X-ray Crystallographic Studies on Hexameric Insulins in the Presence of Helix-Stabilizing Agents, Thiocyanate, Methylparaben, and Phenol

Jean L. Whittingham,* Siddhartha Chaudhuri,‡ Eleanor J. Dodson, Peter C. E. Moody,§ and G. Guy Dodson

Department of Chemistry, University of York, Heslington, York Y01 5DD, England

Received June 23, 1995; Revised Manuscript Received August 25, 1995®

ABSTRACT: Three X-ray crystallographic studies have been carried out on pig insulin in the presence of three ligands, thiocyanate, methylparaben (methyl *p*-hydroxybenzoate), and phenol. In each case, rhombohedral crystals were obtained, which diffracted to 1.8, 1.9, and 2.3 Å, respectively. Each crystal structure was very similar to that of 4-zinc pig insulin, which was used as a starting model for PROLSQ refinement (Collaborative Computational Project, Number 4, 1994). The *R* factors for the refined structures of thiocyanate insulin, methylparaben insulin, and phenol insulin were 19.6, 18.4, and 19.1, respectively. Each crystal structure consists of $T_3R_3^f$ insulin hexamers with two zinc ions per hexamer. In the R_3^f trimer of the thiocyanate insulin hexamer, one thiocyanate ion is coordinated to the zinc ion on the hexamer 3-fold axis, but there is no evidence of zinc ion binding in the off-axis zinc ion sites seen in the 4-zinc pig insulin structure. In the methylparaben insulin and phenol insulin hexamers, the phenolic ligands are bound at the dimer−dimer interfaces in the R_3^f trimers in a manner similar to that of phenol in R_6 phenol insulin. The binding of methylparaben appears to make the hexamer more compact by drawing the A and the B chains closer together in the binding site. In all three structures presented herein, the conformations of the first three residues of the B chain in the R_3^f trimer are extended rather than α-helical, as is seen in R_6 phenol insulin. The energetics of ligand binding in the insulin hexamer are discussed.

Insulin is a peptide hormone, $M_r = 5800$, which promotes transport of glucose across cell membranes of muscle and adipose tissue. Owing to insulin's crucial metabolic role associated with the disease diabetes mellitus, its structure has been, and continues to be, studied extensively both in crystals and in solution. The insulin molecule consists of an A chain of 21 amino acids and a B chain of 30 amino acids. In the crystal structure of 2-zinc pig insulin, the A chain forms two α-helices spanning residues A1−A8 and residues A13−A19, joined by a loop from A9 to A12 (Blundell et al., 1972; Baker et al., 1988). The last two C-terminal residues of the A chain adopt an extended conformation. The B chain consists of an α-helix between residues B9 and B19 and two extended chain regions from B1 to B8 and from B22 to B30. The two chains are covalently linked by disulfide bridges at A7−B7 and A20−B19, and there is an intrachain disulfide bridge joining A6 and A11. Insulin has a quaternary structure in which the monomers associate to form dimers and then hexamers. The 2-zinc insulin hexamer comprises three crystallographically equivalent dimers whose association depends largely on the coordination of two zinc ions located on the 3-fold axis of the hexamer. Each zinc ion has an octahedral coordination sphere consisting of three B10 His side chains, one from each dimer, and three water molecules.

Native insulin has been crystallized in two other hexameric forms which owe their differences to the compositions of the crystallization solutions. While each of these structures is essentially very similar to 2-zinc insulin, there are significant differences at the B chain N-termini and in the metal ion coordination (Figure 1). In the 4-zinc pig insulin hexamer,[1] obtained with the addition of 6% NaCl, the three B chain N-termini in one-half of the hexamer adopt α-helical conformations (Bentley et al., 1976, 1992). The α-helices are stabilized by van der Waals contacts between the side chains of B6 Leu and a chloride ion coordinated to the zinc ion on the hexamer 3-fold axis. This arrangement provides one additional zinc ion binding site at each of the dimer−dimer interfaces, the side chain of residue B10 His assuming two alternative conformations to coordinate to zinc ions either on or off the hexamer 3-fold axis. The other zinc ion on the hexamer 3-fold axis is octahedrally coordinated, similar to those in 2-zinc insulin. In the equivalent human insulin structure (Ciszak & Smith, 1994), there are no off-axis zinc ions, and hence, the B10 His side chains are not disordered.

In the presence of phenol, which is used as a clinical preservative, a third type of insulin hexamer is obtained in which all six B chain N-termini are α-helical (see Figure 1; Derewenda et al., 1989; Smith & Dodson, 1992). Accordingly, there are two off-axis binding sites at each dimer−dimer interface. In this case, each site contains a phenol molecule at van der Waals contact with an adjacent B5 His side chain. In addition, the hydroxyl group of the phenol ligand is hydrogen bonded to the main chain carbonyl oxygen of residue A6 and the main chain amide group of residue A11. The van der Waals contacts to B5 His and the hydrogen bonds to the A chain of the adjacent dimer serve

* Author to whom correspondence should be addressed.
‡ Present address: RSIC, Bose Institute, Calcutta 700 009, India.
§ Present address: Department of Biochemistry, University of Leicester, University Road, Leicester LE1 7RH, U.K.
® Abstract published in *Advance ACS Abstracts,* October 1, 1995.

[1] Early experiments suggested that there were four zinc ions per hexamer, but in fact, the crystal structure analyses revealed that these sites are variably occupied (Bentley et al., 1976; Smith et al., 1984; Ciszak & Smith, 1994).

0006-2960/95/0434-15553$09.00/0

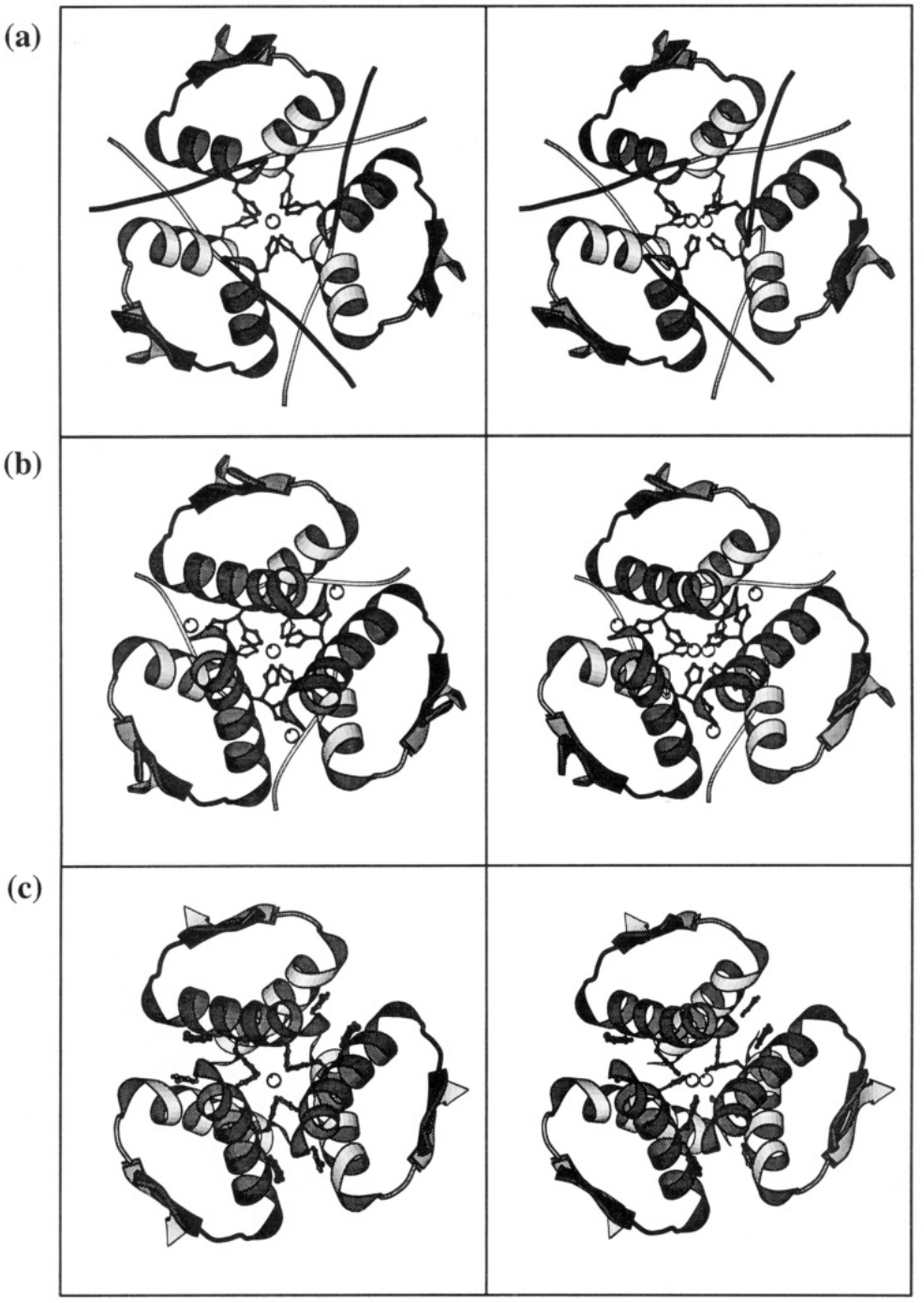

FIGURE 1: Stereographic representations of (a) 2-zinc pig insulin (T_6), (b) 4-zinc pig insulin ($T_3R_3^f$), and (c) rhombohedral phenol insulin (R_6) hexamers viewed approximately parallel to the hexamer 3-fold axis. For clarity, only the B chains are shown, and in each dimer, the two B chains are shaded differently. The positions of the zinc ions (white spheres), B10 His side chains, and phenol ligands (ball and stick) are shown. Notice in (a) that the B10 His side chains coordinate only to two zinc ions on the 3-fold axis, whereas in (b), the B10 His side chains are disordered in one trimer, coordinating zinc ions either on the 3-fold axis or at the off-axis positions between the dimers. In (c), there are two phenol molecules bound at each dimer−dimer interface. These representations were made using MOLSCRIPT (Kraulis, 1991).

to link the dimers together, adding to hexamer stability. This hexamer contains two zinc ions located on the hexamer 3-fold axis. Both are tetrahedrally coordinated to three B10 His side chains and a chloride ion, similar to that in one trimer of 4-zinc pig insulin.

To distinguish the three kinds of hexamers, Kaarsholm et al. (1989) devised a nomenclature in which the terms T and R are used to describe the insulin monomer with B1−B8 extended (unliganded) and α-helical (liganded), respectively. Thus, the 2-zinc, 4-zinc, and phenol insulin hexamers are more systematically named T_6, T_3R_3, and R_6. In addition, the term R^f is used to distinguish the "frayed" B chain N-terminal α-helix, in which residues B1−B3 are in an extended conformation, from that which is fully α-helical and termed R (Ciszak et al., 1995). Herein after, "$T_3R_3^f$" will be used to describe the specific hexamer conformation as seen in a crystal structure, whereas "T_3R_3" will be used to describe the hexamer conformation in more global terms or when the conformations of residues B1−B3 are uncertain. It is useful to consider the hexamer as consisting of two

Table 1: Crystallization Conditions for Thiocyanate Insulin (scn), Methylparaben Insulin (mpb), and $T_3R_3^f$ Phenol Insulin (phn), Adapted from Tolley (1987)

additive	scn	mpb	phn
native (pig) insulin (mg)	10	7	7
0.02 M HCl (mL)	1.0	1.0	1.0
0.15 M zinc acetate (mL)	0.15	0.15	0.15
calcium acetate (mg)	3.6		
acetone (mL)	0.3		
0.2 M trisodium citrate (mL)	0.5	0.5	0.5
0.25% methylparaben (mL)[a]		0.4	
0.25% phenol (mL)[a]			0.4
NaCl (mg)		120	120
KSCN (mg)	40		
pH	6.1−6.5	5.2−5.4	5.2−5.5

[a] Solutions of these ligands were made in acetone; however, their concentrations are given as weight of ligand per equivalent volume of water. Hence, a 0.25% phenol solution contains 0.25 g of phenol per 100 mL of acetone.

trimers, each trimer containing one zinc ion on the hexamer 3-fold axis (Krüger et al., 1990).

Solution studies on the insulin hexamer (Renscheidt et al., 1984; Wollmer et al., 1987; Thomas & Wollmer, 1989; Roy et al., 1989; Brader et al., 1991; Gross & Dunn, 1992; Choi et al., 1993) demonstrate that it exists in a state of dynamic equilibrium according to eq 1, which can be shifted from the T_6 state to the T_3R_3 state by the addition of inorganic anions and to the R_6 state by the addition of phenol-like molecules.

$$T_6 \rightleftharpoons T_3R_3 \rightleftharpoons R_6 \quad (1)$$

The aim of this study is to extend these investigations. Firstly, we were curious to know whether thiocyanate, which is higher in the Hofmeister series than chloride (Riès-Kautt & Ducruix, 1991), has the ability to produce R_6 hexamers in crystals, although neither chloride nor thiocyanate can produce R_6 hexamers in solution. Secondly, it is clear from the crystal structure that the phenol binding pocket in the phenol insulin hexamer is capable of accommodating bulkier ligands. Methylparaben (methyl *p*-hydroxybenzoate), which is also used as a clinical preservative, appeared to represent an ideal candidate with which to explore the ligand binding potential of the local protein envelope. Methylparaben is large enough to fill the phenol binding pocket, and its para ester substituent group might realize additional interactions with surrounding polar protein side chains.

MATERIALS AND METHODS

Crystallization. Crystallization experiments were based on well documented protocols using the batch crystallization technique (Schlichtkrull, 1958; Harding et al., 1966). Pure pig insulin was donated by Novo Nordisk A/S (Denmark), and Analar grade chemical reagents were purchased from Aldrich Chemical Co. Ltd. and Fisons Scientific Apparatus. For each crystallization experiment, an insulin solution was prepared by dissolving the protein in 0.02 M hydrochloric acid prior to adding the other crystallization components in the order given in Table 1. The insulin solution was maintained at approximately 50 °C and the pH adjusted to about pH 8.0 to ensure complete solution. The pH was subsequently lowered to within a range given in Table 1 using either 1 M HCl or NaOH. Warm 2 mL samples were taken at different pH values within the range given and filtered through cellulose nitrate membranes into prewarmed 5 mL test tubes. These were placed in an insulated 50 °C water bath for a one-week, slow-cooling period. The most favorable crystallization conditions are shown in Table 1. All crystals grew in the rhombohedral space group *R*3, having cell dimensions similar to those of 4-zinc pig insulin (Table 2).

Table 2: Cell Dimensions for Thiocyanate Insulin, Methylparaben Insulin, and $T_3R_3^f$ Phenol Insulin[a]

insulin	space group	cell dimensions	hexamer state
thiocyanate	*R*3	$a = b = 80.64$ Å $c = 37.52$ Å	$T_3R_3^f$
methylparaben	*R*3	$a = b = 80.38$ Å $c = 37.00$ Å	$T_3R_3^f$
$T_3R_3^f$ phenol	*R*3	$a = b = 80.70$ Å $c = 37.80$ Å	$T_3R_3^f$
2-Zn (pig)[b]	*R*3	$a = b = 82.50$ Å $c = 34.00$ Å	T_6
4-Zn (pig)[c]	*R*3	$a = b = 80.70$ Å $c = 37.60$ Å	$T_3R_3^f$
tylenol (pH 5.6)[d]	*R*3	$a = b = 80.88$ Å $c = 37.60$ Å	$T_3R_3^f$
phenol[e]	*R*3	$a = b = 79.92$ Å $c = 40.39$ Å	R_6

[a] Those of native 2-zinc pig insulin, 4-zinc pig insulin, and rhombohedral phenol insulin are included for comparison. [b] From Baker et al. (1988). [c] From Bentley et al. (1992). [d] From Smith and Ciszak (1994). [e] From Smith and Dodson (1992).

Data Collection and Data Processing. For both methylparaben insulin and thiocyanate insulin, X-ray diffraction data were collected from a single crystal using a Rigaku 4-circle diffractometer (AFC5R, CuKα rotating anode X-ray source, $\lambda = 1.54$ Å, 50 kV, 100 mA). At the onset of data collection, a psi scan was performed, and the unit cell and orientation matrix were determined from 10−20 reflections in the 2θ range 30.0−30.3°. Measurements were made using ω scans, with a scan width of $0.63 + 0.3 \tan\theta$ degrees and a scan speed of 8° per minute. The data were collected in 2θ shells, radiation damage to the crystal being monitored using three standard reflections. For thiocyanate insulin, the whole of the unique set plus two wedges of data for internal scaling were collected. However, for methylparaben insulin, only the unique data set was collected, owing to limited apparatus time. There was little radiation damage to either crystal in the course of data collection; the crystal of methylparaben insulin diffracted to 1.9 Å spacing, and that of thiocyanate insulin diffracted to 1.8 Å spacing. The raw intensities were corrected for Lorentz, polarization, absorption, and decay effects and then merged using the CCP4 Suite (Collaborative Computational Project, Number 4, 1994). For the methylparaben insulin data set, there were 6620 observations, yielding 6095 unique reflections with an R_{merge}[2] value of 0.03. The thiocyanate insulin data set consisted of 11 175 observations, giving rise to 7807 unique reflections with an R_{merge} value of 0.04.

For the $T_3R_3^f$ phenol insulin crystal, a 2.3 Å data set was collected on a Xentronics area detector mounted on a rotating copper anode X-ray source ($\lambda = 1.54$ Å, 50 kV, 100 mA). A total of 720 frames (each 0.25°) were collected at a 2θ angle of 15°, giving a total of 180° of data. The exposure time per frame was 160 s. The data were measured using XDS (Kabsch, 1988) and merged using AGROVATA (Collaborative Computational Project, Number 4, 1994). The

[2] The value of the merging *R* factor between equivalent measurements, *I*, of the same reflection; $R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$.

Table 3: $\langle I/\sigma I\rangle$ per Resolution Shell for the Thiocyanate Insulin (scn), Methylparaben Insulin (mpb), and $T_3R_3^f$ Phenol Insulin (phn) Data Sets

scn			mpb			phn		
D_{min}	$\langle I/\sigma I\rangle$	fract	D_{min}	$\langle I/\sigma I\rangle$	fract	D_{min}	$\langle I/\sigma I\rangle$	fract
5.64	11.96	0.976	5.95	75.61	0.980	7.17	44.03	0.967
4.01	9.30	1.000	4.23	56.81	1.000	5.11	35.95	0.992
3.28	7.98	1.000	3.46	43.27	0.998	4.18	34.62	0.993
2.84	7.00	0.996	3.00	23.84	0.990	3.63	32.97	1.000
2.54	6.66	1.000	2.68	16.79	0.979	3.25	29.52	1.000
2.32	6.40	0.995	2.45	11.54	0.971	2.97	25.25	0.991
2.15	5.86	0.996	2.27	8.15	0.960	2.75	21.16	0.993
2.01	5.46	0.994	2.12	6.25	0.951	2.57	17.73	0.989
1.90	3.75	0.986	2.00	5.12	0.922	2.42	14.37	0.977
1.80	2.54	0.512	1.90	4.32	0.281	2.30	10.68	0.848
Totals	6.08	0.924	Totals	19.04	0.863	Totals	23.53	0.969

data set consisted of 4058 unique reflections merged from 8990 independent observations. The R_{merge} value for this data set was 0.03. More statistics for the three data sets are listed in Table 3.

Refinement. The starting model for all three refinements was the essentially isomorphous 4-zinc pig insulin structure. At the beginning of each refinement, water molecules and zinc ions were removed from the coordinate file and atomic temperature factors were set to average values determined from Wilson plots (Wilson, 1949). In each case, the initial R factor was in the range of 30−35%. Refinement consisted of rounds of PROLSQ (Collaborative Computational Project, Number 4, 1994) using data between 10.0 Å and the upper limit of resolution, with target rms values listed in Table 4. Between refinement cycles, $2F_o - F_c$ and $F_o - F_c$ maps were calculated using all the data, and water molecules, zinc, chloride, and thiocyanate ions were built into the structures using the program FRODO (Jones, 1978). Thereafter, the zinc ion coordination distances to the NE atoms of B10 His were restrained to 2.0 Å. For the methylparaben insulin structure, a model of methylparaben, assembled using the program QUANTA (MSI), was fitted into positive electron density in the binding site occupied by phenol in monoclinic phenol insulin. Toward the conclusion of refinement, the protein geometry was analyzed using PROCHECK (Laskowski et al., 1993). Refinement statistics are summarized in Table 4. The protein coordinates and structure factors for thiocyanate insulin, methylparaben insulin, and $T_3R_3^f$ phenol insulin have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY (references 2TCI, 3MTH, 1MPJ).

Comparisons. The structure comparisons which follow were made by least-squares aligning hexamers on residues B9−B19, which in all structures form a very similar α-helix, using LSQKAB (Collaborative Computational Project, Number 4, 1994).

RESULTS

Thiocyanate Insulin. To probe whether it is possible to obtain R_6 hexamers in the absence of a phenolic effector, insulin crystals were grown in the presence of thiocyanate, a more powerful promoter of the T → R transition than chloride (De Graaff et al., 1981). The crystallization protocol for 4-zinc pig insulin was modified in various ways using a variety of thiocyanate concentrations and the inclusion of calcium ions to stabilize the hexamers (Xiao, 1990). Different temperatures of crystallization were also explored. Despite these attempts to obtain R_6 hexamers, the resulting crystals were always isomorphous with those of 4-zinc pig insulin (Table 2). The most suitable conditions for the crystallization of thiocyanate insulin are shown in Table 1.

Table 4: Geometrical Restraints, Root Mean Square Deviations (rms Δ), and Other Statistics for the Refined Structures of Thiocyanate Insulin (scn), Methylparaben Insulin (mpb), and $T_3R_3^f$ Phenol Insulin (phn), Containing Output from PROLSQ (Collaborative Computational Project, Number 4, 1994) and PROCHECK (Laskowski et al., 1993)

parameter	target rms	rms Δ for scn	rms Δ for mpb	rms Δ for phn
bond lengths (1−2) (Å)	0.02	0.01	0.01	0.02
angle related distances (1−3) (Å)	0.04	0.03	0.04	0.05
intraplanar distances (1−4) (Å)	0.05	0.03	0.03	0.05
planar groups (Å)	0.02	0.01	0.01	0.01
chiral volumes (Å^3)	0.10	0.07	0.08	0.08
van der Waals' distances:				
single torsion contacts (Å)	0.30	0.18	0.16	0.19
multiple torsion contacts (Å)	0.30	0.27	0.26	0.27
possible hydrogen bonds (Å)	0.30	0.21	0.18	0.20
torsion angles				
planar (0, 180°)	20.00	1.93	2.21	2.54
staggered (60°/120°)	20.00	20.37	22.55	21.07
orthonormal (±90°)	20.00	15.87	15.11	20.47
thermal factors				
main chain bond (1−2) (Å^2)	3.00	1.92	2.21	2.72
main chain angle (1−3) (Å^2)	3.50	3.17	3.80	4.60
side chain bond (Å^2)	3.00	2.33	2.48	2.80
side chain angle (Å^2)	3.50	3.62	3.95	4.61
final R factor[a] (10 Å cutoff) (%)		19.8	18.4	18.9
no. of protein atoms (excluding disordered side chains)		767	750	775
no. of water molecules		107	75	60
average temperature factors				
main chain (rms) (Å^2)		23.00 (1.46)	26.27 (1.72)	29.54 (2.36)
side chain (rms) (Å^2)		30.32 (1.86)	28.01 (2.16)	35.92 (2.10)
residues in allowed regions of Ramachandran plot (%)		100	100	100

[a] Crystallographic R factor $= \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure amplitudes, respectively.

The thiocyanate insulin hexamer has a $T_3R_3^f$ conformation and contains two zinc ions located on the 3-fold symmetry axis. In the R_3^f trimer, a thiocyanate ion is coordinated to the zinc ion, instead of the chloride ion seen in the 4-zinc pig insulin structure (Figure 2). The thiocyanate ligand is sharply defined in the electron density, occupying a hydrophobic channel between three symmetry-related B6 Leu side chains. The Zn−NCS distance, which was not restrained during refinement, is 1.7 Å. For comparison, the Zn−NCS bond distance in the crystal structure of $[N(CH_2CH_2NH_2)_3(NCS)Zn](SCN)$ is 2.0 Å (Andreetti et al., 1969). The zinc ion coordination in the T_3 trimer of the thiocyanate insulin hexamer is also tetrahedral, the coordinating ligand, however, being a water molecule rather than a thiocyanate ion. There is some residual electron density near the water peak which suggests that a small proportion of the zinc ions in this site may be octahedrally coordinated. There is no evidence for calcium ions in the electron density maps even though they were included in the crystallization solution.

Unlike 4-zinc pig insulin, and more akin to $T_3R_3^f$ human insulin (Ciszak & Smith, 1994), the off-axis binding sites in the R_3^f trimer of thiocyanate insulin contain water mol-

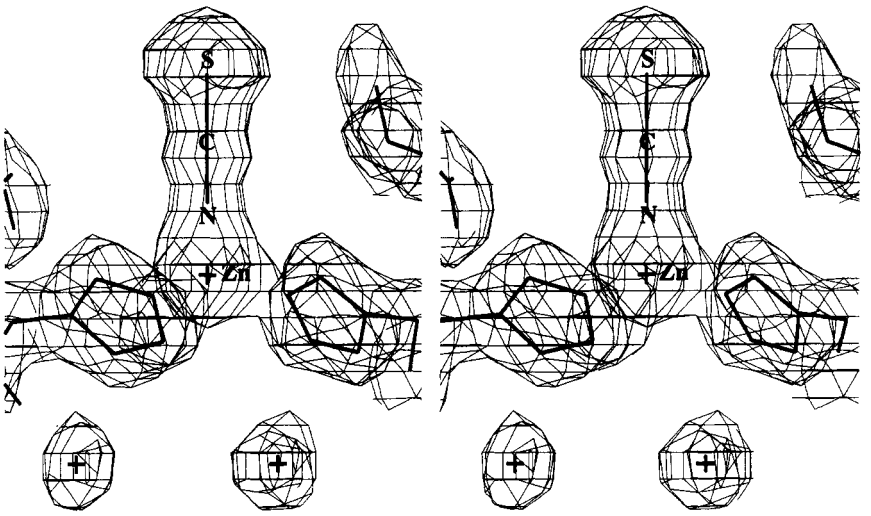

FIGURE 2: Stereoview of a $2F_o - F_c$ electron density map contoured at a level of 1σ, showing a thiocyanate ion coordinated to the zinc ion in the R_3^f trimer of thiocyanate insulin. The Zn−NCS bond distance is 1.7 Å.

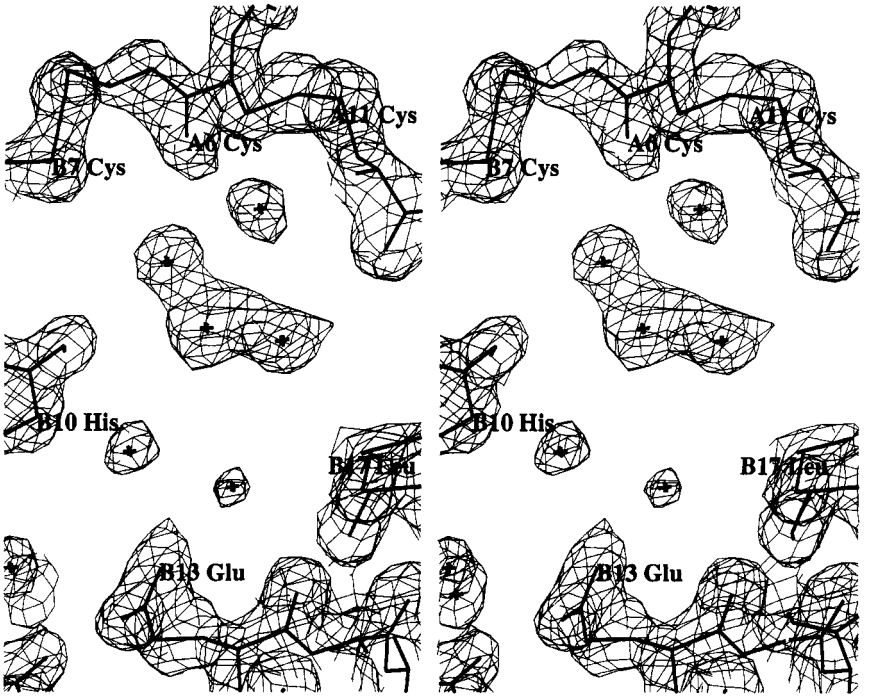

FIGURE 3: Stereoview of a $2F_o - F_c$ electron density map contoured at a level of 1σ, showing water molecules in the off-axis binding site at the dimer−dimer interface of the R_3^f trimer of thiocyanate insulin. In 4-zinc pig insulin, this site contains a coordinated zinc ion.

ecules rather than coordinated zinc ions (Figure 3). In addition, residues B1−B3 have torsion angles corresponding to the β-strand region of the Ramachandran plot, rather than α-helical as seen in R_6 phenol insulin (Derewenda et al., 1989; Smith & Dodson, 1992). A hydrogen bond between the side chain atom B3 OD1 and the main chain atom B5 N ensures that all of the B1−B8 α-helical hydrogen bonds are maintained. This feature also appears in $T_3R_3^f$ human insulin, tylenol insulin (Smith & Ciszak, 1994), and phenol/LysPro human insulin (Ciszak et al., 1995).

Methylparaben Insulin. Methylparaben (methyl *p*-hydroxybenzoate) was cocrystallized with insulin to investigate the capacity of the phenol binding site to accommodate phenol derivatives. A very low concentration of methylparaben had to be used in these crystallization experiments, owing to its limited solubility in aqueous solution. The methylparaben insulin crystals are isomorphous with those of 4-zinc pig insulin (Table 2). The methylparaben insulin hexamer, which has a $T_3R_3^f$ conformation like thiocyanate insulin, contains two zinc ions on the hexamer 3-fold axis with identical tetrahedral coordination spheres comprising three B10 His side chains and a chloride ion. In each of the three off-axis binding sites between the dimers in the R_3^f trimer, there is a methylparaben molecule (Figure 4) which makes similar hydrogen-bonding contacts to phenol in R_6 phenol insulin (Derewenda et al., 1989; Smith & Dodson, 1992). In addition, methylparaben makes van der Waals contacts with residues B5 His, B10 His, and B11 Leu of the R^f state monomer of one dimer and B13 Glu and B17 Leu of the T state monomer of the adjacent dimer. The carbonyl oxygen of methylparaben is hydrogen-bonded to a water molecule, which in turn is within hydrogen bonding distance of the side chains of residues B9 Ser, B10 His, and B13 Glu

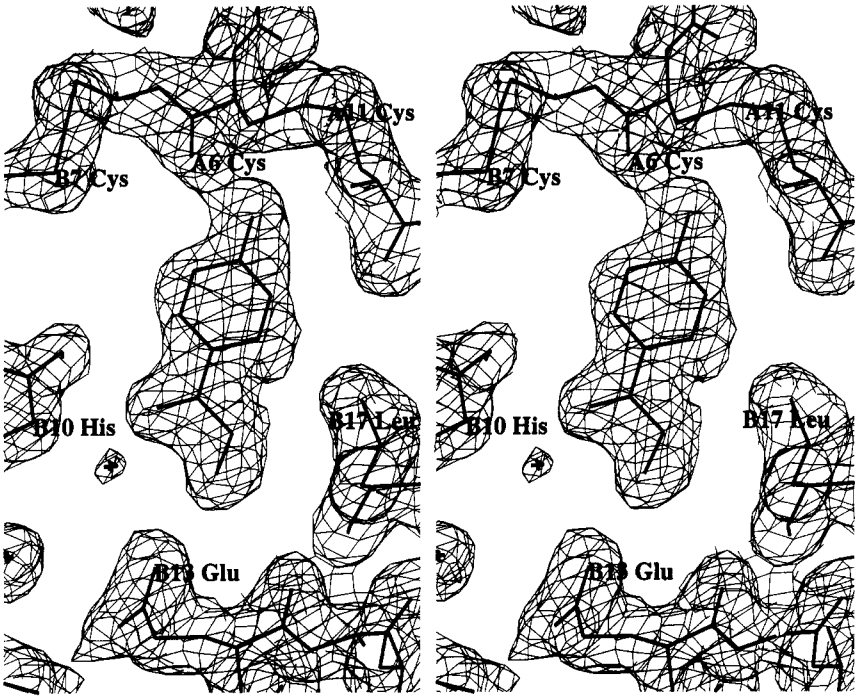

FIGURE 4: Stereoview of a $2F_o - F_c$ electron density map contoured at a level of just below 1σ, showing a methylparaben molecule bound in the R_3^f trimer of methylparaben insulin. A water molecule close to the ester substituent of methylparaben mediates hydrogen bonding between the ligand and some of the lower site residues.

of one dimer and B13 Glu of the adjacent dimer. As the electron density for this water molecule is weak and its temperature factor is relatively high, it is assumed that this water either is very mobile or occupies the site only partially. By contrast, the strong electron density for the methylparaben molecule indicates that it is well-ordered.

The methylparaben insulin hexamer is generally similar to that of 4-zinc pig insulin. However, a comparison of the methylparaben binding site with the off-axis zinc binding site in 4-zinc pig insulin shows that the relative positions of the A and B chains are slightly different in the R_3^f trimers of the two structures (Figure 5a). There is an A chain displacement, which measures approximately 0.9 Å, accompanied by the displacement of the side chains of B17 Leu, B10 His, and B5 His (not shown), owing to the binding of methylparaben. In a similar comparison with one of the phenol binding sites in R_6 monoclinic phenol insulin, the A chain displacement increases to about 1.7 Å (Figure 5b). The side chain conformations of B13 Glu, B17 Leu, and B5 His also appear to be affected by the binding of methylparaben, the position of which is approximately 1.5 Å lower than phenol in the binding site. Comparisons with the other five crystallographically independent phenol binding sites in monoclinic phenol insulin show similar results.

Tylenol (4′-hydroxyacetanilide) is a para-substituted phenol derivative like methylparaben. When methylparaben insulin is compared with tylenol insulin (Smith & Ciszak, 1994), the $T_3R_3^f$ hexamer conformations are found to be very similar. The positions of the methylparaben and tylenol ligands in the binding site are a little different, and there is a small displacement of the A chain (Figure 9). Zinc ion coordination in the R_3^f trimers of the two structures is identical. However, in the T_3 trimer, the zinc ion in the tylenol insulin hexamer is either tetrahedral (pH 5.6) or octahedral (pH 6.4); in each case, the coordination sphere is completed by water molecules.

$T_3R_3^f$ Phenol Insulin. The methylparaben experiment described above led to a number of unexpected observations, including a $T_3R_3^f$ hexamer conformation and tetrahedral zinc−chloride complex in the T_3 trimer. A key difference between the present studies and earlier phenol insulin studies in which R_6 hexamers were obtained (Derewenda et al., 1989; Smith & Dodson, 1992) is the concentration of the phenolic ligand used in, and the pH of, crystallization. To investigate the extent to which these parameters account for the structural differences, phenol was cocrystallized with insulin under the same conditions as those used in the methylparaben insulin study (Table 1).

At a phenol concentration of 0.25% (Table 1), the phenol insulin hexamer is in the $T_3R_3^f$ state. The two zinc ions which are located on the hexamer 3-fold axis appear to have the same tetrahedral coordination as in the methylparaben insulin hexamer. The temperature factors for the zinc ion and the chloride ligand in the R_3^f trimer are 15.7 and 10.3 Å^2, respectively, while those for the zinc ion and the chloride ion in the T_3 trimer are 19.4 and 36.7 Å^2, respectively. The higher temperature factor for the chloride ion in the T_3 trimer is accompanied by a slightly smaller electron density peak, which was originally modeled as a water molecule. The temperature factor for this water molecule refined to an unacceptably low value, however, and in the later stages of refinement, a chloride ion was introduced. Nevertheless, the zinc−chloride distance has refined to 2.1 Å, and we cannot rule out the possibility that a proportion of the chloride sites is occupied by water molecules.

A phenol molecule is bound at each dimer−dimer interface in the R_3^f trimer of the $T_3R_3^f$ phenol insulin hexamer (Figure 6), its binding site being very similar to the off-axis zinc binding sites of 4-zinc pig insulin (Figure 7a). No water molecules were modeled in the lower part of the phenol binding site, owing to the absence of appropriate electron

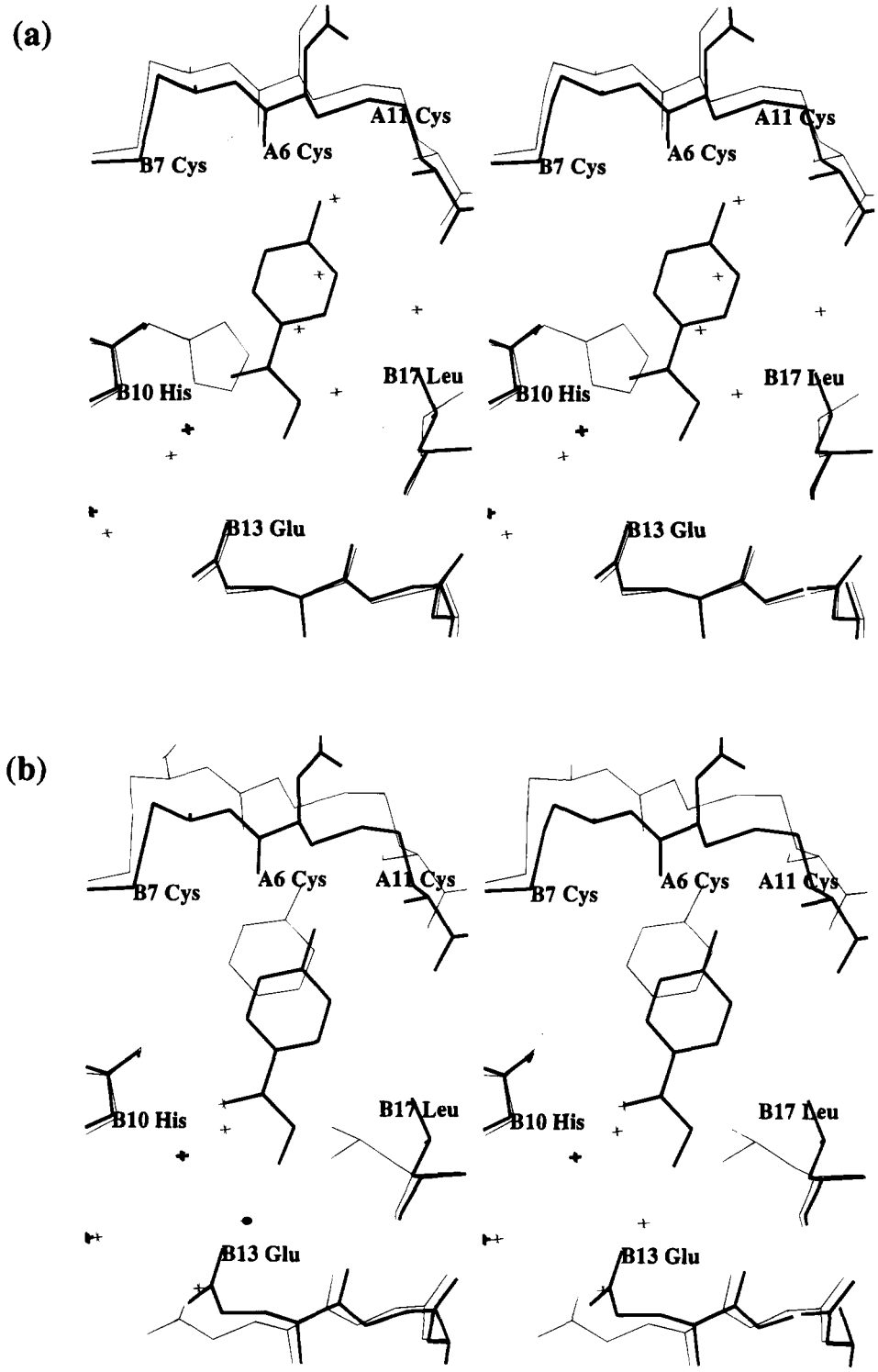

FIGURE 5: Stereoviews of the methylparaben binding site in the $T_3R_3^f$ hexamer (heavy lines) superimposed upon (a) the off-axis zinc ion binding site in the 4-zinc pig insulin hexamer (light lines) and (b) the phenol binding site in monoclinic insulin (light lines). Both comparisons show significant displacements of the A chain seen at the top of each figure.

density. The phenol binding interactions are identical to those of phenol in R_6 monoclinic phenol insulin, even though comparison of these two structures reveals a small but significant displacement of the A chain and adjustments of the side chain positions of residues B13 Glu and B17 Leu in the phenol binding site (Figure 7b). A similar comparison between the ligand binding sites in the methylparaben and $T_3R_3^f$ phenol hexamers shows that methylparaben binding causes a significant displacement of the A chain and that the torsion angles of B17 Leu in the $T_3R_3^f$ methylparaben and phenol binding sites are different (Figure 8).

DISCUSSION

The aim of these studies was to introduce different ligands into the binding sites identified in 4-zinc insulin and phenol insulin and to determine the ways they affect the hexamer structure. As anticipated, the thiocyanate ion replaced a chloride ion in the $T_3R_3^f$ hexamer, stabilizing the R_3^f trimer

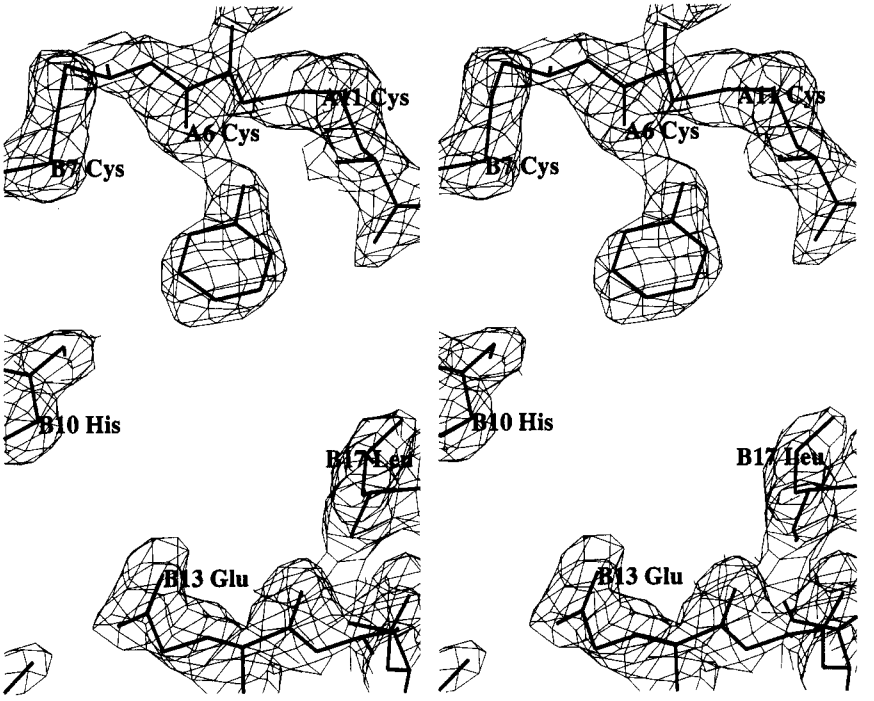

FIGURE 6: Stereoview of a $2F_o - F_c$ electron density map contoured at a level of 1σ, showing a phenol molecule bound in the R_3^f trimer of $T_3R_3^f$ phenol insulin.

by means of zinc ion coordination and van der Waals interactions with three symmetry-related B6 Leu side chains. The absence of coordinated zinc ions at the dimer−dimer interfaces of the thiocyanate insulin hexamer and $T_3R_3^f$ human insulin hexamer (Ciszak & Smith, 1994) shows that the zinc−anion interaction on the hexamer 3-fold axis is on its own sufficient to stabilize the T_3R_3 hexamer. This is confirmed by solution studies on a B5 Ala mutant insulin which also forms T_3R_3 hexamers even though off-axis zinc ion binding cannot take place (Renscheidt et al., 1984). Repeated insulin crystallization experiments over a range of thiocyanate ion concentrations failed to produce R_6 hexamers, reinforcing the observation that inorganic anions cannot stabilize R_6 hexamers of native insulin (Thomas & Wollmer, 1989; Brzovic et al., 1994; Bloom et al., 1995).

The methylparaben binding experiment has produced unanticipated results. First, the methylparaben insulin hexamer has the $T_3R_3^f$ conformation rather than the R_6 conformation of monoclinic phenol insulin. This appears to be due to the low concentration of methylparaben used in the crystallization experiment, which is approximately 10-fold lower than that of phenol in the preparation of R_6 monoclinic phenol insulin (Harding et al., 1966). We have shown that, when a similarly low concentration of phenol is used, $T_3R_3^f$ hexamers are obtained. Similarly, crystallization experiments with resorcinol (D. Smith, E. Ciszak, and J. Whittingham, unpublished results) show that hexamer conformation is determined by the concentration of phenolic additive, which complements many solution studies (Renscheidt et al., 1984; Wollmer et al., 1987; Thomas & Wollmer, 1989; Roy et al., 1989; Brader et al., 1991; Gross & Dunn, 1992; Choi et al., 1993).

A second unusual feature of the methylparaben insulin hexamer is that it has a tetrahedral zinc−chloride complex in the T_3 trimer, even though the equivalent zinc ion in 4-zinc pig insulin is octahedrally coordinated to three histidine side chains and three water molecules. There is no obvious explanation for this, although it is noteworthy that the coordination state of the zinc ion associated with the T_3 trimer in both thiocyanate insulin and $T_3R_3^f$ human insulin (Ciszak & Smith, 1994) is a mixture of octahedral and tetrahedral. Furthermore, the T_3 trimer zinc ion in tylenol insulin can be either tetrahedrally or octahedrally coordinated (Smith & Ciszak, 1994). It is possible that, in methylparaben insulin, the origin of the tetrahedral zinc ion coordination in the T_3 trimer may be electrostatic, given the low pH of crystallization (pH 5.4). It may be energetically favorable to include another anion in the hexamer to balance the charges on the zinc ions and the nearby cluster of B13 Glu side chains at low pH. As $T_3R_3^f$ phenol insulin, crystallized at pH 5.4, also appears to contain a chloride ion bound to the zinc ion in the T_3 trimer, we can at least conclude that this effect is not solely due to the binding of one particular phenolic ligand. To explore this coordination phenomenon further, insulin crystals were prepared in the presence of methylparaben using conditions identical to those used to grow crystals at pH 5.4, but at higher pH (pH 6). In this case, T_3R_3 hexamers in which the zinc ion in the T_3 trimer was octahedrally coordinated were obtained. However, the methylparaben ligand had not bound at the dimer−dimer interface.

By making comparisons of the various $T_3R_3^f$ and R_6 insulin hexamers, we can analyze the local structural perturbations caused by the binding of phenolic ligands. The phenol binding site in $T_3R_3^f$ phenol insulin is remarkably similar to the corresponding site in 4-zinc pig insulin, the only difference being the different conformation of the B10 His side chain which serves as a ligand for the off-axis zinc ion in the latter (Figure 7a). However, in the phenol binding sites of the $T_3R_3^f$ and R_6 insulin hexamers, the A chain positions are somewhat different (Figure 7b). This appears

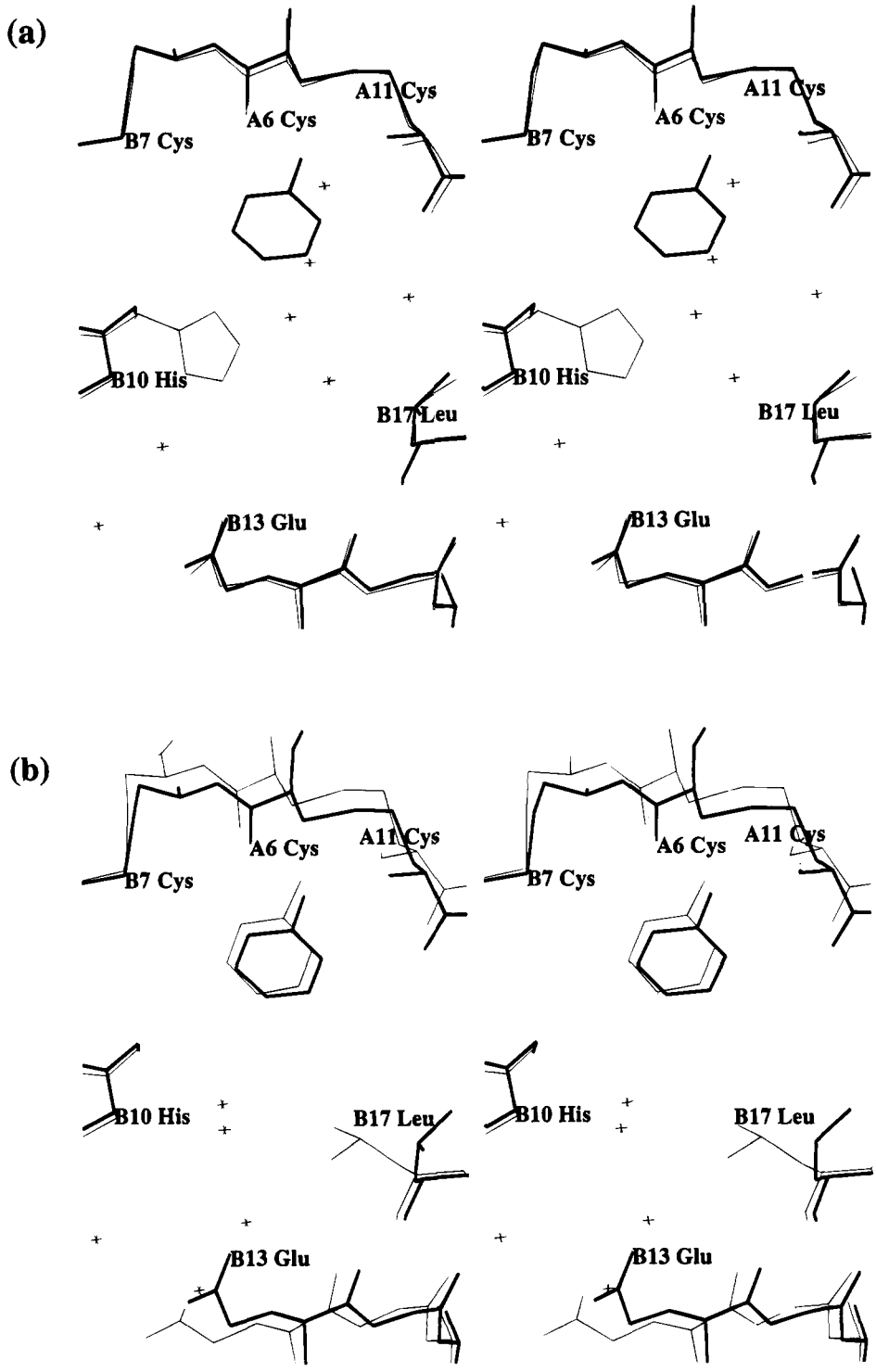

FIGURE 7: Stereoview of the phenol binding site in the $T_3R_3^f$ hexamer (heavy lines) superimposed upon (a) the off-axis zinc ion binding site in the 4-zinc pig insulin hexamer (light lines) and (b) the phenol binding site in monoclinic insulin (light lines).

to be the result of the different conformations of residues B1−B3 in the R_3 trimers of the $T_3R_3^f$ and R_6 hexamers, since the B chain N-terminal residues lie close to the A chain. This A chain movement is facilitated by an alteration in the side chain torsion angle of residue B7 Cys which makes a disulfide bridge with A7 Cys. Hence, B7 Cys C_β and S_γ are displaced by 0.2 and 0.4 Å, respectively, while at the other end of the disulfide bridge, A7 Cys C_α is displaced by 0.8 Å.

With methylparaben bound instead of phenol in the $T_3R_3^f$ hexamer, there is a large A chain displacement, resulting in a more compact binding site (Figure 8). This compactness is reflected in the smaller unit cell dimensions for methylparaben insulin (Table 2). In this case, the A chain displacement is associated with the nature of the phenolic ligand and its position in the binding site. In the methylparaben molecule, the carbonyl group is conjugated with the aromatic ring, making the ligand essentially planar and inflexible. Hence, it is necessary for methylparaben to occupy a lower position in the binding site than phenol, thus preventing interference between the ligand and residue B10 His. This also maximizes van der Waals and hydrogen-

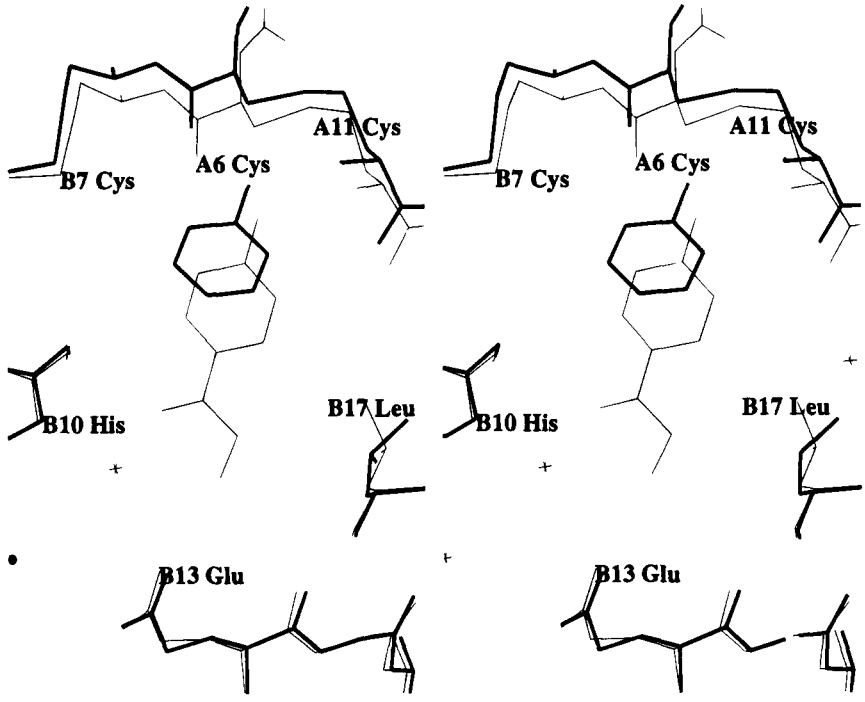

FIGURE 8: Stereoview of the phenol binding site in the $T_3R_3^f$ hexamer (heavy lines) superimposed upon the methylparaben binding site in methylparaben insulin (light lines), showing how the position of the A chain seen at the top of the figure is different in the two structures.

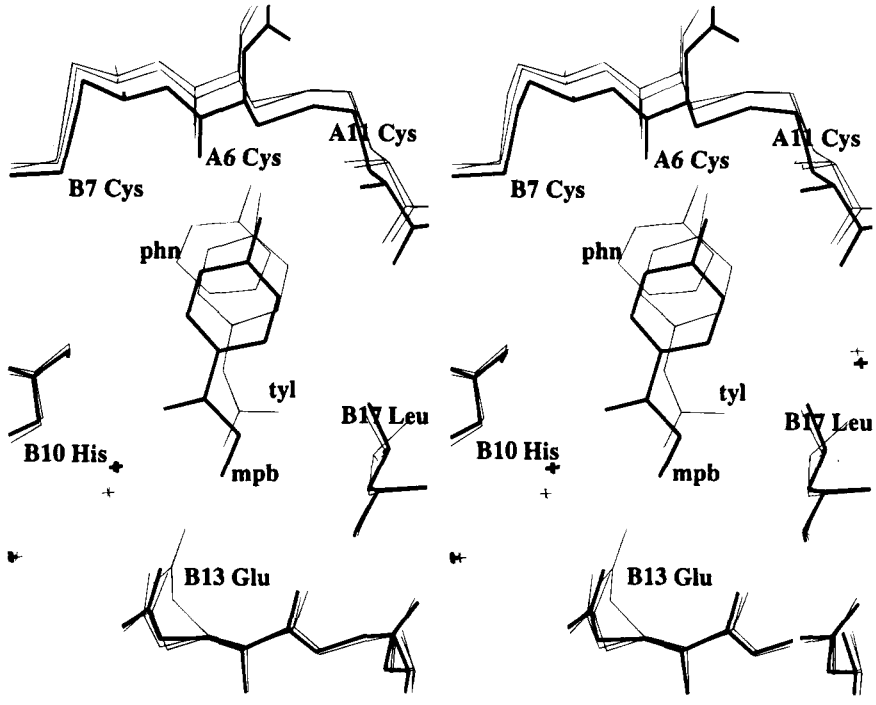

FIGURE 9: Stereoview of the methylparaben (mpb) binding site in the $T_3R_3^f$ hexamer (heavy lines) superimposed upon the ligand binding sites in the phenol (phn) and pH 5.6 tylenol (tyl) $T_3R_3^f$ hexamers (light lines), showing the relative positions of the three phenolic ligands in the binding site.

bonding interactions between methylparaben and the lower site residues. The associated A chain is displaced in order to maintain the hydrogen bonds with the hydroxyl group of methylparaben. A comparison of methylparaben insulin with the pH 5.6 $T_3R_3^f$ tylenol insulin hexamer (Smith & Ciszak, 1994) shows that the tylenol ligand is higher in the binding site than methylparaben although it is a similar size (Figure 9). The explanation for this seems to be that the tylenol molecule is more flexible, a significant rotation of its acetamide group out of the plane of the aromatic ring allowing it to occupy the position shown. The flexibility between the A and the B chains observed in these experiments may well be an inherent feature of the insulin molecule, possibly associated with receptor binding.

A discussion of the effects of ligand binding in the insulin hexamer is not complete without a consideration of the energetics of ligand binding. The binding of anions and phenols can be seen to be both enthalpically and entropically

favorable. The binding of phenol, for instance, results in the displacement of water molecules and the formation of two hydrogen bonds between the ligand hydroxyl group and the main chain atoms of the nearby A chain, together with an aromatic ring−ring interaction with the side chain of B5 His of the adjacent dimer. The binding of methylparaben appears to be even more favorable, involving the displacement of more water molecules and the formation of extensive van der Waals and hydrogen-bonding interactions around the binding site. On the hexamer 3-fold axis, the coordination of a thiocyanate ion to the zinc ion is probably more energetically favorable than the equivalent interaction with a chloride ion because of the "softness" of thiocyanate and its shape (Fridovich, 1962). For this reason, the concentration of thiocyanate required to promote the $T_6 \rightarrow T_3R_3$ transformation is considerably less than that of chloride (De Graaff et al., 1981).

The energy gained from ligand binding in the insulin hexamer is comparable to that required to promote conformational rearrangements within the hexamer. For instance, titration studies using circular dichroism spectroscopy show that the binding energy of phenol or *m*-cresol at the dimer−dimer interface is about −18 kJ/mol, while the $T_6 \rightarrow T_3R_3$ and $T_3R_3 \rightarrow R_6$ transitions have $\Delta G°$ values of about 8 and 21 kJ/mol, respectively (Jacoby et al., 1993). Calculations based on a three-state allosteric model system show similar results (Bloom et al., 1995). These values indicate that the transition of the second trimer requires almost 3 times as much energy as that of the first, i.e. that negative cooperativity exists between the two trimers (Krüger et al., 1990). This information combined with our knowledge of the hexamer structures suggests to us that the binding energy of one anion in the hexamer is sufficient to shift the $T_6 \rightleftharpoons R_6$ equilibrium only as far as the T_3R_3 state, while the energy of binding of six phenolic ligands is sufficient to overcome the negative cooperativity between the trimers to stabilize the R_6 state.

CONCLUSIONS

The crystal structures of thiocyanate insulin, methylparaben insulin, and $T_3R_3^f$ phenol insulin show that the insulin hexamer can accommodate a variety of phenolic and anionic ligands, owing to the flexibility at the dimer−dimer interface and the tolerance of the zinc ion for different coordinating anions. The observed flexibility between the A and B chains in insulin may be necessary *in vivo* as an aid to receptor binding. These experiments and others published recently (Ciszak & Smith, 1994; Smith & Ciszak, 1994) have modified our understanding of the zinc ion coordination in the insulin hexamer, which is now seen to be more complex than was originally thought. They also complement solution studies demonstrating that hexamer conformation is a function of the concentrations and types of additives. The difficulty in inducing R_6 hexamers using only anions such as chloride and thiocyanate suggests that the binding of phenolic molecules in the insulin hexamer is necessary to overcome negative cooperativity between the two trimers of the hexamer, thus promoting the $T_3R_3 \rightarrow R_6$ transition.

ACKNOWLEDGMENT

We are grateful for gifts of insulin from Novo Nordisk. We would like to thank Judy Clarkson for help with crystallization experiments and Anthony Wilkinson and Gideon Davies for helpful discussions during the preparation of the manuscript. We would also like to acknowledge the help of David Edwards, Paul Holden, Richard Lewis, Maria Turkenburg, and Chandra Verma.

REFERENCES

Andreetti, G. D., Jain, P. C., & Lingafelter, E. C. (1969) *J. Am. Chem. Soc. 91*, 4112−4115.
Baker, E. N., Blundell, T. L., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Crowfoot, Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., & Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London B319*, 369−456.
Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., & Mercola, D. (1976) *Nature 261*, 166−168.
Bentley, G. A., Brange, J., Derewenda, Z., Dodson, E. J., Dodson, G. G., Markussen, J., & Wilkinson, A. J. (1992) *J. Mol. Biol. 228*, 1163−1176.
Bloom, C. R., Choi, W. E., Brzovic, P. S., Ha, J. J., Sheng-Tung, H., Kaarsholm, N. C., & Dunn, M. F. (1995) *J. Mol. Biol. 245*, 324−330.
Blundell, T., Dodson, G., Hodgkin, D., & Mercola, D. (1972) *Adv. Protein Chem. 26*, 279−402.
Brader, M. L., Kaarsholm, N. C., Roberts, W.-K., & Dunn, M. F. (1991) *Biochemistry 30*, 6636−6645.
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., & Smith, G. D. (1994) *Biochemistry 33*, 1512−1517.
Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. C., & Smith, G. D. (1995) *Structure 3*, 615−622.
Collaborative Computational Project, Number 4 (1994) *Acta. Crystallogr. D50*, 760−763.
De Graaff, R. A. G., Lewit-Bentley, A., & Tolley, S. P. (1981) in *Structural Studies on Molecules of Biological Interest* (Dodson, G., Glusker, J. P., & Sayre, D., Eds.) pp 547−556, Clarendon Press, Oxford, U.K.
Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swenson, D. (1989) *Nature 338*, 594−596.
Fridovich, I. (1962) *J. Biol. Chem. 238*, 592−598.
Gross, L., & Dunn, M. F. (1992) *Biochemistry 31*, 1295−1301.
Harding, M. M., Hodgkin, D. C., Kennedy, A. F., O'Connor, A., & Weitzmann, P. D. J. (1966) *J. Mol. Biol. 16*, 212−226.
Jacoby, E., Krüger, P., Karatas, Y., & Wollmer, A. (1993) *Biol. Chem. Hoppe-Seyler 374*, 877−885.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.
Kaarsholm, N. C., Ko, H. C., & Dunn, M. F. (1989) *Biochemistry 28*, 4427−4435.
Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
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. Crystallogr. 26*, 283−291.
Renscheidt, H., Strassburger, W., Glatter, U., Wollmer, A., Dodson, G. G., & Mercola, D. A. (1984) *Eur. J. Biochem. 142*, 7−14.
Riès-Kautt, M. M., & Ducruix, A. F. (1991) *J. Cryst. Growth 110*, 20−25.
Roy, M., Brader, M. L., Lee, R. W.-K., Kaarsholm, N. C., Hansen, J. F., & Dunn, M. F. (1989) *J. Biol. Chem. 264*, 19081−19085.
Schlichtkrull, J. (1958) Ph.D. Thesis, University of Copenhagen, Denmark.
Smith, G. D., & Ciszak, E. (1994) *PNAS 91*, 8851−8855.
Smith, G. D., & Dodson, G. G. (1992) *Biopolymers 32*, 441−445.
Thomas, B., & Wollmer, A. (1989) *Biol. Chem. Hoppe-Seyler 370*, 1235−1244.
Tolley, S. P. (1987) Ph.D. Thesis, University of York, U.K.
Wilson, A. J. C. (1949) *Acta. Crystallogr. 2*, 318−321.
Wollmer, A., Rannefeld, B., Johansen, B. R., Hejnaes, K. R., Balschmidt, P., & Hansen, F. B. (1987) *Biol. Chem. Hoppe-Seyler 368*, 903−911.
Xiao, B. (1990) Ph.D. Thesis, University of York, U.K.

BI9514222