counterparts in modern 'protein crystallography', although few today recognize their real origin.

The first of these methods was the use of 'shrinkage' stages (Perutz, 1946; Bragg & Perutz, 1952). It had been noted by Bernal and Crowfoot (Hodgkin) in their study of pepsin that crystals of proteins deteriorate on exposure to air. Perutz examined crystals of horse haemoglobin after they were air-dried for short periods of time and then sealed in capillaries. He found that there were at least seven consecutive discrete shrinkage stages of the unit cell. He realized that each shrinkage stage permitted the sampling of the molecular transform at successive positions, thus permitting him to map the variation of the continuous transform. As he examined only the centric (*h*0*l*) reflections of the monoclinic crystals, he could observe when the sign changed from 0 to π in the centric projection (Fig. 1.2.4.1). Thus, he was able to determine the phases (signs) of the central part of the (*h*0*l*) reciprocal lattice. This technique is essentially identical to the use of diffraction data from different unit cells for averaging electron density in the 'modern' molecular replacement method. In the haemoglobin case, Patterson projections had shown that the molecules maintained their orientation relative to the *a* axis as the crystals shrank, but in the more general molecular replacement case, it is necessary to determine the relative orientations of the molecules in each cell.

The second of Perutz's techniques depended on observing changes in the intensities of low-order reflections when the concentration of the dissolved salts (*e.g.* Cs_2SO_4) in the solution between the crystallized molecules was altered (Boyes-Watson *et al.*, 1947; Perutz, 1954). The differences in structure amplitude, taken together with the previously determined signs, could then map out the parts of the crystal unit cell occupied by the haemoglobin molecule. In many respects, this procedure has its equivalent in 'solvent flattening' used extensively in 'modern' protein crystallography.

The third of Perutz's innovations was the isomorphous replacement method (Green *et al.*, 1954). The origin of the isomorphous replacement method goes back to the beginnings of X-ray crystallography when Bragg compared the diffracted intensities from crystals of NaCl and KCl. J. Monteath Robertson explored the procedure a little further in his studies of phthalocyanines. Perutz used a well known fact that dyes could be diffused into protein crystals, and, hence, heavy-atom compounds might also diffuse into and bind to specific residues in the protein. Nevertheless, the sceptics questioned whether even the heaviest atoms could make a measurable difference to the X-ray diffraction pattern of a protein.* Perutz therefore developed an instrument which quantitatively recorded the blackening caused by the reflected X-ray beam on a film. He also showed that the effect of specifically bound atoms could be observed visually on a film record of a diffraction pattern. In 1953, this resulted in a complete sign determination of the (*h*0*l*) horse haemoglobin structure amplitudes (Green *et al.*, 1954). However, not surprisingly, the projection of the molecule was not very interesting, making it necessary to extend the procedure to noncentric, three-dimensional data. It took another five years to determine the first globular protein structure to near atomic resolution.

In 1950, David Harker was awarded one million US dollars to study the structure of proteins. He worked first at the Brooklyn Polytechnic Institute in New York and later at the Roswell Park Cancer Institute in Buffalo, New York. He proposed to solve the structure of proteins on the assumption that they consisted of 'globs' which he could treat as single atoms; therefore, he could solve the structure by using his inequalities (Harker & Kasper, 1947), *i.e.*, by direct methods. He was aware of the need to use three-dimensional data, which meant a full phase determination, rather than the sign determination of two-dimensional projection data on which Perutz had concentrated. Harker therefore decided to develop automatic diffractometers, as opposed to the film methods being used at Cambridge. In 1956, he published a procedure for plotting the isomorphous data of each reflection in a simple graphical manner that allowed an easy determination of its phase (Harker, 1956). Unfortunately, the error associated with the data tended to create a lot of uncertainty.

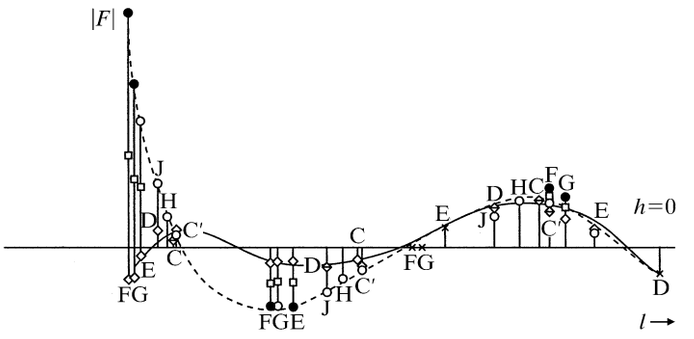

Fig. 1.2.4.1. Change of structure amplitude for horse haemoglobin as a function of salt concentration in the suspension medium of the low-order *h*0*l* reflections at various lattice shrinkage stages (C, C′, D, E, F, G, H, J). Reprinted with permission from Perutz (1954). Copyright (1954) Royal Society of London.

In the first systematic phase determination of a protein, namely that of myoglobin, phase estimates were made for about 400 reflections. In order to remove subjectivity, independent estimates were made by Kendrew and Bragg by visual inspection of the Harker diagram for each reflection. These were later compared before computing an electron-density map. This process was put onto a more objective basis by calculating phase probabilities, as described by Blow & Crick (1959) and Dickerson *et al.* (1961).

One problem with the isomorphous replacement method was the determination of accurate parameters that described the heavy-atom replacements. Centric projections were a means of directly determining the coordinates, but no satisfactory method was available to determine the relative positions of atoms in different derivatives when there were no centric projections. In particular, it was necessary to establish the relative *y* coordinates for the heavy-atom sites in the various derivatives of monoclinic myoglobin and in monoclinic horse haemoglobin. Perutz (1956) and Bragg (1958) had each proposed solutions to this problem, but these were not entirely satisfactory. Consequently, it was necessary to average the results of different methods to determine the 6 Å phases for myoglobin. However, this problem was solved satisfactorily in the structure determination of haemoglobin by using an $(F_{H1} - F_{H2})^2$ Patterson-like synthesis in which the vectors between atoms in the two heavy-atom compounds, H1 and H2, produce negative peaks (Rossmann, 1960). This technique also gave rise to the first least-squares refinement procedure to determine the parameters that define each heavy atom.

Perutz used punched cards to compute the first three-dimensional Patterson map of haemoglobin. This was a tremendous computational undertaking. However, the first digital electronic computers started to appear in the early to mid-1950s. The EDSAC1 and EDSAC2 machines were built in the Mathematical Laboratory of Cambridge University. EDSAC1 was used by John Kendrew for the

* Perutz writes, 'I measured the absolute intensity of reflexions from haemoglobin which turned out to be weaker than predicted by Wilson's statistics. This made me realise that about 99% of the scattering contributions of the light atoms are extinguished by interference and that, by contrast, the electrons of a heavy atom, being concentrated at a point, would scatter in phase and therefore make a measurable difference to the structure amplitudes.'

references