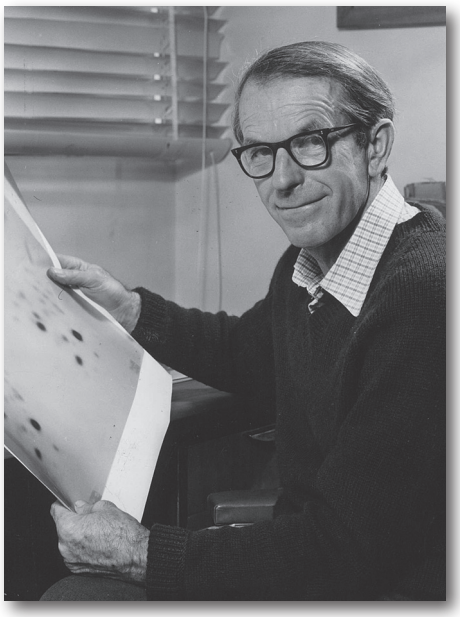
Fred Sanger (1918–2013): a science hero



Fred Sanger in about 1958 (Peter Sanger)

Fred Sanger, who died on 19 November 2013, was one of the most remarkable scientists of our time. He worked at the bench between 1943 and 1983, the first 20 years or so at the Biochemistry Department, University of Cambridge, on amino acids and proteins, resulting in the complete sequencing of insulin in 1955, the first ever protein to be sequenced. It was this achievement that earned him his first (and solo) Nobel Prize for Chemistry in 1958.

Over the next two decades he went on, at the MRC Laboratory for Molecular Biology, to devise methods for sequencing nucleic acids, culminating in 1977 in the development of an ingenious method for sequencing DNA (the 'Sanger dideoxy method'), which opened the way to the complete sequencing of the human genome in 2001. For this he was awarded his second Nobel Prize in Chemistry in 1980 which he shared with Walter Gilbert, "for contributions concerning the determination of base sequences in nucleic acids", the other half going to Paul Berg for his work on recombinant DNA. Fred Sanger was one of only four Nobel laureates ever to be awarded two prizes, and alone in receiving them both for chemistry.

But the outcome of Sanger's life in research was much greater, and, as the late Guy Dodson so aptly put it, "…it was an unequalled foundation for the most prodigious explosion that is modern biology". His unerring chemical judgement, skills as an experimenter and tenacity are legendary, as was his modesty. His many honours included, in addition to the Nobels, an FRS (1954), Foreign Associate of the National Academy of Science, USA (1967), CBE (1963), CH (1981) and OM (1986). He had apparently declined a knighthood because, he said, he didn't like the idea of being called 'Sir'…

His most public legacy is no doubt the Sanger Centre, later the Wellcome Trust Sanger Institute, at Hinxton Hall, Cambridge, where a significant part of the human genome project was carried out. John Sulston, its first Director, vividly recalls that when he asked him over the phone whether it could be called the Sanger Centre, Sanger replied "Well, okay, but it better be good!"

Frederick (Fred) Sanger was born in the Gloucester village of Rendcombe in 1918 where his father, a practising medical doctor who had done research in Cambridge, greatly stimulated his interest in science as well as inculcating in him, through a Quaker upbringing, the principles of pacifism and the value of truth. Although he abandoned the religious aspects of Quakerism by the time he was 19, he remained a conscientious objector and committed supporter of anti-nuclear campaigns throughout his life. A major influence on his choice of career was his chemistry teacher at Bryanston, who encouraged Sanger to 'play around' in the chemistry laboratory where he enjoyed experimenting with dyes and, especially, the growing of beautifully coloured crystals.

Sanger first came across 'biochemistry' as an undergraduate at Cambridge, where he chose it as a final fourth-year option, inspired by talking with Ernest Baldwin, his supervisor at St. John's College. He was apparently quite surprised at obtaining a first class degree in 1940 (qualifying him for postgraduate research) as he 'had not done too well in exams'. Being a Quaker and registered as a conscientious objector, he was exempted from military service and applied to do research with N.W. ('Bill') Pirie, the proteins expert, on obtaining edible protein from grass. By the time the bucketful of ground-up grass Pirie kept in the deep-freeze had thawed, Pirie had left Cambridge for Rothamsted.

Fortunately, Albert Neuberger, who had just arrived in Cambridge as a postdoc, took Sanger on as his PhD student. In 1943, he completed his thesis on 'Lysine metabolism and the more practical problem concerning the nitrogen of potatoes'. As he once put it to me, "he [Neuberger] was the one who really taught me how to do research, both by example and instruction", claiming that he only did what Neuberger told him to do "as it took quite a time – about two years in my case – before you start to think out your own experiments". He also felt that the valuable experience he gained on amino acid chemistry during his doctoral work on lysine was a good introduction to his future work on 'sequences'.

From lysine metabolism to the first complete sequence of a protein

After completing his PhD, Sanger joined A.C. Chibnall's research group before becoming a member of the External Scientific Staff of the Medical Research Council (1951). Chibnall had just succeeded F.G. Hopkins as Professor and Head of Biochemistry at Cambridge in 1943. His group's main expertise was quantitative analysis of the amino acid composition of proteins, especially insulin, one of the few proteins readily available at the time in pure form. According to Sanger, it was Maurice Rees, Chibnall's chief assistant, who discovered that insulin had more free amino acids than could be accounted for by its lysine content, leading Chibnall to suggest that Sanger look into this: the rest is history! Sanger first went on to invent fluorodinitrobenzene (later known as the Sanger reagent) as a new coloured reagent to label the N-terminal groups of insulin and, in 1945, after two years of painstaking work he was able to identify of glycine and phenylalanine as the N-terminal amino acid residues of what were later identified as the two cross-linked chains of insulin. There followed the crucial and serendipitous discovery

(reported in 1949 in the *Biochemical Journal*[1]) that mild (as opposed to complete) hydrolysis of 2,4-dinitrophenol (DNP)-labelled insulin produced short sequences of four or five amino acids near the N-termini that could be separated by paper electrophoresis, producing a distinct pattern which Sanger called a "fingerprint". (Sanger frequently acknowledged his debt to Archer Martin and Richard Synge for introducing the methods of partition chromatography that he was able to exploit so successfully for this work.) Sanger considered his 1949 'peptide paper' as the most significant of his early papers on insulin as it demonstrated for the first time that sequences could, in principle, be located in a protein molecule. That was quite a revolutionary concept at the time, when most biochemists believed the arrangement of different amino acids in a protein to be random. Importantly, partial hydrolysis became a central general approach to the sequencing of proteins, and later of nucleic acids.

By 1955, through work with two of his postdocs, Hans Tuppy and later Ted Thompson, Sanger had succeeded in revealing, as one writer put it, "the linear world of polypeptides", at the same time as Benzer was revealing the linear world of genes. The fact that there was a well-defined sequence in proteins led directly to the idea that there had to be a genetic code*. In addition, in one of several papers published in the *Biochemical Journal* in 1955 completing the work on insulin, Sanger and his colleagues discovered discrete species-specific variations in amino acid sequence within the 'glycyl chain' of insulin. This provided the first evidence for significant similarities between proteins that could be assumed to be related evolutionarily (i.e. homologous), "an insight that led directly to the development of sequence alignment methods, still the most commonly used bioinformatics tools". For this, he has been credited as the "grandfather" of bioinformatics[2].

The second phase: sequencing nucleic acids

Sanger carried out much of his research on proteins in an overcrowded basement laboratory within the Biochemistry Department. Through his enthusiastic encouragement, he was instrumental in getting Max Perutz to persuade the Medical Research Council (MRC) to fund a new building to accommodate both their research groups. And in about 1962, he began the second phase of his career in the newly created MRC Laboratory of Molecular Biology. Here he led the Protein and Nucleic Acid Chemistry Division (initially called the Protein Chemistry Division) until 1983, when his good friend and colleague Cesar Milstein took over.

Although Sanger had previously felt little interest in nucleic acids, after moving to the LMB, he conceded that "with people like Francis Crick around, it was difficult to ignore nucleic acids or to fail to recognize the importance of sequencing them"[3]. An even more important influence was apparently John Smith, a very co-operative colleague who taught him how to fractionate nucleotides.

George Brownlee, who joined him as his first PhD student in 1963, somewhat surprised Sanger by choosing to work on RNA, rather than on proteins – "a bold choice at the time!"†. Sanger considered this a strong factor in getting his whole group to move in that general direction.

Initial investigations on nucleic acids were with defined RNA structures, e.g. tRNA and rRNA, and explored the use of using RNA-cleaving enzymes to prepare partial digests. This led to the discovery that radiolabelled di-, tri- and tetra-nucleotides could be separated by 2D ionophoretic techniques. The first real breakthrough, according to Sanger, came when Brownlee succeeded in deciphering the 120-nucleotide sequence of 5S rRNA from *Escherichia coli*, using novel 2D separation techniques they had developed that allowed 'reading off' sequences directly from their position on autoradiographs. Further refinements of these ingenious techniques led to Sanger and his colleagues to sequence much longer RNA, e.g. the 3300-nucleotide-long bacteriophage R17RNA. As some of the sequences matched the codons predicted from the amino acid sequence of the coat protein (already known), it enabled them for the first time to directly determine (and thus confirm) the genetic code from a nucleotide sequence.

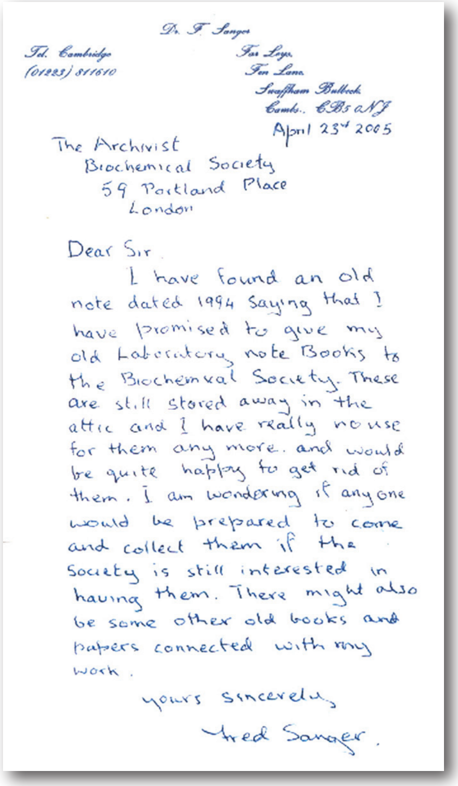
The ultimate challenge: DNA

Then – and apparently to the dismay of some people, including Dorothy Hodgkin who felt he should have returned to proteins – Sanger committed to elucidating the sequence of DNA, that of the bacteriophage ΦX174 to start with. This must surely be counted as one of the boldest leaps forward ever taken in biology. An American reviewer consulted about Sanger's ambitious new programme said, "Fred Sanger continues to bang his head against brick walls. And the brick walls keep falling down."

Sanger initially developed copying methods with DNA polymerase and short nucleotide primers to successfully derive a 50-nucleotide sequence of the single-stranded f1 bacteriophage DNA. He and Alan Coulsen then went on to develop a fast read-out method, the so-called 'plus and minus method' and, shortly thereafter, the game-changing sequencing methodology, the 'dideoxy chain termination method', used to confirm their earlier sequence for the ΦX174 DNA (some 5375 nucleotides long). This was the first whole genome sequence ever produced and Sanger's group went on to sequence the 16 569 bases of the human double-stranded mtDNA (1981): the first extensive human DNA to be sequenced. Sanger then developed the so-called 'whole-genome shotgun method' he

* One of the first persons who made that connection early on was the young Cambridge chemist, John Griffith, who Sanger once told me he considered "a bit of a genius".

† For text references to Sanger and references to key papers, see Guy Dodson's 'BJ Classic Papers' article (doi:10.1042/BJ2005c013) and my 2005 interview at www.biochemistry.org/AboutUs/Archive/Archivecollection/FredSangersequencingpioneer.aspx


Dr. F. Sanger
Tel. Cambridge (01223) 811610
Far Leys, Fen Lane, Swaffham Bulbeck, Cambs., CB5 0NJ
April 23rd 2005

The Archivist
Biochemical Society
59 Portland Place
London

Dear Sir

I have found an old note dated 1994 saying that I have promised to give my old Laboratory note Books to the Biochemical Society. These are still stored away in the attic and I have really no use for them any more. and would be quite happy to get rid of them. I am wondering if anyone would be prepared to come and collect them if the Society is still interested in having them. There might also be some other old books and papers connected with my work.

yours sincerely,

Fred Sanger.

Fred Sanger's letter to the Biochemical Society archivist.

and Coulson used to sequence the genome of the virus bacteriophage λ (48 502 nucleotides) in 1982.

Fred Sanger, a modest genius

Although often referred to as the father of genomics, Sanger always considered his own contribution to nucleic acid sequencing was primarily to the methods. He wrote that he "quite enjoyed the somewhat routine DNA sequencing work….[as] a form of relaxation compared with the more original, exacting, variable, and often frustrating, work on development of methods"[4]. Whatever he felt were his own strengths, Sanger never failed to fully acknowledge the contributions of all of the people he interacted with, at whatever level.

Sanger, once referred to as the most focused of men, was passionate about research and about the intense enjoyment he got from "messing about in the lab". As he once put it, "of the three main activities involved in scientific research, thinking, talking and doing, I much prefer the last and am probably best at it"[3]. John Sulston once remarked in an interview that Sanger "…was a quiet guy, he just got on with it – that was his style …I found him completely inspirational".

He was also most unusual in deciding to give up research completely after retiring from the LMB in 1983, at the statutory age of 65, as he felt his life's work had been completed with the sequencing of DNA. The ending of his life in the laboratory was certainly abrupt and decisive, as related by John Finch in his history of the LMB: "The conventional finishing date in the MRC is 30 September, and on that day Fred was doing experiments – on 1 Oct his laboratory was empty and he was gone"[4].

With characteristic modesty, Sanger later commented that he would have felt guilty at occupying space that could have been available to a younger person and that he looked forward in retirement to spending time enjoying his family, 'messing about with boats' and tending the lovely garden at his home in the Cambridgeshire village of Swaffham Bulbeck.

I had the opportunity of visiting Fred after receiving his letter in 2005 inviting me to collect all (35) of his, mostly hand-written laboratory notebooks which he had decided many years earlier to donate to the Biochemical Society for its archives[‡]. I wish someone had photographed both of us scaling the ladder to the loft and gingerly descending with arms laden with the notebooks!

I later had the privilege of interviewing him at his home, wondering once again how this amazingly self-effacing, quiet and civilized person was indeed the same, some would say visionary, Fred Sanger whose impact on the course of biology has been likened to Charles Darwin's. Contemporary science historians are quite rightly reluctant to speak of heroes, but I hope they might make an exception with Sanger.

Fred Sanger died at Addenbrooke's Hospital on 19 November 2013. He married Joan Howe (who died in 2012) in 1940, and is survived by his three children and two grandchildren. ■

I thank Professor George Brownlee for allowing me to 'mine' the obituary he published in Current Biology **23**, *R1074–R1076*

John Lagnado
(Honorary Archivist)

References

1. Sanger, F. (1949) Biochem. J. **45**, 563–574
2. Samson, C. (2006) The Biochemist, **28** (6), 48–49
3. Sanger, F. (1988) Annu. Rev. Biochem. **57**, 1–29
4. Finch, J. (2008), A Nobel Fellow on Every Floor: a History of the Medical Research Council Laboratory of Molecular Biology, Medical Research Council, Cambridge

[‡] The Biochemical Society archives, including the notebooks and several audio- and video-tape interviews of Sanger, were catalogued in 2009 and are kept on permanent loan at the Wellcome Library. It is particularly pleasing that the Wellcome has included Sanger's notebooks (and related correspondence) as part of their ambitious digitizing project, making their content readily accessible online.