

NMR studies of the aggregation of glucagon-like peptide-1: formation of a symmetric helical dimer

Xiaoqing Chang^{a,1}, Danielle Keller^{a,2}, Seán I. O'Donoghue^{b,3}, Jens J. Led^{a,*}

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

^bEuropean Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany

Received 12 February 2002; accepted 14 February 2002

First published online 27 February 2002

Edited by Thomas L. James

Abstract Nuclear magnetic resonance (NMR) spectroscopy reveals that higher-order aggregates of glucagon-like peptide-1 (7–36)-amide (GLP-1) in pure water at pH 2.5 are disrupted by 35% 2,2,2-trifluoroethanol (TFE), and form a stable and highly symmetric helical self-aggregate. NMR spectra show that the helical structure is identical to that formed by monomeric GLP-1 under the same experimental conditions [Chang et al., *Magn. Reson. Chem.* 37 (2001) 477–483; Protein Data Bank at RCSB code: 1D0R], while amide proton exchange rates reveal a dramatic increase of the stability of the helices of the self-aggregate. Pulsed-field gradient NMR diffusion experiments show that the TFE-induced helical self-aggregate is a dimer. The experimental data and model calculations indicate that the dimer is a parallel coiled coil, with a few hydrophobic residues on the surface that may cause aggregation in pure water. The results suggest that the coiled coil dimer is an intermediate state towards the formation of higher aggregates, e.g. fibrils. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Glucagon-like peptide-1(7–36)-amide; Nuclear magnetic resonance; Aggregation; Intermediate; Modeling

1. Introduction

Glucagon-like peptide-1(7–36)-amide (GLP-1) is a 30 amino acid peptide with the sequence: HAEGTFTSDVSSYLEG-QAAKEFIAWLVKGR-NH₂. It is an important glucocretin hormone that can potentiate glucose-induced insulin secretion, stimulate insulin biosynthesis and inhibit glucagon secretion [1,2]. Consequently, it is a potential drug for the treatment of non-insulin-dependent diabetes mellitus (NIDDM or type 2 diabetes [3]). However, GLP-1 suffers

from both metabolic and physical instability. Recently the metabolic stability and the biological activity of a series of GLP-1 analogues were investigated [4–6]. As for the physical instability, it has been found that the conformation, aggregation, and solubility of GLP-1 depend on the purification procedure and the in-process storage and handling [7,8]. Still, the α -helix seems to be the predominant structural motif of GLP-1 in solution. It was found that GLP-1 can form oligomers with a high helical content [7–9], and that it is mainly helical in membrane-like environments (dodecylphosphocholine micelle) [10]. More recently it was found [11] that monomeric GLP-1 is a random coil in pure water but forms a helical structure in aqueous 2,2,2-trifluoroethanol (TFE) solution (Protein Data Bank at RCSB (PDB) code: 1D0R), similar to the structure observed in the membrane-like environments. However, GLP-1 can also form less soluble β -sheet aggregates [7].

To provide further insight into the self-association behavior of GLP-1 we have studied the conformation of aggregated GLP-1 in water/TFE mixtures using nuclear magnetic resonance (NMR) spectroscopy. TFE is known to stabilize helical conformations in proteins and peptides while disrupting the specific tertiary interactions of native proteins. These effects are ascribed to the ability of TFE to enhance internal hydrogen bonding in polypeptides and to lessen hydrophobic interactions between residues distant in the amino acid sequence [12]. A study of the structure of GLP-1 in water/TFE mixtures seems, therefore, highly interesting in order to get more insight into the complex aggregation propensity of GLP-1 in solution.

2. Materials and methods

2.1. Sample preparation

Samples of aggregated and monomeric recombinant GLP-1 were kindly provided by the pharmaceutical company Novo Nordisk A/S. The lyophilized peptide was dissolved in H₂O (with 10% D₂O) or in 99.96% D₂O. In all samples the concentration of monomeric GLP-1 was approximately 1.4 mM and the pH was 2.5 (meter reading). The concentration of TFE in the water/TFE NMR samples was 35%. Perdeuterated TFE (TFE-d₃) was used in all experiments.

2.2. NMR spectroscopy

The NMR spectra were recorded at 300 K on Varian Inova Unity 500, 750 and 800 spectrometers as described previously [11]. Qualitative discrimination between slowly and fast exchanging amide protons was achieved by recording a series of nuclear Overhauser effect spectroscopy (NOESY) spectra of GLP-1 in 99.96% D₂O with 35% TFE at regular times after the dissolution, and defining the amide protons still giving rise to cross peaks as slowly exchanging. A series of pulsed-field gradient (PFG) experiments with water-suppressed longitudinal encoding decoding (water-sLED) pulse [13,14] was performed to de-

*Corresponding author. Fax: (45)-3535 0609.

E-mail address: led@kiku.dk (J.J. Led).

¹ Present address: Computational Chemistry, GLYCODESIGN Inc., 480 University Ave., Suite 400, Toronto, ON, Canada M5G 1V2.

² Present address: Department of Physics, MEMPHYS, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark.

³ Present address: LION Bioscience AG, Waldhoferstr. 98, Heidelberg 69123, Germany.

Abbreviations: CSI, chemical shift index; GLP-1, glucagon-like peptide-1(7–36)-amide; NOESY, nuclear Overhauser effect spectroscopy; PDB, Protein Data Bank at RCSB; PFG, pulsed-field gradient; TFE, 2,2,2-trifluoroethanol; water-sLED, water-suppressed longitudinal encoding decoding

termine the self-diffusion coefficients and the relative molecular size of monomeric and aggregated GLP-1 in 35% TFE.

2.3. Generation of the GLP-1 coiled coil dimer

An initial model of the backbone coordinates of the GLP-1 dimer was obtained using the program ThreadCoil [15]. ThreadCoil takes as an input an amino acid sequence that is proposed to fold into a coiled coil, and returns a three-dimensional structure (α -carbons only). The most likely structure is found by a grid search over the three coiled coil parameters (radius, pitch, and phase). The likelihood of each structure is evaluated using a pairwise residue potential derived from the PDB database.

The coordinates of the remaining side-chain heavy atoms were introduced by MODELLER4 [16], where the side-chains were energy minimized from initially random positions. The general form of the energy functions and the optimization in MODELLER4 are similar to CHARMM [17]. The hydrogen atoms were then included by the HBUILD feature in X-PLOR, and the full-atom structures were refined against the NMR-derived distance restraints (including the intramonomeric NOEs and two distance restraints for each of the hydrogen bonds observed in the aggregated GLP-1 spectrum, except the one involving the amide proton of residue A2, vide infra) using simulated annealing. The dimer symmetry was imposed using a non-crystallographic symmetry term. The simulated annealing was based on the method described by Nilges and Brünger for automated modeling of coiled coils [18]. During the refinement, the positions of the α -carbons were restrained using a harmonic restraining potential. A calculation of 10 slightly different model structures was carried out starting with different side-chain coordinates, and the model with the lowest total energy was selected as the representative structure.

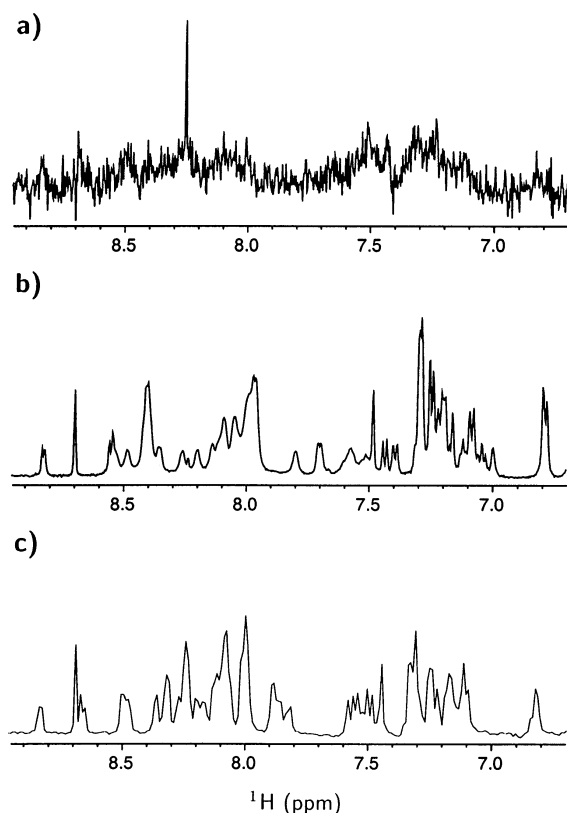


Fig. 1. One-dimensional ^1H NMR spectra of GLP-1. a: Aggregated GLP-1 in pure water. b: The same aggregated form of GLP-1 in water with 35% (v/v) TFE. c: Monomeric, non-aggregated GLP-1 in water with 35% (v/v) TFE.

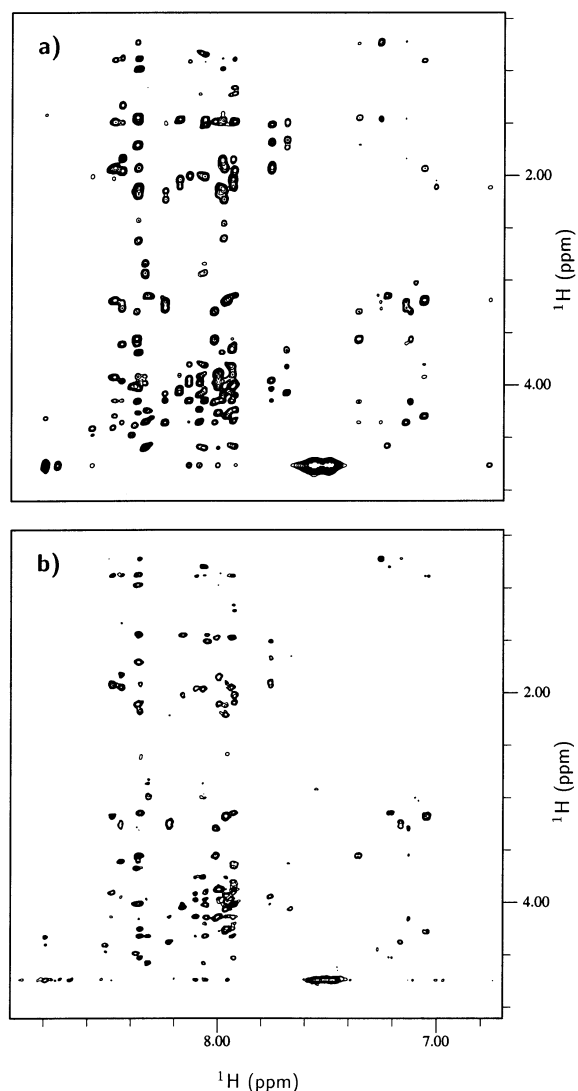


Fig. 2. The amide and aromatic proton region of the NOESY spectra of (a) monomeric and (b) aggregated GLP-1, both in water with 35% (v/v) TFE. The spectra were recorded at (a) 800 MHz and (b) 500 MHz. The same correlations are present in the two spectra, and the chemical shift values are identical.

3. Results and discussion

3.1. Formation of a symmetric helical GLP-1 aggregate

Fig. 1 shows the one-dimensional ^1H NMR spectra of aggregated and monomeric GLP-1 dissolved in pure water and in water with 35% TFE, respectively. The excessive broadening of the resonances of the aggregated form in pure water (Fig. 1a) clearly reveals an extensive aggregation.

In sharp contrast to this, the one-dimensional ^1H spectra of aggregated and monomeric GLP-1 in water with 35% TFE (Fig. 1b,c) are practically identical. This identity is confirmed by the two-dimensional NOESY spectra in Fig. 2, which show that the chemical shifts and the pattern of NOE signals of the two forms are almost identical. Only two extra $\alpha\beta(i,i+3)$ NOEs in the N-terminal end of the peptide appear in the spectrum of the aggregated form (Fig. 3b). Moreover, all the observed signals can be assigned to the same monomeric

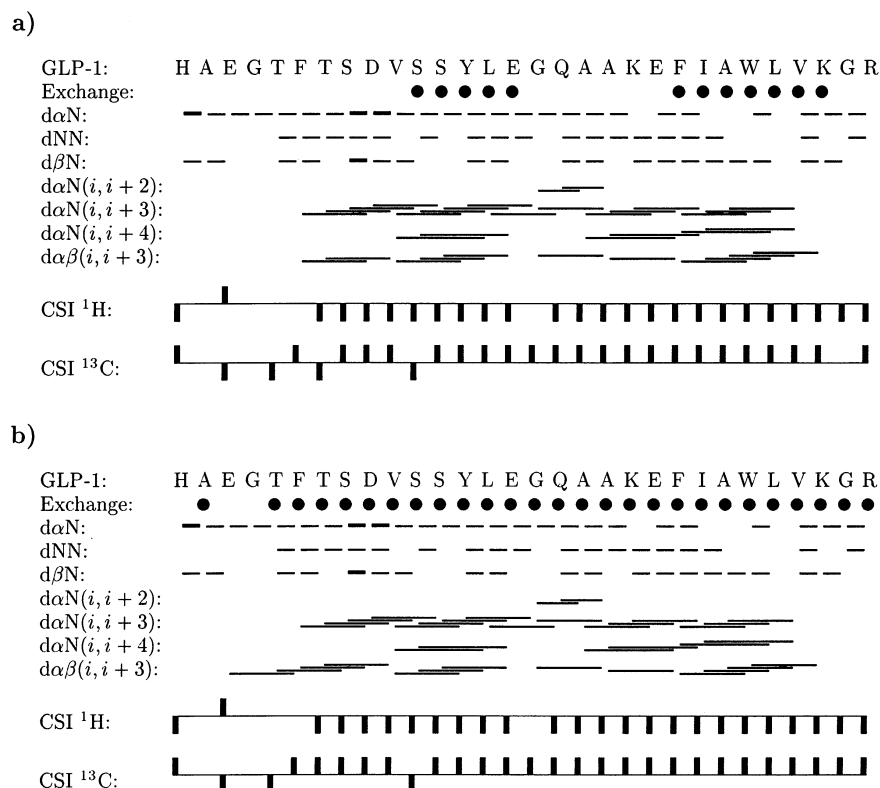


Fig. 3. Summary of sequential and medium-range NOE connectivities, and schematic representation of the $^1\text{H}_\alpha$ chemical shift index (CSI) of (a) monomeric GLP-1 and (b) aggregated GLP-1. In both cases the sample (1.4 mM, pH 2.5, 300 K) was dissolved in H_2O with 35% TFE. The observed NOE connectivities are indicated by bars connecting the two involved residues, and the intensities of the NOEs are indicated by the thickness of the bars. The CSIs were calculated according to Wishart et al. [36]. Slowly exchanging backbone amide protons are indicated by filled circles (●).

structure, i.e. only one set of resonances is present and no long-range NOEs corresponding to a tertiary fold are observed. These results show, unambiguously, that the higher aggregates of GLP-1 have been disrupted by TFE. Furthermore they show that the local geometry of the individual residues of the resulting structure is identical to that in monomeric, helical GLP-1; i.e. the peptide has the same helical structure as the monomeric form under the same experimental conditions, except for a small extension of the α -helix towards the N-terminal end.

However, the amide proton exchanges reveal that the helix is considerably more stable in the aggregated form than in the monomeric form. Thus, the number of slowly exchanging amide protons is increased dramatically in the aggregated form (27 versus 12 in the monomeric form) as shown in Fig. 3b. Furthermore, the intensities of the amide proton signals were unchanged several days after dissolution in D_2O . In fact, most of the signals still remained after months, i.e. the exchange rates of the amide protons are about two orders of magnitude slower in the helical structure obtained from the aggregated form than in the monomeric helix. Consequently, the helix has been stabilized dramatically through a strengthening of its $i, i+4$ hydrogen bonds, while at the same time it has been extended all the way to the N-terminus of the peptide (Fig. 3b). Taken together, the spectra and the amide proton exchange rates of the structure derived from the aggregated form clearly indicate the formation of a highly sym-

metric, highly stable and soluble self-associate consisting of GLP-1 monomers with basically the same extended helical structure as the monomeric GLP-1 described previously [11].

Finally it was found that the helical self-associate can be formed also from monomeric GLP-1 by seeding it with the highly aggregated form. Thus addition of traces of the latter form to a solution of monomeric GLP-1 in water with 35% TFE resulted in the same dramatic decrease of the amide proton exchange rates and the extension of the helix towards the N-terminus of the peptide as described above. This result further emphasizes the stability of the helical self-associate and suggests a role as an intermediate toward the formation of higher aggregates of GLP-1.

3.2. The size of the helical GLP-1 aggregate

The size of the symmetric GLP-1 self-associate in TFE was evaluated by comparing its self-diffusion coefficient with that of monomeric GLP-1 using the PFG NMR self-diffusion experiment, as described in Section 2. The identical experimental conditions that were applied in the two cases (1.4 mM GLP-1 in D_2O and 35% TFE, pH 2.5) and the identical nature of the two macromolecules to be compared ensure a reliable comparison.

The self-diffusion coefficient is sensitive both to the size and to the shape of the molecule, and the oligomerization of several well-characterized proteins has been studied using the approach [19,20] applied here. For the simplest dimerization

coiled coil dimer. As described below the parallel orientation is supported by a possible hydrogen bond between the carboxylic acid side-chains of D9 in the two monomers.

3.3. Modeling the coiled coil dimer

Coiled coil domains occur frequently in native protein (5–10% of all proteins). This supersecondary structure motif can consist of two to five α -helices wound around each other, allowing favorable side-chain meshing without much α -helix distortion. The coiled coil motifs found in oligomerized proteins share a characteristic heptad repeat (*a-b-c-d-e-f-g*) with hydrophobic interactions between *a* and *d* positions, and electrostatic interactions between *e* and *g* positions.

The GLP-1 sequence fits this pattern approximately (Fig. 5a). Only the polar residues D9 and G16 are not usually found in position *a* in leucine zippers or fibrous proteins and will, in general, destabilize the structure. On the other hand, buried polar residues appear to be important for the structure of coiled coils by favoring a specific orientation of the strands through interstrand hydrogen bonding [25–27]. In the case of GLP-1 the D9 residue is uncharged at the experimental conditions (pH 2.5) since its side-chain carboxylic acid has a random coil *pK* value of 4.0 in the absence of any electrostatic interactions [28]. Therefore, the parallel packing of the two helices in the coiled coil GLP-1 dimer may be imparted by an interchain hydrogen bond between the side-chains of the two opposite D9 residues (Fig. 5b).

The coiled coil fold has a relatively simple geometry, with only three unknown parameters – the pitch of the superhelix, the radial separation of the helices, and a phase angle, determining which residues are on the interface. Due to this simplicity, various methods have been developed that can predict coiled coil backbone structures with an accuracy of about 1 Å. Here, ThreadCoil [15] was used with the GLP-1 sequence to perform a wide grid search of the three coiled coil parameters, including the right-handed coiled coil possibility. The structure with the lowest energy score was a left-handed coiled coil, with a pitch, radius, and phase of 136 Å, 3.70 Å, and 40°, respectively. These values are quite similar to those for the leucine zippers. Thus, for example, the Jun homodimer has a pitch, radius, and phase of 147 Å, 4.70 Å, and 34°, respectively [29]. The most significant difference occurs in the supercoil radius, suggesting that the two helices are 2 Å closer in GLP-1 than in the leucine zippers. This may be explained by the relatively small and flexible side-chains that characterize most of the *a*- and *d*-position residues, which allow the α -helices in coiled coils to pack more closely [30,31]. This suggestion is consistent with the very slowly exchanging amide protons observed for the GLP-1 dimer.

The ThreadCoil energy score of the calculated coiled coil backbone structure of GLP-1 (−0.3 kT) was significantly higher than that of leucine zippers (between −1 and −2.5 kT) with the largest contribution stemming from two hydrophobic residues (a leucine and a valine) that are solvent exposed. This indicates that not all of the hydrophobic residues of GLP-1 can be buried in a coiled coil structure, and hence that the structure is not likely to be stable in pure water but will lead to further aggregation. This is consistent with the experimental result obtained here that TFE is necessary for the stabilization of the helical structure.

Starting from this initial backbone model, an all-atom model was calculated as described in Section 2. In accordance with

the experimental observation of the dimerized GLP-1 in solution. The resulting dimer (Fig. 5c) consists of two parallel monomers wound around each other in a left-handed way. The symmetric, parallel arrangement of the two extended helices is in agreement with the observation of only one set of resonances and the lack of long-range NOEs in the molecule.

For the 15 amide protons, that are slowly exchanging only in the coiled coil dimer, the slow exchange can be attributed to their involvement in the helical hydrogen bonding, except for residue A2. The reason for the slow exchange of this residue is not immediately apparent, although side-chain hydrogen bonding involving the amide proton of A2, and the location of the A2 residue on the dimer interface may limit the accessibility of solvent water to its amide proton.

As mentioned above, most of the amide proton signals were still present after a month. However, the intensities of nine of these signals decreased significantly faster than those from the remaining amide protons. This clearly indicates a relatively higher solvent exposure of the nine residues. Seven of these residues (i.e. T7, S8, V10, S11, A18, E21, K28) are located at positions *b*, *c*, *f* and *g* in the helical wheel, i.e. outside the coiled coil dimer interface. The faster exchange rates of their amide protons thus support the conclusion that the GLP-1 dimer is a coiled coil. Similar observations of a relatively faster amide proton exchanges rates for residues in positions *b*, *c*, and *f* have been observed for other coiled coil proteins [25,32,33]. The remaining two residues for which faster amide proton exchanges were observed, i.e. G29 and R30, are both at the C-terminal end and are located at positions *g* and *a*, respectively.

Coiled coils may aggregate to higher-order structures [34]. In the case of GLP-1 further aggregation could take place by intermolecular hydrophobic interactions between the hydrophobic residues located at the *b*, *c*, and *f* positions. However, here the relatively large amount of TFE protects it against further aggregation. In accordance with this, Sykes and co-workers [35] reported that addition of 15% TFE decreased the dimerization and self-aggregation ability of muscle regulatory protein troponin C, without any significant effect on the secondary or tertiary structures. In a similar manner the high content of TFE does not destabilize the supersecondary coiled coil motifs, as observed here for GLP-1 in 35% TFE. However, TFE prevents further aggregation of the peptide by lessening the hydrophobic interaction of solvent-exposed non-polar side-chains [12].

4. Conclusion

The results here show that TFE stabilizes a highly symmetric, helical GLP-1 dimer, when aggregated GLP-1 is dissolved in TFE/water at pH 2.5, while disrupting the higher aggregates. The results indicate that the helical dimer is a parallel coiled coil. The formation of a helix and the absence of any tertiary structures are consistent with the general helix promoting effect of TFE and its destabilizing effect on the tertiary architecture of native proteins. The formation of a coiled coil dimer was indicated by (1) an increase of the number of slowly exchanging amide protons to more than twice the number observed for the single-strand α -helix, and a decrease of the amide proton exchange rates by about two orders of magnitude, (2) insignificant changes in chemical shifts and patterns of signals in the NMR spectra as compared with mono-

meric helical GLP-1, (3) the size of the dimer as determined from its self-diffusion coefficient, and (4) molecular modeling. Although the coiled coil dimer has a propensity for further aggregation due to the hydrophobic side-chains that are located on its surface, the aggregation is prevented by the presence of TFE molecules. Overall the results suggest that the dimeric coiled coil is an intermediate towards the formation of higher-order GLP-1 aggregates, e.g. fibrils, in water.

Acknowledgements: This work was financially supported by the Danish Technical Research Council (J. Nos. 16-5028-1 and 9601137), the Danish Natural Science Research Council (J. 9502759, 9601648 and 9801801), Direktør Ib Henriksens Fond, Carlsbergfondet and Novo Nordisk Fonden. The 750 and 800 MHz NOESY spectra were obtained at The Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules. We are grateful to Dr. Søren M. Kristensen for helpful discussions, and Dr. Jens Duus and Ms. Else Philipp for technical assistance.

References

- [1] Drucker, D.J., Philippe, J., Mosjov, S., Chick, W.L. and Habener, J.F. (1987) *Proc. Natl. Acad. Sci. USA* 84, 3434–3438.
- [2] Drucker, D.J. (1998) *Diabetes* 47, 159–169.
- [3] Adelhorst, K., Hedegaard, B.B., Knudsen, L.B. and Kirk, O. (1994) *J. Biol. Chem.* 269, 6275–6278.
- [4] Deacon, C.F., Knudsen, L.B., Madsen, K., Wiberg, F.C., Jacobsen, O. and Holst, J.J. (1998) *Diabetologia* 41, 271–278.
- [5] Knudsen, L.B., Nielsen, P.F., Huusfeldt, P.O., Johansen, N.L., Madsen, K., Pedersen, F.Z., Thøgersen, H., Wilken, M. and Agerso, H. (2000) *J. Med. Chem.* 43, 1664–1669.
- [6] Xiao, Q., Giguere, J., Parisien, M., Jeng, W., St-Pierre, S.A., Brubaker, P.L. and Wheeler, M.B. (2001) *Biochemistry* 40, 2860–2869.
- [7] Kim, Y., Rose, C.A., Liu, Y., Ozaki, Y., Datta, G. and Tu, A.T. (1994) *J. Pharmacol. Sci.* 83, 1175–1180.
- [8] Senderoff, R.I., Kontor, K.M., Kreilgaard, L., Chang, J.J., Patel, S., Krakover, J., Heffernan, J.K., Snell, L.B. and Rosenberg, G.B. (1998) *J. Pharmacol. Sci.* 87, 183–189.
- [9] Lambert, W.J., Grucza, R.A., Stamper, G.F. and Chrnyk, B. (1994) *Pharm. Res.* 11, S-83.
- [10] Thornton, K. and Gorenstein, D.G. (1994) *Biochemistry* 33, 3532–3539.
- [11] Chang, X., Keller, D., Bjørn, S. and Led, J.J. (2001) *Magn. Reson. Chem.* 39, 477–483.
- [12] Buck, M. (1998) *Q. Rev. Biophys.* 31, 297–355.
- [13] Altieri, A.S., Hinton, D.P. and Byrd, R.A. (1995) *J. Am. Chem. Soc.* 117, 7566–7567.
- [14] Gibbs, S.J. and Johnson Jr., C.S. (1991) *J. Magn. Reson.* 93, 395–402.
- [15] O'Donoghue, S.I. and Nilges, M. (1997) *Fold. Des.* 2, S47–S52.
- [16] Sali, A. and Blundell, T.L. (1993) *J. Mol. Biol.* 234, 779–815.
- [17] Brooks, B., Brucoleri, R., Olafson, B., States, D., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.* 4, 187–217.
- [18] Nilges, M. and Brünger, A.T. (1991) *Protein Eng.* 4, 649–659.
- [19] Dingley, A.J., Mackay, J.P., Chapman, B.E., Morris, M.B., Kuchel, P.W., Hambly, B.D. and King, G.F. (1995) *J. Biomol. NMR* 6, 321–328.
- [20] Ilyina, E., Roongta, V., Pan, H., Woodward, C. and Mayo, K. (1997) *Biochemistry* 36, 3383–3388.
- [21] Eimer, W. and Pecora, R. (1991) *J. Chem. Phys.* 94, 2324–2329.
- [22] Tirado, M.M. and Garcia de la Torre, J. (1979) *J. Chem. Phys.* 71, 2581–2587.
- [23] Brems, D.N., Plaisted, S.M., Kauffman, E.W. and Havel, H.A. (1986) *Biochemistry* 25, 6539–6543.
- [24] Roongta, V., Powers, R., Jones, C., Beakage, M.J., Shields, J.M. and Gorenstein, D.G. (1989) *Biochemistry* 28, 1048–1054.
- [25] Junius, F.K., Mackay, J.P., Bubbs, W.A., Jensen, S.A., Weiss, A.S. and King, G.K. (1995) *Biochemistry* 34, 6164–6174.
- [26] Lumb, K.J. and Kim, P.S. (1995) *Biochemistry* 34, 8642–8648.
- [27] Oakley, M.G. and Kim, P.S. (1998) *Biochemistry* 37, 12603–12610.
- [28] States, D.J. and Karplus, M. (1987) *J. Mol. Biol.* 197, 122–130.
- [29] O'Donoghue, S.I., King, G.F. and Nilges, M. (1996) *J. Biomol. NMR* 8, 193–206.
- [30] Gernert, K.M., Surles, M.C., Labeau, T.H., Richardson, J.S. and Richardson, D.C. (1995) *Protein Sci.* 4, 2252–2260.
- [31] Nooren, I.M.A., Kaptein, R., Sauer, R.T. and Boelens, R. (1999) *Nature Struct. Biol.* 6, 755–759.
- [32] Goodman, E.M. and Kim, P.S. (1991) *Biochemistry* 30, 1615–1620.
- [33] Harbury, P.B., Plecs, J.J., Tidor, B., Alber, T. and Kim, P.S. (1998) *Science* 282, 1462–1467.
- [34] Creighton, T.E. (1993) *Proteins: Structures and Molecular Properties*, Freeman, New York.
- [35] Slupsky, C.M., Kay, C.M., Reinach, F.C., Smillie, L.B. and Sykes, B.D. (1995) *Biochemistry* 34, 7365–7375.
- [36] Wishart, D.S., Sykes, B.D. and Richards, F.M. (1992) *Biochemistry* 31, 1647–1651.
- [37] Stejskal, E.O. and Tanner, J.E. (1965) *J. Chem. Phys.* 42, 288–292.
- [38] Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.* 14, 51–55.