Journal of Biomolecular NMR, **16:** 93–108, 2000.
KLUWER/ESCOM

Unraveling the symmetry ambiguity in a hexamer: Calculation of the R_6 human insulin structure

Seán I. O'Donoghue[a], Xiaoqing Chang[b], Roger Abseher[a], Michael Nilges[a] & Jens J. Led[b,*]
[a]*European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany*
[b]*Department of Chemistry, University of Copenhagen, The H.C. Ørsted Institute, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark*

Received 18 November 1999; Accepted 9 December 1999

Key words: ambiguous distance restraints, insulin hexamer, principal component analysis, solution structure, symmetric oligomers

Abstract

Crystallographic and NMR studies of insulin have revealed a highly flexible molecule with a range of different aggregation and structural states; the importance of these states for the function of the hormone is still unclear. To address this question, we have studied the solution structure of the insulin R_6 symmetric hexamer using NMR spectroscopy. Structure determination of symmetric oligomers by NMR is complicated due to 'symmetry ambiguity' between intra- and intermonomer NOEs, and between different classes of intermonomer NOEs. Hence, to date, only two symmetric tetramers and one symmetric pentamer (VTB, B subunit of verotoxin) have been solved by NMR; there has been no other symmetric hexamer or higher-order oligomer. Recently, we reported a solution structure for R_6 insulin hexamer. However, in that study, a crystal structure was used as a reference to resolve ambiguities caused by the threefold symmetry; the same method was used in solving VTB. Here, we have successfully recalculated R_6 insulin using the symmetry-ADR method, a computational strategy in which ambiguities are resolved using the NMR data alone. Thus the obtained structure is a refinement of the previous R_6 solution structure. Correlated motions in the final structural ensemble were analysed using a recently developed principal component method; this suggests the presence of two major conformational substates. The study demonstrates that the solution structure of higher-order symmetric oligomers can be determined unambiguously from NMR data alone, using the symmetry-ADR method. This success bodes well for future NMR studies of higher-order symmetric oligomers. The correlated motions observed in the structural ensemble suggest a new insight into the mechanism of phenol exchange and the $T_6 \leftrightarrow R_6$ transition of insulin in solution.

Abbreviations: 1AIY, PDB code for previous solution structure of R_6 (Chang et al., 1997); 1ZNJ, PDB code for crystal structure of R_6 (Derewenda et al., 1989); ADR, ambiguous distance restraint; DRS, distance restraint set; DSYM, distance symmetry; MDSA, molecular dynamical simulated annealing; NCS, non-crystallographic symmetry; NOE, nuclear Overhauser enhancement; PCA, principal component analysis; rms, root mean square; rmsd, root mean square deviation; R_6, phenol-induced insulin hexamer; T_6, 2-Zn insulin hexamer; T_3R_3, 4-Zn insulin hexamer.

Introduction

Insulin is a small peptide hormone consisting of two chains, A and B (21 and 30 residues, respectively), joined together by two inter-chain disulphide bonds. Due to the medical importance of insulin, there have been numerous studies of its structure and flexibility in order to get insight into the structural basis for its function and aggregation properties (Blundell et al., 1972; Chothia et al., 1983; Baker et al., 1988; Hua et al.,

*To whom correspondence should be addressed. E-mail: Led@kl5axp.ki.ku.dk.

1991; Jørgensen et al., 1992, 1996). Insulin can dimerize, and, in the presence of Zn^{2+} at neutral pH, it can also form a 36 000 M_r hexamer. In the hexameric state, insulin can make transitions between three conformational states designated T_6, T_3R_3, and R_6 (Kaarsholm et al., 1989). In each of these states, a trimer of dimers is arranged with point group 32 symmetry (for the T_3R_3 state, the dimers are only quasi-symmetric). The $T_6 \rightarrow R_6$ transition can be driven to completion by phenol and other phenolic compounds (Derewenda et al., 1989).

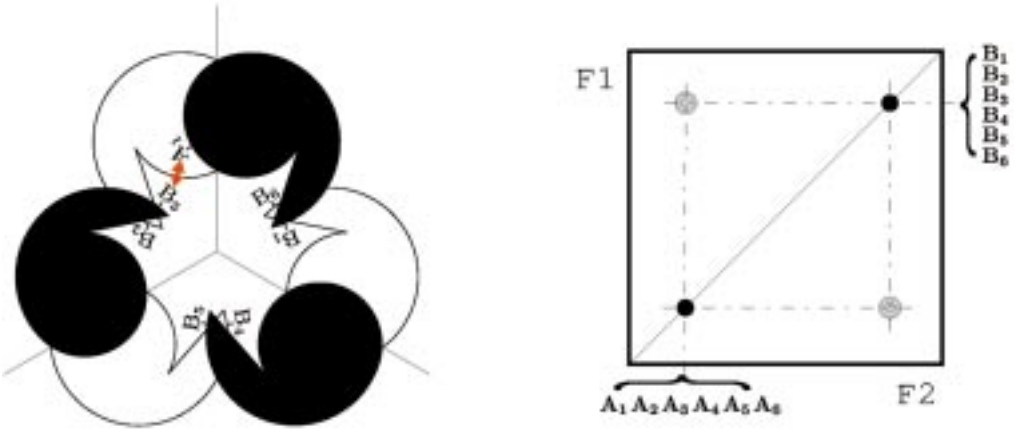
Recently we reported the NMR solution structure of the phenol-stabilized R_6 insulin hexamer (PDB code 1AIY, Chang et al., 1997). However, that study relied on the crystal structure of the R_6 hexamer (1ZNJ, Derewenda et al., 1989) and on previously solved NMR solution structures of a monomer (Jørgensen et al., 1996) and a dimer (Jørgensen et al., 1992), to unravel spectral ambiguities that arise from the symmetry of the hexamer. In NMR spectra of symmetric oligomers, symmetry-related nuclei have degenerate chemical shift due to identical chemical environments. Thus, in NOESY spectra of symmetric oligomers it is intrinsically impossible to distinguish which cross peaks arise from intramonomer interactions, and which arise from the different classes of intermonomer interactions (for a symmetric hexamer, there are five different classes of inter-monomer NOEs; see Figure 1). This symmetry degeneracy simplifies the assignment of resonances; however, it complicates the NOE assignment and the structure calculation considerably. Two further ambiguities complicate the structure calculation even more: firstly, uncertainty about the order of the oligomer (i.e., the number of monomers per molecule); secondly, ambiguity of the point group symmetry for oligomers of even order (i.e., tetramers, hexamers, octamers, etc.) e.g., a symmetric hexamer may have point group 6 (all monomers arranged equally around a single symmetry axis) or point group 32 symmetry (a trimer of dimers). Recently the problems associated with the calculation of structures of symmetric oligomers from NMR data were discussed in detail by O'Donoghue and Nilges (1999).

Several experimental approaches have been proposed to solve the symmetry degeneracy problem, where the basic idea is to break the symmetry by mixing isotopically labelled and unlabelled monomers within the symmetric oligomers (Arrowsmith et al., 1991). However, these methods are sufficient only in the case of dimers; for higher-order oligomers, ambiguity still remains in distinguishing between the different classes of intermonomer NOEs. The order of the oligomer can be determined by a variety of experimental methods: equilibrium ultracentrifugation; dynamic light scattering; osmometry; or NMR line-broadening. However, experimental methods cannot directly resolve point group ambiguity.

An alternative computational solution to the symmetry degeneracy problem is the symmetry-ADR method (Nilges, 1993; O'Donoghue et al., 1993). This method is complementary to the experimental approaches and is more general, as it can be applied to any oligomeric symmetry. In this method, the ambiguous NOEs are expressed as ambiguous distance restraints (ADRs) which are used to drive the calculation of structures with a molecular dynamical simulated annealing (MDSA) protocol (Nilges et al., 1988a, b; Nilges, 1995). The symmetry-ADR method has been applied successfully to solve the structures of several symmetric dimers and to one symmetric tetramer (the p53 tetramerization domain; Lee et al., 1994). While the method has proved to be generally successful, specific difficulties remain. In particular, when many NOEs occur close to the symmetry axes, it can be very difficult to obtain the correct orientation of the monomers (O'Donoghue et al., 1996; O'Donoghue and Nilges, 1999). The ADR method may also be able to determine the point group by building structures with each of the possible point groups and seeing which structure gives the best fit to the ambiguous data.

Here we present a calculation of the R_6 hexamer solution structure, where the symmetry-ADR method has been used to derive a structure from the NMR data alone, without reference to the crystal structure. In this calculation the ambiguity of the NOEs was gradually reduced using the automated iterative assignment scheme of the ARIA program (Nilges et al., 1997). The calculation converged towards a well-defined hexamer structure with the same fold as the crystal structure. The success of this calculation demonstrates that the symmetry-ADR method can be extended to a hexamer, the highest-order symmetric oligomer solved to date by NMR. Further, we show that the ambiguous NOE data fit much better to point group 32 than to point group 6; this demonstrates, for the first time, that NMR data can resolve point group ambiguity of a symmetric oligomer.

The NMR solution structure contains dynamics information which is not available from the crystal structures. Therefore, the study also includes an analy-



$$\bar{D} = [d(\mathbf{A_1}, \mathbf{B_1})^{-6} + d(\mathbf{A_1}, \mathbf{B_2})^{-6} + d(\mathbf{A_1}, \mathbf{B_3})^{-6} + d(\mathbf{A_1}, \mathbf{B_4})^{-6} + d(\mathbf{A_1}, \mathbf{B_5})^{-6} + d(\mathbf{A_1}, \mathbf{B_6})^{-6}]^{-1/6}$$

Figure 1. The NMR symmetry-ambiguity problem for a symmetric hexamer. Above left is shown a schematic diagram of a hexamer with point group 32 symmetry. The three twofold axes are indicated as straight lines; the threefold axis is perpendicular to the page, where the twofold axes intersect. An intermonomer NOE across the trimer interface is shown (indicated by a red arrow). The NOE involves atom $\mathbf{A}_1$ on monomer 1 and atom $\mathbf{B}_3$ on monomer 3. Due to symmetry degeneracy in the corresponding NOESY spectra (above right), it is impossible to distinguish this NOE from the other five classes of NOE which may occur ($\mathbf{A}_1 \leftrightarrow \mathbf{B}_1$, $\mathbf{A}_1 \leftrightarrow \mathbf{B}_2$, $\mathbf{A}_1 \leftrightarrow \mathbf{B}_4$, $\mathbf{A}_1 \leftrightarrow \mathbf{B}_5$, and $\mathbf{A}_1 \leftrightarrow \mathbf{B}_6$). Below the figures, we show the formula used to express this ambiguity as an ambiguous distance restraint (ADR) in the symmetry-ADR method. The formula essentially chooses the minimum distance; during the structure calculation, $\bar{D}$ is restrained to the experimental distance derived from the cross-peak intensity. An additional complication is that the NOE cross peak also includes contributions from the five other possible permutations for this class of NOE, i.e., $\mathbf{A}_2 \leftrightarrow \mathbf{B}_6$, $\mathbf{A}_3 \leftrightarrow \mathbf{B}_5$, $\mathbf{A}_4 \leftrightarrow \mathbf{B}_2$, $\mathbf{A}_5 \leftrightarrow \mathbf{B}_1$, and $\mathbf{A}_6 \leftrightarrow \mathbf{B}_4$. However, this problem is easily overcome since, due to symmetry, we know that each permutation must contribute equally.

sis of the structural diversity within the final solution structure ensemble using a newly developed approach (Abseher et al., 1998a) that is based on the principal component analysis (PCA) method of Amadei et al. (1993). The PCA method, originally developed for analyzing molecular dynamics trajectories, finds a series of modes, each of which corresponds to a collective fluctuation that occurs in the trajectory. The interesting feature of PCA is that, generally, most of the motion in a trajectory can be captured using only the first few modes of highest amplitude (e.g. the first five). This enables a reasonably concise and accurate description of the motion in the trajectory. In several cases, PCA was applied both to NMR structure ensembles and to corresponding molecular dynamics trajectories (Abseher et al., 1998a). In all of these cases, the analyses gave good agreement, suggesting that the large amplitude modes obtained from NMR ensembles do indeed correspond to collective motions which occur in the molecule. The modes obtained here for insulin provide insight into the $T_6 \rightarrow R_6$ transition and the exchange of phenol.

Materials and methods

Sample preparation and spectra acquisition

The phenol-stabilized hexamer of human insulin was prepared according to the experimental procedures (Wollmer et al., 1987; Brader et al., 1991) that resulted in a well-defined R_6 hexamer, and the NOE restraints were obtained from NMR spectra recorded at pH 8.0–8.1 and 310 K, as described previously (Chang et al., 1997).

Monomer distance restraint set

The classification of *inter*- and *intra*-monomeric NOEs and the structure calculation of the insulin monomer within the hexamer were carried out as an iterative process. Despite the complications caused by the high symmetry of the hexamer, most of the NOEs can be unambiguously assigned as intramonomeric. Thus, in the initial structure calculation a reduced NOE data set was applied which only included NOEs that can be unambiguously assigned to an insulin monomer and, therefore, define a monomer structure of the correct overall fold. The identification of this NOE data set was relatively straightforward, since the fold of an insulin monomer is well documented (two

A and B chains connected by three disulfide bridges), and monomer-specific and dimer-specific NOEs have been inferred in the previous NMR study of two engineered insulin mutants (Jørgensen et al., 1992, 1996). Therefore, NOEs also observed as intramonomeric in the isolated des-[Phe(B25)] monomer (Jørgensen et al., 1996) or in Asp(B9) dimer (Jørgensen et al., 1992) were applied as intramonomeric in the initial monomer calculation, together with the sequential NOEs characteristic for the B chain N-terminal helical secondary structure. Once the correct starting structure of a monomer was obtained, the originally ambiguous NOEs which could not be readily assigned were included subsequently in the calculations. The consistency or inconsistency of these NOEs with the starting structure were used as the criterion to distinguish different NOE types, i.e. NOEs giving large violations were classified as intermonomeric. By applying this procedure, 665 of the NOEs could be unambiguously assigned as intramonomeric.

Each monomer unit within the hexamer also includes one phenol molecule non-covalently attached to the protein. NOEs between the phenol molecule and the protein molecules could not be unambiguously assigned at this stage. Hence initially the phenol molecule was included in the calculation by arbitrarily choosing one of the phenol-insulin monomer NOEs (phenol H_2/H_6-protons to the amide proton of Cys(A11)). Thus the initial position of the phenol molecule in the monomer was not exactly defined, but it occurred in about the right position.

Hexamer distance restraint set

The remaining symmetry-ambiguous NOEs were treated as symmetry-ADRs, where all possible contributions to each NOE were considered (the intramonomer contribution and the five possible intermonomer contributions, see Figure 1).

As noted previously (Chang et al., 1997), NOE patterns characteristic of a β-sheet are observed in the B-chain, from B22 to B28. Also the amide protons of residues B24 and B26 exchange slowly, indicating that they are involved in hydrogen bonds. This information was included in the present calculation by specifying an ambiguous intermonomer hydrogen bond from each of the slowly exchanging amides to the backbone oxygen atoms of residues B22 through B28 on the five other B-chains in the hexamer.

In the crystal structure, the zinc atoms co-ordinate to the three His-B10 residues along the threefold axis at the top and at the bottom of the hexamer, respectively. However, since the zinc atoms are not observed in the NMR spectra, they were not included in the calculation. Previously, we showed that excluding the zinc and its coordination with the His residues did not greatly affect the structure (Chang et al., 1997).

Structure calculation

The monomer structures were calculated using the standard sa.inp and refine.inp protocols within X-PLOR 3.1 (Brünger, 1992). Except where noted otherwise, the standard NMR force field (parallhdg) was used throughout.

The calculation of the hexamer structures started from the monomer structures using the MDSA-SO-WDMR-1.1 protocol ('MDSA for symmetric oligomer starting with a well-defined monomer'; O'Donoghue, 1998). This protocol was developed specifically for the present calculation; it is similar to the previously described MDSA-SO-WDMR-1.0 protocol O'Donoghue et al., 1996; O'Donoghue, 1998) except that the initial search and cooling stages use rigid-body dynamics. These protocols are both variations on the MDSA protocols developed by Nilges et al. (1988a,b).

In the protocol, the initial hexamer structure was generated by making six identical copies of the monomer arranged with point group 32 symmetry. The separation between monomer centres was set to twice the radius of gyration of the monomer. The overall symmetric arrangement for the hexamer was maintained by using two symmetry terms. The first term, 'non-crystallographic symmetry' (NCS), kept all the monomers superimposable. In the second term, 'distance symmetry' (DSYM), a subset of the intermonomer distances was chosen, and all symmetry-related distances were restrained to have the same value (Nilges, 1993); this ensured that the relative orientations of the six monomer subunits obeyed point group 32 symmetry. An alternative calculation was also made using point group 6 symmetry for the initial orientation of monomers and for the DSYM term. The protocol we used to generate the DSYM restraints is available (O'Donoghue, 1998).

In the previous calculations, the NCS term was turned off for the side chain of one of the Phe-B25 to account for a minor asymmetry observed in the spectra. However, since this asymmetry occurs in only a small fraction of the molecules (around 10%), only the completely symmetric hexamer was considered here.

The hexamer structure was annealed in four stages. The first stage was a 40 ps high temperature search

stage in which the monomers were constrained to move as rigid bodies. In order to significantly reorientate the (relatively heavy) monomer subunits in this short time, it was necessary to set the temperature very high (2 000 000 K); alternatively, one could also scale down the atomic masses (all atoms were set to a mass of 100 M_r). All the non-bonded interactions were turned off except those involving α-carbons, which were given an increased radius. In the second stage, the temperature of the system was cooled to 1000 K over 30 ps. In the third stage, conventional molecular dynamics was used instead of rigid-body motion, and nonbonded interactions were calculated for all atoms.

The non-crystallographic symmetry (NCS) term was switched on to maintain similarity of the monomer subunits. As with the MDSA-SO-WDMR-1.0 protocol, the weights on the bond and angle terms were set to the final high values to maintain the monomer structure. The temperature of the system was initialized at 2000 K and cooled to 100 K over 40 ps. The final stage consisted of 500 cycles of Powell minimization.

A significant problem at this point in the calculation of symmetric oligomers is in calculating the rmsd between structures. The problem arises because during the annealing procedure, the monomers can become shuffled within the hexamer; in order to calculate a meaningful rmsd, we need to rename the monomers in a standard way. Our approach was to take a standard structure (we used the previous solution structure, 1AIY), and for each new structure check all possible renamings of the monomers; the renaming with the lowest rmsd was then selected.

Once sufficiently many low-energy structures were obtained, the ambiguous NOEs were analysed using the ARIA assignment filter (Nilges et al., 1997; Nilges and O'Donoghue, 1999). The filter loops over all low-energy structures, and calculates the relative contribution of each assignment possibility to the total NOE intensity using r^{-6}-weighting. The assignment possibilities are then sorted in order of their contribution, and the least likely possibilities are excluded, depending on a cut-off parameter p. For the initial iteration, p was set to 0.999, corresponding to a conservative assignment criterion (i.e., the only possibilities excluded are those that together contribute less than 0.1% to the total NOE intensity). A new distance restraints set with slightly reduced total ambiguity was generated by applying the filter to each ambiguous NOE. This DRS was then used to calculate a second round of structures. The procedure was iterated six times, gradually reducing p from 0.999 to 0.93.

The low energy structures were then refined within a shell of water molecules using a full electrostatics force field specifically designed for NMR refinement (Linge and Nilges, 1998). Refined average structures were calculated from ensembles using the probability-map refinement method of DeLano and Brünger (1994).

Principal component analysis

Principal component analysis (PCA) of the final structure ensembles was done using the method of Amadei et al. (1993). In this method, a Cartesian coordinate covariance matrix is calculated and diagonalized. As the large amplitude modes are generally backbone modes, the protein is represented by α-carbons only. For the phenol molecule, two alternative representations were chosen: (i) one atom, i.e., balanced with the representation used for the protein. (ii) All atoms, thus putting greater emphasis on the mobility of the phenol molecule. The second representation was chosen as we were particularly interested in examining the correlation between the phenol molecule and the protein in order to understand how the phenol molecule diffuses in and out.

PCA in Cartesian coordinate space requires the removal of overall motion prior to the analysis, which is done by a least squares fit to a reference structure. The bias introduced by this procedure was kept minimal by using only the least mobile backbone atoms for the fitting (Abseher and Nilges, 1998). The PCA was performed on the hexamer structures. However, due the use of the NCS and DSYM symmetry restraint terms and the symmetric NOE restraints, the eigenvector profiles of each monomer are virtually identical. Hence only one averaged monomer profile is presented here (see Results, Figure 5).

The mobility in the ensemble was analysed by projecting each structure along a selected eigendirection. This projection filters out non-correlated motions, producing a new ensemble of structures which differ only along the given eigendirection. These projected structures offer an instructive way to isolate and study a single correlated mode of motion which occurs in the ensemble.

The atomic coordinates of the final ensemble of 20 structures, the refined average structure of the ensemble, the refined average structures of the 'red' and 'green' substate, and the final assigned distance-restraint set have been deposited in the Brookhaven

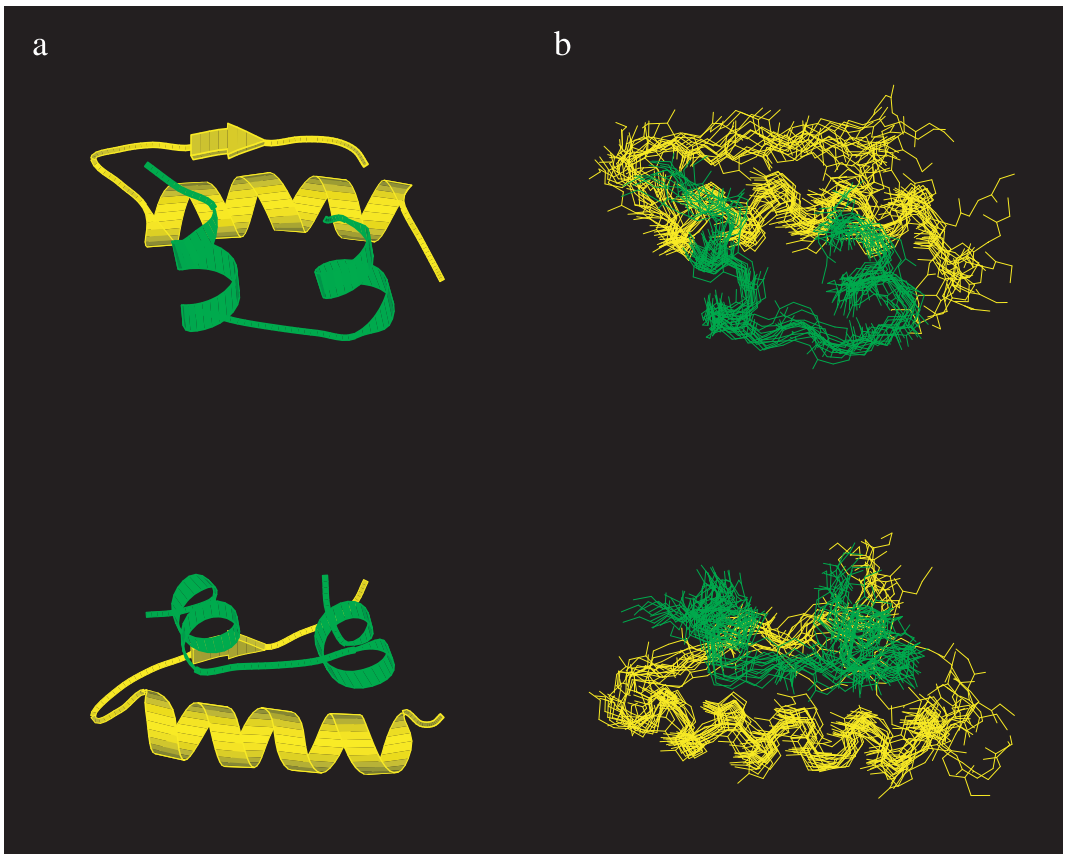


Figure 2. The monomer structure within the R_6 insulin hexamer. (a) Ribbon diagram of the refined average structure. (b) Backbone plot of the 15 lowest energy structures superimposed onto the refined average structure. In both (a) and (b), the top view and the bottom view are related by a 90° rotation along the horizontal axis. The A-chain is shown in green and the B-chain in yellow. The figure was generated using Molscript (Kraulis, 1991).

Protein Data Bank (PDB ID codes: 2aiy, 20 best structures; 3aiy, refined average of 20 best; 4aiy, green substate (refined average structure); 5aiy, red substate (refined average structure)).

Results

Monomer calculation

The structure determination was based on the two-dimensional proton NMR spectra and the NOE assignment described in Materials and methods. The monomer structure was calculated using the 665 NOEs previously assigned as intramonomeric. One of the (ambiguous) phenol-peptide NOEs was arbitrarily chosen to position phenol approximately in the monomer structure. The resulting monomer distance restraint set (DRS) of 666 NOEs was used to calculate 1225 monomer structures using standard MDSA protocols. The monomer structures are well defined (Figure 2), with good covalent geometry (rmsds for bond length, bond angles, and improper dihedral angles of 0.005 Å, 0.6°, and 0.4°, respectively), good agreement with the intramonomer DRS (mean rms deviations for the NOE restraints of 0.02 Å), and relatively good convergence (2.0 Å mean backbone rmsd to the average structure).

Hexamer calculation

Of the total number of NOEs observed in the NOESY spectra, 57 could not be assigned unambiguously by the above method. Hence these NOEs were treated as symmetry-ADRs (Figure 1). The 665 unambiguous intramonomer NOEs were also added to the hexamer DRS. Finally two ambiguous hydrogen bond restraints associated with a β-sheet from residues B22 to B28 and indicated by the NMR data (Chang et al., 1997) were included (see Materials and methods). The re-

Table 1. ARIA iterations for insulin hexamer

Iteration	Structures[a]	p[b]	Inter[c]	Ambig[d]
0			0	57
1	1225	0.999	6	51
2	124	0.999	7	50
3	189	0.99	15	42
4	200	0.98	40	17
5	263	0.96	48	9
6	374	0.93	49	8

[a]For each iteration, the calculation was repeated until at least 40 hexamer structures were obtained with an energy below the threshold set in Figure 3.
[b]Criterion for the ARIA assignment filter.
[c]Number of unambiguous intermonomer NOEs per monomer assigned by the ARIA filter. The 49 intermonomer NOEs assigned by the sixth iteration were all compatible with those made previously (Chang et al., 1997) by reference to the crystal structure.
[d]Number of remaining ambiguous NOEs per monomer. In each iteration, 665 unambiguous intramonomer NOEs were used.

sulting DRS (724 restraints per monomer) was used, together with symmetry restraint terms specific for point group 32, to calculate an initial ensemble of 1225 hexamer structures using the symmetry-ADR method (see Materials and methods). Each structure calculation was started using a different monomer structure with a randomized initial orientation.

Overall, about 12% of the 1225 structures converged to a fold similar to the lowest-energy structure (Figure 3). The 40 structures with lowest energy all have the same general fold (Figure 3) which, furthermore, is the same as the fold of the crystal structure (Derewenda et al., 1989) (1ZNJ). This indicates that the hexamer DRS uniquely identifies the correct fold, i.e., the 57 symmetry-ambiguous NOEs contain enough information to determine the orientation of the monomer within the hexamer.

The search method was inefficient at finding structures both with the correct fold and with low energy. Hence it was necessary to calculate a large number of structures (1225) in order to clearly identify the correct fold based only on total energy. However, the poor convergence does not affect the certainty with which the ambiguous data identify the correct fold; rather, the poor convergence reflects the inefficiency of the search method. While a more efficient search strategy may be possible, the computer time required by the current method was manageable (each hexamer calculation took about 40 min using a single processor Digital Alpha Workstation 500au).

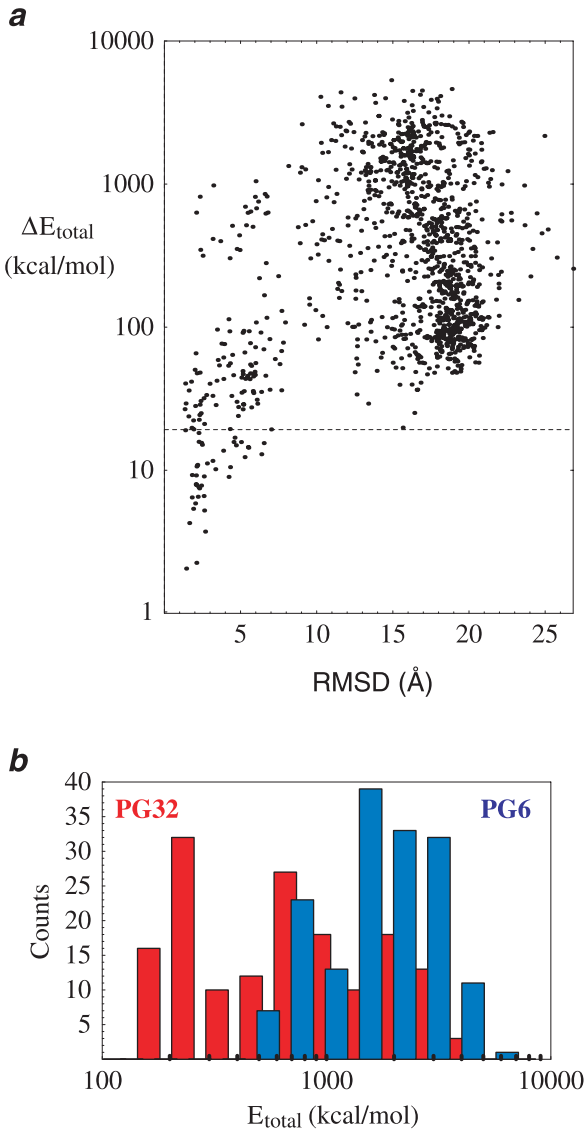


Figure 3. (a) Convergence plot of the initial ensemble of 1225 structures. The total energy and rmsd values of each structure are referenced to the lowest energy structure in the ensemble, i.e., by definition the lowest energy structure has zero total energy and zero rmsd. The plot shows a convincing convergence ('energy-funnel') towards a unique fold, indicating that the data unambiguously define the orientation of the monomer within the hexamer. The fold is the same as that of the crystal structure. The 40 lowest energy structures (below the dashed line at 20 kcal/mol) were selected for the initial assignment of the ambiguous distance restraint set (Table 1). (b) Histogram of total energies for point groups 32 and 6. For each point group, 125 structures were calculated; the energy distribution for point group 32 is clearly better than for point group 6. Hence, although the NOE data are ambiguous, they determine the point group symmetry.

From the NMR spectrum of a symmetric hexamer, it is not possible to distinguish between the two possible point group symmetries – 6 and 32 (O'Donoghue and Nilges, 1999). In the above calculation, we have applied symmetry restraints appropriate for a point group 32 hexamer, since all the other insulin crystal structures under similar conditions have this symmetry (Wollmer et al., 1987). Hence it is quite unlikely that in solution the insulin hexamer has point group 6 – in any case, this symmetry is relatively rare, and would be unlikely to sufficiently bury all the hydrophobic surface area of the monomer. However, to rule out this possibility, we repeated the above calculation using starting conformations and symmetry restraints appropriate for point group 6 symmetry. A total of 125 structures were calculated with point group 6 symmetry. In Figure 3b, the distribution of total energy in these structures is compared with the first 125 structures calculated with point group 32 (above). The best structures with point group 32 have a much lower energy, indicating that this symmetry fits the NOE data significantly better than point group 6.

The 40 lowest energy structures (from the point group 32 calculation) were used to partially assign the ambiguous NOEs in the hexamer DRS. The ARIA assignment filter was applied with a rather conservative criterion ($p = 0.999$, i.e., only very unlikely assignment possibilities are excluded). Six NOEs (per monomer) were assigned as unambiguously intermonomeric in the first calculation (Table 1); five on the dimer interface and one on the trimer interface. The ambiguous hydrogen bonds associated with the β-sheet were also partly assigned. The β-sheet was now unambiguously located close to the dimer axis, although the acceptors of the hydrogen bonds were still ambiguous. The calculation was then iterated a total of six times, while the ARIA assignment criterion was gradually tightened to $p = 0.93$ (Table 1). In each iteration, the calculation was stopped after a sufficient number of low energy structures had been obtained. In these calculations much fewer structures were needed than in the first iteration, indicating that the assignments obtained in the first iteration were already sufficient to specify the correct fold. An additional two iterations, where the assignment criterion was gradually tightened to 0.8 (the usual final value for ARIA calculations), further reduced the ambiguity of the DRS. However, the rmsd convergence of the ensemble did not improve, and three NOEs still remained ambiguous. The final p value used is somewhat arbitrary; since the ensemble did not improve with a lower p value, we chose to stop assigning at the sixth iteration with the slightly more conservative criterion of 0.93.

After the sixth iteration, 49 of the 57 ambiguous NOEs had been assigned as intermonomeric (Table 1) while the hydrogen bonds were completely unambiguously assigned. The 40 lowest energy structures formed a well-defined ensemble (backbone rmsd of < 2 Å to the average structure). These structures were further refined within a shell of water molecules using a full electrostatics force field (Linge and Nilges, 1998), and the 20 lowest energy structures were selected as the final ensemble.

Solution structure of the insulin hexamer

Figure 4b shows the final ensemble superimposed onto the refined average structure (Figure 4a). The ensemble is in good agreement with standard covalent geometry and with all the experimental restraints (Table 2). The convergence of the ensemble is reasonably good, with an average rmsd of 0.96 Å for backbone atoms in the well-ordered regions (A2–A21, B4–B20, B22–B27, and the symmetry-equivalent regions). Analysis by PROCHECK-NMR (Laskowski et al., 1996) found that 78% of the residues are in the most favoured regions of the Ramachandran plot, 17% are in the additionally allowed regions, and 1.6% are in the generously allowed regions.

The B-chain has one extended helix (B4–B19; strictly α-helical only in the region B8–B19). In the hexamer, the six B-chain helices are packed together, both across the dimer axes and around the trimer axis, forming the core of the hexamer. The B-chain helix is followed by a β-turn (B20–B23). The C-terminal region (B24–B26) forms a β-sheet across the dimer axis, with one face exposed to the solvent. The A-chain contains two antiparallel α-helices (A3–A9 and A13–A19) connected by a loop. The A-chains pack outside the B-chain helices, forming the bulk of the hexamer surface. The same overall fold is found in the crystal structure and in the previous solution structure (Derewenda et al., 1989; Chang et al., 1997). The previous solution structure (1AIY) has fewer residues in the most favoured regions of the Ramachandran plot, suggesting that the present structure has the higher quality.

The position of the phenol molecule is well defined (Figure 4b). It is bound to insulin primarily through hydrophobic interactions. However, in 18 of the 20 structures a hydrogen bond occurs between the oxygen atom of phenol and the slowly exchanging amide

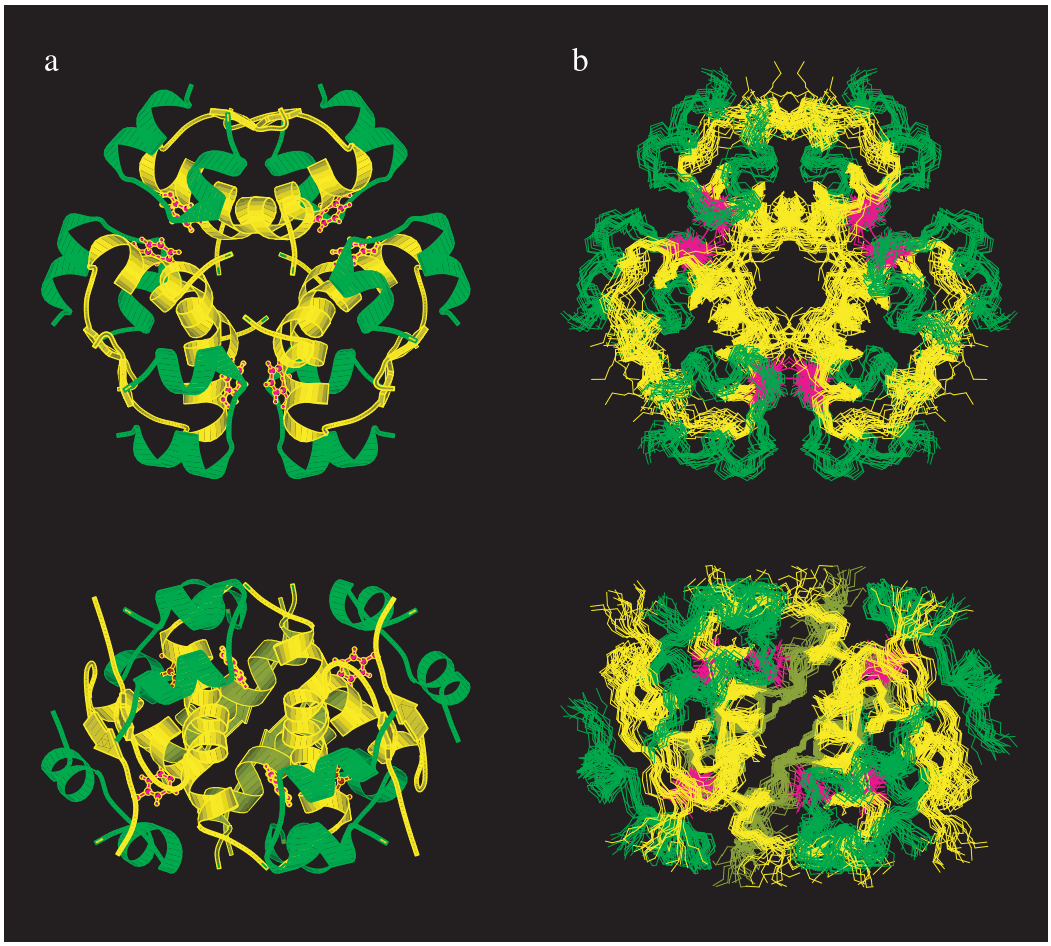


Figure 4. Solution structure of the R_6 insulin hexamer. (a) Ribbon diagram of the refined average structure. (b) Backbone plot of the final ensemble of 20 structures superimposed onto the refined average structure. In both (a) and (b), the top view is down the threefold axis; the bottom view is down the twofold axis (rotated by 90° with respect to the top view). The A-chain is shown in green, the B-chain in yellow, and the phenol molecule in magenta. In (a), the N-terminal residues are indicated by the reverse colouring scheme (yellow for the A-chain, green for the B-chain). The figure was generated using Molscript (Kraulis, 1991).

proton of Cys(A11); in two of the 20 structures, a hydrogen bond occurs between the hydroxyl proton of phenol and the carbonyl oxygen of Cys(A6). These results are in agreement with the previous solution structure (Chang et al., 1997).

Correlated motions in the ensemble

A newly developed principal component analysis (PCA) method (Amadei et al., 1993; Abseher et al., 1998) was applied to the 40 refined structures to detect correlated motions occurring in the ensemble. Two analyses were made: the first 'balanced' analysis was done using only the C_α atoms of the A- and B-chains, and the C_4 of the phenol molecule (Figure 5a); a second 'phenol-centric' analysis used all phenol atoms, but again only the C_α atoms of insulin (Figure 5b). Similar to our experience with other NMR ensembles, in both analyses about 75% of the diversity in the ensemble is accounted for by the first five eigenvectors (Figure 5, top). The contribution of each residue to these eigenvectors is shown in Figure 5 (bottom).

In both analyses, the first eigenvector (corresponding to the largest correlated motion in the ensemble) is dominated by fluctuations of the B-chain N-terminus (B1–B3). This terminus is quite distant from the phenol molecule. Also involved in the first eigenvector are a concerted motion along one face of the B-chain α-helix (peaks at B6, B9, B13, B16, and B20), and a motion of the C-terminal end of the first A-chain helix (A7–A9). Both of these structural elements are

Table 2. Structural statistics for the final ensemble

Constraint classification	
Distance classifications: uniform lower bounds of 2 Å; 13 classes of upper bounds from 2.4 to 5.0 Å.	
Number of constraints per monomer	
Total number of NOEs	722
Intraresidual NOEs	218
Sequential NOEs	174
Medium-range NOEs	198
Long-range intramonomer NOEs	75
Intermonomer NOEs[a]	49
Ambiguous NOEs[a]	8
Dihedral restraints	0
Hydrogen bonds	2
Disulfide bonds	3
Number of refined structures 20	
Residual NOE violations none > 0.3 Å	
Geometry ideality[b]	
Bonds	0.0037 ± 0.0001 Å
Angles	$0.60° \pm 0.02°$
Impropers	$0.47° \pm 0.04°$
Mean rms deviations from experimental restraints	
NOE	0.018 ± 0.002 Å
Comparison of calculated conformers rmsd relative to refined average structure:	
Backbone atoms (A2–A21, B4–B20, B22–B27[c])	0.96 ± 0.23 Å
All heavy atoms (A2–A21, B4-B20, B22–B27)	1.4 ± 0.2 Å
All backbone atoms	1.3 ± 0.3 Å
All heavy atoms	1.8 ± 0.3 Å

[a] Final values – see Table 1.
[b] rmsd from ideal covalent geometry defined in the OPLS/NMR force field (Linge and Nilges, 1998).
[c] These residues, plus their symmetry-relatives, constitute the well-ordered region. This region was defined by iteratively excluding residues with an rmsd greater than two standard deviations from the average.

in contact with the phenol ring (Figure 4b) which participates in the motion with one side (Figure 5b), but not with the other side (including the C_4 atom).

The second eigenvector in both analyses involves a concerted motion of the B-chain C-terminus (B29 and B30) together with the second A-chain helix (A13–A21), and essentially the same face of the B-chain helix as above (B3, B7, B9, B14, B16, and B21). The phenol molecule is even more involved in this motion than in the first eigenvector; it is in contact with A16 and with the B-chain helix face, as mentioned above, but is quite distant from the B-chain C-terminus.

In the third eigenvector of the balanced analysis (Figure 5a), there is a motion involving A14 and A18. These residues lie along the solvent-exposed face of the second A-chain helix. This pattern of a correlated motion along the solvent-exposed face of a helix was also found in other similar analyses of NMR ensembles (R. Abseher, unpublished data). The same motion occurs in eigenvectors 4 and 5 of the phenol-centric analysis (Figure 5b).

The most striking difference between Figures 5a and 5b is that the analysis using all phenol atoms finds an independent collective mode of phenol (eigenvector 3 in Figure 5b), with virtually no involvement of the protein.

The structures in the ensemble cluster into two distinct regions along the first eigenvector (Figure 6a), suggesting that this eigenvector corresponds to a transition of the hexamer between two major substates.

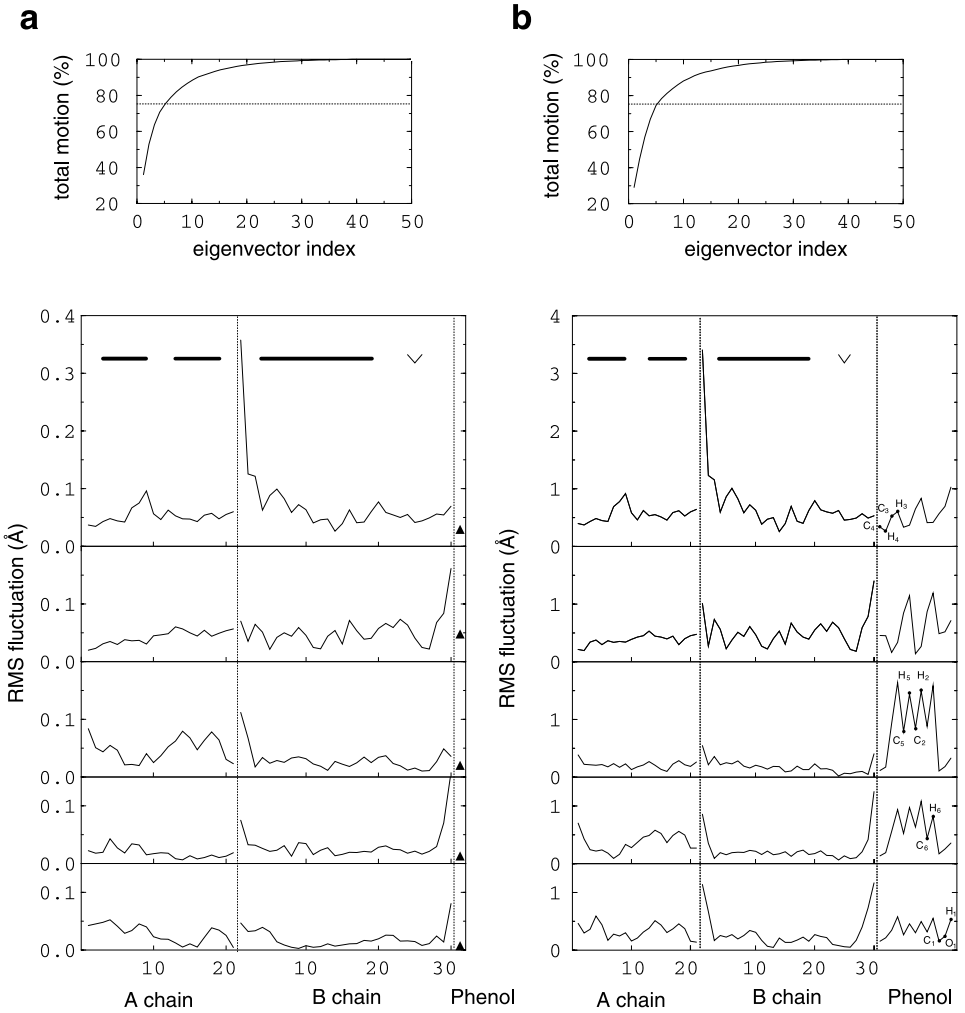


Figure 5. Principal component analysis of the correlated motions in the ensemble. Two analyses are shown: (a) a 'balanced' analysis using only the C_α atoms of the A- and B-chains, and the C_4 of the phenol molecule; and (b) a 'phenol-centric' analysis using all phenol atoms, but again only the C_α atoms of insulin. The top figures show the cumulative extent to which the motion in the ensemble is described by the most significant eigenvectors; in both analyses, the first five eigenvectors account for about 75% of the total motion (indicated by the dotted line). The lower figures show the residue rms fluctuations for each of the first five eigenvectors (from top to bottom). Secondary structure regions are indicated above the first eigenvector plot. In both analyses, the first eigenvector (corresponding to the largest correlated motion in the ensemble) involves the B-chain N-terminus (B1–B3), one face of the B-chain α-helix (peaks at B6, B9, B13, B16, and B20), and the end of the first A-chain helix (A7–A9). One side of the phenol ring is also involved. The phenol atom names follow the PDB naming convention.

The most obvious difference between the two substates is that, in the more populated 'green' substate (Figure 6d), the first three residues in the B-chain are in a helical state, extending the helix all the way to the N-terminal end of the B-chain, whereas in the second 'red' substate (Figure 6c) these residues form an extended structure that points in towards the threefold axis. Another difference is that the red substate is slightly more open along the twofold axis (Figure 6b).

A refined average structure was calculated for each of the two substates, using the corresponding clusters defined by the projection along the first eigenvector (Figure 6a). These average structures were calculated using structures from the original ensemble, not the projected structures presented in Figure 6. Analyzing these average structures, we found that the overall solvent exposed surface of the red substate is slightly increased compared with that of the green substate, particularly for the residues close to the phenol molecule. Also the phenol has a slightly increased exposure in the red substate. Compared with the red substate, the green substate is closer to the crystal structure ac-

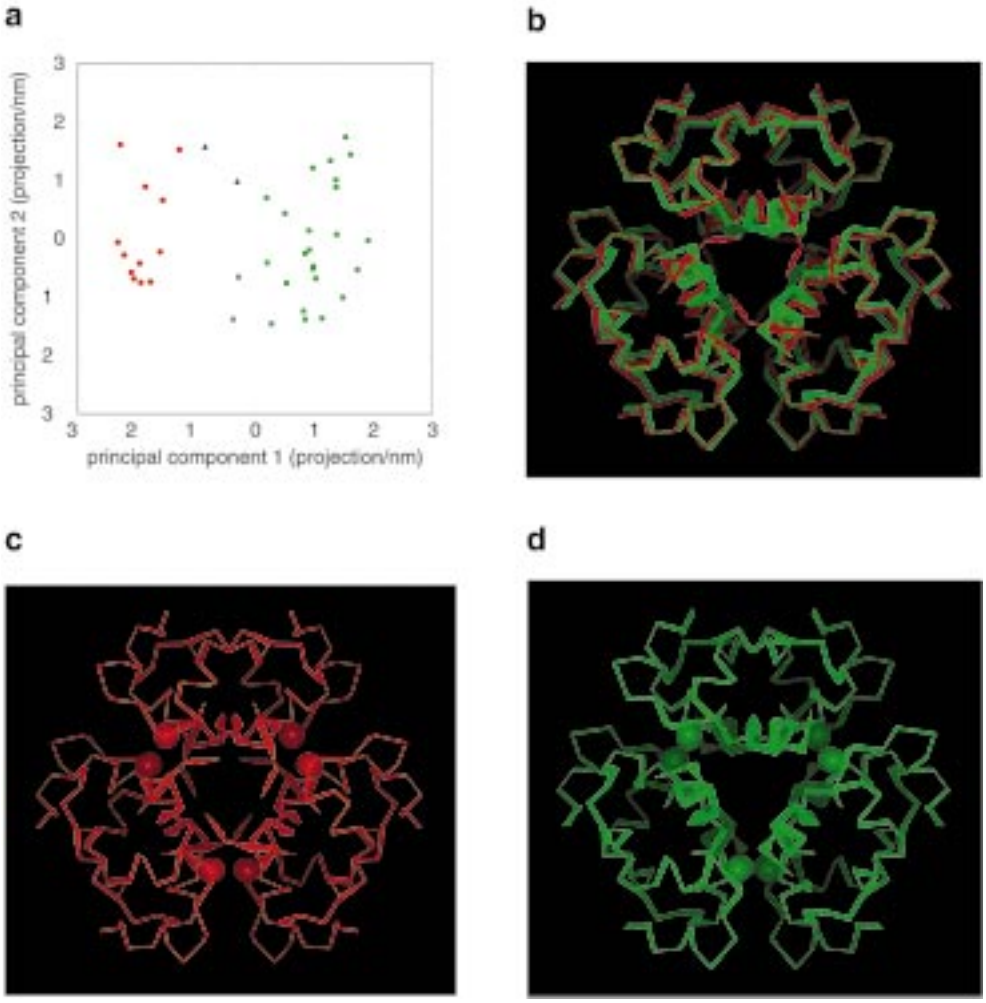


Figure 6. Two major substates of R_6 insulin. (a) Projection of the 40 refined structures onto the principal eigenplane, defined by the first two principal components. The structures cluster into two distinct regions along the first eigenvector, indicating that the hexamer occupies two major substates. The most populated substate is indicated with green squares, the least populated substate with red circles; black triangles are used to indicate points that are intermediate between the two substates. (b) Comparison of the two substates. The structures shown were calculated by projecting along eigenvector 1, hence, these structures are not members of the original ensemble; the projection removes any differences between the structures which are not correlated with the first eigenvector. Only C_α atoms are shown. In the green substate, there is a closing along the twofold axes compared to the red substate. (c) The least populated substate. The phenol atom used in the PCA is shown as a ball; otherwise as per (b). The B-chain N-terminus is in an extended state, pointing in towards the threefold axis. (d) The most populated substate; otherwise as per (c). The B-chain N-terminus forms a helix.

cording to the rms deviations (Figure 7). The refined average structure of all 40 refined structures is close to the green substate, whereas the refined average for the 20 structures of the final ensemble is closer to the red substate (Figure 7).

Discussion

Structure calculation

The results show that the ambiguous NOE data clearly distinguishes the correct point group (32). Indeed, this result is not surprising, since point groups 6 and 32 have very different kinds of intermonomer packing. For insulin, this point is academic anyway; it is well established that a hexamer of point group 32 forms under the conditions studied. However, the result is

significant since it demonstrates that point group ambiguity can be resolved by NMR. The fact that the distinction between the two possible point groups is indeed clear, in spite of the ambiguity of the NOE data, suggests that this method may also work well for other oligomers where the point group is ambiguous (i.e., tetramers, hexamers, octamers, etc.).

The NMR data are also sufficient to define the monomer orientation in the hexamer. This was already clear in the first iteration from the convergence of the lowest energy structures to essentially the same fold, close to the crystal structure. This conclusion is not weakened by the rather low convergence rate to low-energy, correctly folded structures; the low rate reflects the efficiency of the search method, rather than the completeness of the data. In fact, low initial convergence rates are typical in structure calculations where many ADRs are used, since these restraints introduce a large number of local minima into the energy landscape and, hence, make the global minimum search problem more difficult (Nilges and O'Donoghue, 1999).

R_6 insulin is the first symmetric hexamer, and the highest order symmetric oligomer solved by NMR to date. Therefore this calculation sets a new landmark, both for NMR structure determination and for the symmetry-ADR method. The success of this calculation bodes well for future attempts to determine the NMR structures of higher-order symmetric oligomers. Can we expect that the symmetry-ADR method will work as well for other symmetric oligomers; i.e., is the success of this calculation likely to be typical? In the R_6 insulin hexamer, many interfacial residues are located relatively far from a symmetry axis. For this kind of structure, we expect that the symmetry-ADR method will work well (Nilges, 1993; O'Donoghue and Nilges, 1999). On the other hand, difficulties with the method are expected in the case of structures where most of the interface between the monomers occurs close to a symmetry axis, as in the leucine zipper homodimers, or along the dimer–dimer interface in the p53 tetramerization domain. In such cases, X-filtered experiments to determine contacts between symmetry partners are very important. In general, however, most higher-order oligomers, like R_6 insulin, are larger and more globular than leucine zippers or p53, and hence, most interfacial residues are not close to a symmetry axis. Thus we anticipate that the symmetry-ADR method will perform well for many higher-order oligomers.

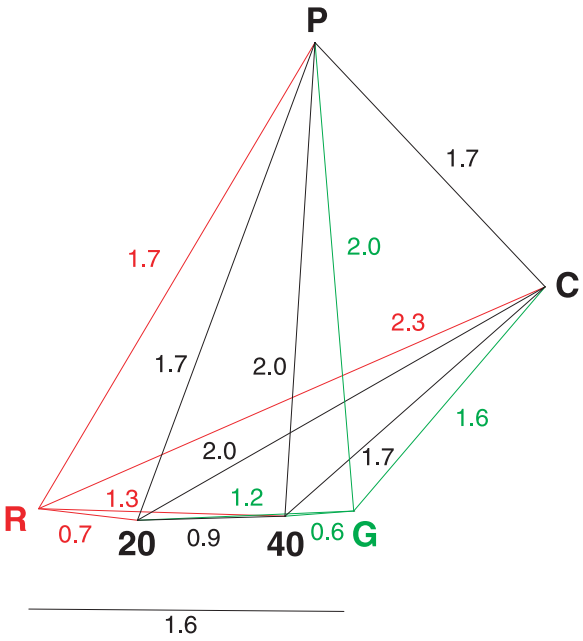


Figure 7. Comparison of different structures for R_6 insulin. The following structures are compared: C, crystal structure (1ZNJ); G, green substate (refined average structure); R, red substate (refined average structure); P, previous solution structure (1AIY); 20, refined average of the final ensemble of 20 structures; and 40, refined average of the 40 refined structures. The plot was generated by measuring backbone rms deviations (using all residues) between each pair of structures, indicated in the figure (in Å). This distance matrix was then embedded into three dimensions using the distance-geometry routine in X-PLOR. In the figure are plotted the obtained x and y Cartesian coordinates, which define the principal eigenplane projection. Of the two substates, the green substate is closer to the crystal structure. The two average structures, 20 and 40, lie almost exactly along the line joining the red and green substates. The previous solution structure (1AIY) is quite distinct. Since all intermonomer assignments made here were compatible with those using 1AIY, both calculations used very similar DRSs, and we would expect the structures to be similar. However, the current calculation differs from that of 1AIY in the following ways: distance symmetry restraints (were not used for 1AIY); different protocol; different force field; refinement in a water shell with full electrostatics.

Symmetry ambiguity of the hexamer NOEs can be partly resolved by the application of mixed-labeling experiments, if isotope labeled protein samples are available. However, in the present study, since we were already able to separate intra- and inter-monomer NOEs by reference to the monomer and dimer mutant structures, the use of mixed-labeling experiments would not bring any substantial new information to reduce the ambiguity problem. These experiments could only confirm the intramonomer/intermonomer separation; they cannot distinguish between different classes of intermonomer NOEs. Furthermore, these experi-

ments require the production of labeled insulin, the costs of which could not be justified given the small benefit that would be obtained. Thus the hexamer calculation was determined using only homonuclear spectra. Generally, however, in solving other symmetric oligomers, where related mutants of known structures are not available, labeling would be essential to complete the intramonomer assignment. Given such data, the oligomer structure could be calculated using the symmetry-ADR method as applied in the present study.

PCA method

The application of a new PCA method to study concerted motions in the structural ensemble is a novel feature of this study. Previously, the method had been applied only to NMR ensembles where molecular dynamics trajectory data were also available. However, each of the principal eigenmodes revealed here by the analysis of the insulin structure ensemble are physically meaningful and insightful. Thus, the motions of the residues that are involved in the first, second, and third eigenvectors (Figure 5a) all make sense in that they are either physically connected in space, or they are located along a single face of a helix. Also the concerted motion of the N-terminus of the B-chain and the phenol molecule has a clear physical interpretation, as discussed above. Finally, the mode in which phenol motion is completely decoupled from that of insulin (the third eigenvector in Figure 5b) also makes sense, as phenol is not covalently bound to insulin. Thus, the results we have obtained with insulin support our previous conclusion that the modes revealed from PCA of NMR ensembles are significantly correlated with real concerted motions of molecules in solution (Abseher et al., 1998), although the correlation will not be perfect in all cases. The use of the first two eigenmodes to define principal eigenplane projections (Figure 6a) gives a clear criterion for elucidating which substates dominated the ensemble, and we suggest that, in general, this approach may be a useful way to get an overview of the major substates that occur in an NMR structural ensemble.

Phenol exchange and flexibility of the R_6 hexamer

The R_6 hexamer structure obtained here is very similar to the structure reported previously, as all 49 intermonomer NOEs assigned by the ARIA filter are identical with those assigned previously by reference to the crystal structures (Chang et al., 1997). Therefore, as in the previous study, the largest difference between the solution and crystal structures is found in the N-terminal end of the B-chain, which is disordered in solution but part of the extended α-helix in the crystal phase (vide supra). However, unlike the previous solution structure, the ensemble obtained here is totally unbiased by the crystal structure and may, therefore, form a better basis for the analysis of the concerted motions in the solution structures using the PCA method.

The structures in the ensemble fall into two relatively distinct substates. We first noticed this by inspecting the ensemble, before the principal component analysis. The principal eigenplane projection (Figure 6a) gives a clear basis for deciding which structures belong to which substate. Do these two substates really occur in solution, or is one of the states an artefact arising from the remaining ambiguous NOEs? Using ADRs, it is always possible that the calculation converges to alternative structures, in addition to the correct structure. With the currently available experimental data, we cannot totally rule out this possibility. However, we note that the PCA method identifies only *correlated* motions in the ensemble, hence the observation of two substates suggests a *correlated* 'lack' of NOEs that would pin down one state. We further note that all cross peaks in the NOESY spectra were well defined, and were used in the structure calculation, i.e., there are no missing NOEs that could be attributed to spectral artefacts. Therefore, since both substates satisfy the NOE data equally well, the data are consistent with the existence of two substates in solution. If both substates do exist, the transition between them must be fast on the NMR time-scale, otherwise we would observe a splitting of resonances. The suggestion that there are two substates in solution is compatible with the observation of only one substate in the crystal structure, since most often, only one of a range of substates present in solution will crystallize.

The NMR data show that the insulin-bound phenol molecules are in fast exchange (Jacoby et al., 1996; Chang et al., 1997). This is remarkable considering the position of insulin-bound phenol molecules, well buried in the trimer interface of the R_6 hexamer (Figure 5). The exit and entry of the phenol molecule would, therefore, seem to require structural fluctuations in 'gatekeeper' residues (Jacoby et al., 1996). Also, in the absence of phenol the hexamer will revert to the T_6-state, where the B-chain helix only includes the region Gly(B8)-Cys(B19), while the region Phe(B1)-Cys(B7) forms an extended loop which runs from the center of the threefold axis and along

the twofold axis, resulting in a displacement of the Phe(B1) by more than 30 Å.

In the light of these experimental observations it is interesting to notice that the red substate deviates more from the crystal structure than the green substate, both in overall rmsd, and in the conformation of the B-chain N-terminus (vide supra). Also, in the red substate the phenol molecules are slightly more exposed to the solvent than in the green substate. Thus, phenol in the red substate may be more likely to exchange; it is possible that switching from the green to the red substate is part of the molecular rearrangement associated with phenol exchange. This suggestion is supported by the extended structure of the B-chain N-terminal found in the red substate but not in the green. Again this may indicate the beginning of a rearrangement of the structure towards the phenol-free T_6-state. It is of particular interest in this context that the transition between these two substates also involves a concerted motion in part of the phenol molecule (Figure 5b, eigenvector 1). It is also interesting that the extended conformation of the B-chain N-terminus in the red substate bears a marked resemblance to the last step in the MD simulation of the $T_6 \rightarrow R_6$ transition (Figure 5 in Schlitter et al. (1993)). All in all, these observations lead us to the suggestion that the red substate may be the first step in the transition back to the T-state, and may also be the state that makes possible the fast exchange of the phenol molecule.

Conclusions

The solution structure of the R_6 insulin hexamer has been derived from NMR data alone; the ambiguity of different intermonomer NOE types was solved using the symmetry-ADR method. This is the highest order symmetric oligomer structure solved to date by NMR. A second novel feature of this study is the use of principal component analysis to examine the correlated motions in the final structure ensemble. The analysis suggested that in solution the hexamer may switch between two major conformational substates. The analysis also suggests a new insight into the mechanism of phenol exchange and the $T_6 \leftrightarrow R_6$ transition. These results clearly demonstrate the value of principal component analysis as a tool for gaining insight into the motions reflected by the diversity in NMR ensembles.

Acknowledgements

The 750 MHz NOESY spectra were obtained at The Danish Instrument Centre for NMR Spectroscopy of Biological Macromolecules. The Copenhagen group was financially supported by the Danish Natural Science Research Council (J. No.'s 9400351, 9502759 and 9601648), the Danish Technical Research Council (J. No.'s 16-5028-1 and 9601137), the Ministry of Industry (J. No. 85886), Julie Damm's Studiefond, Direktør Ib Henriksens Fond, Novo Nordisk Fonden and Carlsbergfondet. The Heidelberg group was financially supported by the Deutsche Forschungsgemeinschaft (Grant No. Ni 499/1-1). R.A. acknowledges a Marie Curie Research Training Grant awarded by the European Commission (project number BIO4CT965088).

References

Abseher, R., Horstink, L., Hilbers, C.W. and Nilges, M. (1998) *Proteins*, **31**, 370–382.

Amadei, A., Linssen, A.B.M. and Berendsen, H.J.C. (1993) *Proteins*, **17**, 412–425.

Arrowsmith, C.H., Pachter, R., Altman, R.B., Iyer, S.B. and Jardetzky, O. (1991) *Biochemistry*, **29**, 6332–6341.

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. and Vijayan, N.M. (1988) *Phil. Trans. R. Soc. Ser. B*, **319**, 369–456.

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

Brader, M.L., Kaarsholm, N.C., Lee, R.W.-K. and Dunn, M.K. (1991) *Biochemistry*, **30**, 6636–6645.

Brünger, A. (1992) *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Chang, X., Jørgensen, A.M.M., Bardrum, P. and Led, J.J. (1997) *Biochemistry*, **36**, 9409–9422.

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

DeLano, W.L. and Brünger, A.T. (1994) *Proteins*, **20**, 105–123.

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

Hua, Q.X. and Weiss, M.A. (1991) *Biochemistry*, **30**, 5505–5515.

Jacoby, E., Hua, Q.X., Stern, A.S., Frank, B.H. and Weiss, M.A. (1996) *J. Mol. Biol.*, **258**, 136–157.

Jørgensen, A.M.M., Kristensen, S.M., Led, J.J. and Balschmidt, P. (1992) *J. Mol. Biol.*, **227**, 1146–1163.

Jørgensen, A.M.M., Olsen, H.B., Balschmidt, P. and Led, J.J. (1996) *J. Mol. Biol.*, **257**, 684–699.

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

Kraulis, P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.

Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–485.

Lee, W., Harvey, T.S., Yin, Y., Yau, P., Litchfield, D. and Arrowsmith, C.H., *Nat. Struct. Biol.*, **1**, 877–890.
Linge, J.P. and Nilges, M. (1998) *J. Biomol. NMR*, **13**, 51–59.
Nilges, M. (1993) *Proteins*, **17**, 297–309.
Nilges, M. (1995) *J. Mol. Biol.*, **245**, 645–660.
Nilges, M. and O'Donoghue, S.I. (1998) *Prog. NMR Spectrosc.*, **32**, 107–139.
Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **229**, 317–324.
Nilges, M., Gronenborn, A.M., Brünger, A.T. and Clore, G.M. (1988) *Protein Eng.*, **2**, 27–38.
Nilges, M., Macias, M.J., O'Donoghue, S.I. and Oschkinat, H. (1997) *J. Mol. Biol.*, **269**, 408–422.
O'Donoghue, S.I. (1998) Protocols for symmetric oligomers. *http://www.EMBL-Heidelberg.DE/nmr/nilges/NMRProtocols/SymmetricOligomers.html.*
O'Donoghue, S.I. and Nilges, M. (1999) In *Biological Magnetic Resonance Vol. 17: Structure Computation and Dynamics in Protein NMR* (Krishna, N.R. and Berliner, J., Eds), Plenum Press, New York, NY, pp. 131–161.
O'Donoghue, S.I., Junius, F.K. and King, G.F. (1993) *Protein Eng.*, **6**, 557–564.
O'Donoghue, S.I., King, G.F. and Nilges, M. (1996) *J. Biomol. NMR*, **8**, 193–206.
Schlitter, J., Engels, M., Krüger, P., Jacoby, E. and Wollmer, A. (1993) *Mol. Simul.*, **10**, 291–308.
Wollmer, A., Rannefeld, B., Johansen, B.R., Hejnaes, K.R., Balschmidt, P. and Johansen, F.B. (1987) *Biol. Chem.*, **368**, 903–911.