Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

R_6 hexameric insulin complexed with *m*-cresol or resorcinol

G. David Smith,[a,b]* Ewa Ciszak,[a]† Lucy A. Magrum,[a] Walter A. Pangborn[a] and Robert H. Blessing[a]

[a]Hauptman–Woodward Medical Research Institute, 73 High Street, Buffalo, NY 14203, USA, and [b]Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA

† Present address: Universities Space Research Association, NASA/Marshall Space Flight Center, Laboratory for Structural Biology, Huntsville, AL 35812, USA.

Correspondence e-mail: smith@hwi.buffalo.edu

Received 5 May 2000
Accepted 17 September 2000

PDB References: monoclinic resorcinol–insulin, 1evr; monoclinic *m*-cresol–insulin, 1ev6; rhombohedral *m*-cresol–insulin, 1ev3.

The structures of three R_6 human insulin hexamers have been determined. Crystals of monoclinic *m*-cresol–insulin, monoclinic resorcinol–insulin and rhombohedral *m*-cresol–insulin diffracted to 1.9, 1.9 and 1.78 Å, respectively, and have been refined to residuals of 0.195, 0.179 and 0.200, respectively. In all three structures, a phenolic derivative is found to occupy the phenolic binding site, where it forms hydrogen bonds to the carbonyl O atom of CysA6 and the N atom of CysA11. Two additional phenolic derivative binding sites were identified within or between hexamers. The structures of all three hexamers are nearly identical, although a large displacement of the N-terminus of one B chain in both monoclinic structures results from coordination to a sodium ion which is located between symmetry-related hexamers. Other minor differences in structure arise from differences in packing in the monoclinic cell compared with the rhombohedral cell. Based upon the differences in conformation of the GluB13 side chains in T_6, $T_3R_3^f$ and R_6 hexamers, the deprotonation of these side chains appears to be associated with the T→R conformational transition.

1. Introduction

Insulin is a polypeptide hormone that is critical for the metabolism of glucose. The biologically active form, monomeric insulin, consists of a 21-residue A chain and a 30-residue B chain linked by two interchain disulfide bonds; one intrachain disulfide bond exists in the A chain. The conformation of the A chain consists of two short α-helical segments linked by an extended section of polypeptide chain, while the B-chain conformation consists of a central α-helical section flanked by two extended regions of polypeptide. In the presence of zinc ions, insulin associates to form a threefold symmetric hexamer (Baker *et al.*, 1988).

X-ray crystallographic studies (Bentley *et al.*, 1976; Ciszak & Smith, 1994; Smith *et al.*, 1996; Whittingham *et al.*, 1995; Derewenda *et al.*, 1989; Smith & Dodson, 1992) and spectroscopic studies (Kaarsholm *et al.*, 1989; Krüger *et al.*, 1990; Choi *et al.*, 1993; Brzovic *et al.*, 1994; Bloom *et al.*, 1995) have shown insulin to be an allosteric molecule; the T and R nomenclature is used to describe the different hexamer types (Kaarsholm *et al.*, 1989). In the absence of high chloride ion concentrations or phenolic derivatives, the first eight residues of the B chain adopt an extended conformation producing a T_6 hexamer (Baker *et al.*, 1988). In the presence of high chloride or thiocyanate concentrations, residues B4 to B8 in only one trimer are transformed to an α-helical conformation, producing a $T_3R_3^f$ hexamer (Whittingham *et al.*, 1995; Ciszak & Smith, 1994; Ciszak *et al.*, 1995). In this hexameric form, a continuous

© 2000 International Union of Crystallography
Printed in Denmark – all rights reserved

Table 1
Crystallization conditions for rhombohedral *m*-cresol (R-Crs), monoclinic *m*-cresol (M-Crs) and monoclinic resorcinol (M-Res) complexes with insulin.

	R-Crs	M-Crs	M-Res
Insulin (mg)	30	30	30
0.02 *M* HCl (ml)	3.0	3.0	3.0
0.15 *M* zinc acetate (ml)	0.3	0.3	0.3
0.2 *M* sodium citrate (ml)	1.5	1.5	1.5
2.5% *m*-cresol in acetone (ml)	1.2		
5% *m*-cresol in ethanol (ml)		1.2	
5% resorcinol in water (ml)			1.2
Sodium chloride (g)	0.36		0.36
pH	8.5	6.5	6.7

Table 2
Data-measurement statistics for rhombohedral *m*-cresol (R-Crs), monoclinic *m*-cresol (M-Crs) and monoclinic resorcinol (M-Res) complexes with insulin.

	R-Crs	M-Crs	M-Res
Space group	*R*3	$P2_1$	$P2_1$
Unit-cell parameters			
a (Å)	78.866	61.247	61.350
b (Å)	78.866	61.739	61.926
c (Å)	39.465	47.467	47.805
α (°)	90.0	90.0	90.0
β (°)	90.0	111.32	110.61
γ (°)	120.0	90.0	90.0
Temperature (K)	100	100	290
B_{iso} (Å^2)	30.6	30.6	33.7
No. of frames	60	85	66
Resolution (Å)	1.70	1.80	1.9
Total data	64359	174925	81864
Unique data	9334	28441	25793
Average redundancy	6.9	6.15	3.2
Completeness (%)	92.8	92.6	96.9
Overall R_{merge}	0.047	0.039	0.048
$\langle F^2\rangle/\langle\sigma(F^2)\rangle$	60.6	40.6	18.4
Resolution (final shell)	1.74–1.70	1.84–1.80	
Completeness (%)	54.1	62.1	
$\langle F^2\rangle/\langle\sigma(F^2)\rangle$	4.8	2.9	
R_{merge}	0.277	0.394	
Resolution (used in refinement)	1.83–1.78	1.94–1.89	2.00–1.90
Completeness (%)	82.6	92.6	93.1
$\langle F^2\rangle/\langle\sigma(F^2)\rangle$	10.0	5.4	2.01
R_{merge}	0.173	0.287	†

† This information is not available with the *R-AXIS* data-processing software.

α-helix is present from GluB4 to CysB19 in each monomer in the R^f-state trimer.

While higher chloride or thiocyanate ion concentrations are ineffective in producing additional conformational transitions in the hexamer, the addition of phenolic derivatives such as phenol, *m*-cresol or resorcinol will drive the T→R transition to completion resulting in monoclinic (Derewenda *et al.*, 1989) or rhombohedral (Smith & Dodson, 1992; Whittingham *et al.*, 1998) crystals containing R_6 hexamers. In contrast to the 'frayed' R^f conformation, the 'pure' R conformation contains a continuous α-helix from PheB1 to CysB19. However, not all phenolic derivatives are capable of driving the T→R transition to completion, as methylparaben (Whittingham *et al.*, 1995), 4′-hydroxyacetanilide (Smith & Ciszak, 1994) and *p*-hydroxybenzamide (Smith *et al.*, 1996) are found as complexes in $T_3R_3^f$ insulin. Hexameric insulin will also crystallize in several different crystalline forms. Rhombohedral crystals in which the threefold axis of the hexamer is coincident with the crystallographic threefold axis have been reported for T_6, $T_3R_3^f$ and R_6 hexamers, while R_6 insulin hexamers will also crystallize in the monoclinic system.

Reported here are the crystal structures of three R_6 human insulin hexamers: two monoclinic forms complexed with either *m*-cresol or resorcinol and one rhombohedral form complexed with *m*-cresol. These structures provide further details of the interactions between phenolic derivatives and the residues within the phenolic binding site and also identify additional sites in which phenolic derivatives can bind to the hexamer.

2. Experimental

2.1. Crystallization

Insulin was provided by Lilly Research Laboratories as crystalline biosynthetic human insulin complexed with zinc. Buffers, salts and other reagents were purchased and used without further purification. The reagents listed in Table 1 were added in the order given, the pH was raised to approximately 9.0 with sodium hydroxide to ensure complete dissolution and hydrochloric acid was used to back-titrate to the final pH. The solutions were transferred to test tubes, warmed to 323 K, placed in a Dewar and allowed to stand undisturbed for one week. The resulting crystals from all three experiments had dimensions ranging from 0.4 to 0.8 mm on an edge. In the resorcinol–insulin crystallization, both rhombohedral and monoclinic crystals were present in the test tubes.

2.2. Data measurements

2.2.1. *m*-Cresol–insulin complexes. Small crystals of both the rhombohedral and monoclinic forms were immersed in paraffin oil in a depression plate. A loop was used to slide the crystal along the bottom of the well, which caused any remaining mother liquor surrounding the crystal to adhere to the well surface. The crystals were picked up in a loop, plunged into liquid nitrogen and transferred to a goniometer head on a Rigaku R-AXIS IV and RU-200 rotating-anode generator equipped with Osmic confocal mirrors and an MSC X-Stream low-temperature system. Integration, scaling and merging of the data were performed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The resulting independent data were then processed further using *SORTAV*, *BAYES* and *LEVY* (Blessing, 1997) to reformat the data, to apply Bayesian statistics to the weak data and to estimate the scale factor and overall isotropic temperature factor, respectively. Unit-cell, space-group and data statistics are given in Table 2.

2.2.2. Resorcinol–insulin complex. A crystal approximately 0.5 mm on an edge was mounted in a glass capillary along with a small drop of mother liquor. Data were measured at room temperature on a Rigaku R-AXIS IIC image-plate system and RU-200 rotating-anode generator with graphite monochromated Cu $K\alpha$ radiation (λ = 1.54178 Å). The integration,

scaling and merging was performed with the *R-AXIS* software. The data were processed using *SORTAV*, *BAYES* and *LEVY* (Blessing, 1997) as described above. The unit-cell, space-group and data statistics are given in Table 2.

2.3. Structure solution and refinement

2.3.1. Monoclinic resorcinol–insulin complex. Both of the monoclinic structures are isomorphous with the monoclinic form of insulin complexed with phenol (Derewenda *et al.*, 1989). The initial model for the resorcinol complex with insulin was constructed by excluding all non-protein atoms from the structure of the phenol complex (PDB code 1znj) as well as B1–B3 and B28–B30 from all six B chains. An initial $F_o - F_c$ map clearly indicated the positions of both zinc ions, both chloride ions and seven resorcinol molecules. These atoms and molecules were included in the model and the refinement was begun with *PROFFT* (Finzel, 1987; Hendrickson & Konnert, 1980) using data in the resolution range 8.0–1.9 Å resolution and $F \geq 2.5\sigma(F)$. As the refinement progressed, $2F_o - F_c$ and $F_o - F_c$ maps were examined with *CHAIN* (Sack, 1988), appropriate adjustments were made to main-chain and side-chain atoms and water molecules were added in accord with good electron density and the formation of hydrogen bonds to other atoms. At the end of the refinement, the residual was 0.190 for 21 837 data. However, a check of the geometry with *CNS* (Brunger *et al.*, 1998) showed that the restraints had been too loose.

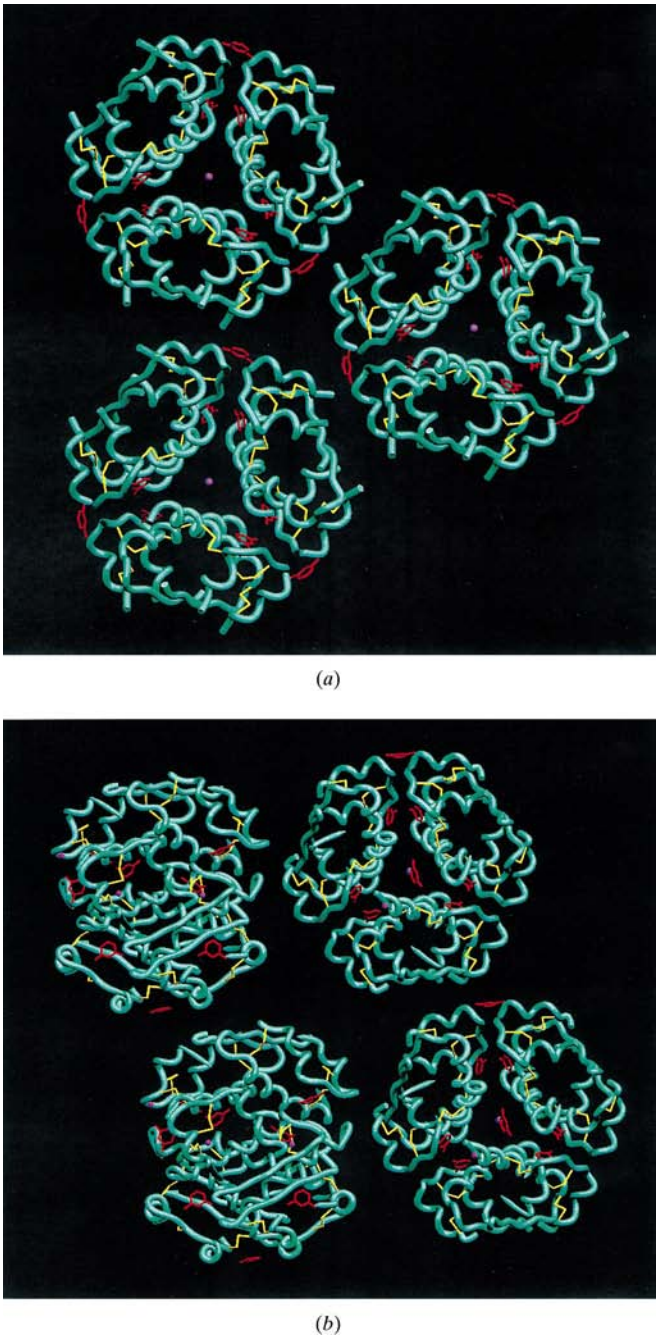


Figure 1
SETOR drawing (Evans, 1993) illustrating the packing of the hexamers in (*a*) the rhombohedral unit cell as viewed parallel to the *c* axis and (*b*) the monoclinic unit cell as viewed parallel to the (110) plane. Zinc ions are colored magenta and phenolic derivatives red.

The refinement was continued with *CNS* using all data, applying an overall anisotropic temperature-factor correction, a bulk-solvent correction and using maximum-likelihood torsion-angle dynamics, conjugate-gradient refinement or individual temperature-factor refinement (Rice & Brünger, 1994; Brunger *et al.*, 1998; Adams *et al.*, 1997; Pannu & Read, 1996). Bond distances between zinc ions and chloride ions and between zinc ions and $N^{\varepsilon 2}$ of the HisB10 residues were restrained. Approximately 10% of the data were reserved for cross validation and were not included in the refinements (Brünger, 1992). σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ maps as implemented in *CNS* (Read, 1986; Brünger *et al.*, 1997) were carefully examined at the end of each round of refinement and the geometry was monitored with *PROCHECK* (Laskowski *et al.*, 1993). Adjustments were made to main-chain and side-chain atoms, additional groups were added if necessary and water molecules were added in accord with good electron density and the formation of hydrogen bonds to existing atoms. Maximum-likelihood torsion-angle dynamics omit maps (Hodel *et al.*, 1992; Rice & Brünger, 1994; Read, 1986; Brünger *et al.*, 1997) were calculated to verify disordered side chains and other portions of the structure that were problematic. Towards the end of the refinement, a persistent 5σ $F_o - F_c$ peak was observed between two symmetry-related hexamers and adjacent to three protein O atoms. Since the only cations present are zinc and sodium, this peak was interpreted as a sodium ion. At the conclusion of the refinement, the final model consisted of 2339 protein atoms, eight resorcinol molecules, two zinc ions, two chloride ions, one sodium ion, 146 waters at unit occupancy, 27 waters with partial occupancy and nine disordered residues containing a total of 77 atoms. Residues which were modeled with two discrete conformations were ValB12.1, PheB25.1, GluA4.2, GlnA5.2, ValB12.4, SerA9.5, GluB13.6, AsnB3.6 and ArgB22.6.[1] The electron density for PheB25.3 and PheB25.4

[1] The decimal portion of the residue number, 0.1 to 0.6, refers to the monomer number in the hexamer and to A chains (*A*, *C*, *E*, *G*, *I* and *K*) and B chains (*B*, *D*, *F*, *H*, *J* and *L*), respectively, in the deposited coordinates.

Table 3
Data-refinement statistics for rhombohedral *m*-cresol (R-Crs), monoclinic *m*-cresol (M-Crs) and monoclinic resorcinol (M-Res) complexes with insulin.

	R-Crs	M-Crs	M-Res
No. of reflections	8552	25127	25784
Resolution range (Å)	19.72–1.78	20.95–1.9	44.75–1.90
R value	0.200	0.195	0.179
R_{free} value	0.266	0.235	0.218
Highest resolution shell			
Resolution range (Å)	1.89–1.78	2.02–1.90	2.02–1.90
No. of reflections	1173	3678	3664
Completeness (%)	88.2	93.9	93.2
R value	0.336	0.271	0.298
R_{free} value	0.371	0.333	0.307
Cross-validated estimate error (Å)	0.22	0.22	0.25
R.m.s. deviations from ideal			
Bond lengths (Å)	0.010	0.009	0.010
Bond angles (°)	1.37	1.38	1.31
Dihedral angles (°)	20.3	20.8	20.9
Improper angles (°)	0.90	0.86	0.84
Isotropic thermal model restraints			
Main-chain bonds (Å^2)	1.78	1.86	2.14
Main-chain angles (Å^2)	2.72	2.74	3.07
Side-chain bonds (Å^2)	2.36	2.59	3.25
Side-chain angles (Å^2)	3.41	3.82	4.67

was very weak and therefore these side chains were assigned occupation factors of 0.5. Owing to high thermal motion and disorder, the side chains of residues TyrA14.1, GluB21.1, LysB29.1, PheB1.2, LysB29.2, LysB29.3, LysB29.4, LysB29.5 and PheB1.6 could not be located. No electron density was found for the N-terminal residue, PheB1.3, or for the six C-terminal ThrB30 residues. Refinement statistics are given in Table 3.

2.3.2. Monoclinic *m*-cresol–insulin complex. The starting coordinates for the refinement of the *m*-cresol complex with insulin consisted of only the protein coordinates from the final refined structure of the resorcinol complex, as described above. 10% of the data were reserved for cross validation and were not used during the refinement (Brünger, 1992). The hexamer was subjected to rigid-body refinement in *CNS* (Brunger *et al.*, 1998; Pannu & Read, 1996; Adams *et al.*, 1997) using an anisotropic overall temperature correction and a bulk-solvent correction and yielded a residual of 0.363. $2F_o - F_c$ and $F_o - F_c\,\sigma_A$-weighted maps (Read, 1986; Brünger *et al.*, 1997) were calculated and clearly showed the positions of the pair of zinc and chloride ions as well as six *m*-cresol molecules. The refinement was continued as described above and bond distances involving zinc were restrained. As was observed in the resorcinol complex, a 5.5σ $F_o - F_c$ peak was interpreted as a sodium ion. At convergence, the final model consisted of 2355 protein atoms, seven *m*-cresol molecules, two zinc ions, two chloride ions, one sodium ion, 218 waters molecules with unit occupancy, 22 water molecules with partial occupancy and seven disordered side chains. No electron density was found for the side chains of any of the B-chain lysyl residues, as well as GluB21.1, GluB21.3, ThrB30.3 and PheB1.6; five of the six C-terminal ThrB30 residues could not be located. Three side chains were modeled in regions of very weak density and were assigned occupation factors of 0.5 and include TyrA14.1, TyrA14.6 and GluB21.6. Seven side chains were modeled in two discrete conformations and include GluB13.1, GlnA5.2, GluB13.2, ArgB22.4, SerA9.5, GluA4.6 and SerB9.6. One *m*-cresol molecule was observed in two orientations. Refinement statistics are given in Table 3.

2.3.3. Rhombohedral *m*-cresol–insulin complex. Although the structure of the rhombohedral phenol complex with insulin has been published (Smith & Dodson, 1992), the resolution of this structure was only 2.5 Å. For this reason, monomers 1 and 2 from the monoclinic resorcinol structure were optimized by a least-squares procedure (Smith, 1993) to the known structure of the rhombohedral insulin–phenol complex. The starting model consisted of an insulin dimer, two zinc ions, two chloride ions and two *m*-cresol molecules. 10% of the data were reserved for cross-validation and were not included as part of the working set of data (Brünger, 1992). Rigid-body refinement using *CNS* (Brunger *et al.*, 1998; Pannu & Read, 1996; Adams *et al.*, 1997), employing an overall anisotropic temperature factor and a bulk-solvent correction, converged at a residual of 0.378. The structure was refined using the protocol described above except that bonds involving zinc were not restrained. At convergence, the final model consisted of 757 protein atoms, two zinc ions, two chloride ions, three *m*-cresol molecules, 86 waters with unit occupancy, four partially occupied water molecules and three residues with disordered side chains. No electron density was found for the side chains of GluA4.1, TyrA14.1, GluB21.1, PheB1.2 and for the C-terminal residues LysB29.1, ThrB30.1, LysB29.2 and ThrB30.2. Residues whose side chains were modeled in two discrete conformations include ValB18.1, ThrA8.2 and ArgB22.2. Refinement statistics are given in Table 3.

3. Discussion

3.1. Hexamer packing

The hexamers in both the monoclinic and rhombohedral complexes adopt R_6 conformations. The packing of the hexamers in the rhombohedral cell (Fig. 1*a*) is quite different from that of the monoclinic unit cell (Fig. 1*b*). In the rhombohedral unit cell, the threefold axis of each hexamer is coincident with the crystallographic threefold axis and each hexamer in Fig. 1(*a*) is related to its neighbors by the rhombohedral centering condition. Because of high thermal motion and disorder, the C-terminal residues of both B chains, LysB29 and ThrB30, could not be located and these residues could potentially form interactions with a translationally related hexamer. In addition to many water-mediated hydrogen bonds between hexamers, two hydrogen bonds are observed between hexamers related by the *c* axis translation, PheB1.2 N to ThrA8.1 O (2.83 Å) and PheB1.1 N to ThrA8.2 O (3.08 Å). An additional hexamer–hexamer contact, related by the 3_1 axis, exists between ThrB27.1 $O^{\gamma1}$ and AsnA18.2 $O^{\delta1}$ (2.67 Å).

Table 4
Interhexamer hydrogen-bonded contacts (Å) in the monoclinic resorcinol (M-Res) and *m*-cresol (M-Crs) complexes with insulin.

Atoms		M-Res	M-Crs	Symmetry
TyrA14.1 O^{η}	GlnA5.5 $O^{\varepsilon1}$		2.77	$x, y, 1 + z$
TyrA14.1 O^{η}	GlnA5.5 $N^{\varepsilon2}$		3.02	$x, y, 1 + z$
AsnB3.1 $N^{\delta2}$	TyrA14.2 O	2.92	2.87	$1 - x, -\frac{1}{2} + y, -z$
AsnB3.1 $N^{\delta2}$	GluA17.2 $O^{\varepsilon1}$	2.84		$1 - x, -\frac{1}{2} + y, -z$
AsnB3.1 N	AsnA18.2 $O^{\delta1}$	3.08	2.96	$1 - x, -\frac{1}{2} + y, -z$
GlnB4.1 $O^{\varepsilon1}$	GlnA15.2 $N^{\varepsilon2}$		2.90	$1 - x, -\frac{1}{2} + y, -z$
GluA17.2 $O^{\varepsilon2}$	AsnB3.3 $N^{\delta2}$	2.88		$1 - x, \frac{1}{2} + y, -z$
AsnA18.2 $N^{\delta2}$	ThrA8.5 O	2.87	2.76	$1 - x, \frac{1}{2} + y, -z$
TyrA14.3 O	AsnB3.4 $N^{\delta2}$	3.05	2.92	$-x, -\frac{1}{2} + y, -z$
AsnA18.3 $O^{\delta1}$	ValB2.4 N	2.95	2.93	$-x, -\frac{1}{2} + y, -z$
AsnA18.3 $O^{\delta1}$	AsnB3.4 N	2.79	2.89	$-x, -\frac{1}{2} + y, -z$
AsnA18.3 $N^{\delta2}$	ThrA8.6 O		2.81	$-x, -\frac{1}{2} + y, -z$
AsnA21.3 OXT	PheB1.6 N	2.76		$-x, -\frac{1}{2} + y, -z$
AsnA21.3 OXT	AsnB3.6 $O^{\delta1}$	2.45		$-x, -\frac{1}{2} + y, -z$
AsnA21.3 OXT	GlnB4.6 $N^{\varepsilon2}$	2.57		$-x, -\frac{1}{2} + y, -z$
ArgB22.3 $N^{\eta2}$	AsnB3.6 $O^{\delta1}$	2.73	3.15	$-x, -\frac{1}{2} + y, -z$
GluA17.4 $O^{\varepsilon2}$	GlyA1.5 N	2.91	3.13	$x, y, 1 + z$
AsnA18.4 $N^{\delta2}$	ThrB27.5 $O^{\gamma1}$		2.88	$x, y, 1 + z$

In contrast, the interactions between hexamers in the monoclinic unit cell (Fig. 1*b*) are much more complicated. The local threefold axis of one hexamer is perpendicular to that of its symmetry-related mate and is nearly parallel to the (110) plane. More than half of the hydrogen-bonded contacts between hexamers involve residues A14.2–A18.2 and A14.3–A21.3. There are no main-chain/main-chain contacts, but 24 of the 36 hydrogen-bonded contacts involve side chains and 17 of these are asparagine side chains. A complete list of hydrogen-bonded contacts is given in Table 4.

An additional interaction in the crystal which has not been observed in insulin structures previously involves a sodium ion which adopts approximately octahedral coordination. The coordination sphere of the sodium ion consists of PheB1.5 O, AsnB3.5 $O^{\delta1}$, a symmetry-related AsnA21.2 OXT and three water molecules; sodium–oxygen distances range from 1.95 to 3.04 Å. Through these interactions, the sodium ion contributes to stabilizing the packing of the hexamers in the monoclinic crystal.

3.2. Monomer conformations

Five of the six ValB2 residues in the monoclinic structures and both of the ValB2 residues in the rhombohedral structure adopt an α-helical conformation, resulting in a continuous α-helix from PheB1 to CysB19. In contrast, ValB2.5 has φ/ψ torsion angles of $-85.1/+40.5^{\circ}$ and $-96.4/+40.8^{\circ}$ for the monoclinic resorcinol and *m*-cresol complexes, respectively, which displaces the PheB1.5 C^{α} by approximately 6 Å relative to the remainder of the B-chain α-helical N-termini. In this orientation, PheB1.5 O is able to coordinate to the sodium ion as described above and the resulting conformation is quite different from the R^{f} conformation (Ciszak *et al.*, 1995) observed in $T_3R_3^{f}$ structures.

The three structures reported here were superimposed in pairs by a least-squares procedure (Smith, 1993), minimizing the displacement of the backbone atoms of A11–A19 and B11–B19 of all six monomers. Additionally, the monoclinic *m*-cresol and resorcinol structures were superimposed on the monoclinic phenol structure (PDB code 1znj) and the rhombohedral *m*-cresol structure was superimposed on the AspB28 insulin mutant complexed with *m*-cresol (Whittingham *et al.*, 1998; PDB code 1zeh). The results shown in Table 5 reveal that the differences between any of the structures are small, with the monoclinic *versus* resorcinol being the smallest. The slightly larger displacements for the monoclinic *m*-cresol complexes are most likely to be a consequence of the presence of the hydrophobic methyl group, which induces displacements of water molecules within the phenolic binding site as well as small perturbations in the structure of the hexamers. The largest difference is observed for the B28 mutant insulin and is most likely to be a consequence of the replacement of proline by aspartate.

Table 5
Mean and root-mean-square displacements following a least-squares superposition of one hexamer onto another.

Values are given for backbone atoms (N_i, C_i^{α} and C_i) used in the least-squares procedure (A11–A19 and B11–B19 for all six monomers) as well as for all backbone atoms in the hexamers of monoclinic *m*-cresol–insulin (M-Crs), monoclinic resorcinol–insulin (M-Res), rhombohedral *m*-cresol–insulin (R-Crs), monoclinic phenol–insulin (M-Phn; PDB code 1znj) and rhombohedral AspB28 insulin mutant (B28B; PDB code 1zeh).

Hexamer	Mean	R.m.s.	Mean	R.m.s.
Types	Displ. LS atoms	Displ. LS atoms	Displ. All atoms	Displ. All atoms
M-Crs *versus* M-Res	0.285	0.372	0.310	0.401
M-Crs *versus* R-Crs	0.326	0.423	0.404	0.600
M-Res *versus* R-Crs	0.367	0.417	0.457	0.606
M-Crs *versus* M-Phn	0.330	0.402	0.349	0.446
M-Res *versus* M-Phn	0.188	0.218	0.231	0.308
R-Crs *versus* B28B	0.496	0.522	0.718	0.861

3.3. Phenolic derivatives

Each hexamer contains six phenolic derivatives in the six binding pockets. The phenolic hydroxy group forms a hydrogen bond to the O atom of CysA6 of each hexamer, with a mean distance of 2.61 (9) Å; a second hydrogen bond exists between the phenolic hydroxy group and the amino group of CysA11 with a mean distance of 2.87 (8) Å. In contrast to that observed for a series of mutant insulin structures (Berchtold & Hilgenfeld, 1999), phenolic hydrogen-bonded distances in trimer I were comparable to those in trimer II. In the case of resorcinol, the second hydroxy group interacts with a water molecule with an average hydrogen-bonded distance of 2.64 (9) Å; this water molecule in turn forms a hydrogen bond to the carbonyl O atom of CysA11, with a mean distance of 2.82 (12) Å. The local environment and electron density for a typical resorcinol and *m*-cresol molecule in the monoclinic structures are shown in Figs. 2(*a*) and 2(*b*), respectively.

Several additional phenolic derivative molecules were located in the three structures. In the resorcinol–insulin complex, one resorcinol molecule resides in the channel that leads from the zinc ion binding site to the surface of the

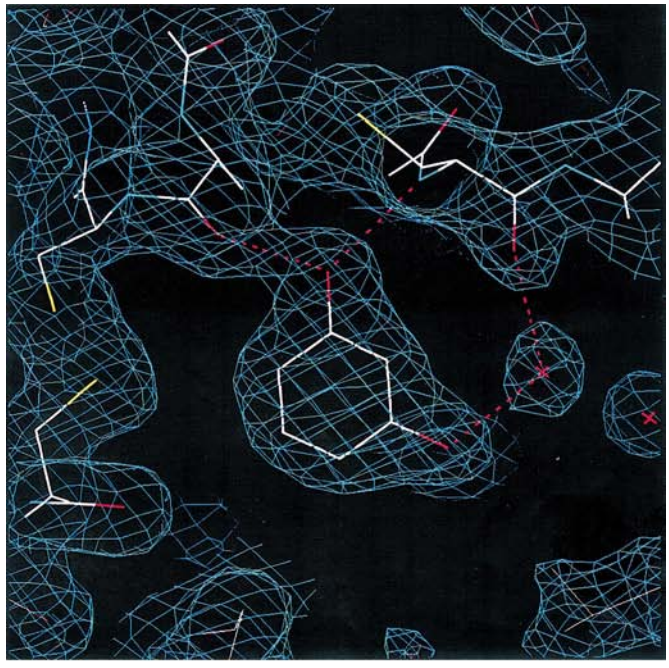
hexamer. Within this channel, the resorcinol molecule forms several hydrogen bonds to water molecules as well as to AsnB3.3 $O^{\delta1}$. The position of this resorcinol molecule is very similar to that reported for a phenol in the GlyA21, ArgB31, ArgB32 (Berchtold & Hilgenfeld, 1999), where the phenol molecule forms a hydrogen bond to the side chain of GluA17 in a neighboring hexamer. This difference in hydrogen bonding may be a consequence in part of the presence of an additional hydroxy group in resorcinol as well as to the presence of mutations to the insulin hexamer. The channel on the opposite side of the hexamer contains only water molecules. In the case of the two *m*-cresol–insulin complexes, both channels contain only water.

An *m*-cresol molecule was found to occupy a position beneath a pair of tyrosine side chains, A14.1 and A14.4, in the monoclinic structure, similar to that observed in the AspB28 insulin mutant complexed with *m*-cresol (Whittingham *et al.*, 1998). A resorcinol molecule in the monoclinic structure and an *m*-cresol molecule in the rhombohedral structure were also observed in this position, even though no electron density was present to model one of the tyrosine side chains in the pair, presumably because of disorder. In all three cases, the electron density for the phenolic derivative was weak and it is likely that the molecule is also disordered. Owing to the threefold symmetry, three such sites exist in the rhombohedral structure. In the complex with *m*-cresol in the monoclinic structure, the other two sites are empty and exposed to the solvent region, owing to a 180° rotation about the χ^1 torsion angle of both pairs of tyrosine side chains. In the complex with resorcinol, both other sites are also empty; three of the tyrosyl side chains adopt conformations similar to those in the *m*-cresol monoclinic structure, while the fourth side chain adopts a conformation similar to that of the rhombohedral structure. However, in this latter case the density for the side chain is poor and it is likely that the side chain is disordered.

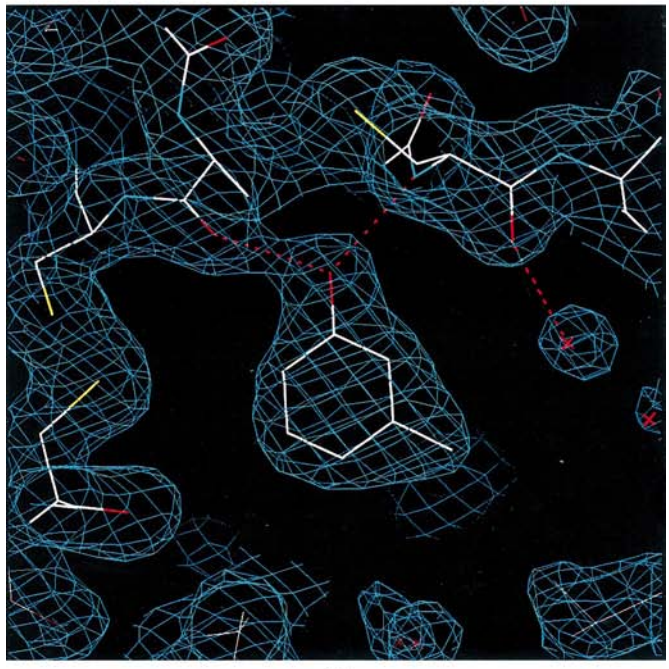
3.4. B13 glutamate

There is evidence to suggest that the GluB13 side chains play an important role in the T→R conformational transition. The mutation of GluB13 to Gln results in the spontaneous formation of a $T_3R_3^f$ hexamer in the absence of zinc or high chloride concentrations (Bentley *et al.*, 1992). In ultrahigh-resolution (1.0 Å) studies of T_6 insulin (Smith, Pangborn & Blessing, unpublished data), the GluB13 side chains are oriented such that the planes of the carboxyl groups are nearly parallel (Fig. 3*a*) to the crystallographic threefold axis. In this orientation, a contact of 2.50 Å exists between a pair of $O^{\epsilon2}$ atoms from two independent monomers, strongly suggesting the presence of a hydrogen bond. While one would expect a glutamic acid residue to be unprotonated at the pH at which the crystals were grown, it has been noted that the pK_a of GluB13 is significantly increased in the hexameric state (Kaarsholm *et al.*, 1990).

Following the conversion of one T-state trimer to R^f, producing a $T_3R_3^f$ hexamer (Ciszak & Smith, 1994), the T-state GluB13 carboxyl groups retain their relative orientation to the threefold axis while the R-state carboxyl groups are directed towards the threefold, shown in Fig. 3(*b*). In this orientation, the closest contact between a pair of GluB13 O atoms is 3.07 Å.



(*a*)



(*b*)

Figure 2
Electron density contoured at 1σ for (*a*) a resorcinol molecule in the monoclinic structure and (*b*) an *m*-cresol molecule in the monoclinic structure. Hydrogen bonds are illustrated as dashed red lines. The water molecule shown in the resorcinol complex (*a*) which bridges the resorcinol molecule and the carbonyl O atom of CysA11 is displaced in the *m*-cresol structure (*b*) by 1.67 Å where it is still capable of forming a hydrogen bond to the carbonyl O atom of CysA11.

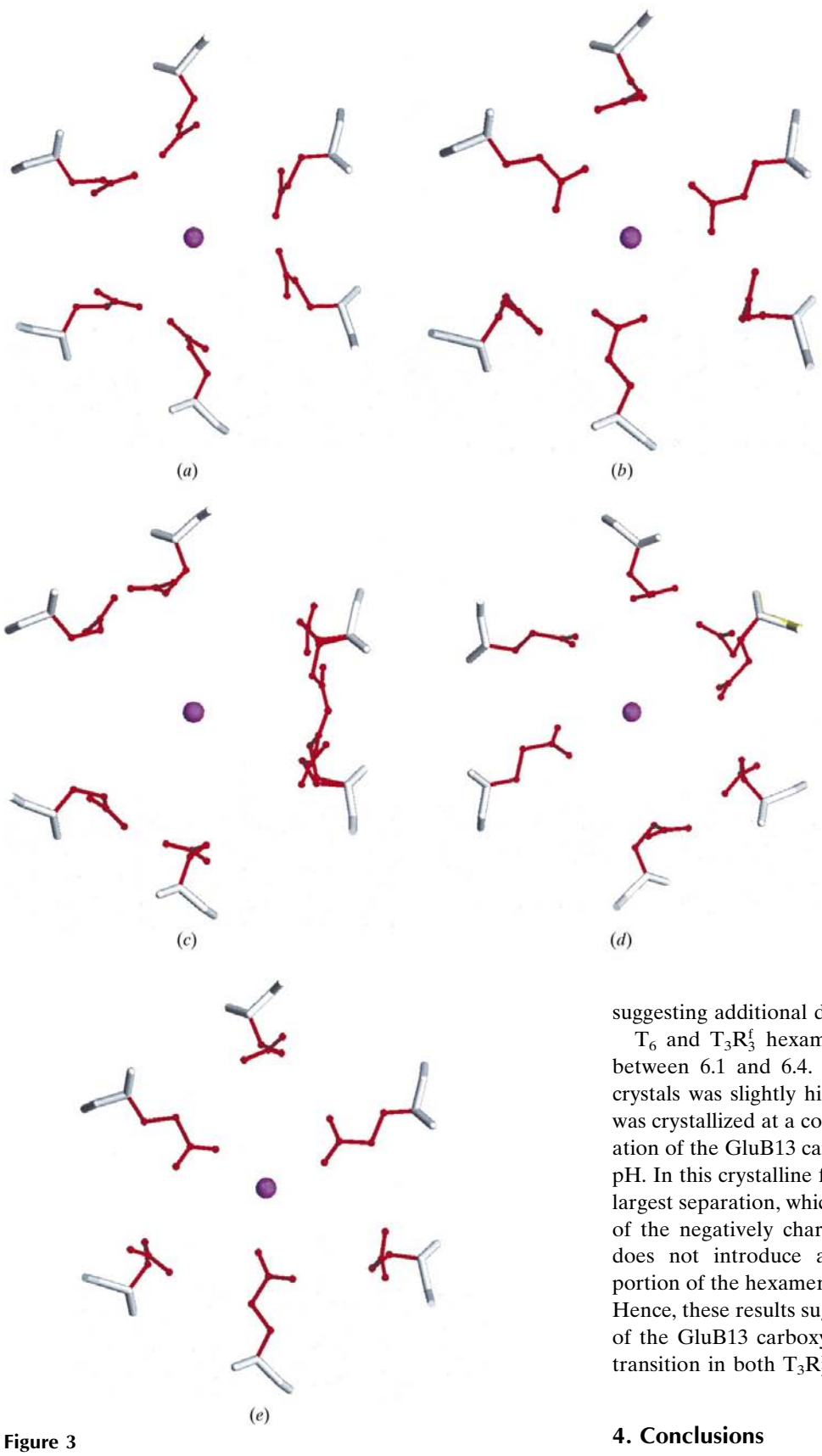


Figure 3
SETOR drawing (Evans, 1993) of B chain Glu13 residues in (*a*) T_6 human insulin, (*b*) $T_3R_3^f$ human insulin, (*c*) monoclinic *m*-cresol human insulin, (*d*) monoclinic resorcinol human insulin and (*e*) rhombohedral *m*-cresol human insulin.

Owing to the loss of crystallographic threefold symmetry in the monoclinic structures the situation is much more complicated, and while several of the GluB13 side chains were modeled in two discrete conformations, the appearance of the electron-density maps suggests that additional disorder is most likely present. In the *m*-cresol–insulin and resorcinol–insulin GluB13 conformations shown in Figs. 3(*c*) and 3(*d*), respectively, the planes of the carboxyl groups are again nearly parallel to the local threefold axis, and in this orientation the groups form hydrogen bonds to the carbonyl O atom and O^{γ} of SerB9 residues in an adjacent monomer. The closest contact between GluB13 O atoms is 3.84 Å for the resorcinol complex and 3.86 Å for the *m*-cresol complex.

The closest contact between GluB13 O atoms in the rhombohedral *m*-cresol complex is 4.36 Å and the overall arrangement of the side chains is similar to that observed for $T_3R_3^f$ hexamers (Fig. 3*e*). As was the case in the monoclinic structures, the electron density in the vicinity of the B13 side chains was less than optimal in spite of the crystallographic threefold symmetry, suggesting additional disorder and high thermal motion.

T_6 and $T_3R_3^f$ hexamers are typically crystallized at a pH between 6.1 and 6.4. While the pH for the monoclinic R_6 crystals was slightly higher (6.6), the rhombohedral R_6 form was crystallized at a considerably higher pH (8.5) and protonation of the GluB13 carboxyl groups would be unlikely at this pH. In this crystalline form the B13 carboxyl groups have the largest separation, which is a consequence of charge repulsion of the negatively charged side chains. The T→R transition does not introduce any steric constraints in the central portion of the hexamer where the B13 carboxyl groups reside. Hence, these results suggest the possibility that deprotonation of the GluB13 carboxyl groups is associated with the T→R transition in both $T_3R_3^f$ and R_6 hexamers.

4. Conclusions

The results presented here clearly show that *m*-cresol or resorcinol are capable of driving the T→R transition to completion to produce an R_6 hexamer in both monoclinic and

rhombohedral space groups. The binding of these phenolic derivatives in the phenol-binding site is accompanied by the typical hydrogen-bonding pattern previously observed between the phenolic derivative and the carbonyl O atom of CysA6 and the N atom of CysA11. In the case of resorcinol, an additional water-mediated hydrogen bond exists between the second hydroxy group and the carbonyl O atom of CysA11. One additional phenolic binding site beneath a pair of TyrA14 side chains is common to all three structures, while a second additional binding site in the zinc ion channel is only observed in the complex with resorcinol. Unlike the $T_3R_3^f$ hexamers, in which the disulfide torsion angles show a wide variation, these torsion angles in the R_6 structures are relatively constant and show no evidence of disorder. Differences in the side-chain orientations of the GluB13 residues in T_6, $T_3R_3^f$ and R_6 hexamers suggest that as a result of the T→R transition the carboxyl group is deprotonated, resulting in a larger separation of the charged groups.

We thank Lilly Research Laboratories for a generous gift of recombinant human insulin. This work was supported by NIH grant GM56829 and in part by a grant from the William G. McGowan Charitable Fund, Inc.

References

Adams, P. D., Pannu, N. S., Read, R. J. & Brünger, A. T. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5018–5023.

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. 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.

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.

Berchtold, H. & Hilgenfeld, R. (1999). *Biopolymers*, **51**, 165–172.

Blessing, R. H. (1997). *J. Appl. Cryst.* **30**, 421–426.

Bloom, C. R., Choi, W., Brzovic, P. S., Huang, J. J., Kaarsholm, N. C. & Dunn, M. F. (1995). *J. Mol. Biol.* **245**, 324–330.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Brünger, A. T., Adams, P. D. & Rice, L. M. (1997). *Structure*, **5**, 325–336.

Brzovic, P. S., Choi, W. E., Borchardt, D., Kaarsholm, N. C. & Dunn, M. F. (1994). *Biochemistry*, **33**, 13057–13069.

Choi, W. E., Brader, M. L., Aguilar, V., Kaarsholm, N. C. & Dunn, M. F. (1993). *Biochemistry*, **32**, 11638–11645.

Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D. & Smith, G. D. (1995). *Structure*, **3**, 615–622.

Ciszak, E. & Smith, G. D. (1994). *Biochemistry*, **33**, 1512–1517.

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

Evans, S. V. (1993). *J. Mol. Graph.* **6**, 244–245.

Finzel, B. C. (1987). *J. Appl. Cryst.* **20**, 53–55.

Hendrickson, W. A. & Konnert, J. H. (1980). *Computing in Crystallography*, edited by R. Diamond, S. Ramaseshan & K. Venkatesan, pp. 13.01–13.23. Bangalore, India: Indian Academy of Sciences.

Hodel, A., Kim, S.-H. & Brünger, A. T. (1992). *Acta Cryst.* A**48**, 851–859.

Kaarsholm, N. C., Havelund, S. & Hougaard, P. (1990). *Arch. Biochem. Biophys.* **283**, 496–502.

Kaarsholm, N. C., Ko, H. & Dunn, M. F. (1989). *Biochemistry*, **28**, 4427–4435.

Krüger, P., Gilge, G., Çabuk, Y. & Wollmer, A. (1990). *Biol. Chem. Hoppe-Seyler*, **371**, 669–673.

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

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Rice, L. M. & Brünger, A.T. (1994). *Proteins Struct. Funct. Genet.* **19**, 277–290.

Sack, J. S. (1988). *J. Mol. Graph.* **6**, 244–245.

Smith, G. D. (1993). *PROFIT. A Program for Orienting one Protein Molecule onto Another by a Least-Squares Procedure.* Hauptman–Woodward Medical Research Institute, Buffalo, NY, USA.

Smith, G. D. & Ciszak, E. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 8851–8855.

Smith, G. D., Ciszak, E. & Pangborn, W. A. (1996). *Protein Sci.* **5**, 1502–1511.

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

Whittingham, J. L., Chaudhurri, S., Dodson, E. J., Moody, P. C. E., Dodson, G. G. (1995). *Biochemistry*, **34**, 15553–15563.

Whittingham, J. L., Edwards, D. J., Antson, A. A., Clarkson, J. M. & Dodson, G. G. (1998). *Biochemistry*, **37**, 11516–11523.