International Tables for Crystallography (2012). Vol. F, Section 1.2.2, pp. 5–6.

Chapter 1.2. Historical background

M. G. Rossmann

1.2.1. Introduction

Crystallography ranks with astronomy as one of the oldest sciences. Crystals, in the form of precious stones and common minerals, have attractive properties on account of their symmetry and their refractive and reflective properties, which result in the undefinable quality called beauty. Natural philosophers have long pondered the unusual properties seen in the discontinuous surface morphologies of crystals. Hooke (1665) and Huygens (1690) came close to grasping the way repeating objects create discrete crystal faces with reproducible interfacial angles. The symmetry of mineral crystals was explored systematically in the 18th and 19th centuries by measuring the angles between crystal faces, leading to the classification into symmetry systems from triclinic to cubic and the construction of symmetry tables (Schoenflies, 1891; Hilton, 1903; Astbury *et al.*, 1935) – the predecessors of today's *International Tables.*

1.2.2. 1912 to the 1950s

It was not until the interpretation of the first X-ray diffraction experiments by Max von Laue and Peter Ewald in 1912 that it was possible to ascertain the size of the repeating unit in simple crystals. Lawrence Bragg, encouraged by his father, William Bragg, recast the Laue equations into the physically intuitive form now known as 'Bragg's law' (Bragg & Bragg, 1913). This set the stage for a large number of structure determinations of inorganic salts and metals. The discovery of simple structures (Bragg, 1913), such as that of NaCl, led to a good deal of acrimony, for crystals of such salts were shown to consist of a uniform distribution of positive and negative ions, rather than discrete molecules. These early structure determinations were based on trial and error (sometimes guided by the predictions of Pope and Barlow that were based on packing considerations) until a set of atomic positions could be found that satisfied the observed intensity distribution of the X-ray reflections. This gave rise to rather pessimistic estimates that structures with more than about four independent atomic parameters would not be solvable.

The gradual advance in X-ray crystallography required a systematic understanding and tabulation of space groups. Previously, only various aspects of three-dimensional symmetry operations appropriate for periodic lattices had been listed. Consequently, in 1935, the growing crystallographic community put together the first set of *Internationale Tabellen* (Hermann, 1935), containing diagrams and information on about 230 space groups. After World War II, these tables were enlarged and combined with Kathleen Lonsdale's structure-factor formulae (Lonsdale, 1936) in the form of *International Tables* Volume I (Henry & Lonsdale, 1952). Most recently, they have again been revised and extended in Volume A (Hahn, 2005).

Simple organic compounds started to be examined in the 1920s. Perhaps foremost among these is the structure of hexamethylbenzene by Kathleen Lonsdale (Lonsdale, 1928). She showed that, as had been expected, benzene had a planar hexagonal structure. Another notable achievement of crystallography was made by J. D. Bernal in the early 1930s. He was able to differentiate between a number of possible structures for steroids by studying their packing arrangements in different unit cells (Bernal, 1933). Bernal ('Sage') had an enormous impact on English crystallographers in the 1930s. His character was immortalized by the novelist C. P. Snow in his book *The Search* (Snow, 1934). By the mid-1930s, J. Monteath Robertson and I. Woodward had determined the structure of nickel phthalocyanine (Robertson, 1935) using the heavy-atom method. This was a major crystallographic success and perhaps the first time that a crystallographer had succeeded in solving a structure when little chemical information was available.

Another event which had a major impact was the determination of the absolute hand of the asymmetric carbon atom of sodium tartrate by Bijvoet (Bijvoet, 1949; Bijvoet *et al.*, 1951). By indexing the X-ray reflections with a right-handed system, he showed that the breakdown of Friedel's law in the presence of an anomalous scatterer was consistent with the asymmetric carbon atom having a hand in agreement with Fischer's convention. With that knowledge, together with the prior results of organic reaction analyses, the absolute hand of other asymmetric carbon atoms could be established. In particular, the absolute structure of naturally occurring amino acids and riboses was now determined.

Until the mid-1950s, most structure determinations were made using only projection data. This not only reduced the tremendous effort required for manual indexing and for making visual estimates of intensity measurements, but also reduced the calculation effort to almost manageable proportions in the absence of computing machines. However, the structure determination of penicillin (Crowfoot, 1948; Crowfoot *et al.*, 1949), carried out during World War II by Dorothy Hodgkin and Charles Bunn, employed some three-dimensional data. A further major achievement was the solution of the three-dimensional structure of vitamin B_{12} by Dorothy Hodgkin and her colleagues (Hodgkin *et al.*, 1957) in the 1950s. They first used a cobalt atom as a heavy atom on a vitamin B_{12} fragment and were able to recognize the 'corrin' ring structure. The remainder of the B_{12} structure was determined by an extraordinary collaboration between Dorothy Hodgkin in Oxford and Kenneth Trueblood at UCLA in Los Angeles. While Dorothy's group did the data collection and interpretation, Ken's group performed the computing on the very early electronic Standard Western Automatic Computer (SWAC). Additional help was made available by the parallel work of J. G. White at Princeton University in New Jersey. This was at a time before the internet, before e-mail, before usable transatlantic telephones and before jet travel. Transatlantic, propeller-driven air connections had started to operate only a few years earlier.

Many technical advances were made in the 1930s that contributed to the rapidly increasing achievements of crystallography. W. H. Bragg had earlier suggested (Bragg, 1915) the use of Fourier methods to analyse the periodic electron-density distribution in crystals, and this was utilized by his son, W. L. Bragg (Bragg, 1929*a*,*b*). The relationship between a Fourier synthesis and a Fourier analysis demonstrated that the central problem in structural crystallography was in the phase. Compu-

tational devices to help plot this distribution were invented by Arnold Beevers and Henry Lipson in the form of their 'Beevers–Lipson strips' (Beevers & Lipson, 1934) and by J. Monteath Robertson with his 'Robertson sorting board' (Robertson, 1936). These devices were later supplemented by the XRAC electronic analogue machine of Ray Pepinsky (Pepinsky, 1947) and mechanical analogue machines (McLachlan & Champaygne, 1946; Lipson & Cochran, 1953) until electronic digital computers came into use during the mid-1950s.

A. Lindo Patterson, inspired by his visit to England in the 1930s where he met Lawrence Bragg, Kathleen Lonsdale and J. Monteath Robertson, showed how to use F^2 Fourier syntheses for structure determinations (Patterson, 1934, 1935). When the 'Patterson' synthesis was combined with the heavy-atom method, and (later) with electronic computers, it transformed analytical organic chemistry. No longer was it necessary for teams of chemists to labour for decades on the structure determination of natural products. Instead, a single crystallographer could solve such a structure in a period of months.

Improvements in data-collection devices have also had a major impact. Until the mid-1950s, the most common method of measuring intensities was by visual comparison of reflection 'spots' on films with a standard scale. However, the use of counters (used, for instance, by Bragg in 1912) was gradually automated and became the preferred technique in the 1960s. In addition, semi-automatic methods of measuring the optical densities along reciprocal lines on precession photographs were used extensively for early protein-structure determinations in the 1950s and 1960s.

1.2.3. The first investigations of biological macromolecules

Leeds, in the county of Yorkshire, was one of the centres of England's textile industry and home to a small research institute established to investigate the properties of natural fibres. W. T. Astbury became a member of this institute after learning about X-ray diffraction from single crystals in Bragg's laboratory. He investigated the diffraction of X-rays by wool, silk, keratin and other natural fibrous proteins. He showed that the resultant patterns could be roughly classified into two classes, α and β, and that on stretching some, for example, wool, the pattern is converted from α to β (Astbury, 1933).

Purification techniques for globular proteins were also being developed in the 1920s and 1930s, permitting J. B. Sumner at Cornell University to crystallize the first enzyme, namely urease, in 1926. Not much later, in Cambridge, J. D. Bernal and his student, Dorothy Crowfoot (Hodgkin), investigated crystals of pepsin. The resultant 1934 paper in *Nature* (Bernal & Crowfoot, 1934) is quite remarkable because of its speed of publication and because of the authors' extraordinary insight. The crystals of pepsin were found to deteriorate quickly in air when taken out of their crystallization solution and, therefore, had to be contained in a sealed capillary tube for all X-ray experiments. This form of protein-crystal mounting remained in vogue until the 1990s when crystal-freezing techniques were introduced. But, most importantly, it was recognized that the pepsin diffraction pattern implied that the protein molecules have a unique structure and that these crystals would be a vehicle for the determination of that structure to atomic resolution. This understanding of protein structure occurred at a time when proteins were widely thought to form heterogeneous micelles, a concept which persisted another 20 years until Sanger was able to determine the unique amino-acid sequences of the two chains in an insulin molecule (Sanger & Tuppy, 1951; Sanger & Thompson, 1953*a*,*b*).

Soon after Bernal and Hodgkin photographed an X-ray diffraction pattern of pepsin, Max Perutz started his historic investigation of haemoglobin.[1] Such investigations were, however, thought to be without hope of any success by most of the contemporary crystallographers, who avoided crystals that did not have a short (less than 4.5 Å) axis for projecting resolved atoms. Nevertheless, Perutz computed Patterson functions that suggested haemoglobin contained parallel α-keratin-like bundles of rods (Boyes-Watson *et al.*, 1947; Perutz, 1949). Perutz was correct about the α-keratin-like rods, but not about these being parallel.

In Pasadena, Pauling (Pauling & Corey, 1951; Pauling *et al.*, 1951) was building helical polypeptide models to explain Astbury's α patterns and perhaps to understand the helical structures in globular proteins, such as haemoglobin. Pauling, using his knowledge of the structure of amino acids and peptide bonds, was forced to the conclusion that there need not be an integral number of amino-acid residues per helical turn. He therefore suggested that the 'α-helix', with 3.6 residues per turn, would roughly explain Astbury's α pattern and that his proposed 'β-sheet' structure should be related to Astbury's β pattern. Perutz saw that an α-helical structure should give rise to a strong 1.5 Å-spacing reflection as a consequence of the rise per residue in an α-helix (Perutz, 1951*a*,*b*). Demonstration of this reflection in horse hair, then in fibres of polybenzyl-L-glutamate, in muscle (with Hugh Huxley) and finally in haemoglobin crystals showed that Pauling's proposed α-helix really existed in haemoglobin and presumably also in other globular proteins. Confirmation of helix-like structures came with the observation of cylindrical rods in the 6 Å-resolution structure of myoglobin in 1957 (Kendrew *et al.*, 1958) and eventually at atomic resolution with the 2 Å myoglobin structure in 1959 (Kendrew *et al.*, 1960). The first atomic resolution confirmation of Pauling's β structure did not come until 1966 with the structure determination of hen egg-white lysozyme (Blake, Mair *et al.*, 1967).

Although the stimulus for the Cochran *et al.* (1952) analysis of diffraction from helical structures came from Perutz's studies of helices in polybenzyl-L-glutamate and their presence in haemoglobin, the impact on the structure determination of nucleic acids was even more significant. The events leading to the discovery of the double-helical structure of DNA have been well chronicled (Watson, 1968; Olby, 1974; Judson, 1979). The resultant science, often known exclusively as molecular biology, has created a whole new industry. Furthermore, the molecular-modelling techniques used by Pauling in predicting the structure of α-helices and β-sheets and by Crick and Watson in determining the structure of DNA had a major effect on more traditional crystallography and the structure determinations of fibrous proteins, nucleic acids and polysaccharides.

Another major early result of profound biological significance was the demonstration by Bernal and Fankuchen in the 1930s (Bernal & Fankuchen, 1941) that tobacco mosaic virus (TMV) had a rod-like structure. This was the first occasion where it was possible to obtain a definite idea of the architecture of a virus.

[1] Perutz writes, 'I started X-ray work on haemoglobin in October 1937 and Bragg became Cavendish Professor in October 1938. Bernal was my PhD supervisor in 1937, but he had nothing to do with my choice of haemoglobin. I began this work at the suggestion of Haurowitz, the husband of my cousin Gina Perutz, who was then in Prague. The first paper on X-ray diffraction from haemoglobin (and chymotrypsin) was Bernal, Fankuchen & Perutz (Bernal *et al.*, 1938). I did the experimental work, (and) Bernal showed me how to interpret the X-ray pictures.'