

Structural stability in the 4-zinc human insulin hexamer

(conformational changes/x-ray crystallography/slow-acting insulin/zinc coordination)

G. D. SMITH*, D. C. SWENSON*, E. J. DODSON†, G. G. DODSON†, AND C. D. REYNOLDS†

*Medical Foundation of Buffalo Inc., Buffalo, NY 14203; and †University of York, Heslington, York, England

Communicated by David Harker, July 23, 1984

ABSTRACT X-ray studies on human insulins prepared by semisynthetic and biosynthetic methods have recently been undertaken. Human insulin differs from porcine insulin only at the COOH terminus of the B-chain. The present study reports the crystal structure of 4-zinc human insulin, which is used clinically as a slow-acting preparation. The structure has been refined, using 1.85-Å resolution data, to a residual of 0.173. The unit cell is rhombohedral, space group R3, with hexagonal cell constants $a = 80.953$ and $c = 37.636$ Å, and it is nearly isomorphous with that of 4-zinc porcine insulin. As a result of a conformational change of the first eight residues of the B-chain of molecule 1 from an extended conformation observed in the 2-zinc structure to an α -helical one, the coordination around one of the zinc ions on the 3-fold axis has changed, an additional zinc ion in a general position is bound by the hexamer, and additional hydrogen-bonded interactions help stabilize dimer and hexamer formation. Unlike the surface of the 2-zinc insulin hexamer, which possesses a shallow depression containing a zinc ion and its coordinating water molecules, the 4-zinc human insulin hexamer contains a zinc and chloride ion at the bottom of an 8-Å tunnel produced by three parallel α -helices. These α -helices shield the zinc ion from the environment, decreasing the rate of dissociation of the hexamer, and provide an explanation for the slow-acting aspect of the 4-zinc crystalline form.

In mammals, insulin is synthesized in the B cells of the pancreas and stored as a zinc-containing hexameric aggregate. However, this hormone, consisting of an A and a B chain, is a monomer as it circulates in the blood stream and when it interacts with its receptor. Binding and activity data on several insulin analogues suggest that the insulin molecule must undergo a substantial conformational change when it binds to its receptor (1, 2).

Crystallographic investigations have been carried out on insulin derived from various species (3), chemically modified insulins (2, 4), and different crystalline modifications (5-7). The most extensive studies have been carried out on porcine insulin, which differs from human insulin by the change of alanine to threonine at B30. As a result of these studies, the nature of the interactions that hold the hexamer together, the interactions between monomers to produce dimers, and finally, the identity and spatial relationship of the residues that are thought to bind to the receptor are known.

Comparison between the human and porcine 2-zinc insulin crystal structures revealed alterations in the hormone's conformation only at B29 and B30 (8). While other studies have shown that the insulin molecule is flexible, the largest changes in conformation have been observed in the 4-zinc insulin structure. This is of particular interest, since a 4-zinc crystalline preparation is used clinically as a "slow-acting" insulin because it is slow to dissolve and enter the bloodstream. The present study was undertaken to determine the

effect of the single residue change at B30, how the conformational change at the NH₂ terminus of the B-chain affects the relative positioning of the receptor binding residues and the stability of the dimer and hexamer, and finally if the slow-acting aspect of 4-zinc insulin can be attributed to structural changes in the molecule.

EXPERIMENTAL

Human insulin was obtained from Lilly Research Laboratories, where it was produced by recombinant DNA techniques (9, 10). Single crystals of 4-zinc human insulin were grown using published techniques (11). This crystalline modification of human insulin, which contains two molecules in the asymmetric unit, crystallizes in space group R3 with hexagonal cell dimensions $a = 80.953$ and $c = 37.636$ Å and is nearly isomorphous with that of 4-zinc porcine insulin ($a = 80.7$ and $c = 37.6$ Å). A total of 12,420 data to a resolution of 1.85 Å were collected from two crystals using CuK α radiation on an Enraf-Nonius CAD-4 diffractometer as θ - 2θ step-scanned reflection profiles. Each of these profiles was then carefully analyzed to determine the optimum width of the peak to minimize the ratio of $\sigma(I)/I$ (12). Intensities were corrected for absorption (13) and Lorentz and polarization factors. Data from the two crystals were combined through the application of a single scale factor calculated on the basis of the reflections common to both data sets. This procedure resulted in a total of 8174 independent data, of which 6642 were found to have $F^2 > 0$. Structure factors for those data with $F^2 < 0$ were equated to $0.5[\sigma(F^2)]^{1/2}$ and, while not included in the refinement, were used in the calculation of the final residuals.

SOLUTION AND REFINEMENT

The starting coordinates used were those of the nearly isomorphous 4-zinc porcine insulin structure, which differs in sequence only at B30 (alanine is replaced by threonine) and has been refined at 1.5-Å resolution (6, 14). Positional and thermal parameters were refined by the fast Fourier block-diagonal least-squares method (15), using data for which $d < 8$ Å. Periodically, the structure was regularized and made to conform to idealized geometry by the use of the computer program MODELFIT (16). At frequent intervals, F_{obs} Fourier and systematic ΔF maps (a series of ΔF maps with approximately 1/10th of the atoms removed at a time) were calculated and carefully examined to verify the positions of existing atoms, to obtain coordinates for the 75 side chain and 22 main chain atoms that were not located in the 4-zinc porcine study (6, 14), and finally to locate additional water molecules in the solvent region.

Although the side chains of B1 phenylalanine, B2 valine, and B30 threonine of molecule 1 and B30 threonine of molecule 2 could not be located, several side chains were found to have two distinct conformations. These were the A4 glutamate, B10 histidine, B21 glutamate, and B29 lysine of molecule 1 and the B29 lysine of molecule 2. The main chain of B30 threonine of molecule 2 was found to occupy two distinct positions.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

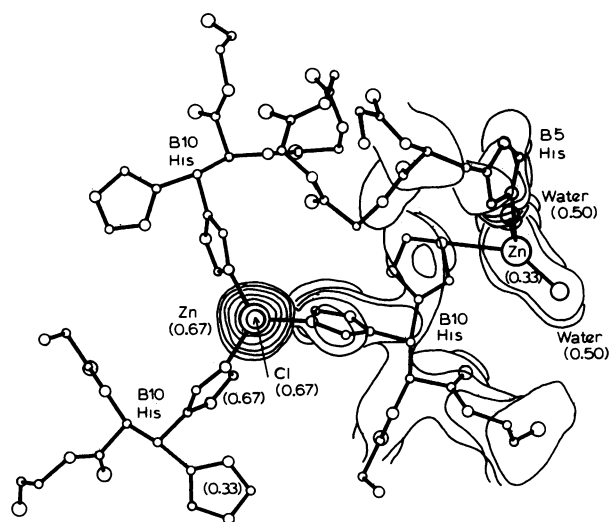


FIG. 1. Zinc coordination as viewed down the 3-fold axis in 4-zinc human insulin, molecule 1 along with the envelope at ≈ 2 standard deviations of the mean systematic ΔF electron density. Occupation factors of the partially occupied atoms and residues are noted in parentheses.

The refinement finally converged for the 826 protein atoms, 3 zinc ions, 1 chloride ion, and 246 water molecules to a residual of 0.173 for the 6548 data used in the refinement (8.0–1.85 Å resolution), 0.180 for all data with $F^2 > 0$ and 0.214 for all 8174 independent data. These residuals were calculated using the regularized model. The rms shift following regularization was 0.15 Å [MODELFIT parameters were $\sigma(\text{bond}) = 0.02$ Å; $\sigma(\text{angle}) = 3.0^\circ$] and the rms deviation of all bonds from their ideal values prior to regularization was 0.184 Å. While one might estimate that the error in the position of most atoms may be ≈ 0.2 Å, it is clear that in some cases, particularly where there are disordered side chains or high thermal motion, the error may be as large as 1.0 Å.

DISCUSSION

The 2-zinc and 4-zinc crystal forms were first prepared by Schlichtkrull (11), who named them according to the analytical ratio of the zinc ions to the insulin hexamer. The 4-zinc form was originally thought to contain four fully occupied zinc sites (5), but more recent results from the 4-zinc porcine analysis (14) and the present study have shown that there are one fully occupied and four partially occupied zinc sites, re-

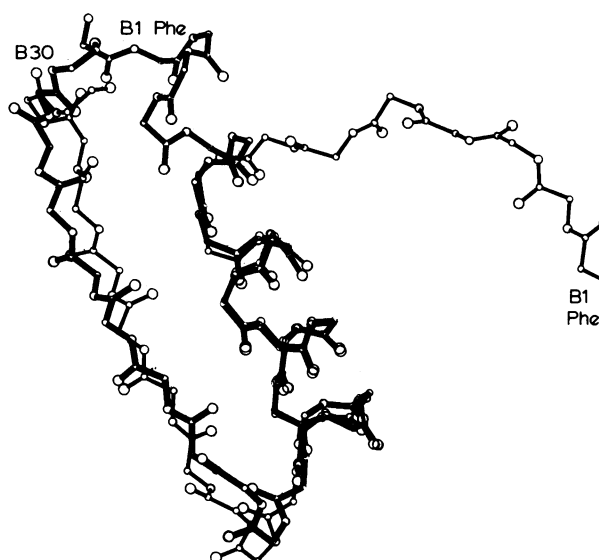


FIG. 2. Comparison of the conformational differences of the backbone atoms of the B chains of molecule 1 of 4-zinc human insulin to that of molecule 2 of 2-zinc porcine insulin. The 4-zinc structure is drawn with wide bonds, the 2-zinc structure is drawn with narrow bonds. Backbone atoms of residues B9–B18 were used in the optimization procedure and resulted in an average displacement of 0.27 Å.

sulting in a zinc content of < 3 per hexamer. Thus, the name 4-zinc insulin is used to describe a particular crystalline form of insulin rather than the zinc content of the hexamer.

Although the 2-zinc human and porcine insulin molecular structures are similar, the 2-zinc porcine insulin structure has been determined at higher resolution (1.5 Å) and will therefore be used as the standard against which all comparisons that follow will be made. A comparison of the structure of 4-zinc human insulin to that of the 2-zinc human and porcine insulins confirms the many substantial changes already noted in the 4-zinc porcine insulin studies. In 2-zinc porcine insulin, two crystallographically independent monomer units are related by a local 2-fold axis producing a dimer. The operation of a crystallographic 3-fold axis generates the hexamer from the dimer. Each of the two monomers, which have approximately the same conformation, consists of an A and a B chain linked by two disulfide bridges, A7–B7 and A20–B19; an additional intrachain disulfide bridge exists between A6 and A11. The conformation of the A chain consists

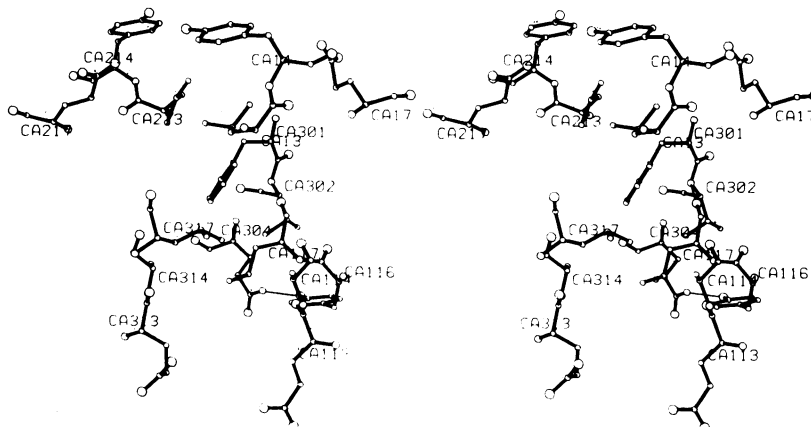


FIG. 3. Stereo drawing of the hexamer-forming residues in 4-zinc human insulin viewed perpendicular to the local 2-fold axis. The C^α carbon atoms are labeled as follows: A chain, molecule 1, 1–21; B chain, molecule 1, 101–130; A chain, molecule 2, 201–221; B chain, molecule 2, 301–330. Hydrogen bonds are shown as thin bonds.

of two antiparallel α -helices, A2–A8 and A13–A20, connected by a section of extended polypeptide chain. The conformation of the B chain consists of two extended sections of polypeptide chain connected by a central piece of α -helix, B9–B19. Each of the two zinc ions, which lie on a crystallographic 3-fold axis, is octahedrally coordinated by B10 histidine side chains of three crystallographically equivalent monomers and three water molecules.

The coordination geometry in the 4-zinc structure is more complex. The coordination of the zinc ion in molecule 2, octahedrally coordinated by three symmetry-related B10 histidines and three water molecules, remains unchanged. In molecule 1, however, there are disordered zinc sites, and the side chain of B10 histidine exists in two distinct orientations, illustrated in Fig. 1. A second zinc ion, which also lies on the 3-fold axis, is coordinated by one orientation of three symmetry-related B10 histidine side chains of molecule 1 and a chloride ion to give tetrahedral rather than octahedral geometry. This is a result of the rearrangement of the NH₂ terminus of the B chain (residues B1–B8) of molecule 1, which brings the B6 leucine side chain close to the B10 histidine so that there is no room for off-axial water molecules in an octahedral coordination. An additional zinc ion in a general position is tetrahedrally coordinated by the second orientation of the disordered B10 histidine side chain, a fully ordered B5 histidine, and two water molecules. Although the zinc ion site in molecule 2 appears to be fully occupied, it is clear from the refined values of the thermal parameters and the appearance of the systematic ΔF maps that the other zinc ion sites are not. Based on the values of the occupation factors that were assigned, a total zinc content of 2.67 ($0.67 + 3 \times 0.33 + 1.00$) is calculated per hexamer.

Throughout the refinement, considerable difficulty was encountered in correctly placing the disordered B10 histidines, the water molecules bound to the zinc ions, and the chloride ion. At the completion of the refinement, systematic ΔF maps were calculated (Fig. 1) and showed that, even though the geometry around the zinc ion in the general position is poor, all atoms lie in optimal positions within the electron density. Since the results of an x-ray analysis produce a static picture of all of the molecules in the unit cell, it is difficult to determine the precise nature and occupancy of the disordered zinc sites. The simplest scheme would require that only one of the two disordered sites in a given hexamer be fully occupied. Thus, 2/3 of the hexamers would possess a fully occupied axial zinc site and 1/3 would possess fully occupied off-axial sites. Other more complicated schemes are also possible and would require identical rotationally disordered hexamers.

The additional zinc ion sites in molecule 1 arise as a result of a dramatic change in conformation of residues 1–8 of the B chain from an extended conformation to an α -helical conformation. This produces a continuous α -helix from B1 phenylalanine through B19 cysteine and involves a shift of >30 Å of the C $^{\alpha}$ carbon of B1 phenylalanine. The B5 histidine undergoes a shift of nearly 13 Å and moves from the surface of the hexamer to the interior, where it coordinates the additional zinc ion. These differences in the B chain are illustrated in Fig. 2, where molecule 1 has been optimized by a least-squares fit to the reference molecule, molecule 2 of 2-zinc porcine insulin. The reference molecule was chosen because of its similarity to one of the molecules in 4-zinc insulin and to the molecule in hagfish insulin (3). Small differences in conformation are noted in the α -helical regions of both A chains.

In the 2-zinc structure, the B1 phenylalanine lies adjacent to the side chains of A13 leucine and A14 tyrosine, and contacts between these residues help to stabilize the formation of the hexamer. In the 4-zinc structure, the NH₂ terminus of the B chain of molecule 2 has also undergone a smaller but

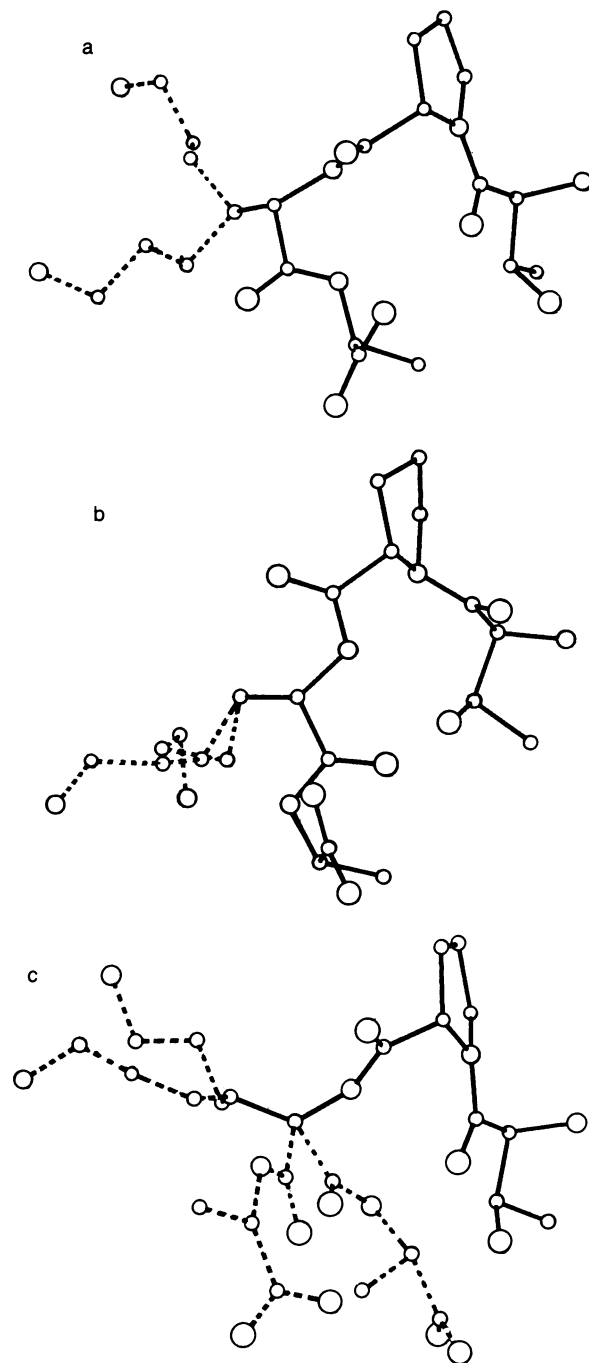


FIG. 4. Conformations of the COOH terminus of the B chains of (a) 2-zinc porcine insulin, molecule 2, (b) 4-zinc human insulin, molecule 1, and (c) 4-zinc human insulin, molecule 2. Fractionally occupied groups are illustrated by dashed bonds.

substantial change in conformation. For example, relative to the 2-zinc structure, the backbone atoms of the first four residues of the B chain of molecule 2 are displaced and range from 4.4 Å for the C $^{\alpha}$ carbon of B2 valine to 1.6 Å for the carbonyl group of B4 glutamine. In addition, the carboxyl group of the side chain of B4 glutamine has shifted 4.9 Å and the side chain of B1 phenylalanine is directed back into the hexamer, occupying a position between the A13 and A17 leucines. This is shown in Fig. 3. As a result of the shift of the backbone atoms of the first four residues and the B4 glutamine side chain, two additional hydrogen bonds result that are not possible in the 2-zinc structure. These occur between the nitrogen of B2 valine of molecule 2 and the carbonyl oxy-

gen of B17 leucine of molecule 1 and between the side chains of B4 glutamine of molecule 2 and B16 tyrosine of molecule 1 and help to further stabilize the hexamer.

Differences in conformation also exist at the COOH termini of both B chains. The side chain of B29 lysine is observed to exist in two different conformations in both molecules of the 4-zinc structure. In addition, the entire B30 threonine residue of molecule 2 is disordered and occupies two distinct positions, a result of a small rotation (25°) of the ϕ torsion angle and a much larger rotation (160°) of the ψ torsion angle of B29 lysine. These two residues lie on the surface of the hexamer and extend into the surrounding solvent region, and it is possible that they may possess more flexibility when the molecule is in the monomeric state. As was observed in 2-zinc human insulin, the conformations of the B29 lysine and B30 threonine residues are observed to be different in the two independent molecules of the 4-zinc structure. Furthermore, neither of these conformations, illustrated in Fig. 4, is the same as that observed in molecule 2 of the 2-zinc porcine insulin structure. This suggests that hydrogen bonding to the solvent is at least as important in establishing the conformation of the COOH terminus as is the identity of the B30 residue.

Five residues on the surface of each monomer interact with each other and stabilize the formation of a dimer. These so-called dimer-forming residues include B12 valine, B16 tyrosine, B24 phenylalanine, B25 phenylalanine, and B26 tyrosine. Four antiparallel hydrogen bonds between B24 phenylalanine and B26 tyrosine of molecule 1 and their counterparts in molecule 2 produce a pleated sheet structure and are very important for the stabilization of the dimer in the 2-zinc structure as well as in the 4-zinc human insulin structure. Again, as a result of the conformational change of the NH_2 terminus of the B chain of molecule 1, additional hydrogen

bonds are formed between the B5 histidine carbonyl oxygen atom of molecule 1 and the OH of the B16 tyrosine side chain of molecule 2 and between the side chains of B9 serine and B13 glutamate. These hydrogen bonds are not possible in the 2-zinc structure, and they help to further stabilize the dimer.

The conformational change in B1–B9 has been discussed in terms of its effect on dimer stabilization, hexamer stabilization, and zinc ion coordination. However, to consider the most profound effect, one must consider the entire hexamer. In the 2-zinc structure, the combination of the crystallographic 3-fold axis and the perpendicular 2-fold axis produce a symmetric hexamer. Each side of the hexamer consists of a broad shallow depression that contains the zinc ion and its three water molecules. The water molecules are exposed to the environment, where they can be approached by other ligands that displace the water, bind to the zinc, and initiate dissociation. This description applies to both faces of the 2-zinc structure, which are similar, as well as to a single face of the 4-zinc structure, illustrated in Fig. 5a.

As a result of the change in conformation of the NH_2 terminus of the B chain, this shallow depression is filled with three parallel α -helices, related by the 3-fold axis. The three water molecules have been replaced by a chloride ion, buried in the hexamer, and are shown in Fig. 5b. The net effect of this conformational change is to protect the chloride ion and, hence, the zinc ion from the environment. The only access to the chloride ion is by traversing a narrow tunnel 8 Å long.

It seems likely that, prior to reverting to the 2-zinc form, which can be accomplished in the crystalline state simply by reducing the chloride concentration in the solution surrounding the crystal, the first eight residues of the B chain must change from an α -helical conformation to an extended one. Since the side chain of A10 isoleucine will lie partially over

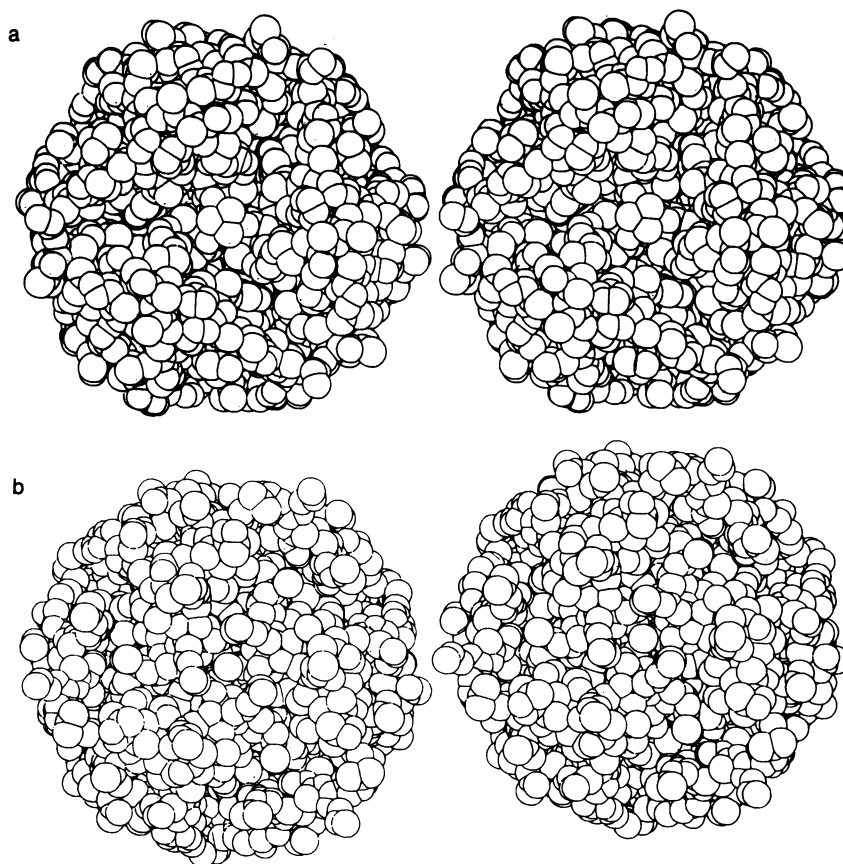


FIG. 5. Space-filling stereo drawing of the 4-zinc human insulin hexamer viewed down the 3-fold axis; (a) "2-zinc face"; (b) "4-zinc face."

the B chain in its extended conformation, some motion of the side chain and main chain atoms in the A chain must occur prior to changes in the B-chain conformation. At the same time, the additional zinc ion and its two water molecules, which lie close to the interface between the dimer units, must be displaced.

4-Zinc insulin has often been called a slow-acting insulin in diabetes therapy because it is slow to dissolve and take effect. Since the insulin hexamer is structurally very stable in the presence of zinc, the rate of dissolution is decreased when zinc cannot be easily removed from the hexamer. This stability can now be understood in terms of the conformational change of the B chain, stabilized by additional hydrogen bonds, which isolates one of the zinc ions from the environment.

The authors wish to express their gratitude to Lilly Research Laboratories for the generous gift of human insulin. This research was supported by grants from the Kroc Foundation, the James H. Cummings Foundation, and the Medical Research Council.

1. Dodson, E. J., Dodson, G. G., Hubbard, R. E. & Reynolds, C. D. (1983) *Biopolymers* **22**, 281–291.
2. Cutfield, J., Cutfield, S., Dodson, E., Dodson, G., Hodgkin, D. & Reynolds, C. (1981) *Hoppe-Seyler's Z. Physiol. Chem.* **362**, 755–761.
3. Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Emdin, S. F. & Reynolds, C. D. (1979) *J. Mol. Biol.* **132**, 85–100.
4. Smith, G. D., Duax, W. L., Dodson, E. J., Dodson, G. G., de Graaf, R. A. G. & Reynolds, C. D. (1982) *Acta Crystallogr. Sect. B* **38**, 3028–3032.
5. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D. & Mercola, D. (1976) *Nature (London)* **261**, 166–168.
6. Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Reynolds, C. D. & Valley, D. (1981) in *Structural Studies on Molecules of Biological Interest*, eds. Dodson, G., Glusker, J. & Sayre, D. (Oxford Univ. Press, Oxford), pp. 527–546.
7. Dodson, E. J., Dodson, G. G., Hodgkin, D. C. & Reynolds, C. D. (1979) *Can. J. Biochem.* **57**, 469–479.
8. Chawdhury, S. A., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Tolley, S. P., Blundell, T. L., Cleasby, A., Pitts, J. E., Tickle, I. J. & Wood, S. P. (1983) *Diabetologia* **25**, 460–464.
9. Chance, R. E., Kroeff, E. P., Hoffmann, J. A. & Frank, B. H. (1981) *Diabetes Care* **4**, 147–154.
10. Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Krasezewski, A., Itakura, K. & Riggs, A. D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 106–110.
11. Schlichtkrull, J. (1958) *Insulin Crystals* (Munksgaard, Copenhagen).
12. Blessing, R. H. & DeTitta, G. T. (1981) *Acta Crystallogr. Sect. A* **37**, C-135, abstr. 06.6-02.
13. North, A. C. T., Phillips, D. C. & Mathews, F. S. (1968) *Acta Crystallogr. Sect. A* **24**, 351–359.
14. Valley, D. (1981) Dissertation (University of York, York, England).
15. Agarwal, R. C. (1978) *Acta Crystallogr. Sect. A* **34**, 791–809.
16. Dodson, E. J., Isaacs, N. W. & Rollett, J. S. (1976) *Acta Crystallogr. Sect. A* **32**, 311–315.