# Solution Structures of the $R_6$ Human Insulin Hexamer[†,‡]

Xiaoqing Chang, Anne Marie M. Jørgensen, Peter Bardrum, and Jens J. Led*

*Department of Chemistry, University of Copenhagen, The H. C. Ørsted Institute, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark*

*Received December 18, 1996; Revised Manuscript Received May 27, 1997*[⊗]

ABSTRACT: The three-dimensional solution structure of the phenol-stabilized 36 kDa $R_6$ insulin hexamer was determined by NMR spectroscopy and restrained molecular dynamics. The hexamer structures were derived using a stepwise procedure. Initially, 60 monomers were obtained by distance geometry from 665 NOE-derived distance restraints and three disulfide bridges. Subsequently, the hexamer structures were calculated by simulated annealing, using 30 hexamers constructed from the best 36 monomer structures as the starting models. The NMR data show that the aromatic ring of residue Phe(B25) can take two different orientations in the solution hexamer: one in which it points inward (molecule 1, about 90%) and one in which it points outward from the surface of the monomer (molecule 2, about 10%). Therefore, two hexamer structures were calculated: a symmetric hexamer consisting of six molecule 1 monomers and a nonsymmetric hexamer consisting of five molecule 1 monomers and one molecule 2 monomer. For each of the six monomers, the restraints used in the calculations of the hexamer structures include, in addition to the intramonomeric restraints, 25 NOEs between insulin and phenol, 23 NOEs and two hydrogen bonds across the dimer interface, nine NOEs across the trimer interface, and five intramonomeric or two intermonomeric NOEs, respectively, specifying the different orientations of the Phe(B25) ring. The coordination of the two Zn atoms was defined by eight distance restraints. Thus, a total of 4394 and 4391 distance restraints, respectively, were used in the two hexamer calculations. The NOE restraints were classified in an iterative process as intra- or intermonomeric on the basis of their consistency or inconsistency with the structure of the monomer. The assignment of the dimer- and trimer-specific NOEs was made using the crystal structure of the $R_6$ hexamer as the starting model. For both solution hexamers, the average backbone rms deviation is 0.81 Å, if the less well-defined N- and C-terminal residues are excluded. The corresponding rms deviations for all heavy atoms are 1.17 and 1.19 Å for the nonsymmetric and symmetric hexamer, respectively. The overall solution structure of the $R_6$ insulin hexamer is compact, rigid, and symmetric and resembles the corresponding crystal structure. However, the extension of the B-chain α-helix, which characterizes the R state, is shorter in the solution structure than in the crystal structure. Also, the study shows that the orientation of the Phe(B25) ring has no effect on the structure of the rest of the molecule, within the uncertainty of the structure determination. The importance of these findings for the current model for the insulin−receptor interaction is discussed.

The determination of the structure and flexibility of the insulin molecule and the elucidation of the importance of these characteristics for the function and aggregation of the hormone have for decades been among the most intriguing problems in structural biology (Blundell et al., 1972; Chothia et al., 1983; Hua et al., 1991).

Within the last few years, studies of the structure−function relationship and flexibility of insulin in solution have gained considerable momentum from NMR[1] spectroscopy. Thus, NMR studies have provided detailed information not only about the three-dimensional solution structures of native insulin and various insulin mutants (Kline & Justice, 1990; Hua & Weiss, 1991; Knegtel et al., 1991; Hua et al., 1991, 1993a; Jørgensen et al., 1992, 1996; Ludvigsen et al., 1994) but also about the flexibility of these insulins in solution and about possible conformational changes that may be associated with the receptor binding of the hormone. Furthermore, insight into the structural basis for the insulin aggregation (Brange et al., 1988; Brems et al., 1992; Jørgensen et al., 1996) has provided a basis for the design of new biologically active mutants that are nonaggregating at pharmaceutical concentrations.

[†] This work was supported by the Danish Technical Research Council (Grants 16-5028-1, 9400446, and 9601137), the Danish Natural Science Research Council (Grants 9400903, 9400351, 9502759, and 9601648), the Ministry of Industry (Grant 85886), Julie Damm's Studiefond, Direktør Ib Henriksens Fond, and Novo Nordisk Fondet.

[‡] The atomic coordinates of the 20 refined hexamer structures and the NMR-derived restraints have been deposited in the Brookhaven Data Bank (PDB ID codes 1AIY and 1AI0 for the symmetric and the nonsymmetric hexamer, respectively).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* July 15, 1997.

[1] Abbreviations: NMR, nuclear magnetic resonance; DQF, double-quantum-filtered; COSY, two-dimensional correlated spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; HSQC, heteronuclear single-quantum coherence; DIPSI, decoupling in the presence of scalar interactions; TPPI, time-proportional phase increment; 2D, two-dimensional; 3D, three-dimensional; FID, free induction decay; $d_{\alpha N}(i,j)$, distance and corresponding NOE connectivity between the α CH proton on residue $i$ and the NH proton on residue $j$ [$d_{\alpha N} = d_{\alpha N}(i,i+1)$, $d_{NN} = d_{NN}(i,i+1)$, $d_{\beta N} = d_{\beta N}(i,i+1)$, and $d_{\alpha\beta}(i,i+3)$ defined accordingly]; $^3J_{H_\alpha H_N}$, three-bond coupling between the α CH proton and the NH proton; CSI, chemical shift index; DGSA, distance geometry-simulated annealing; rms, root mean square; NCS, noncrystallographic symmetry; $T_6$, 2Zn insulin hexamer; $T_3R_3$, 4Zn insulin hexamer; $R_6$, phenol-stabilized 2Zn insulin hexamer; WATERGATE, water suppression by gradient-tailored excitation.

S0006-2960(96)03106-6 CCC: $14.00 

Human insulin consists of a 21-residue A-chain and a 30-residue B-chain, linked by two disulfide bonds. An intrachain disulfide bond also exists in the A-chain. The insulin hexamer, which is formed in the presence of $Zn^{2+}$ ions and at pH values around neutral, shows a considerable flexibility. Thus, it is capable of undergoing transitions between three distinct conformational states (Renscheidt et al., 1984; Wollmer et al., 1987; Brader et al., 1991) designated $T_6$, $T_3R_3$, and $R_6$ (Kaarsholm et al., 1989; Derewenda et al., 1989) on the basis of their ligand binding properties, allosteric behavior, and pseudo point symmetries. In particular, phenol, *m*-cresol, and other phenolic compounds can drive the T → R transition to completion, resulting in an $R_6$ hexamer complexed with the phenols (Derewenda et al., 1989). The principal changes associated with the T → R transition are a transformation of the N-terminal end of the B-chain from extended conformation (T state) to helix (R state) and a movement of the N terminus of the B-chain by more than 30 Å.

The dramatic structural changes that occur in the T → R transition make the $R_6$ insulin hexamer a unique and valuable system in the studies of allosteric conformational changes in proteins in general (Brader et al., 1991) and in insulin in particular. Even though the physiological importance of the T → R transition is unclear, it provides information about the flexibility of the insulin monomer, which may be relevant for its function. Thus, it has been speculated that the T → R transition may mimic changes in the insulin conformation that resemble those associated with its receptor binding (Derewanda et al., 1989; Nakagawa & Tager, 1992; Hua et al., 1993b).

The determination of the three-dimensional structure and the flexibility of the $R_6$ insulin hexamer in solution is, therefore, of great interest. Recently, a provisional DGSA model of an isolated protomer within the $R_6$ solution hexamer was published by Jacoby et al. (1996). However, despite the virtues of NMR spectroscopy as a method for the determination of the solution structures of proteins, it is inherently difficult to determine the solution structures of symmetric multimers such as the $R_6$ insulin hexamer using the method, simply because it is intrinsically impossible to distinguish between intra- and intermonomeric NOEs in regular NOESY spectra. Still, NMR structures have been determined for a series of protein dimers (Clore et al., 1990; Breg et al., 1990; Kay et al., 1991; Jørgensen et al., 1992) either by comparison of the NMR data with the available crystal structures or by using an iterative strategy whereby initial structures are calculated on the basis of a smaller number of unambiguous NOEs. Further alleviation of the ambiguity problem in the studies of protein dimers is provided by the use of asymmetric isotope labeling, either as heterodimers consisting of monomers with different deuterium labeling (Weiss, 1990; Arrowsmith et al., 1990) or, more feasible, by the use of uniformly $^{13}C$-labeled and $^{13}C$-unlabeled monomers (Folkers et al., 1993), and possibly in combination with $^{15}N$ labeling (Clore et al., 1994; Lee et al., 1995). However, even with the latter approach, the determination of the NMR structure of more complex systems, *e.g.* a tetramer, has proven to be difficult (Clore et al., 1995) without prior knowledge of the crystal structure.

Here, we present the three-dimensional solution structure of the complete 36 kDa $R_6$ hexamer of human insulin. The structures of the hexamers were determined in a stepwise process, where first the structure of the monomer and monomer-specific NOEs were obtained in an iterative process and, subsequently, the structure of the $R_6$ hexamer was determined, using the solution structure of an insulin dimer (Jørgensen et al., 1992) and the crystal structure of the $R_6$ insulin hexamer to resolve the NOE ambiguities. Finally, the obtained $R_6$ solution hexamer is compared with the corresponding crystal structure and with the solution structures of dimeric and monomeric mutants of human insulin, and structural changes that are significant for the insulin−receptor interaction are discussed.

## EXPERIMENTAL PROCEDURES

*(a) Sample Preparation.* Biosynthetic 2Zn human insulin was kindly supplied by Novo Nordisk A/S, while the $D_2O$ and perdeuterated phenol was purchased from ISOTEC Inc. and MSD Isotopes, respectively. All other chemicals were obtained from Merck. The $R_6$ insulin hexamer samples were prepared by dissolving the lyophilized Zn insulin in $H_2O$ (with 10% $D_2O$) or in 99.96% $D_2O$. The pH values of the samples were adjusted to 8.0−8.1 (direct meter reading), using NaOH and HCl or NaOD and DCl. The concentration of 2Zn insulin in the NMR samples was 3−5 mM, while the phenol concentration was 21−35 mM, corresponding to a phenol:insulin molar ratio of 7:1. The samples used for the spectral assignment contained perdeuterated phenol, while those used for the identification of NOE correlations between phenol and insulin contained normal [$^1H$]-phenol. In cases where these identifications were hampered by the strong phenol signals, a sample with a phenol:insulin ratio of 3:1 was used. The lower phenol concentration caused no changes in the spectra as compared to those obtained at the higher phenol:insulin ratio, except for the weaker phenol signals. Dioxane (1 mM) was added as an internal reference. The samples were stored at 5 °C between the NMR experiments to prevent degradation. Also, the phenol, which is widely used as a preservative in therapeutic preparations of human insulin, stabilized the samples. Accordingly, no signs of deamidation or other types of degradation of the insulin in the samples were observed even after months.

*(b) NMR Spectroscopy.* Most of the $^1H-^1H$ and $^1H-^{13}C$ chemical shift-correlated 2D NMR spectra were recorded at a $^1H$ frequency of 500 MHz and a $^{13}C$ frequency of 125.7 MHz on a Bruker AM500 spectrometer, with sequential quadrature detection in the $t_2$ dimension (Redfield & Kunz, 1975) and TPPI in the $t_1$ dimension (Drobny et al., 1979; Bodenhausen et al., 1980; Marion & Wüthrich, 1983). The 500 MHz NOESY spectra used to identify the insulin−phenol correlations and the slowly exchanging amide protons and a 750 MHz NOESY spectrum used for a more detailed investigation of the correlation in the N-terminal end of the B-chain were recorded on Varian Unity Inova 500 and 750 spectrometers, respectively, using TPPI−States frequency discrimination (Marion et al., 1989). The temperature was 310 K in most cases, although spectra were also recorded at other temperatures in the range of 298−330 K.

Three different types of 2D experiments were applied: DQF-COSY (Rance et al., 1983), TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985), and NOESY (Jeener et al., 1979; Macura et al., 1981). The TOCSY spectra were recorded with mixing times of 30, 60, and 90 ms in $H_2O$, respectively. A DIPSI-2 (Shaka et al., 1988) spin lock pulse

of 8.3 kHz was applied in combination with two z-filters (Sørensen et al., 1984; Rance, 1987). The NOESY spectra were recorded with mixing times of 40, 150, or 180 ms. In order to obtain correlations in the fingerprint region that involve α-protons normally saturated by the water irradiation, NOESY spectra in $H_2O$ were also acquired using a jump-return read pulse (Driscoll et al., 1989) on the Bruker AM500 spectrometer and the WATERGATE technique (Piotto et al., 1992) on the Varian Unity Inova 500 spectrometer. In these spectra, the residual water resonance was removed by deconvolution before the Fourier transformation (Marion & Bax, 1989), and a linear baseline correction was applied in both frequency dimensions.

Qualitative discrimination between slowly and rapidly exchanging amide protons was achieved by recording a WATERGATE NOESY spectrum during the first 22 h after dissolving the sample in 99.96% $D_2O$ and defining the amide protons giving rise to cross-peaks in this spectrum as slowly exchanging. The sample used in this experiment was obtained by dissolving the $R_6$ insulin from a freeze-dried $H_2O$ sample in $D_2O$ containing the appropriate surplus of perdeuterated phenol.

Two different types of $^1H-^{13}C$ chemical shift-correlated experiments were performed: (1) an HSQC experiment (Bodenhausen & Ruben, 1980) correlating the $^{13}C$ nuclei with the directly bound protons and (2) an HSQC-TOCSY experiment (Norwood et al., 1990; Kessler et al., 1990) correlating $^{13}C$ nuclei with protons further away in the spin systems. The $^1H-^{13}C$-correlated spectra were recorded with $^{13}C$ in natural abundance. The mixing time in the HSQC-TOCSY experiment was 40 ms. The remaining experimental details were as described previously (Kristensen & Led, 1995).

Finally, a few cases of overlap in the 2D spectra that prevented a complete sequential assignment were resolved in a $^1H$ 3D TOCSY-NOESY spectrum recorded as described previously (Badsberg et al., 1996). The mixing times used were 60 ms for the TOCSY part and 180 ms for the NOESY part of the experiment.

The data processing and the digital resolution and window functions applied were as described previously (Kristensen & Led, 1995; Badsberg et al., 1996). The first seven points in each $t_2$ slice were routinely reconstructed by backward linear prediction (Marion & Bax, 1989; Led & Gesmar, 1991). In the $^1H-^{13}C$-correlated spectra, the 512 data points in the $t_1$ dimension were extended to 1024, using the forward linear prediction estimation (Tirendi & Martin, 1989; Led & Gesmar, 1991). The digital resolution of the $^1H-^1H$-correlated 2D spectra was 2.4 Hz/point in the $F_2$ dimension and 4.9 Hz/point in the $F_1$ dimension after zero filling. The digital resolution of the 2D $^1H-^{13}C$-correlated spectra was 1.2 Hz/point in the $F_2$ dimension and 9.8 Hz/point in the $F_1$ dimension. Finally, in the 3D spectrum, the digital resolution was 7.2 Hz/point in the $F_3$ dimension and 14.6 Hz/point in both the $F_2$ and $F_1$ dimensions after zero filling to twice the number of data points in these two dimensions, while the applied window function was a double parabola in all three dimensions. This window function, which has a shape that resembles that of a shifted sine bell, consists of two half-parabola with a common maximum. The first half-parabola defines the window function from the starting point to the maximum, while the second defines the window function from the maximum to the end point.

*(c) Structure Calculations.* The calculation of the hexamer structures was carried out as a stepwise process. First, the structure of the monomers and the assignment of the intramonomeric NOEs were obtained by an iterative procedure to distinguish the intramonomeric NOEs from the intermonomeric NOEs and to overcome complications caused by signal overlap in the NMR spectra. Subsequently, the organization of the six monomers and the phenol molecules in the hexamers were determined using the solution structure of an insulin dimer (Jørgensen et al., 1992) and the crystal structure of the $R_6$ insulin hexamer (Derewenda et al., 1989) to assign the individual NOEs identified as intermonomeric in the iterative process just mentioned. The structure calculations were performed with the program X-PLOR (Brünger, 1992) using distance geometry, simulated annealing, and restrained energy minimization (Nilges et al., 1988). All parameter and topology files used are those given in the X-PLOR program.

The starting point for the structure calculations was a generation of 60 monomer substructures, using the DG-SUB-EMBED protocol. The substructures included the backbone nitrogen, carbon, and hydrogen atoms and the $\beta$- and $\gamma$-carbons. Subsequently, simulated annealing calculations were carried out using the DGSA protocol. The structures were first run through 200 cycles of restrained energy minimization. This was followed by 4 ps of restrained molecular Verlet dynamics at a temperature of 3000 K and 8 ps of cooling to 300 K. The time steps were 0.001 ps in both cases, while the temperature was varied in steps of 50 K during the cooling stage. The van der Waals energy function was represented by a simple repel function. During the cooling stage, the van der Waals interactions were increased by varying the force constant of the van der Waals repel function from 0.003 to 4 kcal $mol^{-1}$ $Å^{-4}$.

In the calculations of the hexamer structures, distance geometry calculations were not performed because of the very large system involved (4802 atoms). Instead, the initial coordinates of the hexamers were taken as those of six different monomer structures chosen randomly among the 36 best monomer DGSA structures. The initial coordinates of the Zn atoms and those of the two Zn-bound water molecules and the phenol molecules (*vide infra*) were taken from a template structure (Brünger, 1992). Subsequently, the relative positions within the hexameric organization of the insulin monomers, the Zn atoms, and the water and phenol molecules were determined in a refinement calculation that included restrained energy minimization and simulated annealing.

Thus, the hexamer structures were first run through 500 cycles of restrained energy minimization and 4000 cycles of restrained molecular Verlet dynamics at a temperature of 6000 K. Subsequently, the structures were slowly (8000 steps) cooled to 100 K, and the 20 structures with the lowest total energy were run through 8000 cycles of restrained molecular dynamics while the temperature was reduced from 6000 to 100 K. In these calculations, the time step was 0.003 ps while the temperature was varied in steps of 50 K. Finally, the structures were run through 200 steps of restrained energy minimization calculations, and the 10 structures with the lowest total energy were selected. Throughout the hexamer calculations, the van der Waals interaction was represented by a simple repel function. In the initial restrained energy minimization, the force constant

of this term was 0.01 kcal mol$^{-1}$ Å$^{-4}$ to allow close contact between the atoms, while in the rest of the DGSA calculations, it was varied as in the monomer calculations. Again during the cooling stage of the refinement, the repel force constant was varied from 0.003 to 4 kcal mol$^{-1}$ Å$^{-4}$.

Two sets of 30 hexamer structures were calculated, the sets differing only in the NOEs that define the orientation of the aromatic ring of the Phe(B25) residue in one of the monomers (*vide infra*). The amount of time used for the simulated annealing calculations and the structure refinement calculations was 3.5 and 0.75 h per structure, respectively, on a DEC3000 AXP 400 workstation.

A number of different calculation strategies were investigated before the final calculation described above. These include a calculation where a number of dimers were constructed on the basis of the monomer structures. Subsequently, these dimers were used as starting coordinates for the hexamer. Both dimer and hexamer structure calculations were carried out as described above, except that the starting temperature was only 5000 K in both calculations. This strategy resulted in hexamer structures very similar to those obtained in the calculations that started directly from the monomers, the rms deviation between the average structures in the two cases being 0.60 Å if the less well-defined terminals were excluded. Also, the respective rms deviations from the average structures were almost identical in the two cases. In the calculation of the dimer, the coordinates of two identical monomers were used as the initial coordinates of the dimers. In the hexamer calculations, however, starting with monomers or dimers with different sets of coordinates proved to be essential. The most important parameter for the structure calculations was found to be the starting temperature. Thus, in the calculations, where the hexamer structures were calculated directly from six monomer structures, an increase of this temperature from 5000 to 6000 K improved the convergence ratio from 1 out of 6 structures to 4 out of 6 structures. Also, the force constant of the NCS term (*vide infra*) has some influence on the calculation of the hexamer from the dimer structures. Thus, a 10-fold increase of the NCS force constant in these calculations improved the convergence ratio. A similar improvement was not observed in the calculation of the hexamers based on the monomer structures.

Finally, it should be mentioned that, although the determination here of the hexameric organization of the six monomers relies on the solution structure of a dimer (Jørgensen et al., 1992) and the crystal structure of the $R_6$ hexamer (Derewenda et al., 1989), a determination of the solution structures of the $R_6$ hexamers from the NMR data alone might well be possible, in the same way as recently demonstrated for the P53 tetramer (Clore et al., 1994; Lee et al., 1995). Moreover, the difficulties of organizing the two dimers encountered in that case (Clore et al., 1995) may not apply here, since the number of intermonomeric NOEs in the $R_6$ insulin hexamers is considerably larger than the number in the P53 tetramer because of tighter packing of the monomers in the former than in the latter multimer. In particular, by employing the dynamic assignment method recently suggested by O'Donoghue et al. (1996), a determination of the structures of the $R_6$ hexamers from NMR data alone might be possible even without the use of isotope labeling.

*(d) Restraints.* Interproton distance restraints were obtained from the NOESY experiments detailed above. The NOEs were identified as intramonomeric NOEs on the basis of their consistency with the calculated monomer structures (*vide infra*). Otherwise, the NOEs were classified as dimer- or trimer-specific NOEs using the crystal structure of the $R_6$ insulin hexamer as the initial model (Derewenda et al., 1989) and subsequently justified in the structure calculations. In a few cases where the observed NOEs were not readily classified as exclusively intramonomeric, but could also result from contacts across the dimer and trimer interfaces according to the calculated structures, they were included only as intramonomeric in the calculation. An example is provided by the NOEs between Val(B12) $\gamma$-$CH_3$ and Phe(B24) $\zeta$-CH, which are found both within the monomer and across the dimer interface.

The phenol−insulin correlations were easily identified by comparison between corresponding spectra of the samples containing perdeuterated phenol and normal [$^1$H]-phenol, respectively.

The NMR spectra contain only a single set of resonances for each of the residues (*vide infra*), except for Phe(B25) and Phe(B24), and only one set of resonances for the phenol molecules. It was, therefore, assumed that the six insulin monomers have identical structures and are organized in the same way within the hexamer together with their associated phenol molecule. Consequently, the same intra- and intermolecular NOEs, as well as NOEs between insulin and phenol molecules, were included in the calculation for each of the six monomers.

The intensities of all NOE cross-peaks were determined quantitatively by a combination of linear prediction analysis and simple least-squares estimation (Gesmar et al., 1994; Kristensen et al., 1996). On the basis of these intensities, the NOEs were classified as "strong", "medium", and "weak" as described previously (Badsberg et al., 1996). The corresponding upper bound distance restraints were 2.85, 3.56, and 5.00 Å, respectively. Correction factors for pseudoatoms were added to these limits as described by Wüthrich et al. (1983). In cases where the NOE intensities were too small to be determined quantitatively by the least-squares procedure, they were classified as weak.

$^3J_{H_\alpha H_N}$ coupling constants could not be obtained from the experimental spectra using standard methods due to overlap caused by the large line widths. Therefore, no backbone torsion angle constraints were included in the calculation.

Two distance restraints were included for each of the intermonomeric hydrogen bonds involving Phe(B24) and Tyr(B26). These distance restraints were from 1.8 to 2.0 Å for the O−H distance and from 2.7 to 3.0 Å for the O−N distance. No other hydrogen bond restraints were included at any stage of the calculations.

Finally, in the distance geometry calculations, one distance restraint was included for each of the three disulfide bonds. Upper and lower limits for the S−S distance across the disulfide bonds were 2.12 and 1.92 Å, respectively. In the simulated annealing and the final energy minimization calculations (*vide infra*), the disulfide bonds were treated as real bonds.

The *topallhdg* and *parallhdg* parameter files were extended to account for the phenol molecules and the coordination to the Zn atoms. Thus, the phenol molecule was included in the topology file as a Tyr residue in which the backbone

and $\beta$-atoms were deleted, while a proton was added in position 4. The charges on the carbon and proton in position 4 were set equal to the charges of the carbons and protons in positions 2, 3, 5, and 6. The binding site of the Zn atoms cannot be identified from the NMR data. It was, therefore, assumed that the two $Zn^{2+}$ ions are tetrahedrally coordinated to the three His(B10) side chains as in the crystal structure of the $R_6$ hexamer (Smith & Dodson, 1992). The Zn atom was introduced in the calculation as described for the Cu coordination in plastocyanin (Badsberg et al., 1996), except that distance restraints between the His(B10) N$\epsilon$2 atoms and the Zn atoms were included instead of real bonds. The distance restraints between the two atoms were kept within the limits of 1.8−2.4 Å, according to the values of the corresponding distance in the crystal structure of the $R_6$ hexamer. The N$\epsilon$2−Zn−N$\epsilon$2 angles were constrained to 110° and the Zn−N$\epsilon$2−C angles to 125°, as found in the crystal structure (Derewenda et al., 1989). For both angles, the force constant was only 10 kcal mol$^{-1}$ rad$^{-2}$ to allow a relatively large variation from the equilibrium value. The fourth group that coordinates to the Zn atom is a water molecule according to the crystal structure. The topology and the parameters describing the water molecules were taken from the files *topoplsx.pro* and *paroplsx.pro*. The mass of the Zn atom was set to 65.37 as in the file *toph19.ion*, while the parameters describing the nonbonding interactions for this atom were taken from *parameters.elements*. To account for the Gly(A1) residue, the file *parallhdg.pro* was extended with one improper angle describing N-terminal glycines.

Throughout the hexamer calculations, the X-PLOR NCS term was included in the empirical energy function to maintain the same structure of the six units within the hexamer, a unit being an insulin molecule together with the nearest phenol molecule. All atoms, except those of Phe(B25) in the nonsymmetric hexamer, were constrained by this term, the force constant being 0.1 kcal mol$^{-1}$ Å$^{-2}$ in the initial restrained energy minimization and 0.001 kcal mol$^{-1}$ Å$^{-2}$ in the high-temperature calculations. During the cooling stage, the NCS force constant was varied from 0.001 to 1.0 kcal mol$^{-1}$ Å$^{-2}$, while in the final refinement, it was 1.0 kcal mol$^{-1}$ Å$^{-2}$.

## RESULTS

*(a) Sequential Assignments.* Essentially complete assignments of the $^1$H and $^{13}$C resonances were achieved by well-established procedures (Wüthrich, 1986) on the basis of 2D $^1$H DQF-COSY, TOCSY, NOESY, and $^1$H−$^{13}$C-correlated HSQC and HSQC-TOCSY spectra and a $^1$H 3D TOCSY-NOESY spectrum. Only one set of resonances was observed for each residue, except for Phe(B25) and Phe(B24) (*vide infra*). Also, the phenol molecules were characterized by only one set of resonances. The quality of the spectra used in the assignment is illustrated in Figures 1 and 2. All sequential NOEs are indicated in Figure 3. In general, the $^1$H assignment is in agreement with the corresponding assignment at 328 K, which was published recently by Jacoby et al. (1996). Only for the amide protons are deviations found systematically, most likely because of the difference in temperature.

As in previous NMR studies of insulin (Kristensen & Led, 1995; Jørgensen et al., 1996), the $^1$H−$^{13}$C-correlated HSQC and HSQC-TOCSY spectra (Figure 2) proved to be valuable

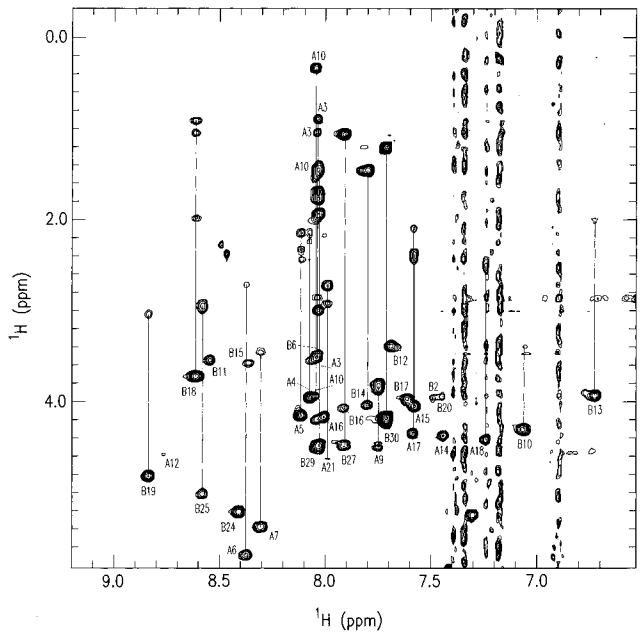


FIGURE 1: Amide proton region of the $^1$H TOCSY spectrum of the phenol-stabilized $R_6$ hexamers of human insulin in $H_2O$ (5 mM, pH 8.0, 310 K) showing the correlations between the amide and aliphatic protons.

in the assignments because of the large dispersion in the $^{13}$C dimension, especially in the aromatic region where the assignment of the $^1$H resonances in proteins is often hampered by the limited dispersion. In particular, the assignment of these protons is important for the determination of the 3D solution structure of the phenol-stabilized $R_6$ insulin hexamer, because of their involvement in the hydrophobic interactions that characterize the dimer formation, and because of the multiple orientations of the phenyl ring of Phe(B25) described below.

The large upfield shifts observed for the protons in the imidazole groups of His(B5) and His(B10) are of special interest. These shifts contrast with those found for most other insulin mutants (Hua & Weiss, 1991; Kristensen et al., 1991; Kristensen & Led, 1995; Jørgensen et al., 1996) but are in agreement with the assignments of the corresponding imidazole protons in the resorcinol-induced $R_6$ hexamer (Brzović et al., 1994) and, most likely, reflect the Zn coordination in the case of His(B10), and an influence from the aromatic alcohol bound in the $R_6$ hexamer in the case of His(B5) (*vide infra*). The absence in the HSQC spectra of the $^1$H−$^{13}$C correlations from the imidazole rings of His(B5) and His(B10) is another indication of the special environments experienced by the two imidazole rings in the phenol-stabilized hexamer. Thus, for other insulin mutants with no phenol molecules and Zn atoms bound, strong and easily identifiable $^1$H−$^{13}$C correlations from the imidazole rings of His(B5) and His(B10) were observed (Kristensen & Led, 1995; Jørgensen et al., 1996).

*(b) Evidence for the Presence of Two Monomer Conformations.* As it appears from Figure 2, the $^1$H−$^{13}$C-correlated $\epsilon$- and $\zeta$-cross-peaks of Phe(B25) are doubled at 310 K. This immediately shows that there is a structural heterogeneity in the vicinity of the aromatic ring of Phe(B25). In the $^1$H−$^1$H-correlated spectra, this is confirmed by the presence of two independent sets of $\epsilon$- and $\zeta$-signals for Phe(B25) belonging to two independent patterns of NOE correlations. Moreover, only one of these patterns of correlations is

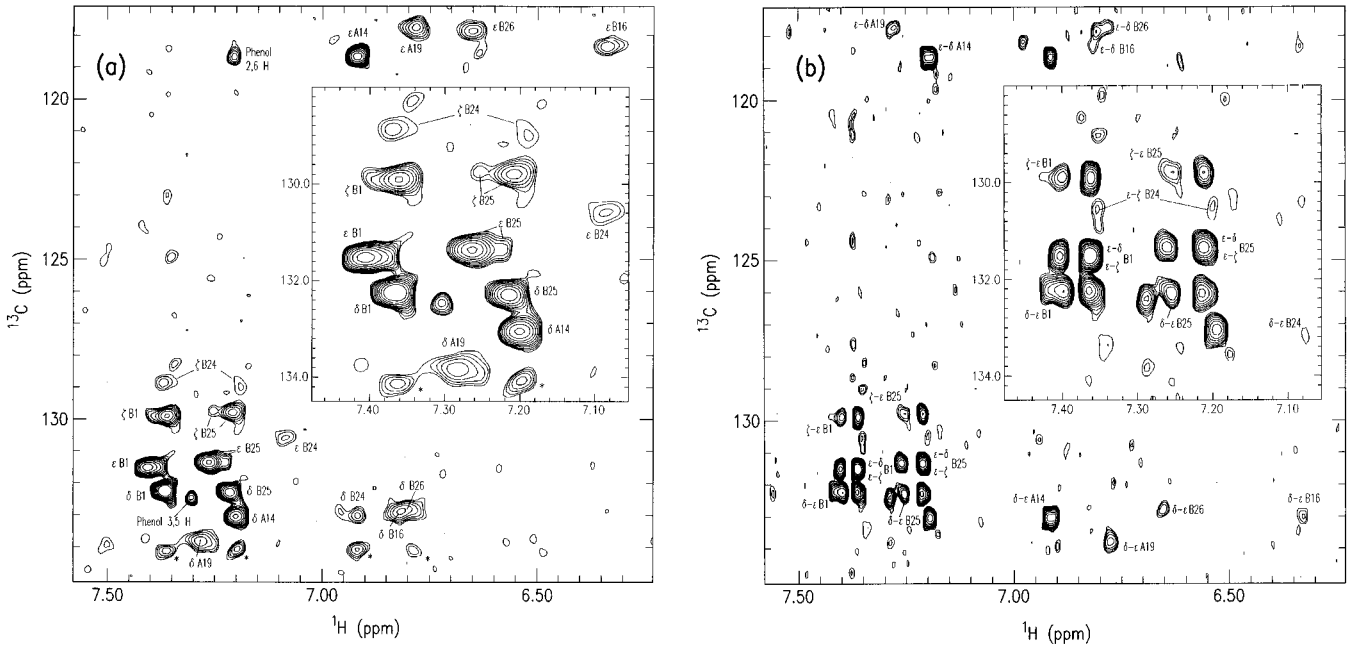


FIGURE 2: Aromatic region of $^{1}H-^{13}C$ chemical shift-correlated 2D NMR spectra of the phenol-stabilized $R_6$ hexamers of human insulin in $D_2O$ (5 mM, pH 8.0, 310 K) showing the correlations in the phenylalanine rings: (a) an HSQC spectrum showing the one-bond correlations and (b) an HSQC-TOCSY spectra showing the two-bond correlations. For the Phe(B25) spin system, two-bond correlations were observed only for the major conformation (molecule 1). The asterisks indicate artifacts arising from the carrier.

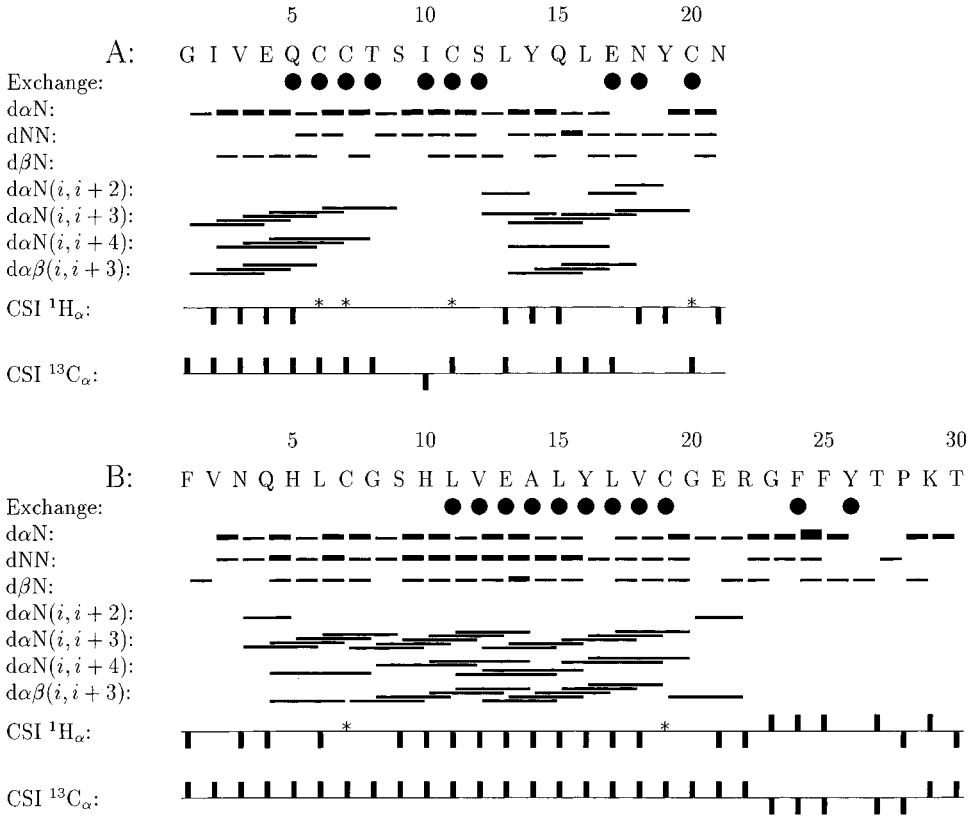


FIGURE 3: Summary of sequential and medium-range NOE connectivities of the phenol-stabilized $R_6$ hexamers of human insulin used in the sequential assignment and to characterize the secondary structure. The observed connectivities are indicated by bars connecting the two involved residues. For Pro(B28), the connectivities to the $\delta$-protons are shown. The intensities of the sequential NOEs are indicated by the thickness of the bars. Backbone amide protons that are exchanging slowly with deuterons in $D_2O$ are indicated by filled circles (●). The CSI of $^{1}H_{\alpha}$ and $^{13}C_{\alpha}$ were calculated as described previously (Jørgensen et al., 1996). No $^{1}H_{\alpha}$ CSIs are given for the six disulfide-bonded cysteine residues, as indicated by the asterisks (*).

exclusively intramonomeric, while the other pattern also includes intermonomeric NOEs. Thus, the first pattern includes connectivities between the following protons: Phe(B25) $\epsilon$-CH and Tyr(A19) $\alpha$-CH, Phe(B25) $\epsilon$-CH and Tyr(A19) $\beta$-$CH_2$, Phe(B25) $\delta$-CH and Thr(B27) $\gamma$-$CH_3$, Phe(B25) $\delta$-CH and Thr(B27) $\beta$-CH, and Phe(B25) $\epsilon$-CH and Thr(B27) $\beta$-CH, all within the same monomer. In contrast, the other pattern includes connectivities between Phe(B25)

ζ-CH and Thr(B27) γ-$CH_3$ and between Phe(B25) ζ-CH and both of the Phe(B25) β-$CH_2$ protons, all across the dimer interface. The distinction between the connectivities belonging to the two individual patterns could in most cases be made by visual inspection of the structures of the monomers. Subsequently, they were all confirmed, unambiguously, by the structure calculations (*vide infra*). Furthermore, the Phe(B25) ζ-CH−Thr(B27) γ-$CH_3$ and Phe(B25) ζ-CH−Phe(B25) β-$CH_2$ correlations gave rise to severe violations when included as intramonomeric NOEs.

The presence of two independent sets of connectivities reveals the fact that the aromatic ring of Phe(B25) exists in two conformations: one in which the aromatic ring of Phe(B25) points toward the other residues within the same monomer and one in which it is turned away from these residues and points outward from the molecule. These orientations closely resemble the orientations of the Phe(B25) phenyl ring in molecule 1 and molecule 2, respectively, of the crystal structure of native insulin (Blundell et al., 1972; Peking Insulin Structure Research Group, 1971). Here, the same designations are adopted for the two conformations of the monomers within the $R_6$ hexamer, even though the other minor structural variations between molecule 1 and molecule 2, which are found in the $T_6$ crystal structure (Chothia et al., 1983), could not be observed in the $R_6$ solution hexamer within the experimental uncertainties.

Also for the ζ-proton of Phe(B24), two signals are observed in the HSQC spectrum. While these signals are well-separated, both of them are correlated to the same, and only, ϵ-resonance in the HSQC-TOCSY spectrum. This difference in multiplicity of the ζ- and ϵ-resonances of Phe(B24) is in agreement with the obtained solution structures of the hexamers (*vide infra*), which indicate that the two conformations of Phe(B25) give rise to a considerably larger difference in the environment for the ζ-proton of Phe(B24) than for the ϵ-proton of Phe(B24). It should be noted that the double signals observed here for the Phe(B24) and Phe(B25) residues still remain at 330 K. This indicates that both conformations exist at this elevated temperature, in contrast with the observations reported by Jacoby et al. (1996).

Another interesting observation is the difference between the intensities of the $^1H-^{13}C$ signals (Figure 2a) corresponding to the molecule 1 and molecule 2 conformation, respectively, which shows that molecule 1 is the dominant conformation, as in the $R_6$ crystal structure (Smith & Dodson, 1992). Thus, at 310 K, the population of the molecule 2 conformation (minor conformation) was found to be 10.1 ± 2.7 and 11.9 ± 3.2%, respectively, of that of the molecule 1 conformation (major conformation), according to the intensities of the two ζ-cross-peaks of Phe(B25) and the two ζ-cross-peaks of Phe(B24) in the $^1H-^{13}C$-correlated HSQC spectra (Figure 2a). The intensities were determined as described for the NOE cross-peaks (*vide supra*). A similar determination of the relative populations of molecule 1 and molecule 2 on the basis of the intensities of the two ϵ-cross-peaks of Phe(B25) in Figure 2a was too uncertain because of severe overlap of the two signals. The different populations of the two conformations also explain why relayed signals correlating the ϵ-protons with the δ- and ζ-protons of Phe(B25) are observed only for molecule 1 (Figure 2b).

It is also interesting that both the ϵ- and the ζ-signals of Phe(B1) are relatively broad, even though most of the signals from the carbon-bound protons in the $R_6$ hexamer have similar line widths as was also observed by Jacoby et al. (1996). Furthermore, the Phe(B1) ϵ-cross-peak in Figure 2a splits into two signals with different $^1H$ frequencies, while the line widths decrease with increasing temperature. Both of these observations indicate a structural inhomogeneity also in this part of the molecule, although no indication of independent spin systems corresponding to the two Phe(B1) ϵ-resonances was found in the $^1H$ NMR spectra.

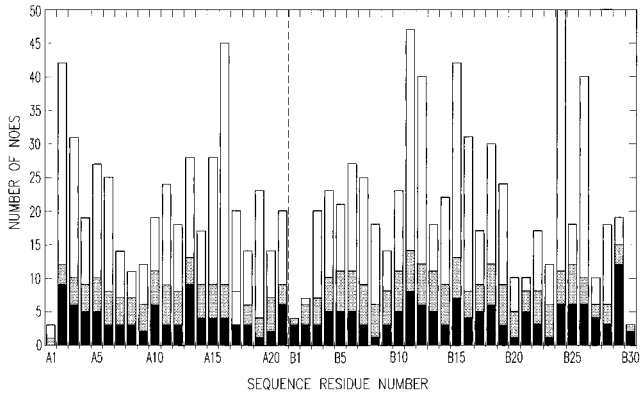


FIGURE 4: Distribution of the NOE connectivities used in the structure calculation of the monomer within the phenol-stabilized $R_6$ hexamers of human insulin. The filled, hatched, and open bars represent intraresidual, sequential, and inter-residual NOEs, respectively. All sequential and inter-residual NOEs are represented twice, once for each of the two involved residues.

*(c) Quality of the Calculated Structures.* The structures of the initial monomers were obtained from 665 intramonomeric NOE distance restraints and the three constraints imposed by the disulfide bridges. The NOE restraints are as follows: 216 intraresidual restraints, 112 sequential restraints, and 337 inter-residual restraints. The distribution of NOEs is shown in Figure 4. Among the intramonomeric NOEs, a total of 480 were found to be nonredundant by the structure calculation program DIANA (Güntert et al., 1991).

Two hexamers were calculated, *i.e.* the structure of a symmetric hexamer consisting of six monomers with the molecule 1 conformation and a nonsymmetric hexamer consisting of five molecule 1 monomers and one molecule 2 monomer. A total of 4394 and 4391 distance restraints were included in these calculations, respectively. Primarily, these restraints include the intramonomeric restraints for each of the six monomers and the restraints that define the two different orientations of the Phe(B25) ring, *i.e.* five intramonomeric restraints for each of the molecule 1 monomers and two intermonomeric restraints for the molecule 2 monomer. Further, the organization of the six insulin monomers and the six phenol molecules in the $R_6$ hexamer was defined by 23 intermonomeric NOEs across the dimer interface and nine NOEs across the trimer interface for each monomer and by 25 NOEs between each phenol molecule and three different monomers in the hexamer. These intermolecular NOEs are summarized in Figure 5, together with the NOEs that define the two different orientations of the Phe(B25) ring. Finally, the intermonomeric hydrogen bonds involving Phe(B24) and Tyr(B26) were included in the calculation of the $R_6$ hexamer, while the Zn coordination was defined by eight distance restraints (see Experimental Procedures).

The convergence of the structure calculations appears in Table 1 and Figures 6 and 7. Thus, the rms deviation of the 10 structures from the mean structure is given in Table 1

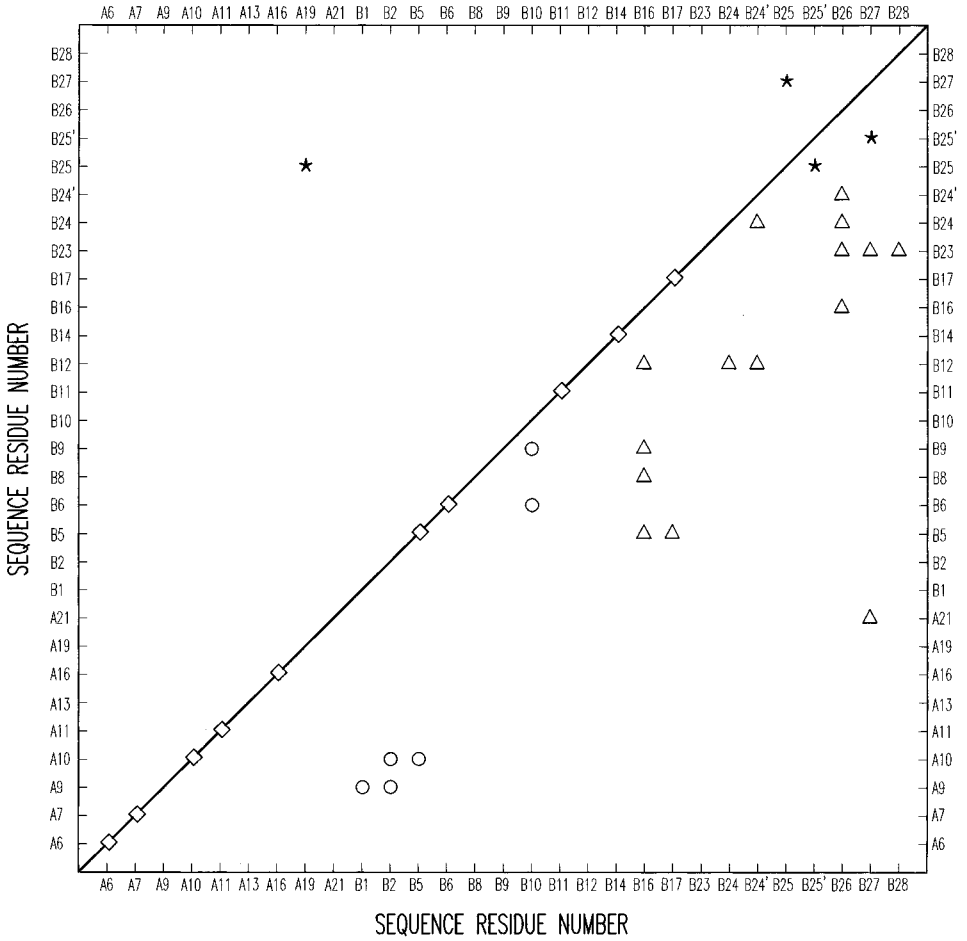


FIGURE 5: Intermolecular NOE contacts and NOE contacts associated with the two different orientations of the phenyl ring of Phe(B25) in the phenol-stabilized $R_6$ hexamers of human insulin. Several of the contacts correspond to more than one NOE connectivity between the indicated residues (see the text and Supporting Information). Below the diagonal: (△) intermonomeric contacts across the dimer interface, (○) intermonomeric contacts across the trimer interface, and (★) intermonomeric contacts observed only for Phe(B25) in molecule 2 (spin system B25′). Above the diagonal: (★) intramonomeric contacts observed only for Phe(B25) in molecule 1 (spin system B25). On the diagonal: (◇) contacts between insulin and phenol.

Table 1: Average rms Deviations[a] (Angstroms) between the Final Structures and the Mean Structures of the $R_6$ Hexamer of Insulin in Solution

| molecule | all atoms[b] | backbone[b] | all helix atoms[c] | helix backbone[c] |
|---|---|---|---|---|
| hexamer (nonsymmetric) | 1.17 | 0.81 | 1.13 | 0.77 |
| hexamer (symmetric) | 1.19 | 0.81 | 1.16 | 0.78 |
| molecule 1[d] | 1.25 | 0.58 | 1.28 | 0.49 |
| molecule 2 | 1.22 | 0.57 | 1.24 | 0.49 |

[a] Fitting was done on the complete hexamer molecules and on the individual monomers within the hexamers, respectively. [b] Fitting was done on all heavy atoms, including the Zn atoms, the phenol and water molecule atoms, and on all N, $C_\alpha$, and C′ backbone atoms, respectively. Residues A1, A21, B1, B2, and B30 were excluded from the rms deviation calculations. [c] Fitting was done on all heavy atoms and on all N, $C_\alpha$, and C′ backbone atoms from the helix region. The helix region included residues A2−A8, A13−A20, and B3−B19. [d] Within the precission of the numbers given here, the same value was found for all the molecule 1 monomers.

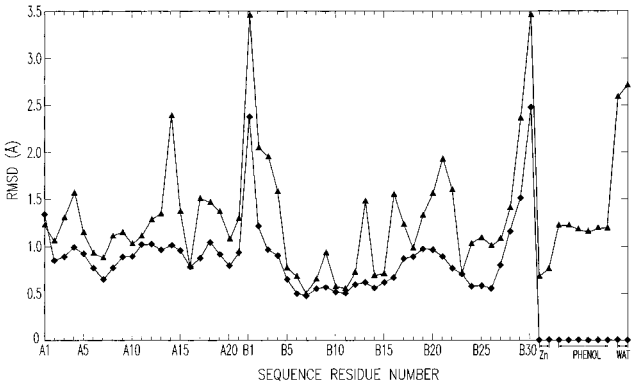


FIGURE 6: Average residual rms deviations (Angstroms) from the mean of the 10 nonsymmetric hexamer structures. The values shown are averages from the six monomers within the hexamer. Also shown are the rms deviations for the two Zn atoms, the six phenol molecules and the two water molecules: (▲) values for the side chain atoms and (◆) values for the N, C′, and $C_\alpha$ backbone atoms.

for both of the two hexamers and for the two monomers (molecule 1 and molecule 2). The average rms deviations from the mean structure are shown in Figure 6 for each of the individual residues in the monomer, as well as for the Zn atoms, the phenol molecules, and the water molecules. For most of the residues, there is a clear correlation between the number of medium-range and long-range NOEs and their rms deviations. A best fit superposition of the backbone of the 20 final structures of the symmetric as well as the nonsymmetric hexamer is shown in Figure 7.

The good agreement between the calculated structures and the experimental data is illustrated by the statistics in Table 2. No NOE violations larger than 0.15 Å were observed in any of the 20 final hexamer structures. The solution structures of the two monomers within the nonsymmetric

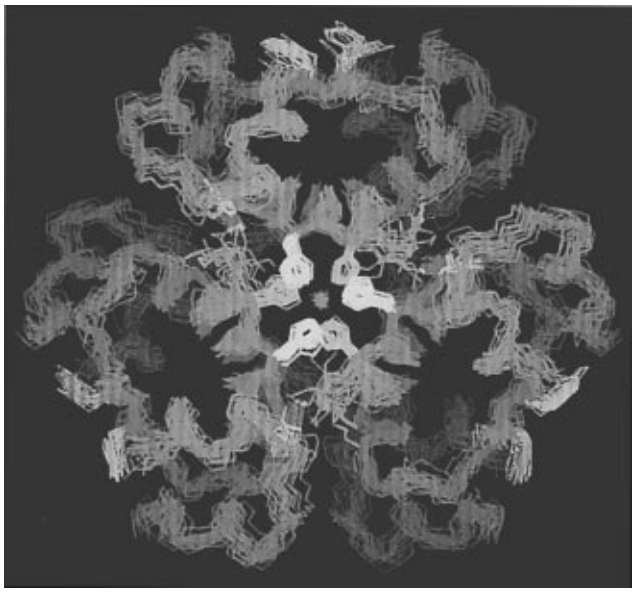

FIGURE 7: Best fit superposition of the 20 selected solution structures of the $R_6$ hexamers of human insulin, *i.e.* 10 structures of the nonsymmetric hexamer and 10 structures of the symmetric hexamer. Only the C′, $C_\alpha$, and N backbone atoms are superimposed, and two amino acid residues in each chain terminal were excluded from the superposition. The backbone atoms are shown together with the aromatic rings of Phe(B25) (orange) and their different orientations, the imidazole rings of His(B10) (yellow), the Zn atoms (blue), and the phenol molecules (purple). The C terminus of the B-chain [Lys(B29) and Thr(B30)] and the two water molecules were excluded from the figure.

hexamer are shown in panels a and b of Figure 8 together with the closest phenol molecule, while the structure of the nonsymmetric hexamer is sketched in Figure 9.

Finally, it should be noted that calculations of the hexamer structures without the Zn atoms and water molecules resulted in overall structures similar to those obtained with these atoms and molecules included, the rms deviation between the average structures in the two cases being only 0.67 Å. This shows that there is no conflict between the solution structure of the peptide chain defined by the NMR data and the Zn coordination found in the crystal structure and indicates that the coordination of the Zn atoms is similar in the two phases.

## DISCUSSION

### *Slowly Exchanging Amide Protons and Hydrogen Bonds*

Although the exchange rates of the amide protons are expected to be relatively fast under the applied experimental conditions (pH 8.0−8.1, 310 K), 21 slowly exchanging amide protons were observed (Figure 3) in the NOESY spectrum recorded during the first 22 h after dissolution in $D_2O$. Among these amide protons, 16 participate in hydrogen bonds in the helices. These are the amide protons of Gln(A5), Cys(A6), Cys(A7), Thr(A8), Glu(A17), Asn(A18), and Cys(A20) in the A-chain and the amide protons of the residues from Leu(B11) through Cys(B19) in the B-chain. Furthermore, the slowly exchanging amide protons of Cys(A11), Phe(B24), and Tyr(B26) are involved in tertiary hydrogen bonds according to the calculated structures. Thus, the amide proton of Cys(A11) is hydrogen-bound to the hydroxyl oxygen of the phenol molecule, while the amide protons of Phe(B24) and Tyr(B26) are both involved in the $\beta$-sheet hydrogen bonding across the dimer interface. The remaining two residues, Ile(A10) and Ser(A12), are buried in the hexamer, which could explain the slow exchange of their amide protons.

### *General Description of the Solution Structure*

*(a) Monomer*. Apart from the different orientations of the phenyl ring of Phe(B25), the structures of two different monomers (molecule 1 and molecule 2) in the $R_6$ hexamer are very similar. Thus, the rms deviation between the N, C′, and $C_\alpha$ backbone atoms of the average molecule 1 and the molecule 2 within the average nonsymmetric hexamer is only 0.04 Å.

The most conspicuous structural feature of the monomer in the R state as compared with monomers in the T state is the extended α-helix of the B-chain (Figure 8), which stretches from Asn(B3) to Cys(B19) in the monomer of the $R_6$ solution hexamer, while it only includes the residues from Gly(B8) to Cys(B19) in the corresponding T state monomers (Jørgensen et al., 1992, 1996). Otherwise, the monomers in the $R_6$ hexamer are structurally similar to those found in other insulin monomers. Thus, the A-chain contains two antiparallel helices, an N-terminal $A_I$-helix from Ile(A2) to Thr(A8) and a C-terminal $A_{II}$-helix from Leu(A13) to Cys(A20), connected by a loop from Ser(A9) to Ser(A12). The $A_I$-helix is an α-helix, whereas the $A_{II}$-helix has both α- and $3_{10}$-helix character. The α-helix of the B-chain is succeeded by a $\beta$-turn from Gly(B20) to Gly(B23). The N and C termini of the B-chain, Phe(B1)-Val(B2) and Thr(B30), respectively, are both disordered. The C-terminal region of the B-chain from Phe(B24) to Pro(B28) forms an extended structure and is involved in the $\beta$-sheet across the dimer interface. The *trans* conformation of the peptide bond to Pro(B28) is stabilized by the packing in the molecule. No indication of a *cis* conformation of this peptide bond was found. The secondary structures obtained in the calculations are supported by the sequential and short-range NOEs summarized in Figure 3. Finally, the patterns of the $^1H_\alpha$ and $^{13}C_\alpha$ CSI (Wishart & Sykes, 1994), also shown in Figure 3, are by and large in agreement with the calculated structures.

*(b) $R_6$ Hexamer*. The calculated hexamer structures show that the helix and the C-terminal end of the B-chain of one monomer pack closely against the same regions of the other monomer in the dimer interface. Primarily, the two monomers are held together by favorable hydrophobic interactions and a small $\beta$-sheet with hydrogen bonds between Phe(B24) and Tyr(B26), as in the Asp(B9) insulin dimer (Jørgensen et al., 1992). However, the packing along the dimer interface is considerably tighter here than in the free Asp(B9) dimer, as shown by the larger number of dimer contacts. Thus, the 12 different residues, which are involved in the intermonomeric contacts across the dimer interface (Figure 5) in the $R_6$ hexamer, have a total of 23 NOE contacts (*vide supra*), whereas the corresponding numbers in the Asp(B9) dimer are only six contacts and eight residues. Among the dimer contacts in the $R_6$ hexamer, the NOEs between His(B5) and Tyr(B16) are unique for the *R* configuration according to the crystal structure, as previously pointed out by Brzović et al. (1994).

The NMR data and the structure calculations show that the hexamers contain one phenol molecule per monomer.

Table 2: Structural Statistics

| | hexamer (nonsymmetric) | hexamer (symmetric) |
|---|---|---|
| rms Deviations from NOE Restraints and from the Idealized Geometry Used within X-PLOR | | |
| NOE (Å) | 0.0101 (0.0012) | 0.0103 (0.0012) |
| bond lengths (Å) | 0.0149 (0.0004) | 0.0149 (0.0004) |
| bond angles (deg) | 0.449 (0.05) | 0.441 (0.04) |
| improper dihedral angles (deg) | 0.218 (0.004) | 0.218 (0.003) |
| X-PLOR Potential Energies (kcal mol$^{-1}$)[a] | | |
| total | 277.334 (11.4) | 277.021 (11.3) |
| repel | 7.313 (2.4) | 7.358 (1.5) |
| Lennard-Jones[b] | −926.052 (55.1) | −953.981 (53.0) |
| NOE restraints | 22.561 (5.4) | 22.393 (5.1) |
| NCS | 45.567 (10.0) | 45.577 (8.7) |

[a] The force constant of the NOE and van der Waals repulsion terms were 50 kcal mol$^{-1}$ Å$^{-2}$ and 4 kcal mol$^{-1}$ Å$^{-4}$, respectively. [b] CHARMM potential (Brooks et al., 1983) used for van der Waals energy calculation.

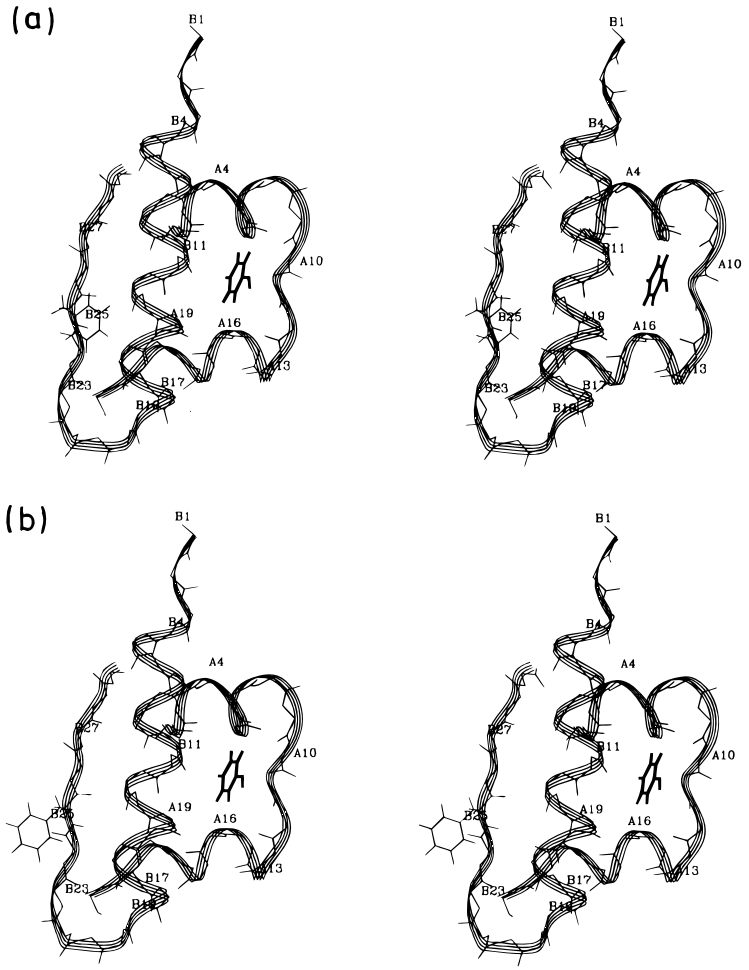


FIGURE 8: Solution structures of the monomers in the nonsymmetric $R_6$ hexamer of human insulin and the associated phenol molecule: (a) molecule 1 and (b) molecule 2. The structures shown are the molecule 2 and one of the molecule 1 monomers within the nonsymmetric hexamer with the lowest total energy. Only backbone atoms and the aromatic ring of Phe(B25) are shown.

The phenol molecule occupies a well-defined position in the trimer interface, as shown in Figure 10. Primarily, the phenol molecule is bound to insulin through hydrophobic interactions, as inferred from the observed NOE contacts between the phenol molecules and a series of hydrophobic residues, and predominantly to the side chains of these residues. Thus, NOEs are observed between phenol and Cys(A6), Cys(A7), Ile(A10), Cys(A11), Leu(A16), Leu(B11), and Ala(B14) in one monomer, His(B5) and Leu(B6) in another monomer, and Leu(B17) in a third monomer. Also, a hydrogen bond between the oxygen atom of phenol and the slowly exchanging amide proton of Cys(A11) was found in 11 of the 20 calculated structures, while a hydrogen bond between the hydroxyl proton of phenol and the carbonyl oxygen of Cys(A6) was found in four of the 20 calculated structures. Further, the extreme upfield chemical shift of the $\delta$-proton of His(B5) is in agreement with the position of the phenol molecule just below this proton and perpendicular to the His(B5) imidazole ring (Figure 10).

In general, the solution structure of the $R_6$ hexamer is highly symmetric, as indicated by the unique set of resonances for each residue, except for Phe(B25) and Phe(B24).

FIGURE 9: Schematic representation of the nonsymmetric $R_6$ hexamer of human insulin, containing five monomers with the molecule 1 conformation and one monomer with the molecule 2 conformation (upper right). The disordered C terminus of the B-chain [Lys(B29) and Thr(B30)] was excluded. Side chains are shown only for the Phe(B25) residues (black) and the His(B10) residues (light gray) bound to the two Zn atoms (white). The six phenol molecules in the trimer interfaces are also shown (gray). The figure was prepared using MOLSCRIPT (Kraulis, 1991).

However, the two different orientations of the side chain of Phe(B25) result in nonsymmetric as well as symmetric hexamers. According to the ratio of about 10:1 between molecule 1 and molecule 2, it is unlikely that the nonsymmetric hexamers contain more than one monomer with the molecule 2 conformation. As shown in Figure 9, the side chain of one of the Phe(B25) residues in the nonsymmetric hexamer is turned outward from the monomer and points across the dimer interface (molecule 2 conformation), whereas the Phe(B25) rings in the other monomers are turned inward (molecule 1 conformation) and make van der Waals contact with the Tyr(A19) of the same monomer.

The presence of both a symmetric and a nonsymmetric hexamer is in qualitative agreement with the crystal structure (Smith & Dodson, 1992), although the fraction of the molecule 1 conformation is somewhat larger in the solution hexamer than in the crystal hexamer. The low ratio of molecule 2 is immediately surprising, considering the essentially identical energies of these two hexamers (Table 2) and the negligible effect of the orientation of the Phe(B25) phenyl ring on the bulk of the structure. However, the difference in population could result from a subtle difference in the local energy of the region around Phe(B25), which is significant for this region but lies within the uncertainties of the total energies calculated here. In support of this, energy minimization calculations indicated (Wodak, 1984) that the molecule 1 conformation is energetically most favorable in the absence of crystal packing forces. Also, molecule 1 is the conformation most commonly found in insulin crystal forms and is totally dominating in the cubic crystal form of insulin (Derewenda et al., 1990).

However, the interesting point here is that the molecule 2 conformation *is* present in solution, since it shows that this conformation is not a result of the crystal packing forces, as suggested previously (Chothia et al., 1983). In this context, it is also interesting that a molecular dynamics study of insulin in water (Mark et al., 1991) indicated a nonsymmetric dimer (one molecule 1 and one molecule 2), in qualitative agreement with the structure obtained here. On the other hand, other minor differences between the two conformations that are found in the crystal structure, *i.e.* primarily a change in the orientation of the two A-chain helices (Chothia et al., 1983), are not observed in solution and could, therefore, result from the crystal packing.

*(c) Rigidity of the Hexamer.* In general, the solution structure of the $R_6$ insulin hexamer is rigid and compact. First of all, this is indicated by the large number of dimer and trimer contacts, by the hydrogen bonds and hydrophobic interactions in the calculated hexamer structures, and by the relatively large number of slowly exchanging amide protons, all features that are observed here at the applied pH values (8.0−8.1). Further, the fact that the line widths of the NMR signals are similar in all regions of the molecule and have the size expected for a 36 kDa molecule is indicative of a rigid structure. The similarity in line width also includes the proton signals that are broadened in the Asp(B9) dimer (Kristensen et al., 1991), as for example the residues near the disulfide bonds Cys(A6)−Cys(A11) and Cys(A7)−Cys(B7). On the other hand, the packing in the trimer interface region seems somewhat looser than that in the dimer interface, as indicated by a correspondingly smaller number of observed NOE contacts across the trimer interface (Figure 5).

The low NOE-violation energy (Table 2) is also in accordance with a rigid hexamer structure. Thus, all the NOE restraints are easily fulfilled by a single structure, indicating a minimal time averaging of the distances. Again, this is in agreement with a back-calculation of the NOESY spectrum of the $R_6$ hexamer (Jacoby et al., 1996), which resulted in a number of predicted NOEs similar to that observed experimentally. In contrast, a back-calculation of the NOESY spectrum of a monomeric insulin (Weiss et al., 1991; Hua et al., 1993a) resulted in a number of predicted NOEs considerably larger than the experimental number. This indicates a much looser structure of the free monomer and led to the suggestion that the functional form of monomeric insulin is a molten globule (Hua et al., 1993a).

### *Comparison with the Crystal Structure*

The overall agreement between insulin monomers in the different phases and states of aggregation is shown in Figure 11, where the structure of the molecule 1 monomer in the solution $R_6$ hexamer is compared with the corresponding monomer in the crystal hexamer and with the monomers in the solution structures of the dimeric Asp(B9) insulin (Jørgensen et al., 1992) and the monomeric His(B16) insulin (Ludvigsen et al., 1994). As it appears from Figure 11, the monomers in the solution and crystal hexamers are very similar. Not only are the same secondary structure elements present in the two structures; the relative orientations of these elements and the tertiary fold in general are almost identical.

The largest difference between the monomers in the two $R_6$ hexamers is the structure of the N-terminal end of the B-chain, which is disordered in solution but part of the extended α-helix in the crystal phase. Thus, while the B-chain α-helix in the $R_6$ crystal hexamer stretches continuously from Cys(B19) to Phe(B1) (Derewenda et al., 1989),

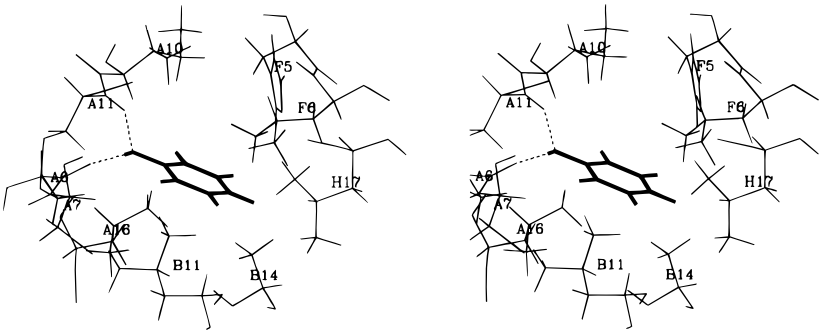


FIGURE 10: Position of the phenol molecule in the trimer interface of the $R_6$ hexamer of human insulin. The two hydrogen bonds observed between phenol and Cys(A6) and Cys(A11) are indicated by dotted lines. The letters B, F, and H are used to distinguish between the B-chains in different monomers.

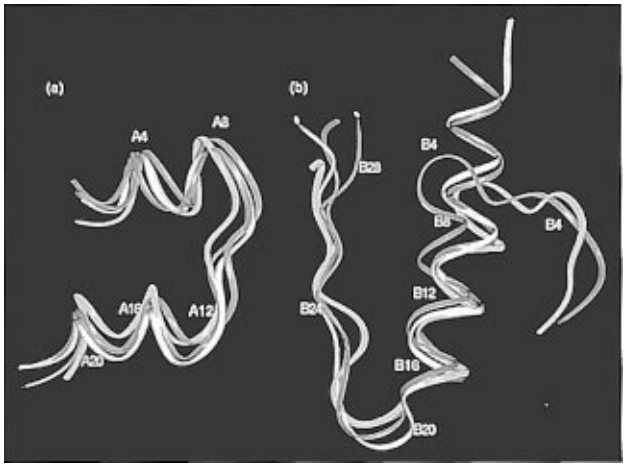


FIGURE 11: Comparison of the molecule 1 within the solution structure of the $R_6$ hexamer of human insulin (light gray) with the molecule 1 in the crystal structure of the $R_6$ hexamer (orange) and with the solution structures of the monomer in the Asp(B9) insulin mutant (yellow) and the monomeric His(B16) insulin mutant (green). The Asp(B9) and the His(B16) solution structures are represented by the structure with the lowest total energy, and the monomer in the $R_6$ hexamer is represented by the average molecule 1 of the symmetric hexamer with the lowest energy. Only the backbone atoms are superimposed, except for the residues from B1 to B8 in Asp(B9) and His(B16). The two chains are shown separately in panels a and b.

the extended B-chain α-helix in the $R_6$ solution hexamer only includes the residues from Cys(B19) through Asn(B3) (Figures 7 and 8). This result is surprising in light of the provisional protomer structure reported by Jacoby et al. (1996), where the extension of the B-chain also includes Val(B2) and Phe(B1), as in the crystal structure. Nevertheless, careful inspection of the 3D TOCSY-NOESY spectrum and the 750 MHz WATERGATE NOESY spectrum (*vide supra*) did not reveal any NOE connectivities, including the helical $d_{\alpha\beta}(i,i+3)$ NOE between Val(B2) and His(B5) reported by Jacoby et al. (1996), which would suggest an extension of the B-chain helix beyond Asn(B3). This holds despite the fact that the part of the NOESY spectrum, which covers the region where these NOEs should be found, is not particularly crowded. Therefore, the results of the present study are compatible only with a B-chain helix from Cys(B19) to Asn(B3).

The different conformations of the N-terminal end of the B-chain in the two phases could be a consequence of the different environments it experiences in these phases. Thus, the N-terminal end of the B-chain is located on the surface of the $R_6$ hexamer, which makes it sensitive to changes in the environment. This explanation was also given for a similar shortening of the extended B-chain α-helices (B3−B19), which was observed in the crystal structures of the $T_3R_3$ hexamers of human insulin (Ciszak & Smith, 1994) and $Lys^{B28}Pro^{B29}$ insulin (Ciszak et al., 1995) and was dubbed a "frayed" α-helix conformation or an $R^f$ conformation. Also in these cases, the conformational differences, as compared to the crystal structure of the $R_6$ insulin hexamer, were attributed to the different environments, here the crystal packing, of the N-terminal end of the B-chain.

*Significance of the Phe(B25) Orientation*

The orientation of the phenyl ring of Phe(B25) plays an important role in the insulin−receptor interaction. Thus, studies of a series of carefully designed insulin mutants (Nakagawa & Tager, 1986; Mirmira & Tager, 1991) showed that the phenyl ring of Phe(B25) is critical for the biological activity of insulin, and it was suggested that Phe(B25) is a potential primary contact between insulin and its receptor. Further, it was suggested that, following initial binding of the Phe(B25) side chain to a specific hydrophobic pocket in the receptor, concerted conformational changes take place in the insulin and the receptor. For the insulin molecule, these changes include a separation of the B-chain C-terminal region from the N-terminal end of the A-chain (Mirmira & Tager, 1989; Derewenda et al., 1991). This leads to a secondary binding event which involves other areas of the insulin monomer, resulting in full biological activity of the insulin molecule and formation of the insulin−receptor complex.

In light of this model for the insulin−receptor interaction, the observation here of the molecule 2 conformation in solution is interesting since this, most likely, is the conformation that interacts with the receptor (Nakagawa & Tager, 1986; Mirmira & Tager, 1991). Thus, the outward orientation of the Phe(B25) phenyl ring in the insulin monomer exposes the ring to the solvent and makes it accessible to the receptor.

Nevertheless, the solution structures of the monomeric His(B16) (Ludvigsen et al., 1994) and des(B26−B30) (DPI) (Knegtel et al., 1991) insulin mutants in water exist only in the molecule 1 conformation, probably because the phenyl ring in this conformation makes a hydrophobic contact to the rest of the molecule, to form a more hydrophilic surface of the monomer. However, in hydrophobic environments, the monomer can take the molecule 2 conformation in solution. This is clearly demonstrated here by the observation of this conformation in the $R_6$ solution hexamer, where

the Phe(B25) phenyl ring of molecule 2 is located in the hydrophobic monomer−monomer interface making hydrophobic contact with the other monomer.

Furthermore, in hydrophobic environments, the monomer can easily exchange between molecule 1 and molecule 2 conformations because of the intrinsic flexibility of the Phe(B25) side chain. This is suggested not only by molecular dynamics simulations (Krüger et al., 1987) but also by the NMR studies of the Asp(B9) insulin dimer (Jørgensen et al., 1992). In this dimer, the two monomers were found to be identical, both having a molecule 1 conformation. However, a number of residues in the dimer interface of the Asp(B9) mutant, including Phe(B24), Phe(B25), Tyr(B26), and Val(B12), show considerable line broadening (Kristensen et al., 1991). This strongly indicates an averaging in this region between different conformations and suggests a larger intrinsic mobility around Phe(B25) in the free solution dimer than in the dimer of the $R_6$ solution hexamer, where the mobility is sufficiently dampened by the hexamer packing to allow individual observation of molecule 1 and molecule 2. The flexibility of the side chain of Phe(B25) is further indicated by the experimental observation that the phenyl ring of Phe(B25) in molecule 2 of the 2Zn crystal structure sometimes occupy a molecule 1 conformation (Baker et al., 1988) and by relaxation studies of insulin in 20% acetic acid (Hua & Weiss, 1991).

Finally, of interest in this context is the suggestion that the molecule 2 is present also in the solution structures of DPI (Hua & Weiss, 1990), human insulin (Hua & Weiss, 1991), and proinsulin (Weiss et al., 1990). This suggestion was based on the fact that no NOE contacts between Phe(B25) and Tyr(A19) were found in these structures. In spite of the observation of only the molecule 1 conformation for other monomeric insulins (Knegtel et al., 1991; Ludvigsen et al., 1994), the suggestion might well be valid, since all three NMR studies mentioned first were carried out in 20% acetic acid, which may stabilize the molecule 2 conformation through hydrophobic interactions between the phenyl ring of Phe(B25) and the organic solvent.

In conclusion, the observation here of the molecule 2 conformation in solution supports the model suggested for the insulin−receptor interaction (Nakagawa & Tager, 1986; Mirmira & Tager, 1991) and outlined briefly above. It suggests, together with other experimental results and theoretical calculations, a large intrinsic flexibility of the Phe(B25) phenyl ring; *i.e.* the orientation of this ring is easily changed without making any significant structural changes in the rest of the molecule. Also, it shows that the phenyl ring of Phe(B25) adopts a molecule 2 conformation to some extent when it can make favorable hydrophobic interactions with another insulin molecule. The same favorable hydrophobic interaction may occur when insulin binds to the receptor.

## ACKNOWLEDGMENT

We are grateful to Per Balschmidt, Novo Nordisk A/S, for the donation of the Zn insulin, Georg Ole Sørensen for computational assistance, and Else Philipp, Søren M. Kristensen, Morten D. Sørensen, Thomas B. Jørgensen, Jan Makropulos, and Aage Nissen for technical assistance during the experimental work. The 750 MHz spectra were obtained at The Danish Instrument Centre for NMR Spectroscopy of Biological Macromolecules.

## SUPPORTING INFORMATION AVAILABLE

Tables containing the $^1H$ and $^{13}C$ chemical shifts of the monomers in the $R_6$ insulin hexamer and the NMR-derived distance restraints used in the structure calculations (22 pages). Ordering information is given on any current masthead page.

## REFERENCES

Arrowsmith, C. H., Pachter, R., Altman, R. B., Iyer, S. B., & Jardetzky, O. (1990) *Biochemistry 29*, 6332−6341.

Badsberg, U., Jørgensen, A. M. M., Gesmar, H., Led, J. J., Hammerstad, J. M., Jespersen, L., & Ulstrup, J. (1996) *Biochemistry 35*, 7021−7031.

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., & Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London, Ser. B 319*, 369−456.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.

Blundell, T. L., Dodson, G. G., Hodgkin, D. M. C., & Mercola, D. A. (1972) *Adv. Protein Chem. 26*, 279−402.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93−106.

Brader, M. L., & Dunn, M. F. (1991) *Trends Biochem. Sci. 16*, 341−345.

Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R., & Voigt, H. O. (1988) *Nature 333*, 679−682.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Breg, J. N., van Opheusden, J. H. J., Burgering, M. J. M., Boelens, R., & Kaptein, R. (1990) *Nature 346*, 586−589.

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., & Frank, B. H. (1992) *Protein Eng. 5*, 527−533.

Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brünger, A. T. (1992) *X-PLOR Manual*, Yale University Press, New Haven, CT.

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

Chothia, C., Lesk, A. M., Dodson, G. G., & Hodgkin, D. C. (1983) *Nature 302*, 500−505.

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

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

Clore, G. M., Appella, E., Yamada, M., Matsushima, K., & Gronenborn, A. M. (1990) *Biochemistry 29*, 1689−1696.

Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E., & Gronenborn, A. M. (1994) *Science 265*, 386−391.

Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E., & Gronenborn, A. M. (1995) *Science 267*, 1515−1516.

Derewanda, U., Derewanda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swenson, D. (1989) *Nature 338*, 594−596.

Derewenda, U., Derewenda, Z. S., Dodson, G. G., & Hubbard, R. E. (1990) in *Handbook of Experimental Pharmacology* (Cuatrecasas, P., & Jacobs, S., Eds.) Vol. 92, pp 23−39, Springer-Verlag, Berlin.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. G., & Markussen, J. (1991) *J. Mol. Biol. 220*, 425−433.

Dodson, E. J., Dodson, G. G., Hubbard, R. E., & Reynolds, C. D. (1983) *Biopolymers 22*, 281−291.

Driscoll, P. C., Clore, G. M., Beress, L., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2178−2187.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. (1979) *Faraday Symp. Chem. Soc. 13*, 49−55.
Folkers, P. J. M., Folmer, R. H. A., Konings, R. N. H., & Hilbers, C. W. (1993) *J. Am. Chem. Soc. 115*, 3798−3799.
Gesmar, H., Nielsen, P. F., & Led, J. J. (1994) *J. Magn. Reson. B103*, 10−18.
Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
Hardaway, L. A., Brems, D. N., Beals, J. M., & MacKenzie, N. E. (1994) *Biochim. Biophys. Acta 1208*, 101−103.
Hua, Q. X., & Weiss, M. A. (1990) *Biochemistry 29*, 10545−10555.
Hua, Q. X., & Weiss, M. A. (1991) *Biochemistry 30*, 5505−5515.
Hua, Q. X., Shoelon, S. E., Kochoyan, M., & Weiss, M. A. (1991) *Nature 354*, 238−241.
Hua, Q. X., Kochoyan, M., & Weiss, M. A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2379−2383.
Hua, Q. X., Kochoyan, M., & Weiss, M. A. (1993a) *Proc. Natl. Acad. Sci. U.S.A. 90*, 582−586.
Hua, Q. X., Jia, W., Frank, B. H., & Weiss, M. A. (1993b) *J. Mol. Biol. 230*, 387−394.
Jacoby, E., Hua, Q. X., Stern, A. S., Frank, B. H., & Weiss, M. A. (1996) *J. Mol. Biol. 258*, 136−157.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
Jørgensen, A. M. M., Kristensen, S. M., Led, J. J., & Balschmidt, P. (1992) *J. Mol. Biol. 227*, 1146−1163.
Jørgensen, A. M. M., Olsen, H. B., Balschmidt, P., & Led, J. J. (1996) *J. Mol. Biol. 257*, 684−699.
Kaarsholm, N. C., Ko, H.-C., & Dunn, M. F. (1989) *Biochemistry 28*, 4427−4435.
Kay, L. E., Forman-Kay, J. D., McCubbin, W. D., & Kay, C. M. (1991) *Biochemistry 30*, 4323−4333.
Kessler, H., Schmieder, P., & Bermel, W. (1990) *Biopolymers 30*, 465−475.
Kline, A. D., & Justice, R. M., Jr. (1990) *Biochemistry 29*, 2906−2913.
Knegtel, R. M. A., Boelens, R., Ganadu, M. L., & Kaptein, R. (1991) *Eur. J. Biochem. 202*, 447−458.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kristensen, S. M., & Led, J. J. (1995) *Magn. Reson. Chem. 33*, 461−470.
Kristensen, S. M., Jørgensen, A. M. M., Led, J. J., Balschmidt, P., & Hansen, F. B. (1991) *J. Mol. Biol. 218*, 221−231.
Kristensen, S. M., Sørensen, M. D., Gesmar, H., & Led, J. J. (1996) *J. Magn. Reson. B112*, 193−196.
Krüger, P., Strassburger, W., Wollmer, A., van Gunsteren, W. F., & Dodson, G. G. (1987) *Eur. Biophys. J. 14*, 449−459.
Led, J. J., & Gesmar, H. (1991) *J. Biomol. NMR 1*, 237−246.
Lee, W., Harvey, T. S., Yin, Y., Yau, P., Litchfield, D., & Arrowsmith, C. H. (1994) *Nat. Struct. Biol. 1*, 877−890.
Ludvigsen, S., Roy, M., Thøgersen, H., & Kaarsholm, N. C. (1994)-*Biochemistry 33*, 7998−8006.
Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259−281.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
Marion, D., & Bax, A. (1989) *J. Magn. Reson. 83*, 205−211.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
Mark, A. E., Berendsen, H. J. C., & van Gunsteren, W. F. (1991) *Biochemistry 30*, 10866−10872.
Mirmira, R. G., & Tager, H. S. (1989) *J. Biol. Chem. 264*, 6349−6354.
Mirmira, R. G., & Tager, H. S. (1991) *Biochemistry 30*, 8222−8229.
Mirmira, R. G., Nakagawa, S. H., & Tager, S. (1991) *J. Biol. Chem. 266*, 1428−1436.
Nakagawa, S. H., & Tager, S. (1986) *J. Biol. Chem. 261*, 7332−7341.
Nakagawa, S. H., & Tager, S. (1992) *Biochemistry 31*, 3204−3214.
Nilges, M. (1993) *Proteins: Struct., Funct., Genet. 17*, 297−309.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., & Campbell, I. D. (1990) *J. Magn. Reson. 87*, 488−501.
O'Donoghue, S. I., King, G. F., & Nilges, M. (1996) *J. Biomol. NMR 8*, 193−206.
Peking Insulin Structure Research Group (1971) *Peking Rev. 40*, 11−16.
Piotto, M., Saudek, V., & Sklenár, V. (1992) *J. Biomol. NMR 2*, 661−665.
Rance, M. (1987) *J. Magn. Reson. 74*, 557−564.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250−254.
Renscheidt, H., Strassburger, W., Glatter, U., Wollmer, A., Dodson, G. G., & Mercola, D. A. (1984) *Eur. J. Biochem. 142*, 7−14.
Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274−293.
Smith, G. D., & Dodson, G. G. (1992) *Biopolymers 32*, 441−445.
Sørensen, O. W., Rance, M., & Ernst, R. R. (1984) *J. Magn. Reson. 56*, 527−539.
Tirendi, C. F., & Martin, J. F. (1989) *J. Magn. Reson. 81*, 577−585.
Weiss, M. A. (1990) *J. Magn. Reson. 86*, 626−632.
Weiss, M. A., Frank, B. H., Khait, I., Pekar, A., Heiney, R., Shoelson, S. E., & Neuringer, L. J. (1990) *Biochemistry 29*, 8389−8401.
Weiss, M. A., Hua, Q. X., Lynch, C. S., Frank, B. H., & Shoelson, S. E. (1991) *Biochemistry 30*, 7373−7389.
Wishart, D. S., & Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
Wodak, S. J., Alard, P., Delhaise, P., & Renneboog-Squilbin, C. (1984) *J. Mol. Biol. 181*, 317−322.
Wollmer, A., Rannefeld, B., Johnsen, B. R., Hejnaes, K. R., Balschmidt, P., & Hansen, F. B. (1987) *Biol. Chem. Hoppe-Seyler 368*, 903−911.
Wüthrich, K. (1986) *N.M.R. of Proteins and Nucleic Acids*, John Wiley and Sons, Inc., New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.

BI9631069