J Biomol NMR (2008) 40:55–64
DOI 10.1007/s10858-007-9206-2

ARTICLE

Structure of human insulin monomer in water/acetonitrile solution

Wojciech Bocian · Jerzy Sitkowski · Elżbieta Bednarek · Anna Tarnowska · Robert Kawęcki · Lech Kozerski

Received: 4 July 2007 / Accepted: 15 October 2007 / Published online: 27 November 2007

Abstract Here we present evidence that in water/acetonitrile solvent detailed structural and dynamic information can be obtained for important proteins that are naturally present as oligomers under native conditions. An NMR-derived human insulin monomer structure in H_2O/CD_3CN, 65/35 vol%, pH 3.6 is presented and compared with the available X-ray structure of a monomer that forms part of a hexamer (Acta Crystallogr. 2003 Sec. D59, 474) and with NMR structures in water and organic cosolvent. Detailed analysis using PFGSE NMR, temperature-dependent NMR, dilution experiments and CSI proves that the structure is monomeric in the concentration and temperature ranges 0.1–3 mM and 10–30°C, respectively. The presence of long-range interstrand NOEs, as found in the crystal structure of the monomer, provides the evidence for conservation of the tertiary structure. Starting from structures calculated by the program CYANA, two different molecular dynamics simulated annealing refinement protocols were applied, either using the program AMBER in vacuum (AMBER_VC), or including a generalized Born solvent model (AMBER_GB).

Keywords Insulin monomer · MD in explicit solvent · NMR structure · Water/acetonitrile solvent

Introduction

Depending on conditions, insulin exists in different association states, from monomer to hexamer, but its biological function, i.e. control of glucose metabolism (Cahill 1971), is attributed to the monomeric structure (Le Roith and Zick 2001) upon binding to the insulin receptor (Luo et al. 1999). A drawback in diabetes treatment with insulin, encountered in preparation and stability of formulations (Lougheed et al. 1983), is also associated with monomer instability and its propensity to form dimers, (Darrington and Anderson 1995; Pocker and Biswas 1981; Vis et al. 1998) higher aggregates, and fibrils.

An insulin monomeric structure in solution and the means to study its aggregation has been a challenge of recent decades. A number of physicochemical techniques have been utilized to characterize the state of insulin aggregation, namely: cryo-electron microscopy (Jimenez et al. 2002) for visualizing the amyloid fibrils, nano-flow ES-MS (Nettleton et al. 2000) for detection of oligomeric states, small angle X-ray scattering (SAXS) (Uversky et al. 2003) for study of aggregation , and spectrophotometric monitoring (by UV/VIS at 350 nm) of turbidity changes (Kwon et al. 2001) to study aggregation at the water-organic solvent interface. Pulsed-Field Gradient NMR spectroscopy has recently been applied (Lin and Larive 1995). UV-CD spectroscopy has been established as a routine tool for inspecting the

Wojciech Bocian contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-007-9206-2) contains supplementary material, which is available to authorized users.

W. Bocian · J. Sitkowski · E. Bednarek · L. Kozerski
National Medicines Institute, Chełmska 30/34,
Warsaw 00-725, Poland

W. Bocian · J. Sitkowski · A. Tarnowska · R. Kawęcki · L. Kozerski (✉)
Institute of Organic Chemistry Polish Academy of Sciences,
Kasprzaka 44, Warsaw 01-224, Poland
e-mail: lkoz@icho.edu.pl

secondary structure motifs of proteins in solution (Rodger and Norden 1997).

Two approaches have been discussed and developed experimentally to achieve a monomeric insulin structure in solution, namely; engineering of the C-terminus of the B chain, or disrupting the hydrophobic dimer interface; (Chiti et al. 2003; Ciszak et al. 1995; Brems et al. 1992; Keller et al. 2001; Ludvigsen et al. 1994; Olsen et al. 1996; Jørgensen et al. 1992, 1996) and addition to water of an organic cosolvent, which is expected to disrupt quaternary structure without affecting the tertiary structure of the protein (Buck 1998). Acetic acid (20%, v/v in water) was used, and the structure of insulin was established (Hua et al. 1995). TFE-d_3 (35% in water) was also applied to derive the structure for a human insulin mutant (Jørgensen et al. 1992). Proton chemical shifts were cited for human insulin dissolved in H_2O/CD_3CN, 65/35 vol%, but except of ^1H chemical shift assignments, details of the structure and analysis of aggregation behavior of insulin in this solvent were not established (Kline and Justice 1990).

In this account we present the NMR-derived human insulin structure in H_2O/CD_3CN, 65/35 vol%, pH 3.6 and compare it with the available X-ray structures and NMR structures in solution (Hua et al. 1995). Detailed analysis using PFGSE NMR, DNMR, dilution experiments and CSI prove the existence of monomeric structures in the concentration range 0.1–3 mM. The presence of long-range interstrand NOEs provides evidence for the conservation of tertiary structure. Therefore the results suggest that this solvent is a suitable medium for studying the native conformation of a protein.

Experimental

Insulin preparation

The human insulin (USP standard) was purchased from Sigma and stored at −20°C.

To remove trace amounts of large soluble aggregates which may influence the diffusion coefficient D_i the zinc-free insulin was filtered on a 25 × 1 cm Sephadex G-25 column. The column was equilibrated with a 1% acetic acid solution. The solution applied to the column was approximately 30 mg/ml in 1% acetic acid. The filtrate was collected by eluting with the same eluent and the fraction was collected which had an absorbance at 276 nm. Before lyophilisation the filtrate was kept overnight in a fridge over Bio-Rad Chelex 100 resin. The concentration was measured in UV using solution taken from NMR tube. A fraction of the solution was used which was obtained after lyophilisation of the column filtrate, dissolving in water/acetonitrile and pH adjustment using NaOH/HCl. The pH electrode was used.

NMR sample preparation

The preparation of the NMR samples was conducted by the same procedure as described by Hua et al. (1995) except that different solvent was used. The solvent used in the present work is H_2O/D_2O:CD_3CN (65/35 vol%). Solutions of at least 1.7 mM concentration, pH 3.6, adjusted by adding aliquots of HCl and NaOH, were used for NMR measurements. Two samples of the same concentration were measured, in H_2O and D_2O. The concentrations of samples were measured by UV. The chemical shifts were measured vs. internal CHD_2CN (2.04 ppm vs. TSPA; 3-trimethylsilylpropionic acid, sodium salt).

NMR Spectra

The following spectra were recorded at 303 K with standard parameters of acquisition and processing used for proteins: NOESY (Jeener et al. 1979) using the States-TPPI method (Bodenhausen et al. 1984; States et al. 1982) (200 ms mixing time and a 2 s delay before each scan). We are aware of a fact that relatively long mixing time can lead to observation of some indirect effects which lower the precision of the structure determination. But, for a small molecule as an insulin (5.8 kD), with fairly narrow lines (1.5 Hz of His aromatic signals) this may not be a serious problem. To minimize zero quantum coherence peaks a small random variation in the mixing time between transients and between t_1 increments was used. (Neuhaus and Williamson 1995); TOCSY (Braunschweiler and Ernst 1983; Griesinger et al. 1988), (mixing times for TOCSY spectra were 90 ms with a spin-lock field of 8 kHz); GHSQC (Bodenhausen and Reuben 1980; Summers et al. 1986); E-COSY (Griesinger et al. 1985, 1987); DQF-COSY (Piantini et al. 1982).

Complete assignments of the ^1H and ^{13}C, ^{15}N resonances of human insulin in 35% CD_3CN, pH 3.6, were achieved using SPARKY (Goddard and Kneller 2004) to analyze 2D NOESY, TOCSY, and $^1H/^{13}C$-, $^1H/^{15}N$-HSQC spectra.

Pulsed Field Gradient Spin Echo (PFGSE) experiments were performed in D_2O according to the following conditions: 24 spectra were acquired using the BPPSTE (Wider et al. 1994; Wu et al. 1995) (stimulated echo sequence incorporating bipolar gradients) sequence. The gradient strengths were incremented as a square dependence in the range from 1 to 50 G/cm. The diffusion time (Δ) and the duration of magnetic field gradients (δ) were 300 ms and 2 ms respectively. The data were processed using Varian DOSY (Johnson 1999) and DECRA (Windig and Antałek 1997; Antałek and Windig 1996) packages.

Amide proton exchange experiments. The experiment was run at 10°C on two independent samples. The reference sample for acquiring the NH integrals at t = 0 was run in H_2O/35% CD_3CN (Fig. 3, left spectrum). The other sample, of comparable concentration, was dissolved in D_2O/35% CD_3CN, pH adjusted and placed in a precooled probe, and acquisition was started 30 min after dissolving and equilibration. The NOESYs, lasting ca.1 hr, were checked after 30,105,180, 255, 650, 1,220, 2,550 and 4,800 min following dissolving. The integrals in both samples were calibrated against a nonexchanging reference signal. The comparison of the integrations between samples in H_2O/35% CD_3CN and in D_2O/35% CD_3CN, after 1 h, served to identify slow exchanging NH resonances using a criterion that half of the initial crosspeak intensity still remained in the latter sample versus the initial integral at t = 0.

Calculation procedures

The 1251 NOE cross-peaks, classified into three classes: backbone-551, sidechain-415 and methyl-285, were manually assigned and converted into upper distance constraints with the *CALIBA* function of CYANA (Güntert et al. 1997). The redundant and meaningless distance restraints were removed and distance limits with diastereotopic groups were adjusted, and pseudo atoms were inserted, if no stereospecific assignment was available. This gave 685 restraints. Angle restraints were generated from 36 ^{3}J(N**H**-C**H**$^{\alpha}$) coupling constants using the Karplus function in CYANA, and additionally allowed ranges of dihedral angles φ, ψ, were obtained by grid search in the function *habas* (CYANA). A total of 250 structures of insulin with random geometry were generated by CYANA and refined iteratively using the simulated annealing protocol (30,000 total steps) of the torsion angle dynamics (version 1.0.6). The 200 structures with lowest CYANA target function values were used for further MD refinement in AMBER 9. (Case et al. 2006) In AMBER calculations additionally to the 685 distance restraints and 72 torsion φ were used 9 disulfide, 51 chirality and 50 trans-ω restraints.

MD refinement

Two different refinement protocols were applied using molecular dynamics simulated annealing: in vacuum and including a generalized Born solvent model (Xia et al. 2002; Onufriev et al. 2000).

For the MD refinement in vacuum (VC), the 200 CYANA structures were first energy minimized in vacuum for 500 steps without any restraints, and then two cycles of 15 ps MD simulated annealing in vacuum were performed by *SANDER* (from AMBER): 0–1,000 steps heating the system from 10 to 1,100 K, 1,001–3,000 steps at 1100 K, 3,000–15,000 steps cooling to 0 K, 0–3,000 steps tight coupling for heating and equilibration (TAUTP = 0.2), 3,001–11,000 steps of slow cooling (TAUTP = 4.0–2.0), 11,000–13,000 steps of faster cooling (TAUTP = 1.0) and 13,000–15,000 steps of fast cooling, like a minimization (TAUTP = 0.5–0.05). In steps 0–3,000 the restraints were slowly increased from 10% to 100% of their final values. The cutoff value was set to 15 Å and DIELC (dielectric constants) was equal to 1.

For the generalized Born (GB) refinement 500 steps energy minimization with the GB model were first performed on the CYANA structures, followed by two cycles of 15 ps simulated annealing with GB and the same protocol as for vacuum. For further analysis from 200 refined structures, both in vacuum and GB, 100 structures were selected with the lowest violation energy and, from these, the 50 structures with the lowest total AMBER energy were chosen as structure representatives.

Results and discussion

Establishing of a monomer structure in solution

A number of physicochemical methods has been utilised to study the aggregation of insulin, and nanospray ESI-MS and PFGSE (*P*ulsed *F*ield *G*radient *S*pin *E*cho) NMR seem to be most promising in providing complementary information on association. The latter method is based on monitoring the translational diffusion coefficients, D_i, of species in solution and is therefore specific for differentiation of various aggregates characterised by different hydrodynamic diameter (Danielsson et al. 2002). During the present study, using PFGSE NMR, we have exploited the fact that on dilution, formerly aggregated insulin will tend to dissociate. This process is monitored by the diffusion coefficient, D_i, attaining a maximum value for the monomer which has a larger diffusion rate than dimer, tetramer, hexamer or higher aggregates. Therefore the dilution experiment is run until a constant value of D_i is attained. Furthermore, using PFGSE NMR one can lift the chemical shift degeneracy between the monomer and dimer which have identical chemical shifts but very different translational diffusion coefficient.

In Figure 1, dilution of human insulin in D_2O is presented. As expected, the diffusion coefficient increases with dilution until a concentration of ca. 40 μM is reached. This concentration was established earlier in the classic work of Kadima et al. (1992) as characteristic for monomer. For sensitivity reasons it was not possible to measure

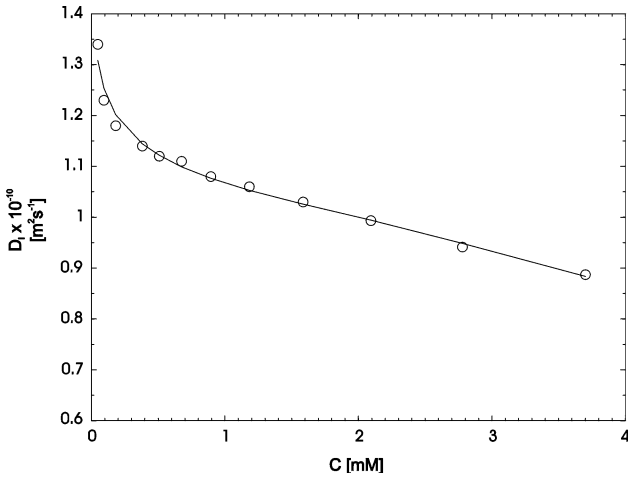

Fig. 1 Concentration depenence of the diffusion coefficient $D_i \times 10^{-10}$ m^2 $s^{-1}$of the Zn^{2+}-free human insulin in D_2O solution, pH 3.6

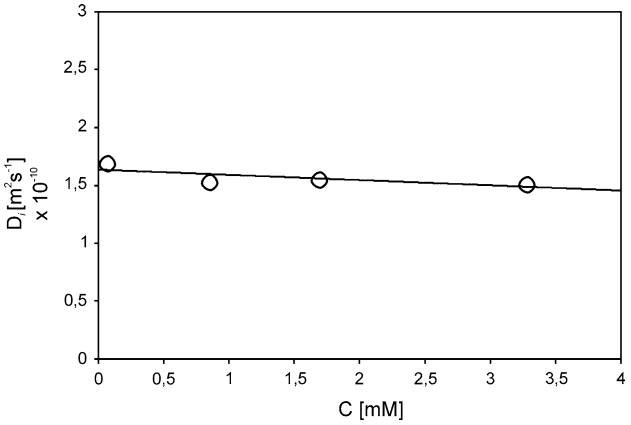

Fig. 2 Concentration dependence of diffusion coefficient for zinc-free human insulin in D_2O/CD_3CN (65/35 v/v %), 30°C, pH 3.6

the diffusion coefficient at higher dilution. Current studies indicate that at this concentration, using the established values of monomer/dimer association constant, $K_{12} \cong 1.0 \times 10^4 M^{-1}$ in gas phase (Nettleton et al. 2000) or 7×10^5 M^{-1} (Kadima et al. 1993) in solution, there is still appreciable amount of dimer present, ca. (50 %). Therefore the largest value of D_i measured, 1.32×10^{-10} $m^2 s^{-1}$, can not be treated as characteristic of Zn-free human insulin monomer in D_2O at 30°C, pH 3.6 in these conditions.

By contrast, if monomeric insulin is present in solution, the diffusion coefficient, D_i, should remain constant on dilution. A small, linear decrease of D_i with increasing concentration is expected because of friction, as a result of crowding of monomeric species (Price et al. 1999). Such a situation is observed in Fig. 2 which evidences the existence of monomeric insulin in D_2O/CD_3CN (65/35 v/v %), pH 3.6 over a wide concentration range. This makes the applied solvent well suited for NMR study, all the more that the ^1H NMR spectrum has sharp lines and shows dispersion of NH chemical shifts characteristic for structured insulin (Fig. 1S). The value of D_i in D_2O/CD_3CN solution is higher for the monomer, 1.63×10^{-10} $m^2 s^{-1}$, than in pure water, 1.53×10^{-10} $m^2 s^{-1}$ (authors' results submitted for publication, dealing with insulin oligomers composition in pure water, monitored by PFGSE). This is accounted by the fact that dilution of water, D_i (25°C) 23.0×10^{-10} m^2 s^{-1}, with acetonitrile, D_i 25°C) 43.7×10^{-10} m^2 s^{-1}, makes the resultant solvent less viscous, thereby facilitating faster translational diffusion of a solute. The correction factor for the change of solvent from D_2O to D_2O/CD_3CN (65:35 vol%) was established in this work 1.07 as measured on the reference crown ether compound which does not aggregate in water. On the other hand, diminished solvation of insulin by water molecules is also expected to influence the diffusion coefficient by affecting its hydrodynamic radius.

We have also measured temperature dependence of the CD spectrum of human insulin in the studied solvent which shows some changes over ca. 30°C. (Fig. 2S). Since various phenomena may account for the observed changes, i.e; loss of secondary structure with temperature, specific solvation of tyrosines in a studied solvent or conformational changes of the tyrosine chromophores, it is less conclusive in assigning the oligomeric state of the insulin.

Establishing secondary structure motifs for human insulin in water/ acetonitrile solvent

The chemical shift index (CSI) has become a useful tool in establishing secondary structure motifs in proteins (Wishart et al. 1992). This requires, however, a knowledge of precise random coil chemical shifts of the amino acids in a given solvent. We have therefore measured the chemical shifts of selected amino acids, which appear in the B chain C-terminus, in the hexapeptide environment proposed by Wishart et al. (1995), H-Gly-Gly-X-Ala-Gly-Gly-OH, to check if a change of the solvent used by Wishart will influence the chemical shifts in water or water/acetonitrile solutions. The results for proton and carbon atoms of Lys, Thr and Phe embedded in hexapeptide in CSI analysis are given in Tables 1S and 2S (Supplementary Information), respectively.

The conclusion from this study is that in both solvents the chemical shifts of H^{α} differ negligibly from values cited by Wishart et al. and therefore these increments can be used in water/acetonitrile and water solution for establishing secondary structure motifs providing the protein is monomeric. This finding of a present work may have a general application to secondary structure motifs assignment in protein studies in cases where tertiary structure can not be solved.

More importantly, both ^{1}H and ^{13}C chemical shifts of protons α,β,γ, δ and ε in lysine residue of a hexapeptide H-Gly-Gly-Lys-Ala-Gly-Gly-OH do not change on going from water to water/acetonitrile solution at the same pH value, 3.6 (comp. data for Lys in Tables 1S and 2S with B29 in Table 7S). While in principle the neat acetonitrile is considered as an excellent hydrogen bond acceptor, this result seems to indicate that binary mixture (water/acetonitrile 65:35 vol. %) largely resembles pure water, this conclusion finds support in a recent vibrational energy relaxation study and hydrogen bonding in water-acetonitrile binary mixtures (Cringus et al. 2004).

In Figs. 3S and 4S the CSI analysis for H^{α} protons in chains A and B in two different sample preparations of human insulin is presented. The CSI graphs for C^{α} carbon atoms in both chains are given in Fig 5S and Fig 6S for the same solutions.

Amide proton exchange and hydrogen bonds

Figure 3 shows the spectra of amide proton resonances before exchange and 10 h after exchange started in D_2O/CD_3CN at 10°C. It is readily seen that substantial amounts of NH resonances still persist. Using criteria described in the Experimental section, fast exchanging NH's were identified. In Table 4S some examples of exchange rates are shown for crosspeaks which did not overlap and were unambiguously assigned. The rates of exchange are compared to the values cited by Keller et al. for engineered insulin in 35% TFE (Keller et al. 2001). It can be seen that the rates found in water/acetonitrile solution for slowly exchanged NH's are similar to those observed in 35% TFE. However, as our results are acquired at 5°C lower temperature, the exchange rates are in fact faster than in TFE at comparable temperature. This makes the water/acetonitrile solvent more like the physiological medium in which the exchange rates are at least order of magnitude larger than observed here and by Keller et al. (2001). The comparison above serves as the evidence that water accessibility to exchanging protons in the interior of the protein is larger in water/acetonitrile than in 35% TFE/water solution.

On the other hand, as will be shown later, all long range interactions found in the crystal structure are also found in NOESY spectra in the studied solution, which strengthens our confidence that the organic cosolvent is not disrupting the tertiary structure present in the native protein conformation.

Tertiary structure calculation

Graphs shown in Fig. 11S present the distribution of restraints used in calculations.

Table 6S gives the listing of the long range, /i–j/ > 5, NOE restraints. These restraints are the most important ones in tertiary structure refinement as they define the hydrophobic core of a protein and the long-range interactions. Their volumes were used as restraints. All of the crucial long-range interactions seen in the monomer from the crystal structure were found in the NOESY spectra, which is considered a necessary precondition for the statement that the water/acetonitrile solvent does not destroy the tertiary structure present in the native conformation.

Furthermore, we have checked in the spectra used to acquire the NOESY data for the presence of the crosspeaks which were considered by Led et al. as the intermolecular ones characterizing the dimer structure (Jørgensen et al. 1992). The following crosspeaks were not found in the spectra; Tyr(B16)GB-Tyr(B26)QD, Gly(B8)HA-Tyr(B16)QE, Val(B12)HA-Tyr(B16)QB, Gly(B23)HA-Tyr(B26)QD, Gly(B23)HA-Tyr(B26)HA, Phe(B24)H-Thr(B26)HA. This is another indication confirming the monomeric structure of the studied insulin with acetonitrile as cosolvent.

Table 5S comprises the data for $^{3}J(H\text{-}N\text{-}C^{\alpha}\text{-}H)$ used to construct torsion restraints, φ, in the CYANA refinement of

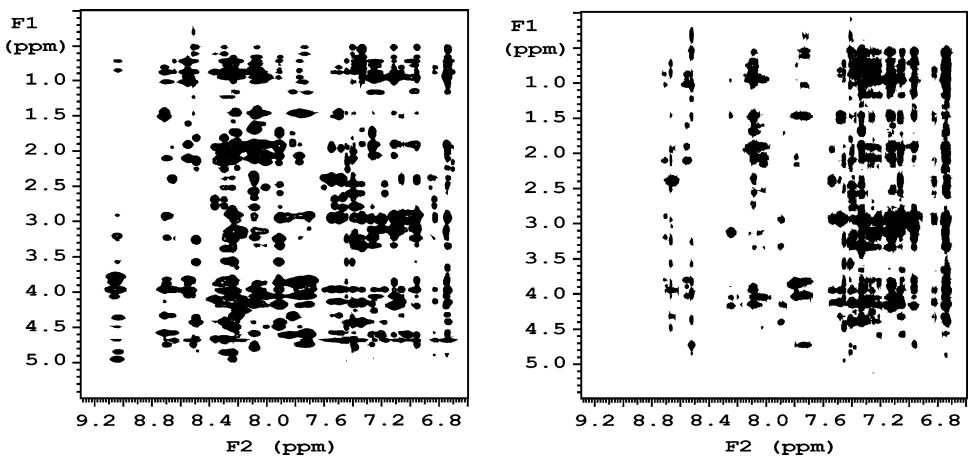

Fig. 3 Amide proton part of a NOESY spectrum, left 0 hr, right 10 hrs, after dissolving the solute in D_2O/CD_3CN (65%/35% v/v), pH 3.6, 10°C

the structure. They were established from NOESY and DQF-COSY spectra. These restraints were used in AMBER refinement methods. Along with our data the values of ^{3}J(H-N-C$^\alpha$-H) vicinal couplings in other insulins studied in water are cited. Inspection of the data shows that vicinal coupling constants established in water/acetonitrile solution are very much the same as the ones in pure water. This is yet another confirmation of the validity of this solvent in structure elucidations.

Starting from structures calculated by the program CYANA, two different refinement protocols used molecular dynamics simulated annealing with the program AMBER; in vacuum (AMBER_VC), and including a generalized Born solvent model (AMBER_GB). Using the available monomer model from X-ray as a starting structure, molecular dynamics in explicit water solvent were also carried out without restraints.

Structure examination of human insulin in 35% CD_3CN / water solution

The calculated structures contain all secondary structure motifs predicted from the CSI analysis, which are confirmed by the pattern of sequential NOEs (Fig.11S) and ^{3}J(HN,H$^\alpha$) coupling constants (Table 5S). From A2 to A8 the sequential NOEs imply a helical structure. In particular, the presence of $d_{\alpha N}$(i, i + 4) together with small ^{3}J(HN,H$^\alpha$) indicate α-helix. In cysteines C6 and C7, this coupling constant is much higher, characteristic for an extended strand. This is also observed for other insulins studied in pure water, (Ludvigsen et al. 1994; Olsen et al. 1996) or with organic cosolvent (Keller et al. 2001). The A9-A12 loop is well characterized by sequential NOEs due to narrow resonances in this part of the structure, except A11CH$^\alpha$ which is rather broad. The structure from A13-A18 is a helix, with a nearly complete set of $d_{\alpha\beta}$(i, i + 3) and slowly exchanging amide protons. The reminder of the A chain forms an extended structure. The B chain character is dominated by the B9-B19 helix. This is followed by a β-turn B20-B23 and random coil B24-B30. Random coil is also observed in the N-terminus of the B chain i.e; B1-B9.

Statistics shown in Table 1 present comparisons of the data acquired for insulins in H_2O/CD_3CN and in H_2O/CD_3COOH solutions. The ensembles were compared to their mean or to the X-ray model.

The torsion restraints were not violated more than an equivalent distance of 0.1 Å. The introduction of side chain ^{3}J(H$^\alpha$,H$^{\beta1,\beta2}$) coupling constants (Table 7S) did not improve the quality of the structure and therefore they were not used in the final run.

The most significant improvement in restraints violation and rmsd was achieved with the AMBER _GB protocol. With the exclusion of ill-defined N- and C-terminal residues (Gly1A, Asn21A; Phe1B, Val2B, Thr30B) the atomic rmsd deviations for 50 conformers were 0.484 and 0.703 for the backbone atoms, from the mean and X-ray, respectively.

The Ramachandran statistics (Fig 10S) (Laskowski et al. 1996) indicate a good quality model has been refined by AMBER _GB. A significant improvement was seen in the percentage of dihedral angles in the most favoured regions of φ, ψ space with a negligible number of structures having torsions in disallowed regions. These are mainly the ultimate or penultimate residues at the N or C termini.

In addition to comparison, in Table 1, with insulin structure established in H_2O/CD_3COOH , Table 9S presents a comparison of selected statistics for the insulins studied in different solvents and computed using various calculation protocols. It is seen in Table 1 that the 50 structure ensemble—(H_2O/CD_3CN) has the same rmsd from the mean as the 10 structure ensemble—(H_2O/CD_3COOH). The important difference is seen in comparison with the X-ray structure. The ensemble studied in H_2O/CD_3CN is much closer to the X-ray structure than the one studied in H_2O/CD_3COOH. This may reflect the fact which we emphasize in this study, i.e. the H_2O/CD_3CN solvent is preserving the native structure more adequately than other water /organic cosolvent media. The binary mixture used in this study largely resembles the pure water. A further confirmation to this is given by the low rmsd from X-Ray for restraints observed for this solvent as compared with the insulins shown in Table 9S, and in particular, in H_2O/CD_3COOH.

Figure 4 presents the comparison of the two structures established in indicated media. For the sake of comparison 10 structures ensemble was selected from 50 structures ensemble from AMBER_GB using criterion of lowest energy and minimum violations energy. In both cases the comparison with X-ray is also shown. It is clearly seen that structures in H_2O/CD_3CN and in H_2O/CD_3COOH differ substantially. There are four main backbone domains of a structure in water/acetic acid which deviate from X-ray, i.e. B chain, N and C-termini, B-β-turn and helix I in A chain. Most significant is the deviation from X-ray structure of the B chain C-terminus which is the interface for a dimer formation. In particular the side chains of B25, B26 of X-ray structure do not fit to any of the conformations in ensemble. The same is observed for A14 side chain. The conformation of the B chain N-terminus also differs largely from the X-ray conformation. There is also seen that B chain terminus is significantly detached from hydrophobic core, mainly due to B26 side chain conformation not found in the X-ray structure. All these deviations are not observed in a structure established in water/acetonitrile. Furthermore, as seen for the B25 and β-turn of a structure in water/

Table 1 Structural statistics of the human insulin structure ensembles computed using various refinement methods. Comparison with structure established in H_2O/CD_3COOH solution

Experimental NOE's[d]	AMBER_VC[a] 1251(551 backbone, 415 side-chain, 285 methyl)	AMBER_GB[b] 1251(551 backbone, 415 side-chain, 285 methyl)	XPLOR[c]
Experimental restraints[e]			
Total inter-proton	685	685	304
Intra-residue	124	124	
Sequential ($\|j-i\| < 2$)	240	240	134
Medium range ($2 < \|j-i\| < 5$)	201	201	92
Long – range ($\|j-i\| > 5$)	110	110	78
Disulfide restraints	9	9	
Chirality restraints	51	51	
Trans-(ω) restraints	50	50	
Torsion restrains (ϕ)	72	72	28ϕ, 15χ^1
Number of distance restraint violations in calculated structures[f]			
Total [max.violation; in no. of struct.]	972 [2.49; 3]	360 [2.40; 3]	–
Intra-residue	349 (0.13 ± 0.03)	167 (0.14 ± 0.04)	–
Short range	533 (0.15 ± 0.05)	133 (0.12 ± 0.03	–
Long range	90 (0.11 ± 0.01)	60 (0.15 ± 0.04)	–
RMSD from mean structure[g]			
All atoms in ensemble	2.01 (1.41)	1.83 (1.27)	1.69 (1.09)
Backbone heavy atoms 153 (CO,CA,N)	1.19 (0.67)	0.99 (0.48)	1.08 (0.49)
Long range distances[h]	0.97 (110)	0.89 (110)	0.73 (78)
RMSD from X-Ray structure[j]			
All atoms in ensemble	2.98 (2.04)	2.53 (1.81)	4.01 (3.51)
Backbone heavy atoms 153 (CO,CA,N)	1.88 (0.89)	1.31 (0.70)	1.92 (1.19)
Long range distances[h]	1.43	1.37	2.14
Ramachandran statistics (% residues included in)[k]			
Most favored regions	87.6	92.4	61.9
Additionally allowed regions	10.3	6.9	31.4
Generously allowed regions	1.5	0.2	6.3
Disallowed regions[m]	0.6	0.5	0.5

[a] In vacuum

[b] H_2O/CD_3CN (65/35 vol.%), in explicit solvent—50 structures ensemble, structure deposited in PDB-**2jv1**

[c] H_2O/CD_3COOH (80/20 vol.%)—10 structures ensemble (Hua et al. 1995) PDB-**2hiu**

[d] The 1521 manually assigned NOE cross-peaks

[e] Transposed by CYANA to distance restraints (**2jv1**)

[f] This work, the 680 restraints considered in 200 models; AMBER statistics with threshold 0.1 Å; values in parentheses show average violation, Å, in total number of violated structures

[g] This work, 50 low energy structures are compared. Values in parentheses refer to statistics with excluded chain ends in both chains, i.e; for residues 2–20 and 3–29 in chain A and B, respectively

[h] Defines the rmsd for the distance between the atoms used in long range restraints

[j] PDB-**1mso** (Smith et al. 2003)

[k] Respective values in X-Ray structure (PDB-**1mso**) are 94.2, 5.8, 0.0, 0.0

[m] This work, Phe(B1), Val(B2), Asn(B3), Tyr(B26), Pro(B28), Lys(B29)

acetonitrile, the ensemble of computed structures contain two different conformations of backbone in these regions, one of which is consistent with the X-ray conformation. This is not so in the ensemble in H_2O/CD_3COOH, there is only one conformation in computed structures, but different from the X-Ray. Basing on the above differences one can suggest that the structure in water/acetonitrile is closer to the X-ray than the other one and therefore there is higher

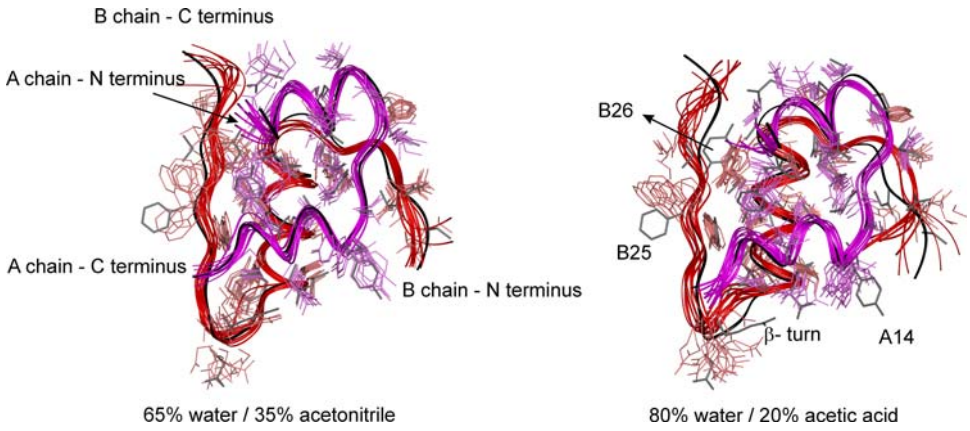

Fig. 4 Left -overlay of 10 AMBER_GB refined lowest energy structures (present work); A chain in magenta, B chain in red, the black string represents the X-Ray structure backbone; right-overlay of 10 structures of Hua (Hua et al. 1995) in the same colors. The side chains were removed for clarity from B1, B28–30, A1, A21 in both cases

probability of finding the native conformation of insulin using experimental restraints acquired in this solvent rather than in water/acetic acid solution.

The regions of increased intramolecular dynamic mobility, the biological relevance of the structure

Figure 5 shows the rmsd of the backbone atoms from the mean of the ensemble and from the X-ray backbone. Excluding the N- and C-termini of both chains it shows higher rmsd, which may possibly imply increased mobility of the B chain C-terminus B24-B28, an observation consistent with the random coil structure found in this part of the backbone. Using the same argument, also the β-turn is clearly visible (Gly 20–Arg23) as a motif of increased mobility. This is an especially important observation confirming the discussed model of insulin interaction with receptor assuming the B-chain terminus (Gly20–Thr 30) detachment from the hydrophobic core of the molecule as a first step in receptor binding (Keller et al. 2001).

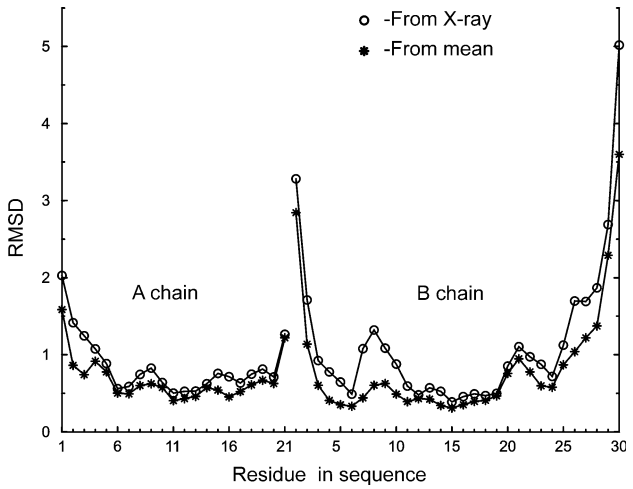

Fig. 5 The backbone rmsd (Å) between the ensemble of 50 models in AMBER_GB and mean structure, and between this ensemble and X-Ray structure

This mobility is also seen in Fig. 4 showing around Gly20 the spread in backbone overlayed structures (left structure). The N-terminus of the B-chain is also flexible as expected for the random coil motif.

The higher rmsd at the A chain N-terminus may also imply increased flexibility, again in agreement with prediction by the cited model of receptor binding, showing a propensity for helix unwinding (Keller et al. 2001). Also remarkable is the similarity of the graphs in Fig. 5, which suggests that the mean NMR structure is very close to the X-ray backbone structure. This is a fact further proving the usefulness of water/acetonitrile in protein structure elucidations, especially in situations (as found for insulins) in which intensive aggregation renders structure elucidations in water difficult or impossible.

Concluding remarks

The aim of the present work was to assess the validity of water /acetonitrile as a proper solvent to conduct structural studies of proteins extensively aggregating in water. It was demonstrated that the quaternary structure is effectively eliminated in this solvent while the native tertiary structure, as modeled by the monomer embedded in the X-ray hexamer structure, is preserved. Aggregation problems can thus be avoided, intermolecular NOEs are not observed, and narrow lines secure the extraction of a complete set of sequential assignments. Structural studies of insulins in a wide range of milimolar concentrations in solution can therefore be successfully conducted by applying water/actonitrile as a medium.

The present structure has been carefully compared with the recently published X-ray crystal structure of the zinc-bound hexamer of the native protein at 1 Å resolution using the coordinates of a single monomer (Smith et al. 2003). The structures of monomer that forms part of a

hexamer in the X-ray structure and in water/acetonitrile solution overlay very well both in terms of backbone and side chains.

Starting from structures calculated by the program CYANA, the two different refinement protocols used molecular dynamics simulated annealing with the program AMBER in vacuum (AMBER_VC), and including a generalized Born solvent model (AMBER_GB).

Acknowledgement A. T. gratefully acknowledges the financial support during this work by means of Ministry of Science and Higher Education grant no. 0278/P01/2006/30.

References

Antałek B, Windig W (1996) Generalized rank annihilation method applied to a single multicomponent pulsed gradient spin echo NMR data set. J Am Chem Soc 118:10331–10332

Bodenhausen G, Kogler H, Ernst RR (1984) Selection of coherence-transfer pathways in NMR pulse experiments. J Magn Reson 58:370–388

Bodenhausen G, Reuben DJ (1980) Structural studies on a-conotoxin SI. Chem Phys Lett 69:185–189

Braunschweiler L, Ernst RR (1983) Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. J Magn Reson 53:521–528

Brems DN, Alter LA, Beckage MJ, Chance RE, DiMarchi RD, Green LK, Long HB, Pekar AH, Shields JE, Frank BH (1992) Altering the association properties of insulin by amino acid replacement Protein Eng 5:527–533

Buck M. (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. Q Rev Biophys 31:297–355

Cahill GF Jr (1971) The Banting Memorial Lecture 1971. Physiology of insulin in man. Diabetes 20:785–799

Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA (2006) AMBER 9. San Francisco, University of California

Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature 424:805–808

Ciszak E, Beals JM, Frank BH, Baker JC, Carter ND, Smith GD (1995) Role of C-terminal B-chain residues in insulin assembly: the structure of hexameric LysB28ProB29-human insulin. Structure 3:615–622

Cringus D, Yeremenko S, Pshenichnikov MS, Wiersma DA (2004) Hydrogen bonding and vibrational energy relaxation in water–acetonitrile mixtures. J. Phys. Chem B 108:10376–10387

Danielsson J, Jarvet J, Damberg P, Gräslund A (2002) Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Aβ(1-40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length. Mag Reson Chem 40:S89–S97

Darrington RT, Anderson BD (1995) Effects of insulin concentration and self-association on the partitioning of its A-21 cyclic anhydride intermediate to desamido insulin and covalent dimer. Pharm Res 12:1077–1084

Goddard TD, Kneller DN (2004) SPARKY 3. San Francisco, University of California

Griesinger C, Otting G, Wüthrich K, Ernst RR (1988) Clean TOCSY for proton spin system identification in macromolecules. J Am Chem Soc 110:7870–7872

Griesinger C, Søerensen OW, Ernst RR (1985) Two-dimensional correlation of connected NMR transitions. J Am Chem Soc 107:6394–6396

Griesinger C, Sørensen OW, Ernst RR (1987) Practical aspects of the E.COSY technique. Measurement of scalar spin-spin coupling constants in peptides. J Magn Reson 75:474–492

Güntert P, Mumenthaler C, Wüthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 273:283–298

Hua QX, Gozani SN, Chance RE, Hoffmann JA, Frank BH, Weiss MA (1995) Structure of a protein in a kinetic trap. Nat Struct Biol 2:129–138

Jeener J, Meier BH, Bachmann P, Ernst RR (1979) Investigation of Exchange Processes by 2-Dimensional NMR-Spectroscopy. J Chem Phys 71:4546–4553

Jimenez JL, Nettleton EJ, Bouchard M, Robinson CV, Dobson CM, Saibil HR (2002) The protofilament structure of insulin amyloid fibrils. Proc Natl Acad Sci USA 99:9196–9201

Johnson CS (1999) Diffusion ordered NMR spectroscopy: principles and applications. Prog Nucl Magn Reson 34:203

Jørgensen AM, Kristensen SM, Led JJ, Balschmidt P (1992) Three-dimensional solution structure of an insulin dimer. A study of the B9(Asp) mutant of human insulin using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. J Mol Biol 227:1146–1163

Jørgensen AM, Olsen HB, Balschmidt P, Led JJ (1996) Solution structure of the superactive monomeric des-[Phe(B25)] human insulin mutant: elucidation of the structural basis for the monomerization of des-[Phe(B25)] insulin and the dimerization of native insulin. J Mol Biol 257:684–699

Kadima W, Øgendal L, Bauer R, Kaarsholm N, Brodersen K, Hansen JF, Porting P (1993) The influence of ionic strength and pH on the aggregation properties of zinc-free insulin studied by static and dynamic laser light scattering. Biopolymers 33:1643–1657

Kadima W, Roy M, Lee RW, Kaarsholm NC, Dunn MF (1992) Studies of the association and conformational properties of metal-free insulin in alkaline sodium chloride solutions by one- and two-dimensional 1H NMR. J Biol Chem 267:8963–8970

Keller D, Clausen R, Josefsen K, Led JJ (2001) Flexibility and bioactivity of insulin: an NMR investigation of the solution structure and folding of an unusually flexible human insulin mutant with increased biological activity. Biochemistry 40:10732–10740

Kline AD, Justice RM Jr (1990) Complete sequence-specific 1H NMR assignments for human insulin. Biochemistry 29:2906–2913

Kwon YM, Baudys M, Knutson K, Kim SW (2001) In situ study of insulin aggregation induced by water-organic solvent interface. Pharm Res 18:1754–1759

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486

Le Roith D, Zick Y (2001) Recent advances in our understanding of insulin action and insulin resistance. Diabetes Care 24:588–597

Lin M, Larive CK (1995) Detection of insulin aggregates with pulsed-field gradient nuclear magnetic resonance spectroscopy. Anal Biochem 229:214–220

Lougheed WD, Albisser AM, Martindale HM, Chow JC, Clement JR (1983) Physical stability of insulin formulations. Diabetes 32:424–432

Ludvigsen S, Roy M, Thogersen H, Kaarsholm NC (1994) High-resolution structure of an engineered biologically potent insulin

monomer, B16 Tyr–>His, as determined by nuclear magnetic resonance spectroscopy. Biochemistry 33:7998–8006

Luo RZ, Beniac DR, Fernandes A, Yip CC, Ottensmeyer FP (1999) Quaternary structure of the insulin-insulin receptor complex. Science 285:1077–1080

Nettleton EJ, Tito P, Sunde M, Bouchard M, Dobson CM, Robinson CV (2000) Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. Biophys J 79:1053–1065

Neuhaus D, Williamson MP (1995) Nuclear overhauser effect in conformational and structural analysis. A guide for Chemists. Wiley-VCH, NY

Olsen HB, Ludvigsen S, Kaarsholm NC (1996) Solution structure of an engineered insulin monomer at neutral pH. Biochemistry 35:8836–8845

Onufriev A, Bashford D, Case DA (2000) Modification of the Generalized Born Model Suitable for Macromolecules. J Phys Chem B 104:3712–3720

Piantini U, Sorensen OW, Ernst RR (1982) Multiple quantum filters for elucidating NMR coupling networks. J Am Chem Soc 104:6800–6801

Pocker Y, Biswas SB (1981) Self-association of insulin and the role of hydrophobic bonding: a thermodynamic model of insulin dimerization. Biochemistry 20:4354–4361

Price WS, Tsuchiya F, Suzuki C, Arata Y (1999) Characterization of the solution properties of Pichia Farinosa killer toxin using PGSE NMR diffusion measurements. J Biomol NMR 13: 113–117

Rodger A, Norden B (1997) Circular dichroism and linear dichroism. Oxford University Press

Smith GD, Pangborn WA, Blessing RH (2003) The structure of T6 human insulin at 1.0 A resolution. Acta Crystallogr D Biol Crystallogr 59:474–482

States DJ, Haberkorn RA, Ruben DJ (1982) A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. J Magn Reson 48:286–292

Summers MF, Marzilli LG, Bax A (1986) Complete proton and carbon-13 assignments of coenzyme B12 through the use of new two-dimensional NMR experiments. J Am Chem Soc 108: 4285–4294

Uversky VN, Garriques LN, Millett IS, Frokjaer S, Brange J, Doniach S, Fink AL (2003) Prediction of the association state of insulin using spectral parameters. J Pharm Sci 92:847–858

Vis H, Heinemann U, Dobson CM, Robinson CV (1998) Detection of a monomeric intermediate associated with dimerization of protein HU by mass spectrometry. J Am Chem Soc 120: 6427–6428

Wider G, Døtsch V, Wüthrich K (1994) Self-compensating pulsed magnetic-field gradients for short recovery times. J Magn Reson Series A 108:255–258

Windig W, Antałek B (1997) Direct exponential curve resolution algorithm (DECRA): A novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles. Chemom Intell Lab Syst 37:241–254

Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. J Biomol NMR 5:67–81

Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31:1647–1651

Wu DH, Chen AD, Johnson CS (1995) An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. J Magn Reson Series A 115:260–264

Xia B, Tsui V, Case DA, Dyson HJ, Wright PE (2002) Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. J Biomol NMR 22:317–331