*J. Natn. Sci. Coun. Sri Lanka.* 1976 4(2) : 87—98

*INVITATION CONTRIBUTION*

# X-rays and the Structure of Insulin*

DOROTHY C. HODGKIN

*Department of Zoology, Laboratory of Molecular Biophysics, Oxford University, England.*

Medical history is to the observer full of events that produce sudden and spectacular changes : before, many die, and afterwards, and suddenly, they live. The isolation of insulin in the summer nearly 50 years ago was an event of this kind. Reading the account of the experiments of Banting and Best[3] in July and August 1921 and of the treatment of the first patients to be given insulin during the following winter is still a moving experience. Old and young who had "melted away" with diabetes, as Aretaeus described hundreds of years earlier, began again to live.

The isolation of insulin was made possible by the recognition that the hormone present in the $\beta$ —islet cells of the pancreas was protein in nature and had to be protected from digestion by the proteolytic enzymes of the pancreas during extraction. It is, in fact, a rather tough protein, as proteins go, and this also helped to preserve its activity intact during the processes first used of acid alcohol extraction. Its purification was greatly helped in 1925 by its crystallization by J. J. Abel, Professor of Pharmacology at Johns Hopkins University.[1] For a little time his method of crystallization—dissolve in acetic acid and add alkali—seemed unreliable until Scott and Fisher (1934) in Toronto showed that Abel's crystals of insulin contained zinc, present also in the pancreas, and that zinc or some other similar metal had to be added to secure crystallization by his techniques.

## Crystals and X-Rays

It was a sample of zinc insulin crystals, prepared for clinical use by Boots Pure Drug Company, that led to my own interest in insulin. The sample was given me by Professor Robert Robinson in 1935 because I was interested in protein crystals. This was the result of experiments carried out by J. D. Bernal at Cambridge the year before, passing X-rays through crystals of the enzyme pepsin. The crystals, 2 mm across, prepared by John Philpot in Uppsala, had been brought over to Cambridge in their mother liquor. Bernal discovered that if they were kept wet when X-rays were passed through them they gave many diffraction spectra indicating that within the pepsin crystals there were very large molecules of definite size and regular arrangements. When the insulin sample came into my hands I looked up all that I could about it, grew large enough crystals by Scott's method—about 0.25 mm across- and found that they also gave good diffraction effects with X-rays.[6]

*Reprinted by kind permission of the distinguished author and the British Medical Journal.

The idea behind our experiments with crystals and X-rays was that we should be able to determine the chemical structures of complicated molecules from the relative positions in space of the atoms within them. Already in 1935 we knew in theory how we might do this. The diffraction spectra produced when X-rays are passed through crystals are the consequence of interference between wave trains scattered by the electrons in the atoms and can formally be recombined as in a lens to show the periodic atomic arrangements responsible for their appearance.

W. H. Bragg pointed out in 1915 that though we have no lens capable of recombining the X-ray waves, they could be recombined mathematically by the use of the relations discovered by Fourier, who represented periodic structures by adding together waves of known phase and amplitude. W. L. Bragg and others made the necessary measurements and calculations on simple crystals of known structure—sodium chloride, a silicate mineral, diopside, and hexachlorobenzene—and showed that the atoms appeared as discrete peaks of electron density. The chief experimental difficulty in extending the results to crystals of unknown structure involved the experimental determination of the phase relations of the scattered waves (readily calculated for crystals of known structure); the amplitudes were easily measured from the intensities of the diffraction spectra. However, in 1927, Cork had proved in a very limited experiment on alums how phases might be found. If the crystal studies could be modified by changing single atoms to heavier atoms in an isomorphous way—that is, without change of the structure as a whole—the difference in the intensities of the diffraction spectra could provide evidence on their phase relations. A process of direct structure analysis was therefore conceivable.

The first X-ray photographs of insulin crystals in 1935 defined the size of the unit rhombohedron—it contained a molecular weight of protein of the order of 36,000. Since the crystals had trigonal symmetry, this unit could be subdivided into three equivalent parts, each with a weight of about 12,000. One could calculate from the empirical analysis of the crystals that such a unit contained about 520 carbon, 130 nitrogen, 150 oxygen and 12 sulphur atoms, and that there were also present a large number of water molecules in the crystal. Before setting out to find the atomic positions in such a complicated crystal, it seemed desirable to try direct X-ray analysis on something simpler. Robertson and Woodward[12] showed that the method worked very beautifully with the phthalocyanines. We set out at the beginning of the war to try to find by this method the structure of penicillin.

**Penicillin and Proteins**

There were complications with penicillin, eventually solved by X-ray measurements involving crystals of sodium, potassium, and rubidium benzylpenicillin, of which the last two were isomorphous. The first three-dimensional map of the electron density that we calculated illustrates very well the character of our evidence. The

map is derived by evaluating the electron density from measurements and calculations at intervals throughout the body of the crystal unit—the figures are generally plotted on sheets perpendicular to one crystal axis. Contours at equal density intervals define the atomic peaks ; they may be drawn on glass or Perspex and the sheets stacked together to show the three-dimensional arrangement of the atoms. So one can see directly from the map in Fig. 1 that there is a four-membered $\beta$-lactam ring in penicillin fused to the five-membered sulphur-containing thiazolidine ring—results that seemed very surprising at the time they were first seen.[7]

Penicillin is still a very small molecule ( $C_{16}\,H_{18}\,O_4\,N_2\,S$ ) compared with insulin, and further improvements in the development of electronic computers and in making large numbers of intensity measurements were desirable before an electron density map could be calculated for any protein. The most difficult problem, however, proved to be the preparation of an isomorphous series of heavy-atom-containing derivatives of a protein crystal, solved first in the case of haemoglobin and myoglobin by Perutz and Kendrew. For haemoglobin, chemical attachment of mercurials to *SH* groups in the molecules proved possible.[11] With myoglobin a more hazardous process of introducing heavy atoms into the spaces filled with solvent in the crystals was found to be successful.[18] In both cases three- dimensional electron density maps were calculated and found to be interpretable—given a good deal of information on the chemical structure of the two proteins studied.

The difficulties of interpreting the electron density maps of protein molecules even when the phase problem has been solved arise from their large size and the fact that in the crystals they form they are almost literally floating in solvent. As a result the X-ray spectra observable fade out at large angles to the X-ray beam and consequently the definition of individual atomic positions decreases. The electron density appears as streaks and peaks of various shapes representing atoms individually unresolved, and chemical information is desirable to help to interpret them.[9] Whereas with penicillin we found the chemical structure from the electron density map, with insulin the knowledge of the chemical structure obtained by Sanger played an essential part in helping us to interpret our map.

**Insulin**

Sanger's work on insulin defined the sequence of residues to be expected in the insulin we studied—pig insulin .[13] The formula shows their arrangement in the two chains A and B and permits us to count the atoms in the molecules exactly to compare with the first crystallographic counts we made. In the crystal it was clear from our early measurements that there were two Sanger molecules in the asymmetric unit. We chose pig insulin for our studies because it had been found to crystallize particularly well by Schlichtkrull in Copenhagen in the course of his work

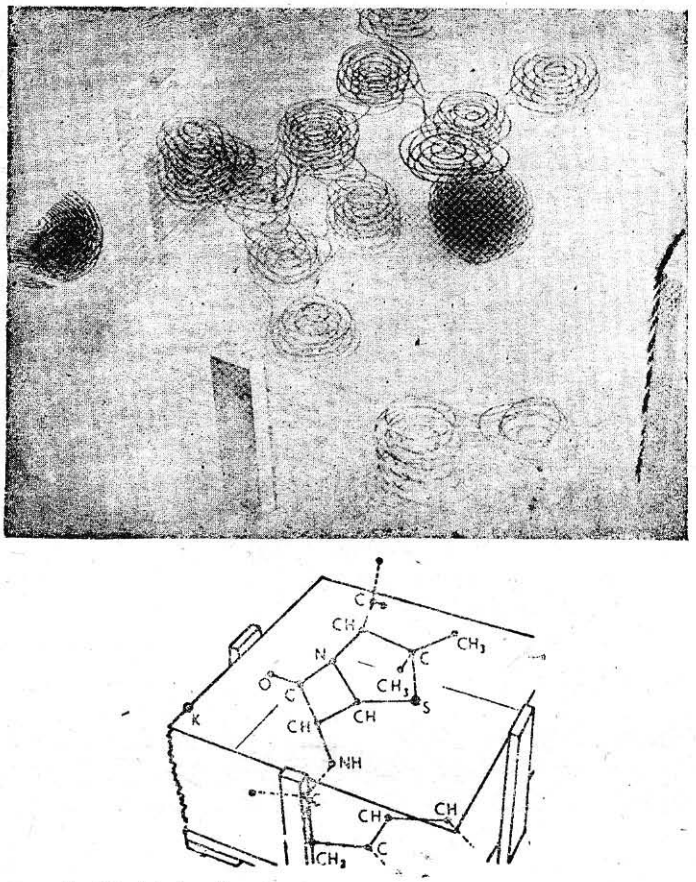
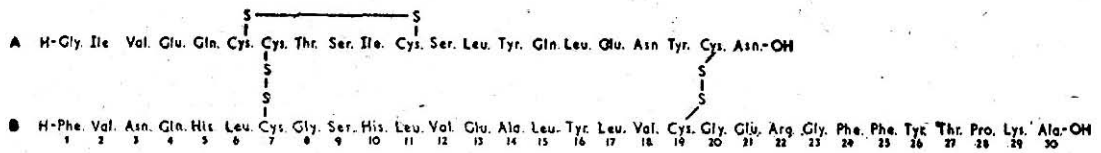



FIG. 1—Model showing the electron density map calculated over part of the crystal unit for potassium benzylpenicillin. As the diagram below indicates, it defines clearly the way in which the atoms are bonded together.

on slow-acting, partly crystalline insulin preparations for clinical use. Schlichtkrull[14] also found that the proportion of zinc in the crystals was two atoms for six Sanger molecules, and this figure defines the zinc atomic positions in our crystals as along the crystal threefold axis.

Our first idea for forming an isomorphous series of insulin crystals was naturally to introduce heavier atoms than zinc by crystallization ; apart from cadmium, not quite heavy enough for our purpose initially, this did not work. However, we did find it possible to replace zinc by lead in the crystals in a rather remarkable way. The crystals were left overnight standing in a dilute solution of a chelating agent, ethylenediamine tetra-acetic acid, which took the zinc out of the crystals. The following night they were left in dilute lead acetate solution. Lead ions occupied

both the old positions of zinc in the crystals and also other sites—one approximately between the two zinc positions, others that proved to be on the outskirts of the molecule. The substitution pattern was quite complicated. But by late 1968 we were able to make much more accurate measurements of the intensities of the X-ray diffraction effects,




The primary structure of pig insulin

and could interpret the small differences between substituted and unsubstituted crystals to place the heavy atoms within the crystals. Further experiments, leaking molecules and ions containing mercury and uranium into the crystals, also proved effective. By the summer of 1969 we had amplitudes measured and phases computed for some 2,000 independent X-ray spectra and so calculated a three-dimensional electron density map for insulin, drew it up on Mylar sheets at a scale of 1 cm to 1 Å, and looked at the answer.[2]

I used to say the evening that I developed the first X-ray photograph I took of insulin in 1935 was the most exciting moment of my life. But the Saturday afternoon in late July 1969, when we realized that the insulin electron density map was interpretable, runs that moment very close. As we expected, the positions of the zinc ions were marked by large electron density peaks but around them were three peaks with a shape that would fit with histidine, each attached to a strand of density, helical in form. We could trace, as in Fig. 2, other peaks attached to the helical chain which had the shape expected from the insulin sequence for the residues neighbouring on histidine B 10.

**Arrangement in Space of Atoms in Molecules**

From this point we began to trace the whole arrangement in space of the atoms in the molecules. There were occasional setbacks and false moves ; during the last 18 months we have built much more accurate models than our first, matching as precisely as we can the form of the electron density against atomic positions. There are still a few obscurities, involving particularly some of the polar residues on the outside of the molecule extending into the solution. But, within a little, we have recorded positions in three dimensions for the now exactly known numbers of atoms in the insulin molecules.[4,5]

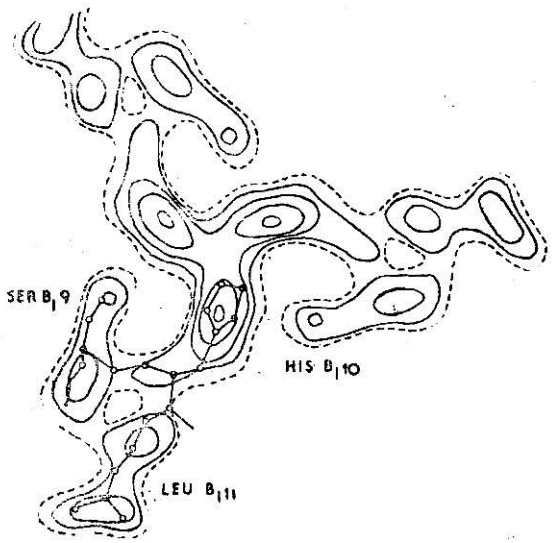



FIG. 2—Electron density contours in a plane perpendicular to the threefold axis, showing superimposed positions of the histidine, leucine, and serine residues. Filled circles— atoms close to section; open circles—atoms within 1 A of section

Owing to the complexity of the crystals studied we see two insulin molecules in our map which are very similar but not exactly identical. They are shown in Fig. 3. Each is a somewhat irregular object over the surface of which extend all the polar residues of the molecule and some non-polar residues. In both, the B chain follows the same general course. The first eight residues are extended ; the chain then passes into a rigid $\alpha$ helix between 9 and 19 ; from 19 to 21 it makes a U turn and again follows an extended form from 22 to 30. The A chain forms a small, compact loop resting within the solid framework of the B chain. There is an initial single turn of $\alpha$ helix followed by the loop closed by the 6-11 disulphide bridge and a later region of somewhat irregularly helical character between residues 12 and 19.

In the crystal, and probably also in many solutions, the two molecules together form a close dimeric unit. Fig. 4 illustrates their relative positions as we observe them. They are related approximately by a twofold axis of symmetry at the position shown. The interactions between them are both non-polar, as between the valine residues, B 14, or phenylalanine residues, B 24, along the lines of the twofold axis, and also polar ; particularly between the peptide groups B 24 and B 26 of neighbouring molecules, which lie anti-parallel to one another, there are hydrogen-bonded contacts within a $\beta$-pleated sheet, which closely combines the terminal parts of the B chains.

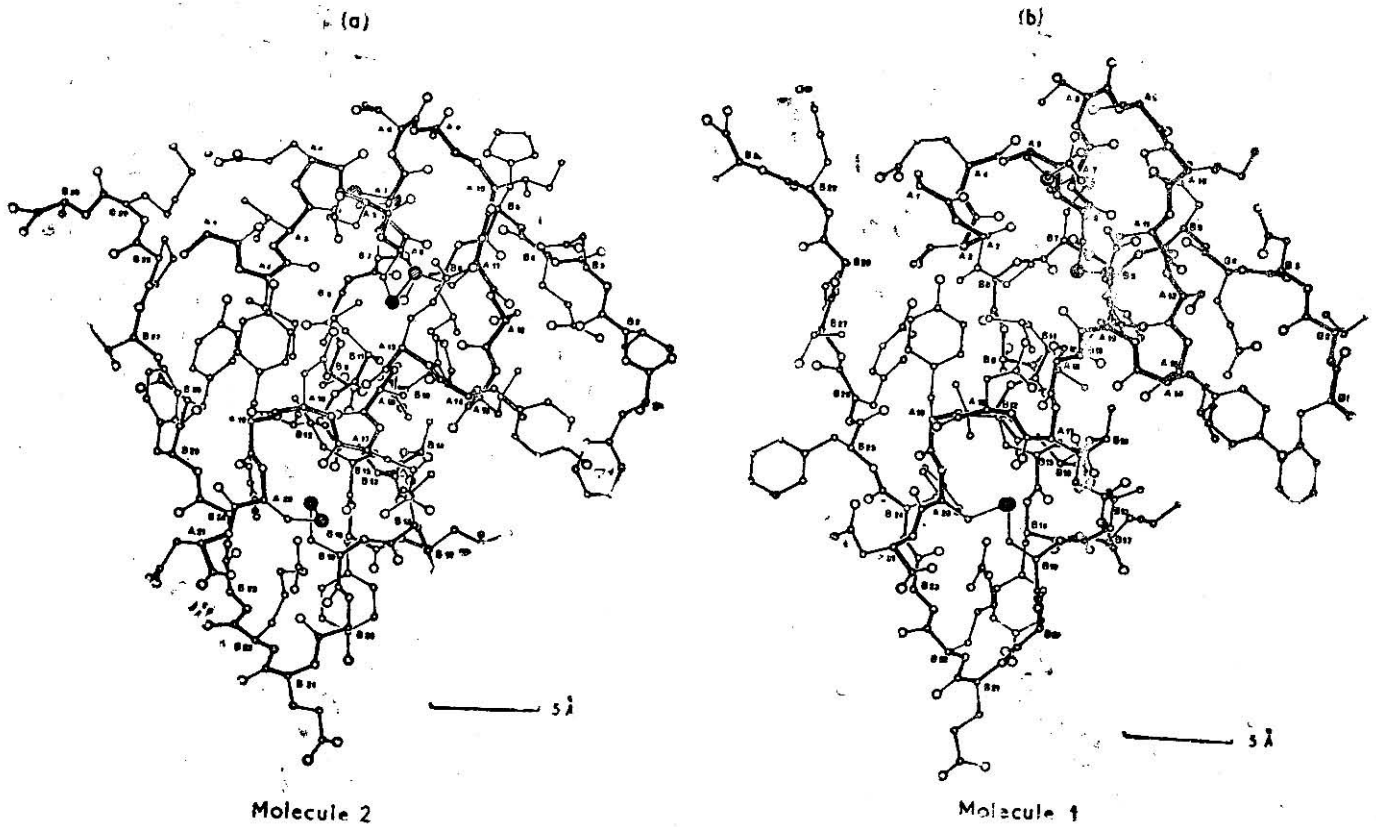



FIG 3—The positions of the atoms found in the two molecules of insulin of the crystal asymmetric unit, projected in equivalent directions normal to the threefold axis. Molecule II has been rotated through 180° from its position in Fig. 4, to compare better with molecule I.

Again in the crystal, and also in solution in the presence of zinc, the three dimers together are connected with two zinc ions to form a hexamer shown in Fig. 5. The main binding forces are the links between the histidine groups and the zinc ions but there are also close contacts, involving polar and non-polar residues, along the region where dimer meets dimer. The hexamer as a whole has a very smooth surface around its circumference, over which many of the polar residues in the molecule are extended. The ends of the dimers, however, project on the upper and lower surfaces of the hexamer as seen in projection in Fig. 4, locking in together when the hexamers pack in the crystal.

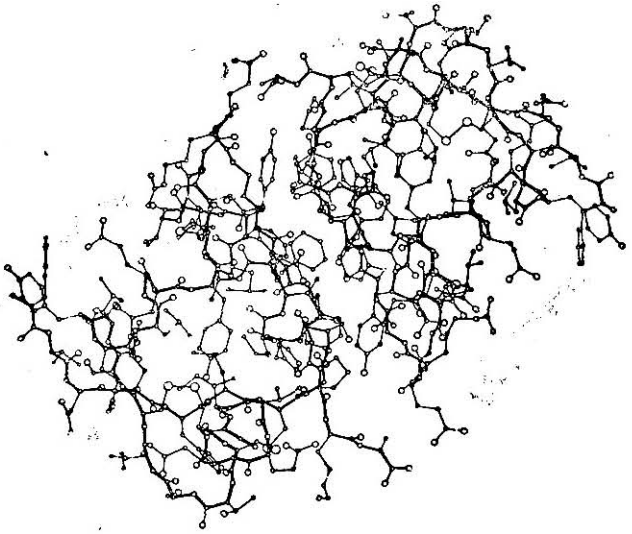


FIG 4—Projection of the atomic positions in the insulin dimer along the approximate twofold axis.

In projection along the threefold axis of the crystals, as in Fig. 5, the hexamer packing simulates close packing of spheres. The molecules make direct contacts with one another yet leaving between them solvent channels, 10 Å or more across, through which other smaller molecules and ions may diffuse (Fig. 6). And with the crystal structure as a whole we return once more to the pancreas. Many past observers have noticed small crystals in the $\beta$ granules of the pancreas of different animals and particularly of the dog. With the electron microscope one can now see roughly spherical units packed within them. A particularly good example, taken of rat islet cells at the University of Sussex, shows lines across the granules representing the packing of particles 50 Å across (Fig. 7).[8] Almost certainly these are insulin hexamers formed around the zinc ions in the pancreas, since 50 Å is very nearly the diameter of the hexamers.

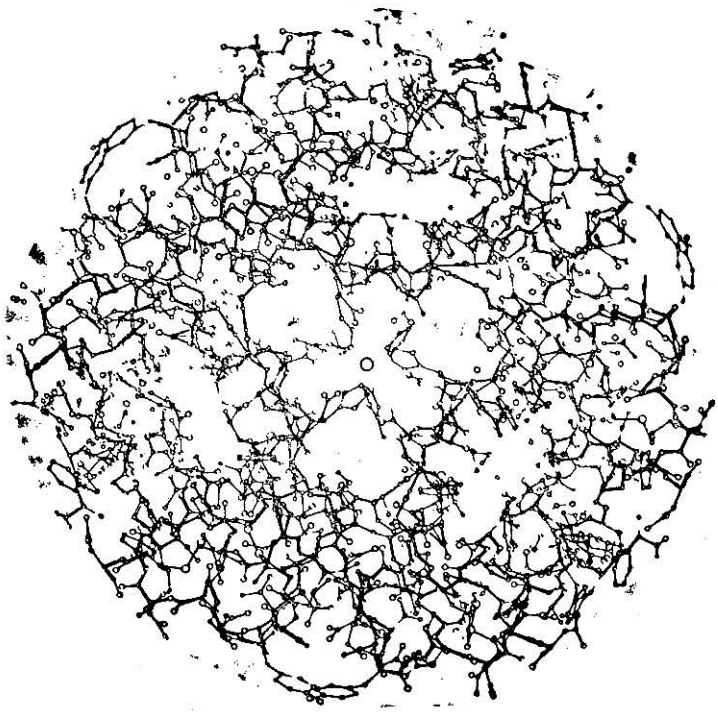


FIG. 5—Projection of the atomic positions in the insulin hexamer, parallel with the threefold axis.

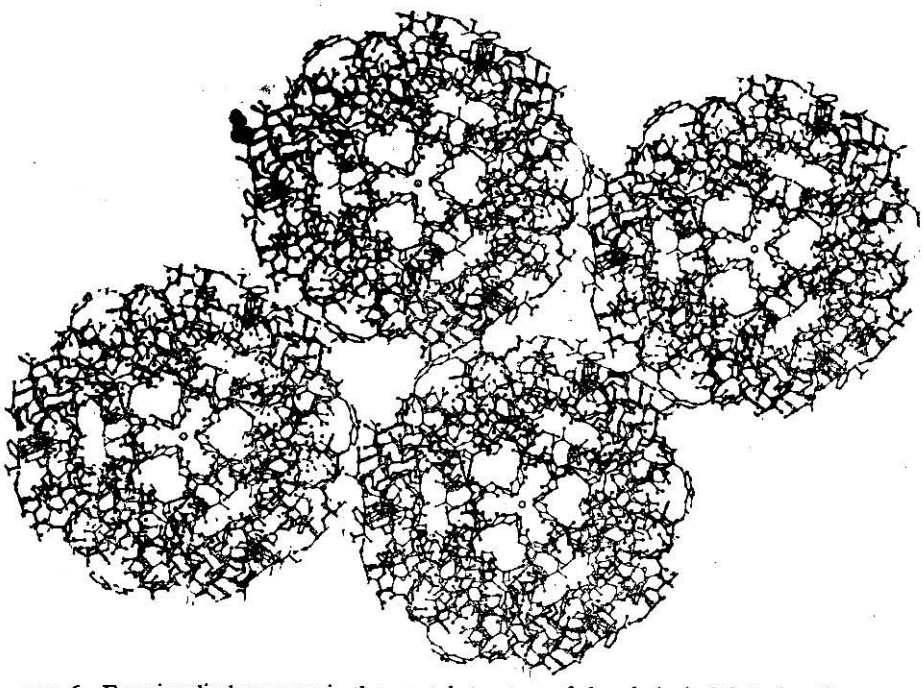


FIG 6—Four insulin hexamers in the crystal structure of rhombohedral 2 Zn insulin, seen projected along the threefold axis.

FIG 7—Negatively stained preparation showing an isolated beta granule. The lines across the granule represent the packing of particles 50 A across—almost certainly insulin hexamers. (Phosphotungstic acid X 200,000.) Reproduced, by permission, from Greider *et al.* (1969).

**Clues in Amino-acid Distribution**

It seems most likely that the formation of hexamers around zinc ions is a way of storing insulin which is found in many creatures, though not in all. From the islet cells the hormone is released for action ; at the dilutions at which it occurs in the body fluids it is very probably present as the dimer or monomer. We can already see certain clues in the amino-acid distribution in these structures that are suggestive in relation to the biological activity of insulin.

First, there are several observations made during the syntheses of insulin carried out in Aachen, Pittsburgh, Shanghai, and Peking. In each case, the syntheses were carried out by making separately the A and B chains, reducing them to the sulphydryl form, and leaving them in solution to unite in the correct order and shape. The A chain alone was observed to have some little activity, the B chain none. The reaction to produce insulin was rather inefficient ; clearly the organization of separate chains into the correct conformations, to make the correct internal links, is not at all automatic. In nature, indeed, a quite different course, via the single-chain precursor proinsulin, is adopted.

On the other hand, some biologically active insulin is formed by the chemical route ; it seems most likely that the B chain tends to fold in the specific form now observed and this may then support the A chain in a biologically active shape through the interaction of certain specific residues.

The nature of these specific residues is suggested by the study of different insulins and the changes that occur with species. The residues that so far are observed as unchanged include all the cystine residues, three glycines, and a number of leucine and isoleucine residues. These are concentrated in the core of the molecule and seem to be largely concerned with maintaining its correct three-dimensional form. Other residues, such as B 24 phenylalanine along the dimer twofold axis or the A chain residues, glycine A 1, glutamine A 5, tyrosine A 19, and asparagine A 21, might, on the other hand, constitute, in some part or other, an active surface ; removal of A 21 or the A1 amino group largely destroys activity. B 24 might well be opened for interaction with a membrane receptor by opening the dimer.

These clues do not yet tell us what it is that insulin does at the molecular level that affects glucose utilization and transport and protein synthesis and so our own continued well-being. But they may help us to devise new experiments so that in time we may understand how this remarkable molecule operates.

## References


1. ABEL, J. J., GEILING, E. M., ROULIER, C. A., BELL, F. K., & WINTERSTEINER, O. (1927). *J. Pharmac. exp. Ther.* **31**: 65.
2. ADAMS, M. J., et al. (1969). *Nature*, **224**: 491.
3. BANTING, F. G., & BEST, C. H. (1921). *J. Lab. clin. Med.* **7**: 464.
4. BLUNDELL, T. L., DODSON, E. J., DODSON, G. G., HODGKIN, D. C., & VIJAYAN, M. (1970). Recent Progress in Hormone Researeh **20**: 1.
5. BLUNDELL, T. L., et al., (1971). *Nature*, In press.
6. CROWFOOT, D., (1935) Nature, **135**: 591.
7. CROWFOOT, D., BUNN, C. W., ROGERS LOW, B. W., & TURNER-JONES, A. (1949). The Chemistry of Penicillin: 310, Princeton, New Jersey, Princeton University Press.
8. GREIDER, M. H., HOWELL, S. L., & LACY, P.E. (1969). *Jnl.Cell.Biol.* **41**: 162.
9. HODGKIN, D. (1960). *Nature*, **188**: 441.
10. KENDREW, J. (1962). Prix Nobel.
11. PERUTZ, M. (1962) Prix Nobel.
12. ROBERTSON, J. M., & WOODWARD, I. (1937). *J. chem. Soc.* 219.
13. RYLE, A. P., SANGER, F., SMITH, L. F., & KITAI, R. (1955). *Biochem. J.* **60**: 541.
14. SCHLICHTKRULL, J. (1956). *Act. chem Scand.* **10**: 1455.
15. SCOTT, D. A., & FISHER, A. M. (1954). *Biochem. J.* **8**: 1592.