Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of an insulin dimer in an orthorhombic crystal: the structure analysis of a human insulin mutant (*B*9 Ser→Glu)

Zhi-Ping Yao,[a]† Zong-Hao Zeng,[a]† Hong-Min Li,[a] Ying Zhang,[a] You-Min Feng[b] and Da-Cheng Wang[a]*

[a]Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, and [b]National Laboratory of Molecular Biology, Shanghai Institute of Biochemistry, Shanghai 200031, People's Republic of China

† These authors made equal contributions to this work.

Correspondence e-mail: wdc@pewdc.ibp.ac.cn

Received 16 February 1999
Accepted 25 June 1999

PDB Reference: Human insulin mutant (Ser*B*9Glu), 1b9e.

The structure of human insulin mutant *B*9 (Ser→Glu) was determined by an X-ray crystallographic method at 2.5 Å resolution with an *R* factor of 0.165 under non-crystallographic restraints. The crystals were grown at low pH (<3.8) and belong to the orthorhombic $P2_12_12_1$ space group with unit-cell dimensions $a = 44.54$, $b = 46.40$, $c = 51.85$ Å and one dimer per asymmetric unit without further aggregation. The structure in this crystal form can be regarded as a model for a discrete insulin dimer and displays the following features compared with the structure of 2Zn insulin. (i) The overall dimer is expanded and more symmetric. The two *A* chains are about 2 Å more distant from each other, while the two *B* chains are about 0.8 Å further apart. Both monomers are more similar to molecule 1 than molecule 2 of the 2Zn insulin dimer. (ii) The dimer structure is stabilized by protonation and neutralization of the carboxyl groups at lower pH and, in addition, by formation of a hydrogen-bond network among the side chains of residues Glu*B*9, His*B*13 and His*B*10 on the dimer-forming surface of both monomers, resulting from a structural rearrangement. (iii) The *B*-chain amino-terminal segment is in an open state (O state), *i.e.* a state different from the well known R and T states found in the insulin hexamer. In the O state, the *B*-chain N-terminal segment is in an extended conformation and is detached from the rest of the molecule. This conformational state has also been observed in the monomeric crystal structure of despentapeptide (*B*26–*B*30) and desheptapeptide (*B*24–*B*30) insulin, as well as in the solution structure of an engineered insulin monomer. It suggests that the O state may be the characteristic conformation of insulin in lower aggregation forms and may be relevant to the formation of insulin fibrils. In addition, based on the crystallization process, the smallest possible building blocks of insulin crystal are also discussed.

1. Introduction

Being used in the treatment of diabetes, insulin is one of the most important protein hormones, stimulating the transport of glucose from blood into cells, as well as related metabolic and growth-promoting effects. Both *in vivo* or *in vitro*, insulin exhibits complicated self-association behaviour from monomer to dimer to hexamer. The self-association of insulin is not only closely related to its storage, transportation and biological function *in vivo*, but is also closely associated with its clinical potency. Insulin is stored in the pancreas as Zn^{2+}-containing hexamers; once diluted in the blood stream, the insulin assembly dissociates rapidly through dimers to biologically active monomers (Blundell *et al.*, 1972). In

medical use, preparations of prolonged-acting or fast-acting insulins are produced by methods which affect the insulin molecules' assembly properties, especially by controlling the dissociation rate of insulin assemblies (Owens, 1986; Markussen, Diers *et al.*, 1987; Markussen, Hougaard *et al.*, 1987; Brange *et al.*, 1988).

One of the methods for controlling dissociation rate is by mutation. The human insulin monomer consists of a 21 amino-acid *A* chain and a 30 amino-acid *B* chain linked by two disulfide bridges, while the insulin dimer is formed from two monomers *via* an antiparallel β-sheet consisting of the two *B*-chain C-terminal residues and by interactions between the two *B*-chain helices. Residue Ser*B*9 is located on one side of the dimer-forming surface and contacts the adjacent residue Glu*B*13 from another monomer on the opposite side. Conversion of serine at position *B*9 to aspartic acid will produce four negatively charged side chains on the dimer-forming surface which are in close contact with each other, consequently reducing the self-association tendency of insulin and destabilizing the dimer (Brange *et al.*, 1988). Experiments on insulin mutant *B*9 Ser→Asp (*B*9D HI) demonstrated that the mutant was essentially monomeric at millimolar concentration and after subcutaneous injection absorption was much faster than for native insulin. At extremely low pH, 1.8–1.9, when the carboxyl groups are protonated and the negative charges are neutralized, *B*9D HI also forms a well defined dimer in solution. NMR study on *B*9D HI indicated that in solution the *B*9D HI dimer retained the major structural features of an assembled dimer in crystal form (Jorgensen *et al.*, 1992).

For the same purpose of producing fast-acting insulin, human insulin mutant Glu*B*9 (*B*9E HI) was prepared by substitution of a glutamic acid residue for neutral serine at position *B*9 by site-directed mutagenesis (Liu *et al.*, 1996). This mutant, like *B*9D HI, was found to diminish the tendency for self-association and was found to exist as monomers in solution at millimolar concentrations. It exhibited 21% of the receptor-binding affinity and 40% of the biological potency of pig insulin. At acidic pH, 3.8–3.9, this mutant also forms a well defined dimer and crystallizes in the orthorhombic $P2_12_12_1$ space group. This new crystal form displayed a structure of discrete insulin dimers; the dimers did not aggregate further into hexamers. The crystal structure of hexameric insulin has been extensively investigated (Baker *et al.*, 1988; Peking Insulin Structure Research Group, 1974; Bentley *et al.*, 1976). Symmetric insulin dimer structures, where the two monomers are related by a twofold crystallographic symmetry operation, have been reported in either tetragonal (Balschmidt *et al.*, 1991) or cubic crystals (Badger *et al.*, 1991). The monomeric insulin analogue, despentapeptide (*B*26–*B*30) insulin (DPI), has also been analyzed (Liang *et al.*, 1985; Bi *et al.*, 1984). Detailed evaluation of the three-dimensional structure of the *B*9E HI dimer will provide a structure of a discrete insulin dimer and comparison with insulin structures in other crystal forms will be of great interest. The results are presented and discussed in this report.

2. Materials and methods

2.1. Sample and crystallization

The *B*9E HI sample provided by the Shanghai Institute of Biochemistry, Chinese Academy of Sciences was prepared by site-directed mutagenesis (Liu *et al.*, 1996). Crystals of *B*9E HI were obtained by the sitting-drop method. Since the mutated residue is located at the monomer interface in the insulin dimer and therefore severely affects the aggregation of insulin molecules in solution, crystallization was only achieved with difficulty. After a systematic search, single crystals were grown from a solution containing 4 µl protein solution (20 mg ml^{-1} *B*9E HI dissolved in 0.03 *M* HCl) and 20 µl buffer solution [0.1 *M* ammonium citrate, 0.12% $CrCl_2$, 10%(*v*/*v*) acetone and 4% dimethylformamide]. The crystal shape and nucleation behaviour varies greatly with the pH value. Under the optimized pH value of 3.8, the number of nuclei could be reduced to 1–2. Application of seeding helped the crystals grow to dimensions of 0.2 × 0.2 × 0.1 mm. The optimized crystal is orthorhombic, with space group $P2_12_12_1$, unit-cell parameters a = 44.54, b = 46.40, c = 51.85 Å and two molecules in the asymmetric unit.

2.2. Diffraction data collection

The diffraction data were collected at room temperature using a MAR Research image-plate system with monochromated Cu $K\alpha$ radiation (λ = 1.5418 Å) generated by a Rigaku rotating-anode generator operating at 2.5 kW. 3338 unique reflections with intensity greater than zero were obtained from a total of 13549 reflections to a resolution of 2.5 Å. The observed reflections are 83.5% complete. The completeness in the 2.54–2.50 Å shell is 87.8%. The overall merging R factor on intensity to a resolution of 2.5 Å is 8.54%, with a multiplicity of 3.83. Since the mutant possesses a strong tendency to monomerization, crystallization was very difficult, and higher resolution and more complete data could not be obtained on the general X-ray source. Great efforts were made to crystallize *B*9E HI at a higher pH where it might form monomers, but all failed.

2.3. Structure determination and refinement

The molecular-replacement method using the program *X-PLOR* (Brunger, 1992) was used to determine the orientation and location of the *B*9E HI molecules in the unit cell. The refined structure of 2Zn pig insulin dimer at a resolution of 1.5 Å (Protein Data Bank, code 4ins) was used as a search model. After Patterson correlation refinement against data with $F/\sigma(F) > 3\sigma$ in the resolution range 6.5–3.0 Å, the highest peak appeared from the 15th rotation peak, with a correlation coefficient of 0.173. The resulting crystal model for *B*9E HI after the translation search showed no overlap between symmetry-related molecules, except for the *B*-chain N-terminal residues.

The program *X-PLOR* (Brunger, 1992) was used for structural refinement and *TURBO* (Roussel & Cambillau, 1991) and *O* (Jones *et al.*, 1991) were used for model

Table 1
Refinement statistics for the *B*9E HI structural model.

Resolution limits (Å)	6.5–2.5
Final *R* factor	0.165
R_{free}	0.264
No. of reflections used	3170
No. of non-H atoms	861
No. of protein atoms (non-H)	816
No. of water molecules	45
Non-crystallographic symmetry restraints	*A*1–*A*21, *B*3–*B*26, exclude side chains of *A*4, *A*17, *B*22 and *B*25
R.m.s. deviations from standard geometry	
Bond lengths (Å)	0.013
Bond angles (°)	1.637
Dihedral angles (°)	24.487
Improper angles (°)	1.422
Average *B* factors (Å^2)	24.35
Protein	24.06
Main chain	22.30
Side chain	25.79
Solvent	29.48
Ramachandran plot: No. of residues in	
Most-favoured region (%)	91.9
Additional allowed region (%)	8.1

rebuilding. The *R* factor decreased from an initial value of 0.514 to 0.216 after one cycle of simulated-annealing refinement in the resolution range 20–3.0 Å. The R_{free} method, with 5% data in the testing set, and non-crystallographic symmetry (NCS) restraints were applied in further refinements in the resolution range 6.5–2.5 Å. NCS restraints on residues *A*1–*A*21 and *B*3–*B*28 reduced R_{free} from 0.374 to 0.344. The proper NCS restraints were obtained by checking R_{free} and the electron density. After repeated rebuilding of the crystal model with adjustment of the NCS restraints, the model with 38 water molecules and NCS (Table 1) gave an *R* factor of 0.168 and an R_{free} of 0.264.

The final refinement was performed over the whole data set in the resolution range 6.5–2.5 Å with $F/\sigma(F) > 2$ (3170 reflections) while retaining the optimized NCS restraints. Seven more waters were added to the model; all the waters are within 4 Å of the nearest protein atom. The final *R* factor was 0.165. Other statistics are listed in Table 1.

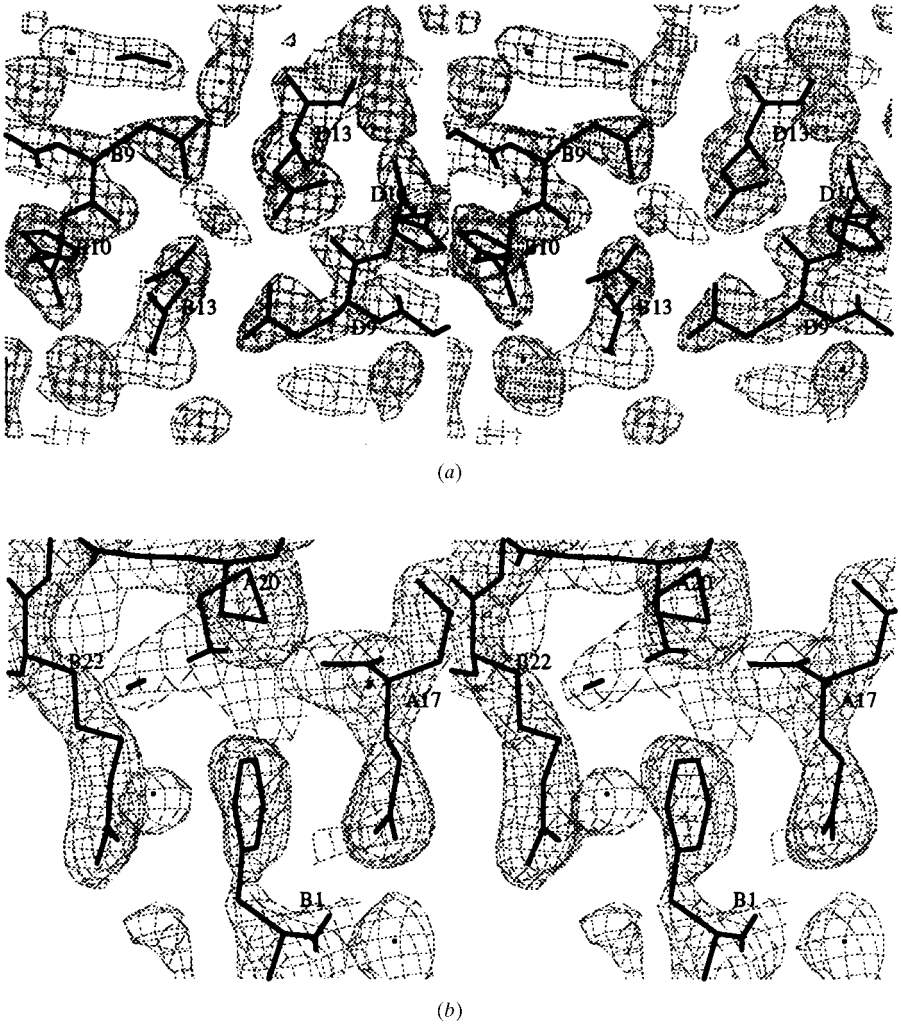

Figure 1
$2F_o - F_c$ electron-density maps with contour level 1.0σ at (*a*) the dimer-forming surface near Glu*B*9, Glu*B*10 and Glu*B*13 of both monomers; (*b*) the vicinity of Phe*B*1, where Glu*A*17 and Arg*B*22 are from molecule 1 of another symmetry-related dimer (a second neighbour). This figure was produced using *O* (Jones *et al.*, 1991).

3. Results and discussion

3.1. Quality of the model

The quality of the refined model is demonstrated by the statistics in Table 1. (Hereafter, the *A* or *B* chains of molecule 2 in a dimer will be referred to as the *C* or *D* chains, respectively.) All residues, except *B*29–*B*30 in both molecules, are built into proper density with no ambiguity, and the positions of the long side chains of the substituted Glu at *B*9 in both molecules are well defined (Fig. 1). The quality of the electron density at residue *D*1 is less definite than the density at *B*1. The model showed no residues located in the generously allowed or disallowed regions (Table 1), as evaluated by *PROCHECK* (Laskowski *et al.*, 1993).

3.2. Packing of *B*9E HI dimers

The crystal is packed with dimers, which do not form hexamers, as building blocks (Fig. 2). The dimer–dimer interaction in this orthorhombic crystal has two characteristics. Firstly, molecule 1 and molecule 2, denoting the two non-crystallographic symmetry-related (a pseudo-twofold axis approximately parallel to the crystal's *b* axis) monomers in a dimer, are arranged alternately in the whole crystal. In other words, the direct neighbours of molecule 1 are all molecule 2, while the direct neighbours of molecule 2 are all molecule 1 (Fig. 2). Secondly, the interactions between

Table 2
Monomer surface areas buried in monomer–monomer interfaces (Å).

Surface areas are calculated per atom by *X-PLOR* as accessible surface areas with a detecting radius of 1.4 Å. The surface areas provided by N or O atoms are taken as hydrophilic and that provided by other atoms as hydrophobic. The surface-area differences between the sum of the two isolated monomers and that of the dimer are taken as buried surface areas.

Interface	Hydrophilic	Hydrophobic	Ratio†	Total
Dimer-forming	329	1028	1/3.12	1357
A	317	950	1/2.99	1267
B	502	628	1/1.25	1131

† Ratio = hydrophilic area/hydrophobic area.

dimers are dominated by interactions on two monomer–monomer interfaces. The two translationally unrelated direct neighbours of a dimer are created by the twofold skew axis parallel to the *a* or the *c* axis and the corresponding dimer–dimer interfaces are denoted as interface *L* and interface *M*, respectively (Fig. 3). Owing to the non-crystallographic symmetry, these dimer–dimer interfaces are in fact monomer–monomer interfaces. Considering the dimer-forming surface, each molecule 1 (or 2) contacts with three molecules 2 (or 1) to form three monomer–monomer interfaces. The monomer surface areas buried in the three monomer–monomer interfaces are comparable (Table 2).

Only a few residues of *B*9E HI, including *A*6, *B*11, *B*15 and *B*19, are completely buried in the monomer (to be as objective as possible, a residue with an accessible surface area of less than 10 Å^2 is regarded as completely buried), and four more residues, *B*12, *B*23, B24 and *B*26, are completely buried in the dimer. The remaining residues can take part in interactions between dimers. On the monomer–monomer interface *L*, residues *A*10–*A*17, (*B*2), *B*3–*B*4, *B*6, *B*10, *B*14 and *B*17–*B*18 have buried surface areas of more that 10 Å^2, while on interface *M*, residues *A*1–*A*2, *A*4–*A*5, *A*15, *A*18–*A*19, (*A*20–*A*21), *B*25, *B*27, *B*29 and (*B*30) have buried surface areas more than 10 Å^2 (with the exceptions of those in parentheses, the residues are from both monomers). The interface *M* is more hydrophilic than the dimer-forming surface and the interface *L*. In fact, the residues buried in interface *M* are those exposed on the hexamer surface of 2Zn insulin. The hydrogen bonds between dimers are listed in Table 3.

Besides the major interfaces *L* and *M*, there is a minor interaction site *N*. The *B*-chain N-terminal residues (*B*1–*B*4) do not interact with the remainder of the monomer; instead, they stretch beyond the direct and second neighbour to reach the third neighbour. The *D*-chain N-terminal residues (*D*1–*D*4) in a fourth dimer stretch across the same area, but from the opposite direction in an antiparallel manner (Figs. 2 and 3). This special arrangement of the *B*-chain N-terminal residues brings together four monomers from four different dimers at the site *N*. The benzene ring of Phe*B*1 points into the space between the side chains of Glu*A*17 and Arg*B*22 of a third monomer (Figs. 1 and 2).

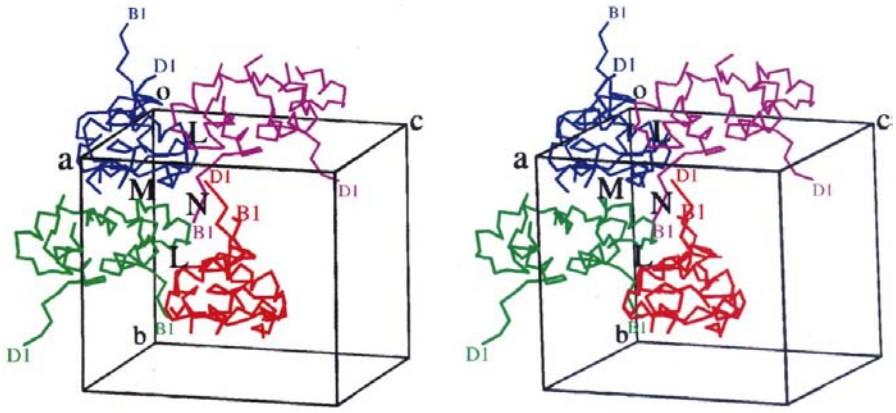

Figure 2
The four dimers packed in an unit cell. The *B*-chain N-termini of the two alternatively distributed monomers are denoted as *B*1 and *D*1, respectively. The dimers are numbered and coloured: 1, blue; 2, purple; 3, green; 4, red. Dimers 1 and 2 are related by the skew axis parallel to the *c* axis, while 1 and 3 are related by the skew axis parallel to the *a* axis. The non-crystallographic twofold axis is approximately parallel to the crystal's *b* axis. *L* and *M* denote the interfaces between neighbouring dimers and the two directly neighbouring dimers. *N* denotes the minor interaction site where the four dimers interact with one another. Figs. 2, 3, 5, 6 and 7 were produced using *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Bacon & Anderson, 1988).

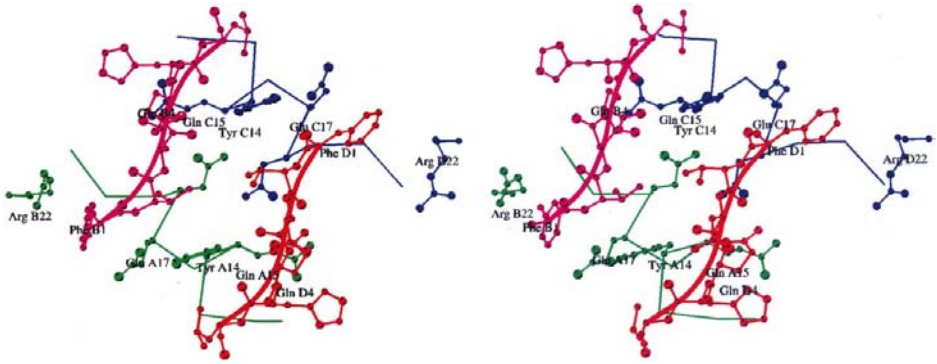

Figure 3
Stereoview of *B*-chain N-terminal residues in crystal packing. Atoms displayed include those of residues *B*1–*B*6 of dimer 4 (red) and *D*1–*D*6 of 2 (purple); C$^\alpha$ atoms in residues *A*11–*A*21 of dimer 3 (green) and *C*11–*C*21 of dimer 1 (blue) and their side chains also take part in interaction with *B*-chain N-terminal residues. The view is roughly towards the crystal's *c* axis.

3.3. Main structural features of *B*9E HI dimer

Compared with the 2Zn insulin dimer in a hexamer (Baker *et al.*,

1988; Chang *et al.*, 1986), the *B*9E HI dimer, which retains the overall fold and major secondary and tertiary structures, possesses the following new structural features.

Firstly, the conformational change of the *B*-chain N-terminal residues move the *B*1 C^{α} atom to a new position 15 Å distant from its position in 2Zn insulin (Fig. 4). *B*1 must jut out as far as possible from the remainder of the monomer while keeping the disulfide bond between *B*7 and *A*7 unbroken, therefore placing the *B*-chain N-terminal residues in an extended conformation roughly perpendicular to the *A*-chain middle segment (Figs. 5 and 9*c*).

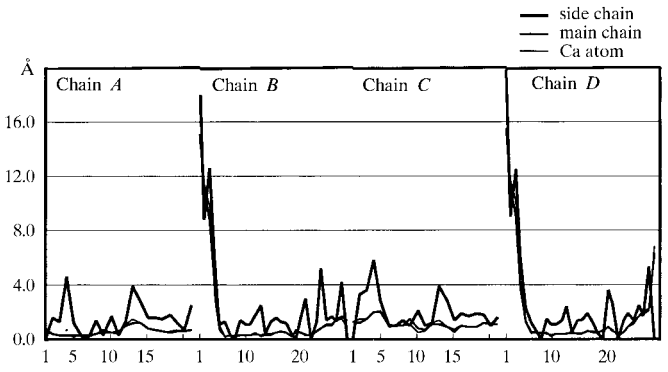

Figure 4
Root-mean-square movements per residue of *B*9E HI dimer relative to 2Zn pig insulin dimer after least-squares fitting.

Secondly, the *B*9E HI dimer is more symmetric and both of the two monomers are more similar to molecule 1 of the 2Zn insulin dimer. The root-mean-square differences in C^{α} atom coordinates are shown in Table 4. In the 2Zn insulin dimer, residues *A*1–*A*5 and the side chain of *B*25 are two major asymmetric parts. In the *B*9E HI dimer, residues *A*1–*A*5 in both monomers adopt the less regular helical conformation of 2Zn insulin dimer molecule 1 (Fig. 5), while the two *B*25 side chains remain asymmetric. The side chain of *B*25 of molecule 1 in the *B*9E dimer, unlike that in the 2Zn insulin dimer, adopts a conformation which brings the benzene ring along the peptide strand towards *B*26 (Fig. 6). This fact confirms the observation that the *B*25 side chain is flexible (Baker *et al.*, 1988).

Finally, there are some overall structural changes. As confirmed by the distances between corresponding segments in molecule 1 and molecule 2 (Table 5), Fig. 5 clearly shows the expansion of the *B*9E HI dimer relative to the 2Zn insulin dimer. At the centre of the dimer, the two strands of the *B*-chain β-sheet became a little closer (Table 6), while the distance between the two *B*-chain α-helices increases slightly; the distance between the two *A* chains increases to a larger extent (from 25.32 to 27.37 Å, an increase of 2 Å). This expansion demonstrates a certain degree of flexibility of the overall structure of the insulin dimer in different crystal forms.

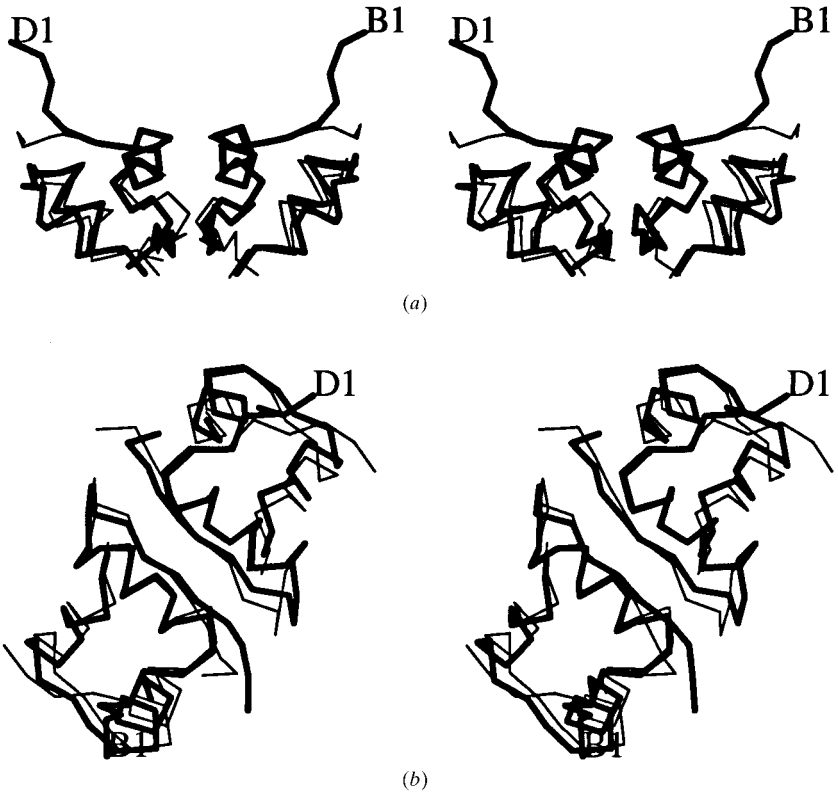

Figure 5
Stereoview of the B9E HI dimer (heavy lines) overlapped with the 2Zn insulin dimer (light lines). Only the C^{α} atom traces are shown. The view is from the direction looking along the dimer-forming surface and (*a*) along and (*b*) perpendicular to the two antiparallel β-strands formed from the two *B*-chain C-terminal segments. From these directions it is clear to see the conformational change of the *B*-chain N-terminal segment and easy to observe that the two monomers, especially the two *A* chains, in the *B*9E HI dimer are becoming more distant from each other.

3.4. Stability of the *B*9E HI dimer

In the native 2Zn insulin structure, the dimers are stabilized by the hydrogen bonds on the antiparallel β-sheet formed by C-terminal residues from the two monomers, and are further stabilized by formation of hexamers, which brings three dimers into one hexamer so as to bury more hydrophobic surface areas in the hexamer's interior. The ratio of hydrophobic surface areas to hydrophilic surface areas falls from 1.93 for the dimer to 1.24 for the hexamer, showing the hydrophobic character of the dimer–dimer interactions in hexamers. Coordination of the six *B*10 side chains to the two zinc ions on the threefold symmetry axis and hydrogen bonds formed between neighbouring dimers make contributions to the stability of dimers while also stabilizing the hexamer (Bentley *et al.*, 1992).

The mutation of Ser to Glu at *B*9 in *B*9E HI brings four glutamic acid residues (*i.e.* the two *B*9 and two *B*13 residues from the two monomers) into a narrow space of 6–7 Å spanning both sides of the dimer-forming

Table 3
Hydrogen bonds between dimers (Å).

Interface	Atom 1		Atom 2		Distance
L	Dimer 2 Mol. 1	Ser*A*12 OG	Dimer 1 Mol. 2	Gln*D*4 NE2	3.19
		Gln*B*4 NE2		Ser*C*12 OG	3.04
		Gln*A*15 NE2		Gln*D*4 OE1	3.30
		Gln*B*4 OE1		Gln*C*15 NE2	3.35
M	Dimer 3 Mol. 1	Gln*A*5 OE1	Dimer 1 Mol. 2	Asn*C*18 O	3.34
		Gln*A*5 NE2		Asn*C*18 O	3.07
		Gln*A*15 O		Asn*C*18 ND2	3.00
		Asn*A*18 ND2		Tyr*C*14 O	3.03
		Asn*A*18 O		Gln*C*5 NE2	3.32
		Cys*A*20 O		Gln*C*5 NE2	3.41
		Pro*B*28 O		Lys*D*29 NZ	3.44
		Lys*B*29 O		Thr*D*30 OG1	2.80
N	Dimer 2 Mol. 1	Phe*B*2 N	Dimer 3 Mol. 1	Glu*A*17 OE1	3.14
	Dimer 4 Mol. 2	Phe*D*1 N	Dimer 1 Mol. 2	Glu*C*17 OE1	2.67

Table 4
R.m.s. differences between individual monomers in *B*9E HI and 2Zn insulin dimers (Å).

Only residues *A*1–*A*21 and *B*6–*B*26 were included in calculations.

	2Zn insulin					
	Molecule 1			Molecule 2		
	C^{α} atom	Side chain	All atoms	C^{α} atom	Side chain	All atoms
*B*9E HI molecule 1	0.654	2.099	1.470	0.990	2.193	1.630
*B*9E HI molecule 2	0.634	2.183	1.517	0.987	2.292	1.686

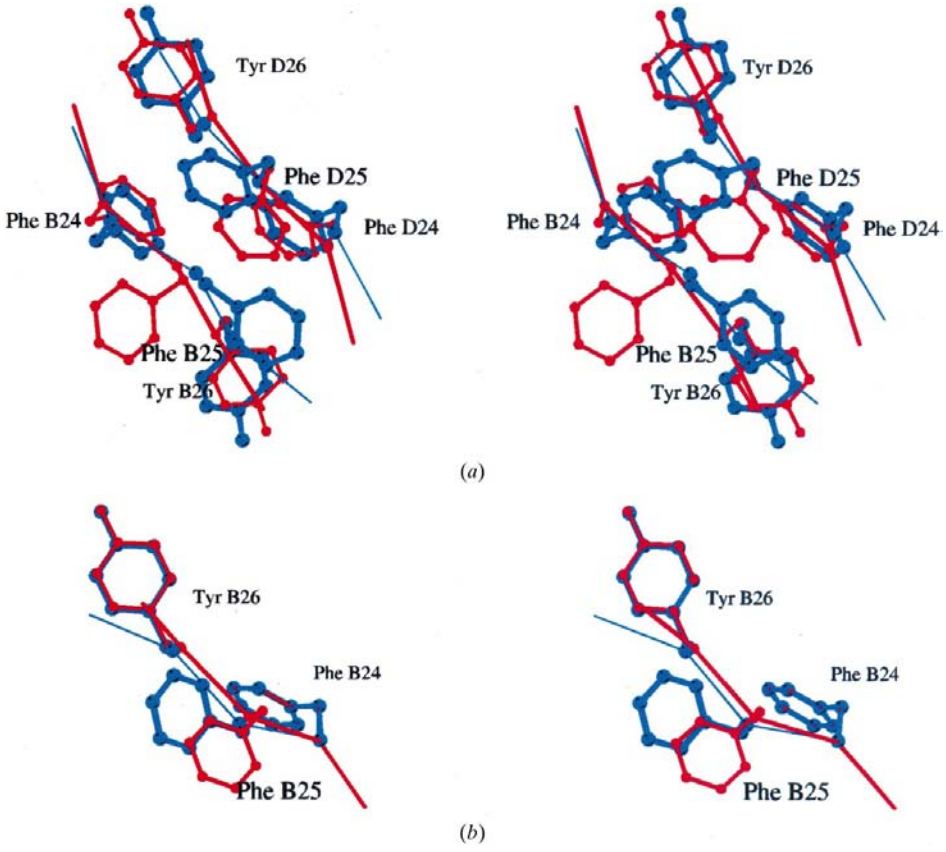

Figure 6
Conformation of *B*25 side chain. (*a*) Comparison with 2Zn insulin. Heavy and light lines represent 2Zn insulin and *B*9E HI, respectively. Residues of *B*9E HI are labelled. (*b*) Comparison of molecule 1 (light) with molecule 2 (bold).

surface. The electronic charges on the side-chain carboxyl groups are great obstacles to dimer formation. However, at low pH, the charges are neutralized or partly neutralized. Furthermore, the slightly increased distance between the two *B*-chain α-helices can accept larger side chains. A concerted rearrangement in the vicinity of *B*9 and *B*13 allows the structure to arrive at a new equilibrium (Fig. 7). The carboxyl group on Glu*B*13, in addition to forming a hydrogen bond with its partner in a dimer, now has the potential to form another hydrogen bond with His*B*10, since the two histidine residues no longer need to interact with the zinc ion. Leaving their old positions, the side chains of *B*10 and *B*13 turn towards each other, moving *B*10 ND1 and *B*13 OE1 to a distance of 3.11 Å (2.99 Å in molecule 2) from each other, while keeping *B*13 OE2 and 13 OE2 within hydrogen-bonding distance (3.20 Å). The carboxyl group on Glu*B*9 stretches out toward Glu*B*13, moving *B*9 OE1 to a distance of 2.93 Å (2.87 Å in molecule 2) away from the atom on *B*13 OE2. In this way, a new hydrogen-bond network is formed (Fig. 7). This hydrogen-bond network is a distinguishing characteristic of the stable *B*9E HI dimer at low pH.

3.5. The open state of *B*-chain N-terminal residues

The protrusion of the *B*-chain N-terminal residues from the monomer (and dimer) surface (Fig. 5) breaks down the monomer's compactness. This open state for *B*-chain N-terminal residues can be seen in the monomeric crystal structures of DPI (Brange *et al.*, 1997) and desheptapeptide (*B*24–*B*30) insulin (DHPI; Bao *et al.*, 1997; Fig. 8*a*), as well as the solution NMR structure of an engineered insulin monomer (Olsen *et al.*, 1996; Fig. 8*b*). This conformation is different from the two

known conformations of insulin, *i.e.* the R and T states (Smith *et al.*, 1984; Smith & Dodson, 1992; Derewenda *et al.*, 1989; Brader & Dunn, 1991). In the R state, the *B*-chain N-terminal residues form an α-helix merged into the *B*-chain middle helix (Fig. 9*a*). In the T state, the *B*-chain N-terminal residues are in a β-strand-like extended conformation and are roughly anti-parallel to the *A*-chain middle extended segment, with hydrogen bonds between *B*4 and *A*11 and between *B*6 and *A*6 to keep the whole monomer in a compact form (Fig. 9*b*). In the new conformational state, the *B*-chain N-terminal residues are also in an extended conformation, but instead of being parallel to the A-chain middle segment, they are perpendicular to it. This open state of the *B*-chain N-terminal residues is stabilized by extensive intermolecular contacts with three neighbouring dimers, including the *B*-chain N-terminal residues of another dimer from an opposite direction (Fig. 3), a situation very similar to that in the DPI crystal (Brange *et al.*, 1997). This new state will be called the open state and abbreviated as O state, as opposed to the T and R states for the *B*-chain N-terminal residues (Fig. 9*c*).

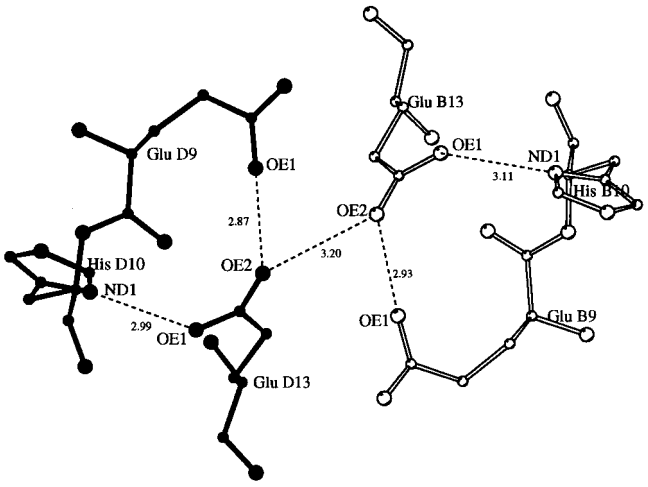

Figure 7
Hydrogen-bond network among side chains of *B*9, *B*10 and *B*13 from both monomers at the dimer-forming surface. Distances are in Å.

Table 5
Distance between peptide segments from molecule 1 and molecule 2 in a dimer (Å).

Distance was calculated as the distance between C^{α} atom geometrical centres of the two segments in molecules 1 and 2.

Peptide segment	*B*9E HI	2Zn insulin	Difference
*B*24–*B*26	4.87	4.98	−0.11
*B*9–*B*19	11.45	10.80	0.65
*B*9–*B*13	9.79	9.11	0.68
*B*6–*B*26	10.38	9.56	0.82
*A*1–*A*21	27.37	25.32	2.05

The crystal structure of DPI (monomeric and monoclinic), DHPI (monomeric and orthorhombic) and *B*9E HI (dimeric and orthorhombic) are all in lower aggregation forms and their *B*-chain N-terminal residues are in the O state. The NMR structure of an engineered insulin monomer at neutral pH displayed a destabilized *B*-chain N-terminal segment which did not pack across the remaining part of the molecule (Olsen *et al.*, 1996). The averaged conformation of the *B*-chain N-terminal residues is also very similar to the O state (Fig. 8*b*).

Based on the high-resolution X-ray structure of DPI, Brange *et al.* (1997) derived a model of insulin fibrils which occasionally presented problems in storage and use of insulin. In this model, the segment *B*1–*B*5 plays a crucial role in the formation of fibrils. There are four peptide amide–carbonyl bonds made between the two strands from symmetry-related molecules. The conformation of this segment *B*1–*B*5 is very similar to the O state described above (Fig. 8*a*). Therefore, we infer that O conformation state of the *B*-chain N-terminal segment may be relevant to the formation of insulin fibrils.

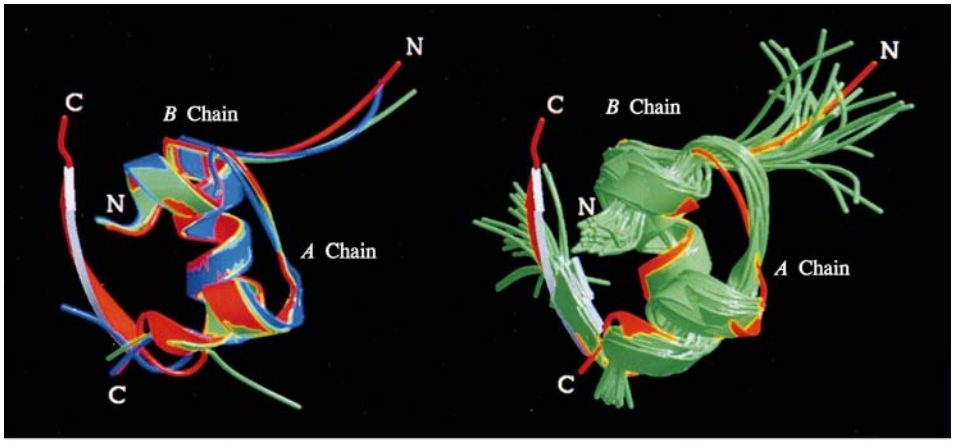

Figure 8
Conformational comparison of *B*-chain N-terminal residues, showing the O state (*a*) in the crystal and (*b*) in solution. (*a*) Overlapped structures of *B*9E HI (red), DPI (blue), DHPI (green). (*b*) The 25 NMR structures (green) of engineered (*B*1, *B*10, *B*16, *B*27)Glu, des *B*30-insulin overlapped on B9E HI (red). The coordinates of DPI, DHPI and (*B*1, *B*10, *B*16, *B*27)Glu, des *B*30-insulin are taken from the Protein Data Bank, with codes 1pid, 1dei and 1hui, respectively.

3.6. The smallest building blocks of insulin crystals

Insulin, either in its native or modified forms, is one of the most extensively crystallographically investigated proteins. In most cases, the crystals have been found to be packed with hexamers as building blocks. So far, no insulin crystals have been found built up of integral insulin molecules with a complete *B*-chain carboxyl segment and not built up of dimers. Only when dramatic modifications are performed, *e.g.* in the case of DPI or DHI, where the *B*-chain C-terminal residues have been removed, can

Table 6
Hydrogen-bond distances in the *B*-chain β-sheet (Å).

	*B*24 N–*D*26 O	*B*24 O–*D*26 N	*B*26 N–*D*24 O	*B*26 O–*D*24 N
2Zn insulin	2.94	2.89	2.89	3.09
*B*9E HI	2.98	2.85	2.60	2.69

monomeric insulin crystals be found (Bi *et al.*, 1984; Liang *et al.*, 1985; Bao *et al.*, 1997; Brange *et al.*, 1997). A natural conclusion from this is that dimerization is unavoidable in the crystallization process when the complete *B*-chain carboxyl segment is present.

In the case of *B*9E HI, the mutation makes aggregation in conventional manner much more difficult, and provides the possibility of growing a monomeric crystal. The failure of attempts to date suggests that the conformational flexibility of the *B*-chain C-terminal residues is a great obstacle to crystallization. This flexibility has been confirmed by the solution structure of an insulin mutant (Hua *et al.*, 1991). The formation of an antiparallel β-sheet on the dimer-forming surface produces a strong tendency to dimer formation while stabilizing the conformation of the *B*-chain C-terminal residues. Therefore, dimer formation seems to be important in crystallization or, in other words, dimers may be the smallest building blocks in insulin crystals when a complete native C-terminal segment is present. The manner of further aggregation depends on complicated intra-dimer interactions and interactions mediated by solvents.

In the crystallization of *B*9E HI, the low pH, which is below the p*K* value for protonation of the carboxyl group of glutamic acid, has a dual effect on the self-association of the mutant. In addition to restoring the dimer-forming ability by neutralizing the negative charges on glutamic acid, the low pH can also reduce the tendency towards hexamer formation by protonating the histidine imidazole ring, which would otherwise be a ligand for the zinc ion in the hexamer. As a result, the low-pH condition brings about orthorhombic crystals where the dimers do not aggregate further into hexamers. It is interesting that in the presence of zinc ion and phenol, hexamers of *B*9D HI have also been crystallized as rhombohedral crystals in our laboratory, though with very poor diffraction ability.

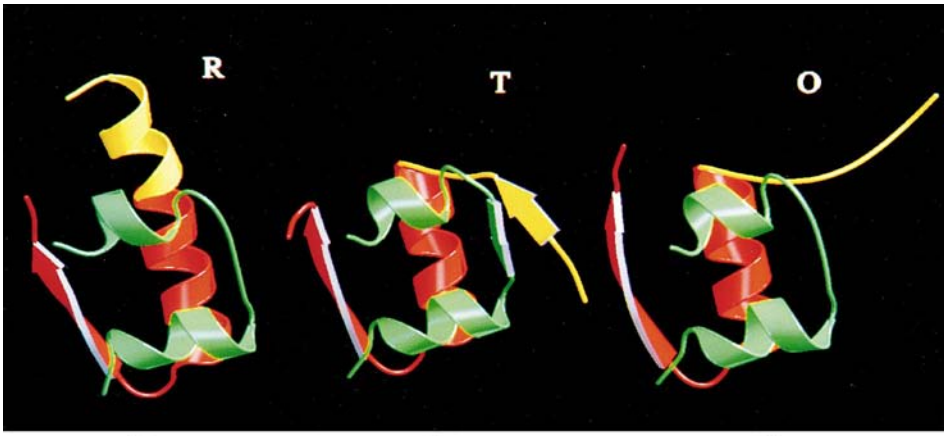

Figure 9
Structural features of the *B*-chain N-terminal residues in different conformational states. (*a*) R state, where the *B*-chain N-terminal residues *B*1–*B*7 are in a helical conformation. (*b*) T state, where the *B*-chain N-terminal residues *B*1–*B*6 are in an extended conformation and approximately form a β-sheet antiparallel to the middle segment of the *A* chain. There are hydrogen bonds between *B*4 and *A*11 and between *B*6 and *A*6 resembling those in an antiparallel β-sheet. (*c*) O state, where the *B*-chain N-terminal residues *B*1–*B*6 are extended and perpendicular to the middle segment of the *A* chain, and protrude out from the monomer's surface without any interaction with the molecule.

4. Conclusions

The structure of the human insulin mutant Ser*B*9Glu was determined by an X-ray crystallographic method. The crystals were grown at pH 3.8 and packed with dimers as building blocks. The *B*9E HI dimer possesses the following structural features compared with the 2Zn insulin dimer. (i) The overall dimer is expanded and more symmetric. The two *B* chains are about 0.8 Å more distant from each other, and the two *A* chains are much more (about 2 Å) further apart. Both monomers are more similar to molecule 1 of the 2Zn insulin dimer. The two *A*1–*A*5 helices become more similar, while the two *B*25 side chains remain asymmetric. (ii) The *B*9E HI dimer structure was stabilized by protonation and neutralization of the carboxyl groups under low-pH conditions and the formation of a hydrogen-bond network among the side chains of residues *B*9, *B*10 and *B*13 of both monomers on the dimer-forming surface, resulting from a structural rearrangement. Under physiological conditions, the dimers would disintegrate owing to repulsion of the negatively charged carboxyl groups of residues *B*9 and *B*13 on both monomers. (iii) The *B*-chain amino-terminal segment is in an open state (O state), a state different from the R and T states found in the insulin hexamer. In the O state, the *B*-chain N-terminal segment is detached from the rest of the molecule. A similar conformational state has been observed in the monomeric crystal structure of despentapetide or desheptapeptide insulin structures (Brange *et al.*, 1997; Bao *et al.*, 1997), as well as in the NMR structure of an engineered insulin monomer (Olsen *et al.*, 1996). This suggests that the O state may be the characteristic conformation of insulin in lower aggregation forms and may play a role in the formation of insulin fibrils.

This work was financially supported by the '863' High-Tech Program of China.

References

Bacon, D. J. & Anderson, W. F. (1988). *J. Mol. Graph.* **6**, 219–220.
Badger, J., Harris, M. R., Reynolds, C. D., Evans, A. C., Dodson, E. J., Dodson, G. G. & North, A. C. T. (1991). *Acta Cryst.* B**47**, 127–136.

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe N. & Vijayan, N. M. (1988). *Philos. Trans. R. Soc. London Ser. B*, **319**, 369–456.

Balschmidt, P., Hansen, F. B., Dodson, E. J., Dodson, G. G. & Korber, F. (1991). *Acta Cryst.* B**47**, 975–986.

Bao, S. J., Xie, D. L., Zhang, J. P., Chang, W. R. & Liang D. C. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 2975–2980.

Bentley, G. A., Brange, J., Derewenda, Z., Dodson, E. J., Dodson, G. G., Markussen, J., Wilkinson, A. J., Wollmer, A. & Xiao, B. (1992). *J. Mol. Biol.* **228**, 1163–1176.

Bentley, G., Dodson, E., Dodson, G., Hodgkin, D. & Mercola, D. (1976). *Nature (London)*, **261**, 166–168.

Bi, R. C., Dauter, Z., Dodson, E., Dodson, G., Giordano, F. & Reynolds, C. (1984). *Biopolymers*, **23**, 391–395.

Blundell, T., Dodson, G., Hodgkin, D. & Mercola, D. (1972). *Adv. Protein Chem.* **26**, 279–402.

Brader, M. L. & Dunn, M. F. (1991). *Trends Biochem. Sci.* **16**, 341–345.

Brange, J., Dodson, G. G., Edwards, D. J., Holden, P. H. & Whittingham, J. L. (1997). *Proteins Struct. Funct. Gen.* **27**, 507–516.

Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sorensen, A. R. & Voigt, H. O. (1988). *Nature (London)*, **333**, 679–682.

Brunger, A. T. (1992). *X-PLOR. Version* 3.1. *A System for X-ray Crystallography and NMR.* Yale University, New Haven, Connecticut, USA.

Chang, W., Stuart, D., Dai, J., Todd, R., Zhang, J., Xie, D., Kuang, B. & Liang, D. (1986). *Sci. Sin. (Engl. Ed.)* **29**, 1273–1284.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Spark, C. & Swenson, D. (1989). *Nature (London)*, **338**, 594–596.

Hua, Q. X., Shoelson, S. E., Kochoyan, M. & Weiss, M. A. (1991). *Nature (London)*, **354**, 238–241.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Jorgensen, A. M., Kristensen, S. M., Led, J. J. & Balschmidt, P. (1992). *J. Mol. Biol.* **227**, 1146–1163.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Liang, D., Stuart, D., Dai, J., Todd, R., You, J. & Lou, M. (1985). *Sci. Sin. B*, **28**, 472–483.

Liu, B., Liang, Z.-H., Tang, Y.-H., Zhang, X-T., Zhu, S. Q. & Feng Y.-M. (1996). *Acta Biochem. Biophys. Sin.* **28**, 245–249.

Markussen, J., Diers, I., Engesgaard, A., Hansen, M. T., Hougaard, P., Langkjaer, L., Norris, K., Ribel, U., Sorenssen, A. R., Sorenssen, E. & Voigt, H. O. (1987). *Protein Eng.* **1**, 215–223.

Markussen, J., Hougaard, P., Ribel, U., Sorensen, A. R. & Sorenssen, E. (1987). *Protein Eng.* **1**, 205–213.

Olsen, H. B., Ludvigsen, S. & Kaarsholm, N. C. (1996). *Biochemistry*, **35**, 8836–8845.

Owens, D. R. (1986). *Human Insulin: Clinical Pharmacological Studies in Normal Man*, pp. 46–49. Dordrecht: Kluwer.

Peking Insulin Structure Research Group (1974). *Sci. Sin. B*, **29**, 1273–1284.

Roussel, A. & Cambillau, C. (1991). *TURBO-FRODO. Silicon Graphics Geometry Partners Directory*, p. 86. Silicon Graphics, Mountain View, California, USA.

Smith, G. D. & Dodson, G. G. (1992). *Proteins*, **14**, 401–408.

Smith, G. D., Swenson, D. C., Dodson, E. J., Dodson, G. G. & Reynolds, C. D. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 7093–7097.