Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric $Lys^{B28}Pro^{B29}$-human insulin

Ewa Ciszak[1], John M Beals[2], Bruce H Frank[2], Jeffrey C Baker[2], Nancy D Carter[2] and G David Smith[1,3]*

[1]Hauptman-Woodward Medical Research Institute, Inc., 73 High Street, Buffalo, NY 14203, USA, [2]Eli Lilly Research Laboratory, Indianapolis, IN 46285, USA and [3]Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA

Background: $Lys^{B28}Pro^{B29}$-human insulin (Humalog™), a fully potent insulin analog in which the prolyl, lysyl sequence at the C-terminal end of the B-chain is inverted, exhibits a decreased association of monomers to dimers leading to rapid *in vivo* absorption. This provides important benefits for the insulin-requiring diabetic. In spite of its monomeric nature, $Lys^{B28}Pro^{B29}$-human insulin can exist as a discrete hexameric structure in the presence of both zinc and phenol. Studies of the crystal structure of $Lys^{B28}Pro^{B29}$-human insulin in a hexameric complex were initiated to gain a molecular understanding of the effect of the sequence inversion on the analog's self-association properties and, consequently, its *in vivo* efficacy.
Results: Under the conditions reported, $Lys^{B28}Pro^{B29}$-human insulin crystallized as a $T_3R_3^f$ hexamer that is isomorphous with the uncomplexed $T_3R_3^f$ native human insulin hexamer previously known as '4Zn insulin'. The three-dimensional structure of the $T_3R_3^f$ hexamer was determined by X-ray crystallographic methods to a resolution of 2.3 Å. The prolyl, lysyl sequence inversion leads to local conformational changes at the C termini of the B-chains which eliminate two critical hydrophobic interactions and weaken two terminal β-sheet hydrogen bonds that stabilize the dimer.
Conclusions: The loss of these native dimer interactions weakens the hexameric $Lys^{B28}Pro^{B29}$-human insulin complex formed in the presence of phenolic ligands. Thus, it is hypothesized that the diffusion of the phenolic ligands from the site of injection results in the dissociation of hexamers directly to monomers, thereby maintaining the rapid time-action of the monomeric analog in spite of the hexameric conformation in therapeutic formulations.

Introduction

The human insulin molecule consists of a 21 amino acid A-chain and a 30 amino acid B-chain, linked by two disulfide bonds. The association of insulin monomers to dimers and larger aggregates is a complex, dynamic equilibrium and is a function of pH, ionic strength and protein concentration. Insulin is capable of binding various divalent metal ions although it is synthesized and stored in the pancreas as a Zn^{2+}-containing hexamer. X-ray crystallographic studies of hexameric insulin have shown the insulin molecule to be very flexible, as discrete conformational states have been observed for the N terminus of the B-chain [1–4]. It has been suggested that conformational changes of the N terminus can be related to loss of Zn^{2+} and the subsequent dissociation of the hexamers to the biologically relevant monomers [2]. Moreover, the C terminus of the B-chain is believed to possess inherent flexibility, which is critical for self-association of insulin molecules and the binding of insulin to its receptor [5–7], as evidenced by the loss of activity of the single-chain B29–A1 peptide-linked insulin. The dimer is stabilized by hydrogen bonds in a short antiparallel β-sheet consisting of the B-chain C-terminal residues Phe^{B24}–Tyr^{B26}. The importance of the C terminus of the B-chain for self-association is shown by removal of residues B26–B30, despentapeptide insulin, which cannot dimerize through this region of the molecule [8]. Although removal of B29 and B30 does not significantly reduce association, the removal or replacement of Pro^{B28} does alter these properties, which strongly suggests that Pro^{B28} is critical to dimer stabilization and hence hexamer formation [9].

$Lys^{B28}Pro^{B29}$-human insulin, an insulin analog created by an amino acid inversion of Pro^{B28} to lysine and Lys^{B29} to proline, has been used to investigate further the role of the C-terminal B-chain residues in self-association. This inversion has been found to diminish the dimerization constant by a factor of 300, and hence hexamer formation [9]. However, in spite of its greatly diminished self-association properties, this monomeric analog can be stabilized in discrete hexameric structures by the addition of both zinc and phenol. Here we report the structure of a hexameric complex generated in the presence of zinc and phenol and identify specific changes in intermolecular interactions in the dimer that are responsible for the monomeric properties of this analog. In addition, these structural results also suggest an explanation for the ability of $Lys^{B28}Pro^{B29}$-human insulin formulations (containing both zinc and phenolic preservatives) to retain the time-action of a monomeric analog, that is, rapid onset of activity, in spite of the hexameric association state in the pharmaceutical preparation.

*Corresponding author.

Results and discussion

Monomer nomenclature

$Lys^{B28}Pro^{B29}$-human insulin crystallizes as a hexamer, illustrated in Figure 1, that contains two zinc ions and three phenol molecules. The hexamer is composed of three insulin dimers arranged about a central crystallographic threefold axis. The two independent monomers which comprise each dimer differ significantly in conformation, especially at the N termini of their B-chains (Fig. 2). The differences in conformation at the first eight residues of the N termini of the B-chains have been described by the T and R nomenclature, where T and R refer to an extended or an α-helical conformation, respectively [10]. Residues $Phe^{B1.1}$–$Gly^{B8.1}$ (where the 1 or 2 in the decimal portion of the chain name and residue number refers to the T-state monomer or the R-state monomer, respectively) of the zinc complex of $Lys^{B28}Pro^{B29}$-human insulin shown in Figure 2a are very similar to those observed in the 2Zn insulin (T_6)
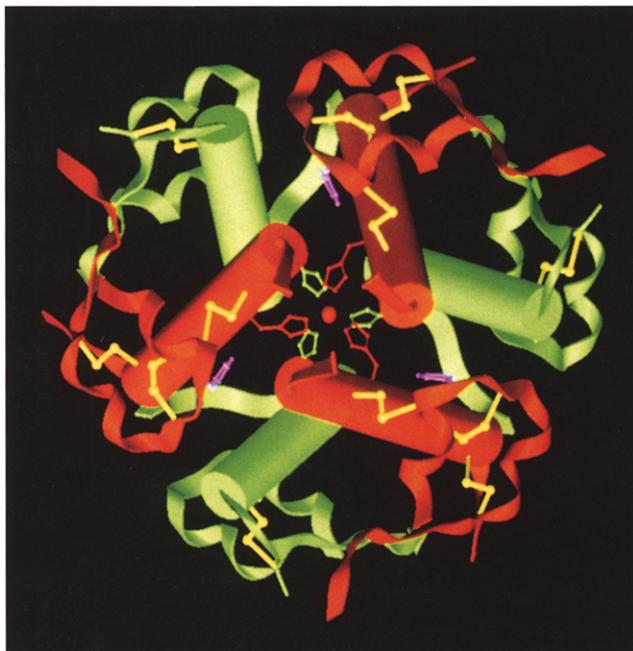
Fig. 1. The $Lys^{B28}Pro^{B29}$-human insulin hexamer viewed along the threefold axis. R^f-state monomers are colored red, T-state monomers green, phenol molecules magenta, disulfide bonds yellow and zinc red. As the hexamer is viewed along the threefold axis, only one zinc ion is visible. (Drawn with SETOR [20].)
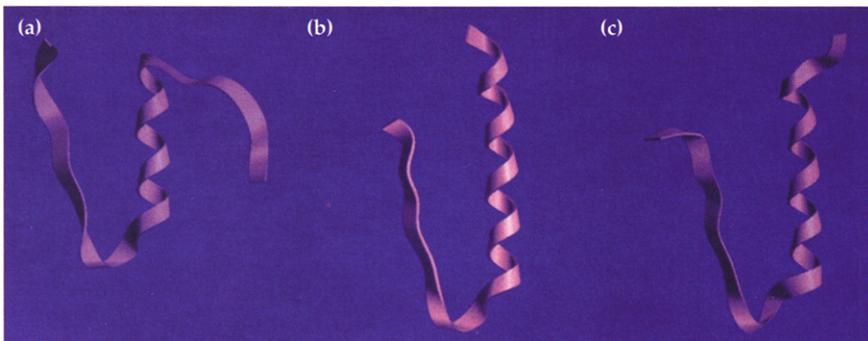
structure [1]. The R-state conformation observed in R_6 insulin hexamers [4,11], in which residues B1–B8 are α-helical, is illustrated in Figure 2b. Unlike the pure R-state conformation, in which there is a continuous α-helix from B1 to B19, a continuous α-helix is present only for residues B4.2 to B19.2 in $Lys^{B28}Pro^{B29}$-human insulin. This 'frayed' α-helical conformation is defined as the R^f-state in order to distinguish it from that observed in R_6 hexamers, and is shown in Figure 2c. The $Lys^{B28}Pro^{B29}$-human insulin hexamer is therefore classified as $T_3R_3^f$. However, this R^f conformational state and the resulting $T_3R_3^f$ hexamer are not unique to this insulin analog. An analogous conformation has been observed in the hexameric insulin previously referred to as 4Zn insulin [3], in the insulin–p-hydroxyacetanilide complex [12], and in the insulin–p-hydroxybenzamide complex [13], all of which are isostructural with the $Lys^{B28}Pro^{B29}$-human insulin analog. The R^f nomenclature had not been explicitly defined at the time these structures were reported, however, and they were classified as simply R. In the R^f-conformational state the α-helix is disrupted at $Asn^{B3.2}$ by a positive ψ angle of $111°$. The $Asn^{B3.2}$ side chain is in a position where it can accept hydrogen bonds from the main-chain nitrogen atoms of both $His^{B5.2}$ and $Leu^{B6.2}$, and thus act as a surrogate for the first two residues of the B-chain in the R conformational state. Therefore, the extended conformation of the first three residues is energetically equivalent to the α-helical conformation, as the number of hydrogen bonds remains unchanged.

Hexamer–hexamer interactions

Each $T_3R_3^f$ $Lys^{B28}Pro^{B29}$-human insulin hexamer can be thought of as constructed from two threefold symmetric trimers, T_3 and R_3^f, which are assembled around the threefold axis. The surface of each trimer is different because of the conformational differences of the N-terminal B-chain residues. The two Zn^{2+} ions (Zn1 and Zn2) lie on the threefold axis and each is coordinated by three symmetry-related His^{B10} residues of the T- or R^f-state monomers. Zn1 lies at the bottom of a shallow depression on the surface of the T_3 trimer, where it is octahedrally coordinated by three $His^{B10.1}$ residues and three water molecules (Zn1–Nϵ2, 2.15 Å; Zn1–OW1, 3.53 Å). In the R^f trimer, Zn2 is also bound by three symmetry-related $His^{B10.2}$ residues (2.10 Å), but as a result of the presence of the α-helix resulting from the T$\rightarrow$R^f transition (due to the presence of chloride ion or

Fig. 2. Conformation of the B-chains in the (a) T-state, (b) R-state and (c) R^f-state monomers.
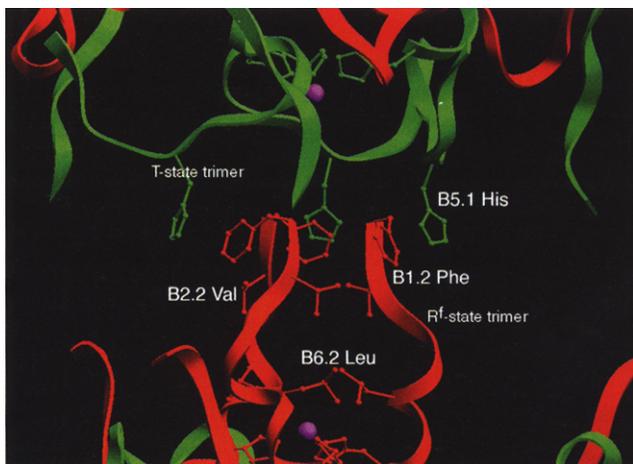
Fig. 3. Ribbon representation of the inter-hexamer interactions at the T and R^f surfaces. T-state monomers and the side chain of His$^{B5.1}$ are illustrated in green. The R^f-state monomers and the side chains of Phe$^{B1.2}$, Val$^{B2.2}$, and Leu$^{B6.2}$ are illustrated in red. Also illustrated are the two independent zinc ions, colored purple.

phenolic derivatives), there is insufficient space for octahedral coordination. Consequently, the coordination at this site is restricted to tetrahedral, the fourth ligand being a chloride ion, which is also located on the threefold axis at a distance of 2.31 Å from Zn2.

As hydrogen bond donation near the N terminus of the R or R^f α-helical segments is satisfied for either conformation, the selection of the R^f conformation by both human insulin and Lys^{B28}Pro^{B29}-human insulin can be best explained on the basis of the packing of the hexamers along the c-axis. The generation of a pure R-state conformation, in which residues Phe$^{B1.2}$ to Asn$^{B3.2}$ are α-helical, would result in unacceptable interactions between the T and R trimers which could' only be relieved by increasing the length of the c-axis and hence increasing the inter-hexamer separation. This increase would in turn expose the side chain of Phe^{B1} to the solvent as well as increase the solvent content of the crystals. In the R^f conformation, however, residues Phe$^{B1.2}$ to Asn$^{B3.2}$ project away from the R$_3^f$ trimer surface where, together with their threefold symmetry-related counterparts, they form a 12 Å deep channel. The three Phe$^{B1.2}$ side chains nestle into the T$_3$ surface of a translationally (c-axis) related hexamer and are nearly parallel to the T-state His$^{B5.1}$ side chains (Fig. 3). These inter-hexamer interactions between the T$_3$ and R$_3^f$ trimers serve to shield the Phe$^{B1.2}$ and Val$^{B2.2}$ side chains from the surrounding solvent and consequently facilitate efficient packing of the T$_3$ and R$_3^f$ surfaces against each other. Although the Phe$^{B1.2}$ and Val$^{B2.2}$ side-chain conformations are stabilized by hydrophobic interactions, no directed interactions (such as hydrogen bonds) stabilize the N terminus of the B-chain in the R^f-state monomer. Consequently, some conformational flexibility results in this region, as shown by the Cα displacements observed in a comparison of the TR^f Lys^{B28}Pro^{B29}-human insulin dimer and the TR^f 'phenol-less' human insulin dimer (Fig. 4a).

Finally, the alignment of the hexamers along the c-axis results in infinite chains of hexamers in which access to either the T or R^f Zn^{2+}-binding site is blocked by the adjacent hexamer.

TR^f dimer

As stated above, Figure 4a shows a comparison of the conformation of the backbone atoms of the TR^f Lys^{B28}Pro^{B29}-human insulin dimer with that of the TR^f human insulin dimer. Their overall similarity is reflected in average and root mean square (rms) displacements of 0.25 Å and 0.28 Å, respectively, minimizing the displacements of the main-chain atoms of A11–A19 and B11–B19. A comparison of the three disulfide bridges in each monomer of the Lys^{B28}Pro^{B29}-human insulin dimer with those observed in the human insulin TR^f dimer reveals no significant differences.

Even though the sequence inversion at B28 and B29 has little effect upon the overall conformation of the dimer, there are five regions in which significant differences are observed: Thr$^{A8.2}$–Ile$^{A10.2}$, Phe$^{B1.2}$–Asn$^{B3.2}$, Gly$^{B20.2}$–Gly$^{B23.2}$ and Thr$^{B27.2}$–Thr$^{B30.2}$ in the R^f-state monomer and Thr$^{B27.1}$–Thr$^{B30.1}$ in the T-state monomer. Minor displacements (less than 1.0 Å) in the Thr$^{A8.2}$–Ile$^{A10.2}$ sequence have been observed in other T$_3$R$_3^f$ insulin hexamers [3,12,13]; to date, a satisfactory explanation for these differences has not yet been made, nor is one obvious in the present work. The origin of the displacements at Phe$^{B1.2}$–Asn$^{B3.2}$ (approximately 2 Å) have been discussed above, under 'Hexamer–hexamer interactions'. Differences in the B27–B30 region (2–4 Å) were expected, as this sequence contains the inversion site. The displacements observed in the B20–B23 region (1–2 Å) are due to the loss of its interactions with residue Pro^{B28}. Moreover, both the B20–B23 and B27–B30 segments lie on the hexamer surface and are responsible for hydrophobic interactions that contribute to the stabilization of the TR^f dimer.

A superposition of the T- and R^f-state monomers of Lys^{B28}Pro^{B29}-human insulin (Fig. 4b), minimizing the displacements of backbone atoms in residues A11–A19 and B11–B19, resulted in an average displacement of 0.54 Å and an rms displacement of 0.61 Å. In Figure 4b, regions Phe^{B1}–Leu^{B6} and Pro^{B29}–Thr^{B30} can easily be identified as having the most significant displacements. The largest conformational difference between the two independent monomers in Lys^{B28}Pro^{B29}-human insulin is associated with the T→R^f transition, in other words, the extension of the α-helix at the N terminus of the B-chain and the 20° rotation of the initial seven residues of the A-chain.

In the human insulin TR^f dimer, the B-chain α-helices terminate at Gly^{B20} and a hydrogen bond associated with the α-helix is observed between the B20 amino nitrogen and the carbonyl oxygen of B16. Residues B20–B23 form a type III β-turn and have nearly identical conformations in both monomers. In contrast, the B20–B23

Fig. 4. A comparison of the displacements of the Cα atoms of (a) the TR^f dimer of Lys^{B28}Pro^{B29}-human insulin with that of the TR^f phenol-less human insulin dimer, and (b) the Lys^{B28}Pro^{B29}-human insulin T-state monomer with the R^f-state monomer.

loops in Lys^{B28}Pro^{B29}-human insulin have significantly different conformations in each of the monomers. In the T-state monomer, this loop adopts a distorted type III β-turn, but in the R^f state, the turn is severely distorted and no hydrogen bond exists between Gly$^{B20.2}$ and Gly$^{B23.2}$ (Table 1). Although the terminal B20–B16 hydrogen bond in the α-helix exists in the T-state monomer, this hydrogen bond is not present in the R^f-state monomer, because of the distortions in the

Table 1. Comparison of torsion angles and hydrogen-bond distances in the B20–B23 loops of Lys^{B28}Pro^{B29}-human insulin with those of human insulin.

	Lys^{B28}Pro^{B29}		Human	
	T-state	R^f-state	T-state	R^f-state
Torsion angles (°)				
φ21	−28	−84	−65	−53
ψ21	−37	−15	−17	−32
φ22	−79	−116	−63	−69
ψ22	−30	−20	−42	−23
Hydrogen-bond distances (Å)				
Gly^{B20} O–Gly^{B23} N	2.72	3.98	3.09	3.19
Tyr^{B16} O–Gly^{B20} N	2.98	3.56	3.13	3.06
Lys$^{B28.1}$ Nζ–Glu$^{B21.2}$ O	2.55		−	

B20–B23 loop, which may be due to the proximity of atoms in an adjacent hexamer or to the formation of a hydrogen bond between Nζ of Lys$^{B28.1}$ and the carbonyl oxygen of Glu$^{B21.2}$ (Table 1).

Although the sequence inversion at B28 and B29 has little effect upon the overall conformation of the dimer, it strongly affects the hydrophobic and hydrophilic interactions between T and R^f monomers. Dimer stabilization in T$_3$R$_3^f$ native insulin is achieved in part by hydrophobic interactions as well as by surface complementarity between the prolyl rings of each monomer and the B20–B23 loops in the other monomer (Fig. 5a). In human insulin, the side chain of Pro^{B28} nestles into the B20–B23 loop of the adjacent monomer, where it makes numerous hydrophobic interactions with main-chain atoms of Gly^{B20}, Glu^{B21} and Gly^{B23}. In Lys^{B28}Pro^{B29}-human insulin, however, the sequence inversion displaces proline towards the C terminus, making hydrophobic interactions between proline and the B20–B23 loop impossible (Fig. 5b). Even if the B-chains made turns at this point as they do in human insulin, the Lys^{B28} side chain is too long to substitute for the prolyl ring and provide the corresponding hydrophobic interactions with the B20–B23 loop of the other monomer. It is interesting to note that the thermal parameters of the T-state

Fig. 5. The interface between the C terminus of the T-state monomer (B26–B30) and the loop of the R^f-state monomer (B19–B24) in **(a)** human insulin where residue label D28PRO refers to $Pro^{B28.2}$ while B20GLY and B23GLY refer to $Gly^{B20.1}$ and $Gly^{B23.1}$, respectively; and **(b)** $Lys^{B28}Pro^{B29}$-human insulin where residue label B228LYS refers to $Lys^{B28.2}$ while B120GLY and B123GLY refer to $Gly^{B20.1}$ and $Gly^{B23.1}$, respectively.

Table 2. Comparison of the β-pleated sheet hydrogen-bond lengths (Å) in the TR^f dimer of $Lys^{B28}Pro^{B29}$-human insulin with those of human insulin.

Monomer I		Monomer II			
Residue	Atom	Residue	Atom	$Lys^{B28}Pro^{B29}$	Human
$Tyr^{B26.1}$	O	$Phe^{B24.2}$	N	3.23 (Å)	2.92 (Å)
$Tyr^{B26.1}$	N	$Phe^{B24.2}$	O	2.89 (Å)	2.93 (Å)
$Phe^{B24.1}$	O	$Tyr^{B26.2}$	N	2.72 (Å)	2.80 (Å)
$Phe^{B24.1}$	N	$Tyr^{B26.2}$	O	3.25 (Å)	2.89 (Å)
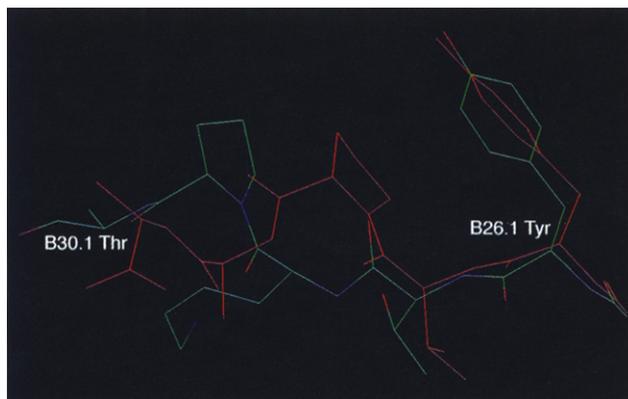
Fig. 6. A superposition of the sequence $Tyr^{B26.1}$–$Thr^{B30.1}$ in human insulin (red) as compared with the same fragment in $Lys^{B28}Pro^{B29}$-human insulin (colored according to atom type).

loop and the R^f-state C terminus are comparable in $Lys^{B28}Pro^{B29}$-human insulin and human insulin, but that there is a substantial increase in the thermal parameters of $Lys^{B28}Pro^{B29}$-human insulin (~12 Å^2) for the R^f-state loop and the T-state C terminus.

The loss of the hydrophobic interactions with the B20–B23 loop also contributes to the weakening of some hydrogen bonds in the β-pleated sheet region (B24–B26) that are critical for stabilizing the dimer. While the inner two hydrogen bonds in the β-pleated sheet are comparable to those observed in human insulin, the outer two hydrogen bonds have lengthened by an average of 0.35 Å (Table 2).

Finally, the displacement of proline eliminates the intramolecular packing of the Pro^{B28} ring against the side chain of Tyr^{B26} in the monomer (Fig. 6). The elimination

of this interaction reduces the conformational stability of the C-terminal residues of both monomers and permits additional flexibility of the C-terminal residues in $Lys^{B28}Pro^{B29}$-human insulin.

Phenol-binding sites

$Lys^{B28}Pro^{B29}$-human insulin binds only three phenol molecules, unlike human insulin, which under similar crystallizing conditions binds six phenol molecules in an R_6 state. Phenol molecules are bound in elliptically shaped binding sites located on the interfaces between R^f monomers through a pair of hydrogen bonds from the phenol hydroxyl group to the carbonyl oxygen of $Cys^{A6.2}$ (2.66 Å) and the nitrogen of $Cys^{A11.2}$ (3.20 Å). This binding pattern is observed in the structures of all reported human insulin–phenol complexes [4,11], and also in complexes with larger phenolic derivatives such as p-hydroxyacetanilide [12] or p-hydroxybenzamide [13]. In $Lys^{B28}Pro^{B29}$-human insulin, that portion of the binding cavity that continues beyond the phenol molecule towards the center of the hexamer is occupied by two water molecules, OW3 and OW33, which form hydrogen bonds to each other (2.77 Å) as well as to the carboxyl group of $Glu^{B13.1}$ (OW33–Oε2, 3.41 Å) and the carbonyl oxygen of $His^{B10.2}$ (OW3–O, 2.95 Å). Both of these water molecules have counterparts in the structure of the phenol-less $T_3R_3^f$ human insulin hexamer [3].
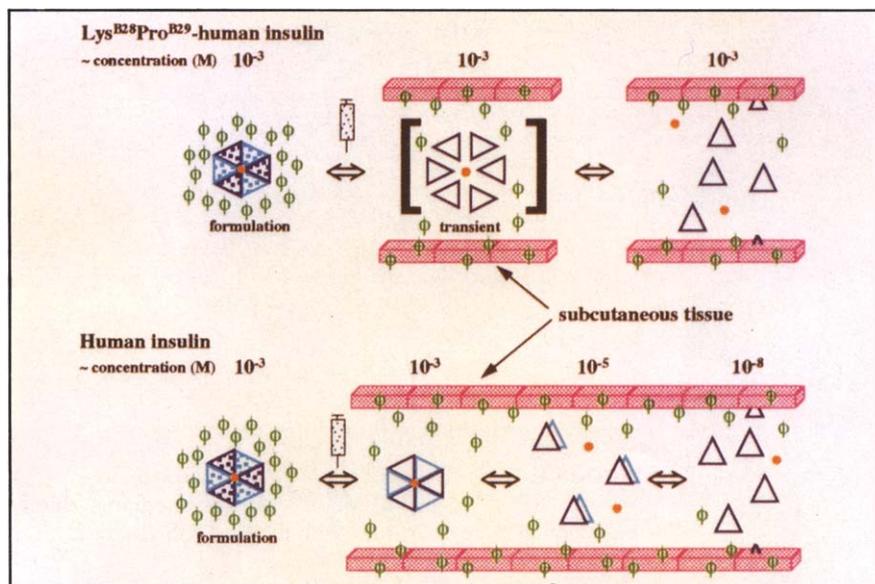
Fig. 7. Schematic comparison of the hypothesized dissociation of Lys^{B28}Pro^{B29}-human insulin and insulin in subcutaneous tissue. The insulin hexamers in the formulation are represented as a dimer of trimers (where a triangle represents an insulin monomer), with one trimer in light blue and the second trimer highlighted as purple. Each trimer coordinates a zinc ion (orange sphere); as the hexamer is viewed along the threefold axis, only one zinc ion is visible. The stippled and unfilled monomers represent the presence and absence of bound phenolic preservative, respectively. The free phenolic preservative is represented as a green ϕ. The subcutaneous tissue is represented as pink boxes. The hexamer dissociates in the subcutaneous tissue and is then absorbed (unpaired triangles) through the capillary membranes in the subcutaneous tissue.

Rapid time-action from hexameric preparations of monomeric analogs

Many therapeutic preparations of human insulin contain Zn^{2+} and a phenolic derivative as a preservative. These preparations are necessarily hexameric (due to the presence of zinc ion) and current models suggest that the rate-limiting step in the absorption of insulin from a subcutaneous injection site is the dissociation of hexamers to the biologically relevant monomers [14]. Clinical studies have verified that monomeric preparations of Lys^{B28}Pro^{B29}-human insulin [15] or other monomeric analogs [14] are more rapidly absorbed into the blood than hexameric human insulin preparations, supporting the hypothesis that the rate-limiting step in insulin absorption is the dissociation of hexamers to dimers and monomers.

However, just as in the case of human insulin therapeutic preparations, monomeric analogs require formulations that yield both suitable chemical and physical stability. Stabilization of Lys^{B28}Pro^{B29}-human insulin is achieved by the addition of zinc and phenol, which act to induce the otherwise monomeric analog to associate into discrete zinc-containing hexameric complexes under pharmaceutically useful concentrations (Bakaysa, D.L., *et al.*, & Radziuk, J., personal communication). We would emphasize that both phenol and zinc are necessary to form discrete Lys^{B28}Pro^{B29}-human insulin hexamers. Surprisingly, in spite of the formation of this hexameric state, recent clinical data show that the time-action of Lys^{B28}Pro^{B29}-human insulin is not significantly affected [15].

The retention of the rapid time-action of this analog can be explained in part by comparing the subtle structural differences that are observed between the $T_3R_3^f$ crystal structures of Lys^{B28}Pro^{B29}-human insulin and native human insulin. In the latter structure, two sets of interactions involving residues B20–B23 of each monomer contribute to the stabilization of the dimer: firstly,

hydrophobic interactions between each Pro^{B28} and the B20–B23 loop of the other monomer in the dimer, and secondly, four hydrogen bonds across the β-pleated sheet. The sequence inversion at B28–B29 results in poorer surface complementarity, which disrupts the hydrophobic stabilizing effect of Pro^{B28} and contributes to a lengthening of the hydrogen bonds at each end of the β-pleated sheet by approximately 0.35 Å. Consequently, these weakened dimer-stabilizing interactions coupled with the requirement of both phenolic preservative and zinc for the stabilization of the Lys^{B28}Pro^{B29}-human insulin hexamer explain how a monomeric analog formulated as a hexamer can retain its rapid time-action, that is, the ability to readily dissociate into a rapidly absorbed monomeric compound (Fig. 7).

We hypothesize that after a subcutaneous injection the excess, unbound phenolic preservatives are rapidly absorbed into the surrounding tissue. The loss of this excess preservative, coupled with the dynamic structural changes at the N terminus of the B-chain ($R^f{\rightarrow}T$), shifts the phenol-binding equilibrium, destabilizing the Lys^{B28}Pro^{B29}-human insulin hexamer. The hexamers then readily dissociate into monomers as a result of the weakened dimer-stabilizing interactions. The dissociation of the Lys^{B28}Pro^{B29}-human insulin hexamers differs from that of human insulin (because of their differential stability in the absence of phenol) such that the diffusion of phenolic preservatives from a human insulin hexamer would not alter the association state of the protein, but only the conformational state ($R{\rightarrow}T$). Thus, in the case of native human insulin, elimination of the zinc and dilution of the insulin is required to dissociate the hexamer first to dimers and subsequently to monomers.

Finally, the T and R^f conformations, which fall within the envelope of structures described for the insulin molecule, show that the flexibility of the insulin molecule is retained in this analog. This is important not only from a

structural perspective but also a biological perspective as insulin's flexibility has been related to its activity [6]. Based on this crystal structure, it can be seen that none of the residues implicated in receptor binding undergoes changes in conformation and consequently the analog maintains full biological potency, that is, the ability to elicit a full biochemical response [1].

Biological implications

The dissociation of zinc–insulin hexamers into monomers is hypothesized to be the rate-limiting step in the absorption of insulin from a subcutaneous injection. Thus, to accelerate this absorption process, a monomeric insulin analog was developed by inverting the prolyl, lysyl sequence at residues 28 and 29 of the human insulin B-chain, yielding $Lys^{B28}Pro^{B29}$-human insulin. This insulin analog displays monomeric properties under *in vivo* conditions as a result of a decreased propensity of the analog to dimerize. However, in spite of a ~300-fold decrease in the dimerization constant, $Lys^{B28}Pro^{B29}$-human insulin can still associate into discrete stable hexamers under conditions used to stabilize the therapeutic preparation (i.e. in the presence of both phenolic preservatives and zinc).

Interestingly, clinical studies have shown that the time-action of $Lys^{B28}Pro^{B29}$-human insulin remains unaltered in spite of the formation of discrete hexamers. The solution of the $T_3R_3^f$ crystal structure (where T refers to an extended conformation of the first eight residues of the B-chain, and R^f describes an extended conformation for the first three residues and an α-helical one for residues four to eight), prepared in the presence of zinc and phenol, has provided an explanation as to why the rapid time-action of $Lys^{B28}Pro^{B29}$-human insulin in the hexameric state is unaltered in the therapeutically acceptable preparation. The $T_3R_3^f$ crystal structure, although isomorphous with that of $T_3R_3^f$ human insulin, clearly exhibited localized structural differences in the C-terminal region of the B-chain. These differences result in the elimination of two critical hydrophobic interactions involving Pro^{B28} and a weakening of two β-pleated sheet hydrogen bonds that stabilize the dimer subunits that compose the hexamer. These localized structural changes coupled with the requirement of phenol and zinc for the stabilization of the hexamer explain why the formulated $Lys^{B28}Pro^{B29}$-human insulin hexamer can readily dissociate in the subcutaneous tissue, yielding the monomeric form of the analog which is rapidly absorbed. We hypothesize that the absorption of the phenolic preservative destabilizes the hexamer, allowing it to dissociate into monomers at millimolar concentrations.

Materials and methods

Crystallization

Biosynthetic $Lys^{B28}Pro^{B29}$-human insulin was provided by the Biosynthetic Development Division of the Lilly Research Laboratories. Single crystals from which the 2.5 Å data were measured were grown from a solution containing 0.27 mM $Lys^{B28}Pro^{B29}$-human insulin, 0.12 M acetic acid, 0.05 M sodium citrate in the presence of Zn^{2+} and phenol at a pH of 5.9, by slow cooling from 50°C to room temperature. As crystallization experiments were continued, larger single crystals grew under similar conditions, in which the 0.12 M acetic acid was replaced by 0.12 M hydrochloric acid; these crystals were used to measure the 2.3 Å data.

X-ray data collection

The two sets of diffraction data were measured using an R-AXIS II-C image plate system and RU-200 rotating anode generator with CuKα radiation. Integration, interframe scaling, and merging were performed with the R-AXIS software.

The 2.5 Å data were measured from a single crystal of dimensions 0.20 mm × 0.15 mm × 0.15 mm and a total of 3306 measurements yielded 2653 unique data $[R_{merge}(F^2)=0.069]$. The crystals belong to space group R3 and were indexed in a hexagonal unit cell. The volume of the unit cell of dimensions a=b=79.62 Å, c=37.78 Å is consistent with the presence of a dimer in the asymmetric unit, and the crystals are nearly isomorphous with crystals of uncomplexed $T_3R_3^f$ human insulin [3]. A second set of diffraction data was measured from a crystal of dimensions 0.20 mm × 0.30 mm × 0.30 mm, to a resolution of 2.3 Å. A total of 19299 data yielded 3910 unique data $[R_{merge}(F^2)=0.060]$.

Structure solution and refinement

The coordinates of a $T_3R_3^f$ human insulin dimer [3], excluding residues B1, B2 and B27–B30 of both B-chains, zinc, chloride ions, and water molecules, were used to calculate an initial difference electron-density map using the 2.5 Å data. The largest peaks from this map corresponded to a pair of zinc ions and a chloride ion located on the threefold axis. The next largest region of electron density was found at the positions expected for residues Val^{B2} and Thr^{B27}. Initial refinement of the model was performed using the programs X-PLOR [16] and PROFFT [17,18]. When the 2.3 Å data became available, the refinement was stopped at a residual of 0.181.

Table 3. Refinement statistics.

	Model	Target σ
Distances (Å)		
Bond 1–2	0.016	0.020
Angle 1–3	0.054	0.050
Planar 1–4	0.046	0.050
Chiral volume (Å^3)	0.175	0.150
Planar groups (Å)	0.013	0.020
Thermal parameters (Å^2)		
<ΔB> Main chain	2.156	2.500
<ΔB> Side chain	1.555	1.500
Non-bonded distances (Å)		
Single torsion	0.235	0.500
Multiple torsion	0.268	0.500
Possible H-bonds	0.412	0.500
Torsion angles (°)		
Planar	2.1	2.5
Staggered	18.4	20.0
Orthonormal	15.7	20.0

Fig. 8. $2F_o-F_c$ electron density maps contoured at 1σ in the vicinity of the B-chain C-terminal residues of **(a)** the T-state monomer and **(b)** the R^f-state monomer.

A structure factor calculation, using the second data set and the refined coordinates from the 2.5 Å refinement, produced a residual of 0.219. The program PROFFT [17,18] supported by the graphics program CHAIN [19] was used to refine both the positional and thermal parameters of the structure. The missing fragments (Phe^{B1} and Lys^{B28}–Thr^{B30} of both monomers) were located during the refinement based on the 1σ $(2F_o-F_c)$ electron-density maps and were subsequently included into the refinement. Appropriate adjustments were made to main and side chains, and water molecules were added in accord with the criteria of good electron density and acceptable hydrogen bonds to other atoms. No interpretable electron density was observed for the side-chain atoms of Tyr$^{A14.1}$, Thr$^{B30.1}$,

Glu$^{B21.2}$, Thr$^{B27.2}$ and Thr$^{B30.2}$. The refinement converged at a residual of 0.161 for 3548 data $[F_o>2\sigma(F_o)]$ between 8.0 Å and 2.3 Å resolution, and 0.178 for all 3860 data in this range. Details of the refinement are given in Table 3. Illustrated in Figure 8 is the $2F_o-F_c$ electron density in the vicinity of the C termini of the B-chains of both monomers.

The coordinates have been deposited with the Brookhaven Protein Data Bank (entry code 1LPH).

References

1. Baker, E.N., *et al.,* & Vijayan, N.M. (1988). The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Phil. Trans. R. Soc. Lond. [Biol.]* **319**, 369–456.
2. Smith, G.D., Swenson, D.C., Dodson, E.J., Dodson, G.G. & Reynolds, C.D. (1984). Structural stability in the 4-zinc human insulin hexamer. *Proc. Natl. Acad. Sci. USA* **81**, 7093–7097.
3. Ciszak, E. & Smith, G.D. (1994). Crystallographic evidence for dual coordination around zinc in the T_3R_3 human insulin hexamer. *Biochemistry* **33**, 1512–1517.
4. Derewenda, U., *et al.,* & Swenson, D. (1989). Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. *Nature* **338**, 594–596.
5. Mirmira, R.G., Nakagawa, S.H. & Tager, H.S. (1991). Importance of the character and configuration of residues B24, B25, and B26 in insulin–receptor interactions. *J. Biol. Chem.* **266**, 1428–1436.
6. Dodson, E.J., Dodson, G.G., Hubbard, R.E. & Reynolds, C.D. (1983). Insulin's structural behavior and its relation to activity. *Biopolymers* **22**, 281–291.
7. Derewenda, U., Derewenda, Z., Dodson, E.J., Dodson, G.G., Bing, X. & Markussen, J. (1991). X-ray analysis of the single chain B29–A1 peptide-linked insulin molecule. A completely inactive analogue. *J. Mol. Biol.* **220**, 425–433.
8. Bi, R.C., Dauter, Z., Dodson, E.J., Dodson, G.G., Giordano, F. & Reynolds, C. (1984). Insulin's structure as a modified and monomeric molecule. *Biopolymers* **23**, 391–395.
9. Brems, D.N., *et al.,* & Frank, B.H. (1992). Altering the association properties of insulin by amino acid replacement. *Protein Eng.* **6**, 527–533.
10. Kaarsholm, N.C., Ko, H.-C. & Dunn, M.F. (1989). Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin and miniproinsulin. *Biochemistry* **28**, 4427–4435.
11. Smith, G.D. & Dodson, G.G. (1992). Structure of a rhombohedral R6 insulin/phenol complex. *Proteins* **14**, 401–408.
12. Smith, G.D. & Ciszak, E. (1994). The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide. *Proc. Natl. Acad. Sci. USA* **91**, 8851–8855.
13. Smith, G.D., Ciszak, E. & Pangborn, W. (1995). Crystallographic studies of insulin crystals grown in microgravity. In *Proceedings of Conference on NASA Centers for Commercial Development of Space (NASA-CCDS).* (El-Genk, M.S. and Whitten, R.P., eds), pp. 177–182, American Institute of Physics, New York.
14. Brange, J., Owens, D.R., Kang, S. & Volund, A. (1990). Monomeric insulins and their experimental and clinical implications. *Diabetes Care* **13**, 923–954.
15. Howey, D.C., Bowsher, R.R., Brunelle, R.L. & Woodworth, J.R. (1994). [Lys(B28), Pro(B29)]-human insulin: a rapidly absorbed analogue of human insulin. *Diabetes* **43**, 396–402.
16. Brünger, A.T. (1992). *X-PLOR Version 3.1. A system for X-ray crystallography and NMR.* Yale University Press, New Haven and London.
17. Hendrickson, W.A. & Konnert, J.H. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography.* (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp. 13.01–13.25, Indian Academy of Sciences, Bangalore.
18. Finzel, B.C. (1987). Incorporation of fast Fourier transforms to speed restrained least-squares refinement of protein structures. *J. Appl. Crystallogr.* **20**, 53–55.
19. Sack, J.S. (1988). CHAIN — a crystallographic modelling program. *J. Mol. Graphics* **6**, 244–245.
20. Evans, S.V. (1993). SETOR: hardware-lighted three dimensional solid model representations of macromolecules. *J. Mol. Graphics* **11**, 134–138.
21. Connolly, M.L. (1983). Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548–558.