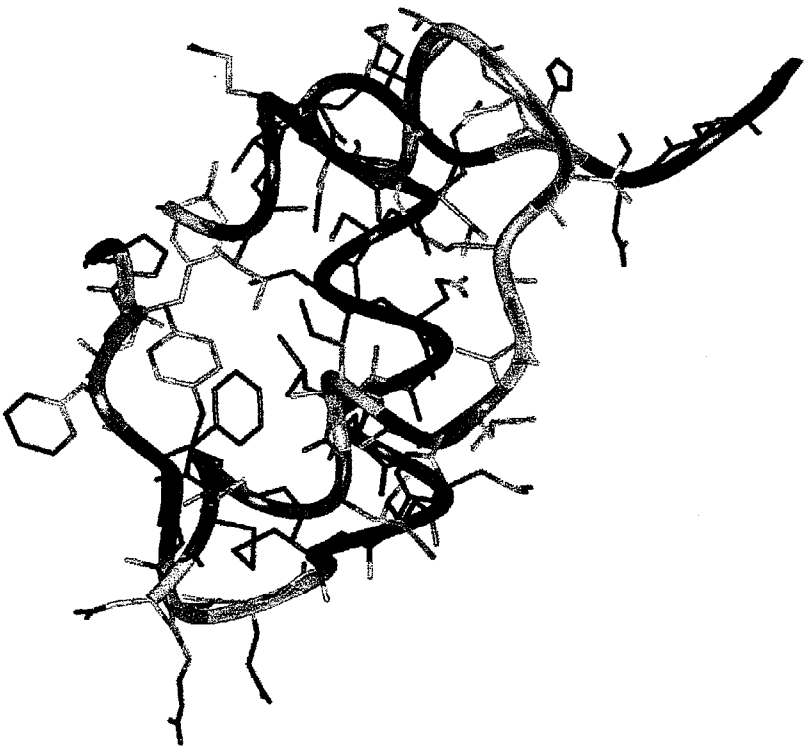
Investigations of Structure and Dynamics of Insulin Mutants using NMR Spectroscopy

Birk Olsen, Helle

Publication date:
1996

Investigations of Structure and Dynamics of Insulin Mutants using NMR Spectroscopy



Helle Birk Olsen

Roskilde University,

Novo Nordisk A/S & ATV

february 1996

Investigations of Structure and Dynamics of Insulin Mutants using NMR Spectroscopy

- The three dimensional solution structure of the biological active (B1, B10, B16, B27) Glu, des-B30 mutant of human insulin at neutral pH.
- Backbone dynamics of the (B1, B10, B16, B27) Glu, des-B30 and B16 His, des-B30 insulin mutants.
- Preliminary structure investigations of two insulin mutants with reduced and increased, respectively, biological activity.

This thesis is submitted in partial fulfilment of the requirements for the degree of Ph.D. at the Department of Life Sciences and Chemistry, Roskilde University.

The project was a collaboration between Novo Nordisk A/S and Roskilde University as part of the Industrial Research Programme administered by ATV (Danish Academy of Technical Sciences), EF 454.

Acknowledgments

The project presented in this thesis was a collaboration between Novo Nordisk A/S, Roskilde University and ATV (Danish Academy of Technical Sciences) in the Industrial Research Programme (EF454). The work was done in the period october 1992 to february 1996 interrupted by a maternity leave.

I would like to thank the staff of the Insulin Research Department and the director Dr. techn. Jan Markussen who gave me the oppurtunity to work on this project. Besides the indispensable and continued efforts of my supervisors at Novo Nordisk A/S Lic. scient. Svend Ludvigsen and Lic. techn. Niels C. Kaarsholm as well as Professor, Lic. scient. Poul Erik Hansen, RUC, I would like to acknowledge the work of several people. Ane Blom for careful preparation of samples, assistance with operation and maintenance of the NMR spectrometer, encouragements and never failing high spirits. Ivan Diers, Svend Havelund, Anne-Marie Kolstrup, and Lene G. Andersen for fermentation and purification of mutant insulins, and Susan E. Danielsen for CD measurements. Jean Wittingham for permitting me to reproduce her figure (Figure 2.2). Last but certainly not least my family for their patience and encouragements all the way through, my husband Peter Fæster Nielsen for 'graphical advices' and proof reading.

february 13, 1996

Helle Birk Olsen

Table of contents

Abbreviations

CD, circular dichroism.

DPI, des-(B26-B30) pentapeptide human insulin.

DQF-COSY, double quantum filtered two-dimensional correlated spectroscopy.

FID, free induction decay.

HI, human insulin.

HMQC, hetero-nuclear multiple-quantum coherence.

NMR, nuclear magnetic resonance.

NOE, nuclear Overhauser enhancement.

NOESY, two-dimensional nuclear Overhauser enhanced spectroscopy.

TOCSY, two-dimensional total correlation spectroscopy.

rms, root mean square.

1. Introduction

The insulin dependent diabetes mellitus (IDDM) is characterized by a reduced or completely absent production of endogenous insulin. Patients suffering from this chronic disease are treated with daily subcutaneous injections of insulin preparations to keep their blood glucose level under control. The efficiency of this insulin therapy to imitate a normal insulin level in the blood is of vital importance to the diabetics in ensuring their daily health and suppressing the development of late complications. At Novo Nordisk A/S ongoing research aims at optimizing this therapy as to mimick the 'normal' timing and profile of insulin in the blood.

An attractive longterm goal is, however, to replace the injection therapy by insulin-mimetica to be administered peroral. Thus, in recent years much effort has been put into characterizing the insulin-binding regions of the insulin receptor and into elucidating the relationship between the spatial structure of the insulin molecule and its affinity for the insulin receptor. The three dimensional structure of insulin in the hexameric units of insulin crystals has been intensively investigated by X-ray crystallography during the recent decades (Adams et al., 1969, Smith et al., 1984, Derewenda et al., 1989), whereas the structure of the insulin receptor with or without insulin bound is unknown. Until now, structural determinations of insulin species in aqueous solution by NMR spectroscopy of the biological active form of the molecule, the monomer, has been conducted at acidic pH or in the presence of organic co-solvents (Knegtel et al.; 1991, Hua et al., 1991, 1992a, 1993b; Ludvigsen et al., 1994; Jørgensen et al., 1996). A prerequisite for detailed investigations of the structure/receptor affinity relationship of insulin is to determine the three dimensional structure of the monomeric insulin species in aqueous solution at physiological pH. Despite its small size (51 amino acids, 5808 Dalton) a complicated pattern of aggregation and precipitation properties of insulin hampers the structural investigations of the native species in the neutral pH range. Recently, the access to genetical engineering has facilitated the production of insulin species with mutations of specific residues and it has proved possible to manipulate the aggregation properties and the biological activity independently.

The introduction of two dimensional ^{1}H experiments in the early 80's enabled the use of NMR for determination of the three dimensional structure of proteins in solution. Using the information from a series of 2D experiments it became possible to assign the spectra of proteins (up to approximately 100 residues) and to extract the information about inter-proton distances from NOEs that is the input to structure calculations. The development of 3D and 4D experiments correlating protons with ^{15}N and ^{13}C nuclei together with the ability to enrich the proteins with these NMR active isotopes of nitrogen and carbon pushed the limits to even larger proteins (up to approximately 300 residues). The ^{15}N and ^{13}C enrichments also opened the possibility to investigate the dynamics of the backbone and the sidechains on the picosecond-nanoseconds timescale. Compared to structure determination in crystals by X-ray crystallography the NMR method is still more time-

consuming and outnumbered, however the information concerning the dynamics on different time scales that is obtainable from the NMR data is extremely valuable and not available from X-ray crystallography.

The work presented in this thesis is centered on the determination of the three dimensional structure of a biological active mutant of human insulin at neutral pH by NMR spectroscopy. The backbone dynamics of this mutant and of a mutant of which the structure is known at acidic pH are examined and compared. Finally the sequential assignment and secondary structural elements of a pair of insulin mutants with manipulated biological activity is presented.

Chapter 2 of this thesis describes the determination by NMR of the three dimensional structure of the (B1, B10, B16, B27) Glu, des-B30 mutant of human insulin in aqueous solution at pH 6.5. The mutant is monomeric in the neutral pH range at the millimolar concentration needed for structural investigations by NMR. The biological activity of the mutant is 47 % of that of the native species. The work has been submitted for publication in Biochemistry in the form enclosed in Appendix A. Besides the addition of the Figures 2.1 and 2.2 in Chapter 2 and insertion of Figure 4 of the appendix into Figure 2.6 in Chapter 2 the contents of Chapter 2 and Appendix A are identical.

Chapter 3 presents the examination of the backbone dynamics of the (B1, B10, B16, B27) Glu, des-B30 mutant and the B16 His des-B30 mutant of which the structure has been published at acidic pH (Ludvigsen et al., 1994). Using samples uniformly enriched with the NMR active isotope of Nitrogen, ^{15}N, the relaxation parameters of the backbone ^{15}N nuclei are determined and translated to a generalized order parameter. This order parameter describes the spatial restriction of the picosecond-nanosecond motions of the $^{15}N^{1}H$-vector of each residue. Millisecond timescale dynamics are evaluated as well. The contents of this Chapter is under preparation for publication.

Chapter 4 describes the preliminary structural investigations of two insulin mutants with manipulated biological activity achieved by point mutations of the (B1, B10, B16, B27) Glu, des-B30 mutant. Both mutants are monomeric in the millimolar range and at neutral pH. The biological activity of the A3 Gly (B1, B10, B16, B27) Glu, des-B30 mutant is reduced to 0.1 % compared to the native species whereas that of the A8 His (B1, B10, B16, B27) Glu, des-B30 mutant is increased to 143%.

The efforts of engineering, expressing , fermenting and purifying the insulin mutants. were completed before the initiation of my work. My contributions have been to record and assign the spectra as well as to calculate the three dimensional structure of the (B1, B10, B16, B27) Glu, des B30 mutant. Furthermore the extraction of ^{15}N relaxation parameters and development of the software necessary to translate these to the parameters describing the backbone dynamics is my work as well. In parallel to the work described in the present thesis, I have been involved in the development of new heteronuclear pulse-sequences as published in Olsen et al., 1993a,b. in the same period.

2. Structure of an Engineered Insulin Monomer at Neutral pH

2.1 Abstract

Insulin circulates in the bloodstream and binds to its specific cell-surface receptor as a 5808 Dalton monomeric species. However, studies of the monomer structure and dynamics in solution are severely limited by insulin self-association into dimers and higher oligomers. In the present work we use site-directed mutagenesis of the dimer- and hexamer-forming surfaces to yield the first insulin species amenable for structure determination at neutral pH by nuclear magnetic resonance (NMR) spectroscopy. The preferred insulin mutant, i.e., (B1, B10, B16, B27) Glu, des-B30 insulin retains 47% biological potency, and remains monomeric at millimolar concentrations in aqueous solution at pH 6.5-7.5 as judged by NMR and near-UV circular dichroism (CD) spectroscopy. From a series of 2D ^{1}H-NMR spectra collected at pH 6.5 and 34°C, the majority of the resonances are assigned to specific residues in the sequence, and nuclear Overhauser enhancement (NOE) cross-peaks are identified. NOE-derived distance restraints in conjunction with torsion restraints based on measured coupling constants, $^3J_{H^N H^\alpha}$, are used for structure calculations using the hybrid method of distance geometry and simulated annealing. The calculated structures show that the major part of the insulin mutant is structurally well-defined with an average root mean square (rms) deviation between the 25 calculated structures and the mean coordinates of 0.66 Å for backbone atoms (A2-A19 and B4-B26) and 1.31 Å for all backbone atoms. The A-chain consists of two antiparallel helices, A2-A7 and A12-A19, connected by a loop. The B-chain contains a loop region (B1-B8), an α-helix (B9-B19), a type I turn (B20-B23), and terminates as an extended strand (B24-B29). The B1-B4 and B27-B29 regions are disorded in solution. The structure is generally similar to crystal structures and resembles a crystalline T-state more than an R-state in the sense that the B-chain helix is confined to residues B9-B19.

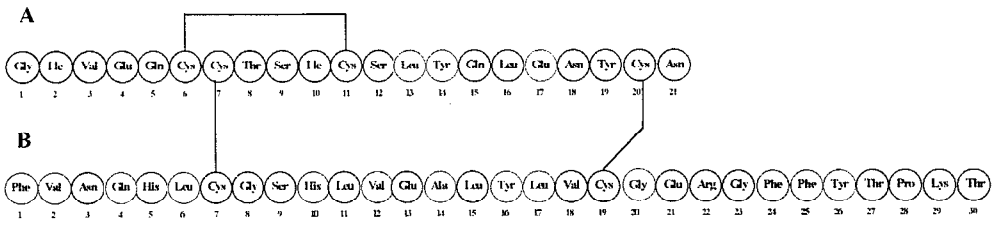


Figure 2.1 The primary structure of human insulin. Layout adapted from Brange et al., (1990). The lines indicate disulfide bridges. The colored residues are located in the aggregation surfaces. Green residues are involved in dimer formation, and blue residues are involved in the assembly of dimers into hexamers.

2.2 Introduction

Insulin is central to the hormonal control of metabolism. Due to its importance as a pharmaceutical preparation for the treatment of diabetes mellitus, much effort has been directed towards understanding the structural basis for insulin bioactivity. The protein is composed of two polypeptide chains, the A-chain (21 residues) and the B-chain (30 residues). The two chains are covalently linked by disulfide bridges at A7-B7 and A20-B19, and an intrachain disulfide bridge is joining A6 and A11. The primary structure of insulin is shown in Figure 2.1.

The three dimensional structure of insulin has been characterized in detail by X-ray analysis of aggregated species, notably zinc insulin hexamers (Adams et al., 1969; Smith et al., 1984; Derewenda et al., 1989). The three principal hexamer conformations have been designated the T_6, T_3R_3, and R_6 forms, respectively (Kaarsholm et al., 1989). In each of these structures, the A-chain folds into a helix-loop-helix motif with helical stretches located in A2-A8 and A12-A19. The B-chain can assume two distinct conformations. In the T-state, the B-chain contains an extended N-terminal arm, central α-helix (B9-B19), type I turn (B20-B23), and C-terminal β-strand. In the R-state, residues B1-B9 take up a helical conformation to form a region of α-helix contiguous from B1-B19. Figure 2.2 gives a stereographic representation of the three hexamer conformations (Wittingham et al., 1995) 2-zinc (T_6), 4-zinc ($T_3R_3^f$) pig insulin (the R^f form is similar to the R form, but with residues B1-B3 disordered) and rhombohedral phenol insulin (R_6). The inter-conversion between the T_6, T_3R_3, and R_6 states of the insulin hexamer is modulated by a set of homotropic and heterotropic ligand binding interactions and has been shown to take place in solution (Kaarsholm et al., 1989; Brader et al., 1991; Bloom et al., 1995).

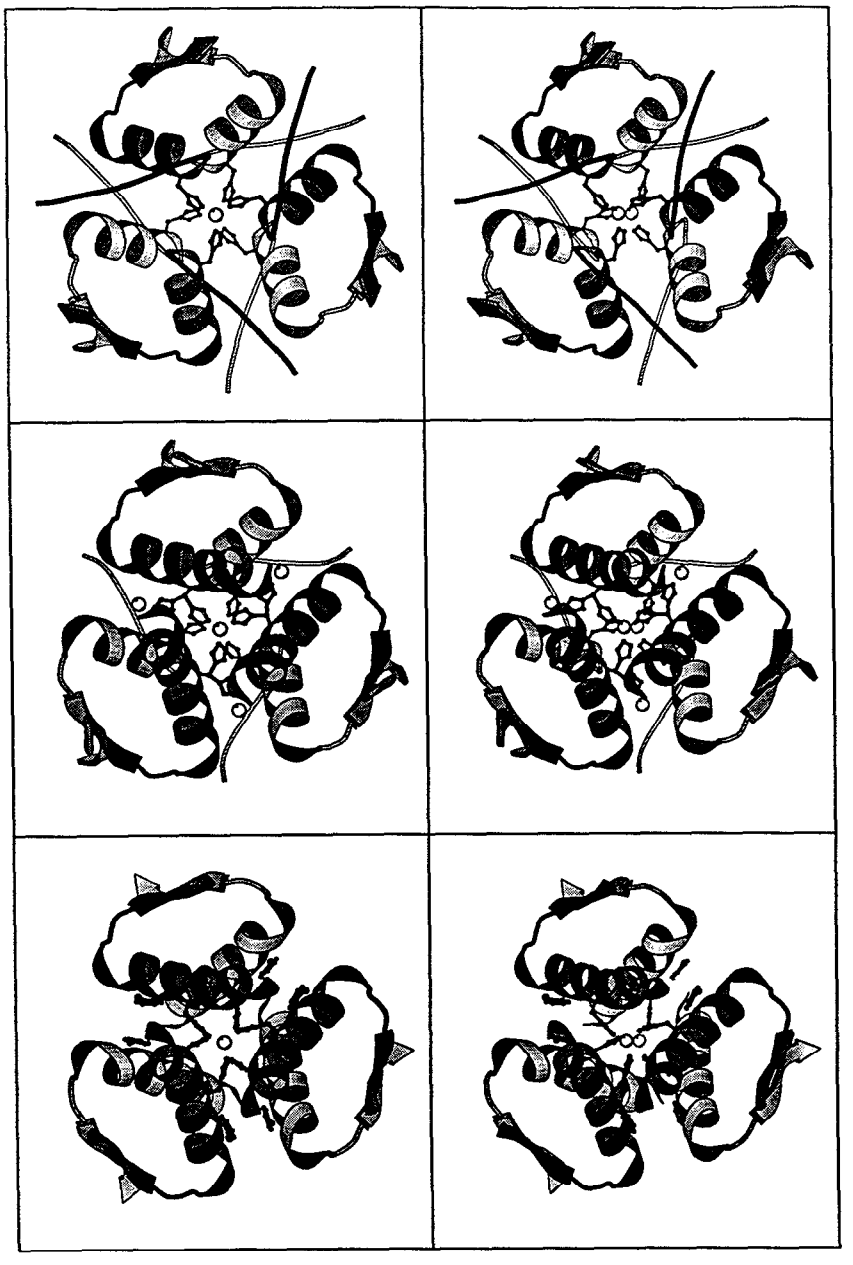


Figure 2.2 Stereographic representation of (upper panel) 2-zinc pig insulin (T_6), (middle panel) 4-zinc pig insulin ($T_3R_3^f$) and (lower panel) rhombohedral phenol insulin (R_6). For clarity only the B-chains are shown, and in each dimer, the two B-chains are shaded differently. The positions of the zinc ions (white spheres), B10 His sidechains, and phenol ligands (ball and stick) are shown. Adapted from Wittingham et al. (1995) with authors permission. The program MolScript was used to make the representations (Kraulis, 1992).

Although insulin hexamers are extensively characterized by X-ray crystallography, the physiologically active form of the hormone is the 5808 Dalton monomer. Hence, in any discussion of structure-activity relationships for insulin, it is necessary to consider whether crystal packing forces have modified the structure from that required for biological action. Due to the complicated pattern of

insulin self-association in solution, detailed NMR analysis of the insulin monomer has often been ambiguous. Accordingly, several groups have reported results of NMR studies carried out at low pH (i.e., pH 1.8-3.5) using either modified insulins, organic co-solvents, or both, in an effort to counteract self-association (see e.g. Kline & Justice, 1990; Knegtel et al., 1991; Hua et al., 1991, 1992a, 1993b; Jørgensen et al., 1992; Ludvigsen et al., 1994). The reported structures generally agree that the secondary structure of the insulin monomer in solution is similar to that of the crystallographically identified T-state. However, considerable differences are apparent in terms of the structural resolution expressed e.g. by the atomic rms values.

In one extreme, insulin structures with properties similar to a 'molten-globule' state have been reported. On the basis of these structures determined in the presence of 20% acetic acid, it has been postulated that the lack of tertiary structural detail is intrinsic to the native insulin monomer and that the phenomenon per se is important for the interactions of insulin with its receptor (Hua et al., 1991, 1992a,b, 1993a,b). In contrast with these low resolution structures, our NMR studies on the biologically active B16 Tyr → His mutant in water at low pH have revealed a well-defined solution structure (Ludvigsen et al., 1994; Kaarsholm & Ludvigsen, 1995). Because the B16 His mutant remains essentially monomeric at millimolar concentrations in aqueous solution at low pH, these results identify sample homogeneity and aggregation state as major determinants for the quality of the NMR-derived structure.

The high resolution solution structure of the insulin monomer at low pH provides an important prerequisite for the understanding of the interaction between insulin and its receptor. However, a long-term goal for these NMR studies is to determine the structure and dynamics of engineered monomers under physiologically conditions as a base line for further structural examination of mutations associated with enhanced or diminished affinity for the insulin receptor.

Detailed NMR studies of the insulin monomer at neutral pH have so far been limited by protein solubility and self-association in addition to the inherently faster rate of NH exchange in this pH region. In the present work we use site-directed mutagenesis to manipulate the solubility and aggregation pattern of insulin in the neutral pH region. These experiments identify a mutant and a set of conditions where the monomer is amenable for high resolution NMR structural analysis. The preferred mutant has glutamate residues substituted into four positions, i.e., B1 Phe, B10 His, B16 Tyr, and B27 Thr and the C-terminal B30 Thr is removed. This mutant retains 47% biological potency and is monomeric at millimolar concentrations in the pH range 6.5-7.5. The resulting structure is generally well-defined as evidenced by a high number of sequential NOEs as well as many long range NOEs. Apparent disorder is observed near the termini of the B-chain, e.g. B1-B4 and B27-B29, and is to a large extent ascribed to local effects of the mutations.

2.3 Materials and methods

2.3.1 Materials

Native and mutant insulins were constructed by oligonucleotide-directed mutagenesis, fermented in yeast, and purified as described (Markussen et al., 1987; Brange et al., 1988). In a typical small-scale preparation, the mutant is expressed and partly purified as a single-chain mini-proinsulin precursor, B1..B29Lys-Ala-Ala-Lys-A1Gly..A21. Prior to the final purification step, the connecting peptide is cleaved off by treatment with a lysyl endopeptidase (Achromobacter Protease I, EC 3.4.21.50; Wako Inc., Japan), and the mutant insulin is isolated in the form of des-B30 Thr insulin. The removal of the B30 Thr residue has no effect on the biological potency of the molecule.

2.3.2 CD Spectroscopy

CD spectra were recorded with a Jobin Yvon Mark V dichrograph calibrated with (+)-10-camphorsulfonic acid as described (Kaarsholm et al., 1993). Near-UV CD spectra were recorded between 250 and 350 nm using an appropriate combination of cell pathlength and protein concentration to yield an absorbancy of less than 1. Protein concentrations were determined by UV absorbance using $\varepsilon_{276}= 6.2 \times 10^3$ $M^{-1}cm^{-1}$. The same extinction coefficient was used for estimation of the concentration of mutant species with the assumption that each of the four tyrosines in human insulin contributes 25% to ε_{276}

2.3.3 NMR Spectroscopy

Samples were prepared by dissolving the lyophilized protein powder in 10/90 D_2O H_2O or 99.8 % D_2O and adjusting the pH as desired by additions of small amounts of 1 M DCl or NaOD. All pH meter readings are without correction for isotope effects. For all NMR experiments reported herein a temperature of 307 K was used.

Two dimensional ^1H-^1H NMR spectra, DQF-COSY (Piantini et al., 1982; Rance et al., 1983). NOESY (Jeener et al., 1979; Anil-Kumar et al., 1980, 1981) and TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) were recorded on a Bruker AMX600 spectrometer. For TOCSY and NOESY spectra, mixing times were between 40-90 and 120-180 ms, respectively. All spectra had a spectral width in both dimensions of 6579 Hz; 1024 t_1 increments were acquired each with a size of 2048 real datapoints. The spectra were recorded in the phase-sensitive mode using the time proportional phase incrementation scheme (TPPI, Marion & Wüthrich, 1983). The carrier was placed on the water resonance to enable irridation of the water during a period of 1.5 s between the individual scans.

Prior to Fourier transformation the FID's were zero filled once in both dimensions. For resolution enhancement of the DQF-COSY spectra, a squared sine-bell shifted 90° was used in both dimensions, whereas for the NOESY and TOCSY spectra, a Gaussian function with an exponential line broadening of -7 Hz and a factor of 0.15 was applied. Data processing was performed using the MNMR package (PRONTO Software Development and Distribution, Copenhagen, Denmark) or the Bruker UXNMR software on a Silicon Graphics Indigo computer. Exchange of amide protons was followed

by a series of 8K datapoint one-dimensional spectra. These spectra were zero filled once, and a 1.5 Hz line broadening was used as a window multiplication prior to the Fourier transformation.

The program PRONTO (PRONTO software Development and Distrubution; see Kjær et al., 1991) was used to keep track of spectral assignments, cross-peak integration, sequence-specific and stereo-specific assignments and related bookkeeping during spectral analysis. $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$ coupling constants were measured by the facility in the PRONTO software which uses a combined analysis of COSY and NOESY spectra (Ludvigsen et al., 1991). Chemical shifts were measured in parts per million as observed relative to dioxane (3.75 ppm).

2.3.4 Structure Calculations

The program X-PLOR (Brünger, 1992) was used to calculate structures based on distance- and dihedral angle restraints derived from the NMR spectra. Integrated NOESY cross peaks were divided into three classes of distance restraints using 1.0 Å as lower limit and upper limits of 2.7, 3.3 and 5.5 Å, respectively. For restraints involving methyl groups, an additional 0.5 Å was added to the upper limit. The measured $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$ coupling constants were converted into ϕ-angle restraints as follows: -60°± 30° (2-4 Hz); -70° ± 30° (4-6 Hz); -120° ± 60 (6-8 Hz); and -120° ± 35° (8-10 Hz). The stereospecific assignment of several β methylene protons was obtained by combined analysis of COSY cross peak patterns (Hybert et al., 1987) and of the intraresidual NOE intensities between the methylene protons, the α- and the amide protons, respectively (Wagner et al., 1987). This method allowed assignment of the side-chain to one of the three staggered conformations, χ_1, assuming the values -60° ± 60°, 60° ± 60°, or 180° ± 60°. Stereospecific assignment of δ methyl groups of Leu residues were found and converted to χ_2 restraints in a similar fashion. The structure calculations were performed using a combination of distance geometry (Crippen & Havel, 1988; Kuszewski et al., 1992) and simulated annealing as proposed by Nilges et al. (1988).

2.3.5 Comparison of Solution Structures

From the Protein Data Bank (PDB), Brookhaven (January 1996) all files containing structures of monomeric native and mutant insulins determined by NMR were collected. Each file contains an ensemble of structures as well as a representative structure. For the purpose of comparison of the different structures, geometric average structures were calculated from these ensembles of structures following the alignment procedures presented in the literature references as listed in the header of the files. In order to obtain a reasonable comparison, we have made one exception to this procedure and used both chains rather than just the B-chain helix in the alignment of the DPI structures. Table 2.3 at the end of this Chapter gives a summary of the experimental conditions and details concerning the structural calculations as referenced in the PDB files and in the original publications. The program X-PLOR (Brünger, 1992) was used to calculate the average structures.

2.4 Results

2.4.1 Design of a Monomeric Insulin Suitable for 2D NMR at Neutral pH

The aggregation and precipitation pattern of insulin is a complex function of protein concentration, pH, temperature, metal ions, ionic strength, and solvent composition. In the millimolar concentration range usually required for high-resolution NMR structural work, metal-free insulin exists predominantly as dimers and higher oligomers at pH 2.0-3.5. At higher pH aggregation increases, which in turn leads to precipitation in the pH region between 4.2 and 6.6. Above pH 6.6 insulin dissolves, presumably as a mixture of dimers and higher aggregates, which then gradually dissociate as the pH increases. As the pH is raised above 10-11, the monomeric state is finally reached (Jeffrey & Coates, 1966; Pekar & Frank, 1972; Lord et al., 1973; Goldman & Carpenter, 1974; Jeffrey et al., 1976; Pocker & Biswas, 1981; Strazza et al., 1985; Mark et al., 1987; Kaarsholm et al., 1990; Roy et al., 1990a; Kadima et al., 1992). The X-ray structure of the T_6 zinc insulin hexamer shows that distinct interfaces are involved in insulin dimer and hexamer packing (Baker et al., 1988). The published equilibrium constants for solution aggregation along the monomer/monomer interface are similar at pH 2 and pH 7 (i.e., $\approx 10^5$ M^{-1}). However, the corresponding aggregation along the dimer/dimer interface appears much stronger at pH 7 than at pH 2 (Mark et al., 1987). Hence, in contrast with the situation at low pH (Ludvigsen et al., 1994), insulin self-association must be inhibited along two distinct interfaces in order to obtain monomers at millimolar concentrations at neutral pH. At the same time, the resulting mutant should retain near-native biological potency and the solubility pattern must allow NMR measurement in a pH range where the exchange rate of backbone NH is sufficiently slow for observation with conventional ^{1}H-NMR techniques (Wüthrich, 1986).

Aggregation properties of various insulin mutants were evaluated by near-UV circular dichroism (CD) and NMR resonance line widths. The near-UV CD spectrum (350-250 nm) of insulin reflects the environment of the tyrosine chromophore. The signal is very sensitive to aggregation and may be used to monitor the formation of dimer interface (Morris et al., 1968; Goldman & Carpenter, 1974; Wood et al., 1975; Strickland & Mercola, 1976). NMR spectra are sensitive to insulin self-association because longer rotational correlation times lead to line broadening and because dynamic equilibria between oligomeric states can lead to line broadening in the intermediate-exchange regime.

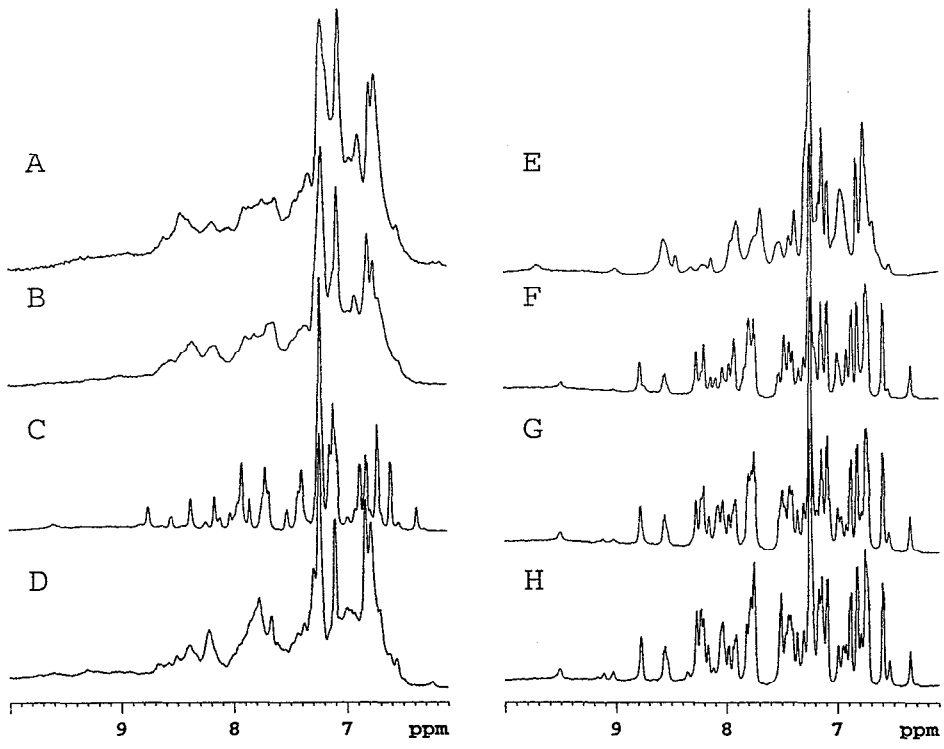


Figure 2.3 1D NMR spectra showing the aromatic and amide proton resonances of native and mutant insulins at 1 mM protein concentration in 10/90 D_2O/H_2O, 307 K. Panels (A): human insulin, pH 7.5; (B): B27 Thr → Glu , pH 7.5; (C): (B16, B27) Glu, pH 7.5; (D): (B1, B27) Glu, pH 7.5; (E): B10 Glu, des-B30, pH 7.5; (F)-(H): (B1, B10, B16, B27) Glu, des-B30 insulin at pH 7.5, 7.0 and 6.5, respectively.

Figure 2.3 shows the aromatic and amide proton resonances of the 1D ^{1}H-NMR spectra for human insulin (panel A) and a series of mutants (panels B-F) at 1 mM protein in 10/90 D_2O/H_2O, pH 7.5. Under these conditions, human insulin exists as a mixture of dimers and higher aggregates and the resulting spectrum is very poorly resolved. The X-ray structure of the T_6 insulin hexamer shows that dimer formation is dominated by a series of non-polar contacts contributed by B-chain residues, notably B12 Val and B16 Tyr in the central helix, and residues B23-B28 in the extended chain. Panels (B-C) illustrate the effects of introducing charged residues at two different positions in the monomer/monomer interface. As shown in Panel (B), the B27 Thr → Glu mutation at the edge of the monomer/monomer interface effects little or no improvement in spectral resolution relative to human insulin. This mutation has a minor inhibitory effect on dimer formation (Brange et al., 1988); however, in the present study the modification is preferred due to its positive effect on the expression level in yeast and hence on the overall fermentation yield during small-scale mutant preparation. Introduction of charge into the B16 Tyr position has previously been shown to provide an efficient means of inhibiting the formation of dimer interface at low pH (Ludvigsen et al., 1994; Kaarsholm & Ludvigsen, 1995). In accordance with this result, panel (C) shows that the B16 Tyr → Glu mutation strongly enhances the spectral resolution at

pH 7.5. While the (B16, B27) Glu species is nearly monomeric at 1 mM protein, the exchange rate of several NH-protons is too fast for observation in 2D spectra. At slightly lower pH, extensive line broadening is observed in concert with aggregation along the dimer/dimer interface. Accordingly, the number of cross-peaks observed in 2D NOESY spectra of the (B16, B27) Glu mutant is significantly smaller at pH 7.0 than at pH 7.5 (data not shown).

When the T_6 hexamer is assembled from dimers through the coordination of Zn^{2+} to the B10 His residues, both polar and non-polar residues are buried between the dimers. The packing is correspondingly much looser than in the monomer/monomer interface within each dimer. A set of important contacts across the dimer-dimer interface involves the N-terminal part of the B-chain. Of particular interest is the B1 Phe residue, which fit into a pocket between the main A-chain and A14 Tyr residue of its neighbor. Because the two A14 Tyr are also in contact across the dimer-dimer interface, a close aggregate involving four aromatic residues is formed (Baker et al., 1988). Comparison of panels (B) and (D) in Figure 2.3 shows that the B1 Phe → Glu mutation leads to somewhat improved resolution of the 1D NMR spectrum. A slightly better improvement is obtained with the B10 His → Glu mutation positioned at the edge of the dimer/dimer interface, viz. panel (E).

The B10 His → Glu mutation also has the effect of increasing the biological potency of the insulin molecule 4-fold as determined by the ability to incorporate [2-^{3}H]-glucose in isolated mouse adipocytes according to Moody et al. (1974). In the same assay, the B27 Glu, (B1, B27) Glu, (B16, B27) Glu and des-B30 insulins exhibit potencies of 1.07, 0.97, 0.13 and 1.00, respectively, relative to a value of

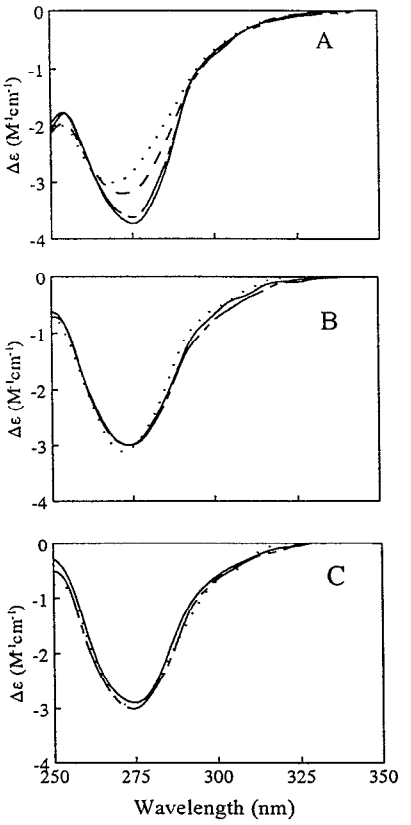


Figure 2.4 Near-UV CD spectra as a function of protein concentration. (A): zinc-free human insulin in 10 mM Tris/ClO_4^-, pH 8.0, and (· · · ·) 5 μM, (– – –) 20 μM, (– — –) 200 μM, and (——) 2 mM protein. (B) and (C): (B1, B10, B16, B27) Glu, des-B30 mutant insulin in 25 mM Tris/ClO_4^-, pH 7.5 (B) and 25 mM phosphate, pH 6.5 (C), and (· · · ·) 50 μM, (– — –) 600 μM and (——) 2.6 mM protein.

1.00 found for human insulin. For the mutant carrying all modifications, i.e. (B1, B10, B16, B27) Glu, des-B30 insulin, a relative potency of 0.47 is found indicating that the substitutions affect the biological activity in a nearly independent manner. Panel (F) in Figure 1 shows that the 1D spectrum of (B1, B10, B16, B27) Glu, des-B30 insulin is well-resolved and retains the dispersion of resonance lines characteristic for globular proteins. Furthermore, as the pH is adjusted down to 7.0-6.5, the high resolution is maintained (panels (G) and (H)), while selected NH

resonances become sharper as expected due to the slowed exchange rate.

Figure 2.4 shows near-UV CD spectra as a function of protein concentration for native insulin at pH 8.0 and the (B1, B10, B16, B27) Glu, des-B30 mutant at pH 6.5 and 7.5. For native insulin, the progressive increase in intensity of the negative signal around 274 nm indicates the expected increase in association with increasing concentration. In contrast, the near-UV CD spectrum of the mutant is independent of the protein concentration in the 60 μM to 2.6 mM range at both pH 6.5 and 7.5. Although the mutant is missing one tyrosine reporter group (B16), Tyr B26 is expected to be strongly affected by monomer/monomer interactions. Hence, these results provide further evidence that the (B1, B10, B16, B27) Glu, des-B30 mutant is monomeric under conditions necessary for NMR structural studies at neutral pH, and this species was selected for detailed characterization.

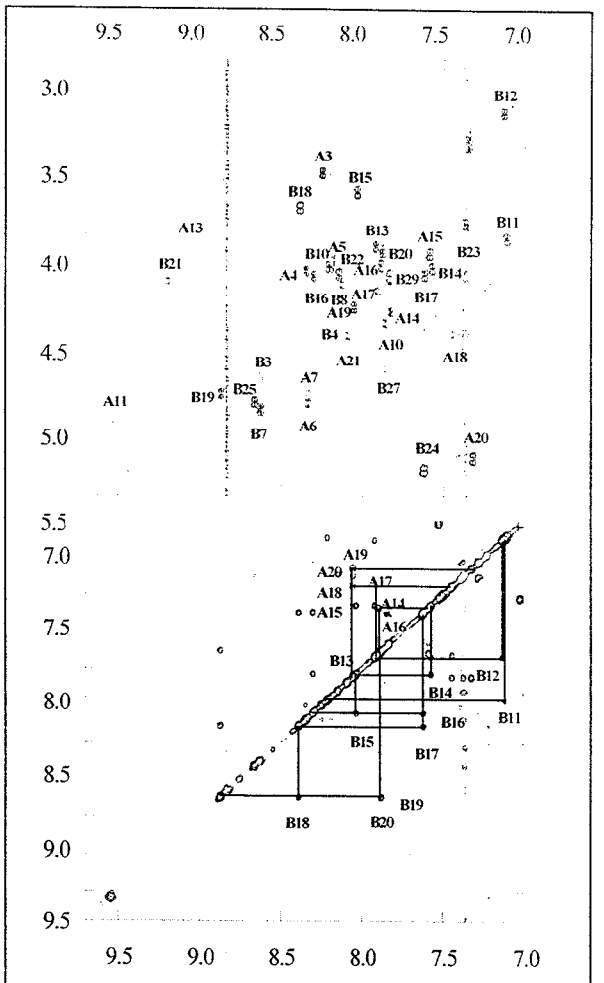


Figure 2.5 Sections from the DQF COSY spectrum (top) and NOESY spectrum (mixing-time 180 ms) (bottom). The fingerprint region of the DQF COSY spectrum shows the major part of the correlations between α-protons and amide protons that was identified. The correlations between amide protons indicative of helical areas in the structure are annotated in the bottom panel, and lines are drawn to follow the amide protons throughout the helices.

2.4.2 Assignment of Spin Systems

NMR spectra were assigned using the standard procedures outlined by Wüthrich (1986). The fingerprint region of the DQF-COSY spectrum is shown in the upper panel of Figure 2.5. The chemical shifts are dispersed and well-resolved as expected for a structured globular protein, and 37 of the 47 possible H^NH^α cross peaks are annotated in the plot. Among the remaining H^NH^α cross peaks, six were assigned using TOCSY spectra. These are the resonances that are either close to the water line or unusually broad, i.e., A8, A9, A12, B5, B6, and B26. The linewidth of the CysA11 amide proton is unusually broad, but identification was possible in DQF-COSY spectra. Amide protons from residues A2, B6 and B9 were not identified. Finally, for the side-chains of Glu and Gln residues, considerable overlap of cross peaks in the H^β-H^γ area was resolved using TOCSY spectra occasionally supported by NOESY.

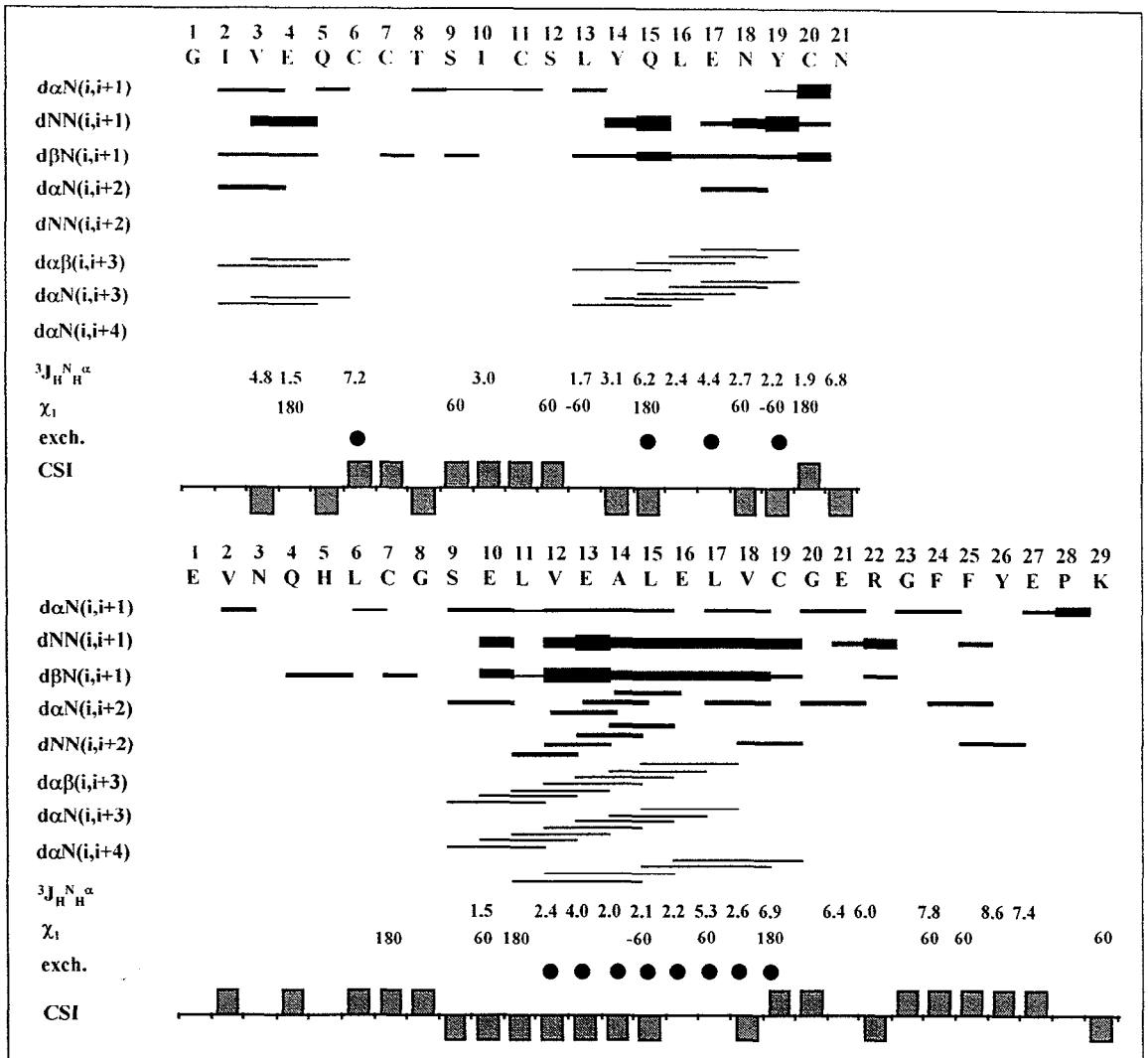


Figure 2.6 Summary of the data used for sequential assignment and determination of secondary structural elements. The thickness of the bars indicates the intensity of the NOEs as measured in a NOESY spectrum with a mixing-time of 150 ms. The coupling constants and the χ_1 dihedral angles were measured as desribed in Materials and Methods. The filled circles indicate that the amide proton signal is observable in a 1D spectrum at least 12 minutes after dissolution of the sample in D_2O. The last row shows the chemical shift index, CSI, (Wishart et al., 1992) of the α-protons.

Table 2.1 at the end of this Chapter contains a list of the ^1H chemical shifts of the insulin mutant.

2.4.3 Sequential Assignment and Secondary Structure

The sequential assignment was straightforward for most parts of the insulin mutant. Including the regions B9-B29, A9-A12, and A13-A21. The assignment of the stretch from A4 to A8 was complicated by overlapping H^N and H^α resonances; this is reflected in the lack of sequential NOEs as well as NOEs indicative of secondary structure in this region. In the N-terminus of the B-chain, the NOEs are sparse and the assignments here were made by exclusion as the very last part of the procedure.

The first indications of the secondary structures come from inspection of the chemical shifts of the α-protons compared to the random coil values (Wishart et al., 1992). The chemical indices shown in Figure 2.6 suggest helical structures characterized by upfield shifts of the α-protons in the second part of the A-chain (the A(II)-helix) and in the central part of the B-chain (the B-helix), whereas the helix expected in the N-terminus of the A-chain (the A(I)-helix) seems less well defined. These observations are substantiated by the NOEs assigned in the same regions as shown in Figure 2.6 and in the NOE matrix in the lower panel of Figure 2.7.

The B-helix, stretching from B9 to B19, is a well defined helix characterized by a dense network of NOEs between the α-protons in position i and the amide protons in position i+3 and position i+4 as well as the β-protons in position i+3. Furthermore, stronger NOEs between amide protons 2 residues apart in the helix are found. The C-terminal A(II)-helix comprising the amino acids from A13 to A20 has a less dense network of NOEs, i.e., only between α-protons in position i and β- and amide protons in position i+3. The N-terminal A(I)-helix stretches from A2 to A6 and is characterized by the same type of NOEs as the A(II)-helix. In the lower panel of Figure 3, the H^N to H^N part of the NOESY spectrum is shown with annotations for the sequential assignment of the helical areas.

An amide proton characterized by an exchange rate slower than the average is a good evidence for participation in a hydrogen bond either in a secondary structural element or as part of the tertiary fold of the protein. At near neutral

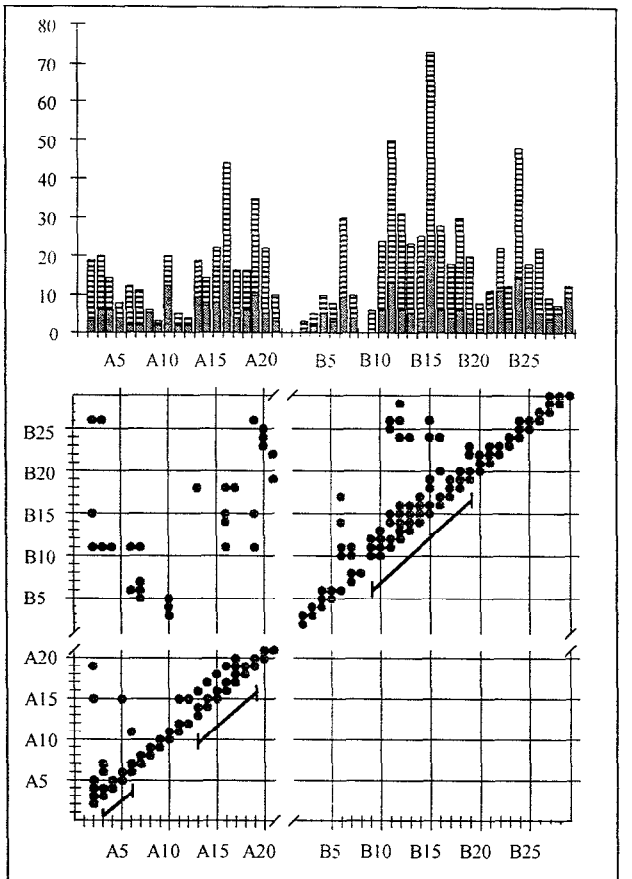


Figure 2.7 The upper panel shows the distribution of NOEs by residue, intraresidual NOEs are marked with filled boxes and interresidual NOEs with vertical lines. In the matrix in the lower panel filled circles indicate the presence of at least one NOE between the connected residues. Multiple occurences of NOEs on both sides of the diagonal in a spectrum or in different spectra were filtered out before the conversion to distance restraints. The presence of helical regions, as expected from the patterns of NOEs, is indicated.

pH, the exchange rate of amide protons is in the order of 10^2 min^{-1} (Englander et al., 1972). From 1D proton spectra recorded in the period between 12 and 60 minutes after dissolution of the insulin mutant in D_2O, it is possible by visual inspection to follow the decay of amide proton resonances belonging to the B-helix and the A(II)-helix as well as a single amide proton in the A(I)-helix. After 50 minutes, all amide protons are exchanged by deuterons. The locations of the slow exchanging

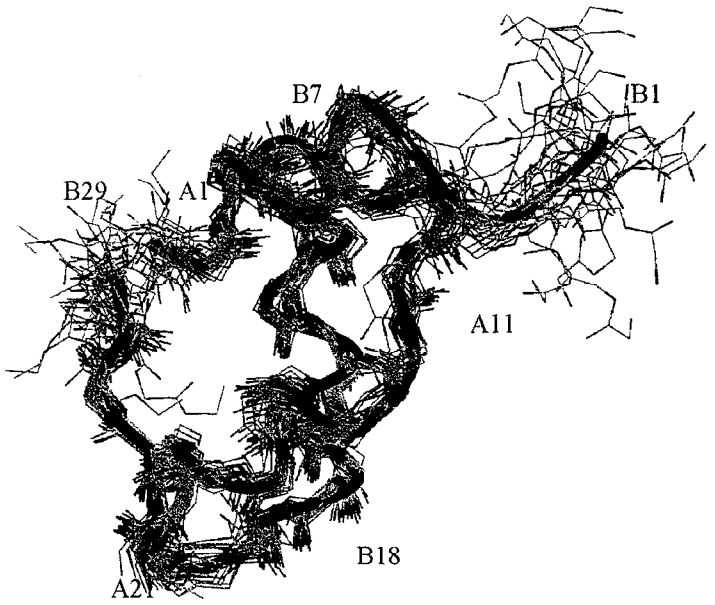


Figure 2.8 The backbone of 25 converged structures of the insulin mutant superimposed together with their geometric average. The average structure is represented by a ribbon, the A-chain is shown in red, the B-chain in blue. The structures are aligned in the regions B4-B26, A2-A19.

amide protons according to this criterion are shown in Figure 2.6.

2.4.4 Structure Calculations

Table 2.2 lists the number of distance restraints and dihedral angle restraints used for the final structure calculations. In addition, the sulfide bridges were added as distance restraints, while no restraints were added for hydrogen bonds. The matrix in the lower panel of Figure 2.7 shows the distribution of NOEs between residues. Multiple occurences of NOEs from both sides of the diagonal in a spectrum or from different spectra were filtered out before the conversion of NOEs to distance restraints. As a supplement, the upper panel of the same Figure depicts the number of NOEs by residue. The distribution of dihedral angle restraints on residues is shown in Figure 2.6. The calculation of structures using X-PLOR 3.0 proceeded as described in Materials and Methods starting from a reduced set of distance restraints. In a sequence of iterations, this set was slowly expanded as ambiguities could be resolved, and finally the dihedral angle restraints were introduced. A total of 100 structures were calculated 25 of which were characterized by distance restraints violations below 0.3 Å, dihedral angle restraints violations below 2°, and a low total energy. Table 2.2 summarizes the structural statistics for this set of structures.

The set of 25 converged structures is depicted in Figure 2.8 (backbone representation) together with their geometric average represented by a ribbon. In Figure 2.9, the side-chains have been included, and the average structure shown

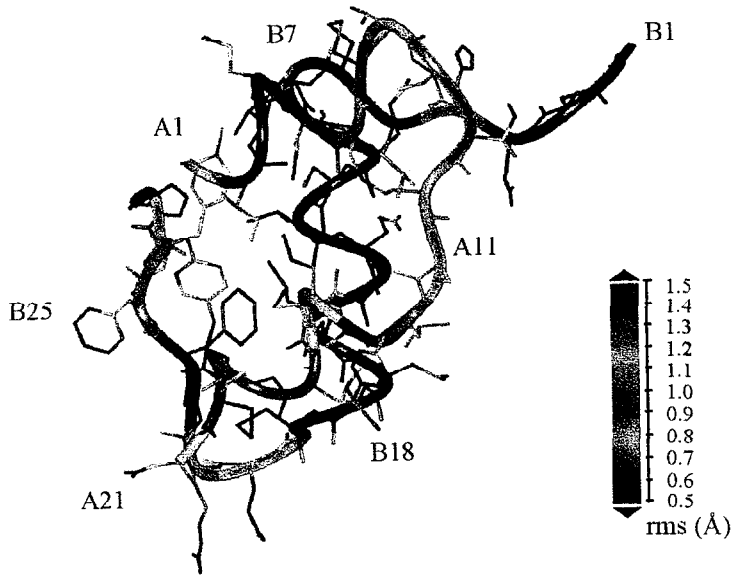


Figure 2.9 The geometric average structure of the insulin mutant colored according to atomic rms. The average structure was calculated based on alignment of residues B4-B26 and A2-A19.

here is colored according to the atomic rms deviation.

2.4.5 Description of the Structures

The major part of the mutant structure is well-defined. This goes indeed for the B-helix (B9-B19) which exhibits low rms deviation among the structures as expected from the large number of structural NOEs as well as the small coupling constants, indicative of helix structure. The hydrogen bonds fit a regular α-helix pattern all the way starting from B9(CO)···B13(H^N) and ending at the last donor B19(H^N).

At the C-terminal end of the helix, a type I turn from B20 to B23 turns into a β strand that stretches along the central helix. The three C-terminal residues are disordered. As is evident from Figure 4, no NOE structural information was obtained in the B1-B4 region of the B-chain. This fact is reflected in the random distribution of the termini among the 25 structures presented in Figure 2.8 and implies that the average structure calculated in this area is physically meaningless. The A(II)-helix (A13-A20) shows a high degree of accordance within the bundle of structures, and the hydrogen bonding pattern fits an α-helix A13(CO)···A17(H^N), A15(CO)···A19(H^N), and A16(CO)···A20(H^N), starting out with the irregular A12(CO)···A15(H^N) hydrogen bond. The A(I) helix (A2-A8) is a little less well-defined, a single hydrogen bond between A2(CO) and A6(H^N) is established. The loop area (A9-A12) connecting the two A-chain helices shows a larger spread of conformations. In 20 out of 25 calculated structures the interchain hydrogen bond A19(CO)···B25(H^N) is established, in agreement with crystal structures. The atomic rms deviations calculated are presented

in Table 2.2 along with the general structural statistics.

2.5 Discussion

The (B1, B10, B16, B27) Glu, des-B30 mutant is the first example of an insulin species that remains monomeric at millimolar concentration in aqueous solution at neutral pH. Hence, by tailored mutations in two dinstict protein/protein interfaces, the inherent self-association of insulin is strongly inhibited, while the biological potency is essentially retained. As a result, samples may be prepared under conditions where the exchange rate of backbone amide protons is sufficiently slow for detailed NMR analysis. Previous NMR investigations of native and mutant insulins at near-neutral pH have been limited by incomplete assignments of the spectra. Nevertheless, on the basis of 1D spectra of human insulin at high dilution and of the B9 Ser → Asp monomeric mutant in the pH 8-9.5 range, Roy et al. (1990a,b) were able to show that the association of insulin monomers into dimers is accompanied by a change in conformation involving the relative position of residues B15 Leu and B24 Phe. In their NMR analysis of the B10 Asp, B28 Lys, B29 Pro (DKP) insulin monomer at pH 8.0, Weiss et al. (1991) also reported partial assignments and employed selective ^{2}H and ^{13}C labeling of the B23-B26 residues to confirm that this region is indeed in close contact with residues from the central B-chain helix.

The (B1, B10, B16, B27) Glu, des-B30 mutant is structurally well-ordered with the exception of the B-chain termini, i.e., residues B1-B4 and B27-B29. The observed disorder in the N-terminus is most reasonably ascribed to opposing effects of the B1 and B10 mutations. Removal of the hydrophobic B1 Phe side-chain destabilizes the packing of the B1 residue against the A13 Leu region (Ludvigsen et al., 1994), while the B10 Asp side-chain caps the central B9-B19 helix and hence stabilizes the T-state relative to the R-state (Kaarsholm et al., 1993). In any event, because the B-chain helix is clearly confined to residues B9-B19, the overall structure resembles a crystallographic T-state more than a R-state. In the C-terminal B29-B30 residues, disorder is observed in the crystal structure as well as in the low-pH solution structure (Baker et al., 1988; Ludvigsen et al., 1994). In the present work at neutral pH, the lack of structural definition extents to the last three residues, B27-B29. Again, the effect is most likely due to the modifications employed, i.e., the introduction of negative charge close to the shortened C-terminus (B27 Thr → Glu, des-B30). In accordance with previous work (Kline & Justice, 1990; Ludvigsen et al., 1994), broad amide proton resonances are observed in the A-chain loop region suggesting that conformational substates exchange on a millisecond time scale in this particular region.

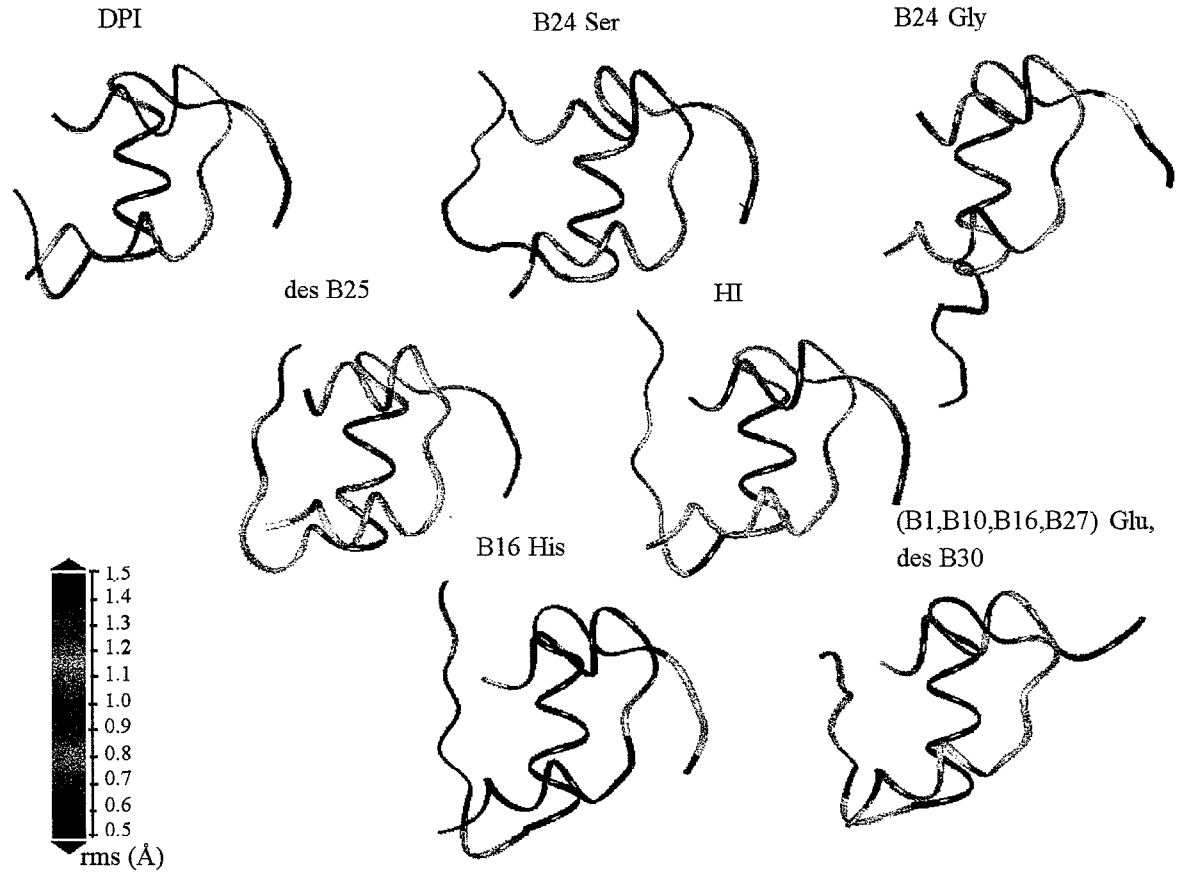


Figure 2.10 Ensemble averages of three dimensional structures of native (HI) and mutant insulins determined by NMR. The structures are color coded according to the atomic root mean square deviation of the backbone. For further details, see Table 2.3.

2.5.1 Comparsion with Other Monomer Solution Structures

Several NMR investigations of insulin monomers have been carried out at low pH and/or in the presence of organic co-solvent. Table 2.3 compares experimental conditions and details concerning the structure calculations for the (B1, B10, B16, B27) Glu, des-B30 mutant and for other insulin monomer structures deposited in the Brookhaven Data Bank. In addition to these monomers, the structure of a B9 Asp insulin dimer has been described (Jørgensen et al., 1992), and the low-pH structure of DPI has been determined by Knegtel et al. (1991) as well as by Hua et al. (1992a). Figure 2.10 compares the three dimensional structure of (B1, B10, B16, B27) Glu, des-B30 insulin with those of the other monomers.

The structures are color coded according to the atomic rms deviations characterizing each of these ensemble average structures. Table 2.3 contains details concerning the average structure calculations, alignment procedures and rms deviations of the backbone calculated by using the same alignment procedure for all mutants. Note that for certain parts of these structures, distance restraints containing information about secondary and tertiary fold are either absent or very few in number. This is particularly true for the C-terminal B22-B30 residues of the B24 Gly and B24 Ser mutants (Hua et al., 1991, 1993b). As a result, considerable atomic rms deviations are obtained which make the average structures physically meaningless in these parts colored red in Figure 2.10. Paradoxically, a distinct bulge appears on the C-terminal β-strand of the B24 Ser structure.

The precision of the individual structures is quantified by the ensemble average rms deviations in Table 2.3. A comparison of the two columns of backbone rms deviations in the Table reveals that most of the differences are caused by imprecision of the C-terminal of the B-chain. The B16 His structure shows the highest degree of overall accordance between the individual structures in the ensemble. In this case, the presicion is directly correlated with the number of interresidual distance restraints used in the structure calculation. However, due to differences in the procedures for counting of the NOEs and for the translation of NOEs into distance restraints, the correlation between structural precision and the number of NOEs is generally not straightforward. The high precision of side-chain structures in B16 His insulin are closely related to the number of assignments of χ_1 and χ_2 dihedral angles. The individual monomer structures are compared one by one in Table 2.4 at the end of this Chapter. It is clear that the B16 His insulin has the closest structural resemblance to the (B1, B10, B16, B27) Glu, des-B30 mutant. The major difference between these two structures is the lack of definition of residues B1-B4 in (B1, B10, B16, B27) Glu, des-B30 insulin. In B16 His insulin, the close contact between B-chain N-terminus and the A-chain loop region provides the set of NOEs defining the spatial arrangement of the loop. Thus, the lack of structure in residues B1-B4 propagates to an increased inpresicion of the A-chain loop region in (B1, B10, B16, B27) Glu, des-B30 insulin.

Finally, we note that with the exception of the B-chain termini, the structures depicted in Figure 9 share common features including the secondary structure elements and the overall spatial arrangement of these elements into the tertiary structure. As would be expected for a globular protein, Table 2.3 shows that each structural core is characterized by a low rms deviation (high precision) relative to the precision of the overall structure. In contrast with these calculations, Hua et al. (1992a, 1993a) reported that variations in the mutual spatial arrangement of the two chains in the case of DPI span the differences reported among various insulin crystal forms. The phenomenon was ascribed to the absence of restraints and interpreted as evidence that DPI exists in a 'molten globule' state under native-like conditions (Hua et al., 1992a, 1993a). A molten globule state is usually characterized as a compact denatured state arising as an intermediate in the pathway from native to unfolded protein under slightly denaturing conditions (see e.g. reviews by Dobson (1991,1992)). The state is characterized by the presence of secondary structural domains, but absence of tertiary interactions, i.e., the conformation of the side-chains is essentially random. In a NMR spectrum, one feature

of a molten globule state would be a reduced dispersion of side-chain resonances in comparison with the native state (Goto & Fink, 1989). The original alignment of the DPI ensemble presented by Hua et al. (1992a) was based on a subset of the structure, i.e., the central B-chain helix (residues B9-B19). This procedure per se leads to an accumulation of the overall structural imprecision in the parts that are left out of the alignment, in this case the entire A-chain and the B-chain termini. In contrast, our alignment procedure (Table 2.3) is based on both chains leaving out only those parts of the molecule where the number of restraints per residue is well below the average, i.e., the outermost parts of both chains. The same procedure was used by Knegtel et al. (1991) for the alignment of DPI. With this procedure, the major part of all molecules (including DPI) exhibit reasonably well-ordered secondary and tertiary structure. For the (B1, B10, B16, B27) Glu, des-B30 mutant at neutral pH and the B16 His mutant at low pH, the dispersion of side-chain resonances corresponds to that expected for a packed, globular protein. Furthermore, both mutants exhibit a distinct near-UV CD spectrum very similar to that of native human insulin (Figure 2 and Ludvigsen et al., 1994). Hence, there seems to be no evidence coming from NMR or CD to support the idea that insulin mutants exist in a molten globule state characterized by variations in the mutual arrangement of the chains.

2.6 Supplements

Table 2.1 ^{1}H chemical shifts of the insulin mutant at pH 6.5 and 307 K measured in ppm relative to dioxane (3.75 ppm)[a]

Residue		H_N	H_α	H_β	Others
A1	Gly				
A2	Ile		3.98	1.23	$H_{\gamma 1}$ 1.22/0.95 $H_{\gamma 2}$ 0.77 H_δ 0.69
A3	Val	8.24	3.66	2.03	$H_{\gamma 1}$ 0.97 $H_{\gamma 2}$ 1.03
A4	Glu	8.34	4.21	2.23,2.11	H_γ 2.55/2.33
A5	Gln	8.18	4.15	2.22	H_γ 2.48/2.56 H_ε 7.01/7.52
A6	Cys	8.33	4.96	2.98/3.37	
A7	Cys	8.32	4.91	3.82/3.38	
A8	Thr	8.18	4.20	4.49	$H_{\gamma 2}$ 1.320
A9	Ser	7.43	4.80	4.09,3.95	
A10	Ile	7.85	4.51	1.70	$H_{\gamma 1}$ 1.29/0.72 $H_{\gamma 2}$ 0.78 H_δ 0.66
A11	Cys	9.53	5.08	3.43/3.14	
A12	Ser	8.36	4.77	4.05,4.19	
A13	Leu	9.05	4.08	1.60,1.75	H_γ 1.76 $H_{\delta 1}$ 1.00 $H_{\delta 2}$ 0.91
A14	Tyr	7.81	4.44	3.11/3.13	H_δ 7.23 H_ε 6.96
A15	Gln	7.58	4.12	2.37,2.16	H_γ 2.48/2.32 H_ε 6.85/7.45
A16	Leu	7.88	4.18	1.51/2.07	H_γ 1.85 H_δ 0.80/0.84
A17	Glu	7.91	4.32	2.11	H_γ 2.38/2.45
A18	Asn	7.44	4.57	2.65,2.71	H_δ 7.24/6.68
A19	Tyr	8.05	4.42	3.55,2.89	H_δ 7.37 H_ε 6.87
A20	Cys	7.33	5.29	2.90,3.43	
A21	Asn	8.08	4.58	2.87/2.75	H_δ 6.45/7.49
B1	Glu				
B2	Val		4.19	2.17	H_γ 1.01/1.04
B3	Asn	8.61	4.82	2.87/2.84	H_δ 6.89/7.59
B4	Gln	8.09	4.54	2.00/2.20	H_γ 2.34 H_δ 7.49/6.91
B5	His	8.42	4.71	3.26/3.57	$H_{\delta 2}$ 7.24 $H_{\varepsilon 1}$ 8.29
B6	Leu	8.81	4.61	1.00/1.76	H_γ 1.64 $H_{\delta 1}$ 0.85/0.91
B7	Cys	8.61	5.00	3.05,3.27	
B8	Gly	8.13	4.28		
B9	Ser		4.23	4.09/4.03	
B10	Glu	8.21	4.19	2.32,2.16	H_γ 2.45/2.52
B11	Leu	7.11	4.02	1.89,1.24	H_γ 1.32 H_δ 0.76
B12	Val	7.13	3.31	2.18	$H_{\gamma 1}$ 1.04 Hγ2 1.09
B13	Glu	7.91	4.07	2.10	H_γ 2.51/2.33
B14	Ala	7.56	4.19	1.49	
B15	Leu	8.03	3.76	0.28,1.09	H_γ 1.33 $H_{\delta 1}$ 0.35 $H_{\delta 2}$ 0.66
B16	Glu	8.29	4.23	2.32/2.69	H_γ 2.47/2.12
B17	Leu	7.61	4.23	1.89,1.77	$H_{\delta 1}$ 0.99 $H_{\delta 2}$ 0.74
B18	Val	8.38	3.85	2.09	$H_{\gamma 1}$ 1.12 $H_{\gamma 2}$ 0.97
B19	Cys	8.87	4.92	3.01,3.40	
B20	Gly	7.87	4.11/4.09		
B21	Glu	9.19	4.27	2.17	H_γ 2.27/2.45
B22	Arg	8.14	4.23	2.16/2.29	H_γ 1.97 H_δ 3.49/3.43 H_ε 7.35
B23	Gly	7.36	3.94/4.23		
B24	Phe	7.61	5.36	3.42,3.02	H_δ 6.90 H_ε 7.28 H_ζ 7.38
B25	Phe	8.65	4.96	3.25/3.32	H_δ 7.37 H_ε 7.44
B26	Tyr	8.07	4.72	3.06/3.00	H_δ 7.01 H_ε 6.74
B27	Glu	7.84	4.77	2.10/1.94	H_γ 2.30
B28	Pro		4.37	2.25	$H_{\gamma 1}$ 2.03 $H_{\gamma 2}$ 2.12 $H_{\delta 1}$ 3.79 $H_{\delta 2}$ 3.63
B29	Lys	7.82	4.24	1.77,1.90	H_γ 1.00/1.46 H_δ 1.74 H_ε 3.02

[a] Methylene protons with stereospecific assignment are seperated by a comma, $H_{\beta 1}$ is listed first.

Table 2.2 Structural statistics

Distances and torsion angle restraints

intra residual	297	ϕ	34
short range	193	χ_1	14
long range	51	χ_2	5
cross chain	56		

Energy statistics after simulated annealing (kcal/mol)

Total energy	43.8	±	3.8
Bonds	2.2	±	0.2
Angles	25.5	±	2.2
Repel	7.4	±	1.2
NOE[a]	4.1	±	1.3
Torsion	0.3	±	0.1
Impropers	3.3	±	1.3

van der Waal energy measured with CHARMM[b] potential (kcal/mol)

van der Waal	-70.4	±	11.8

Deviations from Ideal Geometry

Bond (Å)	0.0017	±	0.0001
Angle (deg)	0.35	±	0.02
Improper (deg)	0.26	±	0.04
NOE (Å)	0.012	±	0.002
Torsion (Å)	0.29	±	0.05

Average no. of NOE violations

0.0-0.1 Å	31.8	±	3.1
0.1-0.2 Å	2.0	±	1.0
0.2-0.3 Å	0.16	±	0.37
> 0.3 Å	0		

Atomic rms Values for 25 converged structures vs their geometric average

backbone(all)	1.31	±	0.19
heavy atoms(all)	1.88	±	0.23
backbone(A2-A19) [c]	0.67	±	0.13
backbone(B4-B26) [c]	0.64	±	0.09
backbone(A2-A19,B4-B26) [c]	0.66	±	0.09

[a] Force constants for distance and torsion angle restraints are 50 kcal $mol^{-1}Å^{-2}$ and 200 kcal mol^{-1} $radian^{-2}$

[b] CHARMM potential (Brooks et al., 1983) used for van der Waal energy calculation

[c] Alignment of A2-A19 and B4-B26 backbone

Table 2.3 Details of the three dimensional structures of human insulin and mutants determined by NMR [a]

Mutations	Experimental conditions [b]				Structure calculations [c]				No of	Alignment	rms [e] (Å)		
	pH□	T(C°)	Solvent	Conc(mM)	NOE	φ	χ_1	χ_2	struct	procedure [d]	bb	h	
des-(B26-B30)	1.8	25	20% HAc/80% H_2O	?	368	?	10	0	15	A2-A19,B9-B19 [g]	0.78	1.41	0
B24 Ser	1.9	25	20% HAc/80% H_2O	1.5	323	27	14	0	10	A2-A19, B4-B19	1.76	2.30	0
B24 Gly	1.9	25	20% HAc/80% H_2O	1.5	290	27	14	0	9 [h]	A2-A19, B4-B19	3.01	3.53	0.
des B25	3.0	37	H_2O	4.6	368	10	0	0	20	A2-A20, B2-B28	0.87	1.58	0.
III	1.9	25	20% HAc/80% H_2O	1	422	28	15	0	11	A2-A19, B4-B28	0.97	1.57	0.
B16 His	2.4	24	H_2O	2-5	479	44	23	6	20	A2-A19, B4-B28	0.42	0.85	0.
(B1,B10,B16,B27) Glu, desB30	6.5	34	H_2O	3	300	34	14	5	25	A2-A19, B4-B26	0.59	1.16	0.

(a) Structures of monomeric human insulin and mutants of this determined by NMR and available in the Protein Data Bank, Brookhaven
(b) The experimental conditions includes pH, temperature, solvent, and insulin concentration.
(c) The number of NOEs (intraresidual NOEs excluded) and the number of dihedral angle restraints used in the calculations.
(d) The alignment procedures used in the original publication, and reproduced in the ensemble averages.
(e) The rms deviations listed are calculated as ensemble averages using similar alignments for all molecules, the basis has been backbone heavy atoms (h) respectively. Alignment used A2-A8, A12-A19, B5-B28 (-B23 for DPI), (-B27 for (B1,B10,B16,B27) Glu, des B30).
(f) Alignment used A2-A8, A12-A19, B5-B19
(g) Originally only the B chain helix was aligned. In order to make this average structure comparable to the others alignment of the A ch
(h) 11 structures was published, 9 was deposited.

Table 2.4 Comparison of solution structures of human insulin and mutants[a].

rms (Å)[(a)]	DPI	B24 Ser	B24 Gly	des B25	HI	B16 His	(B1,B10,B16,B27) Glu, des-B30
DPI	-	1.07	1.38	1.26	0.33	1.12	1.15
B24 Ser	0.97	-	1.54	1.14	1.10	1.54	1.35
B24 Gly	1.41	1.54	-	1.36	1.37	1.14	1.18
des B25	1.23	1.14	1.36	-	1.27	1.49	1.35
HI	0.33	1.10	1.37	1.27	-	1.09	1.08
B16 His	1.20	1.54	1.14	1.49	1.09	-	0.84
(B1,B10,B16,B27) Glu, des-B30	1.16	1.35	1.18	1.35	1.08	0.84	-

[a] The root mean square deviations of the backbone were determined aligning the regions A2-A19, B4-B19 above the diagonal and the helix regions A2-A8, A13-A19, B9-B19 below the diagonal. The structures are ensemble averages aligned as listed in Table 2.3.

3. Backbone Dynamics

The three dimensional solution structure of the (B1, B10, B16, B27) Glu, des-B30 insulin mutant at neutral pH was determined and compared to the solution structure of several other insulin mutants determined at acidic pH in Chapter 2. Common of all these insulin structures is the conservation of the secondary and tertiary structure in the core of the molecule. However, the overall precision of the structures in the conserved region as well as in the rest of the molecule, the B-chain terminal parts, differs widely. The neutral pH structure as compared with the other structures showed the resemblance to the B16 His mutant determined at acidic pH (Ludvigsen et al., 1994, Kaarsholm & Ludvigsen, 1995) to be the best. However, although the overall precision of the neutral pH structure is second best among all structures (0.52 Å for the backbone in the A2-A8, A12-A19, B5-B27 region) there is still a gap to the best, the B16 His mutant (0.33 Å in the same region). The difference in number of distance restaints (observed NOEs) used in the two structure calculations explains part, at least, of the difference between the overall precisions of the final structures. To investigate whether this lack of NOEs is directly related to a more pronounced inherent internal dynamic of the insulin mutant at neutral pH, the relaxation properties of the backbone ^{15}N nuclei of both mutants are examined here. In 'the model free approach' introduced by Lipari & Szabo (1982a,b) the ^{15}N relaxation parameters, T_1, T_2, and the $\{^1H^{15}N\}$ NOE, are translated to a generalized order parameter, S^2, describing the spatial restriction of $^1H^{15}N$-bond motions on the picoseconds-nanoseconds time scale. Motions on the milliseconds timescale are mapped as well.

3.1 Theory of relaxation

The relaxation mechanism of the amide ^{15}N nucleus in a protein backbone is a combination of dipolar interactions with the directly attached amide proton and chemical shift anisotropy. Compared to the relaxation mechanism of the proton it is pleasingly simple. The longitudinal relaxation rate R_1 and the transversal relaxation rate R_2, the reciprocal of T_1 and T_2 respectively, are simply sums of the terms for dipolar interaction (DD) and chemical shift anisotropy (CSA), cross correlation between the two mechanisms is neglected.

$$R_1 = \frac{1}{T_1} = \frac{1}{T_1^{DD}} + \frac{1}{T_1^{CSA}} \qquad (1)$$

$$R_2 = \frac{1}{T_2} = \frac{1}{T_2^{DD}} + \frac{1}{T_2^{CSA}} + \pi\,\Delta_{ex} \qquad (2)$$

The extra term added in eq 2 accounts for contributions to the transversal relaxation from pseudo first order processes, i.e., conformational exchange, the time scale is milliseconds to seconds. The relaxation rates and the ^1H-^{15}N nuclear Overhauser effect, NOE, are related to the spectral density function, J (Abragam, 1961, Lipari & Szabo, 1982a,b). The spectral density function is the

Fourier transformed correlation function of the motion of a unit vector along the NH bond. The spectral density function is thus the "spectrum" of the characteristic reorientations giving the frequency of these.

$$\frac{1}{T_1} = \frac{d^2}{4}(J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)) + \frac{c^2}{3}J(\omega_N) \quad (3)$$

$$\frac{1}{T_2} = \frac{d^2}{8}(4J(0) + 4J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N))$$

$$+ \frac{c^2}{18}(3J(\omega_N) + 4J(0)) \quad (4)$$

$$NOE = 1 + T_1 \frac{\gamma_H}{\gamma_N}\frac{d^2}{4}(6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)) \quad (5)$$

with the constants c and d given by

$$d = \frac{h\mu_0}{8\pi^2}\frac{\gamma_H\gamma_N}{r_{NH}^3}$$

$$c = \frac{\omega_N(\sigma_{||} - \sigma_\perp)}{3} \quad (6)$$

h is Planck's constant ($6.626*10^{-34}$ Js), μ_0 is the vacuum permeability ($4\pi*10^{-7}$ TmA^{-1}), γ_H and γ_N are the gyromagnetic ratios of the two nuclei ($2.6752*10^8$ and $-2.7108*10^7$ s^{-1} T^{-1}), r_{NH} is the bond length (1.02 Å), and $\sigma_{||} - \sigma_\perp$ is the difference between the components of the chemical shift tensor parallel and perpendicular respectively to the field ($-160*10^{-6}$, Hiyama et al., 1988).

In the model free approach introduced by Lipari & Szabo (1982a,b) J is approximated by introduction of a generalized order parameter, S^2, and an effective correlation time for the fast internal motion, $\tau_e < 0.3$ nanoseconds. The overall tumbling of the molecule is assumed to be isotropic, i.e., characterized by a single correlation time, τ_r, in the nanoseconds time scale.

$$J = \frac{2}{5}\left(S^2\frac{\tau_r}{1+\omega^2\tau_r^2} + (1-S^2)\frac{\tau}{1+\omega^2\tau^2}\right)$$

$$\text{where } \tau = \frac{\tau_r\tau_e}{\tau_r + \tau_e} \quad (7)$$

The generalized order parameter is a measure of the degree of spatial restriction of the NH-vector motion. S^2 takes the value of 1 for complete restriction of the motion and the value of 0 for completely unrestricted motion. The strength of this approach is that these parameters are introduced without employing a specific model for the internal motion. Figure 3.1 is a graphical

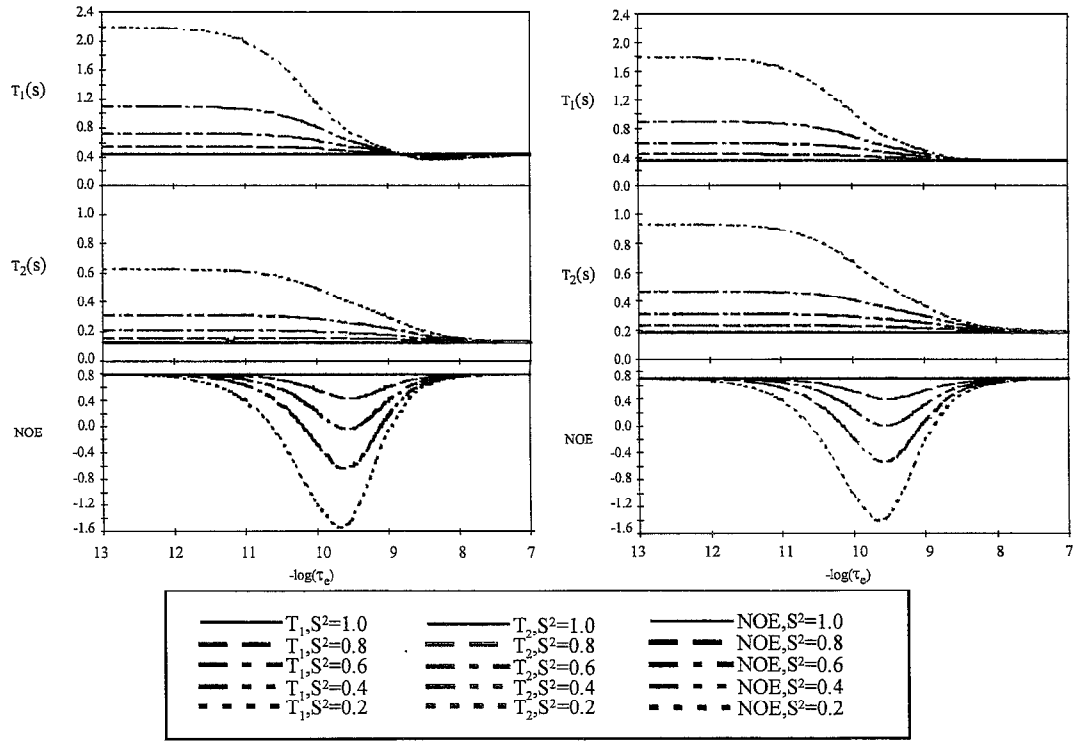


Figure 3.1 T_1, T_2, and NOE vs. the correlation time, τ_e, for internal motion of the NH-vector. Each graph shows the dependency on the general order parameter, S^2, at five different values. The left hand side of the figure corresponds to τ_r=4.98 ns, as for the insulin mutant at acidic pH, the right-hand side corresponds to τ_r=2.86 ns, as for the insulin mutant at neutral pH.

representation of the relaxation rate constants T_1 and T_2 and the $\{^1H^{15}N\}$ NOE vs. the effective correlation time for internal motion, τ_e. Two different correlation times for overall rotational tumbling are shown corresponding to the values found for the two insulin mutants examined in the following section. The data has been calculated (eq 3-5,7), disregarding conformational exchange line broadening, for five different values of S^2 ranging from completely restricted motion, S^2=1.0, to almost unrestricted motion, S^2=0.2.

Inspection of the graphs shows that in the time range from 1 to about 50 picoseconds and for non-extreme S^2 values (0.6-1.0), the T_1 and T_2 graphs are independent of the exact value of τ_e and the NOE is at a maximal value dependent only on τ_r. This situation corresponds to a reduced form of the spectral density function where the dependence of the correlation time for the internal motion, τ_e, is lost.

$$J_{red} = \frac{2}{5}\left(\frac{S^2\tau_r}{1+\omega^2\tau_r^2}\right) \quad (8)$$

Note again that the simplified spectral density function is applicable in the time range <50 ps, giving T_1 and T_2 values dependent only on S^2 (and τ_r) and a NOE at its theoretical maximum independent of S^2. Figure 3.3 is a slice of Figure 3.1 in the 1-50 picoseconds region in the S^2 dimension to show T_1 and T_2 (eq 3-4) calculated using the reduced spectral density function (eq 8). The NOE is not shown.

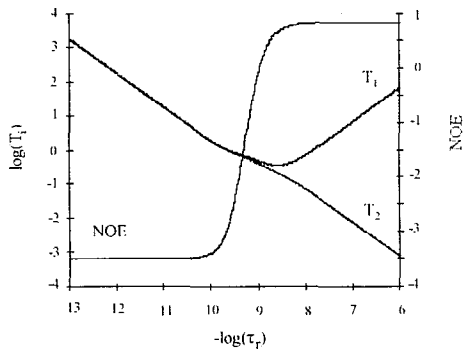


Figure 3.2 Relaxation rate constants T_1 and T_2 and ^1H-^{15}N NOE for the ^{15}N nucleus (eq 3-5), vs. the correlation time, τ_r, for overall isotropic tumbling of the molecule. T_1, T_2, and τ_r are measured in seconds. The reduced spectral density function (eq 8) was used.

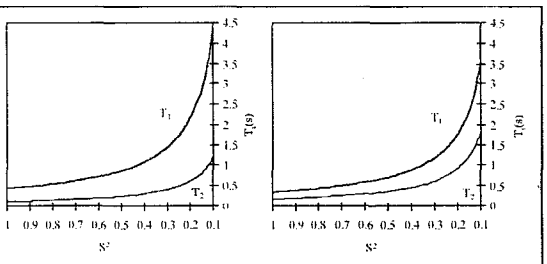


Figure 3.3 Relaxation rate constants T_1 and T_2 vs. the general order parameter, S^2, for the internal motion of the NH-vector. The reduced spectral density function, (eq 8) was used for two different τ_r values. The left hand side corresponds to τ_r =4.98 ns, as for the insulin mutant at acidic pH, giving a theoretical $^1H^{15}N$ NOE of 0.79. The right hand side corresponds to τ_r =2.86 ns, as for the insulin mutant at neutral pH, giving a theoretical $^1H^{15}N$ NOE of 0.70.

Figure 3.2 is the classical picture of the relaxation times and the NOE vs. the correlation time for overall isotropic tumbling (eq. 3-5,8) fixing S^2 to 1, i.e., assuming total restriction of the internal motions. The correlation times for proteins assume nanoseconds values giving a positive NOE and T_1 and T_2 in the range of 0.1-1.0 s.

An extension to the model free approach was introduced by Clore et al. (1990a,b) taking internal motions on two different time scales into account. The slower motion is described by a generalized order parameter, S_s^2, and a correlation time, τ_s, the rapid motion is described by a generalized order parameter, S_f^2.

$$J_{S_f} = \frac{2}{5} S_f^2 \left(\frac{S_s^2 \tau_r}{1+\omega^2\tau_r^2} - \frac{(1-S_s^2)\tau}{1+\omega^2\tau^2} \right)$$

$$\text{where } \tau = \frac{\tau_r \tau_s}{\tau_r + \tau_s} \qquad (9)$$

In order to be resolvable in an NMR experiment, the two correlation times must differ by at least 1-2 orders of magnitude, i.e., τ_s is in the nanoseconds time scale whereas the faster motion is characterized by a correlation time < 10 ps. Note that this extended spectral density function (eq 9) reduces to the original function (eq 6) if S_f^2 takes the value of 1.

3.1.1 Models

The model free approach extracts dynamic information from the relaxation rate constants and the NOE. The output is an order parameter, S^2 generalized to cover the broad range of possible motions, and a corresponding correlation time, τ_e. A feasible interpretation of S^2 in the relaxation of backbone amides is the random motion of the NH-vector within a cone of semiangle θ as shown in Figure 3.5, (Clore et al., 1990a). S^2 is related to θ through

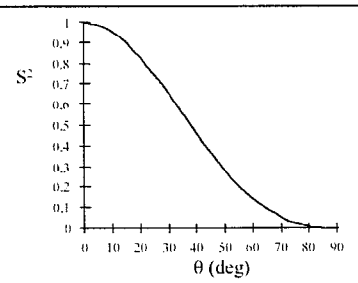


Figure 3.4 Generalized order parameter S^2 vs cone semiangle θ according to eq 10.

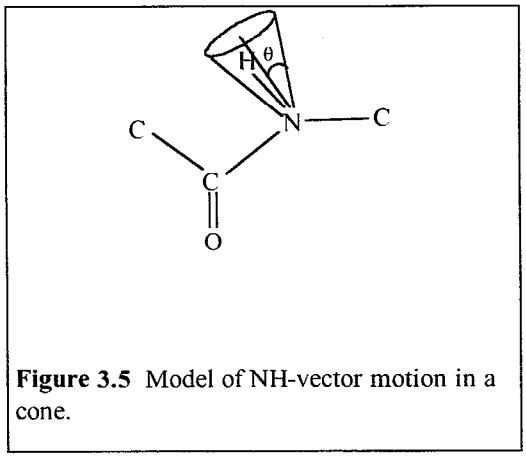


Figure 3.5 Model of NH-vector motion in a cone.

$$S^2 = (0.5\cos\theta(1+\cos\theta))^2 \quad (10)$$

A graphical representation of eq 10 is shown in Figure 3.4.

The extended spectral density function (eq 9) introduces an extra order parameter. Clore et al. (1990a) suggest that the fast motion corresponds to the free diffusion in a cone (eq 10) and the slower motion to jumps between different orientations of the cone.

3.2 Materials and Methods

3.2.1 Materials

Two mutants of human insulin were constructed by oligonucleotide-directed muta-genesis. The rationale behind engi-neering these mutants has been described elsewhere (Ludvigsen et al., 1994, Chapter 2) The fermentation in yeast, using $^{15}NH_4Cl$ as the source of ^{15}N, gave uniformly enriched samples which was purified as described (Markussen et al., 1987; Brange et al., 1988). The mutants, B16 His, des-B30 and (B1, B10, B16, B27) Glu, des-B30 were dissolved in 80% H_2O / 20% D_2O to a concentration of approximately 2.5 mM at pH=2.4, 24°C, and at pH=6.5, 34°C, respectively.

The three dimensional structure of the B16 His mutant has been determined by NMR (Ludvigsen et al., 1994, Kaarsholm & Ludvigsen, 1995) and the result is deposited in The Protein Data Bank, Brookhaven. The three dimensional structure of the (B1, B10, B16, B27) Glu, des-B30 mutant was presented in Chapter 2. The average structures shown here has been calculated from ensembles of structures aligning the regions A2-A19 and B4-B27 simultaneously.

3.2.2 NMR Spectroscopy

The measurement of relaxation rate constants of the insensitive ^{15}N nucleus and $\{^1H^{15}N\}$ NOE is accomplished by aquisition of series of two dimensional heteronuclear experiments. The pulse sequences used herein were first introduced by Kay et al. (1989) as a variation of the original $^1H^{15}N$ correlation experiment by Bodenhausen & Ruben (1980), with an improvement to avoid cross correlation between the two relaxation mechanisms, i.e., dipolar coupling

and chemical shift anisotropy, proposed by Kay et al. (1992). The experiments are designed to avoid the insensitivity of the ^{15}N nucleus, the magnetogyric ratio is about 1/10 of that of ^{1}H (and negative), to prolong the experimental time. Refocused INEPT sequences (Morris & Freeman 1979, Burum & Ernst 1980) use the spin-spin coupling to transfer magnetization originating at the ^{1}H to the directly bound ^{15}N in the preparation period and to return it to ^{1}H for observation . In the t_1 period, the magnetization is modulated with the ^{15}N chemical shift. Detection of ^{1}H therefore results in a correlation of the ^{15}N and ^{1}H frequencies in the Fourier transformed spectrum. For the measurement of relaxation rate constants, the relaxation delay in the pulse sequences is varied through the series of spectra, resulting in a single exponential decaying intensity of each ^{1}H ^{15}N crosspeak with a time constant of T_1 and T_2, respectively. In the T_2 series the relaxation delay is of the Carr-Purcell-Meiboom-Gill (CPMG) type (Carr & Purcell, 1954; Meiboom & Gill, 1958). The $\{^{1}H^{15}N\}$ NOE is measured as the ratio of signal intensities from spectra recorded in the presence and absence of ^{1}H presaturation, respectively.

The assignments of the ^{15}N resonances of the two insulin mutants were based on three dimensional HMQC-TOCSY (Marion et al., 1989a, Zuiderweg et al., 1989) and HMQC-NOESY spectra (Marion et al., 1989b). These experiments are the 3D versions of the original TOCSY (Braunschweiler & Ernst, 1983, Bax & Davis, 1985) and NOESY (Jeener et al., 1979; Anil-Kumar et al., 1980, 1981) experiments. The addition of an ^{15}N dimension enables the correlation of each amide proton to the backbone ^{15}N nucleus and overcomes part of the problems with overlapping crosspeaks in the 2D spectra. The magnetization originates at the protons and is modulated by their chemical shift in the t_1 period, the TOCSY or NOESY mixing time is followed by a ^{15}N chemical shift modulation in the t_2 period, and finally the amide proton chemical shifts are observed. Thus, the three dimensions of these spectra are, amide protons, ^{15}N nuclei and all protons.

The data was recorded on a Bruker AMX600. For the 2D experiments 256 t_1 slices consisting of 2048 real data points were sampled using the time-proportional phase incrementation scheme (TPPI) (Marion & Wüthrich, 1983). With 16 scans in a t_1 slice, the experimental time was approximately 4 hours for each spectrum. The spectral width in the ^{1}H dimension was 6579 Hz and the carrier was placed exactly on the water resonance to enable irradiation between the scans. In the ^{15}N dimension, the spectral width was 1225 Hz. The 3D experiments were sampled as cubes of 1024 data points in the t_3 dimension, amide protons, 64 data points in the t_2 dimension, ^{15}N nuclei, and 256 data points in the t_1 dimension, all protons.

For the T_2 series the following relaxation delays were used 9, 17, 34, 43, 56, 69, 77, 90, 98, 116, 150 and 300 ms. For the T_1 series the relaxation delays employed were 10, 50, 100, 200, 300, 400, 500, 700 and 1000 ms. T_1 as well as T_2 were determined as the average of two series. The NOE was determined as an average of three pairs of spectra. For the INEPT sequences, the delays were adjusted to give maximal transfer of magnetization, i.e., $(^{1}/_{4} \times {}^{3}J_{^{1}H^{15}N})$, ${}^{3}J_{^{1}H^{15}N}$ at a value of 92 Hz.. The mixing time of the HMQC-

Model	Spectral density function	Input parameters	Output parameters
A	J (7)	T_1,T_2,NOE	S^2, τ_e
B	J (7), $R_2=R_2+\pi\Delta_{ex}$	T_1,T_2,NOE	S^2, τ_e, Δ_{ex}
C	J_{sf} (9)	T_1,T_2,NOE	S_f^2, S_s^2, τ_s
D	J_{red} (8)	T_1,T_2	S^2
E	J_{red} (8), $R_2=R_2+\pi\Delta_{ex}$	T_1,T_2	S^2, Δ_{ex}

Table 3.1 Description of the five different models used to fit the generalized order parameters, S^2 or S_s^2 and S_f^2 and the correlation times, τ_e or τ_s, for the ^{15}N-^1H vector motion as well as the exchange linebroadening of the ^{15}N resonance to the ^{15}N relaxation times, T_1 and T_2, and the $\{^1H^{15}N\}$ NOE. The spectral density functions listed refer to the text. Each fit was performed minimizing a target function, a sum of squared residuals of the parameters involved, i.e., for the models A, B, C $\left(\frac{T_1^{calc}-T_1^{obs}}{T_1^{obs.err}}\right)^2+\left(\frac{T_2^{calc}-T_2^{obs}}{T_2^{obs.err}}\right)^2+\left(\frac{NOE^{calc}-NOE^{obs}}{NOE^{obs.err}}\right)^2$ and for the models D, E $\left(\frac{T_1^{calc}-T_1^{obs}}{T_1^{obs.err}}\right)^2+\left(\frac{T_2^{calc}-T_2^{obs}}{T_2^{obs.err}}\right)^2$.

NOESY experiment was 120 ms and that of the HMQC-TOCSY was 65 ms.

The processing of the data was performed using the processing package MNMR (PRONTO Software Development and Distribution, Copenhagen, Denmark) or UXNMR (Bruker) on a SGI Indigo^2. Prior to Fourier Transformation, the data was zero filled once in each dimension and a gaussian window function (-7 Hz, factor of 0.15) was applied.

3.2.3 Analysis of the relaxation times, T_1 and T_2, and the NOE

Eq 3-5 relate the generalized order parameter and the correlation time for internal motion to the reciprocal relaxation rates T_1, T_2 and the $\{^1H^{15}N\}$NOE through the spectral density function. The spectral density function is defined in three different levels in eq 7-9. The first step in the analysis is to determine the overall rotational correlation time, τ_r. Under the assumption that the overall motion of the protein is the same for all residues, the ratio T_1/T_2 depends only on the correlation time for the overall motion, τ_r and is constant throughout the molecule (Kay et al., 1992),. Residues with the line broadened by conformational exchange are disregarded. The next step is to fit the parameters describing the dynamic of each NH-vector to the values measured for T_1, T_2, and NOE through the spectral density function. The generalized order parameter, S^2, is determined for all residues, and in most cases the correlation time for internal motion, τ_e. For some residues a linebroadening factor is determined in addition. The five different models employed are listed in Table 3.1.

The first choice of model is A. where all available data are fitted simultaneously to a generalized order parameter, S^2, and a corresponding correlation time, τ_e. For a broad range of τ_e values, the NOE is invariant at its maximal value and the values of T_1 and T_2 are independent of the correlation time, τ_e. This situation is allowed for in model D. The models B or E are used for residues characterized by a small T_2, i.e., a T_1/T_2 ratio well above the average. Both models include a linebroadening term in the T_2 equation. The last model, C, covers motions in two different time scales, a fast time scale characterized by the order parameter S_f^2, on the assumption that the corresponding correlation time is < 10 ps, and a slower time scale characterized by the order parameter S_s^2 and the correlation time τ_s.

The least squares fits were performed using a grid search algorithm programmed in Fortran 77. For each residue a grid spanning the range of possible values for each parameter was searched to determine the global minimum of the penalty function. The penalty function for each model is listed in the legend to Table 3.1. The overall rotational correlation time, τ_r, was kept at a constant value in the grid search. The standard deviations of the results were calculated as described in the supplements section at the end of this Chapter.

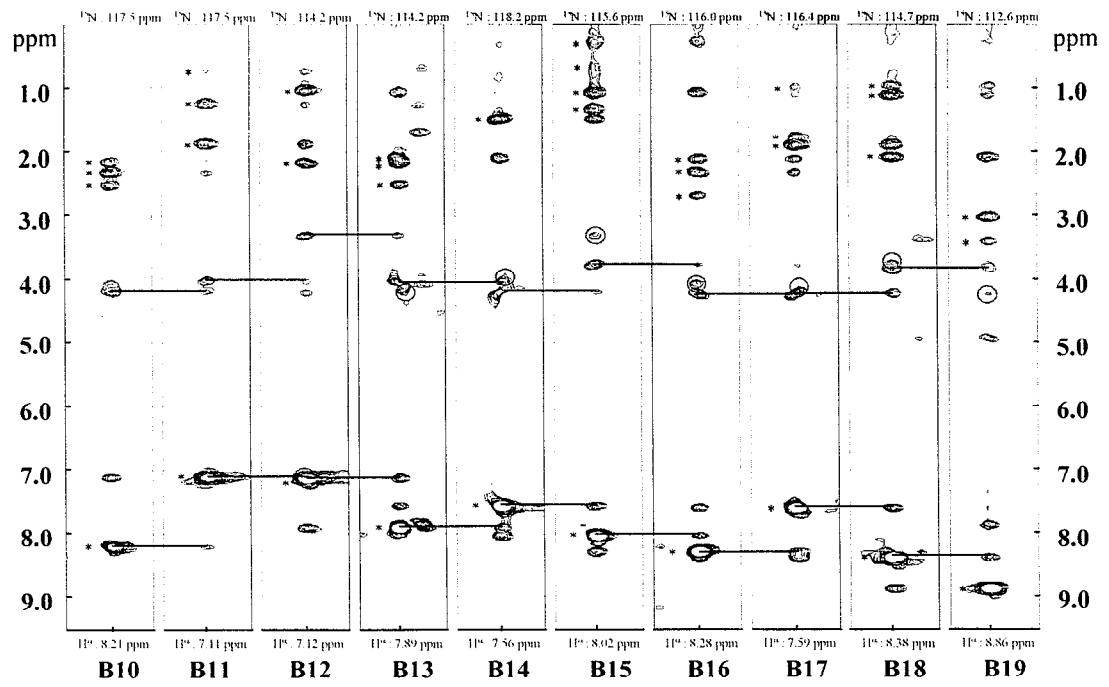


Figure 3.6 Slices from a 3D HMQC-NOESY spectrum (mixing time 150 ms) of the insulin mutant at neutral pH. The slices show the intra- and interresidual correlations to the amide proton in the ^{15}N resonance slice of each residue. The lines indicate sequential contacts H^N-H^N and H^α-H^N, intra-residual crosspeaks are marked with * and blue circles surround the helix-characteristic crosspeaks correlating H^α of residue i to H^N of residue i+1 NOEs.

3.3 Results

Analysis of the 3D HMQC-TOCSY and HMQC-NOESY spectra gave a straigthforward assignment of the ^{15}N spectra. Figure 3.6 shows ^{15}N slices from the HMQC-NOESY spectrum of the insulin mutant at neutral pH in the helix-region B10-B19. The sequential assignment of residues in this region was based on the strong H^N-H^N connectivities as well as the weaker H^α-H^N connectivities. The ^{15}N chemical shifts are included in Table 3.2 at the end of this Chapter.

The relaxation rate times and the $\{^1H\ ^{15}N\}$NOE were measured for most residues of the two mutants of human insulin. Figure 3.8 depicts data from a series of spectra in the T_2 measurement series of the insulin mutant at neutral pH. Each excerpt shows the signals of the residues B11 and B12. The intensities were extracted as maximal peakheights in an area covering each signal giving the single-exponential decaying curves shown in Figure 3.7.

Figure 3.11 gives the complete set of T_1, T_2, and NOE measured for the two insulin mutants. The numbers are listed in Table 3.2. For the mutant at neutral pH the N-terminal residues of the B-chain showed cross peaks just above the noise level of the spectra making peak intensity measurements impossible.

Figure 3.9 shows the T_1/T_2 -ratio by residue for the two insulin mutants. Note the different scales of the two data sets. Marked with black are those residues with a ratio falling more than one standard deviation outside the average of all residues. Above the average are residues with

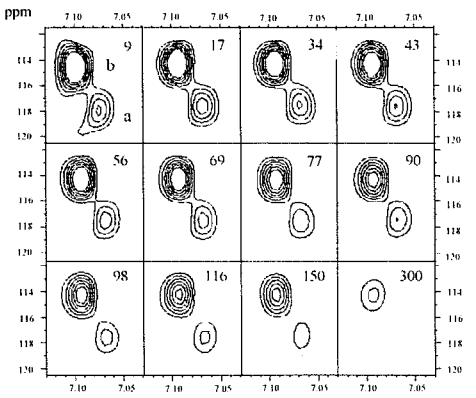


Figure 3.8 Sections from 12 ^{1}H-^{15}N 2D spectra in a T_2 measurement series marked with the delay time (ms) used in each experiment. Each section shows the amide peak from the two residues B11 (a) and B12 (b). The contouring levels are equidistant starting from the same lower level in all spectra.

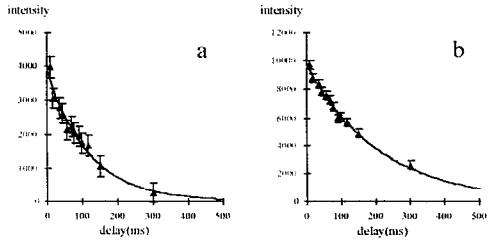


Figure 3.7 Peakheights extracted from the T_2 series shown in Figure 3.8. The same arbitrary units are used in the intensity scales of both graphs and the noise level is indicated as error bars. The lines show the single exponential fit employed to extract T_2 . Residue B11 is shown in panel (a) and residue B12 in panel (b).

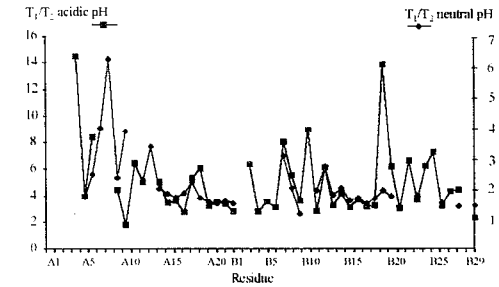


Figure 3.9 The ratio T_1/T_2 for the two insulin mutants. Marked with black are those residues for which the ratio falls outside the average by more than 1 standard deviation.

lines broadened by conformational exchange and below the average are residues with unusually low T_1 values.

For the insulin mutant at neutral pH, the τ_r value was calculated from the average of all residues excluding the ones marked with black in Figure 3.9. Inspection of the data of the insulin mutant at acidic pH shows a wider dispersion of the T_1/T_2 -ratio, as a consequence of a larger number of residues involved in conformational exchange. All of these are excluded from the average calculation. The τ_r values determined for the two insulin mutants was 2.86 ± 0.21 ns and 4.98 ± 0.49 ns at neutral and acidic pH, respectively.

Figure 3.10 is a graphical representation of a grid search using Model A. The area spans S^2 and τ_e and the penalty function (the third dimension) involves T_1, T_2, and NOE. Obviously, for this residue the global minimum of S^2 is well defined in contrast to the global minimum of τ_e. The proper choice in this case is Model D, corresponding to a one dimensional search in the S^2 dimension.

In all attempts to use model C the correlation coefficients between the three parameters were ≈ 1.0, consequently these results were discarded. The ratio between DD and CSA was approximately 0.8/0.2 for all residues.

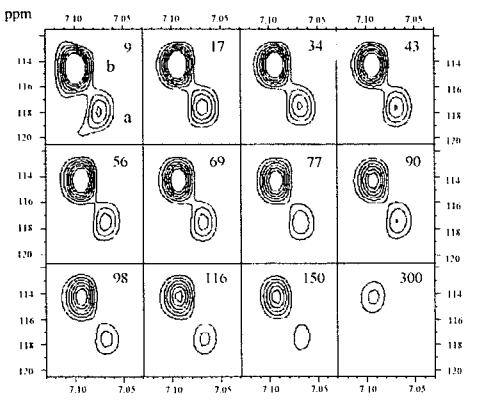


Figure 3.8 Sections from 12 ^{1}H-^{15}N 2D spectra in a T_2 measurement series marked with the delay time (ms) used in each experiment. Each section shows the amide peak from the two residues B11 (a) and B12 (b). The contouring levels are equidistant starting from the same lower level in all spectra.

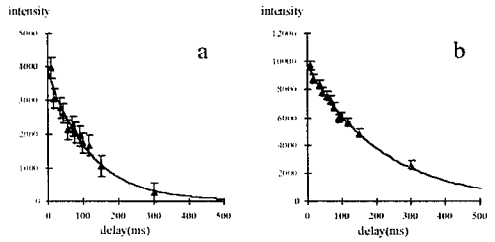


Figure 3.7 Peakheights extracted from the T_2 series shown in Figure 3.8. The same arbitrary units are used in the intensity scales of both graphs and the noise level is indicated as error bars. The lines show the single exponential fit employed to extract T_2 . Residue B11 is shown in panel (a) and residue B12 in panel (b).

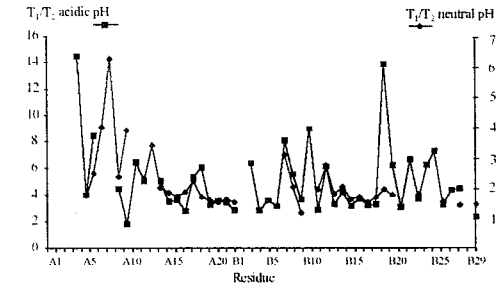


Figure 3.9 The ratio T_1/T_2 for the two insulin mutants. Marked with black are those residues for which the ratio falls outside the average by more than 1 standard deviation.

lines broadened by conformational exchange and below the average are residues with unusually low T_1 values.

For the insulin mutant at neutral pH, the τ_r value was calculated from the average of all residues excluding the ones marked with black in Figure 3.9. Inspection of the data of the insulin mutant at acidic pH shows a wider dispersion of the T_1/T_2 -ratio, as a consequence of a larger number of residues involved in conformational exchange. All of these are excluded from the average calculation. The τ_r values determined for the two insulin mutants was 2.86 ± 0.21 ns and 4.98 ± 0.49 ns at neutral and acidic pH, respectively.

Figure 3.10 is a graphical representation of a grid search using Model A. The area spans S^2 and τ_e and the penalty function (the third dimension) involves T_1, T_2, and NOE. Obviously, for this residue the global minimum of S^2 is well defined in contrast to the global minimum of τ_e. The proper choice in this case is Model D, corresponding to a one dimensional search in the S^2 dimension.

In all attempts to use model C the correlation coefficients between the three parameters were ≈ 1.0, consequently these results were discarded. The ratio between DD and CSA was approximately 0.8/0.2 for all residues.

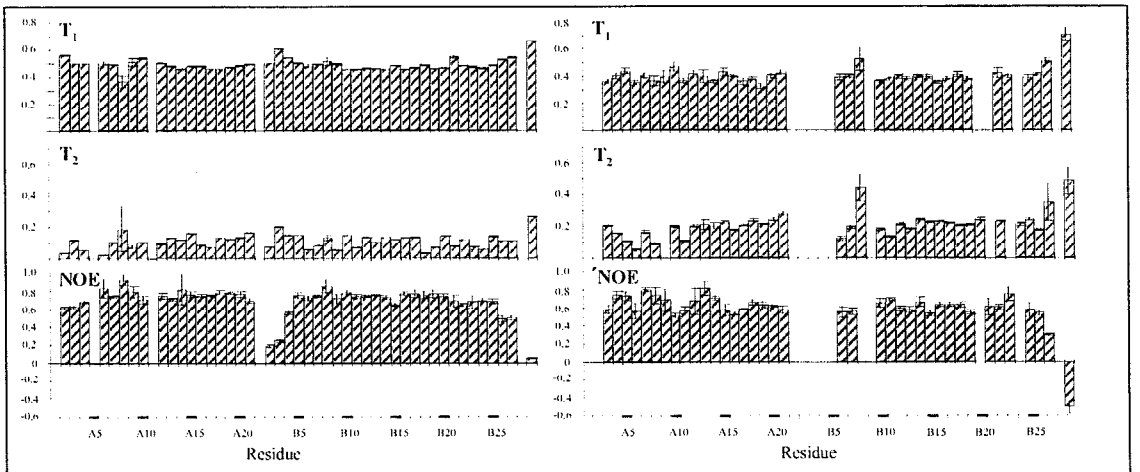


Figure 3.11 ^{15}N relaxation parameters, T_1 and T_2 measured in seconds, and $^1H^{15}N$ NOE for two insulin mutants. The relaxation times are measured in seconds. The left-hand side panel shows the insulin mutant at acidic pH, and the right-hand side panel shows the insulin mutant at neutral pH.

The results of the fit covering two different time scales are tabulated in Table 3.3. Figure 3.12 depicts the picosecond dynamics represented by the generalized order parameter together with the rms deviations of the backbone of an ensemble of structures. The eye is immediately caught by the fine accordance between the the two mappings.

For the B-chain residues the picoseconds dynamics in the termini are reflected in the precision of the calculated structures. Moreover, the decline of the curves in the loop regions flanking the central B-helix, B19-B22 and B7-B9, is a common trend. The overall level of the order parameters for the mutant at neutral pH is below that of the mutant at acidic pH, the same thing goes for an comparison of the rms deviations of the two mutants. The B-chain helix shows a region of higher mobility around residue B15, which is not reflected in rms deviations.

The two A-chain helices share the same characteristics, i.e, higher mobility in the termini than in the middle of the helix. The loop region is

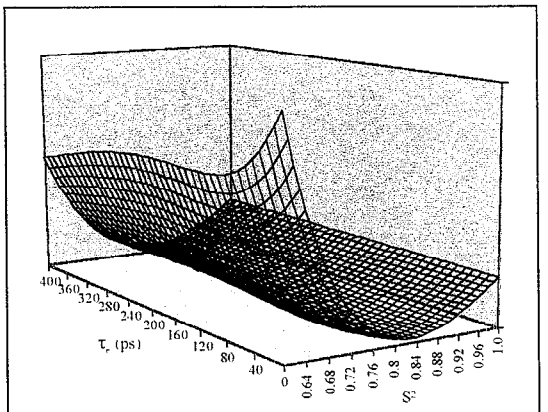


Figure 3.10 Section of a S^2 ,τ_c surface spanning the global minimum area. The vertical axis shows the penalty function. The data represents the fit of T_1, T_2, and NOE for ^{15}N in residue A15 of the insulin mutant at neutral pH using model A.

more mobile than the helices. The most striking difference between the two mutants is the overall stability of the helices. For both mutants the order parameter level of the A(II)-helix and the B-helix is similar. However, the mobility of these helices compared to the A(I)-helix is not the same. At acidic pH, the A(I)-helix shows higher mobility whereas at neutral pH the mobility of the A(I)-

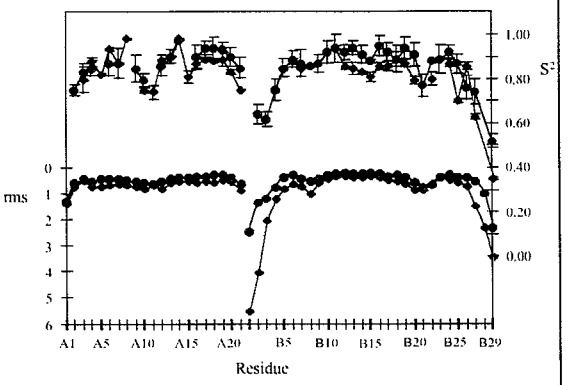


Figure 3.12 Dynamics in the picoseconds time scale and precision of the three dimensional structure. General order parameter for internal motion of the NH-vectors (top) and rms of the backbone (bottom). Data from two insulin mutants is shown, B16 His, des-B30 mutant at acidic pH (red) and (B1, B10, B16, B27) Glu, des-B30 mutant at neutral pH (green).

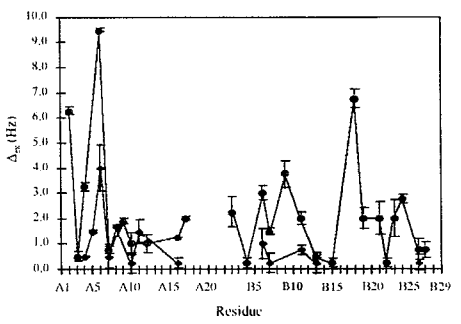


Figure 3.13 Dynamics in the milliseconds time scale. Linebroadening of the ^{15}N resonance caused by conformational exchange. Data from two insulin mutants is shown, B16 His, des-B30 mutant at acidic pH (red) and (B1, B10, B16, B27) Glu, des-B30 mutant at neutral pH (green).

helix is lower. As was the case for residue B15, the region around residue A15 also shows unexpectedly high mobility (when comparing to the rms).

The millisecond dynamics represented by broadening of the ^{15}N resonances (Figure 3.13) show resemblances between the two pH values in the A(I)-helix and in the A-chain loop region both of which are characterized by linebroadenings. At acidic pH, additional linebroadenings are scattered throughout the structure.

3.4 Discussion

The relaxation properties of the ^{15}N nuclei in the backbone of two mutants of human insulin B16 His, des-B30 and (B1, B10, B16, B27) Glu, des-B30 in aqueous solution have been under investigation. The biological activity of the mutants is 43 % and 47 %, respectively. These two mutants have been engineered to overcome complications resulting from aggregation of insulin molecules in different pH regions. Thus the (B1, B10, B16, B27) Glu mutant was investigated at near neutral pH 6.5 (34°C), and the B16 His mutant at acidic pH 2.5 (24°C). The three dimensional structures of the mutants have been determined prior to this investigation (Ludvigsen et al., 1994, and Chapter 2). The structures show slight differences mostly confined to the termini of the B-chain. These differences arise due to the insertion of the charged Glu sidechains in the (B1,B10,B16,B27) Glu, des-B30 mutant and are therefore only indirectly an

effect caused by a change in pH. The difference of overall precision between the two structures are more pronounced. The aim of the analysis of the ^{15}N relaxation properties of the two insulin mutants has been to map the backbone dynamics at different solution acidity and to examine the relation to the overall precision of the calculated structures.

The first result was the determination of the correlation time, τ_r, of the overall tumbling of the protein. Stokes law relates the frictional constant, f, to the effective radius of the protein, r_{eff}, through $f=6\Pi r_{eff}\eta$, where η is the viscosity of the solvent. τ_r is proportional to f and therefore dependent on the effective radius of the protein as well as on the viscosity of the solvent. τ_r for the two insulin mutants at neutral pH, 24°C and acidic pH, 34°C was surprisingly different, 2.86 ± 0.21 ns and 4.98 ± 0.48 ns, respectively. Part of the difference of τ_r is caused by the different temperatures employed. The viscosity of the solvent (water) at the two temperatures employed here are $\eta_{24°C}$=0.911cp and $\eta_{34°C}$=0.7340 cp (Hand-book of Chemistry and Physics, Chemical Rubber Company). With offset in τ_r measured for the insulin mutant at neutral pH, 34°C, τ_r for the mutant at acidic pH corrected to the same temperature takes a value of approximately 3.55 ns. The pH of the solvent influences the viscosity although this contribution is minor. Upon dissolution of the lyophilized insulin mutant in water to a concentration of 2.5 mM the difference in addition of HCl required to adjust the pH to 2.5 and 6.5, respectively, is approximately 15 mM. The effect of this is a slightly increased viscosity of the solvent in the acidic solution, < 1 % (Handbook of Chemistry and Physics, Chemical Rubber Company). Thus the difference in viscosity of the solvent of the two investigations can not explain all of the difference between the measured τ_r values. The attention is therefore concentrated on the effective radii of the two insulin mutants. Apparently r_{eff} of the acidic pH mutant is larger than r_{eff} of the neutral pH mutant. As mentioned above, the structural studies revealed only small differences in the three dimensional structure of the two mutants. The most prominent difference is the poorly defined termini of the B-chain of the mutant observed at neutral pH (Figure 3.15), which qualitatively preserves or even increaces the radius of this mutant. The net relative charge of the two mutants is the same but oppositely signed (-5.5 at neutral pH and +6 at acidic pH), and the distribution of the charges in the three dimensional space are almost identical, but the effect of the different ionic strengths of the solutions on the hydrodynamic volume of the two mutants is not straightforward to account for. Thus, from NMR data alone, it is not possible to explain the different correlation times measured here.

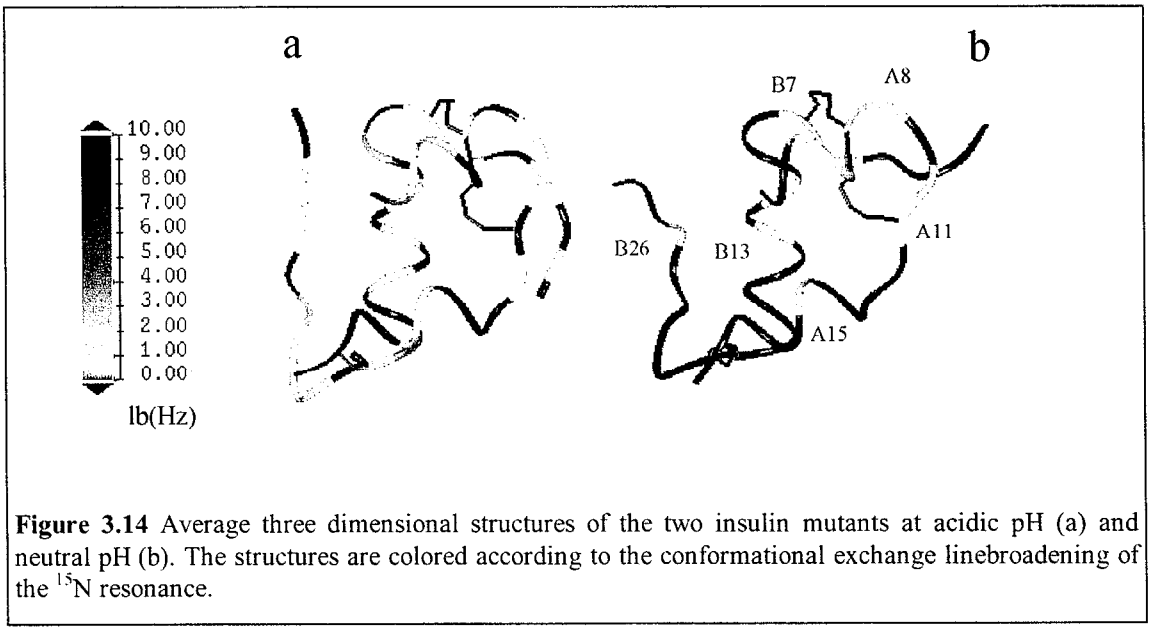


Figure 3.14 Average three dimensional structures of the two insulin mutants at acidic pH (a) and neutral pH (b). The structures are colored according to the conformational exchange linebroadening of the ^{15}N resonance.

As was shown in Figure 3.13 the two insulin mutants show different patterns in the linebroadening of the ^{15}N resonances. At both pH values the A(I)-helix, and especially the A5 and the A6 residues as well as part of the loop region connecting the two helices in the A-chain show millisecond time scale dynamics. Figure 3.14 shows the line-broadening data as colorcodes for the three dimensional structures. Focusing on the neutral pH structure the linebroadenings are obviously located in the distinct three dimensional area surrounding the inter-chain disulfide bridge A7-B7 and the intra-chain disulfide bridge A6-A11. The origin of these linebroadenings is therefore proposed to be disulfide bond isomerizations taking place in the millisecond to second time scale (see e.g. Otting et al., 1993) thus giving rise to broadening of the ^{15}N resonance lines. The ^{1}H resonance lines in the same areas are broadened as well (Ludvigsen et al., 1994). At acidic pH, the A-chain dynamics propagate to the well defined part of the N-terminus of the B-chain which is close in space at this pH. Apparently, the milliseconds time scale dynamic is more pronounced at acidic pH and furthermore all secondary structural elements of the molecule are involved. The fact that the residues exhibiting dynamics in the milliseconds regime are not concentrated near the mutation sites but scattered throughout the structure implies that this phenomenon is intrinsic to the pH (or the temperature).

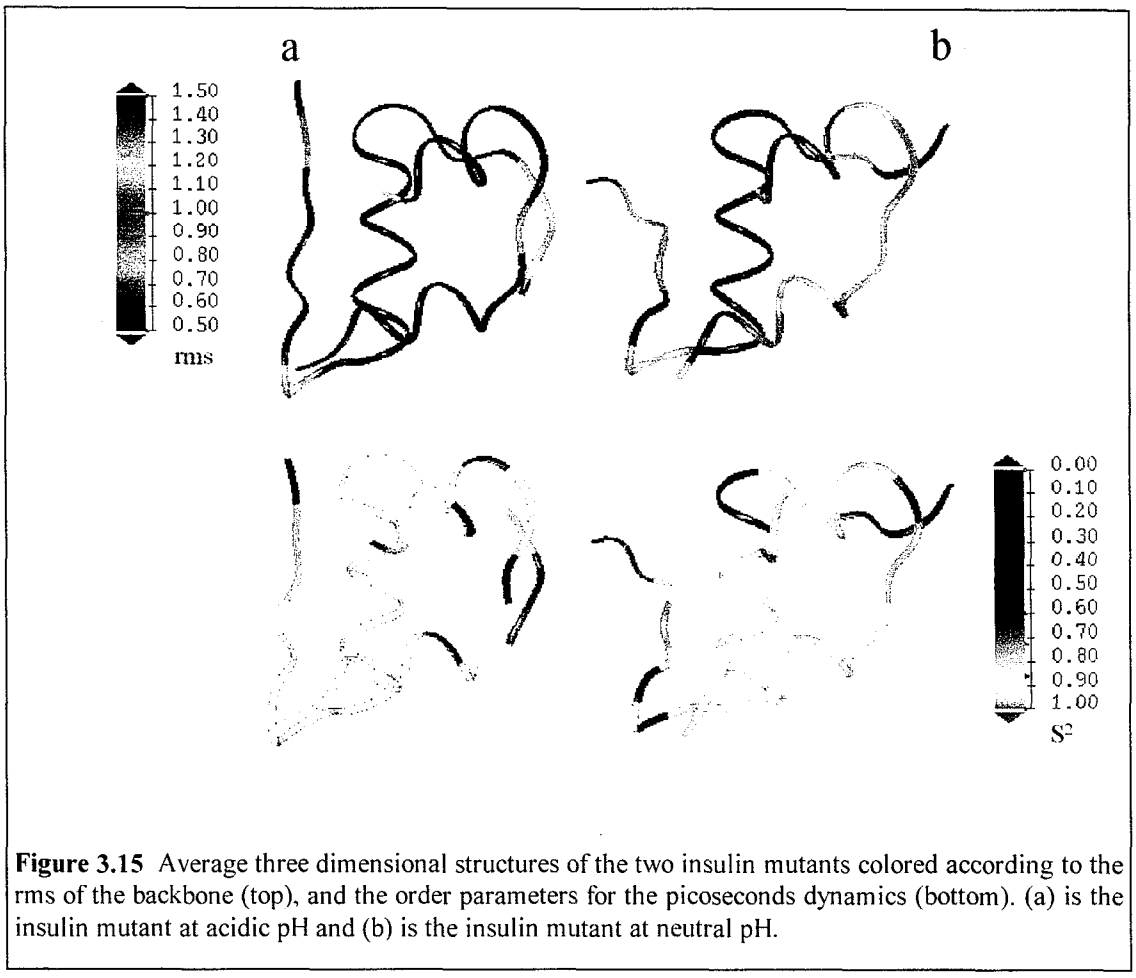


Figure 3.15 Average three dimensional structures of the two insulin mutants colored according to the rms of the backbone (top), and the order parameters for the picoseconds dynamics (bottom). (a) is the insulin mutant at acidic pH and (b) is the insulin mutant at neutral pH.

With references to Figure 3.12 and to the alternative view in Figure 3.15, the accordance between the generalized order parameter and the rms is striking for almost all residues at both pH values. Furthermore, the difference of overall precision between the two structures is mirrored perfectly in the overall levels of the generalized order parameter. Obviously picoseconds regime dynamics are responsible for the quenching of NOEs in the termini of the B-chain, as it is reflected in the increased rms deviations in these areas. The more pronounced detachment of the termini from the core of the molecule at neutral pH, caused by introduction of the charged glutamate residues (see Chapter 2), is reflected perfectly in the decline of the order parameter in these regions. The characteristic shape of the order parameter plot in the A-chain helices suggests the nature of the motions in these regions to be a bend of the helix (Pervushin et al., 1995). The dispersion of order parameter levels between the three helices are, however, not identical in the two investigations. The B-chain helix and the A(II)-helix share a common level in both investigation. Compared to this level, the A(I)-helix is relatively more mobile at acidic pH than at neutral pH. The importance of the A(I)-helix as to biological activity (examined in Chapter 4) makes this result worth noting.

3.5 Supplements

Standard deviation calculation, model A

Given are two relaxation rate times, T_1 and T_2, with standard deviations σ_{T_1} and σ_{T_2} and the NOE with standard deviation σ_{NOE}. From these figures the two parameters S^2, the generalized order parameter, and τ_e, the correlation time for internal movement, are calculated.

$$\begin{pmatrix} S^2 \\ \tau_e \end{pmatrix} = f\begin{pmatrix} T \\ T_2 \\ NOE \end{pmatrix} \qquad (11)$$

The variance matrix of the relaxation rate constants and the NOE contains the squared standard deviations of the three parameters on the diagonal, the off diagonal elements contain the correlations between the parameters. Assuming no correlation between the relaxation rate constants and the NOE the variance matrix is

$$\mathrm{var}\begin{pmatrix} T_1 \\ T_2 \\ NOE \end{pmatrix} = \begin{pmatrix} \sigma_{T_1}^2 & 0 & 0 \\ 0 & \sigma_{T_2}^2 & 0 \\ 0 & 0 & \sigma_{NOE}^2 \end{pmatrix} \qquad (12)$$

The variance matrix of the resulting S_2 and τ_e is calculated from

$$\mathrm{var}\begin{pmatrix} S^2 \\ \tau_e \end{pmatrix} = D \cdot \mathrm{var}\begin{pmatrix} T_1 \\ T_2 \\ NOE \end{pmatrix} \cdot D^{\mathrm{T}} \qquad (13)$$

where D is

$$D = \begin{pmatrix} \frac{\partial S^2}{\partial T_1} & \frac{\partial S^2}{\partial T_2} & \frac{\partial S^2}{\partial NOE} \\ \frac{\partial \tau_e}{\partial T_1} & \frac{\partial \tau_e}{\partial T_2} & \frac{\partial \tau_e}{\partial NOE} \end{pmatrix} \qquad (14)$$

and D^{T} is the transposed D matrix. The D matrix elements are determined by numerical diferentiation.

The elements in the resulting variance matrix for S^2 and τ_e are given by

$$\mathrm{var}\begin{pmatrix} S^2 \\ \tau_e \end{pmatrix} = \begin{pmatrix} \sigma_{S^2}^2 & \rho(\sigma_{S^2}\sigma_{\tau_e}) \\ \rho(\sigma_{S^2}\sigma_{\tau_e}) & \sigma_{\tau_e}^2 \end{pmatrix} \qquad (15)$$

the diagonal elements are the squared standard deviations of S^2 and τ_e, and the off-diagonal elements are correlations between the corresponding diagonal elements.

Table 3.2 ^{15}N relaxation data T_1, T_2, and NOE and ^{15}N chemical shift measured in ppm relative to the external reference 22.7 % NH_3 in aqueous solution.

	Insulin mutant at acidic pH							Insulin mutant at neutral pH						
Residue	^{15}N	T_1	σ_{T1}	T_2	σ_{T2}	NOE	σ_{NOE}	^{15}N	T_1	σ_{T1}	T_2	σ_{T2}	NOE	σ_{NOE}
A1	-							-						
A2	119.6	0.5651	0.0040	0.0386	0.0008	0.6205	0.0129	-						
A3	121.2	0.5055	0.0020	0.1243	0.0012	0.6224	0.0175	119.4	0.3660	0.0149	0.2070	0.0050	0.5850	0.0440
A4	119.5	0.5048	0.0034	0.0589	0.0009	0.6868	0.0150	117.7	0.4035	0.0205	0.1595	0.0045	0.7480	0.0440
A5	-							108.1	0.4385	0.0211	0.1082	0.0056	0.7280	0.0480
A6	108.3	0.5041	0.0153	0.0272	0.0042	0.8312	0.0994	108.2	0.3560	0.0168	0.0563	0.0063	0.5670	0.0900
A7	116.3	0.4969	0.0024	0.1075	0.0022	0.7441	0.0118	114.5	0.4060	0.0163	0.1682	0.0104	0.8030	0.0210
A8	112.2	0.3709	0.0417	0.1952	0.1450	0.9262	0.0580	108.2	0.3685	0.0336	0.0936	0.0077	0.7390	0.0870
A9	114.3	0.5146	0.0292	0.0783	0.0103	0.7912	0.0639	113.5	0.3980	0.0416			0.6870	0.1270
A10	115.2	0.5462	0.0062	0.1050	0.0025	0.7067	0.0431	114.6	0.4725	0.0313	0.2009	0.0065	0.5140	0.0230
A11	-							122.3	0.3670	0.0163	0.1067	0.0059	0.5710	0.0420
A12	114.7	0.5065	0.0054	0.0976	0.0034	0.7466	0.0343	112.5	0.4205	0.0218	0.2044	0.0081	0.6720	0.1480
A13	119.3	0.4840	0.0023	0.1357	0.0015	0.7129	0.0127	119.6	0.3975	0.0476	0.2115	0.0301	0.8160	0.0750
A14	113.1	0.4600	0.0076	0.1246	0.0033	0.8165	0.1663	112.3	0.3635	0.0133	0.2057	0.0165	0.7070	0.0220
A15	115.5	0.4811	0.0023	0.1674	0.0019	0.7588	0.0310	115.0	0.4335	0.0298	0.2261	0.0068	0.5760	0.0660
A16	117.6	0.4830	0.0024	0.0881	0.0012	0.7486	0.0198	115.9	0.4035	0.0097	0.1781	0.0030	0.5330	0.0170
A17	111.5	0.4636	0.0020	0.0743	0.0008	0.7516	0.0067	112.2	0.3625	0.0244	0.2091	0.0051	0.5810	0.0070
A18	114.7	0.4597	0.0018	0.1380	0.0014	0.7843	0.0227	108.3	0.3825	0.0099	0.2399	0.0080	0.6600	0.0300
A19	114.4	0.4750	0.0023	0.1294	0.0016	0.7827	0.0159	114.2	0.3280	0.0185	0.2145	0.0040	0.6310	0.0360
A20	113.5	0.4850	0.0032	0.1378	0.0025	0.7681	0.0343	112.5	0.4045	0.0152	0.2448	0.0115	0.6200	0.0180
A21	118.8	0.4967	0.0029	0.1693	0.0030	0.6958	0.0230	122.4	0.4290	0.0146	0.2790	0.0107	0.5810	0.0420
B1	-							-						
B2	121.7	0.5055	0.0020	0.0772	0.0015	0.1967	0.0161	-						
B3	121.1	0.6105	0.0018	0.2056	0.0019	0.2552	0.0160	119.3						
B4	117.6	0.5485	0.0024	0.1501	0.0020	0.5596	0.0197	107.9						
B5	116.8	0.5064	0.0031	0.1539	0.0026	0.7543	0.0284	-						
B6	124.4	0.4856	0.0120	0.0593	0.0052	0.7223	0.0255	122.6	0.3935	0.0202	0.1263	0.0103	0.5530	0.0580
B7	116.4	0.4969	0.0024	0.0873	0.0015	0.7441	0.0118	115.9	0.4015	0.0103	0.1949	0.0102	0.5620	0.0350
B8	110.1	0.5201	0.0293	0.1388	0.0158	0.8559	0.0685	123.2	0.5270	0.0842	0.4421	0.0810		
B9	120.6	0.4973	0.0071	0.0543	0.0020	0.7162	0.0524	-						
B10	116.9	0.4596	0.0019	0.1552	0.0016	0.7863	0.0212	117.5	0.3615	0.0107	0.1821	0.0061	0.6530	0.0510
B11	119.8	0.4590	0.0054	0.0730	0.0017	0.7399	0.0245	117.5	0.3785	0.0110	0.1370	0.0059	0.6920	0.0140
B12	115.1	0.4684	0.0033	0.1392	0.0020	0.7491	0.0154	114.2	0.4005	0.0138	0.2154	0.0066	0.5940	0.0320
B13	115.2	0.4660	0.0019	0.1070	0.0009	0.7537	0.0048	114.2	0.3865	0.0126	0.1868	0.0028	0.5800	0.0380
B14	119.6	0.4639	0.0018	0.1414	0.0012	0.7312	0.0255	118.2	0.4005	0.0101	0.2432	0.0070	0.6650	0.0510
B15	115.4	0.4862	0.0018	0.1264	0.0009	0.6484	0.0143	115.6	0.3975	0.0163	0.2277	0.0041	0.5410	0.0270
B16	115.0	0.4567	0.0021	0.1378	0.0016	0.7799	0.0230	116.0	0.3570	0.0125	0.2320	0.0052	0.6350	0.0240
B17	117.4	0.4717	0.0030	0.1399	0.0022	0.7697	0.0454	116.4	0.3835	0.0129	0.2229	0.0046	0.6340	0.0250
B18	113.5	0.4891	0.0041	0.0349	0.0007	0.7556	0.0380	114.7	0.4130	0.0189	0.2072	0.0051	0.6330	0.0310
B19	113.5	0.4648	0.0023	0.0736	0.0010	0.7655	0.0389	112.6	0.3845	0.0141	0.2139	0.0051	0.5470	0.0230
B20	105.7	0.4704	0.0028	0.1464	0.0026	0.7390	0.0297	105.5			0.2450	0.0113		
B21	121.5	0.5516	0.0086	0.0815	0.0024	0.6885	0.0643	-					0.6100	0.0890
B22	115.8	0.4862	0.0018	0.1264	0.0009	0.6484	0.0143	114.5	0.4245	0.0343	0.2318	0.0056	0.6100	0.0290
B23	101.4	0.4785	0.0062	0.0749	0.0023	0.6825	0.0724	109.9	0.4015	0.0144			0.7510	0.0730
B24	112.4	0.4676	0.0031	0.0630	0.0010	0.6979	0.0140	110.3			0.2144	0.0081		
B25	116.7	0.4895	0.0033	0.1450	0.0026	0.6837	0.0242	115.8	0.3875	0.0170	0.2455	0.0075	0.5770	0.0770
B26	119.2	0.5314	0.0025	0.1187	0.0019	0.4941	0.0401	118.1	0.4125	0.0110	0.1750	0.0061	0.5450	0.0270
B27	117.8	0.5503	0.0040	0.1186	0.0024	0.5039	0.0301	120.8	0.5165	0.0165	0.3516	0.1149	0.3020	0.0090
B28	-							-						
B29	120.3	0.6712	0.0018	0.2728	0.0023	0.0538	0.0100	123.2	0.7110	0.0432	0.4833	0.0851	-0.5090	0.0630

Table 3.3 Generalized order parameters and exchange line broadening

	Insulin mutant at acidic pH				Insulin mutant at neutral pH			
Residue	S^2	σ_{S2}	Δ_{ex}(Hz)	σ_{ex}	S^2	σ_{S2}	Δ_{ex}(Hz)	σ_{ex}
A1								
A2	0.75	0.02	6.3	0.14				
A3	0.83	0.04	0.5	0.18	0.80	0.06		
A4	0.85	0.04	3.3	0.16	0.88	0.02	0.5	0.03
A5					0.82	0.01	1.5	0.04
A6	0.87	0.06	9.5	0.06	0.94	0.01	4.0	0.92
A7	0.87	0.07	0.8	0.13	0.87	0.01	0.5	0.47
A8					0.98	0.01	1.7	0.06
A9	0.85	0.06	1.9	0.09				
A10	0.79	0.03	1.0	0.44	0.75	0.02	0.3	0.39
A11					0.74	0.03	1.5	0.45
A12	0.86	0.05	1.0	0.35	0.89	0.02		
A13	0.90	0.03			0.90	0.00		
A14	0.97	0.00			0.98	0.00		
A15					0.81	0.03		
A16	0.90	0.06	1.3	0.00	0.87	0.02	0.3	0.20
A17	0.94	0.03	2.0	0.09	0.89	0.01		
A18	0.94	0.05			0.88	0.01		
A19	0.93	0.04			0.89	0.03		
A20	0.90	0.04			0.83	0.01		
A21	0.85	0.05			0.75	0.01		
B1								
B2	0.64	0.05	2.3	0.58				
B3	0.62	0.03						
B4	0.75	0.05	0.3	0.21				
B5	0.85	0.04						
B6	0.89	0.04	3.0	0.29	0.88	0.01	1.0	0.62
B7	0.87	0.06	1.5	0.13	0.85	0.00	0.3	0.37
B8	0.86	0.00						
B9	0.87	0.04	3.8	0.53				
B10	0.92	0.05						
B11	0.94	0.06	2.0	0.26	0.94	0.01	0.8	0.18
B12	0.92	0.04			0.86	0.02		
B13	0.94	0.03	0.5	0.07	0.85	0.03	0.3	0.40
B14	0.91	0.04			0.83	0.00		
B15	0.88	0.01	0.3	0.15	0.81	0.02		
B16	0.95	0.04			0.86	0.02		
B17	0.92	0.04			0.86	0.02		
B18	0.89	0.06	6.8	0.39	0.89	0.03		
B19	0.94	0.05	2.0	0.39	0.87	0.02		
B20	0.91	0.05			0.79	0.02		
B21	0.77	0.05	2.0	0.65				
B22	0.88	0.01	0.3	0.15	0.80	0.03		
B23	0.89	0.06	2.0	0.73				
B24	0.92	0.03	2.8	0.15	0.87	0.02		
B25	0.87	0.04			0.70	0.01		
B26	0.76	0.05	0.8	0.41	0.86	0.02	0.3	0.29
B27	0.74	0.06	0.8	0.31	0.63	0.01		
B28								
B29	0.52	0.03			0.35	0.01		

4. Investigating structure function relationship of the A(I)-helix

This section presents the preliminary assignments of the NMR spectra of two mutants of human insulin. Both mutants are engineered in continuation of the (B1, B10, B16, B27) Glu, des-B30 mutant adding a point mutation in the A(I)-helix in order to enhance or diminish, respectively, the biological activity. The A3 Gly, (B1, B10, B16, B27) Glu, des-B30 mutant shows a biological activity of 0.1 % and the A8 His, (B1, B10, B16, B27) Glu, des-B30 mutant exhibits 143 % biological activity relative to the native species. The assignment of spin systems has been accomplished using DQF-COSY and TOCSY spectra. The NOESY spectra have been assigned to the point of sequential arrangement of the spin systems and secondary structure and compared to the (B1, B10, B16, B27) Glu, des-B30 NOESY spectrum on selected NOEs indicative of the tertiary fold of the molecule.

4.1 The A(I)-helix region

Throughout all determinations of insulin structures by NMR so far (see references in Table 2.3) the A2-A8 region takes a helical conformation. This is in agreement with the structures known from crystals. Although the structural definition of this region is poor, compared to the B-chain-helix (B9-B19) and the A(II)-helix (A12-A19), the helix propensity is clear. The A1 Gly, A2 Ile, and the two Cys residues A6 and A7 are invariant in naturally occuring insulins. The Cys residues are involved in two disulfide bridges, the A6-A11 intrachain- and the A7-B7 interchain-connection. The backbone dynamics in the millisecond timescale of the (B1, B10, B16, B27) Glu, des-B30 mutant (Chapter 3) concentrate in the three dimensional region of these disulfide bridges. This is interpreted as an indication of disulfide bond isomerization.

The A-chain N-terminus is part of one of the putative receptor binding sites of insulin (Pullen et al., 1976). This is evidenced by the marked influence of mutations and deletions in this region on the biological activity of the species. The relationship of specific sidechain conformations in the region to the biological activity has not been completely mapped. In the systematic approach, 'Ala-scan', each residue in turn is mutated by Alanin. Replacement of a residue by Ala corresponds to a deletion of the sidechain causing minimal disruption of the backbone conformation. This is in contrast to the backbone flexibility imposed by a Gly replacement. Part of this 'Ala-scan' can be pieced together from various publications. Compared to native insulin, the biological activities of the mutants with Ala point mutations are : A1 Gly → Ala 9 % (Cosmatos et al., 1978), A2 Ile → Ala 0.6 %, A3 Val → Ala 1.8 % (Nakagawa & Tager, 1992), A5 Gln → Ala 49 % (Märki et al., 1979), the double substitution of A6 Cys and A11 Cys by Ala 8 % (Märki et al., 1979), and finally A8 Thr → Ala 112 % (Kaarsholm et al., 1993). In addition to this, several publications present mutations in the A(I)-helix region with other amino acids as summareized here. In position A1, mutations with D-amino acids gives a slight reduction or a retention of the biological activity, this is in

contrast to mutations with the same amino acids in the naturally occurring L conformation, which causes 10-100 fold reduction (Cosmatos et al., 1978, Geiger at al., 1980 and 1982). In position A2, the mutation Ile → Val has only little effect on biological activity, whereas any other mutation causes a 10-100 fold reduction. In position A3, the biological activity is reduced by 100-1000, regardless of the mutation applied (Nakagawa & Tager, 1992). Furthermore, deletion of A1, A1-A2, A1-A3 (Nakagawa & Tager, 1992), or A1-A4 (Ferderigos et al., 1983) reduces the biological activity to < 0.003 % compared to the native species. In position A5 a substitution with Leu gives 30 % biological activity (Ferderigos et al., 1983). Introduction of the D form of Cys in positions A6 and A7 causes a reduction of biological activity to < 0.002 %. In contrast to these reductions, an increased biological activity is obtained by a stabilization of the A(I)-helix (as well as the B-chain helix, e.g. B10 Glu) (Kaarsholm et al., 1993). The introduction of the positively charged His or Arg residues to replace the Thr in position A8, the C cap of the helix, i.e. the negatively charged end of the helix-dipole, causes a 3 fold increase in bioactivity in concert with an enhancement of the folding stability. Note also that, 'miniproinsulin', i.e., insulin with a peptide crosslink between the termini of the two chains, B29 Lys to A1 Gly, shows a 1000 fold decrease of receptor binding potency (Markussen et al., 1985). The structure of 'miniproinsulin' is close to the structure of the native species, only slightly more compact (Derewenda et al., 1991).

As summarized above the biological activity of insulin is highly sensitive to mutations in the N-terminal A-chain helix. Substitutions of residues in the A2-A7 region cause a reduction of the biological activity, most of them even a dramatic reduction. In contrast, the introduction of His or Arg in the C cap of the helix, residue A8, increases the biological activity to approximately 300 % compared to the native species.

The relationship between the effect on biological activity and the perturbations of the three dimensional structure imposed by these mutations is, however, not examined in detail. An investigation of two insulin mutants with reduced, A3 Val → Gly, and increased, A8 Thr → His, biological activity has been initiated to adress this issue. In continuation of the (B1, B10, B16, B27) Glu, des-B30 mutant (biological activity 43 %), of which the structure is determined at physiological pH (Chapter 2) , the point mutations have been added to this species giving biological activities of 143 % for the A8 His, (B1, B10, B16, B27) Glu, des-B30 mutant ($^{+}$A8 His) and 0.1 % for the A3 Gly, (B1, B10, B16, B27) Glu, des-B30 mutant ($^{+}$A3 Gly) respectively.

4.2 Materials and methods

The two insulin mutants, $^{+}$A8 His and $^{-}$A3 Gly, were constructed, fermented and purified as desribed for the (B1,B10,B16,B27) Glu, des-B30 mutant in Chapter 2. The experimental conditions were duplicated, i.e., the samples were prepared at a concentration of 2.5 mM in 10/90 D_2O/H_2O and pH was adjusted to a direct meter reading of 6.5. All NMR experiments were recorded at a temperature of 307 K. The series of NMR experiments recorded were similar to the series for

the (B1,B10,B16,B27) Glu, des-B30 mutant, refer to Chapter 2 for the details.

4.3 Results

The 1D ^{1}H spectra of the three insulin mutants (B1,B10,B16,B27) Glu, des-B30 (basis), $^+$A3 Gly and $^+$A8 His are shown in Figure 4.1. The monomeric character of the $^+$A3 Gly and $^+$A8 His species is evident from the high resolution of the spectra, which is comparable to the resolution of the basis mutant spectrum.

The near-UV CD spectrum of the basis mutant at increasing protein concentration was shown in Figure 2.4. The spectra of the $^+$A3 Gly and $^+$A8 His mutants (not shown) share the same characteristics, i.e., independence of protein concentration. This is a further evidence of the monomeric character of both mutants.

Figure 4.2 gives an impression of the fine quality of the DQF-COSY and NOESY spectra of both mutants. Both set of spectra are in close resemblance to the spectra of the basis mutant (Figure 2.5). This applies to the dispersion as well as the linewidths of the crosspeaks in the fingerprint region of the DQF-COSY spectra. The similarities are even more clear in the excerpts from the NOESY spectra. Apparently the effect of the mutations on the amide proton resonances in the B-chain helix and the A(II)-helix are minor. The dispersion of the resonances in the fingerprint region of the DQF-COSY spectra made the assignment of the spectra, using the method of Wüthrich (1986), fairly straightforward. The major part of the H^NH^α crosspeaks were assigned in the DQF-COSY spectra, and annotated on the plot. The TOCSY spectra allowed the assignment of the

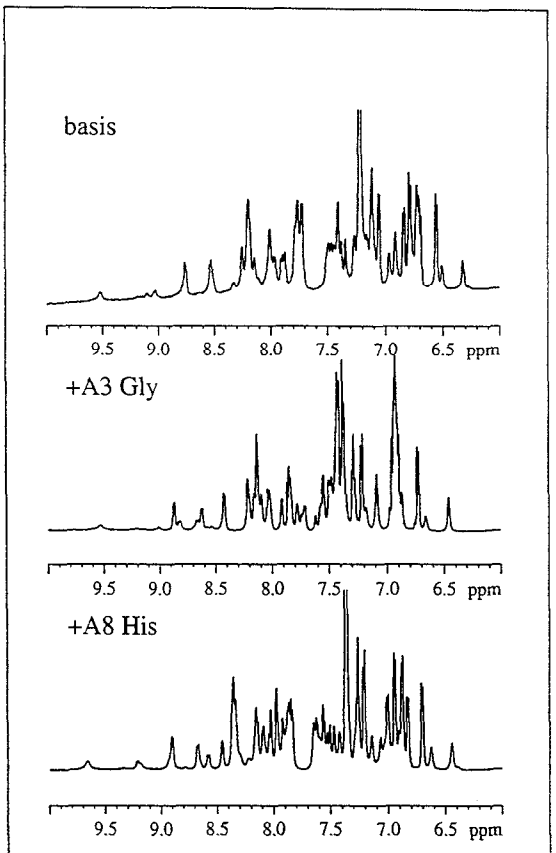


Figure 4.1 1D proton spectra of three insulin mutants, region of amide and aromatic protons.

rest, but still the N-terminal residues of both chains, A1 Gly and B1 Glu, as well as B8 Gly are unassigned. For the $^+$A8 His mutant the A9 Ser misses assignment as well. During the sequential analysis the major part of the B-chain, B2-B7 and B9-B29, and the A-chain stretches A2-A8 and A10-A21 was assigned. Table 4.1 at the end of this chapter lists the chemical shifts of both mutants.

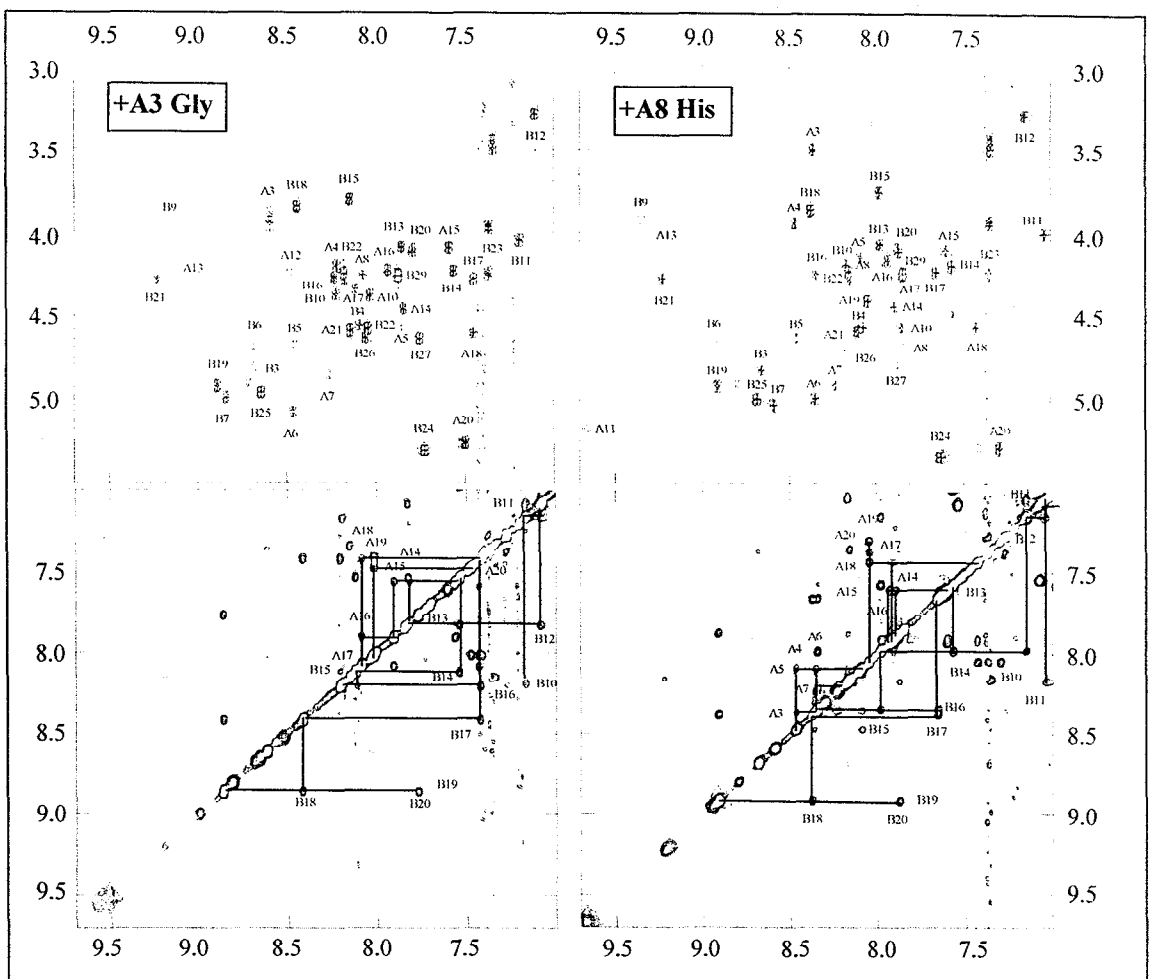


Figure 4.2 Sections from DQF-COSY spectra (top) and NOESY spectra (mixingtime 120 ms) (bottom). In the fingerprint region of the DQF COSY spectrum the major part of the identified correlations between α-protons and amide protons is annotated. The correlations between amide protons of neighboring residues, indicative of helical areas in the structure, are annotated in the bottom panel and lines are drawn to follow the amide protons throughout the helices.

Part of the secondary structural elements of the mutants, the helix regions, are indicated in the excerpts from the NOESY spectra in the bottom panel of Figure 4.2. Intense H^NH^N connectivities are assigned for both mutants in the B-chain helix region, B10-B20, and in the A(II)-helix region, A14-A20. In contrast to the $^+$A3 Gly mutant (and the basis, Figure 2.4), the spectrum of the $^+$A8 His mutant shows intense H^NH^N connectivities in the A(I)-helix, A3-A7. This is an evidence of the helix stabilization caused by the A8 Thr → His mutation in the helix C cap. Note that the basis mutant spectra do show H^NH^N connectivities in the A3-A5 region, these peaks are less intense and visible, but not annotated, in Figure 2.5.

Further indications of secondary structure come from comparison of the α-proton chemical shift and the random coil values, the secondary chemical shifts of the α-protons. Data for all three mutants is shown in the left-hand column of Figure 4.3. The helix elements predicted by the extents of

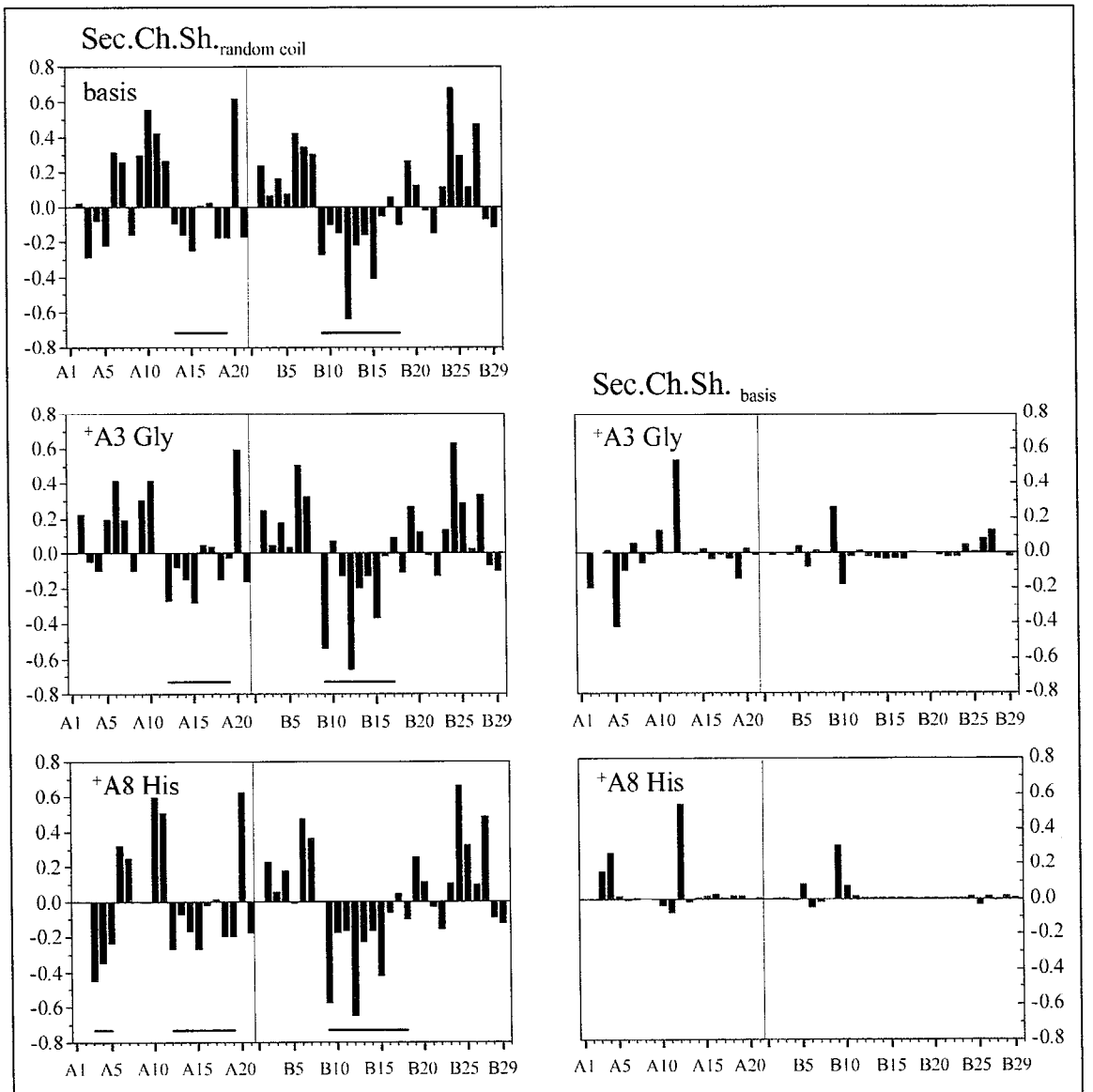


Figure 4.3 Secondary chemical shifts of the α-protons for three insulin mutants. The left hand column contains the difference by residue between the measured chemical shift and the random coil value. Regions predicted to be helix from these data are marked with lines. The right hand column contains the differences between the measured chemical shifts of the $^{+}$A3 Gly and $^{+}$A8 His mutants and those of the basis mutant, the points of mutation are left out.

secondary chemical shifts (Wishart et al., 1992) are annotated in the plots. The B-chain and the A(II)-helix regions are predicted for all three mutants, and the differences between the three mutants in these regions are only minor.

The N-terminal region of the A-chain, however, shows marked differences among the mutants. The size of the upfield shifts of the A3, A4, A5 α protons of the $^{+}$A8 His mutant compared to the basis mutant indicate the A(I) helix to be better defined in the $^{+}$A8 His mutant. The $^{+}$A3 Gly mutant, on the other hand shows no sign of helix content in this region. These predictions are in agreement with the observation of intense $H^{N}H^{N}$

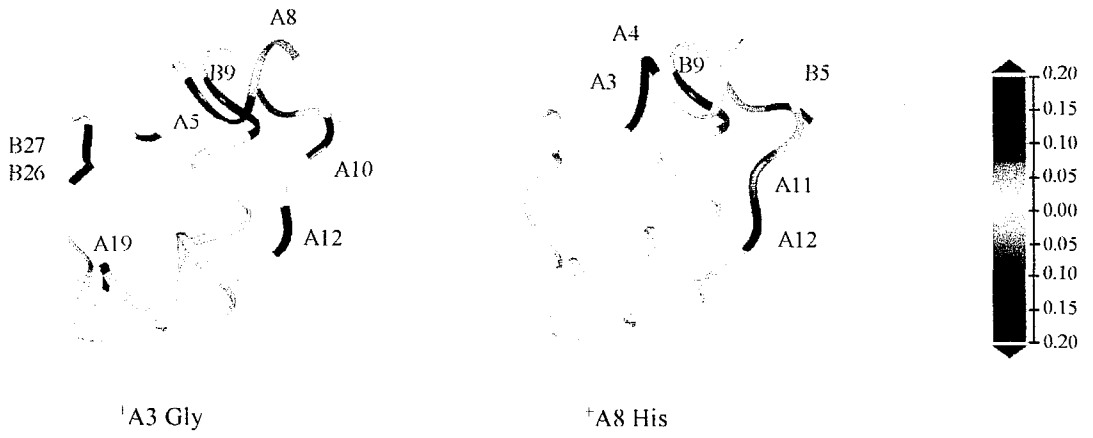


Figure 4.4 Perturbation of the three dimensional structure caused by point mutations. The structure of the basis mutant is color coded according to difference in α-proton chemical shifts between the two mutants ⁱA3 Gly and ⁺A8 His and the basis mutant.

connectivities in the A3-A7 region of the ⁺A8 His mutant, and the absence of the same connectivities in the specctra of the basis and the ⁺A3 Gly mutants.

The right-hand column of Figure 4.3 shows another secondary chemical shift, the difference in chemical shift of the ⁺A3 Gly and ⁺A8 His α-protons, respectively, and the basis mutant α-protons. The experimental conditions, i.e., temperature, pH, and protein concentration, of the three investigations is the same, i.e., the major influence on this secondary chemical shift comes from differences in the secondary and tertiary structure caused by the point mutations.

For the ⁺A3 Gly mutant, secondary chemical shifts are observed throughout the A2-A10 region. In addition to this, the loop region of the B-chain, which is connected to A-chain N-terminal region by the A7-B7 disulfide bridge, as well as the A19 residue show secondary chemical shifts. The C-terminus of the B-chain is characterized by secondary chemical shifts increasing towards the end.

The extent of secondary chemical shifts in the ⁺A8 His mutant is less than in the ⁺A3 Gly mutant. The primary structure around the point mutation is affected, the shift of the A3 and A4 protons can be subscribed to the strengthening of the helix. The spatially close B-chain residues, B5-B10, are affected, as was the case for the ⁺A3 Gly mutant. The secondary chemical shifts in the C-terminus of the B-chain are negligible.

The close similarity between the spectra of the three mutants and the identical A(II) and B-chain helix regions as well suggest the three dimensional structure of the core of the molecule to be conserved throughout. In order to confirm these observations, a few preliminary assignments of NOEs, based on NOEs found in the spectra of the basis mutant, has been accomplished in the spectra of both mutants. The internal arrangement of the A-chain elements are substantiated by NOEs from the sidechain of residue Ile A2 to the aromatic part of residue Tyr A19. NOEs from the sidechain of Leu B15 to the aromatic part of Phe B24 indicate the conservation of the antiparallel helix-strand

motif in the B-chain. Finally, NOEs from the sidechain of Leu B15 to the aromatic part of Tyr A19 as well as from the sidechain of Ile A10 to the sidechains of the Asn B3, Gln B4, and His B5 residues confirm the conservation of the mutual arrangement of the chains.

4.4 Conclusion

In continuation of the (B1,B10,B16,B27) Glu, des-B30 insulin mutant of which the three dimensional structure is known, two mutants employing point mutations in the A(I)-helix has been engineered. The A(I)-helix is part of one of the putative receptor binding sites of insulin, and the aim of the mutations has been to manipulate the biological activity. The NMR spectra of these two insulin mutants A3 Gly, (B1,B10,B16,B27) Glu, des-B30 (biological activity of 0.1 %) and A8 His, (B1,B10,B16,B27) Glu, des-B30 (biological activity of 143 %) has been examined as to the sequential assignment of spinsystems and secondary structure. The helix regions in the B-chain (B9-B19) and in the C-terminus of the A-chain (A13-A19) is conserved as evidenced by close contacts between neighboring amide proton. The chemical shift of the α-protons compared to random coil values give indications of the B-chain and A(II)-helices. Furthermore a lack of helix structure in the A2-A8 region of the $^{+}$A3 Gly mutant and a strengthening of the helix in the same region (the A(I)-helix) in the $^{+}$A8 His mutant is evidenced by these secondary chemical shifts. A few NOEs indicative of the tertiary fold of the molecule assigned for the basis mutant have been found in the spectra of both mutants. These assignments indicate a conservation of the overall spatial arrangement of the secondary structural elements in both chains as well as the mutual arrangement of the chains. A comparison of the chemical shift of the α-protons of the two mutants to those of the basis mutant gives a direct measure of the propagation of the structural changes caused by the point mutations. The conformational changes of the +A8 His mutant is confined to the A(I)-helix especially residues A3 and A4, as a consequence of the stabilized helix, and the spatially close B-chain loop area. The apparent unravelling of the A(I)-helix of the +A3 Gly mutant affects the B-chain loop area as well and furthermore to the C terminus of the B-chain.

The complete assignment of the NOESY spectra of the two mutants as a prerequisite of the structure determination will add details to the observations listed above. The precise structure of the backbone and indeed the presentation of sidechains in the A(I)-helix region will be important results in the examination of the relationship between structure and biological activity of insulin. Moreover, measurements of the relaxation parameters of the backbone ^{15}N nuclei will allow a detailed description of the millisecond and picosecond dynamics of the backbone of the mutants in comparison to the basis mutant.

4.5 Supplements

^{1}H chemical shifts of the A3 Gly, (B1, B10, B16, B27) Glu, des B30 insulin mutant at pH 6.5 and 307 K measured in ppm relative to dioxane (3.75 ppm).

Residue		H_N	H_α	H_β	Others
A1	Gly				
A2	Ile		4.18	1.68	$H_{\gamma1}$ 1.39/1.09 $H_{\gamma2}$ 0.85 H_δ 0.77
A3	Gly	8.59	3.88/3.96		
A4	Glu	8.22	4.19	2.54/2.28	H_γ 2.44/2.16
A5	Gln	7.85	4.57	2.95/3.44	
A6	Cys	8.46	5.07	3.00/3.56	
A7	Cys	8.25	4.85	3.73/3.28	
A8	Thr	8.06	4.25	4.47/1.31	
A9	Ser		4.81	4.06/3.97	
A10	Ile	8.03	4.37	1.79	$H_{\gamma1}$ 0.80 $H_{\gamma2}$ 1.36 H_δ 0.81/0.67
A11	Cys	9.53		3.28/3.18	
A12	Ser	8.54	4.23	4.08	
A13	Leu	9.02	4.09	1.64/1.74	H_δ 1.02/ 0.93
A14	Tyr	7.84	4.45	3.12	H_δ 7.23 H_ϵ 6.96
A15	Gln	7.59	4.09	2.38	H_γ 2.50/2.30
A16	Leu	7.93	4.22	2.08/1.47	H_γ 1 .85 H_δ 0.83
A17	Glu	8.11	4.33	2.12	H_γ 2.33/2.48
A18	Asn	7.46	4.60	2.69/2.48	H_δ 7.30/6.68
A19	Tyr	8.04	4.57	3.04	H_δ 7.44 H_ϵ 6.92
A20	Cys	7.50	5.25	2.95/3.46	
A21	Asn	8.14	4.59	2.77/2.90	H_δ 6.48/7.52
B1	Glu				
B2	Val		4.20	2.19	H_γ 1.05/1.01
B3	Asn	8.67	4.80	2.90/2.86	
B4	Gln	8.09	4.55	2.01/2.21	H_γ 2.34
B5	His	8.45	4.67	3.23/3.53	$H_{\delta2}$ 7.09 $H_{\epsilon1}$ 8.15
B6	Leu	8.67	4.68	1.76/0.94	H_γ 1.66 H_δ 0.86
B7	Cys	8.83	4.98	3.01/3.30	
B8	Gly				
B9	Ser	9.15	3.96	4.13	
B10	Glu	8.22	4.37	2.27/2.21	H_γ 2.46/2.37
B11	Leu	7.20	4.04	1.84/1.27	H_γ 1.27 H_δ 0.68/0.73
B12	Val	7.11	3.29	2.21	H_γ 1.04/1.11
B13	Glu	7.85	4.09	2.10	H_γ 2.50/2.34
B14	Ala	7.56	4.22	1.50	
B15	Leu	8.15	3.80	0.29/1.05	H_γ 1.42 H_δ 0.68
B16	Glu	8.23	4.27	2.33	H_γ 2.69/2.13
B17	Leu	7.45	4.27	1.87/1.92	H_γ 1.79 H_δ 1.02/0.99
B18	Val	8.44	3.84	2.08	H_γ 0.96/1.11
B19	Cys	8.89	4.92	3.04/3.41	
B20	Gly	7.80	4.12/4.08		
B21	Glu	9.22	4.28	1.92/2.18	H_γ 2.46/2.27
B22	Arg	8.18	4.25	2.17/2.29	H_γ 1.98 H_δ 3.50/3.43 H_ϵ 7.34
B23	Gly	7.37	4.24/3.97		
B24	Phe	7.73	5.30	3.47/2.99	H_δ 6.94 H_ϵ 7.30 H_ζ 7.41
B25	Phe	8.64	4.95	3.26/3.32	H_ϵ 7.39
B26	Tyr	8.06	4.63	2.89/3.01	H_δ 6.94 H_ϵ 6.74
B27	Glu	7.75	4.63	1.88/2.02	H_γ 2.23
B28	Pro		4.37	2.27	H_γ 2.10/2.00 H_δ 3.53/3.73
B29	Lys	7.87	4.26	1.79	H_γ 1.49/1.92 H_δ 1.76 H_ϵ 3.06

^{1}H chemical shifts of the A8 His, (B1, B10, B16, B27) Glu, des B30 insulin mutant at pH 6.5 and 307 K measured in ppm relative to dioxane (3.75 ppm).

Residue		H_N	H_α	H_β	Others
A1	Gly				
A2	Ile				$H_{\gamma1}$ 0.85/1.08 $H_{\gamma2}$ 0.70 H_δ 0.49
A3	Val	8.36	3.50	1.94	H_γ 1.04/0.75
A4	Glu	8.47	3.94	2.26/2.11	
A5	Gln	8.09	4.13	2.17	
A6	Cys	8.35	4.98	2.95	
A7	Cys	8.23	4.90	3.84/3.23	
A8	His	7.84	4.64	3.45/3.49	$H_{\delta1}$ 6.94 $H_{\epsilon2}$ 8.02
A9	Ser				
A10	Ile	7.85	4.55	1.67	$H_{\gamma1}$ 0.62/1.24 $H_{\gamma2}$ 0.77
A11	Cys	9.66	5.16	3.51/3.14	
A12	Ser	8.30	4.23	4.06	
A13	Leu	9.19	4.10	1.61/1.75	H_γ 1.76 H_δ 1.00/0.91
A14	Tyr	7.89	4.43	3.13/3.08	H_δ 7.22 H_ϵ 6.95
A15	Gln	7.60	4.10	2.37	H_γ 2.49/2.33 H_δ 7.54/7.07
A16	Leu	7.93	4.15	2.08/1.51	H_γ 1.84 H_δ 0.78/0.84
A17	Glu	7.91	4.31	2.10	H_γ 2.38/2.46
A18	Asn	7.43	4.55	2.71/2.62	H_δ 7.21/6.63
A19	Tyr	8.05	4.40	3.55/2.85	H_δ 7.37
A20	Cys	7.30	5.28	2.89/3.43	
A21	Asn	8.10	4.57	2.75/2.87	H_δ 6.45/7.51
B1	Glu				
B2	Val		4.18	2.16	H_γ 1.04/1.00
B3	Asn	8.65	4.81	2.89/2.84	H_δ 7.63/6.91
B4	Gln	8.07	4.55	2.01	H_γ 2.21/2.34
B5	His	8.45	4.62	3.25/3.58	$H_{\delta1}$ 7.26 $H_{\epsilon2}$ 8.35
B6	Leu	8.93	4.65	1.77	H_γ 1.63 H_δ 0.94
B7	Cys	8.58	5.02	3.01/3.27	
B8	Gly				
B9	Ser	9.34	3.92	4.09	
B10	Glu	8.17	4.11	2.35/2.16	H_γ 2.52/2.45
B11	Leu	7.04	4.00	1.89/1.17	
B12	Val	7.15	3.30	2.16	H_γ 1.02/1.08
B13	Glu	7.98	4.06	2.10	
B14	Ala	7.57	4.18	1.48	
B15	Leu	7.98	3.75	0.24/1.07	H_γ 1.28 H_δ 0.34/0.62
B16	Glu	8.34	4.23	2.32/2.12	H_γ 2.70/2.47
B17	Leu	7.66	4.22	1.89/1.76	
B18	Val	8.38	3.85	2.08	H_γ 0.96/1.12
B19	Cys	8.91	4.91	3.01/3.40	
B20	Gly	7.87	4.11/4.07		
B21	Glu	9.22	4.26	2.17/2.26	H_γ 2.45
B22	Arg	8.15	4.22	2.16/2.28	H_γ 1.97 H_δ 3.49/3.42 H_ϵ 7.35
B23	Gly	7.35	4.23/3.93		
B24	Phe	7.63	5.33	3.44/3.02	H_δ 6.89 H_ϵ 7.27 H_ζ 7.37
B25	Phe	8.68	4.99	3.26/3.33	H_δ 7.37
B26	Tyr	8.17	4.70	2.99/3.09	H_δ 7.01 H_ϵ 6.71
B27	Glu	7.87	4.78	1.94	H_γ 2.09/2.29
B28	Pro		4.35	2.23	H_γ 2.11/2.02 H_δ 3.60/3.79
B29	Lys	7.84	4.23	1.89/1.75	H_γ 1.44 H_δ 1.72 H_ϵ 3.01

Table 4.1 ^{1}H chemical shifts of the $^+$A3 Gly and $^+$A8 His insulin mutants.

5. Conclusion

This Chapter gives a summary of the results described in the preceeding Chapters as well as the concluding remarks of the present thesis.

The three dimensional structure of the (B1, B10, B16, B27) Glu, des -B30 mutant of human insulin in aqueous solution at neutral pH was determined by ^{1}H two-dimensional NMR. The aggregation problems that characterize millimolar solutions of the native species are succesfully overcome by site-directed mutagenesis of the dimer- and hexamer-forming surfaces with only minor effect of the biological activity (47 % compared to that of native insulin). The insulin mutant is monomeric in the neutral pH range, and the three dimensional structure is the first insulin structure determined at physiological pH. The core of the molecule is structurally well-defined with an rms between the 25 calculated structures and their average of 0.66 Å for the backbone atoms (A2-A19, B4-B26). The A-chain consists of the antiparallel helices A2-A8 and A12-A19 connected by a loop, the B-chain residues B1-B4 are disordered, residues B5-B8 are in a loop region flanking the central B-chain helix that comprises residues B9-B19, residues B20-B23 forms a type I turn continuing in the N-terminal strand, B27-B29 are disordered. The disorder in the B-chain termini is subscribed to the insertion of the charged Glu sidechains. The structure of this insulin mutant was compared to the structures of other monomeric insulin mutants determined at acidic pH. Apart from the differences in overall precision of the structures a common feature is the spatial arrangement of the well defined part of the structure, the core of the molecule, comprising the A-chain and the central B-chain helix, backbone as well as sidechains. The (B1, B10, B16, B27) Glu, des-B30 showed closest structural resemblance to the B16 His mutant, the best defined structure in the set, rms of the backbone for this mutant is 0.46 Å in the region A2-A19, B4-B28.

In order to elucidate the inherent dynamics of the (B1, B10, B16, B27) Glu, des-B30 mutant of human insulin at neutral pH as compared to the dynamics of the B16 His, des-B30 mutant of which the three dimensional strucuture has been determined at acidic pH, the backbone dynamics of the ^{15}N enriched mutants were examined. The relaxation parameters, T_1, T_2, and NOE of the amide ^{15}N nuclei, are translated to a generalized order parameter, S^2, describing the spatial restriction of each NH-vector. A fine overall coherence between the rms and the spatial restriction of the NH-vector motion of each residue is observed for both mutants. This points to the pico-nanoseconds timescale motions to be responsible for the lack of NOEs in the termini of both chains that leads to disorder in the calculated structures. Moreover the difference in rms levels between the two mutants is reflected in a similar difference in NH-vector motion level. The millisecond time scale motions observed in the spatial surroundings of the two disulfide bridges A6-A11 and A7-B7 at neutral pH are interpreted as evidence for disulfide bond isomerization. The millisecond time scale motions at acidic pH are more pronounced and involves all secondary structural elements.

The preliminary assignments of the NMR spectra of the two insulin mutants A3 Gly, (B1, B10, B16, B27) Glu, des-B30 and A8 His, (B1, B10, B16, B27) Glu, des-B30 with biological

activities of 0.1 % and 143 % respectively, compared to native insulin lead to conclusions concerning the secondary and tertiary structure of these mutants. The immediate resemblance of these spectra to the spectra of the (B1, B10, B16, B27) Glu, des-B30 mutant indicates the overall three dimensional structure to be conserved, this is confirmed by the assignment of a subset of the NOEs indicative of the tertiary structure. The sequential assignment and the α-proton chemical shifts as compared to their random coil values and to the chemical shift of the (B1, B10, B16, B27) Glu, des B30 mutant α-protons revealed similarities as well as differences. The helix in the A2-A8 region of the A8 His, (B1, B10, B16, B27) Glu, des-B30 is strengthened, this change in conformation propagates (through the disulfide bridges) to the spatially close B-chain loop region especially B9 and to the residues A11 and A12. In the A3 Gly, (B1, B10, B16, B27) Glu, des-B30 mutant the helix propensity of the A2-A8 region is lost, the structural disorder spreads to the same regions as for the other mutant, but also to the N-terminal part of the B-chain and to some extent to the B-chain helix.

The determination of the three dimensional structure of the (B1, B10, B16, B27) Glu, des-B30 mutant at physiological pH and the knowledge of the inherent dynamics provides the basis for future investigations of the relationship between structure and biological activity of insulin species. The detailed structure determination of the two mutants with manipulated biological activity of which the preliminary assignments have already revealed local changes in secondary structure and located the regions of structural perturbations, will undoubtedly lead to a closer insight into this complicated area. Furthermore, if the future brings fragments of the soluble parts of the insulin receptor that bind the insulin molecule, of size and properties suitable for binding studies by NMR, the (B1, B10, B16, B27) Glu, des-B30 mutant will be an obvious candidate to be used in this context.

6. Danish summary / Dansk sammenfatning

Nærværende ph.d. afhandling omhandler NMR (kerne magnetisk resonans) spektroskopiske undersøgelser af struktur og dynamik af en række mutanter af human insulin. Projektet blev udført som en erhvervsforskeruddannelse i et samarbejde mellem Novo Nordisk A/S, Roskilde Universitetscenter og Akademiet for de Teknisk Videnskaber (ATV).

Den insulinkrævende (Type 1) diabetes mellitus er karakteriseret ved en stærkt nedsat eller fraværende produktion af endogen insulin. Patienter med denne kroniske sygdom behandles med subkutane injektioner af insulinpræparater en eller flere gange dagligt for at opretholde et nogenlunde normalt blodsukkerniveau. På Novo Nordisk forskes til stadighed i at udvikle denne injektionsterapi til så præcist som muligt at efterligne kroppens normale insulinprofil. Et meget attraktivt langsigtet mål er at udvikle et insulinmimetika, som kan indgives peroralt. Derfor er der i de seneste år udført forskning med henblik på at karakterisere bindingsområderne på insulinreceptoren. Parallelt hermed søges sammenhængen mellem forskellige insulin mutanters variable affinitet for insulinreceptoren og deres tre-dimensionelle struktur klarlagt, herunder hører undersøgelse af de strukturændringer som følger af insulins binding til receptoren. Dette projekt er iværksat som en del af denne klarlægning.

Insulin består af 51 aminosyre rester fordelt på to kæder, A-kæden består af 21 rester og B-kæden består af 30 rester. De to kæder er indbyrdes forbundet af to svovlbroer, A7-B7 og A20-B19, herudover findes en svovlbro internt i A-kæden, A6-A11. Den rumlige struktur af insulin molekylet er grundigt undersøgt vha. Røntgen-krystallografi. I nærvær af zinc krystalliserer molekylet i hexamer-aggregater sammensat af tre ens dimerer. Den biologisk aktive konformation af insulin er imidlertid monomeren.

En række komplicerede aggregerings- og opløseligheds-forhold vanskeliggør detaljerede strukturstudier vha. NMR spektroskopi af den native monomer i det fysiologiske pH område. Ved at indføre substitutioner af aminosyre rester i begge aggregeringsflader har det imidlertid vist sig muligt at fremstille en mutant af human insulin, som er monomer netop i det ønskede pH område. Den biologiske aktivitet af mutanten er 47 % af den oprindelige. Den tre-dimensionelle struktur af (B1, B10, B16, B27) Glu, des-B30 mutanten blev bestemt ved pH 6.5. Strukturen er generelt veldefineret, rms (root mean square) for backbone i områderne A2-A19 og B4-B26 er 0.66 Å. Strukturen i opløsning adskiller sig kun lidt fra krystalstrukturerne. A-kæden betår af to antiparallelle helix områder, A(I)-helixen, A2-A8, og A(II)-helixen, A12-A19, forbundet af et loop. I B-kæden er B1-B3 uordnede, B4-B8 danner et loop til den centrale α-helix, B9-B19, herefter ses et type I turn, B20-B23, som fortsætter i en β-struktur, resterne B27-B29 er uordnede. Fælles for denne struktur og strukturen af andre insulin mutanter bestemt i sur opløsning, er et veldefineret kerneområde, som består af B-kæde helixen og hele A-kæden. Den sekundære såvel som den tertiære struktur i dette kerneområde er bevaret gennem hele serien af strukturer. Forskelle mellem strukturerne findes primært i de terminale dele af B-kæden. Sammenlignet med strukturen af B16 His mutanten, som denne struktur ligner mest

i serien af mutanter, bevirker indsættelsen af de ladede Glu sidekæder en ringere præcision af strukturen i begge B-kæde terminaler.

Mærkning af de to insulin mutanter (B1, B10, B16, B27) Glu, des-B30 og B16 His, des-B30 med den NMR-aktive nitrogen-isotop ^{15}N muliggør bestemmelse af relaxationsparametrene, T_1 og T_2, samt NOE'en for ^{15}N kernerne i backbone. Disse relaxationsparametre kan oversættes til en generaliseret ordens parameter, som er et udtryk for den rumlige begrænsning af NH bindingens bevægelse på pico-nanosekund tidsskalaen. For begge mutanter ses en fin korrelation mellem præcisionen af den beregnede strukturs backbone og dynamikken her. Terminalerne i B-kæden udviser stigende dynamik mod enderne, hvilket reflekteres direkte i faldende strukturel definition i de samme områder. De to mutanters forskellige generelle strukturelle præcision afspejles perfekt i deres overordnede dynamik niveauer. Endvidere observeres millisekund tidsskala dynamik i området omkring svovlbroerne A7-B7 og A6-A11 ved neutral pH, dette tages som en indikation for isomeri omkring svovlbindingerne i broerne. Ved surt pH er dynamikken på denne tiddskala spredt ud over samtlige sekundære strukturelementer.

A(I)-helix området er en del af et af insulins formodede receptor bindings-områder, mutationer i dette område indvirker kraftigt på den biologisk aktivitet. I fortsættelse af (B1, B10, B16, B27) Glu, des-B30 insulin mutanten er to mutanter med en ekstra mutation, A3 Gly hhv. A8 His, som giver 0.1 % hhv. 143 % biologisk aktivitet fremstillet. Begge mutanter er monomere under eksperimentelle betingelser svarende til basis mutantens. Den tre-dimensionelle struktur af de to mutanter er endnu ikke beregnet, men resultaterne af de indledende NMR undersøgelserne viser som ventet markante forskelle i A2-A8 området. For A8 His mutanten ses en klar stabilisering af helixen i området, herudover ses små strukturelle ændringer i den del af B-kæden, som ligger runligt tæt herpå, især B9. For A3 Gly mutanten er alle tegn på helix i A2-A8 området forsvundet, dette indvirker ikke kun på strukturen i området omkring B9, men også på strukturen i den C-terminale del af B-kæden og i nogen grad på den centrale B-helix.

References

* Abragam, A. (1961) *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford.
* Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Harding, M. M., Hodgkin, D. C., Rimmer, R, & Sheet, S. (1969) *Nature 224*, 491-496.
* Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Comm. 95*, 1-5.
* Anil-Kumar, Wagner, G., Ernst, R. R., & Wüthrich, K. (1981) *J. Am. Chem. Soc. 103*, 3654-3658.
* Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., & Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London ser. B 319*, 369-456.
* Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355-360.
* Bloom, C. R., Choi, W. E., Brzovic, P. S., Ha, J. J., Huang, S.-T., Kaarsholm, N. C., & Dunn, M. F. (1995) *J.Mol.Biol. 245*, 324-330.
* Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett.* 69, 185.
* Brader, M. L., Kaarsholm, N. C., Lee R., W.-K., & Dunn, M. F. (1991) *Biochemistry 30*, 6636-6645.
* Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R., & Voigt, H. O. (1988) *Nature, 333*, 679-682.
* Brange, J., Owens, D. R., Kang, S., Vølund, Aa. (1990) *Diabetes Care 13*, 923-954.
* Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521-528.
* Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaninathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187-217.
* Brünger, A. T. (1992) *X-PLOR manual version 3.0*, Yale University, New Haven, CT.
* Burum, D. P., & Ernst, R. R. (1980) *J. Magn. Reson.* 39, 163.
* Carr, H. Y., & Purcell, E. M. (1954) *Phys. Rev.* 94, 630.
* Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C. & Gronenborn, A. M. (1990a) *J. Am. Chem. Soc. 112*, 4989-4991.
* Clore, G. M., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990b) *Biochemistry 29*, 7387-7401.
* Cosmatos, A., Cheng, K., Okada, Y., & Katsoyannis, P. G. (1978) *J. Biol. Chem. 253*, 6586-6590.
* Crippen, G. & Havel, T. (1988) *Distance Geometry and Molecular Conformation.* Research Studies Press, Taunton, Somerset, England.
* Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swensen, D. (1989) *Nature 338*, 594-596.
* Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. G., & Markussen, J. (1991) *J. Mol. Biol. 220*, 425-433.
* Dobson, C. M. (1991) *Current Opinion in Structural Biology 1*, 22-27.

* Dobson, C. M. (1992) *Current Opinion in Structural Biology 2*, 6-12.

* Englander, S. W., Downer, N. W., & Teitelbaum, H. (1972) *Annu. Rev. Biochem. 41*, 903-924

* Ferderigos, N., Burke, G. T., Kitagawa, K. & Katsoyannis, P. G. (1983) *J. Protein Chem. 2*, 147-170.

* Geiger, R., Geisen, K., Regitz, G., Summ, H.-D., & Langner, D. (1980) *Hoppe-Seyler's Z. Physiol. Chem. 361*, 563-570.

* Geiger, R., Geisen, K., & Summ, H.-D. (1982) *Hoppe-Seyler's Z. Physiol. Chem. 363*, 1231-1239.

* Goldman, J., & Carpenter, F. H. (1974) *Biochemistry 13*, 4566-4574.

* Goto, Y., & Fink, A. L. (1989) *Biochemistry 28*, 945-952.

* Hiyama, Y., Niu, C., Silverton, J. V., Bavoso, A., & Torchia, D. A. (1988) *J. Am. Chem. Soc. 110*, 2378.

* Hua, Q. X., Shoelson, S. E., Kochoyan, M., & Weiss, M. A. (1991) *Nature 354*, 238-241.

* Hua, Q. X., Kochoyan, M., & Weiss, M. A. (1992a) *Proc.Natl.Acad.Sci. USA* 89, 2379-2383.

* Hua Q. X., Shoelson, S. E., & Weiss, M. A. (1992b) *Biochemistry 31*, 11940-11951.

* Hua, Q. X., Ladbury, J. E., & Weiss, M. A. (1993a) *Biochemistry* 32, 1433-1442.

* Hua, Q. X., Shoelson, S. E., Inouye, K., & Weiss, M. A. (1993b) *Proc. Natl. Acad. Sci. 90*, 582-586.

* Hybert, S. G., Märki, W., and Wagner, G. (1987) *Eur. J. Biochem. 164*, 625-635.

* Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546-4553.

* Jeffrey, P. D., & Coates, J. H. (1966) *Biochemistry 5*, 489-498.

* Jeffrey, P. D., Milthorpe, B. K., & Nichol, L. W. (1976) *Biochemistry 15*, 4660-4665.

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

* Jørgensen, A. M., Olsen, H. B., Led, J. J., & Balschmidt, P. (1996) *J. Mol. Biol. in press*

* Kaarsholm, N. C., Ko, H.-C., & Dunn, M. F. (1989) *Biochemistry 28*, 4427-2235.

* Kaarsholm, N. C., Havelund, S., & Hougaard, P. (1990) *Arch. Biochem. Biophys. 283*, 496-502.

* Kaarsholm, N. C., Norris, K., Jørgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sørensen, A. R., & Havelund, S. (1993) *Biochemistry 32*, 10773-10778.

* Kaarsholm, N. C., & Ludvigsen, S. (1995) *Receptor 5*, 1-8

* Kadima, W., Roy, M., Lee, R. W.-K., Kaarsholm, N. C., & Dunn, M. F. (1992) *J. Biol. Chem. 267*, 8963-8970.

* Kay, L. E., Torchia, D. A. & Bax, A. (1989) *Biochemistry 28*, 8972-8979.

* Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A. & Torchia, D. A. (1992) *J. Magn. Reson. 97*, 359-375.

* Kjær, M., Andersen, K. V., Shen, H., Ludvigsen, S., Windekilde, D., Sørensen, B., & Poulsen, F. M. (1991) *NATO ASI Series* (Hoch, J. C., Redfield C., & Poulsen, F. M., Eds.) Plenum, New York.

* Kline, A. D. & Justice, R. M. Jr. (1990) *Biochemistry 29*, 2906-2913.

* Knegtel, R. M. A., Boelens, R., Ganadu, M. L., & Kaptein, R. (1991) *Eur. J. Biochem. 202*, 447-458.

* Kraulis, P. J., (1991) *J. Appl. Cryst. 24*, 946-950.

* Kuszewski, J., Hilges, M., & Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33-56.

* Lipari, G., & Szabo, A. (1982a) *J. Am. Chem. Soc. 104*, 4546-4559.

* Lipari, G., & Szabo, A. (1982b) *J. Am. Chem. Soc. 104*, 4559-4570.

* Lord, R. S., Gubensek, F., & Rupley, J. A. (1973) *Biochemistry 12*, 4385-4392.

* Ludvigsen, S., Andersen, K. V., & Poulsen, F. M. (1991) *J. Mol. Biol. 217*, 731-736.

* Ludvigsen, S., Roy, M., Thøgersen, H., & Kaarsholm, N C. (1994) *Biochemistry 33*, 7998-8006.

* Marion, D. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Comm. 117*, 486-492.

* Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A.M., & Clore, G. M. (1989a) *Biochemistry 28*, 6150-6156.

* Marion, D., Kay, L. E., Sparks, S. W:, Torchia, D. A., & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515-1517.

* Mark, A. E., Nichol, L. W., & Jeffrey, P. D. (1987) *Biophys. Chem. 27*, 103-117.

* Märki, F., Gasparo, M. D., Eisler, K., Kamber, K., Riniker, B., Rittel, W., & Sieber, P. (1979) Hoppe-Seyler's Z. Physiol. Chem. 360, 1619-1632.

* Markussen, J., Jørgensen, K. H., Sørensen, A. R., & Thim, L. (1985) *Int. J. Peptide Protein Res. 26,* 70-77.

* Markussen, J., Diers, I., Engesgaard, A., Hansen, M. T., Hougaard, P., Langkjaer, L., Norris, K., Ribel, U., Snel, L., Sørensen, A. R., Sørensen, E., & Voigt, H. O. (1987) *Protein Eng. 1*, 215-223.

* Meiboom, S., & Gill, D. (1958) *Rev. Sci. Instrum.* 29, 688.

* Moody, A. D., Stan M. A., Stan, M. & Gliemann, J. (1974) *Hormone Metab. Res. 6*, 12-16.

* Morris, J. W. S., Mercola, D., & Arquilla, E. R. (1968) *Biochim. Biophys. Acta 160*, 145-155.

* Morris, G. A., & Freeman, R. (1979) *J. Am. Chem. Soc.* 101, 760.

* Nakagawa, S. H., & Tager, H. S. (1992) *Biochemistry 31*, 3204-3214.

* Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 139*, 317-324.

* Olsen, H. B., Ludvigsen, S., & Sørensen, O. W. (1993a) *J. Magn. Reson. A104*, 226-230.

* Olsen, H. B., Ludvigsen, S., & Sørensen, O. W. (1993b) *J. Magn. Reson. A105*, 321-322.

* Otting, G., Liepinsh, E., Wüthrich, K. (1993) *Biochemistry 32*, 3571-3582.

* Pekar, A. H., & Frank, B. H. (1972) *Biochemistry 11*, 4013-4016.

* Pervushin, K. V., Orekhov, V. Y., Korzhnev, D. M., & Arseniev, A. S. (1995) *J. Biomol. NMR 5*, 383-396.

* Piantini, U., Sørensen O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.

* Pocker, Y., & Biswas, B. (1981) *Biochemistry 20*, 4354-4361.

* Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blundell, T. L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., Gliemann, J., Gammeltoft, S. (1976) *Nature 259*, 369-373.

* Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Comm. 117*, 479-485.

* Roy, M., Lee, R. W. K., Brange, J., & Dunn, M. F. (1990a) *J. Biol. Chem. 265*, 5448-5452.

* Roy, M., Lee, R.W.-K., Kaarsholm, N.C., Thøgersen, H., Brange, J., & Dunn, M.F. (1990b) *Biochim. Biophys. Acta* 1053, 63-73.

* Smith, G.D., Swenson, D.C., Dodson, E.J., Dodson, G.G., & Reynolds, C.D. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7093-7097.

* Strazza, S., Hunter, R., Walker, E., & Darnall, D. W. (1985) *Arch. Biochem. Biophys. 238*, 30-42.

* Strickland, E. H., & Mercola, D. (1976) *Biochemistry 15*, 3875-3884.

* Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gõ, & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611-639.

* Weiss, M. A., Hua, Q.-X., Lynch, C. S., Frank, B. H., & Shoelson, S. E. (1991) *Biochemistry 30*, 7373-7389.

* Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647-1651.

* Wittingham, J. L., Chaudhuri, S., Dodson, E. J., Moody, P. C. E., & Dodson, G. G. (1995) *Biochemistry 34*, 15553-15563.

* Wood, S. P., Blundell, T. L., Wollmer, A., Lazarus, N. R., & Neville, R. W. J. (1975) *Eur. J. Biochem. 55*, 531-542.

* Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

* Zuiderweg, E. P. R., & Fesik, S. W. (1989) *Biochemistry 28*, 2387-2391.

Appendix A

Solution Structure of an Engineered Insulin Monomer at Neutral pH

Helle B. Olsen*, Svend Ludvigsen, and Niels C. Kaarsholm

Novo Research Institute, Novo Nordisk A/S, Novo Allé 6A,
DK-2880 Bagsværd, Denmark

Running title : Insulin Monomer Structure at Neutral pH

FOOTNOTES

[1]Abbreviations: NMR, nuclear magnetic resonance; CD, circular dichroism; COSY, two-dimensional correlated spectroscopy; DQF, double quantum filtered; FID, free induction decay; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhanced spectroscopy; TOCSY, two-dimensional total correlation spectroscopy; rms, root mean square.

ABSTRACT

Insulin circulates in the bloodstream and binds to its specific cell-surface receptor as a 5808 Dalton monomeric species. However, studies of the monomer structure and dynamics in solution are severely limited by insulin self-association into dimers and and higher oligomers. In the present work we use site-directed mutagenesis of the dimer- and hexamer-forming surfaces to yield the first insulin species amenable for structure determination at neutral pH by nuclear magnetic resonance (NMR) spectroscopy. The preferred insulin mutant, i.e., (B1, B10, B16, B27) Glu, des-B30 insulin retains 47% biological potency, and remains monomeric at millimolar concentrations in aqueous solution at pH 6.5-7.5 as judged by NMR and near-UV circular dichroism (CD) spectroscopy. From a series of 2D ^{1}H-NMR spectra collected at pH 6.5 and 34°C, the majority of the resonances are assigned to specific residues in the sequence, and nuclear Overhauser enhancement (NOE) cross-peaks are identified. NOE-derived distance restraints in conjunction with torsion restraints based on measured coupling constants, $^3J_{H^N H^\alpha}$, are used for structure calculations using the hybrid method of distance geometry and simulated annealing. The calculated structures show that the major part of the insulin mutant is structurally well-defined with an average root mean square (rms) deviation between the 25 calculated structures and the mean coordinates of 0.66 Å for backbone atoms (A2-A19 and B4-B26) and 1.31 Å for all backbone atoms. The A-chain consists of two antiparallel helices, A2-A7 and A12-A19, connected by a loop. The B-chain contains a loop region (B1-B8), an α-helix (B9-B19), a type I turn (B20-B23), and terminates as an extended strand (B24-B29). The B1-B4 and B27-B29 regions are disorded in solution. The structure is generally similar to crystal structures and resembles a crystalline T-state more than a R-state in the sense that the B-chain helix is confined to residues B9-B19.

Insulin is central to the hormonal control of metabolism. Due to its importance as a pharmaceutical preparation for the treatment of diabetes mellitus, much effort has been directed towards understanding the structural basis for insulin bioactivity. The protein is composed of two polypeptide chains, the A-chain (21 residues) and the B-chain (30 residues). The two chains are covalently linked by disulfide bridges at A7-B7 and A20-B19, and an intrachain disulfide bridge is joining A6 and A11. The three-dimensional structure of insulin has been characterized in detail by X-ray analysis of aggregated species, notably zinc insulin hexamers (Adams et al., 1969; Smith et al., 1984; Derewenda et al., 1989). The three principal hexamer conformations have been designated the T_6, T_3R_3, and R_6 forms, respectively (Kaarsholm et al., 1989). In each of these structures, the A-chain folds into a helix-loop-helix motif with helical stretches located in A2-A8 and A12-A19. The B-chain can assume two distinct conformations. In the T-state, the B-chain contains an extended N-terminal arm, central α-helix (B9-B19), type I turn (B20-B23), and C-terminal β-strand. In the R-state, residues B1-B9 take up a helical conformation to form a region of α-helix contiguous from B1-B19. The interconversion between the T_6, T_3R_3, and R_6 states of the insulin hexamer is modulated by a set of homotropic and heterotropic ligand binding interactions and has been shown to take place in solution (Kaarsholm et al., 1989; Brader et al., 1991; Bloom et al., 1995).

Although insulin hexamers are extensively characterized by X-ray crystallography, the physiologically active form of the hormone is the 5808 Dalton monomer. Hence, in any discussion of structure-activity relationships for insulin, it is necessary to consider whether crystal packing forces have modified the structure from that required for biological action. Due to the complicated pattern of insulin self-association in solution, detailed NMR analysis of the

insulin monomer has often been ambiguous. Accordingly, several groups have reported results of NMR studies carried out at low pH (i.e., pH 1.8-3.5) using either modified insulins, organic co-solvents, or both, in an effort to counteract self-association (see e.g. Kline & Justice, 1990; Knegtel et al., 1991; Hua et al., 1991, 1992a, 1993b; Jørgensen et al., 1992; Ludvigsen et al., 1994). The reported structures generally agree that the secondary structure of the insulin monomer in solution is similar to that of the crystallographically identified T-state. However, considerable differences are apparent in terms of the structural resolution expressed e.g. by the atomic rms[1] values.

In one extreme, insulin structures with properties similar to a 'molten-globule' state have been reported. On the basis of these structures determined in the presence of 20% acetic acid, it has been postulated that the lack of tertiary structural detail is intrinsic to the native insulin monomer and that the phenomenon per se is important for the interactions of insulin with its receptor (Hua et al., 1991, 1992a,b, 1993a,b). In contrast with these low resolution structures, our NMR studies on the biologically active B16 Tyr → His mutant in water at low pH have revealed a well-defined solution structure (Ludvigsen et al., 1994; Kaarsholm & Ludvigsen, 1995). Because the B16 His mutant remains essentially monomeric at millimolar concentrations in aqueous solution at low pH, these results identify sample homogencity and aggregation state as major determinants for the quality of the NMR-derived structure.

The high resolution solution structure of the insulin monomer at low pH provides an important prerequisite for the understanding of the interaction between insulin and its receptor. However, a long-term goal for these NMR studies is to determine the structure and dynamics of

engineered monomers under physiologically conditions as a base line for further structural examination of mutations associated with enhanced or diminished affinity for the insulin receptor.

Detailed NMR studies of the insulin monomer at neutral pH have so far been limited by protein solubility and self-association in addition to the inherently faster rate of NH exchange in this pH region. In the present work we use site-directed mutagenesis to manipulate the solubility and aggregation pattern of insulin in the neutral pH region. These experiments identify a mutant and a set of conditions where the monomer is amenable for high resolution NMR structural analysis. The preferred mutant has glutamate residues substituted into four positions, i.e., B1 Phe, B10 His, B16 Tyr, and B27 Thr and the C-terminal B30 Thr is removed. This mutant retains 47% biological potency and is monomeric at millimolar concentrations in the pH range 6.5-7.5. The resulting structure is generally well-defined as evidenced by a high number of sequential NOEs as well as many long range NOEs. Apparent disorder is observed near the termini of the B-chain, e.g. B1-B4 and B27-B29, and is to a large extent ascribed to local effects of the mutations.

MATERIALS AND METHODS

Materials. Native and mutant insulins were constructed by oligonucleotide-directed mutagenesis, fermented in yeast, and purified as described (Markussen et al., 1987; Brange et al., 1988). In a typical small-scale preparation, the mutant is expressed and partly purified as a single-chain mini-proinsulin precursor, B1..B29Lys-Ala-Ala-Lys-A1Gly..A21. Prior to the final purification step, the connecting peptide is cleaved off by treatment with a lysyl endopeptidase (Achromobacter Protease I, EC 3.4.21.50; Wako Inc., Japan), and the mutant insulin is isolated in the form of des-B30 Thr insulin. The removal of the B30 Thr residue has no effect on the biological potency of the molecule.

CD Spectroscopy. CD spectra were recorded with a Jobin Yvon Mark V dichrograph calibrated with (+)-10-camphorsulfonic acid as described (Kaarsholm et al., 1993). Near-UV CD spectra were recorded between 250 and 350 nm using an appropriate combination of cell pathlength and protein concentration to yield an absorbancy of less than 1. Protein concentrations were determined by UV absorbance using ε_{276}= 6.2 × 10^3 $M^{-1}cm^{-1}$. The same extinction coefficient was used for estimation of the concentration of mutant species with the assumption that each of the four tyrosines in human insulin contributes 25% to ε_{276}.

NMR Spectroscopy. Samples were prepared by dissolving the lyophilized protein powder in 10/90 D_2O/H_2O or 99.8 % D_2O and adjusting the pH as desired by additions of small amounts of 1 M DCl or NaOD. All pH meter readings are without correction for isotope effects. For all NMR experiments reported herein a temperature of 307 K was used.

Two dimensional ^1H-^1H NMR spectra, DQF-COSY (Piantini et al., 1982; Rance et al., 1983), NOESY (Jeener et al., 1979; Anil-Kumar et al., 1980, 1981) and TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) were recorded on a Bruker AMX600 spectrometer. For TOCSY and NOESY spectra, mixing times were between 40-90 and 120-180 ms, respectively. All spectra had a spectral width in both dimensions of 6579 Hz; 1024 t_1 increments were acquired each with a size of 2048 real datapoints. The spectra were recorded in the phase-sensitive mode using the time proportional phase incrementation scheme (TPPI, Marion & Wüthrich, 1983). The carrier was placed on the water resonance to enable irridation of the water during a period of 1.5 s between the individual scans.

Prior to Fourier transformation the FID's were zero filled once in both dimensions. For resolution enhancement of the DQF-COSY spectra, a squared sine-bell shifted 90° was used in both dimensions, whereas for the NOESY and TOCSY spectra, a Gaussian function with an exponential line broadening of -7 Hz and a factor of 0.15 was applied. Data processing was performed using the MNMR package (PRONTO Software Development and Distribution, Copenhagen, Denmark) or the Bruker UXNMR software on a Silicon Graphics Indigo computer. Exchange of amide protons was followed by a series of 8K datapoint one-dimensional spectra. These spectra were zero filled once, and a 1.5 Hz line broadening was used as a window multiplication prior to the Fourier transformation.

The program PRONTO (PRONTO software Development and Distrubution; see Kjær et al., 1991) was used to keep track of spectral assignments, cross-peak integration, sequence-

specific and stereo-specific assignments and related bookkeeping during spectral analysis. $^3J_H{}^N{}_H{}^\alpha$ coupling constants were measured by the facility in the PRONTO software which uses a combined analysis of COSY and NOESY spectra (Ludvigsen et al., 1991). Chemical shifts were measured in parts per million as observed relative to dioxane (3.75 ppm).

Structure Calculations. The program X-PLOR (Brünger, 1992) was used to calculate structures based on distance- and dihedral angle restraints derived from the NMR spectra. Integrated NOESY cross peaks were divided into three classes of distance restraints using 1.0 Å as lower limit and upper limits of 2.7, 3.3 and 5.5 Å, respectively. For restraints involving methyl groups, an additional 0.5 Å was added to the upper limit. The measured $^3J_H{}^N{}_H{}^\alpha$ coupling constants were converted into ϕ-angle restraints as follows: -60°± 30° (2-4 Hz); -70° ± 30° (4-6 Hz); -120° ± 60 (6-8 Hz); and -120° ± 35° (8-10 Hz). The stereospecific assignment of several β methylene protons was obtained by combined analysis of COSY cross peak patterns (Hybert et al., 1987) and of the intraresidual NOE intensities between the methylene protons, the α- and the amide protons, respectively (Wagner et al., 1987). This method allowed assignment of the side-chain to one of the three staggered conformations, χ_1, assuming the values -60° ± 60°, 60° ± 60°, or 180° ± 60°. Stereospecific assignment of δ methyl groups of Leu residues were found and converted to χ_2 restraints in a similar fashion. The structure calculations were performed using a combination of distance geometry (Crippen & Havel, 1988; Kuszewski et al., 1992) and simulated annealing as proposed by Nilges et al. (1988).

Comparison of solution structures. From the Protein Data Bank (PDB), Brookhaven (January 1996) all files containing structures of monomeric native and mutant insulins

determined by NMR were collected. Each file contains an ensemble of structures as well as a representative structure. For the purpose of comparison of the different structures, geometric average structures were calculated from these ensembles of structures following the alignment procedures presented in the literature references as listed in the header of the files. In order to obtain a reasonable comparison, we have made one exception to this procedure and used both chains rather than just the B-chain helix in the alignment of the DPI structures. Table III gives a summary of the experimental conditions and details concerning the structural calculations as referenced in the PDB files and in the original publications. The program X-PLOR (Brünger, 1992) was used to calculate the average structures.

RESULTS

Design of Monomeric Insulin Suitable for 2D NMR at Neutral pH. The aggregation and precipitation pattern of insulin is a complex function of protein concentration, pH, temperature, metal ions, ionic strength, and solvent composition. In the millimolar concentration range usually required for high-resolution NMR structural work, metal-free insulin exists predominantly as dimers and higher oligomers at pH 2.0-3.5. At higher pH aggregation increases, which in turn leads to precipitation in the pH region between 4.2 and 6.6. Above pH 6.6 insulin dissolves, presumably as a mixture of dimers and higher aggregates, which then gradually dissociate as the pH increases. As the pH is raised above 10-11, the monomeric state is finally reached (Jeffrey & Coates, 1966; Pekar & Frank, 1972; Lord et al., 1973; Goldman & Carpenter, 1974; Jeffrey et al., 1976; Pocker & Biswas, 1981; Strazza et al., 1985; Mark et al., 1987; Kaarsholm et al., 1990; Roy et al., 1990a; Kadima et al., 1992). The X-ray structure of the T_6 zinc insulin hexamer shows that distinct interfaces are involved in insulin dimer and hexamer packing (Baker et al., 1988). The published equilibrium constants for solution aggregation along the monomer/monomer interface are similar at pH 2 and pH 7 (i.e., $\approx 10^5\ M^{-1}$). However, the corresponding aggregation along the dimer/dimer interface appears much stronger at pH 7 than at pH 2 (Mark et al., 1987). Hence, in contrast with the situation at low pH (Ludvigsen et al., 1994), insulin self-association must be inhibited along two distinct interfaces in order to obtain monomers at millimolar concentrations at neutral pH. At the same time, the resulting mutant should retain near-native biological potency and the solubility pattern must allow NMR measurement in a pH range where the exchange rate of backbone NH is sufficiently slow for observation with conventional ^{1}H-NMR techniques (Wüthrich, 1986).

Aggregation properties of various insulin mutants were evaluated by near-UV circular dichroism (CD) and NMR resonance line widths. The near-UV CD spectrum (350-250 nm) of insulin reflects the environment of the tyrosine chromophore. The signal is very sensitive to aggregation and may be used to monitor the formation of dimer interface (Morris et al., 1968; Goldman & Carpenter, 1974; Wood et al., 1975; Strickland & Mercola, 1976). NMR spectra are sensitive to insulin self-association because longer rotational correlation times lead to line broadening and because dynamic equilibria between oligomeric states can lead to line broadening in the intermediate-exchange regime.

Figure 1 shows the aromatic and amide proton resonances of the 1D ^{1}H-NMR spectra for human insulin (panel A) and a series of mutants (panels B-F) at 1 mM protein in 10/90 D_2O/H_2O, pH 7.5. Under these conditions, human insulin exists as a mixture of dimers and higher aggregates and the resulting spectrum is very poorly resolved. The X-ray structure of the T_6 insulin hexamer shows that dimer formation is dominated by a series of non-polar contacts contributed by B-chain residues, notably B12 Val and B16 Tyr in the central helix, and residues B23-B28 in the extended chain. Panels (B-C) illustrate the effects of introducing charged residues at two different positions in the monomer/monomer interface. As shown in Panel (B), the B27 Thr → Glu mutation at the edge of the monomer/monomer interface effects little or no improvement in spectral resolution relative to human insulin. This mutation has a minor inhibitory effect on dimer formation (Brange et al., 1988); however, in the present study the modification is preferred due to its positive effect on the expression level in yeast and hence on the overall fermentation yield during small-scale mutant preparation. Introduction of charge into

the B16 Tyr position has previously been shown to provide an efficient means of inhibiting the formation of dimer interface at low pH (Ludvigsen et al., 1994; Kaarsholm & Ludvigsen, 1995). In accordance with this result, panel (C) shows that the B16 Tyr → Glu mutation strongly enhances the spectral resolution at pH 7.5. While the (B16, B27) Glu species is nearly monomeric at 1 mM protein, the exchange rate of several NH-protons is too fast for observation in 2D spectra. At slightly lower pH, extensive line broadening is observed in concert with aggregation along the dimer/dimer interface. Accordingly, the number of cross-peaks observed in 2D NOESY spectra of the (B16, B27) Glu mutant is significantly smaller at pH 7.0 than at pH 7.5 (data not shown).

When the T_6 hexamer is assembled from dimers through the coordination of Zn^{2+} to the B10 His residues, both polar and non-polar residues are buried between the dimers. The packing is correspondingly much looser than in the monomer/monomer interface within each dimer. A set of important contacts across the dimer-dimer interface involves the N-terminal part of the B-chain. Of particular interest is the B1 Phe residue, which fit into a pocket between the main A-chain and A14 Tyr residue of its neighbor. Because the two A14 Tyr are also in contact across the dimer-dimer interface, a close aggregate involving four aromatic residues is formed (Baker et al., 1988). Comparison of panels (B) and (D) in Figure 1 shows that the B1 Phe → Glu mutation leads to somewhat improved resolution of the 1D NMR spectrum. A slightly better improvement is obtained with the B10 His → Glu mutation positioned at the edge of the dimer/dimer interface, viz. panel (E).

The B10His → Glu mutation also has the effect of increasing the biological potency of the insulin molecule 4-fold as determined by the ability to incorporate [2-^{3}H]-glucose in isolated mouse adipocytes according to Moody et al. (1974). In the same assay, the B27 Glu, (B1, B27) Glu, (B16, B27) Glu and des-B30 insulins exhibit potencies of 1.07, 0.97, 0.13 and 1.00, respectively, relative to a value of 1.00 found for human insulin. For the mutant carrying all modifications, i.e. (B1, B10, B16, B27) Glu, des-B30 insulin, a relative potency of 0.47 is found indicating that the substitutions affect the biological activity in a nearly independent manner. Panel (F) in Figure 1 shows that the 1D spectrum of (B1, B10, B16, B27) Glu, des-B30 insulin is well-resolved and retains the dispersion of resonance lines characteristic for globular proteins. Furthermore, as the pH is adjusted down to 7.0-6.5, the high resolution is maintained (panels (G) and (H)), while selected NH resonances become sharper as expected due to the slowed exchange rate.

Figure 2 shows near-UV CD spectra as a function of protein concentration for native insulin at pH 8.0 and the (B1, B10, B16, B27) Glu, des-B30 mutant at pH 6.5 and 7.5. For native insulin, the progressive increase in intensity of the negative signal around 274 nm indicates the expected increase in association with increasing concentration. In contrast, the near-UV CD spectrum of the mutant is independent of the protein concentration in the 60 μM to 2.6 mM range at both pH 6.5 and 7.5. Although the mutant is missing one tyrosine reporter group (B16), Tyr B26 is expected to be strongly affected by monomer/monomer interactions. Hence, these results provide further evidence that the (B1, B10, B16, B27) Glu, des-B30 mutant is monomeric under conditions necessary for NMR structural studies at neutral pH, and this species was selected for detailed characterization.

Assignment of Spin Systems. NMR spectra were assigned using the standard procedures outlined by Wüthrich (1986). The fingerprint region of the DQF-COSY spectrum is shown in the upper panel of Figure 3. The chemical shifts are dispersed and well-resolved as expected for a structured globular protein, and 37 of the 47 possible H^NH^α cross peaks are annotated in the plot. Among the remaining H^NH^α cross peaks, six were assigned using TOCSY spectra. These are the resonances that are either close to the water line or unusually broad, i.e., A8, A9, A12, B5, B6, and B26. The linewidth of the CysA11 amide proton is unusually broad, but identification was possible in DQF-COSY spectra. Amide protons from residues A2, B6 and B9 were not identified. Finally, for the side-chains of Glu and Gln residues, considerable overlap of cross peaks in the H^β-H^γ area was resolved using TOCSY spectra occasionally supported by NOESY.

Sequential Assignment and Secondary Structure. The sequential assignment was straightforward for most parts of the insulin mutant. Including the regions B9-B29, A9-A12, and A13-A21. The assignment of the stretch from A4 to A8 was complicated by overlapping H^N and H^α resonances; this is reflected in the lack of sequential NOEs as well as NOEs indicative of secondary structure in this region. In the N-terminus of the B-chain, the NOEs are sparse and the assignments here were made by exclusion as the very last part of the procedure.

The first indications of the secondary structures come from inspection of the chemical shifts of the α-protons compared to the random coil values (Wishart & Sykes, 1992). These chemical shift differences shown in Figure 4 suggest helical structures characterized by upfield shifts of the α-protons in the second part of the A-chain (the A(II)-helix) and in the

central part of the B-chain (the B-helix), whereas the helix expected in the N-terminus of the A-chain (the A(I)-helix) seems less well defined. These observations are substantiated by the NOEs assigned in the same regions as shown in Figure 5 and in the NOE matrix in the lower panel of Figure 6. The B-helix, stretching from B9 to B19, is a well defined helix characterized by a dense network of NOEs between the α-protons in position i and the amide protons in position i+3 and position i+4 as well as the β-protons in position i+3. Furthermore, stronger NOEs between amide protons 2 residues apart in the helix are found. The C-terminal A(II)-helix comprising the amino acids from A13 to A20 has a less dense network of NOEs, i.e., only between α-protons in position i and β- and amide protons in position i+3. The N-terminal A(I)-helix stretches from A2 to A6 and is characterized by the same type of NOEs as the A(II)-helix. In the lower panel of Figure 3, the H^N to H^N part of the NOESY spectrum is shown with annotations for the sequential assignment of the helical areas.

An amide proton characterized by an exchange rate slower than the average is a good evidence for participation in a hydrogen bond either in a secondary structural element or as part of the tertiary fold of the protein. At near neutral pH, the exchange rate of amide protons is in the order of 10^2 min^{-1} (Englander et al., 1972). From 1D proton spectra recorded in the period between 12 and 60 minutes after dissolution of the insulin mutant in D_2O, it is possible by visual inspection to follow the decay of amide proton resonances belonging to the B-helix and the A(II)-helix as well as a single amide proton in the A(I)-helix. After 50 minutes, all amide protons are exchanged by deuterons. The locations of the slow exchanging amide protons according to this criterion are shown in Figure 5.

Structure Calculations. Table II lists the number of distance restraints and dihedral angle restraints used for the final structure calculations. In addition, the sulfide bridges were added as distance restraints, while no restraints were added for hydrogen bonds. The matrix in the lower panel of Figure 4 shows the distribution of NOEs between residues. Multiple occurences of NOEs from both sides of the diagonal in a spectrum or from different spectra were filtered out before the conversion of NOEs to distance restraints. As a supplement, the upper panel of the same Figure depicts the number of NOEs by residue. The distribution of dihedral angle restraints on residues is shown in Figure 5. The calculation of structures using X-PLOR 3.0 proceeded as described in Materials and Methods starting from a reduced set of distance restraints. In a sequence of iterations, this set was slowly expanded as ambiguities could be resolved, and finally the dihedral angle restraints were introduced. A total of 100 structures were calculated 25 of which were characterized by distance restraints violations below 0.3 Å, dihedral angle restraints violations below 2°, and a low total energy. Table II summarizes the structural statistics for this set of structures.

The set of 25 converged structures is depicted in Figure 7 (backbone representation) together with their geometric average represented by a ribbon. In Figure 8, the side-chains have been included, and the average structure shown here is colored according to the atomic rms deviation.

Description of the Structures. The major part of the mutant structure is well-defined. This goes indeed for the B-helix (B9-B19) which exhibits low rms deviation among the structures as expected from the large number of structural NOEs as well as the small coupling

constants, indicative of helix structure. The hydrogen bonds fit a regular α-helix pattern all the way starting from B9(CO)···B13(H^N) and ending at the last donor B19(H^N).

At the C-terminal end of the helix, a type I turn from B20 to B23 turns into a β strand that stretches along the central helix. The three C-terminal residues are disordered. As is evident from Figure 4, no NOE structural information was obtained in the B1-B4 region of the B-chain. This fact is reflected in the random distribution of the termini among the 25 structures presented in Figure 7 and implies that the average structure calculated in this area is physically meaningless. The A(II)-helix (A13-A20) shows a high degree of accordance within the bundle of structures, and the hydrogen bonding pattern fits an α-helix A13(CO)···A17(H^N), A15(CO)···A19(H^N), and A16(CO)···A20(H^N), starting out with the irregular A12(CO)···A15(H^N) hydrogen bond. The A(I) helix (A2-A8) is a little less well-defined, a single hydrogen bond between A2(CO) and A6(H^N) is established. The loop area (A9-A12) connecting the two A-chain helices shows a larger spread of conformations. In 20 out of 25 calculated structures the interchain hydrogen bond A19(CO)···B25(H^N) is established, in agreement with crystal structures. The atomic rms deviations calculated are presented in Table II along with the general structural statistics.

DISCUSSION

The (B1, B10, B16, B27) Glu, des-B30 mutant is the first example of an insulin species that remains monomeric at millimolar concentration in aqueous solution at neutral pH. Hence, by tailored mutations in two dinstict protein/protein interfaces, the inherent self-association of insulin is strongly inhibited, while the biological potency is essentially retained. As a result, samples may be prepared under conditions where the exchange rate of backbone amide protons is sufficiently slow for detailed NMR analysis. Previous NMR investigations of native and mutant insulins at near-neutral pH have been limited by incomplete assignments of the spectra. Nevertheless, on the basis of 1D spectra of human insulin at high dilution and of the B9Ser → Asp monomeric mutant in the pH 8-9.5 range, Roy et al. (1990a,b) were able to show that the association of insulin monomers into dimers is accompanied by a change in conformation involving the relative position of residues B15 Leu and B24 Phe. In their NMR analysis of the B10 Asp, B28 Lys, B29 Pro (DKP) insulin monomer at pH 8.0, Weiss et al. (1991) also reported partial assignments and employed selective ^2H and ^{13}C labeling of the B23-B26 residues to confirm that this region is indeed in close contact with residues from the central B-chain helix.

The (B1, B10, B16, B27) Glu, des-B30 mutant is structurally well-ordered with the exception of the B-chain termini, i.e., residues B1-B4 and B27-B29. The observed disorder in the N-terminus is most reasonably ascribed to opposing effects of the B1 and B10 mutations. Removal of the hydrophobic B1 Phe side-chain destabilizes the packing of the B1 residue against the A13 Leu region (Ludvigsen et al., 1994), while the B10 Asp side-chain caps the

central B9-B19 helix and hence stabilizes the T-state relative to the R-state (Kaarsholm et al., 1993). In any event, because the B-chain helix is clearly confined to residues B9-B19, the overall structure resembles a crystallographic T-state more than a R-state. In the C-terminal B29-B30 residues, disorder is observed in the crystal structure as well as in the low-pH solution structure (Baker et al., 1988; Ludvigsen et al., 1994). In the present work at neutral pH, the lack of structural definition extents to the last three residues, B27-B29. Again, the effect is most likely due to the modifications employed, i.e., the introduction of negative charge close to the shortened C-terminus (B27 Thr → Glu , des-B30). In accordance with previous work (Kline & Justice, 1990; Ludvigsen et al., 1994), broad amide proton resonances are observed in the A-chain loop region suggesting that conformational substates exchange on a millisecond time scale in this particular region.

Comparison with Other Monomer Solution Structures. Several NMR investigations of insulin monomers have been carried out at low pH and/or in the presence of organic co-solvent. Table III compares experimental conditions and details concerning the structure calculations for the (B1, B10, B16, B27) Glu, des-B30 mutant and for other insulin monomer structures deposited in the Brookhaven Data Bank. In addition to these monomers, the structure of a B9Asp insulin dimer has been described (Jørgensen et al., 1992), and the low-pH structure of DPI has been determined by Knegtel et al. (1991) as well as by Hua et al. (1992a). Figure 9 compares the three dimensional structure of (B1, B10, B16, B27) Glu, des-B30 insulin with those of the other monomers. The structures are color coded according to the atomic rms deviations characterizing each of these ensemble average structures. Table III contains details concerning the average structure calculations, alignment procedures and rms deviations of the

backbone calculated by using the same alignment procedure for all mutants. Note that for certain parts of these structures, distance restraints containing information about secondary and tertiary fold are either absent or very few in number. This is particularly true for the C-terminal B22-B30 residues of the B24 Gly and B24 Ser mutants (Hua et al., 1991, 1993b). As a result, considerable atomic rms deviations are obtained which make the average structures physically meaningless in these parts colored red in Figure 9. Paradoxically, a distinct bulge appears on the C-terminal β-strand of the B24 Ser structure.

The precision of the individual structures is quantified by the ensemble average rms deviations in Table III. A comparison of the two columns of backbone rms deviations in the Table reveals that most of the differences are caused by imprecision of the C-terminal of the B-chain. The B16 His structure shows the highest degree of overall accordance between the individual structures in the ensemble. In this case, the presicion is directly correlated with the number of interresidual distance restraints used in the structure calculation. However, due to differences in the procedures for counting of the NOE's and for the translation of NOE's into distance restraints, the correlation between structural precision and the number of NOE's is generally not straightforward. The high precision of side-chain structures in B16 His insulin are closely related to the number of stereospecific assignments of χ_1 and χ_2 dihedral angles. The individual monomer structures are compared one by one in Table IV. It is clear that the B16 His insulin has the closest structural resemblance to the (B1, B10, B16, B27) Glu, des-B30 mutant. The major difference between these two structures is the lack of definition of residues B1-B4 in (B1, B10, B16, B27) Glu, des-B30 insulin. In B16 His insulin, the close contact between B-chain N-terminus and the A-chain loop region provides the set of NOEs defining the spatial

arrangement of the loop. Thus, the lack of structure in residues B1-B4 propagates to an increased inpresicion of the A-chain loop region in (B1, B10, B16, B27) Glu, des-B30 insulin.

Finally, we note that with the exception of the B-chain termini, the structures depicted in Figure 9 share common features including the secondary structure elements and the overall spatial arrangement of these elements into the tertiary structure. As would be expected for a globular protein, Table III shows that each structural core is characterized by a low rms deviation (high precision) relative to the precision of the overall structure. In contrast with these calculations, Hua et al. (1992a, 1993a) reported that variations in the mutual spatial arrangement of the two chains in the case of DPI span the differences reported among various insulin crystal forms. The phenomenon was ascribed to the absence of restraints and interpreted as evidence that DPI exists in a 'molten globule' state under native-like conditions (Hua et al., 1992a, 1993a). A molten globule state is usually characterized as a compact denatured state arising as an intermediate in the pathway from native to unfolded protein under slightly denaturing conditions (see e.g. reviews by Dobson (1992,1993)). The state is characterized by the presence of secondary structural domains, but absence of tertiary interactions, i.e., the conformation of the side-chains is essentially random. In a NMR spectrum, one feature of a molten globule state would be a reduced dispersion of side-chain resonances in comparison with the native state (Goto & Fink, 1989). The original alignment of the DPI ensemble presented by Hua et al. (1992a) was based on a subset of the structure, i.e., the central B-chain helix (residues B9-B19). This procedure per se leads to an accumulation of the overall structural imprecision in the parts that are left out of the alignment, in this case the entire A-chain and the B-chain termini. In contrast, our alignment procedure (Table III) is based on both chains leaving out only those parts

of the molecule where the number of restraints per residue is well below the average, i.e., the outermost parts of both chains. The same procedure was used by Knegtel et al. (1991) for the alignment of DPI. With this procedure, the major part of all molecules (including DPI) exhibit reasonably well-ordered secondary and tertiary structure. For the (B1, B10, B16, B27) Glu, des-B30 mutant at neutral pH and the B16 His mutant at low pH, the dispersion of side-chain resonances corresponds to that expected for a packed, globular protein. Furthermore, both mutants exhibit a distinct near-UV CD spectrum very similar to that of native human insulin (Figure 2 and Ludvigsen et al., 1994). Hence, there seems to be no evidence coming from NMR or CD to support the idea that insulin mutants exist in a molten globule state characterized by variations in the mutual arrangement of the chains.

Acknowledgments

We thank I. Diers, S. Havelund, A.-M. Kolstrup and L.G. Andersen for fermentation and purification of mutant insulins, we thank A. R. Sørensen for insulin potency measurements, and we thank A. Blom for assistance with operation and maintenance of the NMR spectrometer.

Reference

* Adams, M.J, Blundell, T.L., Dodson, E.J., Dodson, G.G., Vijayan, M., Baker, E.N., Harding, M.M., Hodgkin, D.C., Rimmer, R, & Sheet, S. (1969) *Nature 224*, 491-496.
* Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Comm. 95*, 1-5.
* Anil-Kumar, Wagner, G., Ernst, R. R., & Wüthrich, K. (1981) *J. Am. Chem. Soc. 103*, 3654-3658.
* Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M. Cutfield, Dodson, E. J., Dodson, G. G., D. M. C. Hodgkin, Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., & Vijayan, N. M. (1988) *Phil. Trans. Roy. Soc. ser. B 319*, 369-456.
* Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355-360.
* Bloom, C. R., Choi, W .E., Brzovic, P. S., Ha, J. J., Huang, S.-T., Kaarsholm,N. C., & Dunn, M. F. (1995) *J.Mol.Biol. 245*, 324-330.
* Brader, M. L., Kaarsholm, N. C., Lee R., W.-K., & Dunn, M. F. (1991) *Biochemistry 30*, 6636-6645.
* Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R., & Voigt, H. O. (1988) *Nature, 333*, 679-682.
* Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521-528.

* Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaninathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187-217.
* Brünger, A. T. (1992) *X-PLOR manual version 3.0*, Yale University, New Haven, CT.
* Crippen, G. & Havel, T. (1988) *Distance Geometry and Molecular Conformation.* Research Studies Press, Taunton, Somerset, England.
* Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swensen, D. (1989) *Nature 338*, 594-596.
* Dobson, C. M. (1991) *Current Opinion in Structural Biology 1*, 22-27.
* Dobson, C. M. (1992) *Current Opinion in Structural Biology 2*, 6-12.
* Englander, S. W., Downer, N. W., & Teitelbaum, H. (1972) *Annu. Rev. Biochem. 41*, 903-924
* Goldman, J., & Carpenter, F.H. (1974) *Biochemistry 13*, 4566-4574.
* Goto, Y., & Fink, A. L. (1989) *Biochemistry 28*, 945-952.
* Hua, Q. X., Shoelson, S. E., Kochoyan, M., & Weiss, M. A. (1991) *Nature 354*, 238-241.
* Hua, Q. X., Kochoyan, M., and Weiss, M. A. (1992a) *Proc.Natl.Acad.Sci. USA* 89, 2379-2383.
* Hua Q. X., Shoelson, S. E., & Weiss, M. A. (1992b) *Biochemistry 31*, 11940-11951.
* Hua, Q. X., Ladbury, J. E., and Weiss, M. A. (1993a) *Biochemistry* 32, 1433-1442.
* Hua, Q. X., Shoelson, S. E., Inouye, K., & Weiss, M. A. (1993b) *Proc. Natl. Acad. Sci. 90*, 582-586.
* Hybert, S. G., Märki, W., and Wagner, G. (1987) *Eur. J. Biochem. 164*, 625-635.

* Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546-4553.
* Jeffrey, P. D., & Coates, J. H. (1966) *Biochemistry 5*, 489-498.
* Jeffrey, P. D., Milthorpe, B. K., & Nichol, L. W. (1976) *Biochemistry 15*, 4660-4665.
* Jørgensen, A. M., Kristensen, S. M., Led, J. J., & Balschmidt, P. (1992) *J. Mol. Biol. 227*, 1146-1163.
* Kaarsholm, N. C., Ko, H.-C., & Dunn, M. F. (1989) *Biochemistry 28*, 4427-2235.
* Kaarsholm, N. C., Havelund, S., & Hougaard, P. (1990) *Arch. Biochem. Biophys. 283*, 496-502.
* Kaarsholm, N. C., Norris, K., Jørgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sørensen, A. R., & Havelund, S. (1993) *Biochemistry 32*, 10773-10778.
* Kaarsholm, N. C., & Ludvigsen, S. (1995) *Receptor 5*, 1-8
* Kadima, W., Roy, M., Lee, R. W.-K., Kaarsholm, N. C., & Dunn, M. F. (1992) *J. Biol. Chem. 267*, 8963-8970.
* Kjær, M., Andersen, K. V., Shen, H., Ludvigsen, S., Windekilde, D., Sørensen, B., & Poulsen, F. M. (1991) *NATO ASI Series* (Hoch, J. C., Redfield C., & Poulsen, F. M., Eds.) Plenum, New York.
* Kline, A. D. & Justice, R. M. Jr. (1990) *Biochemistry 29*, 2906-2913.
* Knegtel, R. M. A., Boelens, R., Ganadu, M. L., & Kaptein, R. (1991) *Eur. J. Biochem. 202*, 447-458.
* Kuszewski, J., Hilges, M., & Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33-56.
* Lord, R. S., Gubensek, F., & Rupley, J. A. (1973) *Biochemistry 12*, 4385-4392.

* Ludvigsen, S., Andersen, K. V., & Poulsen, F. M. (1991) *J. Mol. Biol. 217*, 731-736.
* Ludvigsen, S., Roy, M., Thøgersen, H., & Kaarsholm, N C. (1994) *Biochemistry 33*, 7998-8006.
* Marion, D. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Comm. 117*, 486-492.
* Mark, A. E., Nichol, L. W., & Jeffrey, P. D. (1987) *Biophys. Chem. 27*, 103-117.
* Markussen, J., Diers, I., Engesgaard, A., Hansen, M. T., Hougaard, P., Langkjaer, L., Norris, K., Ribel, U., Snel, L., Sørensen, A. R., Sørensen, E., & Voigt, H. O. (1987) *Protein Eng. 1*, 215-223.
* Moody, A. D., Stan M. A., Stan, M. & Gliemann, J. (1974) *Hormone Metab. Res. 6*, 12-16.
* Morris, J. W. S., Mercola, D., & Arquilla, E. R. (1968) *Biochim. Biophys. Acta 160*, 145-155.
* Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 139*, 317-324.
* Pekar, A. H., & Frank, B. H. (1972) *Biochemistry 11*, 4013-4016.
* Piantini, U., Sørensen O. W.,& Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.
* Pocker, Y., & Biswas, B. (1981) *Biochemistry 20*, 4354-4361.
* Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Comm. 117*, 479-485.
* Roy, M., Lee, R. W. K., Brange, J.,& Dunn, M. F. (1990a) *J. Biol. Chem. 265*, 5448-5452.
* Roy, M., Lee, R.W.-K., Kaarsholm, N.C., Thøgersen, H., Brange, J., & Dunn, M.F. (1990b) *Biochim.Biophys.Acta* 1053, 63-73.

* Smith, G.D., Swenson, D.C., Dodson, E.J., Dodson, G.G., & Reynolds, C.D. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7093-7097.
* Strazza, S., Hunter, R., Walker, E., & Darnall, D. W. (1985) *Arch. Biochem. Biophys. 238*, 30-42.
* Strickland, E. H., & Mercola, D. (1976) *Biochemistry 15*, 3875-3884.
* Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gō, & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611-639.
* Weiss, M. A., Hua, Q.-X., Lynch, C. S., Frank, B. H., & Shoelson, S. E. (1991) *Biochemistry 30*, 7373-7389.
* Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647-1651.
* Wood, S. P., Blundell, T. L., Wollmer, A., Lazarus, N. R., & Neville, R. W. J. (1975) *Eur. J. Biochem. 55*, 531-542.
* Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Table I ^{1}H chemical shifts of the insulin mutant at pH 6.5 and 307 K measured in ppm relative to dioxane (3.75 ppm)[a]

Residue		H_N	H_α	H_β	Others
A1	Gly				
A2	Ile		3.98	1.23	$H_{\gamma 1}$ 1.22/0.95 $H_{\gamma 2}$ 0.77 H_δ 0.60
A3	Val	8.24	3.66	2.03	$H_{\gamma 1}$ 0.97 $H_{\gamma 2}$ 1.03
A4	Glu	8.34	4.21	2.23,2.11	H_γ 2.55/2.33
A5	Gln	8.18	4.15	2.22	H_γ 2.48/2.56 H_ϵ 7.01/7.52
A6	Cys	8.33	4.96	2.98/3.37	
A7	Cys	8.32	4.91	3.82/3.38	
A8	Thr	8.18	4.20	4.49	$H_{\gamma 2}$ 1.320
A9	Ser	7.43	4.80	4.09,3.95	
A10	Ile	7.85	4.51	1.70	$H_{\gamma 1}$ 1.29/0.72 $H_{\gamma 2}$ 0.78 H_δ 0.66
A11	Cys	9.53	5.08	3.43/3.14	
A12	Ser	8.36	4.77	4.05,4.19	
A13	Leu	9.05	4.08	1.60,1.75	H_γ 1.76 $H_{\delta 1}$ 1.00 $H_{\delta 2}$ 0.91
A14	Tyr	7.81	4.44	3.11/3.13	H_δ 7.23 H_ϵ 6.96
A15	Gln	7.58	4.12	2.37,2.16	H_γ 2.48/2.32 H_ϵ 6.85/7.45
A16	Leu	7.88	4.18	1.51/2.07	H_γ 1.85 H_δ 0.80/0.84
A17	Glu	7.91	4.32	2.11	H_γ 2.38/2.45
A18	Asn	7.44	4.57	2.65,2.71	H_δ 7.24/6.68
A19	Tyr	8.05	4.42	3.55,2.89	H_δ 7.37 H_ϵ 6.87
A20	Cys	7.33	5.29	2.90,3.43	
A21	Asn	8.08	4.58	2.87/2.75	H_δ 6.45/7.49
B1	Glu				
B2	Val		4.19	2.17	H_γ 1.01/1.04
B3	Asn	8.61	4.82	2.87/2.84	H_δ 6.89/7.59
B4	Gln	8.09	4.54	2.00/2.20	H_γ 2.34 H_δ 7.49/6.91
B5	His	8.42	4.71	3.26/3.57	$H_{\delta 2}$ 7.24 $H_{\epsilon 1}$ 8.29
B6	Leu	8.81	4.61	1.00/1.76	H_γ 1.64 $H_{\delta 1}$ 0.85/0.91
B7	Cys	8.61	5.00	3.05,3.27	
B8	Gly	8.13	4.28		
B9	Ser		4.23	4.09/4.03	
B10	Glu	8.21	4.19	2.32,2.16	H_γ 2.45/2.52
B11	Leu	7.11	4.02	1.89,1.24	H_γ 1.32 H_δ 0.76
B12	Val	7.13	3.31	2.18	$H_{\gamma 1}$ 1.04 Hγ2 1.09
B13	Glu	7.91	4.07	2.10	H_γ 2.51/2.33
B14	Ala	7.56	4.19	1.49	
B15	Leu	8.03	3.76	0.28,1.09	H_γ 1.33 $H_{\delta 1}$ 0.35 $H_{\delta 2}$ 0.66
B16	Glu	8.29	4.23	2.32/2.69	H_γ 2.47/2.12
B17	Leu	7.61	4.23	1.89,1.77	$H_{\delta 1}$ 0.99 $H_{\delta 2}$ 0.74
B18	Val	8.38	3.85	2.09	$H_{\gamma 1}$ 1.12 $H_{\gamma 2}$ 0.97
B19	Cys	8.87	4.92	3.01,3.40	
B20	Gly	7.87	4.11/4.09		
B21	Glu	9.19	4.27	2.17	H_γ 2.27/2.45
B22	Arg	8.14	4.23	2.16/2.29	H_γ 1.97 H_δ 3.49/3.43 H_ϵ 7.35
B23	Gly	7.36	3.94/4.23		
B24	Phe	7.61	5.36	3.42,3.02	H_δ 6.90 H_ϵ 7.28 H_ζ 7.38
B25	Phe	8.65	4.96	3.25/3.32	H_δ 7.37 H_ϵ 7.44
B26	Tyr	8.07	4.72	3.06/3.00	H_δ 7.01 H_ϵ 6.74
B27	Glu	7.84	4.77	2.10/1.94	H_γ 2.30
B28	Pro		4.37	2.25	$H_{\gamma 1}$ 2.03 $H_{\gamma 2}$ 2.12 $H_{\delta 1}$ 3.79 $H_{\delta 2}$ 3.63
B29	Lys	7.82	4.24	1.77,1.90	H_γ 1.00/1.46 H_δ 1.74 H_ϵ 3.02

[a] Methylene protons with stereospecific assignment are seperated by a comma, $H_{\beta 1}$ is listed first.

Table II Structural Statistics

Distances and torsion angle restraints

intra residual	297	ϕ	34
short range	193	χ_1	14
long range	51	χ_2	5
cross chain	56		

Energy statistics after simulated annealing (kcal/mol)

Total energy	43.8	±	3.8
Bonds	2.2	±	0.2
Angles	25.5	±	2.2
Repel	7.4	±	1.2
NOE[a]	4.1	±	1.3
Torsion	0.3	±	0.1
Impropers	3.3	±	1.3

van der Waal energy measured with CHARMM[b] potential (kcal/mol)

van der Waal	-70.4	±	11.8

Deviations from Ideal Geometry

Bond (Å)	0.0017	±	0.0001
Angle (deg)	0.35	±	0.02
Improper (deg)	0.26	±	0.04
NOE (Å)	0.012	±	0.002
Torsion (Å)	0.29	±	0.05

Average no. of NOE violations

0.0-0.1 Å	31.8	±	3.1
0.1-0.2 Å	2.0	±	1.0
0.2-0.3 Å	0.16	±	0.37
> 0.3 Å	0		

Atomic rms Values for 25 converged structures vs their geometric average

backbone(all)	1.31	±	0.19
heavy atoms(all)	1.88	±	0.23
backbone(A2-A19)[c]	0.67	±	0.13
backbone(B4-B26)[c]	0.64	±	0.09
backbone(A2-A19,B4-B26)[c]	0.66	±	0.09

[a] Force constants for distance and torsion angle restraints are 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} radian^{-2}

[b] CHARMM potential (Brooks et al., 1983) used for van der Waal energy calculation

[c] Alignment of A2-A19 and B4-B26 backbone

Table III Details of the three dimensional structures of human insulin and mutants determined by NMR[a]

Mutations	Experimental conditions[b]			Structure calculations[c]				No of struct.	Alignment procedure[d,e]	rms[f](Å)		rms[g](Å)		Litterature reference
	pH	T(C°)	Solvent	Conc(mM)	NOE	φ	χ_1	χ_2						
des-(B26-B30)	1.8	25	20% HAc80% H_2O	7	368	?	10	0	15	A2-A19 B9-B19[h]	0.71	1.41	0.60	1.11	Hua et al., 1992a
B24 Ser	1.9	25	20% HAc80% H_2O	1.5	323	27	14	0	10	A2-A19, B4-B19	1.74	2.30	0.70	1.18	Hua et al., 1993
B24 Gly	1.9	25	20% HAc80% H_2O	1.5	290	27	14	0	5[h]	A2-A19, B4-B19	3.01	3.53	0.67	1.13	Hua et al., 1991
des B25	3.0	37	H_2O	4.6	368	10	0	0	20	A2-A20, B2-B28	0.87	1.58	0.65	1.31	Jorgensen et al., 1996
HI	1.9	25	20% HAc80% H_2O	1	422	28	15	0	11	A2-A19, B4-B28	0.97	1.57	0.63	1.13	Hua et al., 1991
B16 His	2.4	24	H_2O	2-3	479	44	23	6	20	A2-A19, B4-B28	0.42	0.85	0.33	0.69	Ludvigsen et al. 1994
(B1,B10,B16,B27) Gln, des B30	6.5	34	H_2O	3	300	34	14	5	25	A2-A19, B4-B26	0.59	1.16	0.52	1.07	present work

(a) Structures of monomeric human insulin and mutants of this determined by NMR and available in the Protein Data Bank, Brookhaven, january 1996.

(b) The experimental conditions includes pH, temperature, solvent, and insulin concentration.

(c) The number of NOEs (intraresidual NOEs excluded) and the number of dihedral angle restraints used in the calculations.

(d) The alignment procedures used in the original publication, and reproduced in the ensemble averages.

(e) The rms deviations listed are calculated as ensemble averages using similar alignments for all molecules, the basis has been backbone atoms (bb) and all heavy atoms (h) respectively. Alignment used A2-A8, A12-A19, B5-B28 (-B23 for DPI), (-B27 for (B1,B10,B16,B27) Glu, des B30).

(f) Alignment used A2-A8, A12-A19, B5-B19

(g) Originally only the B chain helix was aligned. In order to make this average structure comparable to the others alignment of the A chain has been added.

(h) 11 structures was published, 9 was deposited.

Table IV Comparison of solution structures of human insulin and mutants[a].

rms (Å)[(a)]	DPI	B24 Ser	B24 Gly	des B25	HI	B16 His	(B1,B10,B16,B27) Glu des B30
DPI	-	1.07	1.38	1.26	0.33	1.12	1.15
B24 Ser	0.97	-	1.54	1.14	1.10	1.54	1.35
B24 Gly	1.41	1.54	-	1.36	1.37	1.14	1.18
des B25	1.23	1.14	1.36	-	1.27	1.49	1.35
HI	0.33	1.10	1.37	1.27	-	1.09	1.08
B16 His	1.20	1.54	1.14	1.49	1.09	-	0.84
(B1,B10,B16,B27) Glu des B30	1.16	1.35	1.18	1.35	1.08	0.84	-

[a] The root mean square deviations of the backbone were determined aligning the regions A2-A19, B4-B19 above the diagonal and the helix regions A2-A8, A13-A19, B9-B19 below the diagonal. The structures are ensemble averages aligned as listed in Table III.

Figure Legends

Figure 1 1D NMR spectra showing the aromatic and amide proton resonances of native and mutant insulins at 1 mM protein concentration in 10/90 D_2O/H_2O, 307 K. Panels (A): human insulin, pH 7.5; (B): B27 Thr → Glu , pH 7.5; (C): (B16, B27) Glu, pH 7.5; (D): (B1, B27) Glu, pH 7.5; (E): B10 Glu, des-B30, pH 7.5; (F)-(H): (B1, B10, B16, B27) Glu, des-B30 insulin at pH 7.5, 7.0 and 6.5, respectively.

Figure 2 Near-UV CD spectra as a function of protein concentration. (A): zinc-free human insulin in 10 mM Tris/ClO_4^-, pH 8.0, and (·····) 5 μM, (– – –) 20 μM, (–·–·–) 200 μM, and (——) 2 mM protein. (B) and (C): (B1, B10, B16, B27) Glu, des-B30 mutant insulin in 25 mM Tris/ClO_4^-, pH 7.5 (B) and 25 mM phosphate, pH 6.5 (C), and (·····) 50 μM, (–·–·–) 600 μM and (——) 2.6 mM protein.

Figure 3 Sections from the DQF COSY spectrum (top) and NOESY spectrum (mixing-time 180 ms) (bottom). The fingerprint region of the DQF COSY spectrum shows the major part of the correlations between α-protons and amide protons that was identified. The correlations between amide protons indicative of helical areas in the structure are annotated in the bottom panel, and lines are drawn to follow the amide protons throughout the helices.

Figure 4 Difference between chemical shifts of the α-protons and the random coil values (Wishart et al., 1992). Stretches where the secondary chemical shifts indicate an α-helix are marked.

Figure 5 Summary of the data used for sequential assignment and determination of secondary structural elements. The thickness of the bars indicates the intensity of the NOEs as measured in a NOESY spectrum with a mixing-time of 150 ms. The coupling constants and the χ_1 dihedral angles were measured as desribed in Materials and Methods. The filled circles indicate that the amide proton signal is observable in a 1D spectrum at least 12 minutes after dissolution of the sample in D_2O.

Figure 6 Panel A shows the distribution of NOEs by residue, intraresidual NOEs are marked with filled boxes and interresidual NOEs with vertical lines. In the matrix in panel B filled circles indicate the presence of at least one NOE between the connected residues. Multiple occurences of NOEs on both sides of the diagonal in a spectrum or in different spectra were filtered out before the conversion to distance restraints. The presence of helical regions, as expected from the patterns of NOEs, is indicated.

Figure 7 The backbone of 25 converged structures of the insulin mutant superimposed together with their geometric average. The average structure is represented by a ribbon, the A-chain is shown in red, the B-chain in blue. The structures are aligned in the regions B4-B26, A2-A19.

Figure 8 The geometric average structure of the insulin mutant colored according to atomic rms. The average structure was calculated based on alignment of residues B4-B26 and A2-A19.

Figure 9 Ensemble averages of three dimensional structures of native (HI) and mutant insulins determined by NMR. The structures are color coded according to the atomic root mean square deviation of the backbone. For further details, see Table III.

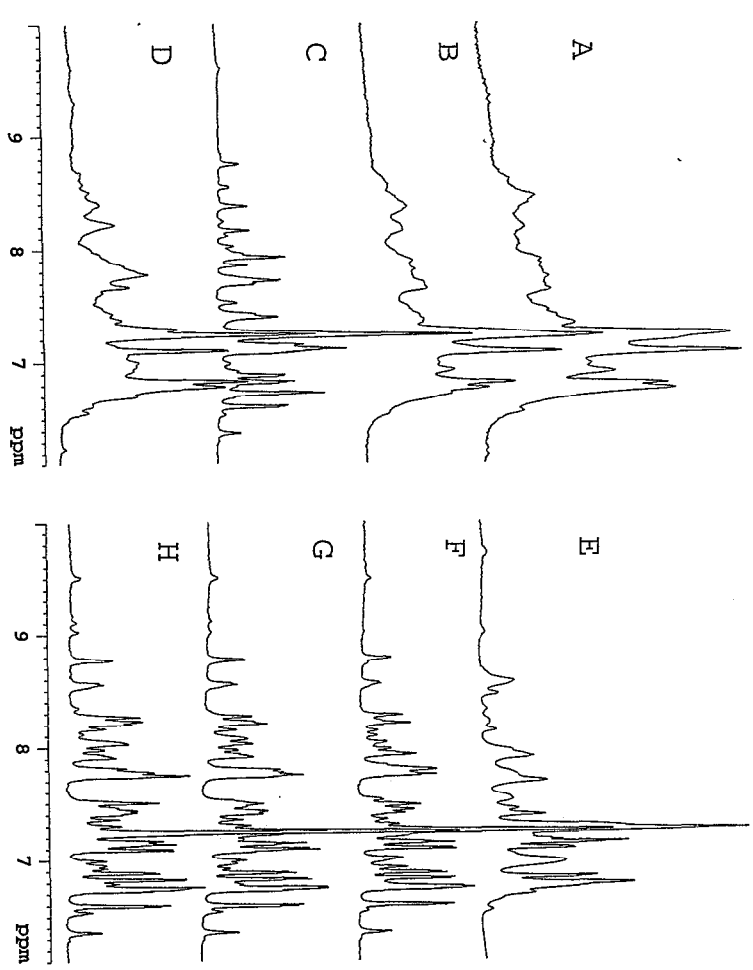
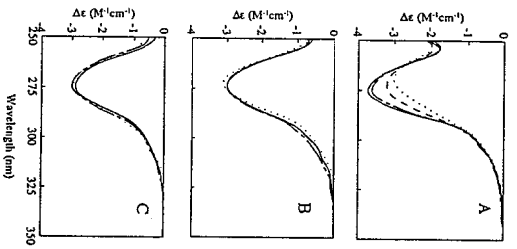
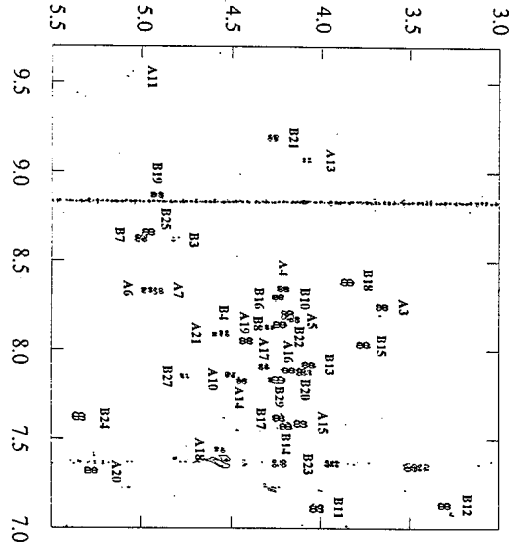


Figure 2



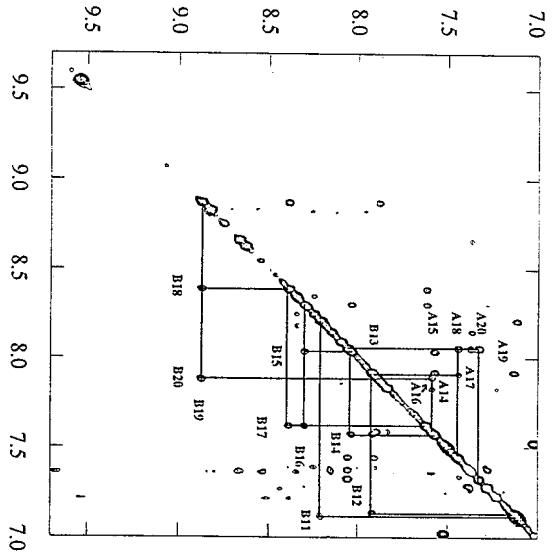
37

Figure 3 a



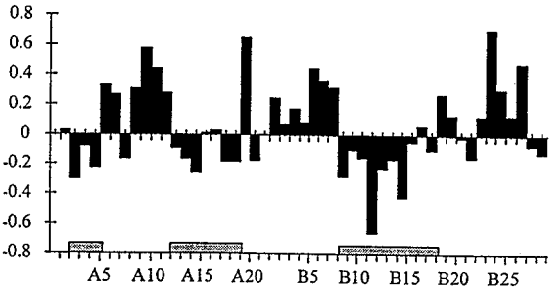
38

Figure 3 b



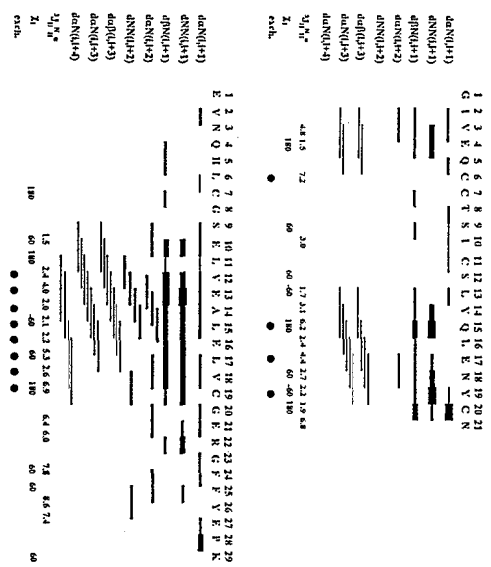
39

Figure 4



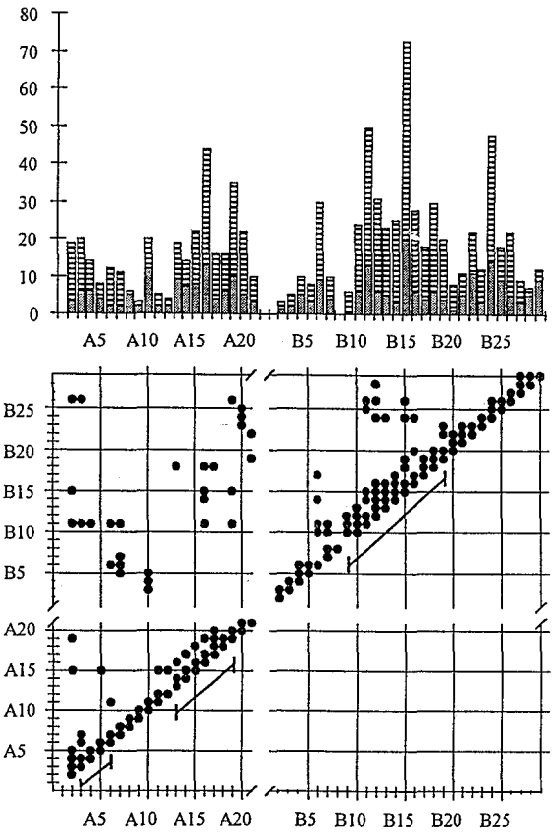
40

Figure 5



41

Figure 6



42

Figure 7

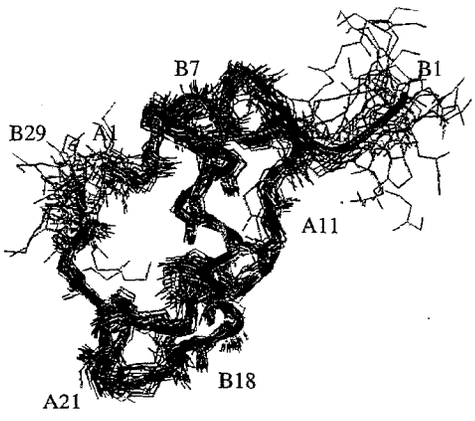


Figure 8

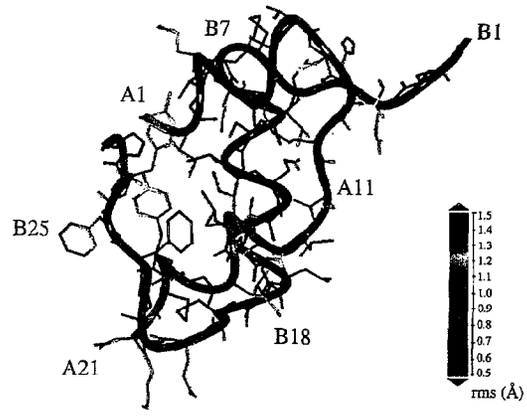
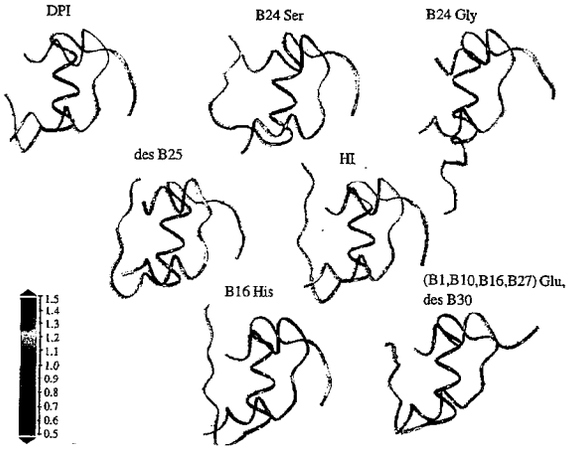


Figure 9



resonances become sharper as expected due to the slowed exchange rate.

Figure 2.4 shows near-UV CD spectra as a function of protein concentration for native insulin at pH 8.0 and the (B1, B10, B16, B27) Glu, des-B30 mutant at pH 6.5 and 7.5. For native insulin, the progressive increase in intensity of the negative signal around 274 nm indicates the expected increase in association with increasing concentration. In contrast, the near-UV CD spectrum of the mutant is independent of the protein concentration in the 60 μM to 2.6 mM range at both pH 6.5 and 7.5. Although the mutant is missing one tyrosine reporter group (B16), Tyr B26 is expected to be strongly affected by monomer/monomer interactions. Hence, these results provide further evidence that the (B1, B10, B16, B27) Glu, des-B30 mutant is monomeric under conditions necessary for NMR structural studies at neutral pH, and this species was selected for detailed characterization.

2.4.2 Assignment of Spin Systems

NMR spectra were assigned using the standard procedures outlined by Wüthrich (1986). The fingerprint region of the DQF-COSY spectrum is shown in the upper panel of Figure 2.5. The chemical shifts are dispersed and well-resolved as expected for a structured globular protein, and 37 of the 47 possible H^NH^α cross peaks are annotated in the plot. Among the remaining H^NH^α cross peaks, six were assigned using TOCSY spectra. These are the resonances that are either close to the water line or unusually broad, i.e., A8, A9, A12, B5, B6, and B26. The linewidth of the CysA11 amide proton is unusually broad, but identification was possible in DQF-COSY spectra. Amide protons from residues A2, B6 and B9 were not identified. Finally, for the side-chains of Glu and Gln residues, considerable overlap of cross peaks in the H^β-H^γ area was resolved using TOCSY spectra occasionally supported by NOESY.

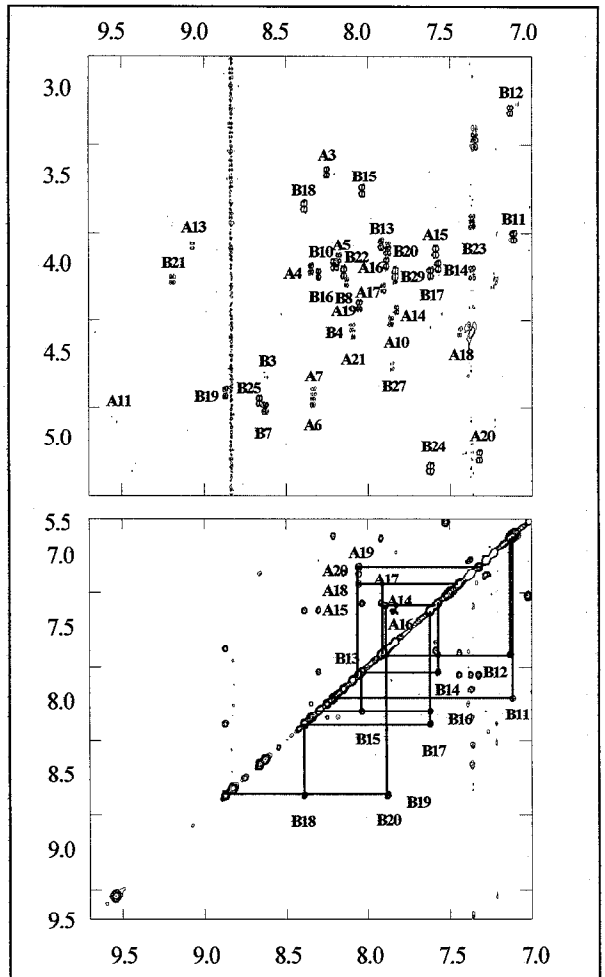


Figure 2.5 Sections from the DQF COSY spectrum (top) and NOESY spectrum (mixing-time 180 ms) (bottom). The fingerprint region of the DQF COSY spectrum shows the major part of the correlations between α-protons and amide protons that was identified. The correlations between amide protons indicative of helical areas in the structure are annotated in the bottom panel, and lines are drawn to follow the amide protons throughout the helices.

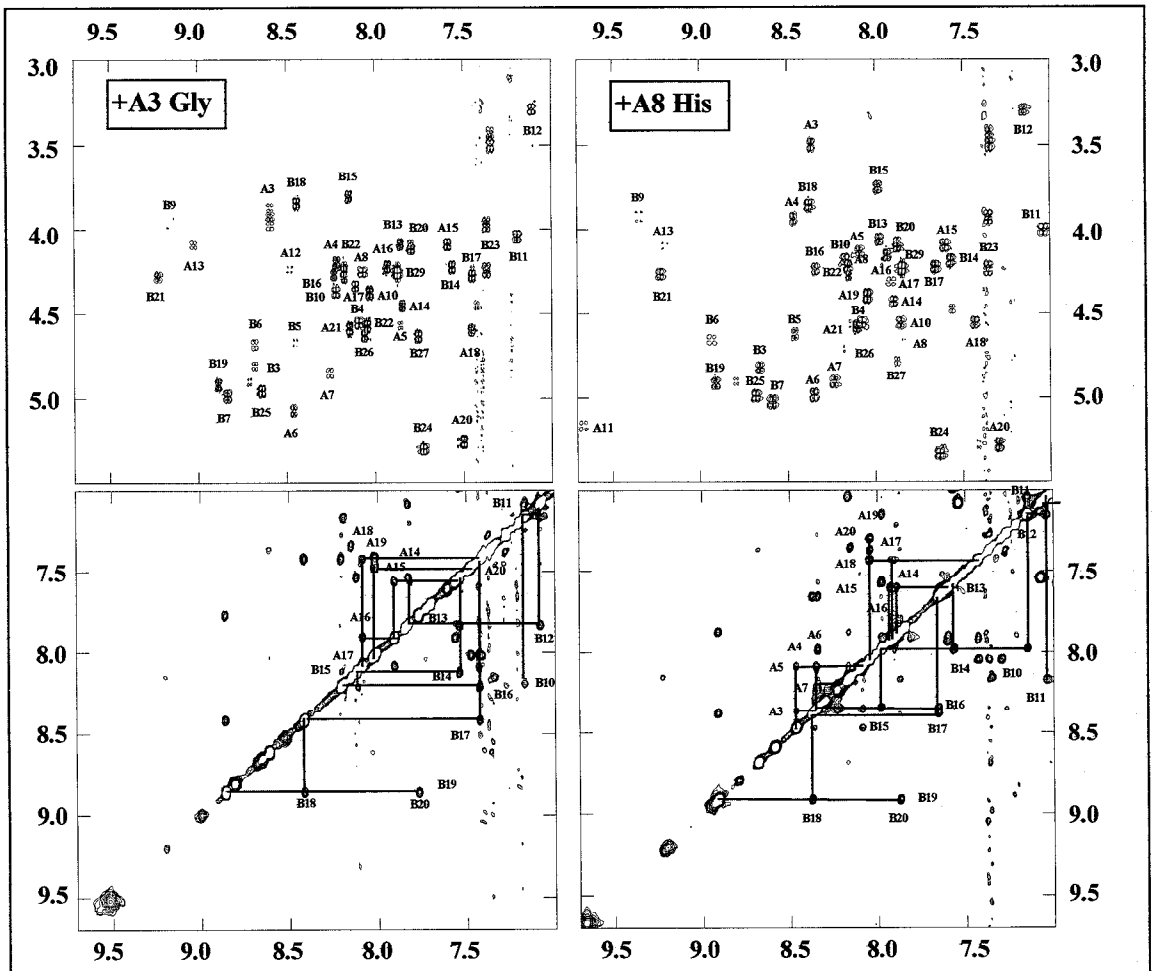


Figure 4.2 Sections from DQF-COSY spectra (top) and NOESY spectra (mixingtime 120 ms) (bottom). In the fingerprint region of the DQF COSY spectrum the major part of the identified correlations between α-protons and amide protons is annotated. The correlations between amide protons of neighboring residues, indicative of helical areas in the structure, are annotated in the bottom panel and lines are drawn to follow the amide protons throughout the helices.

Part of the secondary structural elements of the mutants, the helix regions, are indicated in the excerpts from the NOESY spectra in the bottom panel of Figure 4.2. Intense H^NH^N connectivities are assigned for both mutants in the B-chain helix region, B10-B20, and in the A(II)-helix region, A14-A20. In contrast to the $^+$A3 Gly mutant (and the basis, Figure 2.4), the spectrum of the $^+$A8 His mutant shows intense H^NH^N connectivities in the A(I)-helix, A3-A7. This is an evidence of the helix stabilization caused by the A8 Thr → His mutation in the helix C cap. Note that the basis mutant spectra do show H^NH^N connectivities in the A3-A5 region, these peaks are less intense and visible, but not annotated, in Figure 2.5.

Further indications of secondary structure come from comparison of the α-proton chemical shift and the random coil values, the secondary chemical shifts of the α-protons. Data for all three mutants is shown in the left-hand column of Figure 4.3. The helix elements predicted by the extents of