# Insulin structure and diabetes treatment

**Jens Brange*, Jean Whittingham†, David Edwards†, Zhang You-Shang§, Axel Wollmer‡, Dietrich Brandenburg‡, Guy Dodson† and John Finch****

†Department of Chemistry, The University of York, Heslington, York YO1 5DD, UK
*Novo Nordisk a/s, DK-Bagsvaerd, Denmark
§Shanghai Institute of Biochemistry, 320 Yue-Yong Road, Shanghai 200031, China
‡Institute fur Biochemie, RWTH Klinikum, Pauwelsstrasse 30, D-52057 Aachen, Germany
**Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

**Insulin fibrils, formed by native and by monomeric molecules, have been prepared. When examined by electron microscopy they are indistinguishable. The packing of the monomeric insulin in the crystal lattice suggests how insulin fibres are constructed; the model explains the accelerated fibre formation by monomeric insulins and identifies mutations to probe the mechanism further.**

THE discovery of insulin in 1921 by Banting and Best and its subsequent use in controlling diabetes was a major advance in medicine[1]. The molecule itself, naturally enough, was intensively studied. It was the first protein whose self-assembly was characterized in detail[2]. Dorothy Hodgkin's own crystallographic studies had demonstrated that the rhombohedral crystal unit cell contained a species of about 36,000 molecular weight and the asymmetric unit a species one third of this, corresponding nicely to Svedberg's measurements. It was the first protein to be sequenced[3] and its structural determination in 1969 by Dorothy Hodgkin and her colleagues in Oxford was amongst the early successes in protein crystallography[4]. During this time major advances in understanding the hormone's biochemistry and its role in biological processes were elucidated; its precursor proinsulin was discovered[5] and its mechanism of action via a membrane-bound receptor was recognized[6] – these were discoveries of general importance.

The solution of the crystal structure of 2 Zn insulin gave Dorothy Hodgkin the opportunity she had been waiting for some 34 years, to establish the structural basis of the hormone's self-assembly to dimer and to hexamer, and to rationalize the vast literature on its chemical and biological behaviour[7]. It was much more difficult, however, to locate the molecule's active surface which, through its contact with the receptor, stimulated glucose transport and other metabolic pathways in the cell. Nonetheless, review of the known insulin sequences and the knowledge of the pathway of various modified (and later mutated) insulin made it possible to identify, at least in part, the receptor-binding surfaces, providing a framework for considering the receptor interactions and, more speculatively, the synthesis of smaller insulin analogues[7-9].

In diabetes treatment, insulin is injected into the musculature where it typically forms a deposit. It then dissolves and travels to the blood where it circulates and can act on its cellular receptors. Although insulin injections have proved generally remarkably effective in controlling diabetes, there are several requirements for therapeutic preparations of insulin. First, they must be stable to temperature and to agitation; if suitable precautions are not taken the protein will precipitate, forming insoluble fibrils. Secondly, because of these phenomena insulin in therapeutic preparations is normally in the form of zinc-containing hexamers, since in this state it is much more stable and resistant to fibril formation. Unfortunately, the large size of the insulin hexamer prevents its absorption, since the hexamers or microcrystals have first to be broken down to release dimers and then monomers, which are small enough to get through the tissue. This leads to a sluggish response to the changes in the blood sugar levels which follow ingestion of food or glucose (Brange *et al.*[10] and references therein). Possibly the slow and inappropriate insulin response experienced by diabetics relying on the injected hormone is responsible for the deterioration of small capillaries and blood vessels, leading to some of the eye, kidney and other circulatory complications that a proportion of patients develop after long periods of therapy.

## Monomeric insulin

It was realized at the time of the solution of insulin's crystal structure that knowledge of its 3-dimensional structure offered the promise of improving the molecule's therapeutic properties. Protein engineering techniques made this hope a practicable possibility and programmes to engineer insulins with altered aggregation properties, but unaltered biological activity, were now feasible. One target was the creation of a monomeric insulin[11]. Monomeric insulins, it was argued, would be rapidly absorbed by the tissues and offered an obvious solution to the hormone's slow release from the injection deposit. The surfaces that govern dimer formation and hexamer formation had been defined from the structure of the zinc-containing hexamer in the 2 Zn insulin crystal. The

close nature of the dimer-forming contacts and some of the hexamer-forming contacts suggested their modification could destabilize hexamers and dimers and thus lead to a stable monomeric molecule. Using protein engineering, two successful strategies were followed to develop monomeric insulins. These were undertaken[11,12] in the 1980s first by Novo Nordisk and later by Lilly.

Novo Nordisk used the strategy to mutate residues that lay opposite to negatively charged carboxylate groups, to carboxylate groups. Thus when dimers formed a charge repulsion would be generated destabilizing this structure[11]. Lilly's approach was based on mutation of the B28 proline which removed favourable van der Waals' contacts between the monomers in the dimer, thus greatly reducing the dimer's stability[13]. These approaches have led to the development of monomeric insulins with improved therapeutic properties, now in clinical trials. However, these monomeric insulins are less stable and much more prone to form fibrils.

## Insulin fibril formation

Insulin fibres were first characterized by Waugh in the 1940s in his extensive studies on insulin stability and precipitation[14]. The conditions that led to fibril formation were identified; they included low pH, elevated temperature, organic solvents and agitation. Various factors such as their ready preparation, reversibility and the dimensions suggested to Waugh that the fibres were constructed from either monomers or dimers and that the molecule was largely unaltered in the process[14,15]. Although the fibrils have been rather little studied, there is a considerable body of spectroscopic evidence to show that they contain $\beta$-sheet – which of course in the globular molecule is a structural feature of the dimer and not of the monomer[16] (Figure 1). The conclusion that the $\beta$-sheet originated from the native dimer has however been undermined by the observation that the monomeric insulins can form fibrils, and that the monomeric insulin without its C-terminal residues formed fibrils more rapidly than native insulin[10].

The role of the B chain C-terminal residues in affecting fibril formation is seen from the behaviour of an insulin cross linked by the peptide bond between B29 and A1. This analogue does not form fibrils, but forms dimers and hexamers as readily as native insulin[17]. These

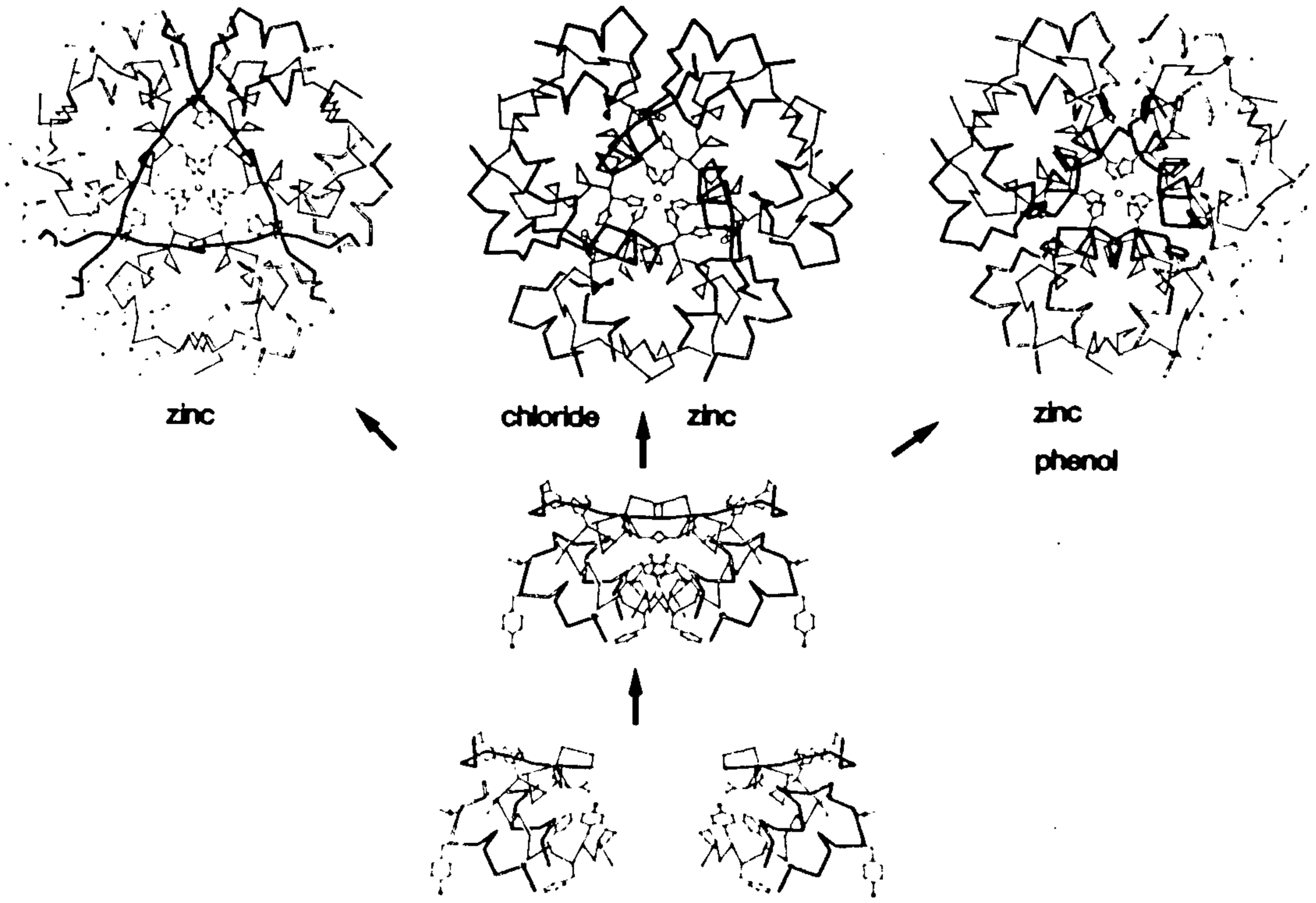


**Figure 1.** Insulin assembly from the monomer to the dimer and then to the hexamer. The molecules are represented as $C\alpha$ traces with selected sidechains. The A chain backbone is shown as a double line, the B chain as a thin line. Note that the dimer-forming contacts are almost exclusively made by B chain interactions.

observations suggest that the A1–B29 cross-link prevents movement of the B chain C-terminus that needs to occur for fibril formation. This would explain the ready ability of insulins whose B chain is truncated at the C-terminus (such as des-B30–B26 insulin discussed below) to form fibrils – the B chain C-terminal residues that need to move away are absent.

## Crystal structure of monomeric des-pentapeptide (B30–B26) insulin

The monomeric nature of des-pentapeptide insulin (DPI) (which lacks the B-chain C-terminal residues involved in the dimer's $\beta$-sheet), its accurate crystal structure and its ready ability to form fibres make it an excellent candidate for investigating the nature of fibre formation.

In contrast to DPI, the native insulin dimer is stable – it is assembled from two monomers, which are packed about a local two-fold axis. The two B chain C-terminal residues form an antiparallel $\beta$-sheet with hydrogen bonds connecting residues B24 Phe and B26 Tyr and their two-fold equivalents (Figure 1). In DPI the removal of the residues B30–B26 has little effect on the rest of the molecule but abolishes its ability to form dimers. The near-normal potency of this monomeric insulin[18,19] shows that the B chain C-terminal residues are not important in receptor binding, and are probably displaced during contact with the receptor. This is strikingly reminiscent of insulin and DPI sharing the ability to form fibrils.

The DPI molecule forms unusual crystals for a protein – they contain only 17% solvent by volume[18]. The removal of residues B30–B26 has exposed a good deal of non-polar surface previously covered, and this exposure is increased by the movement of B25 (now C-terminal) away from the molecule. This new surface is largely aliphatic, consisting of A2 isoleucine, A3 valine, B15 leucine together with B12 valine, B24 and B25 phenylalanine, which are on the native molecule surface and are part of the dimer forming surface. These non-polar residues make extensive and important lattice contacts in the DPI crystal with the non-polar sidegroups of the B chain B6, B15, B17 leucines and B14 alanine (Figures 2 *a*, *b*).

It has been proposed that with insulin (and DPI), fibril formation is driven by assembly along these surfaces[20]. Further inspection of the DPI crystal lattice reveals a set of antiparallel $\beta$-sheet interactions between the B1 and B4 segments of the two-fold related molecules (see Figure 2 *c*). It is therefore proposed that polymerization of the fibrils through these contacts is responsible for the further polymerization of the fibril and that it

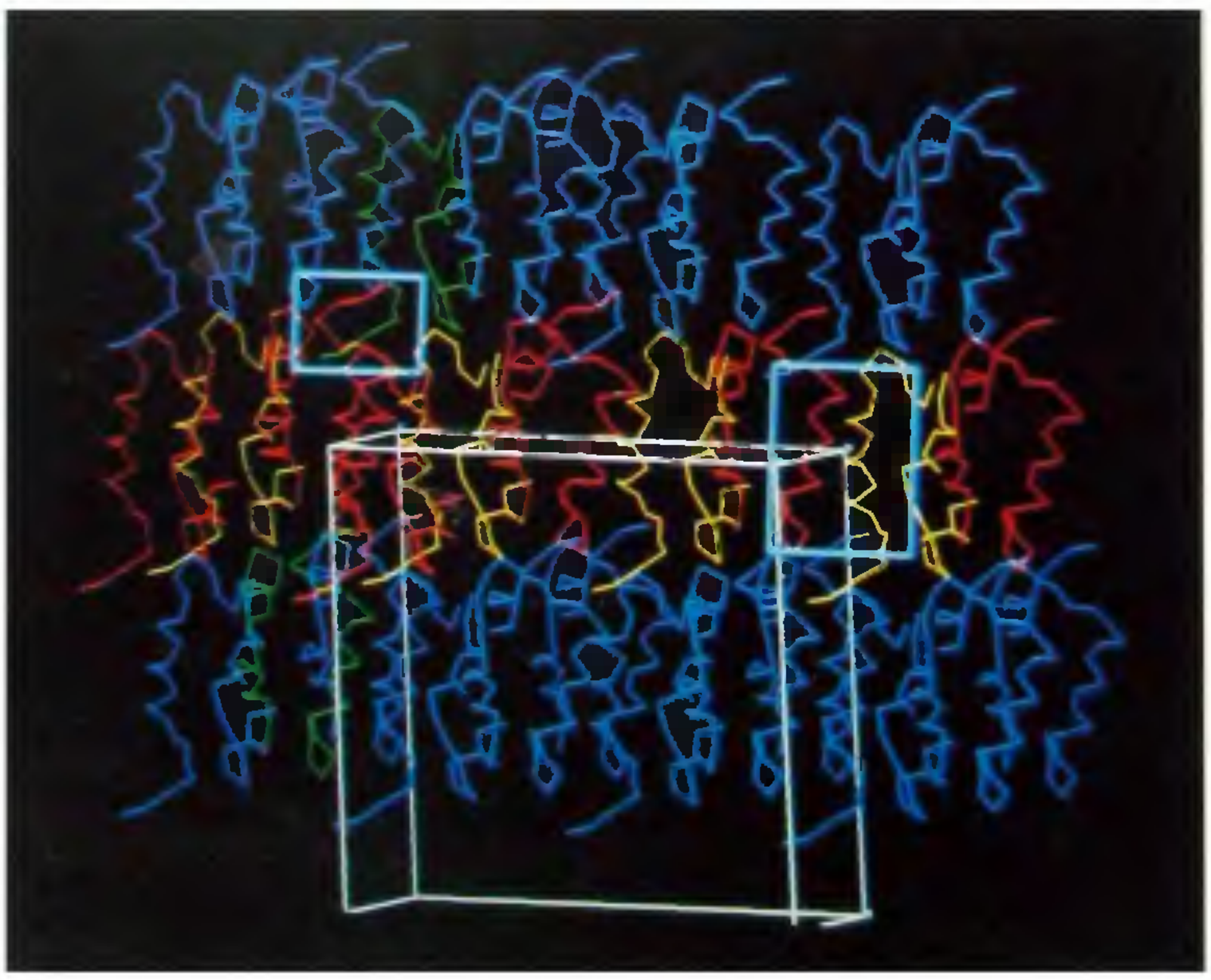

**Figure 2 *a*.** A schematic view of the DPI crystal packing. The unit cell is outlined in white and the DPI molecules are represented as C$\alpha$ traces. The red and yellow molecules are linked by non-crystallographic two-fold symmetry axes. Each two-fold related structure is then translated. The two boxes outline respectively Figures 2 *b* and 2 *c*, highlighting the contacts in the crystal.

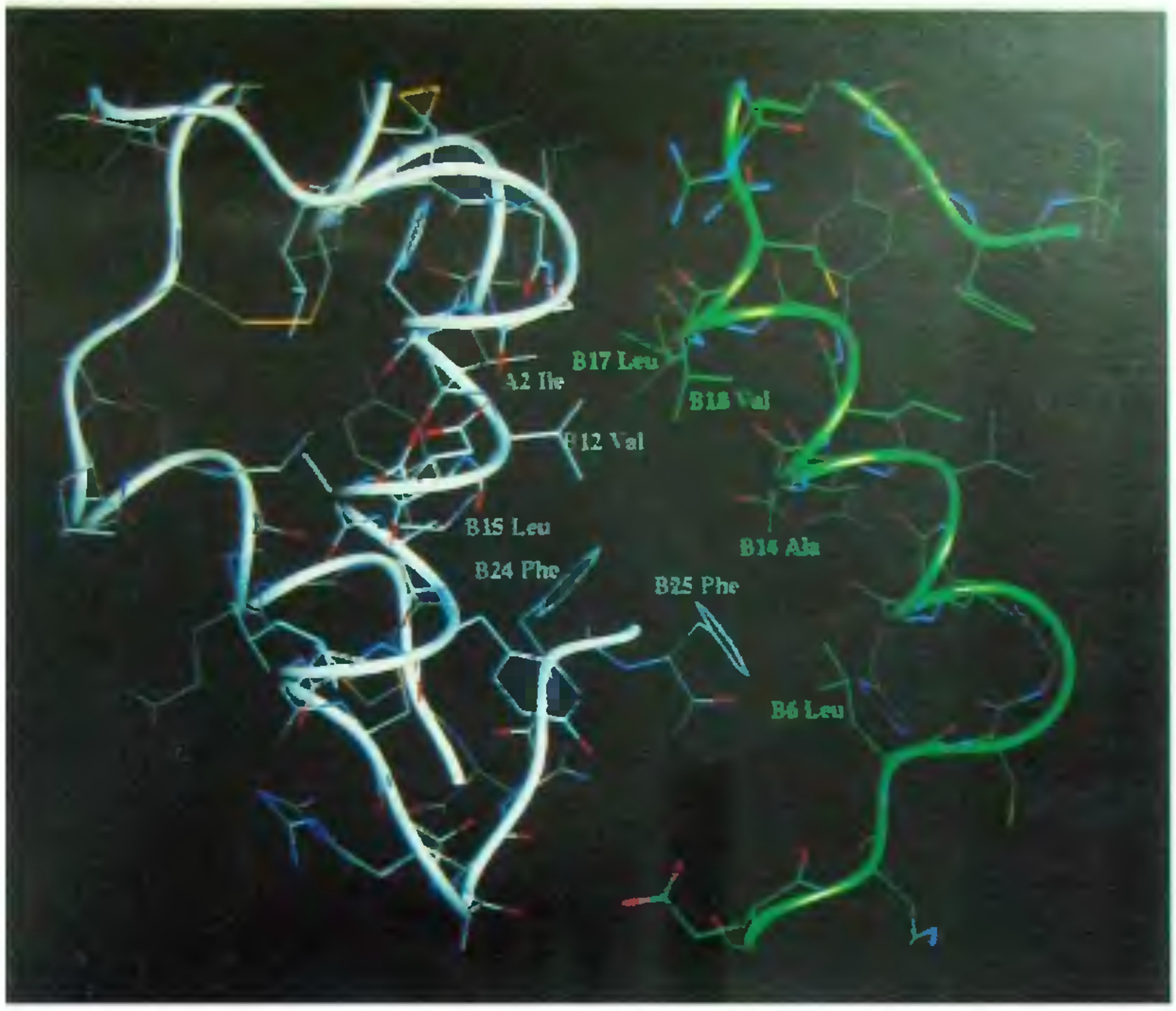


**Figure 2 *b*.** Contacts between two DPI molecules in the crystal. These contacts are non-polar and involve both A chain and B chain residues. This figure distinguishes aggregation between DPI molecules in the crystal from insulin aggregation to the dimer.

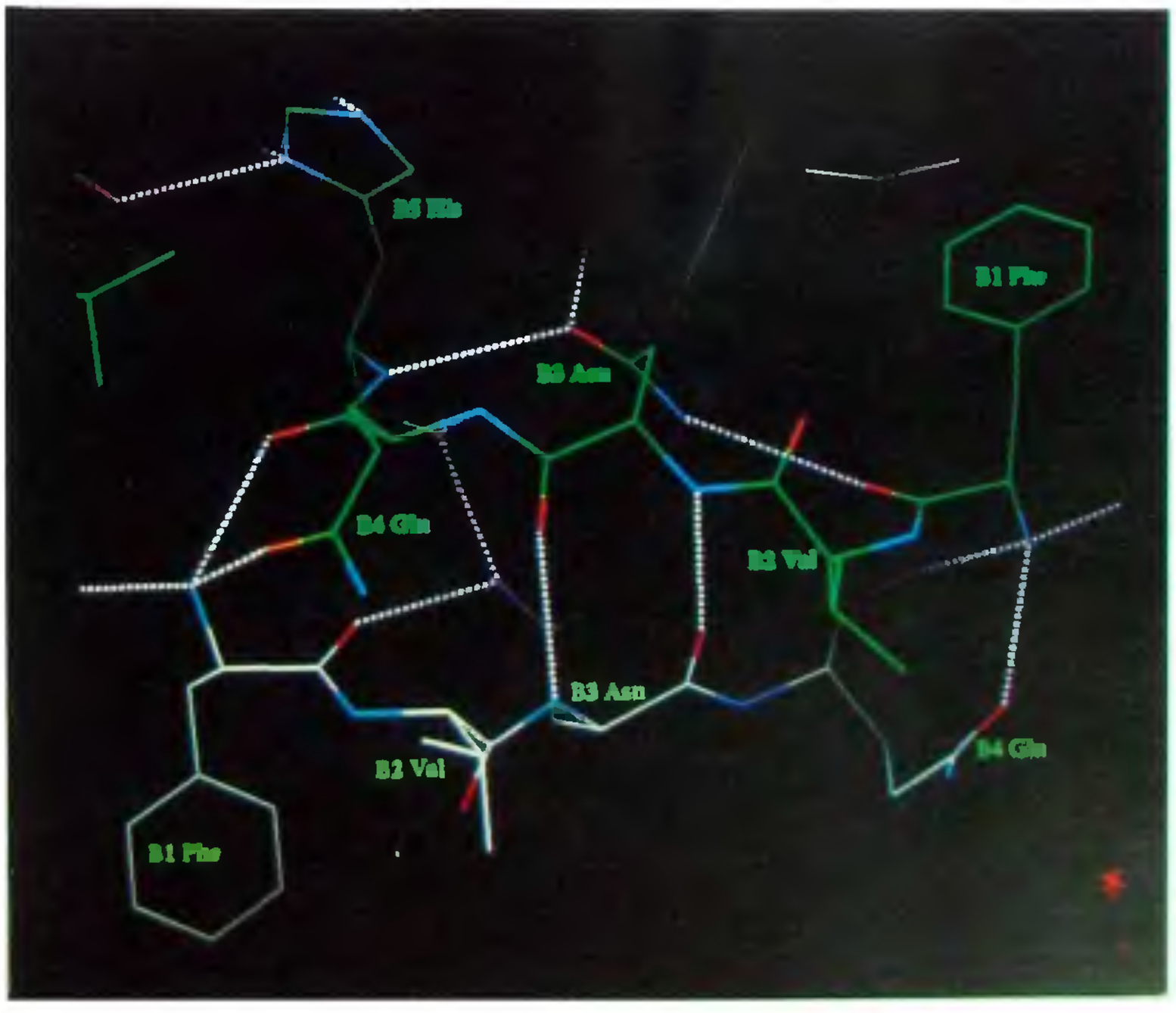


**Figure 2 *c*.** The anti-parallel $\beta$-sheet contacts between residues B1 and B4 in the DPI crystal. A local two-fold axis of symmetry exists centred at B3. Note the involvement of B3 and B4 sidechains in important hydrogen bonding contacts.

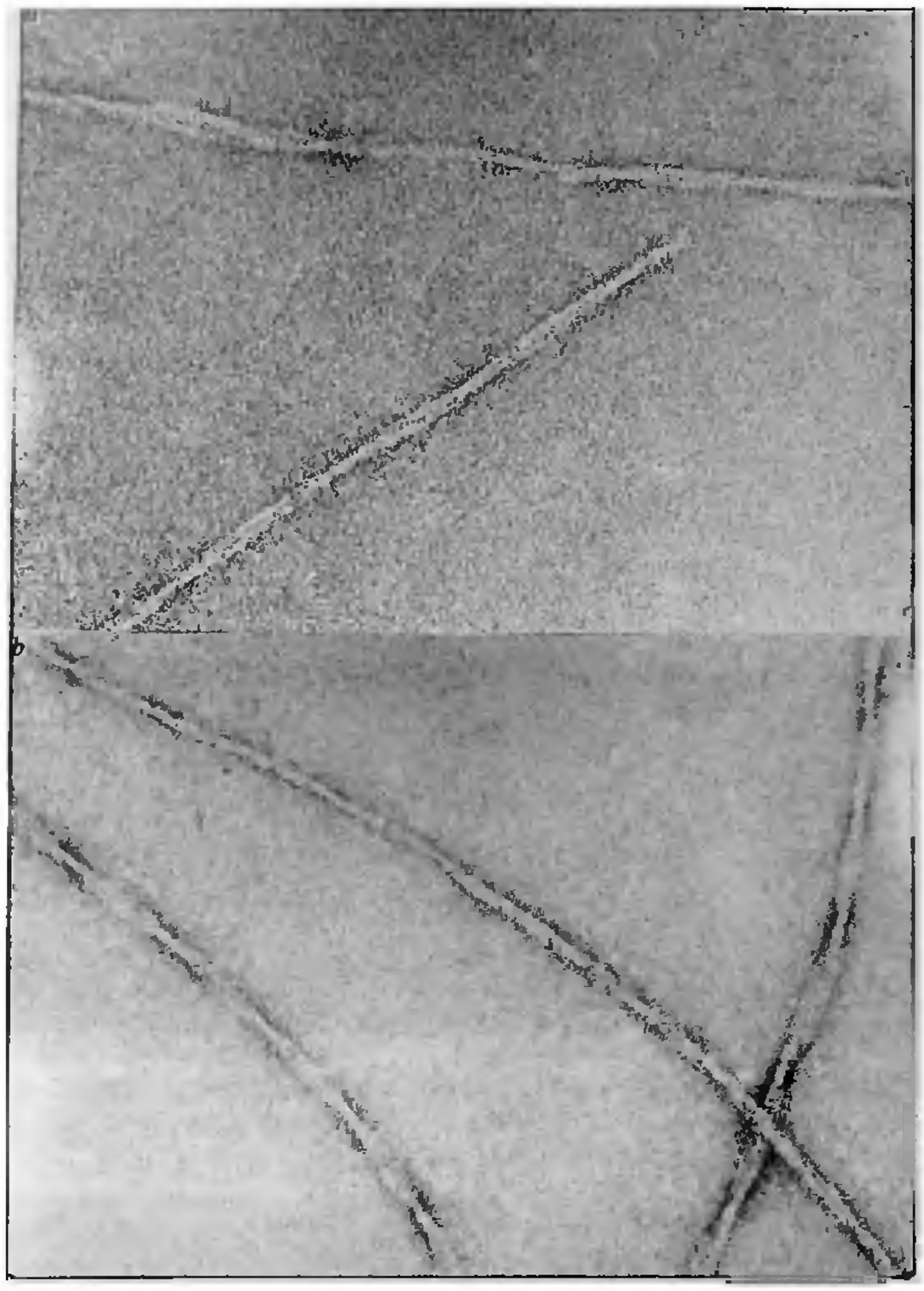

**Figure 3.** *a*, Insulin fibrils formed with HCl. *b*, DPI fibrils formed with HCl (details in text).

is the presence of this anti-parallel $\beta$-sheet structure at B1–B4 that is detected in the insulin fibres. The third surface on the DPI molecule is less likely to aggregate strongly. It includes polar and charged groups and it is largely covered by water structure which helps to mediate the lattice interactions in the crystal. In the conditions of fibril formation tighter interactions between this surface are probable, and may well involve rather different contacts to those seen in the crystal.

## Experiments on DPI fibre formation

### *Fibril formation*

In order to compare the fibril characteristics of insulin and DPI, preparations of the fibres were made by two protocols and examined by electron microscopy. The fibril-forming protocols were based on established procedures described originally by Krogh and Hemmingsen[21]; Vigneaud *et al.*[22] and later by Waugh[14].

(i) Samples of native insulin and DPI were made up to millimolar concentration in 1 ml volumes with 0.01 N HCl and heated for 10 min at 100°C and then left to stand. These are referred to as HCl-treated samples.
(ii) Samples of native insulin and DPI were dissolved and then dialysed first against 0.01 N HCl, and then against sulphuric acid (pH = 2). The solution was then heated for 10 min at 100°C and then left to stand. These are referred to as $H_2SO_4$-treated samples.

The samples were allowed to cool and were left to stand for several days. It was noted that fibril formation, as indicated by viscosity, appeared earlier in the DPI solutions. There were no accurate measurements made on this timing owing to the approximate estimates for determining the state of the solutions.

### *Electron microscopy*

The four samples were examined in the electron microscope after negative staining with uranyl acetate.

Figures 3 and 4 show the electron micrographs of the insulin and DPI samples respectively. It can be seen

that both insulin and DPI when treated only with HCl form fibrils of essentially identical dimensions and character. These appear to consist of two twisted individual fibrils each with dimensions of about 30 Å across. Treatment with $H_2SO_4$ leads to shorter but stubbier fibres. In these the fibrils have apparently assembled together side-by-side into thicker arrays.

## Discussion

The ability of DPI to form fibrils provides conclusive proof that fibril formation proceeds through the monomer and its interactions with other monomers. The fibril forming behaviour of DPI also indicates that a principal mode of contact in fibre formation does not depend on the B-chain C-terminal residues. Indeed the model of the protofibril structure derived from the DPI crystal suggests that it depends on their removal from the contacting surfaces. Analysis of the non-polar contacts in the DPI crystal suggests how these residues would be displaced in the growing fibril without significantly affecting the rest of the insulin molecule's structure[20].

The two-stranded fibre seen in the HCl-treated samples is, we propose, generated by an anti-parallel $\beta$-sheet contact between fibrils, leading to the structure described in Figure 2 *a*. In $H_2SO_4$-treated samples, these contacts are more extensive leading to stubbier fibres, an observation which needs explanation.

This pattern of assembly of the insulin fibre from a correctly folded molecule explains naturally many of the characteristics of insulin fibre formation. Thus, the stronger tendency of DPI to form fibres plus the even stronger tendency seen in des-octapeptide insulin (DOI), (Brange, in press[20]) is explained by their monomeric nature, and perhaps by the absence in these molecules of C-terminal residues which would simplify packing along the non-polar surface. The inability of the A1–B29 cross linked insulin to form fibres is equally accounted for by the covalent bond to the B chain C-terminus preventing their displacement. Thus the non-polar contacts that generate the fibril cannot be formed. Finally the dimensions of the protofibril are in cross-section about 30 Å by 50 Å, very much the same as the linear assembly in the DPI crystal (Brange *et al.*, in press[20]).

If the model is correct a number of modifications which would interfere with fibre formation suggest themselves. Thus, substitution of the non-polar residues A13 Leu, B14 Ala and B17 Leu by charged sidechains would impair the non-polar packing made by these residues along the fibre axis. By contrast, removal of the B-chain N-terminal residues should prevent assembly of the fibrils into fibres. The experiments by Waugh and later workers[23–26] have demonstrated that heat, acid pH, agitation and organic solvents all accelerate fibril formation. All these conditions increase monomer formation by mechanical, thermal, chemical and entropic effects, and have been presented as arguments for the fibril process depending on the monomer. By contrast, other factors, such as zinc and calcium ions and phenol, reduce fibre formation. These conditions favour the molecule's assembly to the hexamer, the most stable species in solution, reducing the monomer population and hence fibril formation.

Protein fibres have been a matter for speculation for many years. The early X-ray diffraction studies, electron microscopy and the various spectroscopic analyses that came later, have all been interpreted as indicating the presence of extensive $\beta$-sheet structure. A general similarity in the orientation of the $\beta$-sheets (the $\beta$-strands are perpendicular to the fibre axis) and their characteristic reaction with Congo Red (associated with anti-parallel $\beta$-sheet)[27] has led to the proposal that the ultrastructure of amyloidal fibres is similar. Because amyloidal proteins, when globular, have different structures it has been

**Figure 4.** *a*, Insulin fibrils formed with HCl followed by $H_2SO_4$ treatment. *b*, DPI fibrils formed with HCl followed by $H_2SO_4$ treatment (details in text).

suggested that the $\beta$-sheet structure in these fibres is produced by restructuring of the proteins. By this mechanism the general apparent similarity in their ultrastructure can be explained[28]. It seems from these rather preliminary studies on insulin fibres, however, that they are built from properly folded molecules, which are little altered in conformation by assembly, as Waugh himself proposed[14]. These studies raise the possibility that the apparent common features of amyloid fibres arise from a common, very limited local $\beta$-sheet motif formed at the interface of adjacent molecules, as they polymerize. Further research on insulin fibres is essential firstly to establish the structure of their constituent molecules and which features of that structure govern fibre formation, and secondly to relate these findings to the X-ray and spectroscopic evidence.

These studies on insulin fibres were stimulated by the unusual packing of DPI molecules in the crystal. Dorothy Hodgkin would have been delighted but not surprised by the possible correlation between the crystalline state and fibre formation and would have regarded it as the kind of outcome a crystallographer should always be ready for.


1. Banting, F. G. and Best, C. H., *J. Lab. Clin. Med.*, 1922, **7**, 464–472.
2. Svedberg, T., *Nature*, 1937, **139**, 160.
3. Brown, H., Sanger, F. and Kitai, R., *Biochem. J.*, 1955, **60**, 556–565.
4. Adams, M. J., Blundell, T. L., Dodson, E. J. *et al.*, *Nature*, 1969, **224**, 491–495.
5. Steiner, D. F., Cunningham, D. D., Spigelman, L. and Aten, B., *Science*, 1967, **157**, 697–700.
6. Cuatracasas, P., *Proc. Nat. Acad. Sci. USA*, 1969, **63**, 450–457.
7. Blundell, T., Dodson, G., Hodgkin, D. and Mercola, D., *Adv. Prot. Chem.*, 1972, **26**, 279–402.
8. Pullen, R. A., Li, D. G., Wood, S. P. *et al.*, *Nature*, 1976, **259**, 369–373.
9. Baker, E. N., Blundell, T. L., Cutfield, J. F. *et al.*, *Philos. Trans. R. Soc. London*, 1988, **319**, 369–456.
10. Brange, J., Hansen, J. F., Havelund, S. and Melberg, S. G., in *Advanced Models for the Therapy of Insulin-Dependent Diabetes* (eds. Brunetti, P. and Waldhäusl, W. K.), Raven, New York, 1987, pp. 85–90.
11. Brange, J., Ribel, U., Hansen, J. F. *et al.*, *Nature*, 1988, **333**, 679–682.
12. Brems, D. N., Alter, L. A., Beckage, M. J. *et al.*, *Prot. Eng.*, 1992, **5**, 527–533.
13. Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D. and Smith, G. D., *Structure*, 1995, **3**, 615–622.
14. Waugh, D. F., *J. Am. Chem. Soc.*, 1946, **68**, 247–250.
15. Koltun, W. L., Waugh, D. F. and Bear, R. S., *J. Am. Chem. Soc.*, 1954, **76**, 413–417.
16. Yu, N. T., Jo, B. H., Chang, R. C. C. and Huber, J. D., *Arch. Biochem. Biophys.*, 1974, **160**, 614–622.
17. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Xiao, B. and Markussen, J., *J. Mol. Biol.*, 1991, **220**, 425–433.
18. Bi, R.-C., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Giordano, F., Reynolds, C. D. and Tolley, S. P., *Acta Crystallogr.*, 1983, **B39**, 90–98.
19. Taylor, R., Hetherington, C. S., Tolley, S. and Dodson, G., *Diabetologia*, 1984, **27**, 337A.
20. Brange, J., Dodson, G. G., Edwards, D. J., Holden, P. H. and Whittingham, J. L., *Proteins*, 1997 (in press).
21. Krogh, A. and Hemmingsen, A. M., *Biochem. J.*, 1928, **22**, 1231–1238.
22. Vigneaud, V., du, Sifferd, R. H. and Sealock, R. R., *J. Biol. Chem.*, 1933, **102**, 521–533.
23. Waugh, D. F., Wilhelmson, D. F., Commerford, S. L. and Sackler, M. L., *J. Am. Chem. Soc.*, 1953, **75**, 2592–2600.
24. Schlichtkrull, J., Brange, J., Christiansen, A. H., Hallund, O., Heding, L. G. and Jørgensen, K. H., *Diabetes*, 1972, **21 (suppl. 2)**, 649–656.
25. Sluzky, V., Kilbanov, A. M. and Langer, R., *Biotechnol. Bioeng.*, 1992, **40**, 895–903.
26. Sluzky, V., Tamada, J. A., Klibanov, A. M. and Langer, R., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 9377–9381.
27. Turnell, W. G. and Finch, J. T., *J. Mol. Biol.*, 1992, **227**, 1205–1233.
28. Blake, C. and Serpell, L., *Structure*, 1996, 989–998.