleavages produced by the Type-II restriction endonucleases and of the associated cloning procedures pioneered by Paul Berg. The new armory of nucleic acid technology has indeed provided yet another revolution in biology, for it is now possible to select required segments of a genome and to sequence them. Almost overnight, it seems, DNA sequencing has been transformed from an esoteric end in itself to an invaluable tool for the next wave of biochemistry.