

## Structural changes in the monomeric despentapeptide (B30-B26) insulin crystal

RU CHANG BI\*, ZBIGNIEW DAUTER<sup>†</sup>, ELEANOR DODSON, GUY DODSON, FEDERICO GORDIANO<sup>‡</sup>, ROD HUBBARD, and COLIN REYNOLDS

Department of Chemistry, University of York, Heslington, York, YO1 5DD, England.

<sup>†</sup> Università di Napoli, Via Mezzocannone 4, 80134 Napoli, Italy.

\*Permanent address: Institute of Biophysics, Academia Sinica, Peking, China.

<sup>‡</sup>Permanent address: Department of Biochemistry, Technical University of Gdansk, Gdansk, Poland.

**Abstract.** Insulin modified by the removal of its 5 B chain C terminal residues is monomeric but remains substantially potent. The crystal structures of the beef and insulin (DPI) with two molecules in the asymmetric unit has been determined by x-ray analysis. The 3-dimensional structure of DPI proves to be generally similar to that of native molecule in 2 Zn insulin. More detailed comparison reveals that the slight differences in the two independent molecules of beef DPI are distributed uniformly throughout the structure in contrast to insulin in 2 Zn insulin, where the structural changes are concentrated in specific regions.

The loss of symmetry in the DPI crystal appears to be the inability of the A9 serine to pack effectively in the C2 cell. The efficient packing of the sheep DPI molecule whose crystal structure has also been determined and where A9 is glycine supports this conclusion.

**Keywords.** Insulin; despentapeptide insulin; crystal structure; monomer.

### 1. Introduction

Insulin molecule has been at the centre of biochemical research ever since its discovery in the early 1920s. Its medical importance in controlling sugar levels in the blood and alleviating diabetes has stimulated study of its chemical properties, its behaviour in solution and its biological effects. The molecule's sequence, which was determined by Sanger and his colleagues, provided the first proper chemical description of proteins (Ryle *et al* 1955). Research into the hormone's crystal structure was begun by Dorothy Hodgkin (then Crowfoot) in 1935 and consequently insulin crystals were the subject of the earliest protein crystallographic investigations (Hodgkin and Riley 1968).

A number of experimental techniques through which the phase could be derived were investigated over the years. One very important development in which Professor S. Ramaseshan played a leading part, was the exploitation of anomalous scattering differences produced by heavy atoms introduced into the crystal structure (Ramaseshan and Narayan 1981).

The 3-dimensional structure of the insulin molecule was first determined in 1969 (Adams *et al* 1969). In the 2 Zn insulin crystal, the molecule's A chain of 21 residues was found to fold compactly with an N terminal helix extending from A<sub>1</sub>–A<sub>8</sub> running adjacent and antiparallel to a C terminal helix between A<sub>12</sub>–A<sub>21</sub>. The B chain of 30 residues was seen to contain a central section of helix between B<sub>9</sub>–B<sub>19</sub>; from each end of this the N and C terminal arms extended around the A chain.

In 2 Zn insulin the hormone is organized as a hexamer. Since 1969 the crystal structures of another hexamer, 4 Zn insulin and a dimeric insulin (Cutfield *et al* 1981) and, most recently, of a modified monomeric insulin have been established. Crystallographic studies on normal insulins are restricted to aggregated forms of the molecule owing to its property of self-assembly to dimer and hexamers at low concentrations. The obstacle to the study of the insulin monomer presented by its property of self-assembly has been overcome by the preparation and crystallisation of a modified molecule from which the 5 B chain C terminal residues (B26–B30) have been removed. These 5 residues complete the surface which is buried by dimer formation and provide a complementary pair of H bonds to the partner molecule. Their removal abolishes the molecule's capacity to dimerise and to hexamerise but leaves it with substantial potency (Shanghai Insulin Research Group 1976). The Chinese in their systematic analyses of those regions of insulin have studied the pig DPI molecules in solution and in the crystal co-crystallised with cadmium ions. Comparison between the pig and beef and sheep DPI crystal structures shows their conformations are generally very similar but there are significant differences which will need to be examined (D Stuart and D C Liang, private communication).

The structure of the monomer is of great interest for three principal reasons. First, the hormone is active as a monomer. Secondly, insulin's aggregation could bring about some rearrangements of the structure which in turn could affect our ideas of the relationship between the hormone's structure and activity. Thirdly the molecule possesses considerable flexibility when organised as a hexameric structure. Since there is evidence that conformational flexibility is important in the expression of the hormone's biological activity, crystallographic study of the monomer free of the constraints imposed by aggregation should give an insight into its structural and conformational behaviour relevant to its biological action. It has emerged that the mechanism by which the molecule adjusts is by relative movements of its helical structures; these behave as essentially rigid entities, linked by more flexible extended pieces of polypeptide chain (Chothia *et al* 1983). This behaviour explains nicely the structural changes seen in the insulin molecules in the insulin hexamers where the differences are associated with the crystal-packing requirements forcing one molecule to alter. The despentapeptide insulin molecule provides a further and similar example of this phenomenon which will be the subject of this paper.

## 2. Experimental

### 2.1 Preparation and crystallisation of DPI

The DPI from beef and sheep insulin was prepared by controlled cleavage at the B25–B26 peptide bond using pepsin at 4°C according to the procedure described by Gattner (1975). The product was purified by column chromatography to homogeneity. The beef and sheep DPI samples were crystallized using citrate/tris buffer with  $(\text{NH}_4)_2\text{SO}_4$  as a precipitant (Bi *et al* 1983).

Several crystal forms of beef and sheep DPI have been obtained (Bi *et al* 1983) but only two have been refined. Their main crystallographic details are given in table 1. The structures were initially determined by the molecular replacement method and subsequently refined by using the Agarwal (1978) fast Fourier least squares procedure. Because the resolution of the x-ray diffraction data in these studies is limited, the

**Table 1.** Crystallographic details of beef and sheep DPI.

	Space group	<i>a</i>	<i>b</i>	<i>c</i>	$\beta$	Molecule in ass. unit	X-ray data resolution	<i>R</i>
Beef DPI	C2	52.7	26.1	51.6	93.4	2	1.5 Å	17.0%
Sheep DPI	C2	53.9	25.9	25.6	93.8	1	2.1 Å	21.2%

calculated shifts are inaccurate and consequently the atomic positions have been constrained to conform to proper peptide geometry after each cycle (Dodson *et al* 1976).

The refinement of the beef DPI is now complete and the well-defined atoms are placed with an accuracy between 0.05 and 0.1 Å. Inevitably more poorly ordered atoms (including of course many of the solvent water molecules) are less accurately positioned; their errors range between 0.2–0.3 Å. The sheep DPI refinement is not yet finished but the well-defined protein atoms are nonetheless determined with sufficient accuracy for useful comparison with the beef DPI molecule.

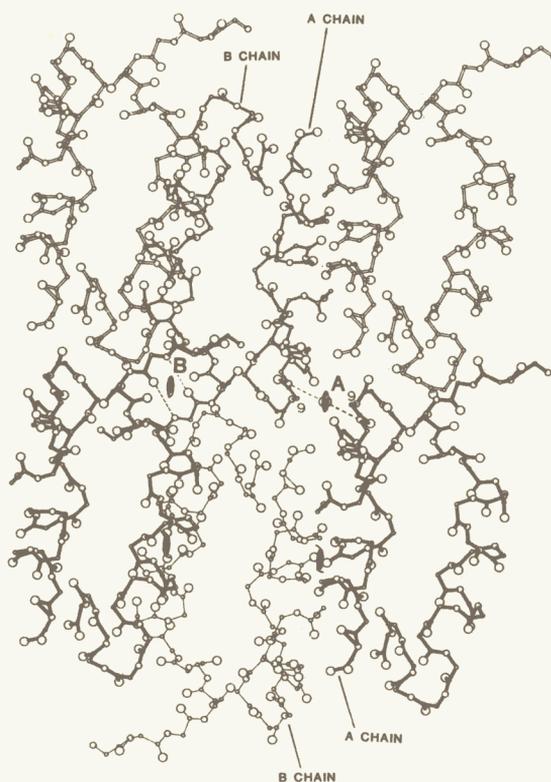
## 2.2 Comparison of beef DPI and sheep DPI crystal organisation

In figures 1(a)–(d) the packing of the beef and sheep DPI molecules in their unit cells shows how one of the crystal axes (labelled A) in the sheep DPI crystal is not exactly obeyed in the beef DPI crystal. The reason for this is clear. At A9 beef DPI has a serine whose sidechains lie very near the now approximate axis of two fold symmetry (A) relating the molecules. In figure 1(c) two sheep DPI molecules related by the operation of the exact two fold axis (A) are shown in a view perpendicular to the axis. The two beef DPI molecules related by the now pseudo axis (A) in their unit cell are illustrated in figure 1(d) in a view equivalent to that in figure 1(c). The mismatching illustrated in the figures shows that in order to accommodate the bulk of the A9 serine on the axis, the equivalent structures in the two beef DPI molecules are displaced by about 4 Å along the direction of the *b* axis.

Inspection of the approximate diad A in figure 1(b) shows that the overall movement which is produced by the need to pack the A9 serines is however accompanied by practically no change in the conformation at A9 and only limited changes at A8 and A10. The ability of the seryl sidechain to resist the packing pressures is, in part at least, a consequence of the well-defined H bond between the A9 seryl OG and A10 NH, an interaction which is also present in 2 Zn insulin, together with the inherent stability of the helical organisation. In figure 2 the two beef DPI molecules related by the pseudo diad are overlapped, giving an impression of the internal changes in the molecules associated with the loss of symmetry.

## 2.3 Comparison between the structures of 2 Zn insulin and beef and sheep DPI

The comparison between the insulin molecule in 2 Zn insulin and the DPI molecules illustrated in figure 3 shows that they are all generally similar except at the B chain N terminal residues B1–B4. These residues extend away from the molecule in DPI in contrast to the situation in 2 Zn insulin, where they wrap around the A chain. There are also significant structural changes at B25 phenylalanine, a C terminal residue in DPI. In



**Figure 1.** Crystal packing of the DPI molecules shown as mainchain for all residues except A9 which in sheep DPI is glycine and beef DPI is serine. The two-fold axis in the sheep DPI are labelled A & B; B remains a crystal axis in beef DPI while its two-fold symmetry is broken. (a) Sheep DPI. Note the water on the axis at A, H bonded to the 2-fold related A10 NH groups. The A & B chains of the two 2-fold related molecules are indicated; A9 is labelled 9.

DPI there are no contacts between B25 and the rest of the molecule; in 2 Zn insulin the B25 residue makes a well-defined H bond interaction to A19 O (Dodson *et al* 1979).

More detailed analyses (table 2) reveals that the DPI molecules from beef and sheep are much nearer to the molecule 1 structure (Chothia *et al* 1983) in 2 Zn insulin, although there is some relative movement in the A and B chains (see table 2). This molecule's structure also occurs in 4 Zn insulin and in the dimeric pig and hagfish insulin crystals (Cutfield *et al* 1981; Reynolds and McCall 1983). For this reason and because of its larger number of internal H bonds, molecule 1 has been regarded as likely to be near the structure of the monomer, a proposal that the DPI analyses have quite confirmed.

Figure 4(a) shows the overlap of the two beef DPI molecules optimised on the B chain helix structure. It can be seen that while the B chain helices are very similar the A chain positioning relative to the B chain is evidently different in the two molecules, indicating there is some movement of the two chains (rms  $\Delta = 0.63$  Å). When the helical segments A1–A9 and A12–A20 are each optimally overlaid, their structural similarity becomes

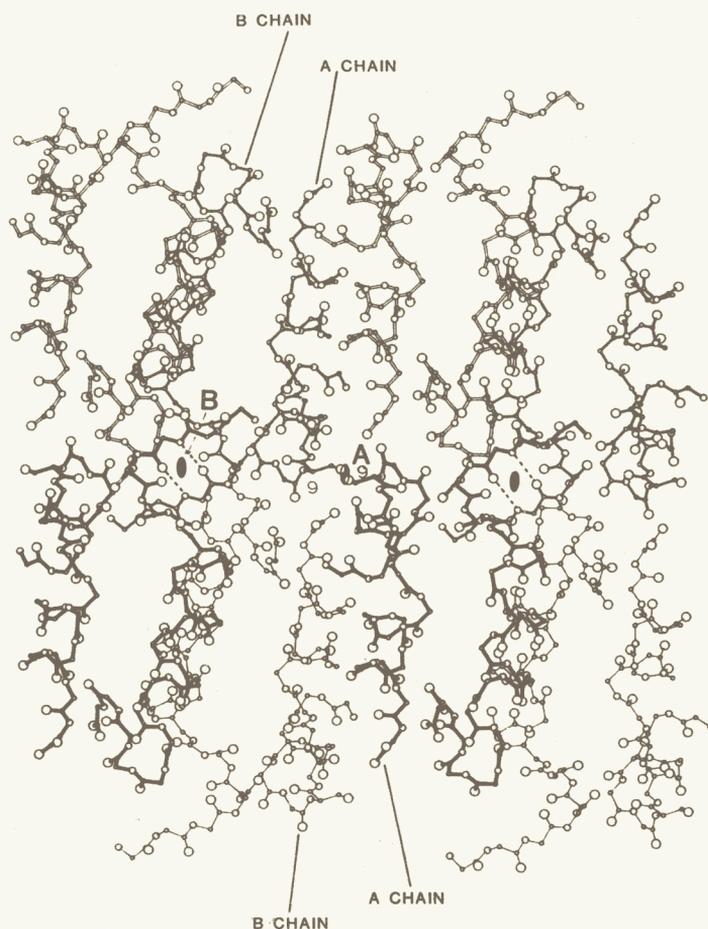


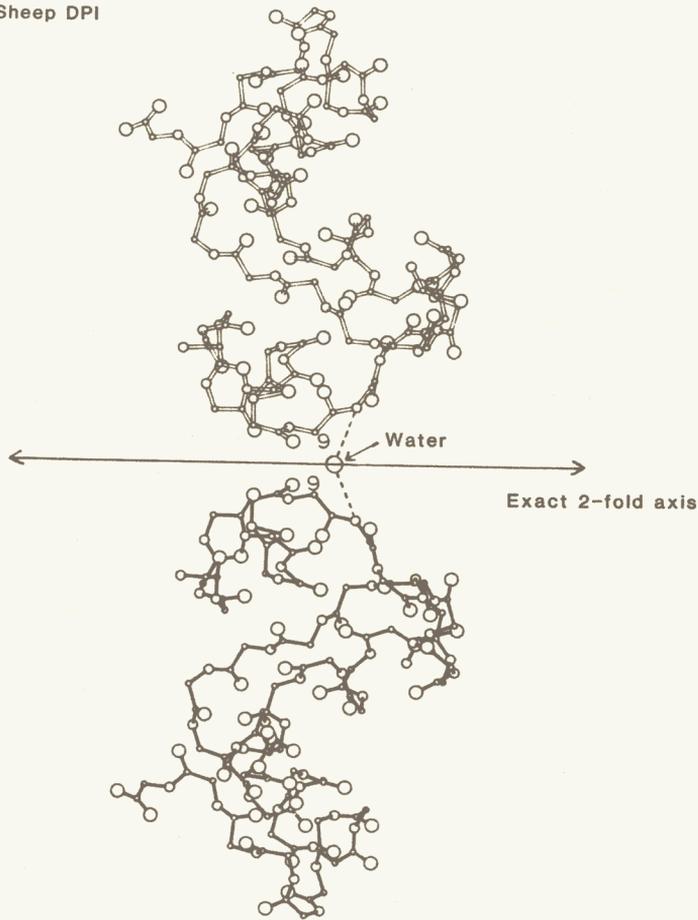
Figure 1. (Contd.)

(b) Beef DPI. Note the close approach by the A9 seryl sidechains to the line of the local axis A. The A & B chains of the two molecules related by the local axis at A are indicated.

apparent from figures 4(b) and (c) and from the magnitudes of the rms discrepancies between the matched structures (table 2). The N terminal B chain residues B1–B5 correlate better with the A chain helical segments than with the B9–B19 helix. The disulphide bond at A7–B7 between the chains is presumably responsible for this linked behaviour. By contrast the structural changes in insulin in 2 Zn insulin, shown in figure 4(d), are focussed at A1–A8 and A19 & B25 sidechains.

This picture of roughly equal distribution of structural movement in the helices in the two beef DPI molecules by which they have adjusted to their crystal environment is confirmed by comparison with the sheep DPI. As mentioned earlier this molecule crystallises with only one molecule in the asymmetric unit permitted because its A9 residue is glycine which, having no side chain, allows symmetrical packing about the crystallographic two fold axis (see figure 1a). Analysis of the two independent beef DPI

## Sheep DPI



**Figure 1.** (Contd.)

(c) Sheep DPI. The two molecules related by A are viewed perpendicular to the axis. The water on the axis A between the two A10 NH groups is indicated.

molecules listed in table 2 shows that they are almost equally different from the sheep DPI. There is however a significantly larger difference in the two beef and the sheep DPI molecules at the A1–A8 helix. In part this is because the A1 residue in sheep DPI is still not satisfactorily determined; there is furthermore rather low electron density at the residues A1–A4 suggesting they are less well-ordered than the subsequent structure. Other differences however occur at A5–A10 which apparently indicate this helix does alter in its internal structure unlike the others. There is on the other hand an important structural difference; the loss of the A9 serine in sheep insulin changes the *main chain* H bonding pattern (between A9 OG and A10 NH). Because the A10 NH no longer interacts with a serine OG (from A9 in beef DPI) it makes H bonds to a water molecule trapped on the local axis (A) between a pair of two fold related molecules (figure 1b). The adjustment of the peptide conformation needed to make this interaction then alters

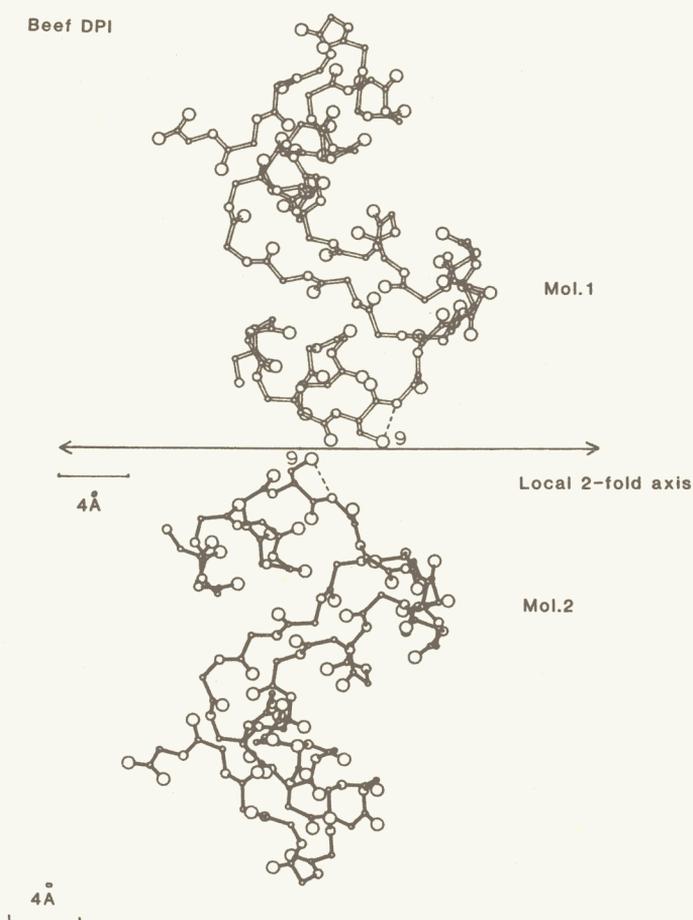


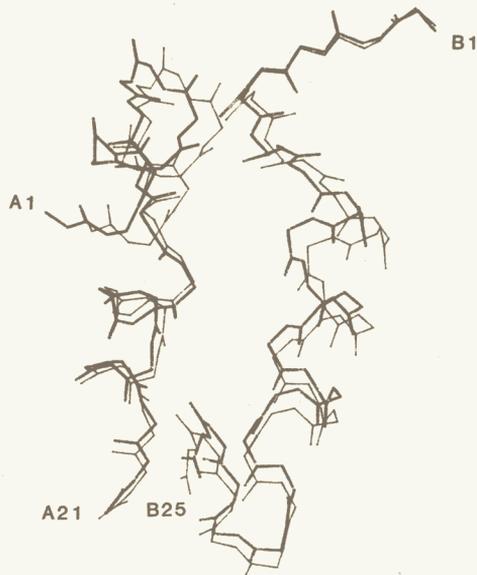
Figure 1. (Contd.)

(d) Beef DPI viewed in the same direction as (c). The two molecules are no longer exactly related by an axis at A; the effect of the A9 serine sidechain is a displacement between the two molecules of about 4 Å in the direction of the local axis.

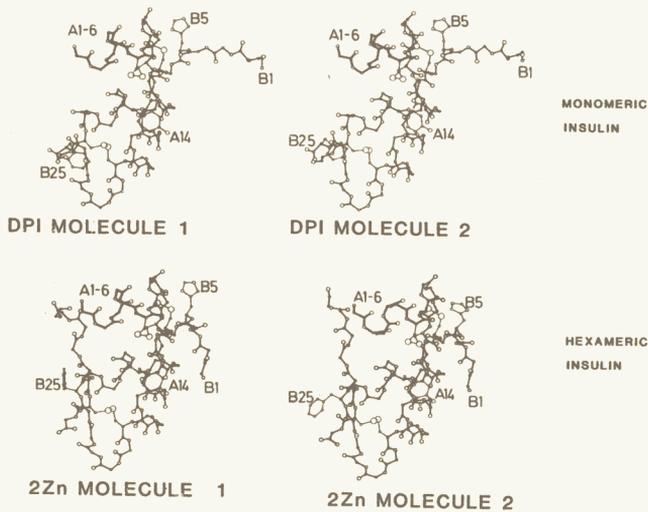
the mainchain H bond contacts between A90 and B5 imidazole ND and the A8NH and A30 disturbing the structure of the helix.

### 3. Conclusion

The loss of two fold symmetry in the 2 Zn insulin hexamer is associated with the packing of the B5 histidines in adjacent hexamers stacked along the three-fold axis. These would intersect in the 2 Zn insulin crystal organisation if the molecules both had the conformation seen in molecule 1 (Cutfield *et al* 1981). It appears that in concert with structural changes at the A chain N terminals and at B28–B30 the two opposing B5 histidines can make satisfactory contacts in the 2 Zn insulin crystals but they lose two-



**Figure 2.** The overlap of the two beef DPI mainchain structures when related by the local diad A. The A and B chains are labelled. Note the movement in the A chain N terminal helix.



**Figure 3.** The pig 2 Zn insulin and beef DPI molecules all shown in an equivalent view which is perpendicular to the direction of the local axis in 2 Zn insulin. There are two crystallographically independent pig insulin molecules in 2 Zn insulin and two independent beef DPI molecules in the C2 beef DPI crystal, and in each case both structures are shown. Only selected sidechains are included to simplify the figure.

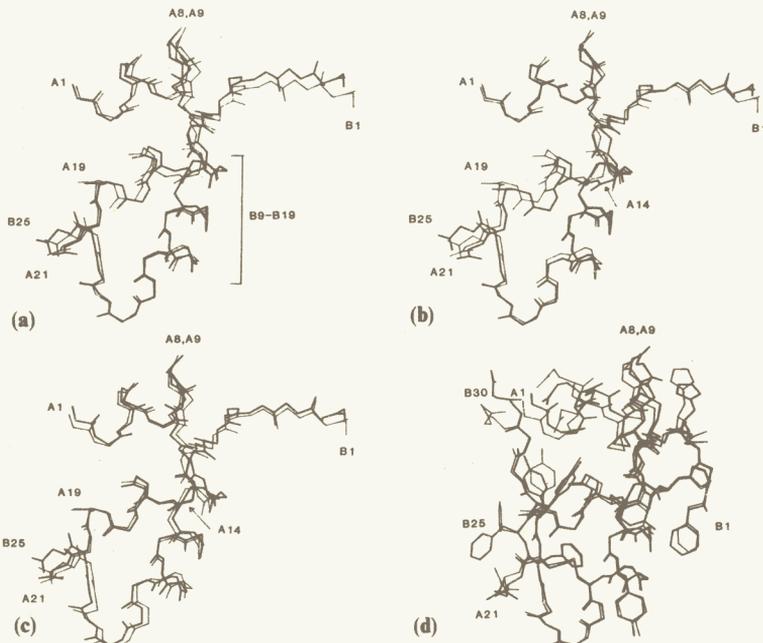
The general similarity in the chain folding of the pig 2 Zn insulin and the DPI molecules is clear; the changes in structure at A1–A9 in 2 Zn pig insulin (molecule 1) are evident.

**Table 2.** The rms discrepancies between the helical segments in Beef and sheep DPI and 2 Zn insulin.

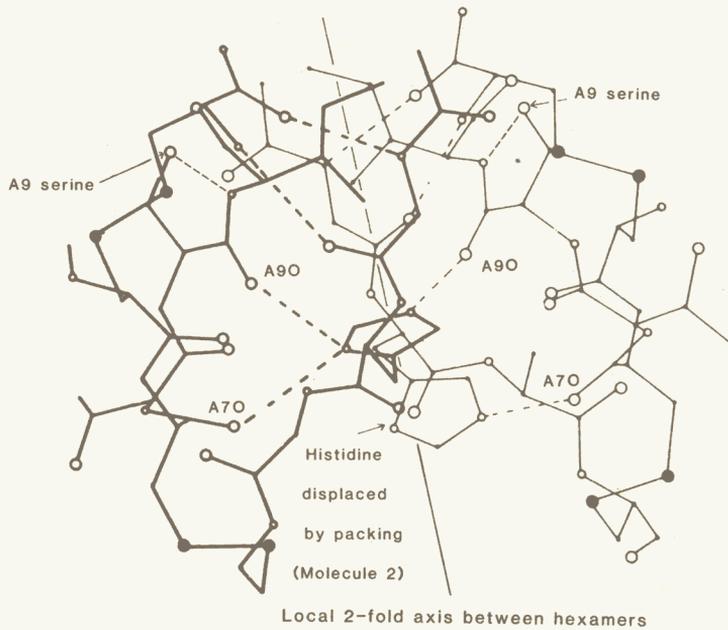
Insulin	Helical Segments	Beef DPI (1)	Beef DPI (2)	Sheep DPI	2 Zn insulin molecule 1	2 Zn insulin molecule 2
beef DPI (1)	B9-B19		0.29	0.29	0.24	0.24
	A1-A8		0.31	0.56*	0.27	0.94
	A12-A20		0.27	0.36	0.24	0.41
	All		0.63	0.43*	0.88	1.04
beef DPI (2)	B9-B19			0.33	0.24	0.27
	A1-A8			0.60*	0.26	1.15
	A12-A20			0.36	0.24	0.28
	All			0.52*	0.61	1.27
sheep DPI	B9-B19				0.38	0.38
	A1-A8				0.59*	1.28*
	A12-A20				0.42	0.44
	All				0.77*	1.09*
2 Zn insulin molecule 1	B9-B19					0.12
	A1-A8					1.10
	A12-A20					0.15
	All					0.46†

\* These overlapped segments include residues A2-A8 only owing to the poor definition of A1.

† These overlapped segments include residues A12-A20 and B9-B19 only owing to the large structural differences at A1-A5, B1-B8 and B27-B30.



**Figure 4.** The overlap of the helical segments in the two beef DPI molecules in the C2 crystal asymmetric unit illustrating their close structural similarity. Only some selected sidechains are shown. (a) B9-B19. (b) A1-A8. (c) A12-A20. (d) The overlap of the two insulin molecules in 2 Zn pig insulin carried out by matching the mainchain between A10-A20 & B1-B27. There is a close similarity in much of the two molecules' structure, except at A1-A9 and B5 and at tyrosyl & phenylalanyl sidechains A19 & B25. More selected sidechains are included than in a, b & c in order to illustrate that they also possess close similarity in conformation.



**Figure 5.** The crystal packing around the B5 histidines in pig 2 Zn insulin. In accommodating to the packing requirements the histidine of molecule 2 has changed its conformation and makes different internal contacts to the B5 histidine in molecule 1 and no contacts with the adjacent molecule. The histidine of molecule 2 makes 2 well-defined internal H bonds (to A7 and A9 carbonyl O), and a good inter hexamer H bond to A9O which has allowed it to retain its original conformation while its companion alters in structure.

fold symmetry in doing so (figure 5). There are obvious parallels between the situation in 2 Zn insulin and the beef  $\text{DPI}$  crystals with its packing together of the A9 serines. But the structural response of the molecules in the  $\text{DPI}$  crystal is quite different. The 2 Zn insulin changes of structure are large but are limited essentially to A1–A5 and B28–B30 (& B25 sidechains). The 2 Zn insulin hexamer probably contains exact two-fold symmetry in solution; the perturbations needed to allow satisfactory crystal packing have to overcome a pre-existing symmetrical structure. It thus appears that the integration of the hexamer prevents much deviation from symmetry except on the surface where the adjustments are relatively large. Those within the hexamer are much smaller, of the order 0.2–0.3 Å and the segments themselves change very little, by between 0.1 and 0.2 Å. The absence of structural change at A9 serine and the structural change in the beef  $\text{DPI}$  molecules in response to the requirements of crystal packing illustrates the expected greater flexibility of a monomeric insulin species. This property, however, derives mainly from the ability of the various segments of secondary structure, such as the helices, to move relative to one another through the flexibility of the extended pieces of main chain that link them (Chothia *et al* 1983).

## References

- Adams M J, Blundell T L, Dodson G G, Vijayan M, Baker E N, Harding M M, Hodgkin D, Rimmer B and Sheat S 1969 *Nature (London)* **224** 491

- Agarwal R C 1968 *Acta Crystallogr.* **A34** 791
- Bi R C, Cutfield S M, Dodson E J, Dodson G G, Giordano F, Reynolds C D and Tolley S P 1983 *Acta Crystallogr.* **B39** 90
- Chothia C, Lesk A, Dodson G G and Hodgkin D 1983 *Nature (London)* **302** 500
- Cutfield J F, Cutfield S M, Dodson E J, Dodson G G, Reynolds C D and Valley D 1981 In *Structural studies on molecules of biological interest* (eds) G Dodson, J Glusker and D Sayre (Oxford: University Press) pp. 527
- Dodson E J, Isaacs N W and Rollet J R 1976 *Acta Crystallogr.* **A32** 311
- Dodson E J, Dodson G G, Hodgkin D C and Reynolds C D 1979 *Can. J. Biochem.* **57** 469
- Gattner H G 1975 *Hoppe Seyler's Z. Physio. Chem.* **386** 1397
- Hodgkin D C and Riley D P 1968 In *Structural chemistry and molecular biology* (eds) A Rich and N Davidson (San Francisco: W H Freeman) p. 15
- Ramaseshan S and Narayan R 1981 In *Structural studies in molecules of biological interest* (eds) G Dodson, J Glusker and D Sayre (Oxford: University Press) p. 233
- Reynolds C D and McCall M Private communication
- Ryle A P, Sanger F, Smith L F and Kitai R 1955 *Biochem. J.* **60** 541
- Shanghai Insulin Group 1976 *Sci. Sin.* **19** 351