Dorothy Hodgkin, protein crystallography and insulin

G. G. Dodson

Department of Chemistry, University of York, Heslington, UK

Dorothy Hodgkin, or Crowfoot as she then was, began her research into insulin's crystal structure in 1935, soon after moving to Oxford. The crystals were rhombohedral and easily grown following Abel's recipe published a little earlier. Insulin, being a protein and a hormone, was a natural study for her to begin following the experiments that she and J. D. Bernal had carried out on pepsin crystals and on the analysis of steroids. In these pepsin experiments they had established the techniques for obtaining X-ray diffraction patterns from protein crystals. These relied on placing the wet crystal in a sealed glass capillary with a reservoir of mother liquor. Under these conditions the protein crystal lattice retains its waters of crystallization which typically constitute 50% of the cell volume. This technique is still almost universally used today, and only the recent introduction of the cryoscopic method is likely to replace it in any significant way. With this strikingly simple experimental trick, Dorothy in Oxford was able to carry out preliminary X-ray investigations on a number of proteins and characterized crystals of β lactoglobulin, tobacco seed globulin and of course insulin. Most remarkably she studied crystals of tobacco necrosis virus reporting the results in *Nature* after the war.

The insulin studies were Dorothy's major commitment in protein structural research. Once the problems of mounting the crystal and recording their diffraction patterns on film were resolved, the challenge of determining the structure of protein crystals presented itself. At this time, the 1940s, a few small molecule crystal structures had been solved, some of these by Dorothy herself. The problems faced were truly formidable—not even the exact molecular weight of the hormone was known. The evidence from Svedburg was of a 36,000 molecular weight species which could dissociate to a species 12,000 molecular weight. This observation corresponded nicely to the constitution of the rhombohedral crystals whose unit cell contained the 36,000 Dalton species, and the asymmetric unit the 12,000 Dalton species. Not daunted by the complexities Dorothy went ahead and calculated the Patterson function almost as soon as it was published in 1935 for both the wet crystals and the partially dried crystal, for which data extended to 6 Å spacing. The attraction of this calculation was that it required only experimental observations not phases. The disadvantage was that the function generated the vector peak between each pair of atoms in the cell. For the insulin crystal the complexity of the unit cell made structural interpretation of the resultant maps inconceivable. But the Patterson maps of the wet and dry crystal revealed that the two series exhibited marked similarities, and that their main features were preserved. There were, however, relative movements of these features. These experiments suggested that loss of water and the accompanying loss of crystal quality was associated with rearrangements in the crystal lattice but that this phenomenon did not alter the protein structure significantly. These were important experiments in understanding the properties of proteins; they required incidentally considerable skill and a bit of luck. Not all air dried protein crystals diffract, and it was fortunate that the rhombohedral crystals survived this treatment well enough to yield 6 Å data.

With purified insulin and well-defined recipes from Jorgan Schlichtkrull for a second rhombohedral insulin crystal, which also contained zinc, Dorothy returned to concentrate on the crystal structure in the mid 1950s. This step was stimulated by the determination of the chemical sequence of insulin by Fred Sanger and his colleagues, and by the success of John Kendrew and Max Perutz in preparing heavy atom derivatives in myoglobin and haemoglobin. A research group which over the years got quite large and very international, grew up around the insulin crystallographic research.

There were numerous difficulties presented by these rhombohedral crystals, the most serious being their poor ability to react with heavy atoms and the absence of centric reflections which were very convenient for analysing heavy atoms at that time. Their absence meant the heavy atom scattering in the derivatives was not directly measurable from the isomorphous differences, limiting the Patterson functions and the refinement of heavy atom atomic parameters. These difficulties encouraged other approaches. Probably the most important of these were the Patterson based rotation and translation functions being developed in Cambridge by Michael Rossmann and David Blow. They applied these methods to the two different but related rhombohedral insulin crystals, both of which contained two independent molecules in their asymmetric units.

The solution to the insulin structural analysis problems was to exploit anomalous scattering effects. Their measurement made it possible to construct the heavy atom scattering for all reflections, albeit inaccurately. (The

mathematical formulations for the expression of the heavy atom scattering were first devised by Luke Hodgkin, Dorothy's eldest son.) They were applied first with partial success to the insulin data by Dorothy and Marjorie Aitken, later Harding, and Margaret Adams. The visit of Sivaraj Ramaseshan in 1964 led to his developing more powerful ways of utilizing anomalous effects in neutrons and most significantly, synchrotron radiation, which has become a major tool in protein crystallography. It was only when the anomalous scattering differences were measured accurately on the Hilger-Watts 4-circle diffractometer by Vijayan, Tom Blundell and myself, however, that the anomalous differences could be exploited effectively in calculations carried out largely by Eleanor Dodson. These analyses revealed that the pattern of heavy atom substitution was surprisingly complex and they needed both accurate and relatively high resolution data (better than 3 Å spacing) to be resolved.

Throughout the period 1955–1969 when the development of the techniques and technologies seemed so slow, Dorothy remained confident that the insulin structure could be solved. She provided a continuous flow of ideas and suggestions. When the structures of myoglobin and lysozyme were determined, for example, she immediately went to see for herself (taking the laboratory with her) what the molecules looked like and what might be learned from heavy atom experiments.

The structure of the 2 Zn insulin crystal was determined in 1969, the culmination of 34 years of research by Dorothy and many individuals. The interpretation of the electron density map was completed in an intensely busy and absorbing weekend. There were many moments of excitement – one was the confirmation of the disulphide bond arrangements defined in Sanger's experiments. The zinc coordination was readily established; and the hormone was seen to be organized in the hexamer as three well defined dimers, held together by H-bonds in β pleated sheet arrangements.

The three-dimensional structure of the two crystallographically independent insulin monomers in the dimer was very similar, though some significant differences exist. The molecule's architecture and the spatial arrangements of the sidechains deduced from the electron density largely explained the hormone's chemical properties (for which there was a huge literature). It was also possible now to begin the task of relating insulin's three-dimensional structure to its biological history and to its action on the cell. In spite of the complex nature of the evidence, progress was made and a region identified as a possible binding surface to the insulin receptor. The widespread use of insulin in the treatment of diabetes added a medical dimension to the structure which was immediately applied to understanding better the design of therapeutic preparations. And of course the possibility of identifying the molecule's active surface raised the prospect, very distant, of synthesizing drug molecules which could act properly on the insulin receptor and remove the burden of injections.

Studies on the other rhombohedral crystals which were grown in concentrated chloride were continued, reflecting curiosity about the origin of the chemical and the structural differences in the two closely related cells. To Dan Mercola's and Graham Bentley's and everyone's surprise the insulin hexamer in this second rhombohedral form contained marked structural differences – including a new zinc coordination site. It was not obvious then – and it is still not obvious why this structural rearrangement should be generated by chloride ions. The large extent of the conformational change demonstrated the striking flexibility of the molecule and was at the time the most dramatic example of the intrinsic capacity for structural alterations in proteins.

The hope that the knowledge of the three-dimensional structure would give clear indications of the insulin's biologically active surface were only partially realized. From the very beginning we and many in other laboratories thought about how the binding surface of the hormone might be identified. The evidence was often suggestive but never conclusive, and some of our ideas were wrong. The essential framework for the developments in thinking was, however, clearly defined in the early structural papers. It turns out, however, from recent research at York and elsewhere that the main problem was failing to appreciate fully enough that the molecule can undergo extensive conformational changes, and that some of these undoubtedly occur at the receptor, and in doing this reveals previously buried surfaces.

In Dorothy's research there was a continuing interest in crystallography and its application to protein crystals. During the heavy atom experiments she actively encouraged Michael Rossmann and David Blow to help with their rotation and translation calculation through which clues to the organization of the molecules in the cell could be obtained. Although no phasing information was obtained from these calculations they helped to demonstrate the value of the methods which are now very widely used. Dorothy, building on her experience with vitamin B_{12}, pursued the exploitation of phase determining methods by anomalous scattering. Her contacts with David Sayre and her interest in solving structures led to their studies with Neil Isaacs on applying direct methods on insulin crystals, and particularly in the later years gave great attention to refinement methods.

The most consistent feature of the insulin research after the solution of the crystal structure was the steady extension of the resolution from 2.8 Å to 1.5 Å spacing. Dorothy always intended to complete the 2 Zn insulin structure as a properly refined molecule. This took some time. It involved developing new methods and endless

hours of examining maps. In the end the 2 Zn insulin crystal structure was completed, with all protein atoms identified and a comprehensive and detailed description of the water structure. A good deal of this methodological research was never published in detail, which is a pity, but a reflection of the pressures successful protein crystallographic groups often experience.

Dorothy's influence in insulin research was important. It was of great significance in China where an insulin crystallographic programme was initiated in the mid 1960s following the successful accomplishment of the hormone's chemical synthesis. The research led by Liang Dong-cai, was greatly encouraged by Dorothy, and she had great pleasure comparing the insulin structures determined in Oxford and Beijing in 1971. The subsequent growth of protein crystallography in China owes a great deal to the success of their insulin study and Dorothy's support for it. Other examples of protein crystallographic research being started from the Oxford research are the laboratories of Ted Baker in Palmerston North (New Zealand), Guy and Eleanor Dodson (in York), Tom Blundell (in London) and Vijayan (in India) all of whom worked on insulin with Dorothy at Oxford.

After her retirement in 1977 Dorothy continued to work on insulin, giving lectures and in her research concentrating on the refinement of the 2 Zn insulin crystal with the 1.5 Å spacing data. The approach was strictly crystallographic and the protein atoms and the solvent molecules were analysed and refined as rigorously as possible. This study keeps a standard for protein structure analysis and its completion 50 years after the initial experiments illustrates the persistence Dorothy showed in her research and the immense achievement it represents.

Chemical crystallography – Past, present and future

Jack D. Dunitz

Organic Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zurich, Switzerland

We must not forget that the name of the Oxford laboratory was the Chemical Crystallography Laboratory. I think that may have been one of the reasons I wanted to go there. What I knew about chemistry was based mainly on Pauling's book *The Nature of the Chemical Bond*, and what I knew about crystallography came from the first volume of *The Crystalline State* written by Bragg. A Chemical Crystallography Laboratory seemed like the sort of place where I could combine my interest in these two fields, which was unbounded, with my scanty knowledge.

At that time, remember, fifty years ago, chemists were still arguing and even sometimes quarreling about the molecular structures of simple natural products. A would show that the structure of some compound deduced in B's laboratory was incorrect, and a year later B would turn the tables on A by showing that the revised structure was also wrong. In contrast, the structures determined by X-ray crystallography had a satisfying impression of definiteness about them. Molecules were revealed to correspond to objects of definite size and shape, not just intellectual constructs designed to explain chemical reactivity. There was a price to be paid, of course. In the course of determining the structure of a compound by crystallography, no new chemistry was done, nothing new was learned about the chemical reactivity of the compound in question. It seemed to me then, naive as I was, that as one learned more about the systematics of molecules, it ought to be possible to fill this gap, to deduce chemical properties from molecular structure.

One problem was that, in those days, it took a lot of time and work to determine the structure of even a single relatively simple organic molecule. Also some luck. We knew then practically every organic structure that had ever been determined; who had done it, how it had been solved – there was not much choice, trial and error, or some variant of the heavy-atom method. Two-dimensional Fourier projections, using Beevers–Lipson strips or Robertson boards, were about the limit of computational practicability. Even if one went to the trouble of collecting three-dimensional diffraction data, it was a daunting task to do the necessary calculations. This is why most crystal structures determined in those days had a short cell dimension, short enough so that there was not too much atomic superposition in the corresponding two-dimensional projection.

Chemical crystallography? Ionic structures were reasonably well understood; Pauling's Rules were enough to rationalize them and even to predict unknown structures. In the organic chemistry area there was less to boast about, although there were one or two notable achievements. Through J. M. Robertson's work, for example, the bond lengths in aromatic hydrocarbons could be correlated with simple theoretical models – resonance structures. But for the most part, one had to