Acta Cryst. (1997). D53, 507–512

Structure of Monomeric Porcine DesB1–B2 Despentapeptide (B26–B30) Insulin at 1.65 Å Resolution

JIA-SHENG DIAO, ZHU-LI WAN, WEN-RUI CHANG AND DONG-CAI LIANG*

National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. E-mail: dcliang@sun5.ibp.ac.cn

(*Received 5 February 1996; accepted 13 March 1997*)

Abstract

Insulin has a concentration of 10^{-8}–10^{-11} *M* in the blood which ensures that it circulates and exerts its physiological functions *in vivo* as a monomer. The crystal structure of monomeric porcine desB1-B2 despentapeptide (B26–B30) insulin (DesB1-2 DPI) with $M_r = 4934$ Da has been determined at 1.65 Å resolution using the molecular replacement method. A structural comparison between DesB1-2 DPI and 2Zn insulin reveals that the conformation of DesB1-2 DPI is more similar to molecule I than molecule II of 2Zn insulin. The remarkable conformational difference between *B*25-Phe in DesB1-2 DPI and *B*25-Phe in despentapeptide (B26–B30) insulin (DPI) indicates that the residue *B*25-Phe possesses great flexibility and mobility.

1. Introduction

Insulin, which is the principal hormone controlling blood glucose level, acts by stimulating glucose influx and metabolism in muscles and adipocytes and inhibiting gluconeogenesis by the liver. In addition insulin modifies the expression or activity of a variety of enzymes and transport systems in nearly every kind of cell. Insulin activity is mediated through insulin receptors, transmembrane glycoproteins with intrinsic protein tyrosine kinase activity. The insulin receptor has a molecular mass of 300 000 and includes two α- and two β-chains organized in a way similar to immunoglobin molecules. Insulin binds to the receptor α-chain to initiate the phosphorylation of β-subunit tyrosine residues which then influence enzymes involved in phosphorylation and desphosphorylation in the cell which may, in turn, activate the processes of glycolysis, lipogenesis and protein synthesis (Kasuga *et al.*, 1982; Van Obberghen, Rossi, Kowalski, Gazzano & Ponzio, 1983).

Sequence comparison has shown that *A*1–*A*4 (Gly-Ile-Val-Glu or Asp) and *B*23–*B*26 (Gly-Phe-Phe-Tyr) exhibit extended interspecies homology and they are preserved in almost all insulin species (Blundell & Wood, 1975). In fact, amino-acid substitutions within these regions have been identified in the insulins from three individuals in whom genetic mutations resulted in the secretion of abnormal hormones associated with diabetes (Tager *et al.*, 1979, 1980; Shoelson, Haneda *et al.*, 1983; Shoelson, Fickova *et al.*, 1983; Haneda, Chan, Kowk, Rubenstein & Steiner, 1983), insulin Chicago (*B*25-Phe→*B*25-Leu), insulin Los Angeles (*B*24-Phe→*B*24-Ser), and insulin Wakayama (*A*3-Val→*A*3-Leu). The structural analysis of a *A*1–*B*29 cross-linked insulin, which was completely inactive, indicates that its conformation is nearly same as that of native insulin (Derewenda *et al.*, 1991). The lower binding potency of the cross-linked insulin results from the restraining effect of the cross-link that impedes the conformational changes necessary for generation of the binding conformation. In the possible binding-interaction mechanism for insulin with its receptor, when the monomeric insulin molecule comes close to its receptor molecule, the C-terminal segment of the *B* chain moves away from its original position through the hinge peptide at *B*24 and thus exposes the hydrophobic surface for recognition and binding with a corresponding hydrophobic surface on the receptor. *B*25-Phe, which has a strong conformational mobility, may orient itself for the recognition and interaction of the π-bonds between aromatic rings when the insulin molecule comes close to and binds with the receptor (Liang, Chang, Zhang & Wan, 1992; Liang, Chang, Wang & Zhang, 1992; Wang & Liang, 1985; Liang, Chang & Wan, 1994). Additional studies on the insulin analogues of the *B*-chain N-terminus reveal that the removal of the N-terminal *B*1-Phe does not lead to any change in the biological activity but greatly reduces the immunoactivity (Brandenburg, 1969). Investigations of insulin analogues with amino acids removed from the *B*-chain N-terminus (Schwartz & Katsoyannis, 1978; Chu, Chu & Chang, 1984; Lei, Dong, Xie & Yuan, 1983), have indicated that the immunoactivity and biological activity both decrease when the peptide chain is shortened.

Several three-dimensional structures of insulins from different species, chemically modified insulin derivatives, and mutant insulins have been described, including some high-resolution X-ray diffraction analyses (Adams *et al.*, 1969; Cutfield *et al.*, 1981; Chang *et al.*, 1986). These studies have produced a wealth of data

leading to various models describing the interaction of insulin with its receptor. DesB1-2 DPI which retains 10.8% receptor-binding activity of insulin but a very low immunoactivity (2.4%) (Chu *et al.*, 1984; Chu, 1987) is one kind of insulin derivative having a modified *B*-chain N-terminus and C-terminus that exists in monomeric form. The structure of DesB1-2 DPI will be used for comparison with the structures of 2Zn insulin, DPI and other insulin derivatives to further analyze the insulin structure–function relationship.

2. Experimental

2.1. *Crystal and data*

DesB1-2 DPI was crystallized as described by Wan, Diao, Chang & Liang (1994). The best crystals were obtained from a 5 mg ml^{-1} protein solution in 0.05 *M* citrate buffer (pH 8.4) containing 0.025% zinc chloride and 10% acetone. Crystals of DesB1-2 DPI belong to space group $P2_1$, with cell dimensions $a = 26.9$, $b = 24.5$, $c = 28.1$ Å, $\beta = 103.8°$. There is one molecule in each crystallographic asymmetric unit according to the crystal densities of insulin analogues. The unit-cell dimensions correspond to a specific volume V_m of 1.82 Å^3 Da^{-1}, implying a solvent content of 32% (Matthews, 1968).

Reflection data for DesB1-2 DPI crystals were measured at room temperature on a Siemens X-200B area detector using graphite-monochromated Cu $K\alpha$ radiation generated by a Rigaku RU-300 rotating anode operating at 200 kV and 50 mA. The crystal-to-detector distance was fixed at 8 cm. Two data sets were collected at 2θ angles of 0 and 25° to obtain as many reflections as possible. The two data sets were merged using the *XENGEN* package (Howard, Gilliland, Finzel & Poulos, 1987) with R_{sym} equal to 2.72%. There were 3829 unique reflections, 3739 of these were greater than 5σ and 3789 were greater than 2σ. The completeness of the total data set for DesB1-2 DPI in the range of 27.3–1.65 Å was 87.1%.

2.2. *Structural determination and refinement*

The cross-rotation function from *X-PLOR* (Brünger, 1992) was used for the cross-rotation search. The 1.5 Å DPI crystal structure model (Dai, Lou, You & Liang, 1987) with the residues *B*1 and *B*2 cut off was used as the search model. Patterson maps of the model were calculated by placing the probe model into a cubic box with cell edges 60 Å long for the structure calculation. The first peak of the rotation function always appeared at Eulerian angles (248.7, 15.0, 288.7°). The rotation search values showed apparent discrimination; using an integration radius of 20 Å and the data between 7 and 3 Å resolution, the highest value from the rotation search was 8.7σ. The second highest was 6.3σ. The 100 highest peaks from the cross-rotation search were subjected to Patterson correlation (PC) refinement treating the entire DesB1-2 DPI as a rigid body. PC refinement produced a maximum peak that corresponded to the highest solution from the cross-rotation function search. The correlation coefficient of this solution was 0.34. The next highest peak had a correlation coefficient of only 0.13. The model was then rotated by the angles (248.7, 15.0, 288.7°) corresponding to the best rotation solution. The *X-PLOR* translation function was used for the translation search. The translation solution $X = 0.038$, $Y = 0.000$, $Z = 0.192$ (fractional coordinates) always appeared at the first position. Using data from 8 to 4 Å, the translation

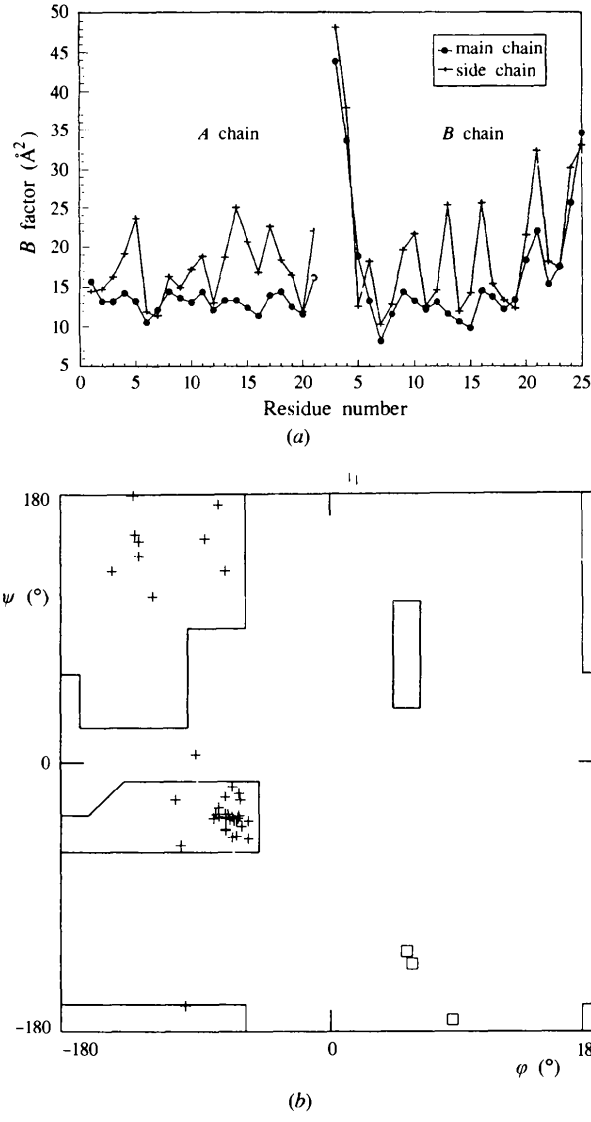

Fig. 1. (*a*) Average isotropic temperature factor for main-chain and side-chain non-H atoms from DesB1-2 DPI. (*b*) Ramachandran plot. Non-glycine residues are indicated by + and glycine residues are indicated by □.

function gave the largest signal-to-noise ratio, 1.47, and the lowest R factor, 40.9%.

The rotation function and the translation function gave the orientation angles (248.7, 15.0, 288.7°) and the position (0.038, 0.000, 0.192) of the DesB1-2 DPI molecule in the unit cell. Rigid-body refinement, using all the data between 8 and 3 Å resolution, reduced the R factor to 40.6%. The rigid-body refined DesB1-2 DPI model was the starting point for the restrained least-squares refinement using *GPRLSA* (Furey, Wang & Sax, 1982). The refinement started at low resolution with a small number of reflections and used tight stereochemical restraints to avoid destroying the original model. Seven rounds of refinement were executed using data in the range 10–3 Å. The structure was then refined with higher 2.5 Å resolution data. After each round of refinement, the model was adjusted on a graphics workstation using omit maps and $2F_o - F_c$ maps. The individual thermal-factor refinement was initiated during the positional refinement at 2.2 Å which reduces the R factor to 25.4%. The resolution was then continually extended to 2, 1.8 and 1.65 Å with adjustments to the model at each round which further reduces the R factor to 22.8%. The resolution was extended to 10–1.65 Å to include 10–5 Å data closely concerned with solvent reflections. 33 water molecules with full occupancy were added into the final structure. At the end of the crystallographic refinement, the R factor calculated on the basis of 378 non-H atoms of protein and solvent molecules was 19.7%, for the 3774 reflections in the 10–1.65 Å resolution range (86.3% completeness in the shell). The r.m.s. deviations from ideality were 0.017 Å for bond lengths, 1.9° for bond angles and 0.021 Å for planar groups. Figs. 1(*a*) and 1(*b*) display the B-factor distribution and stereochemistry of DesB1-2 DPI.

The atomic coordinates and structure factors of DesB1-2 DPI have been deposited in the Protein Data Bank, Brookhaven National Laboratory.*

3. Results and discussion

3.1. *Structural description*

The DesB1-2 DPI monomer contains two chains, A and B, containing 21 and 23 residues, respectively, linked by disulfide bridges, $A7$–$B7$ and $A20$–$B19$. A third disulfide bridge lies within the A chain at $A6$–$A11$. The A chain contains two helices, $A1$–$A9$ and $A13$–$A19$ which lie antiparallel to each other. Residues $A10$–$A12$ constitute a loop region and residues $A20$–$A21$ form an extended peptide. The first five residues of the B chain, $B3$–$B7$ have an extended conformation. At residue $B8$, the chain takes a sharp turn and residues $B9$–$B19$ form an α-helix. Residues $B20$ and $B23$ are in a U-shaped configuration but the remaining residues $B24$–$B25$ form part of an extended chain.

The electron density for the DesB1-2 DPI molecule is very well defined (Fig. 2) while the density for the side chain of residue $B3$-Asn is somewhat weak. A total of 33 well defined water molecules form a network of hydrogen bonds, with 25 of them directly hydrogen bonded with the protein molecule. The average temperature factor of the solvent molecules is roughly comparable to that of the other atoms in the structure. This substantial order of water molecules in the structure may arise from the relatively low solvent content of the DesB1-2 DPI crystals (32%), so that most of the 33 solvent molecules are located within the first or second contact shell of the protein molecule.

3.2. *Comparison with 2Zn insulin*

Molecule I of the dimer of 2Zn insulin is very similar in conformation to the corresponding molecule I of the dimer of the 4Zn form. The 2Zn insulin coordinates were derived from the structure refined at 1.2 Å resolution with an R factor of 12.8% (Chang *et al.*, 1986).

The superposition is based on 44 residues common to DesB1-2 DPI and to molecule I or molecule II of 2Zn insulin. The r.m.s. deviations of DesB1-1 DPI from molecule I and molecule II for the superimposed atoms of all residues were 1.46 and 1.96 Å, respectively. The corresponding deviations for main-chain atoms were 0.64 and 1.06 Å. As previously noted (Stuart *et al.*, 1984), these results confirm that the insulin molecule I represents a more typical insulin structure, which is also in accord with studies of 4Zn and hagfish insulins (Cutfield *et al.*, 1981). The detailed information was obtained from the comparison by dividing the DesB1-2 DPI molecule into the A chain, the B chain, and several individual segments of $A1$–$A9$, $A13$–$A18$, $B3$–$B8$, $B9$–$B19$ and $B20$–$B25$ (Table 1). The r.m.s. deviations of the 2Zn insulin molecule I from DesB1-2 DPI for all atoms of the A chain, the B chain, and individual segments were almost all less than the r.m.s. deviations from comparisons of DesB1-2 DPI with molecule II of 2Zn insulin, giving further evidence of the similarities between molecule I of 2Zn insulin and DesB1-2 DPI.

Structural comparison of either the entire molecule, the A chain, the B chain, or the individual segments of monomeric DesB1-2 DPI with corresponding parts of the 2Zn insulin molecule reveal that the conformation of molecule I in the dimer of 2Zn insulin is quite close to the structure of the monomeric DesB1-2 DPI molecule,

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1SBD and R1SBDSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0615). At the request of the authors the structure factors will remain privileged until 1 January 2000.

Table 1. *Atomic r.m.s differences for the DesB1-2 DPI and 2Zn insulin structures*

	Molecule I		Molecule II	
Residue range*	All atoms†	Main chain‡	All atoms	Main chain
*A*1–*B*25	1.46	0.64	1.96	1.06
*A*1–*A*21	1.03	0.46	1.64	1.16
*A*1–*A*9	0.63	0.20	1.69	1.12
*A*13–*A*18	1.23	0.25	1.25	0.19
*B*3–*B*25	1.67	0.67	2.13	0.82
*B*3–*B*8	1.17	0.55	1.73	0.77
*B*9–*B*19	1.80	0.17	1.78	0.20
*B*20–*B*25	1.18	0.82	2.14	0.79

* Superimposed fragments. † All non-H atoms. ‡ Atoms N, C and Cα.

implying that the conformation of molecule I is more similar to the monomeric molecule of DesB1-2 DPI than that of molecule II in 2Zn insulin.

3.3. *Comparison with DPI*

The crystal structure of a monomeric porcine despentapeptide (B26–B30) insulin (DPI) was first determined at 2.4 Å resolution (Liang *et al.*, 1983; Stuart *et al.*, 1984) and then refined at 1.5 Å resolution with an *R* factor of 14.4% (Dai *et al.*, 1987). DPI was found to have only one molecule in the asymmetric unit. A beef DPI structure containing two molecules in the asymmetric unit was also reported at high resolution (Bi *et al.*, 1984). The structural comparison of the modified insulin monomers of DesB1-2 DPI and porcine DPI at high resolution gives some insight into the structural features of the free insulin monomer.

Considering the main-chain atoms of the *A* chain and the *B* chain, the optimal superposition produced r.m.s. deviations of 0.38 and 1.07 Å, respectively. These results showed that the *A*-chain conformations of the two monomers are more closely related than those of the *B* chain. Further structural comparison was based on the individual segments of DesB1-2 DPI and the equivalent segments of DPI. The main-chain r.m.s. deviations of individual segments *A*1–*A*9 (0.18 Å), *A*10–*A*12 (0.14 Å) and *A*13–*A*18 (0.25 Å) were all less than that of the *A* chain as a whole (0.38 Å), which implies that despite the similar folding topology, the individual segments of the *A* chain have moved relative to one other. The main-chain atoms of *B*3–*B*8 and *B*9–*B*19 of DesB1-2 DPI and DPI fit with r.m.s. deviations of 0.17 and 0.32 Å, whereas the C-termini of the *B* chain have quite different conformations with r.m.s. deviations of 1.42 Å for residues *B*20–*B*25. *B*25-Phe in the DesB1-2 DPI molecule had moved a significantly amount from its position in the DPI molecule (Fig. 3). The mobility of *B*25 in a despentapeptide insulin is perhaps affected by its being the C-terminal residue. It is apparent in Fig. 1(*a*) where the residue *B*25 and as well as N-terminal residues *B*3 and *B*4 have high *B* factors.

In summary, the general foldings of DesB1-2 DPI and DPI are very similar. The most stable parts of the two monomers are the three helices *A*1–*A*9, *A*13–*A*18 and *B*9–*B*19, which have almost identical conformations in the two structures with some relative movement around certain peptide bonds. The great variation occurs in the region of residues *B*20–*B*25, in which the *B*25-Phe residue is markedly different.

Studies of the model proposed to describe the interaction between the insulin molecule with its receptor (Liang *et al.*, 1994) showed that the hydrophobic residues on the amphipathic binding surface consisted mainly of *A*2-Ile, *A*3-Val, *A*19-Tyr, *B*11-Leu,

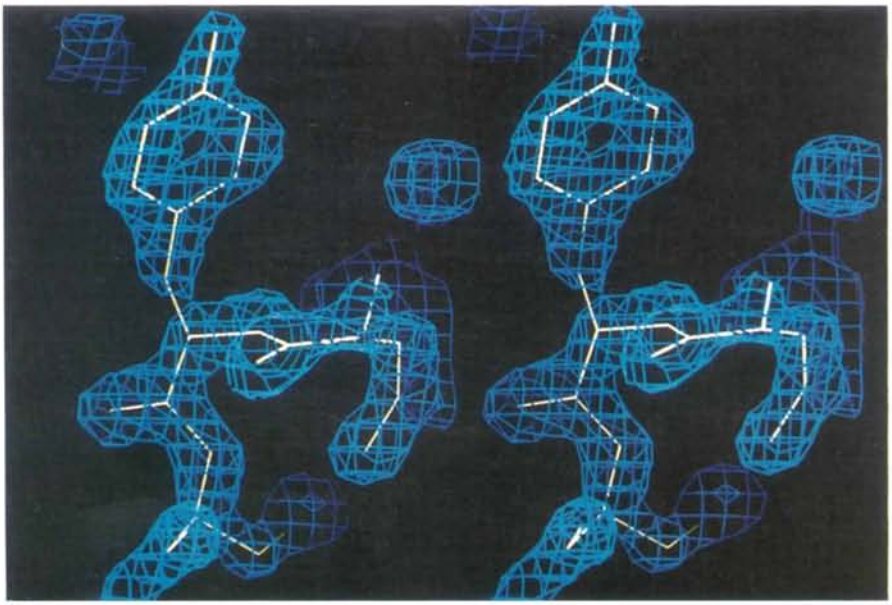

Fig. 2. Stereoview of residue *A*19-Tyr superimposed with the corresponding $2F_o - F_c$ map contoured at 2.0σ.

*B*12-Val, *B*15-Leu, *B*16-Tyr, *B*24-Phe, *B*25-Phe, and perhaps *B*26-Tyr. Residue *B*25-Phe was one of the crucial residues having a bulky aromatic ring on the hydrophobic surface. In porcine DPI crystals, the orientation of residue *B*25-Phe was opposite to that of the 2Zn porcine insulin and the main chain had been turned (Liang *et al.*, 1992), indicating that the conformation of residue *B*25–Phe was very flexible. The position of residue *B*25-Phe in the DesB1-2 DPI structure has been determined precisely (Fig. 4). Superposition of the main chains of DesB1-2 DPI and DPI (Fig. 3) shows that the conformation of residue *B*25-Phe in the *B*-chain C-terminus is significantly different in the two molecules providing further evidence that the conformation of residue *B*25-Phe is very flexible. This conformational flexibility may guarantee the conformational changes of the *B*-chain C-terminus, which is essential for insulin recognition with its receptor (Wang & Liang, 1985).

The authors are grateful for financial support from the Foundation of the Chinese Academy of Sciences and the National Natural Science Foundation of China.

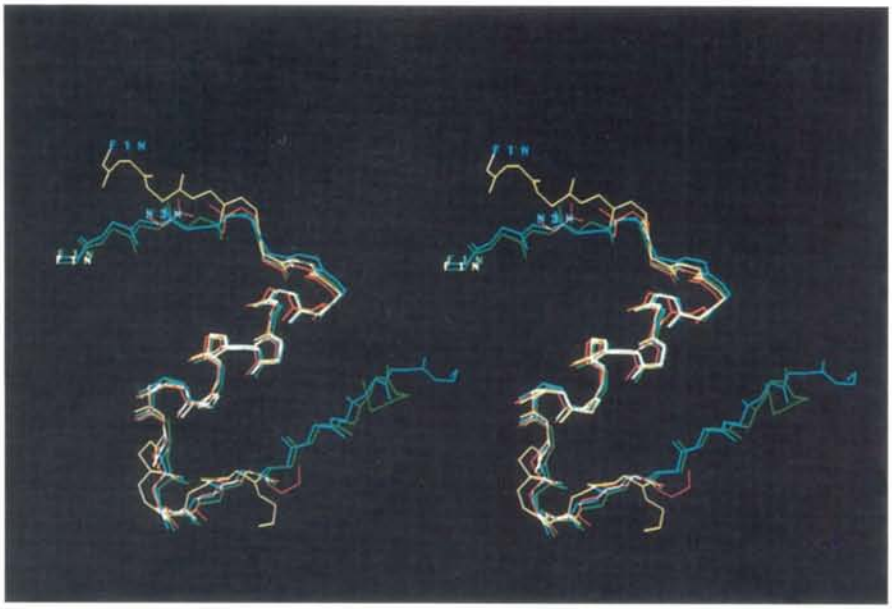

Fig. 3. Stereoview of least-squares superposition of the main-chain atoms of the *B* chain of DesB1-2 DPI (red Line). DPI (yellow line), molecule I (green line) and molecule II (blue line) of 2Zn insulin.

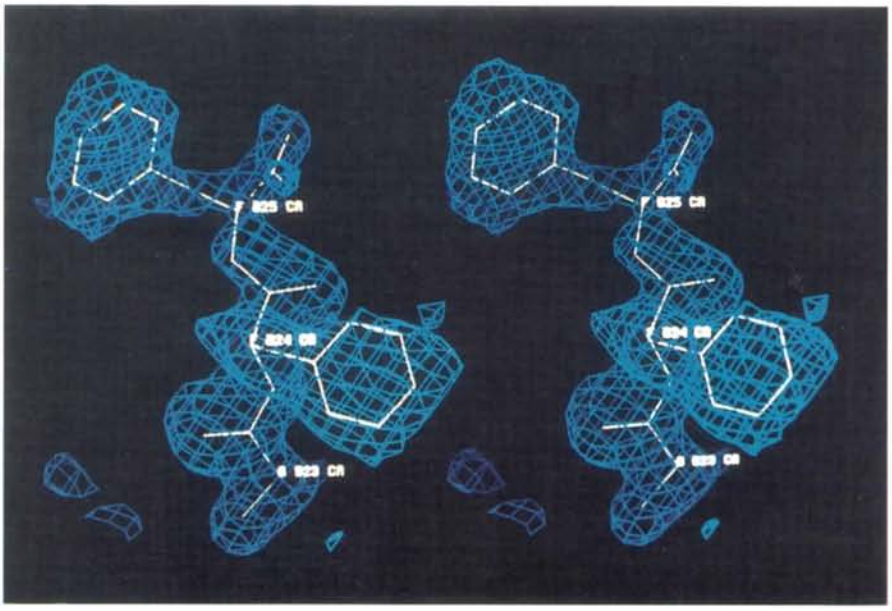

Fig. 4. Stereoview of the $F_o - F_c$ omit map of residues *B*23–*B*25 contoured at 2.0σ.

References

Adams, M. J., Blundell, T. J., Dodson, E. L., Dodson, G. G., Vijayan, M., Baker, E. N., Harding, M. M., Hodgkin, D. C., Rimmer, R. & Sheet, S. (1969). *Nature (London)*, **224**, 491–495

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

Blundell, T. L. & Wood, S. P. (1975). *Nature (London)*, **257**, 197–203.

Brandenburg, D. (1969). *Hoppe-Seyler's Z. Physiol. Chem.* **350**, 741–747.

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

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

Chu, J. H., Chu, S. C. & Chang, Y. S. (1984). *Acta Biochim. Biophys. Sin.* **16**, 264–268.

Chu, S. C. (1987). *Acta Biochim. Biophys. Sin.* **19**, 99–104.

Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G., Reynolds, C. D. & Valley, D. (1981). *Structural Studies on Molecules of Biological Interest*, edited by G. Dodson, J. P. Glusker & D. Sayre, pp. 527–546. Oxford University Press.

Dai, J. B., Lou, M.Z., You, J. M. & Liang, D. C. (1987). *Sci. Sin. B*, **30**, 55–65.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Xiao, B. & Markussen, J. (1991). *J. Mol. Biol.* **220**, 425–433.

Furey, W., Wang, B. C. & Sax, M. (1982). *J. Appl. Cryst.* **15**, 160–166.

Haneda, M., Chan, S. J., Kowk, S. C. M., Rubenstein, A. H. & Steiner, D. F. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 6366–6370.

Howard, A. J., Gilliland, G. L., Finzel, B. C. & Poulos, T. L. (1987). *J. Appl. Cryst.* **20**, 383–387.

Kasuga, K., Zick, Y., Blith, D. L., Karlsson, F. A., Haring, H. U. & Kahn, C. R. (1982). *J. Biol. Chem.* **257**, 9891–9894.

Lei, K. J., Dong, B., Xie, D. L.& Yuan, H. S. (1983). *Acta Biochim. Biophys. Sin.* **15**, 547–461.

Liang, D. C., Chang, W. R. & Wan, Z. L. (1994). *Biophys. Chem.* **50**, 63–71.

Liang, D. C., Chang, W. R., Wang, D. C. & Zhang, G. P. (1992). *Sci. China, B*, **35**, 418–428.

Liang, D. C., Chang, W. R., Zhang, G. P. & Wan, Z. L. (1992). *Sci. China B*, **35**, 547–557.

Liang, D. C., Dai, J. B., Stuart, D., Wan, Z. L., Todd, R., Yun, J. M. & Lou, M. Z. (1983). *Biochem. Soc. Trans.* **11**, 419–425.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Schwartz, G. & Katsoyannis, P. G. (1978). *Biochemistry*, **17**, 4550–4556.

Shoelson, S., Fickova, M., Haneda, M., Nahum, A., Musso, G., Kaiser, E. T., Rubenstein, A. & Tager, H. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 7390–7394.

Shoelson, S., Haneda, M., Blix, P., Nanjo, A., Sanke, T., Inouye, K., Steiner, D., Rubenstein, A. & Tager, H. (1983). *Nature (London)*, **302**, 540–543.

Stuart, D., Dai, J. B., Lou, M. Z., You, J. M., Todd, R., Li, J. Y., Wang, J. H. & Liang, D. C. (1984). *Sci. Sin. B*, **27**, 1265–1272.

Tager, H., Given, B., Baldwin, D., Mako, M., Markese, J., Rubenstein, A., Olefsky, J., Koboyashi, M., Kotterman, O. & Poucher, R. (1979). *Nature (London)*, **281**, 122–125.

Tager, H., Thomas, N., Assoian, R., Rubenstein, A., Saekow, M., Olefsky, J. & Kaiser, T. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 3181–3185.

Van Obberghen, E., Rossi, B., Kowalski, A., Gazzano, H. & Ponzio, G. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 945–949.

Wan, Z. L., Diao, J. S., Chang, W. R. & Liang, D. C. (1994). *Chin. Sci. Bull.* **39**, 1130–1132.

Wang, C. C. & Liang, D. C. (1985). *Chin. Biochem. J.* **1**, 9–12.