

## Interactions of Phenol and *m*-Cresol in the Insulin Hexamer, and Their Effect on the Association Properties of B28 Pro → Asp Insulin Analogues<sup>‡</sup>

Jean L. Whittingham,\* David J. Edwards,<sup>§</sup> Alfred A. Antson, Judy M. Clarkson, and G. Guy Dodson

*Department of Chemistry, University of York, Heslington, York YO10 5DD, England*

*Received April 10, 1998; Revised Manuscript Received June 5, 1998*

**ABSTRACT:** Insulin's natural tendency to form dimers and hexamers is significantly reduced in a mutant insulin B28 Pro → Asp, which has been designed as a monomeric, rapid-acting hormone for therapeutic purposes. This molecule can be induced to form zinc hexamers in the presence of small phenolic derivatives which are routinely used as antimicrobial agents in insulin preparations. Two structures of B28 Asp insulin have been determined from crystals grown in the presence of phenol and *m*-cresol. In these crystals, insulin exists as R<sub>6</sub> zinc hexamers containing a number of phenol or *m*-cresol molecules associated with aromatic side chains at the dimer–dimer interfaces. At the monomer–monomer interfaces, the B28 Pro → Asp mutation leads to increased conformational flexibility in the B chain C termini, resulting in the loss of important intermolecular van der Waals contacts, thus explaining the monomeric character of B28 Asp insulin. The structure of a cross-linked derivative of B28 Asp insulin, containing an Ala-Lys dipeptide linker between residues B30 Ala and A1 Gly, has also determined. This forms an R<sub>6</sub> zinc hexamer containing several *m*-cresol molecules. Of particular interest in this structure are two *m*-cresol molecules whose binding disrupted the β-strand in one of the dimers. This observation suggests that the cross-link introduces mechanical strain on the B chain C terminus, thereby weakening the monomer–monomer interactions.

The continuing efforts to analyze and understand the structure of insulin are motivated partly by the need to produce modified insulins for use in the treatment of diabetes. Currently, daily variations in blood glucose levels are controlled by administering a combination of rapid- and prolonged-acting insulins. The rapid-acting insulin provides a high dose of insulin at meal times, while the prolonged-acting insulin ensures that there is a prevailing basal level of insulin between meals, particularly overnight. The design of these analogues is based on the fact that insulin, which is active as a monomer, has an inherent tendency to form hexamers in the presence of zinc ions. Although the pharmacokinetics of insulin absorption after subcutaneous injection are very complex, it can be simply stated that the rate of hexamer dissociation is proportional to the rate of insulin action (*1*). Crystal structure determinations have helped to explain why this is the case, and have provided the impetus for the design of new insulins which might help to reduce the complications associated with the lack of normoglycaemia in diabetics.

The first insight into the assembly of insulin was revealed when the structure was determined in the 1960s (*2, 3*). The insulin monomer contains two polypeptide chains, an A chain of 21 amino acids and a B chain of 30 amino acids, which

contain α-helices between residues A1 and A8, A13 and A19, and B9 and B19. The two chains are covalently linked by two interchain disulfide bridges between A7 and B7 and A20 and B19. Additional secondary structure exists in the dimer in which identical, hydrophobic surfaces of two monomers come together forming an antiparallel β-strand between residues B23 and B30 of each monomer, characterized by four main chain–main chain hydrogen bonds. Residue B28 Pro is very significantly placed, making van der Waals contacts with residues B20–B23 of the adjacent monomer. Experiments with B chain C terminus substitutions and truncations have shown that residues B26–B30 are essential for dimer formation, and that a single substitution at residue B28 is sufficient to give the insulin a markedly increased tendency to form monomeric units upon dilution (*4*).

The insulin hexamer consists of three dimers aggregated around two zinc ions forming a soluble, globular protein structure (*3*). The hexamer has been seen to adopt one of three conformational states known as T<sub>6</sub>, T<sub>3</sub>R<sub>3</sub>, and R<sub>6</sub>, which differ in their symmetry and the arrangements of the first eight residues of the B chain N terminus of each monomer; in the T state, these residues adopt an elongated conformation, while in the R state, they form an α-helix. The R<sub>6</sub> hexamer is obtained in the presence of phenol molecules which are routinely added to insulin preparations as antimicrobial agents. Crystallographic studies have shown that the phenol molecules participate in specific interactions at the dimer–dimer interface which stabilize the B1–B8 α-helices (*5, 6*). The stability of the R<sub>6</sub> hexamer as a whole is enhanced by these interactions, since the diffusion of the

<sup>‡</sup> The coordinates for the three B28 Asp insulin analogue structures have been deposited in the Brookhaven Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY (references 1ZEG, 1ZEH, and 1ZEI for B28Asp-*phn*, B28Asp-*mcr*, and B28Asp-X-*mcr*, respectively).

\* To whom correspondence should be addressed.

<sup>§</sup> Current address: Molecular Simulations Inc., 9685 Scranton Rd., San Diego, CA 92121-3752.

Table 1: Crystallization Conditions

additive	B28Asp- <i>phn</i>	B28Asp- <i>mcr</i>	B28Asp- <i>X-mcr</i>
insulin (mg)	3.5	3.5	3.5
0.02 M HCl ( $\mu$ L)	500	500	500
0.12 M zinc acetate ( $\mu$ L)	50	50	—
0.15 M zinc acetate ( $\mu$ L)	—	—	50
0.2 M trisodium citrate ( $\mu$ L)	250	250	360
2.5% phenol in ethanol ( $\mu$ L)	200	—	—
2.5% <i>m</i> -cresol in ethanol ( $\mu$ L)	—	200	200
NaCl (mg)	60	60	60
pH	6.2–6.8	5.8–6.8	5.4–6.4

zinc ions out of the hexamer, which leads to its breakdown into dimers, is hindered owing to the close proximities of the B1–B8  $\alpha$ -helices to the metal ions (7).

An understanding of the forces which stabilize the dimer and the hexamer has made it possible to manipulate the rate of insulin action for therapeutic purposes. For example, prolonged-acting preparations consist of amorphous or crystalline preparations of zinc hexamers which dissolve slowly after injection (8). On the other hand, rapid action can be achieved by introducing a mutation which disturbs the monomer–monomer contact in the dimer (9, 10), and in this respect, a mutation at position B28 is particularly effective. In this study, the crystal structure of the rapid-acting B28 Asp mutant insulin has been determined to investigate its ability to aggregate in the presence of zinc ions and to determine how preservatives such as phenol and *m*-cresol affect the structure. The crystal structure of the Asp<sup>B28</sup>-Lys<sup>B29</sup>-Ala-Ala-Lys-Gly<sup>A1</sup>, a cross-linked low-potency precursor of B28 Asp insulin, has also been determined to extend the commentary on the B28 Asp mutation and *m*-cresol binding. The cross-link prevents movement of the B chain C terminus which is essential for receptor binding (11).

## MATERIALS AND METHODS

### Crystallization

The B28 Pro  $\rightarrow$  Asp mutant insulin and the cross-linked analogue were donated by Novo Nordisk A/S (Denmark). Crystals were grown by the batch method (12). For each of the three crystallizations, a protein solution was made according to the conditions listed in Table 1. While the temperature of the solution was maintained at approximately 50 °C, the pH was increased to 8.0 using 1 M NaOH, to ensure that all the components were completely dissolved, and then gradually reduced by addition of 1 M HCl. Samples (100  $\mu$ L) were taken over the pH range given in Table 1. The filtered samples were sealed in 1 mL glass tubes and placed in an insulated 50 °C water bath for a slow cooling period of 1 week to allow crystal formation.

### Data Collection and Processing

Data sets were collected on two crystal forms of B28 Asp insulin, one with phenol (B28Asp-*phn*) and the other with *m*-cresol (B28Asp-*mcr*). Data for B28Asp-*mcr* insulin were collected from two crystals from the same crystallization tube (pH 6.8) using a Marresearch scanner mounted on a Rigaku RU200 rotating anode X-ray source ( $\lambda = 1.5418$  Å). Each crystal was initially soaked in a cryosolution consisting of

40% glycerol and 60% crystallization solution, before being mounted in a rayon loop and vitrified at 120 K. The two crystals were used to collect low- and high-resolution data sets, which were then processed with DENZO and scaled and merged with SCALEPACK version 1.6.1 (13). Although the unit cells for the two frozen crystal were slightly different, the two sets of data merged without problems. The data for B28Asp-*phn* insulin were collected in exactly the same manner, although in this case, only a single crystal (pH 6.5) was needed to collect both the high- and low-resolution data. Data processing statistics for both crystal forms are given in Table 2.

Data for the cross-linked Asp<sup>B28</sup>-Lys<sup>B29</sup>-Ala-Ala-Lys-Gly<sup>A1</sup> insulin (B28Asp-*X-mcr*) were obtained at room temperature. An initial data set was collected at the Daresbury SRS, station 9.5 ( $\lambda = 0.876$  Å), using a Marresearch scanner. This data set was collected to 1.9 Å resolution, with 2° oscillation per image and a total of 70° rotation of the crystal. The rest of the data were collected from three crystals (pH 6.4) using an Rigaku R-AXIS IIC imaging plate mounted on a Rigaku RU200 rotating anode X-ray source ( $\lambda = 1.5418$  Å). These data sets were collected to 2.4 Å resolution, with 1.5° oscillation per image and total rotations of 10.5, 129.0, and 60.0°, respectively, for the three crystals. The crystals were highly sensitive to X-rays; nevertheless, data from four crystals could be successfully processed and merged using DENZO and ROTAVATA/AGROVATA (14) (Table 2).

### Structure Solution and Refinement

**B28Asp-*mcr* Insulin.** The cell dimensions for this B28Asp-*mcr* insulin crystal suggested that it was isomorphous with one of the predetermined rhombohedral insulins. Since the T<sub>3</sub>R<sub>3</sub> hexamer of four-zinc insulin (15) seemed to give the best initial *R*-factor, this structure was tried as a starting model for refinement having removed water molecules, counterions, and residues B27–B30. Atomic *B*-factors were each given an average value of 20.0 Å<sup>2</sup> in accordance with the Wilson plot of the data. During refinement, which was carried out using REFMAC (16), a 5% sample set of data was excluded and used to calculate an *R*-free value. Both the conventional *R*-factor and *R*-free values decreased satisfactorily, although the electron density maps soon gave clear indications that the structure was in fact that of an R<sub>6</sub> hexamer. After this, considerable rebuilding of the B chain N terminus of one of the monomers of the dimer and also the position of the nearby A chain took place using the X-BUILD option in the QUANTA molecular modeling system (17). As the refinement continued, the positions of the mutated side chains became clear in the electron density maps, and five *m*-cresol molecules per dimer were gradually built into the model. Refinement details and statistics for the final structure are given in Table 3.

**B28Asp-*phn* Insulin.** The refinement of B28Asp-*phn* insulin was carried out along the same lines as above using the refined B28Asp-*mcr* insulin coordinates as a starting model. To reduce bias, water molecules, counterions, *m*-cresol molecules, and residues B27–B30 were removed from the initial set of coordinates, and all atomic *B*-factor values were set to a value of 20.0 Å. After the first few cycles of refinement, the positions of two zinc ions and two phenol molecules were very clear in both  $2F_o - F_c$  and  $F_o$

Table 2: Data Statistics

	B28Asp- <i>phn</i>	B28Asp- <i>mcr</i>	B28Asp-X- <i>mcr</i>
space group	<i>R</i> 3	<i>R</i> 3	<i>P</i> 2 <sub>1</sub>
cell dimensions (Å)	<i>a</i> = 77.73, <i>c</i> = 39.16	<i>a</i> = 77.49, <i>c</i> = 38.91	<i>a</i> = 48.91, <i>b</i> = 64.77, <i>c</i> = 53.95, $\beta$ = 109.8°
resolution limits (Å)	14.70–1.60	19.39–1.51	19.92–1.90
no. of observations	74 966	102 174	79 085
no. of independent reflections	11 469	13 601	21 942
highest-resolution shell (Å)	1.63–1.60	1.53–1.51	2.00–1.90
completeness (%)			
all data	98.7	97.5	87.8
highest-resolution shell	97.4	54.6	77.0
<i>R</i> <sub>merge</sub> (on <i>I</i> ) <sup>a</sup>			
all data	0.023	0.070	0.054
highest-resolution shell	0.038	0.169	0.216
reflections where <i>I</i> > 3σ (%)			
all data	98.9	91.8	80.8
highest-resolution shell	96.7	81.0	49.6

<sup>a</sup>  $R_{\text{merge}} = \sum |I - I_i| / \sum I$ , where *I* is the mean intensity of the *N* reflections with intensities *I<sub>i</sub>*.

Table 3: Refinement and Model Statistics

parameter	target rms	rms Δ		
		B28Asp- <i>phn</i>	B28Asp- <i>mcr</i>	B28Asp-X- <i>mcr</i>
bond lengths (1–2) (Å)	0.020	0.014	0.014	0.019
angle-related distance (1–3) (Å)	0.040	0.032	0.036	0.041
intraplanar distances (1–4) (Å)	0.050	0.034	0.036	0.044
peptide planar groups (Å)	0.030	0.027	0.027	0.028
chiral volumes (Å <sup>3</sup> )	0.100	0.084	0.117	0.116
van der Waals distances				
single torsion contacts (Å)	0.300	0.165	0.169	0.171
multiple torsion contacts (Å)	0.300	0.278	0.276	0.272
possible hydrogen bonds (Å)	0.300	0.134	0.101	0.169
torsion angles				
planar (0°, 180°)	7.0	5.7	6.0	4.9
staggered (60°, 120°)	15.0	11.1	12.9	17.9
orthonormal (±90°)	20.0	15.3	16.6	16.5
thermal factors				
main chain bond (1–2) (Å <sup>2</sup> )	2.000	1.589	1.564	2.393
main chain angle (1–3) (Å <sup>2</sup> )	3.000	2.363	2.359	3.460
side chain bond (Å <sup>2</sup> )	2.000	2.314	2.061	3.252
side chain angle (Å <sup>2</sup> )	3.000	3.402	2.987	4.944
final <i>R</i> -factor, <sup>a</sup> all data ( <i>R</i> -free)		0.1449 (0.1923)	0.1590 (0.1930)	0.1762 (0.2316)
no. of protein atoms in asymmetric unit		812	812	2502
no. of water molecules in asymmetric unit		106	106	151
average temperature factors				
main chain (rms) (Å <sup>2</sup> )		15.313 (1.136)	16.643 (1.176)	26.851 (1.697)
side chain (rms) (Å <sup>2</sup> )		20.823 (1.594)	22.589 (1.481)	31.497 (2.517)
residues in allowed regions of Ramachandran plot		100%	100%	100%

<sup>a</sup> Crystallographic *R*-factor (*R*-free) =  $\sum |F_o| - |F_c| / \sum |F_o|$ .

– *F<sub>c</sub>* electron density maps, and as the refinement continued, two chloride ions and another phenol molecule were also built into the maps. At this point, residues B27–B30 of molecule 2 were also built into the model, but it was only toward the end of the refinement that the positions of residues B27–B30 of molecule 1 could be determined with any conviction. More refinement details are given in Table 3.

**B28Asp-X-*mcr* Insulin.** This insulin crystallized in a monoclinic unit cell containing a hexamer in the asymmetric unit. The dimer from native monoclinic phenol insulin, which had a different unit cell (5), was used to solve the structure with AMoRe (18), and after rigid body refinement with XPLOR (19), the *R*-factor dropped from 50.07 to 42.35 while the *R*-free value calculated with 5% of the data fell to 44.88%. REFMAC and X-BUILD were employed for the refinement and model building. After the first few cycles of refinement, zinc and chloride ions and six *m*-cresol molecules were built into 2*F<sub>o</sub>* – *F<sub>c</sub>* and *F<sub>o</sub>* – *F<sub>c</sub>* electron

density maps. As the refinement continued, water molecules were added using the X-SOLVATE option in the QUANTA molecular modeling system (17), and electron density for parts of the cross-linked regions became evident. Two extra *m*-cresol molecules also became apparent at one of the monomer–monomer interfaces. The asymmetry of this arrangement was unusual, and the structure was checked carefully using omit maps. For refinement details and statistics for the final structure, see Table 3.

#### Structure Comparisons

In the following discussion, dimers and hexamers have been compared by superposing them using the program LSQKAB (14), using a least-squares alignment of the main chain atoms of the B9–B19 α-helix of each of the insulin molecules concerned. For each of the structures, the two monomers in the dimer have been labeled molecule 1 and

molecule 2, the numbering being an arbitrary assignment within the dimer but consistent from structure to structure.

## RESULTS

*Structures of B28 Asp Insulin with Phenol (B28Asp-phn) and with m-Cresol (B28Asp-mcr).* Both structures consist of  $R_6$  insulin hexamers whose crystal packing is described by the rhombohedral space group  $R3$ , in which the asymmetric unit contains an insulin dimer. Three dimers are arranged around a crystallographic 3-fold axis to form a compact hexamer. In each hexamer, two zinc ions are located about 15 Å apart on a central channel which is collinear with the crystallographic 3-fold axis. The zinc ions mediate important interactions between the dimers, each one forming a tetrahedral coordination sphere involving three B10 His side chains (one from each dimer) and a chloride ion at a distance of 2.2 Å. Both structures contain a number of phenolic ligands, most of which are lodged at the dimer-dimer interfaces.

The two structures are on the whole very similar, the overall main chain rms difference for the dimers being 0.18 Å. However, residues B1, B2, and B27–B30 (excluded from this calculation) show considerable variation. In molecules 1 and 2 of B28Asp-mcr, residues B1–B8 form a complete  $\alpha$ -helix, while in B28Asp-phn, the B1–B8  $\alpha$ -helices are slightly “frayed” because of a 90° rotation of residue B1 Phe about the C–C $\alpha$  bond of B2 Val. As a result, the main chain oxygen of B1 Phe, rather than making an internal hydrogen bond with the main chain nitrogen of B5 His, makes a crystal contact with B1 N in an adjacent, symmetry-related hexamer. These observations illustrate the inherent flexibility at the beginning of the B chain which has also been noted in earlier structure determinations (20, 21), and in solution (22).

When the B chain C termini of the two structures are compared, those in molecule 2 adopt strikingly similar conformations with the side chain of B28 Asp forming a 2.8 Å salt bridge with B29 Lys NZ and a 2.7 Å hydrogen bond with the hydroxyl group of B26 Tyr (Figure 1). In molecule 1, however, the two structures are very different. In B28Asp-phn, the last four residues of the B chain adopt a  $\beta$ -bend conformation, allowing the side chain of B30 Thr to hydrogen bond to B27 N. The side chain of B28 Asp, which is on the turn itself, shows some sign of disorder. In B28Asp-mcr, residues B29 and B30 are disordered and the position of B28 Asp is completely different from that in B28Asp-phn, presumably because of the presence of an *m*-cresol molecule near B28 Asp (Figures 2 and 3). It is worth noting that since the B chain N and C termini are located toward the outside of the hexamer their differences in conformation have no impact on the quaternary structure of the hexamer.

*Phenol and m-Cresol Binding.* The B28Asp insulin hexamers are host to a number of phenolic ligands, occupying a variety of binding sites. In the structure of B28Asp-phn, there are two different types of phenol binding sites, which for clarity have been named type I and type II sites. The type I site is located in a hydrophobic pocket at the dimer interface and has been seen in many previously determined structures (Figure 2) (5, 6, 21, 23, 24). Here, the buried phenol molecule hydrogen bonds to the main chain atoms

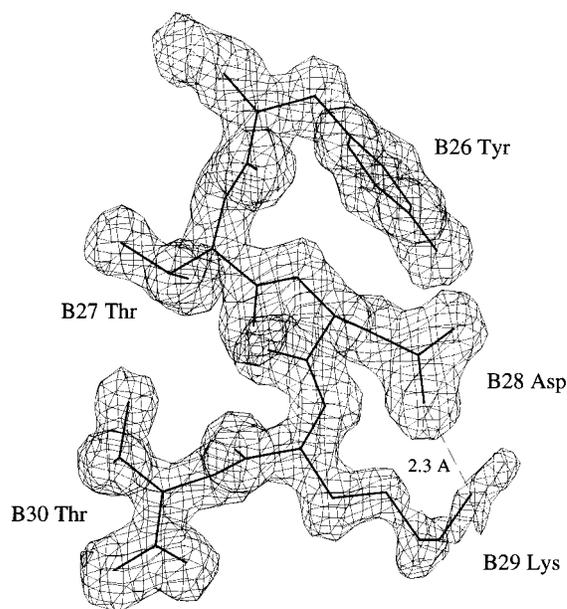


FIGURE 1: Electron density map showing the position of the mutated B28 Asp side chain in molecule 2 of B28Asp-phn insulin. There is a 2.3 Å salt bridge between the side chains of B28 Asp and B29 Lys (dashed line).

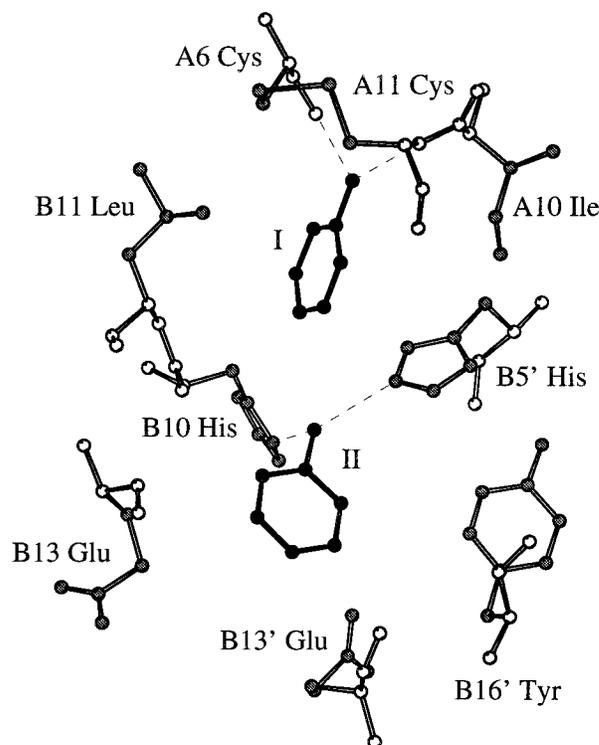


FIGURE 2: Type I and type II phenol binding in the B28Asp-phn insulin hexamer. The two phenol molecules, shown in black, are surrounded by hydrophobic and polar side chains at the dimer-dimer interface, shown in gray (main chain atoms in white). The residue labels are either primed or unprimed to show residues from different dimers. Hydrogen bonds are indicated by dashed lines. This figure was made using MOLSCRIPT (30).

A6 O and A11 N of one dimer and forms a ring-stacking interaction with the side chain of B5 His in the neighboring dimer. There are two type I phenol molecules per dimer-dimer interface associated with two B1–B8  $\alpha$ -helices. The type II phenol binding site is located near the type I site at the dimer-dimer interface, closer to the center of the

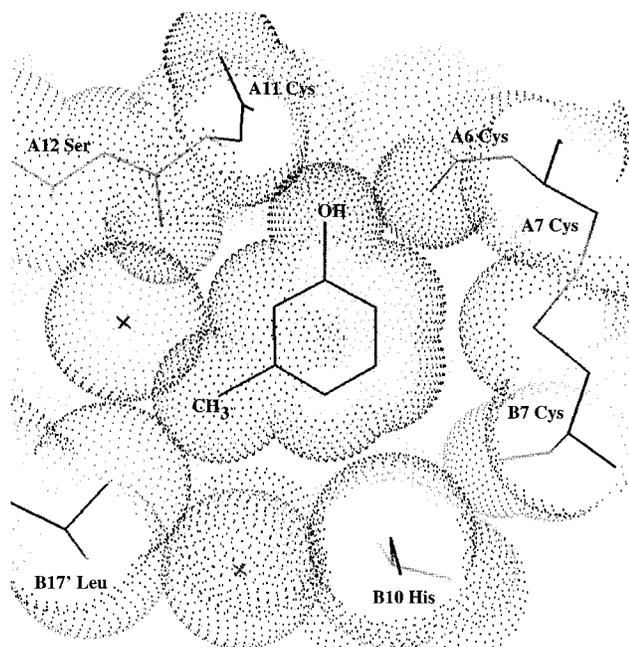


FIGURE 3: Type I *m*-cresol binding in the B28Asp-*mcr* insulin hexamer, showing the close packing interactions between the ligand and residues and water molecules (marked by  $\times$ ) at the dimer-dimer interface. Note in particular the van der Waals contact between the methyl group of *m*-cresol and the side chain of residue B17 Leu (primed to show that it is from a different dimer compared to the other residues).

hexamer (Figure 2). The position of the phenol molecule is determined by two hydrogen bonds to B5 His NE2 (molecule 2) and B10 His ND1 (molecule 2) of adjacent dimers. Although there is a convenient space for the phenol molecule to bind, its presence brings about some small alterations to the side chains positions of nearby B13 Glu side chains of molecules 1 and 2, B17 Leu (molecule 1) and B5 His (molecule 2). Furthermore, the limited size of the cavity does not allow for the binding of an equivalent phenol molecule associated with the other molecule in the dimer, so there is just one type II phenol molecule per dimer.

The hexamer of B28Asp-*mcr* contains both type I and type II *m*-cresol binding sites. In the type I site, the position and interactions of the *m*-cresol molecule are very similar to those of phenol in B28Asp-*phn*, although in molecule 2 the *m*-cresol ligand is displaced by 0.6 Å relative to the equivalent phenol molecule, placing the methyl group in van der Waals contact with the side chains of B14 Ala and B17 Leu from adjacent dimers. In molecule 1, the *m*-cresol ligand makes similar interactions but is displaced only slightly (Figure 3). In the type II binding site, the *m*-cresol molecule occupies the same position as the equivalent phenol molecule, albeit a little disordered. This disorder may be correlated with that of a nearby B13 Glu side chain which adopts two conformations, one directed away from the *m*-cresol molecule (as in the equivalent phenol binding site) and the other occupying the same space as *m*-cresol.

There are two additional ligand binding sites which are unique to the B28Asp-*mcr* structure. One of these, named type III, is located at the B chain C terminus of molecule 1 in the proximity of the mutated residue B28 Asp. In this position, the *m*-cresol molecule is exposed to solvent, its hydroxyl group hydrogen bonding to two first shell water

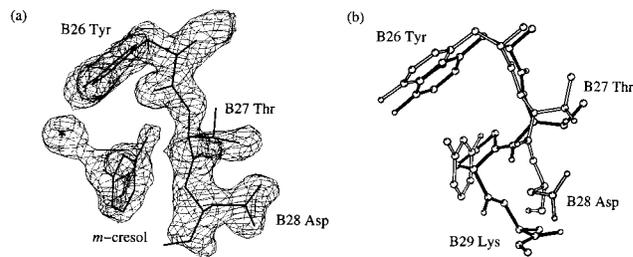


FIGURE 4: Type III *m*-cresol binding in B28Asp-*mcr* insulin. (a) Electron density map showing the proximity of the *m*-cresol molecule to the mutated side chain B28 Asp. The hydroxyl group of *m*-cresol is hydrogen bonded to a water molecule (marked by an asterisk). (b) Comparison of the dimers of B28 Asp-*mcr* insulin (white bonds) and the native R<sub>6</sub> insulin hexamer (black bonds), showing the displacement of the B chain C-terminal residues in molecule 1 as a result of *m*-cresol binding. This figure was made using BOBSRIPT (R. Esnouf, unpublished), a modified version of MOLSCRIPT (30).

molecules. Nevertheless, it makes some very significant interactions with the protein, the aromatic ring making van der Waals contact with the main chain of residue B28 Asp, and the methyl group making hydrophobic interactions with the side chains of nearby B26 Tyr (about 4.0 Å) and A2 Ile and A3 Val of the adjacent monomer in the dimer (Figure 4a). The remaining *m*-cresol binding site, type IV, is situated between the dimers toward the surface of the hexamer. A ring-stacking interaction between *m*-cresol and the side chain of A14 Tyr (molecule 1) shields the ligand from the solvent continuum, and hydrogen bonds to a water molecule and the main chain oxygen of B18 Val (molecule 2) fix the position of the hydroxyl group. This binding site is probably only half-occupied, and there is only one type IV site per dimer since the side chain of A14 Tyr (molecule 2) is disordered. In total, there are nine phenol molecules in the B28Asp-*phn* hexamer and a maximum of 15 *m*-cresol molecules in the B28Asp-*mcr* hexamer.

*Structure of Asp<sup>B28</sup>-Lys<sup>B29</sup>-Ala-Ala-Lys-Gly<sup>A1</sup> Insulin with m-Cresol (B28Asp-X-mcr).* Using crystallization conditions similar to those of B28Asp-*phn* and B28Asp-*mcr* insulins (Table 1), this cross-linked insulin analogue crystallizes in space group  $P2_1$  with one zinc-containing hexamer per asymmetric unit. As in the other structures, this R<sub>6</sub> hexamer accommodates six *m*-cresol molecules, occupying type I binding sites at the dimer-dimer interfaces, which interact in a fashion similar to that for the molecules in B28Asp-*mcr* insulin. The two zinc ions on the hexamer 3-fold axis are approximately 15 Å apart and exhibit tetrahedral coordination geometry consisting of three B10 His side chains and a chloride ion at a distance of 2.2 Å.

Two of the dimers in the hexamer are very similar to each other and to those of the un-cross-linked B28 Asp insulin described above. The rms difference on all atom positions (excluding residues B1 and B2) for these two dimers is 0.7 Å. However, in the third dimer, residues B23–B30 of both monomers are displaced by 1.4–5.3 Å because of the binding of two additional *m*-cresol molecules at the monomer-monomer interface (Figure 5). This mode of binding, referred to as type V, is characterized by many van der Waals contacts between the *m*-cresol molecules and the dimer-forming residues, including ring-stacking interactions with the side chains of residues B16 Tyr, B24 Phe, and B26 Tyr, and nonpolar contacts between the methyl groups of the

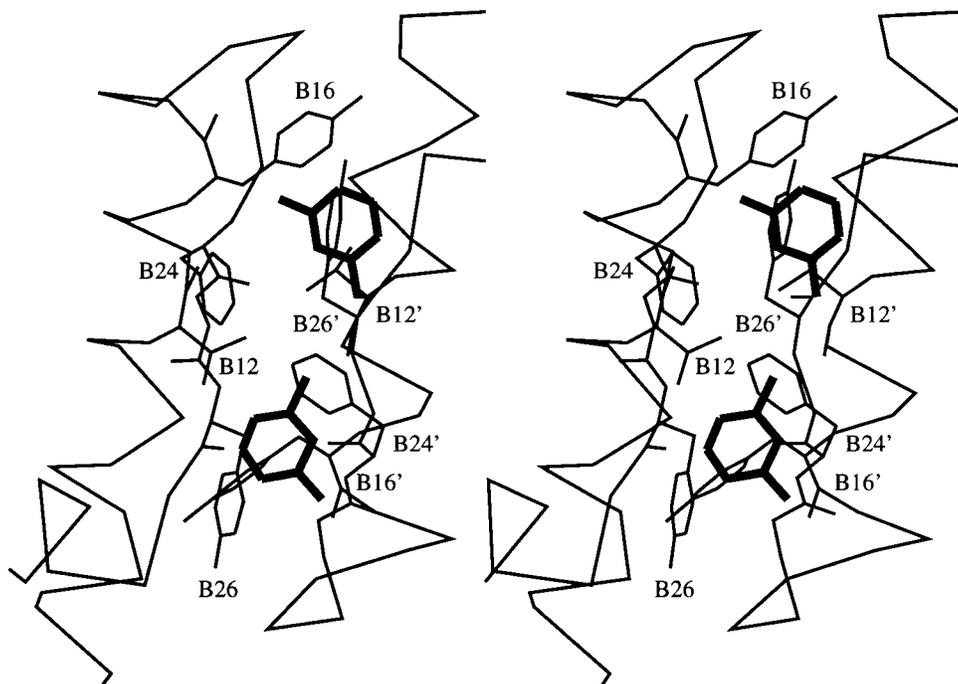


FIGURE 5: Type V *m*-cresol binding in one of the dimers of the B28Asp-X-*mcr* (cross-linked) insulin hexamer. The *m*-cresol molecules, shown in heavy black lines, are surrounded by hydrophobic side chains at the monomer–monomer interface, including residues B12 Val, B16 Tyr, B24 Phe, and B26 Tyr. The residue labels are either primed or unprimed to distinguish between the two monomers.

*m*-cresol molecules and the side chains of residues A12 Val, B12 Val, B15 Leu, B24 Phe, and B26 Tyr. As a result of these interactions, the four  $\beta$ -strand hydrogen bonds thought to be crucial for dimer formation no longer exist.

In each of the six monomers, an Ala-Lys bridge joining residues B30 Ala and A1 Gly adopts a helical conformation, forming an extension to the A chain N terminus  $\alpha$ -helix. From dimer to dimer, the variation in quality of the electron density associated with the cross-links suggests that some are more ordered than others. Those in the third dimer are particularly well-organized and most  $\alpha$ -helical in form as a result of the displacement of the B chain C termini associated with *m*-cresol binding. Near each of the six cross-links, residue B28 Asp is situated in an exposed position, marking the transition from  $\beta$ -strand to  $\alpha$ -helix in this part of the molecule. A comparison of the B28 Asp insulin cross-links with those in an equivalent B28 Pro insulin (unpublished results) shows that the additional chain flexibility introduced by the B28 Pro  $\rightarrow$  Asp substitution is very important in facilitating helix formation and ordering of the cross-link. The fact that the data for this structure were collected at room temperature, rather than at 120 K, might have had some influence on the quality of the electron density for the flexible cross-link regions, although it is likely that the overall observations would be the same at either temperature.

## DISCUSSION

An amino acid substitution B28 Pro  $\rightarrow$  Asp at the dimer-forming surface of the insulin monomer has proved to be effective in destabilizing the dimer, giving rise to an essentially monomeric insulin at physiological concentrations. According to solution studies, the association constant for B28 Asp insulin is reduced 200-fold relative to that of native insulin (4). Furthermore, this mutation has the advantage of leaving the insulin in an active form since residues B26–

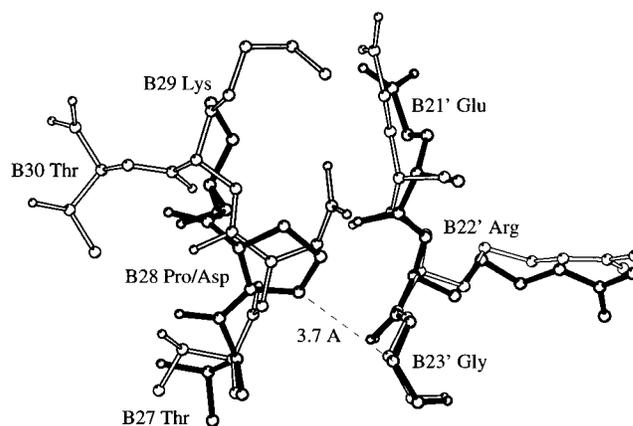


FIGURE 6: Comparison of the dimers of B28 Asp-*phm* insulin (white bonds) and the native R<sub>6</sub> insulin hexamer (black bonds), showing the effect of the B28 Pro  $\rightarrow$  Asp mutation. The dashed lines shows the close van der Waals contact between residues B28 Pro and B23 Gly in the native insulin dimer. This figure was made using BOBSCRIPT (R. Esnouf, unpublished), a modified version of MOLSCRIPT (30).

B30 are not implicated in receptor binding. The crystal structures of B28 Asp insulin show that the introduction of an aspartic acid residue at position B28 allows more flexibility at the end of the B chain and removes vital van der Waals contacts between B28 Pro and B23' Gly at the monomer–monomer interface, while not affecting the overall insulin fold (Figure 6). Other modifications of the B chain C terminus have proved to be very effective in disrupting dimer formation. For instance, in monomeric insulin Lys<sup>B28</sup>-Pro<sup>B29</sup> human insulin, the lysine side chain at position B28 introduces steric hindrance between the monomers (20). Despite the monomeric character of B28 Asp insulin, it forms hexamers in the presence of zinc ions and phenolic molecules, which has considerable advantages for therapy. The phenolic additives provide protection against microbes and,

Table 4: Summary of Phenol and *m*-Cresol Binding in the B28 Asp Insulin Hexamer

type of site	location of site	ligand (occupancy) <sup>a</sup>	insulin
I	dimer–dimer interface, near B5 His	phenol (1.0)	B28Asp- <i>phn</i>
		<i>m</i> -cresol (1.0)	B28Asp- <i>mcr</i>
		<i>m</i> -cresol (1.0)	B28Asp-X- <i>mcr</i>
II	dimer–dimer interface, near B5 His and B10 His	phenol (1.0)	B28Asp- <i>phn</i>
		<i>m</i> -cresol (0.5)	B28Asp- <i>mcr</i>
III	near B28 Asp (molecule 1)	<i>m</i> -cresol (1.0)	B28Asp- <i>mcr</i>
IV	dimer–dimer interface, near A14 Tyr (molecule 1)	<i>m</i> -cresol (0.5)	B28Asp- <i>mcr</i>
V	monomer–monomer interface, between $\beta$ -strand residues	<i>m</i> -cresol (1.0)	B28Asp- <i>mcr</i>

<sup>a</sup> 1.0 represents a fully occupied site; 0.5 represents a half-occupied site.

combined with zinc ions, promote hexamer formation, thereby protecting the insulin molecules from physical and chemical degradation and inhibiting the formation of insoluble fibrils (25). Kinetic solution studies have shown that the R<sub>6</sub> hexamer has far greater stability than the T<sub>6</sub> or T<sub>3</sub>R<sub>3</sub> hexamers (7, 26), and in the case of genetically engineered monomeric insulins, phenols assist in hexamer assembly, their stabilizing effect on the hexamer compensating for the destabilizing effect of the mutation (22).

These studies reveal a complex pattern of phenol and *m*-cresol binding which has been summarized in Table 4. All three structures consist of R<sub>6</sub> hexamers as a result of phenol or *m*-cresol in both type I binding sites at the dimer–dimer interface. Type II binding, which sometimes accompanies type I binding, was first observed in the structure of T<sub>3</sub>R<sub>3</sub> 4-hydroxybenzamide insulin (24); this ligand conveniently fills a small cavity at the dimer–dimer interface. Its buried location suggests that it enters the site during formation of the hexamers in solution; however, it is not likely that type II phenol binding is essential for the formation of the R<sub>6</sub> hexamer since this ligand is not as closely associated with the B1–B8  $\alpha$ -helix as the type I ligand. Like the type I binding site, the type II site has the versatility to accommodate different ligands, including phenol, *m*-cresol, and larger derivatives such as 4-hydroxybenzamide. Type III and type IV *m*-cresol binding interactions occur on the surface of the hexamer, and consequently, the ligands are a little less well ordered (as indicated by their temperature factors) than those at the dimer–dimer interfaces. In the case of type III binding, the extra chain flexibility introduced by the B28 Pro  $\rightarrow$  Asp mutation allows the *m*-cresol to interact with the surface-accessible aromatic side chain B26 Tyr, causing a considerable displacement of the nearby residues (Figure 4). The van der Waals contacts between the methyl group of the ligand and the side chain of B26 Tyr appear to be required to make the chain movement worthwhile, since type III binding does not occur in the equivalent phenol-containing structure, B28Asp-*phn*.

In the cross-linked insulin hexamer, the binding of two *m*-cresol molecules at one of the monomer–monomer interfaces (type V binding) involves the disruption of some key hydrogen bonding interactions fundamental to dimer formation. This is compensated for by extensive ring stacking and van der Waals contacts between the *m*-cresol molecules and the surrounding residues (Figure 5). A

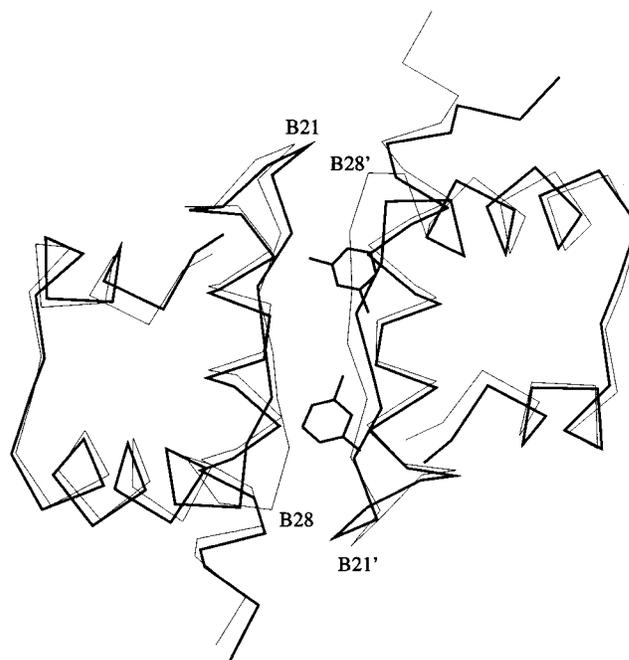


FIGURE 7: Comparison of two B28Asp-X-*mcr* dimers showing how the binding of two *m*-cresol molecules at the monomer–monomer interface of one of the dimers disrupts the  $\beta$ -strand. Residues B21 and B28 have been labeled for reference (primed or unprimed to show different monomers). Note the variations in the positions or the cross-links in the region of B28 Asp.

comparison of the three dimers of B28Asp-X-*mcr* insulin, with and without *m*-cresol, reveals significant displacements of the cross-links which in the presence of *m*-cresol gives rise to well-ordered  $\alpha$ -helical cross-links (Figure 7). In the absence of *m*-cresol, the position of the B chain C terminus is such that the associated cross-link is strained and disordered. These displacements probably reflect dynamic variations of the cross-link in solution, and suggest that in the cross-linked insulin the monomer–monomer interactions are weakened not only by the B28 Pro  $\rightarrow$  Asp mutation but also by a spring-like strain on the B chain C terminus produced by the cross-link. This would explain why type V *m*-cresol binding occurs in B28Asp-X-*mcr* insulin but not in B28Asp-*mcr* insulin. Another significant effect of the cross-link is related to activity. For normal insulin activity, it is necessary for the B chain C terminus to move extensively, revealing internal hydrophobic residues (11). However, the cross-link, by imposing a tight restraint on the B chain C terminus, reduces the activity of the cross-linked insulin to less than 1% of that of native insulin. This is in contrast to the B28 Pro  $\rightarrow$  Asp substitution which has no significant effect on activity.

The diversity of phenol and *m*-cresol binding interactions described in these studies is not surprising given the size and amphipathic nature of these molecules. Indeed, their resemblance to a tyrosine side chain equips them for interacting with protein structures. Of the two ligands, *m*-cresol shows the greatest versatility in binding to the protein (Table 4), yet solution studies have found that phenol is more effective than *m*-cresol in driving the T  $\rightarrow$  R transition (27, 28). This apparent contradiction is a reflection of the complexities of a system in which both thermodynamics and kinetics are important. In these studies, the interactions of phenol and *m*-cresol with the protein have

served to illuminate the effect of the B28 Pro → Asp substitution and the cross-link with regard to chain flexibility and have demonstrated that in the case of insulin weak association properties can be overcome with the use of appropriate ligands. Recent solution studies have shown that more complex phenols, such as 2,6- or 2,7-dihydroxynaphthalene, bind in the hexamer more tightly than phenol, thereby enhancing the stability of the R<sub>6</sub> hexamer still further (29). This result may have applications in the future in the continued development of monomeric, rapid-acting insulins.

#### ACKNOWLEDGMENT

We acknowledge Jens Brange as the source of the original idea for the B28 Pro → Asp mutation, and we thank Jan Markussen and Svend Havelund for provision of the B28 Asp insulin. We are also indebted to Xiao Bing, Vibeke Larsen, and Wendy Offen for carrying out the initial crystallization trials and to Anthony Wilkinson and Leo Caves for useful discussions during preparation of the manuscript.

#### REFERENCES

1. Brange, J. (1997) *Diabetologia* 40, S48–S53.
2. Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Harding, M. M., Hodgkin, D. C., Rimmer, B., and Sheat, S. (1969) *Nature* 224, 491–495.
3. Baker, E. N., Blundell, T. L., Cutfield, J. F., 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., and Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London* 319B, 369–456.
4. Brems, D. N., Alter, L. A., Beckage, M. J., Chance, R. E., DiMarchi, R. D., Green, L. K., Long, H. B., Pekar, A. H., Shields, J. E., and Frank, B. H. (1992) *Protein Eng.* 5, 527–533.
5. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., and Swenson, D. (1989) *Nature* 338, 594–596.
6. Smith, G. D., and Dodson, G. G. (1992) *Biopolymers* 32, 441–445.
7. Kaarsholm, N. C., Hui-Chong, K., and Dunn, M. F. (1989) *Biochemistry* 28, 4427–4435.
8. Markussen, J., Hougaard, P., Riebel, U., Sørensen, A. R., and Sørensen, E. (1987) *Protein Eng.* 1, 205–213.
9. Brange, J., Ribel, U., Hansen, J. F., Dodson, G. G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R., and Voigt, H. O. (1988) *Nature* 333, 679–682.
10. Brange, J., Owens, D. R., Kang, S., and Vølund, A. (1990) *Diabetes Care* 13, 923–954.
11. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Xiao, B., and Markussen, J. (1991) *J. Mol. Biol.* 220, 425–433.
12. Harding, M. M., Hodgkin, D. C., Kennedy, A. F., O'Connor, A., and Weitzmann, P. D. J. (1966) *J. Mol. Biol.* 16, 212–226.
13. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, in *Methods in Enzymology, Volume 276: Macromolecular Crystallography, Part A* (Carter, C. W., and Sweet, R. M., Eds.) pp 307–326, Academic Press, New York and London.
14. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D* 50, 760–763.
15. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. (1976) *Nature* 261, 166–168.
16. Murshudov, G. N., Dodson, E. J., and Vagin, A. A. (1997) *Acta Crystallogr. D* 53, 240–255.
17. Molecular Simulations Inc. (1996) *QUANTA*, 9685 Scranton Rd., San Diego, CA 92121-3752.
18. Navaza, J. (1994) *Acta Crystallogr. A* 50, 157–163.
19. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science* 235, 458–460.
20. Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D., and Smith, G. D. (1995) *Structure* 3, 615–622.
21. Whittingham, J. L., Chaudhuri, S., Dodson, E. J., Moody, P. C. E., and Dodson, G. G. (1995) *Biochemistry* 34, 15553–15563.
22. Bakaysa, D. L., Radziuk, J., Havel, H. A., Brader, M. L., Shun, L., Dodd, S. W., Beals, J. M., Pekar, A. H., and Brems, D. N. (1996) *Protein Sci.* 5, 2521–2531.
23. Smith, G. D., and Ciszak, E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 8851–8855.
24. Smith, G. D., Ciszak, E., and Pangborn, W. (1996) *Protein Sci.* 5, 1502–1511.
25. Brange, J. (1994) Studies on insulin stability: C. Physical stability, in *Stability of Insulin*, pp 18–23, Kluwer Academic Publishers, Lancaster, U.K.
26. Rahuel-Clermont, S., French, C. A., Kaarsholm, N. C., and Dunn, M. F. (1997) *Biochemistry* 36, 5837–5845.
27. Krüger, P., Gilge, G., Cabuk, Y., and Wollmer, A. (1990) *Biol. Chem. Hoppe-Seyler* 371, 669–673.
28. Choi, W. E., Brader, M. L., Aguilar, V., Kaarsholm, N. C., and Dunn, M. F. (1993) *Biochemistry* 32, 11638–11645.
29. Bloom, C. R., Heymann, R., Kaarsholm, N. C., and Dunn, M. F. (1997) *Biochemistry* 36, 12746–12758.
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950.

BI980807S