Solution Structure and Backbone Dynamics of the Human α3-Chain Type VI Collagen C-Terminal Kunitz Domain[†,‡]

Morten Dahl Sørensen,[§] Søren Bjørn,[∥] Kjeld Norris,[∥] Ole Olsen,[∥] Lars Petersen,[∥] Thomas L. James,[⊥] and Jens J. Led*[,§]

Department of Chemistry, University of Copenhagen, The H. C. Ørsted Institute, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark, Health Care Discovery, Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd, Denmark, and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446

Received March 11, 1997; Revised Manuscript Received June 4, 1997[⊗]

ABSTRACT: The solution structure and backbone dynamics of the 58-residue C-terminal Kunitz domain fragment [α3(VI)] of human α3-chain type VI collagen has been studied by two-dimensional ^{1}H−^{1}H and ^{1}H−^{15}N nuclear magnetic resonance spectroscopy at 303 K. The solution structure is represented by an ensemble of 20 structures calculated with X-PLOR using 612 distance and 47 dihedral angle restraints. The distance restraints were obtained by a complete relaxation matrix analysis using MARDIGRAS. The root mean squared (rms) deviation is 0.91 Å for the backbone atoms of the residues Thr2(8)−Gly12(18), Arg15(21)−Tyr35(41), and Gly40(46)−Pro57(63). The central β-sheet [residues Ile18(24)−Tyr35(41)] and the C-terminal α-helix [residues Gln48(54)−Cys55(61)] are better defined with a backbone rms deviation of 0.46 Å. The solution structure of α3(VI) is virtually identical to the crystal structure of α3(VI) and to the solution structure of bovine pancreatic trypsin inhibitor (BPTI). The ^{15}N spin−lattice and spin−spin relaxation rates and the ^{1}H−^{15}N heteronuclear nuclear Overhauser enhancement (NOE) were analyzed using both the "model-free" formalism [Lipari, G., & Szabo, A. (1982) *J. Am. Chem. Soc. 104,* 4546−4559 and 4559−4570] and the reduced spectral density mapping procedure [Farrow, N. A., Szabo, A., Torchia, D. A., & Kay, L. E. (1995) *J. Biomol. NMR 6*, 153−162]. The results obtained from the "model-free" analysis include an overall correlation time τ_c of 3.00 ns and backbone order parameters S^2 in the range from 0.28 to 0.93. The necessity of including an exchange term in the analysis of the relaxation data from 14 residues indicated that these residues are involved in motions on the micro- to millisecond time scale. The majority of the 14 residues are located in the vicinity of the Cys14(20)−Cys38(44) disulfide bond, suggesting the presence of a disulfide bond isomerization similar to the one observed in BPTI [Otting, G., Liepinsh, E., & Wüthrich, K. (1993) *Biochemistry 32*, 3571−3582]. It is suggested that this disulfide bond isomerization is the main reason for the surprisingly small effect on trypsin inhibition observed when Thr13(19) of α3(VI) is substituted with Pro.

The bovine pancreatic trypsin inhibitor (BPTI)[1] is a very efficient inhibitor of several plasma serine proteinases, including kallikrein, plasmin, and protein C, a precursor of a plasma serine proteinase (Fritz & Wunderer, 1983; Norris et al., 1990; Westaby, 1993). Kallikrein, plasmin, and protein C are all very important players in the process of blood coagulation, and BPTI (often also referred to as aprotinin) is, therefore, highly interesting from a pharmaceutical point of view. Thus, BPTI has been used as a therapeutic agent in connection with cardiopulmonary bypass surgery, where it leads to improved hemostasis through the inhibition of kallikrein, plasmin, and protein C (Westaby, 1993). One drawback in BPTI therapy is that the protein is a bovine product and may, therefore, cause a primary allergic reaction. To circumvent this problem, considerable attention has been focused on human homologues with corresponding properties.

Sequences of approximately 60 residues, which are highly similar to the sequences of serine proteinase inhibitors, have been found in large proteins, *e.g.*, in a precursor protein of Alzheimer β-amyloid protein (Ponte et al., 1988; Tanzi et al., 1988; Kitaguchi et al., 1988) and in the α3 chain of human type VI collagen (Chu et al., 1990). Alignment of a 58-residue sequence from the C-terminal domain, C5, of the human type VI collagen α3 chain with the sequence of BPTI

[†] This work was supported by the Danish Technical Research Council. (16-4963-1, 16-5028-1, and 9400446), the Danish Natural Science Research Council (9400903), the Ministry of Industry (85886), Julie Damm's Studiefond, Direktør Ib Henriksen's Fond, and Novo Nordisk Fondet.

[‡] Coordinates of the 20 refined structures and the NMR-derived restraints have been deposited in the Brookhaven Protein Data Bank (PDB ID code 1KUN).

* Author to whom correspondence should be addressed.

[§] University of Copenhagen.

[∥] Novo Nordisk A/S.

[⊥] University of California.

[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1997.

[1] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; C5, C-terminal domain of the human type VI collagen α3-chain; NMR, nuclear magnetic resonance; α3(VI), the fragment from residue 7 to 64 of the 76-residue C5 Kunitz domain from the α3-chain of human type VI collagen; TPPI, time-proportional phase increment; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement correlated spectroscopy; DANTE, delays alternating with nutation for tailored excitation; DQF, double-quantum filtered; COSY, two-dimensional correlated spectroscopy; E.COSY, exclusive two-dimensional correlated spectroscopy; J_{NHH^α}, three-bond coupling between the α proton and the amide proton; $J_{H^\alpha H^\beta}$, three-bond coupling between the α proton and the β proton; HSQC, heteronuclear single quantum coherence; DGSA, distance geometry simulated annealing; CSA, chemical shift anisotropy; ppm, parts per million; rms, root mean squared; APPI, Alzheimer precursor protein inhibitor domain.

S0006-2960(97)00557-6 CCC: $14.00

```
                5         10        15        20        25
α3(VI): E T D I C K L P K D E G T C R D F I L K W Y Y D P N T K S
BPTI  : R P D F C L E P P Y T G P C K A R I I R Y F Y N A K A G L
               30        35        40        45        50        55
α3(VI): C A R F W Y G G C G G N E N K F G S Q K E C E K V C A P V
BPTI  : C Q T F V Y G G C R A K R N N F K S A E D C M R T C G G A
```

FIGURE 1: Alignment of the amino acid sequence of the 7−64 fragment [α3(VI)] of the human type VI collagen C5 domain and BPTI. Identical residues are boxed. Note that the numbering of α3(VI) starts at 1 to maintain direct comparison with BPTI.

reveals (Figure 1) that the two sequences have a 43% homology including 18 identical residues, most notably the six cysteines. It is also noteworthy that both sequences have a basic residue in position 15, the active inhibitor (P1) site.

Recently the NMR solution structure of the 76-residue complete C5 domain (Zweckstetter et al., 1996) and the X-ray crystal structure of the 58-residue fragment from residue 7 to residue 64 (the 7−64 fragment) shown in Figure 1 (Arnoux et al., 1995) were published. The solution structure of the complete C5 domain contains a well-defined 55-residue core (the residues from 8 to 62) and disordered N- and C-terminals (7 and 12 residues, respectively). The global folds of the X-ray structure of the 7−64 fragment and the well-defined core of the complete C5 domain were found to be very similar to the global fold of BPTI.

The 7−64 fragment of the C5 domain has no inhibitory activity toward trypsin, kallikrein, plasmin, and other proteinases (Arnoux et al., 1995; Zweckstetter et al., 1996) despite its structural similarity to BPTI and despite the basic residue in the active inhibitor site. The primary cause of the different inhibitory activity of the 7−64 fragment and BPTI is most likely the aspartate in position 16(22) of the 7−64 fragment (Dufton, 1985). Thus, both Arnoux et al. (1995) and Zweckstetter et al. (1996) found that substitution of this aspartate with alanine results in partial recovery of antiprotease activity. However, it was clear that other factors also contribute to the difference in inhibitory activity (Arnoux et al., 1995).

The tight binding of an inhibitor to the active site of the target enzyme occurs in a reversible way (Laskowski & Kato, 1980) and is expected to depend not only on the structural properties but also on the dynamical properties of both the inhibitor and the enzyme. Internal motions on a time scale ranging from picoseconds to seconds have been proposed to play an important role in protein functions such as ligand binding, enzyme catalysis, etc. (Karplus & McCammon, 1983). However, so far the correlation between protein function and protein dynamics is not completely understood.

Here we present the results of an independent NMR study of the 7−64 fragment of the C5 domain of the human type VI collagen α3 chain. The results include a determination of the three-dimensional solution structure as well as a description of the backbone dynamics obtained from the ^{15}N longitudinal and transverse relaxation rates and the $^{1}H-^{15}N$ heteronuclear NOEs. This combination of the solution structure and the backbone dynamics provides new insight into the factors important for the inhibitory activity. Also, it allows further investigation of the different dynamical behavior observed for the 7−64 fragment in the studies of Arnoux et al. (1995) and Zweckstetter et al. (1996).

In the following the 7−64 fragment is referred to as α3(VI). Throughout the paper sequence numbers of α3(VI) will start with residue 1 to maintain direct comparison with the well-known BPTI. Sequence numbers in parentheses refer to the corresponding sequence number in the complete C5 domain.

EXPERIMENTAL PROCEDURES

(a) Sample Preparation. Recombinant α3(VI) was produced in transformed yeast as previously described (Arnoux et al., 1995). Uniform isotropic labeling was obtained by use of ^{15}N-ammonium sulfate as the only nitrogen source during fermentation. The unlabeled α3(VI) was purified as described by Arnoux et al. (1995), while the ^{15}N-labeled α3(VI) was purified as follows.

The yeast supernatant, approximately 2 L, was adjusted to pH 3.0 with 5% acetic acid and phosphoric acid, and applied to a 20 mL column of S-Sepharose Fast Flow (Pharmacia) equilibrated with 50 mM formic acid, pH 3.7. After thorough washing with equilibrated buffer, the column was developed with a linear gradient over 100 min from 0 to 1 M NaCl in the formic acid buffer. The flow was kept at 4 mL/min, and fractions of 8 mL were collected. Aliquots were analyzed by reverse-phase HPLC, and fractions 21−36 were combined for final purification on a Vydac C4 (22 × 250 mm) preparative reverse-phase HPLC column (214TP1022, The Separation Group). The flow rate was 4 mL/min. The column was developed with a gradient from 5% to 75 % acetonitrile in 0.1 M trifluoroacetic acid. Aliquots were analyzed by HPLC, and fractions containing the purified inhibitor were combined.

The yield of essentially pure recombinant α3(VI) was 84.9 mg. The primary structure was confirmed by N-terminal sequence analysis. Electrospray mass spectrometry (Sciex API III) demonstrated that α3(VI) had been uniformly ^{15}N-labeled with 100% efficacy (found MW 6699.6, calculated 6699.4). The purified product was concentrated and the acetonitrile was evaporated by vacuum centrifugation. The final concentration of α3(VI) was 44.9 mg/mL as determined by quantitative amino acid analysis.

In the NMR samples the concentration of either ^{15}N-labeled or unlabeled α3(VI) was 2 mM, while the pH was in the range from 2.8 to 3.0 (direct meter reading).

(b) Trypsin Inhibition. Kinetic studies of the inhibition of trypsin by α3(VI) and the [D16(22)A], [D16(22)A, R15(21)K], and [T13(19)P, D16(22)A, R15(21)K] mutants of α3(VI) were performed using UV spectroscopy. The studies were carried out in 50 mM Tris-HCl and 150 mM NaCl, pH 7.4 at 298 K in 96-well microtiter plates. The change in absorbance was measured at 405 nm by means of a SLT ELISA reader. Porcine trypsin (8.2 nM) was incubated with various fixed concentrations (0−300 nM) of wild-type α3(VI) or one of the mutants for 15 min. Substrate (0.5 mM D-valyl-L-leucyl-L-lysine-*p*-nitroanilide) was then added and residual activity was measured.

The chromogenic substrate was obtained from Chromogenix (Molndal, Sweden). Porcine trypsin was from Novo Nordisk A/S (Bagsværd, Denmark). The α3(VI) mutants were obtained by site-directed mutagenesis and purified as described for α3(VI) (Arnoux et al., 1995).

(c) NMR Experiments. A series of $^{1}H-^{1}H$ and $^{1}H-^{15}N$ chemical shift correlated two-dimensional (2D) NMR spectra were recorded at a ^{1}H frequency of 500 MHz and a ^{15}N frequency of 50.68 MHz on a Bruker AM500 spectrometer equipped for inverse detection. The temperature was 303

K in all experiments. All spectra were recorded with sequential quadrature in the t_2 dimension (Redfield & Kunz, 1975) and time-proportional phase increment (TPPI) in the t_1 dimension (Drobny et al., 1979; Bodenhausen et al., 1980; Marion & Wüthrich, 1983), allowing the spectra to be represented in the phase-sensitive mode.

Five NOESY (Jeener et al., 1979; Macura et al., 1981) spectra were recorded for the generation of distance restraints. Four were recorded with a prescan delay of 3.75 s. Two of these were recorded on a sample in H_2O using mixing times of 100 and 170 ms, respectively, while two were recorded on a sample in D_2O using the same mixing times. Furthermore, a NOESY spectrum with a mixing time of 170 ms was recorded on the sample in H_2O using a prescan delay of 1.55 s. The five NOESY spectra each consisted of 1024 t_1 data points and 3000 t_2 data points. The sweep width was 10 000 Hz in both dimensions. The total acquisition time for each spectrum was 41 h. Water suppression was achieved using a DANTE pulse (Bodenhausen et al., 1976) during a 1.55 s delay (included in the prescan delay) immediately prior to the first 90° pulse. Additional water suppression was achieved by placing a 180° composite pulse in the middle of the mixing period. In the fingerprint region, correlations that involve α-protons affected by the DANTE pulse were identified in a NOESY jump-return spectrum. This spectrum was used only for identification purposes and not in the subsequent (*vide infra*) MARDIGRAS calculations.

Three spectra were recorded for the determination of coupling constants used for generation of dihedral angle restraints. A DQF-COSY (Rance et al., 1983) spectrum was recorded on the sample in H_2O to obtain the J_{NNH^α} coupling constants. A DQF-COSY spectrum and an E.COSY (Griesinger et al., 1985, 1986, 1987) spectrum were recorded on the sample in D_2O for the determination of the $J_{H^\alpha H^\beta}$ coupling constants. The DQF-COSY spectrum in H_2O consisted of 808 t_1 data points and 3000 t_2 data points, whereas the DQF-COSY spectrum in D_2O consisted of 1024 t_1 data points and 3000 t_2 data points. For both spectra the sweep width was 7353 Hz in both dimensions. The E.COSY spectrum consisted of 2048 t_1 data points and 8192 t_2 data points. The sweep width was 6250 Hz in both dimensions. In all three spectra, water suppression was achieved using a DANTE pulse during a 1.2 s prescan delay. The total acquisition times ranged from 23 to 112 h.

The digital resolution of the NOESY, DQF-COSY, and E.COSY spectra was 4.88, 3.59, and 0.76 Hz/point in the F_2 dimension and 9.76, 7.18, and 3.05 Hz/point in the F_1 dimension, respectively, after zero-filling and Fourier transformation. The window functions used for resolution enhancement in the NOESY spectra were, in the t_2 dimension, a Gaussian line-broadening function of −15 Hz and a scaling factor of 0.15 and, in the t_1 dimension, a squared sine bell function shifted 70°. In the DQF-COSY and E.COSY spectra a combination of a sine bell function and a squared cosine bell function was used in both dimensions.

The ^{15}N relaxation rates, R_1 and R_2, were measured using the pulse sequences suggested by Kay et al. (1992). A total of seven spectra were recorded in both the R_1 and R_2 series. The relaxation delays used were, in the R_1 series, 20, 100, 250, 400, 600, 850, and 1100 ms and, in the R_2 series, 18.9, 56.8, 94.6, 189.2, 302.7, 425.7, and 851.4 ms. The Carr−Purcell−Meiboom−Gill spin−echo period was 0.9 ms. Water suppression was achieved using a DANTE pulse during a 1.2 s prescan delay.

The $^1H-^{15}N$ nuclear Overhauser effects (NOEs) were measured using the jump-return pulse sequence previously described (Sørensen et al., 1995). Two sets of spectra were recorded in order to cover the complete amide region. The jump-return delay was in one set 150 μs and in the other 120 μs, corresponding to excitation maxima at 8.1 and 8.9 ppm, respectively. The set using a jump-return delay of 120 μs was recorded twice to evaluate of the uncertainty of the NOEs. A prescan delay of 4 s was used in all $^1H-^{15}N$ NOE experiments. Finally, a $^1H-^{15}N$ HSQC (Bodenhausen & Ruben, 1980) spectrum was recorded for the assignment of the ^{15}N frequencies.

All $^1H-^{15}N$ correlated spectra consisted of 256 t_1 data points and 3000 t_2 data points, the sweep width being 10 000 Hz in the ^1H dimension and 1388.89 Hz in the ^{15}N dimension. The ^1H carrier was positioned on the water resonance at 4.71 ppm, whereas the ^{15}N carrier was positioned at 117.9 ppm. The total recording time ranged from 14 to 42 h.

The digital resolution was 4.88 and 5.42 Hz/point in the ^1H and ^{15}N dimensions, respectively, after zero-filling and Fourier transformation. The applied window functions were identical to those used for the $^1H-^1H$ NOESY spectra.

(d) Structure Determination: (1) Restraints. Based on the ^1H assignments published previously (Sørensen et al., 1996), interproton distance restraints were obtained from the NOESY experiments using the newly developed procedure RANDMARDI (Liu et al., 1995) and the program MARDIGRAS (Borgias & James, 1990). The volumes of the NOESY peaks, used in the complete relaxation matrix analysis, were obtained by fitting a Gaussian line shape function to the peaks using the program SPARKY (University of California, San Francisco). Distance restraints were established with RANDMARDI using the peak volumes, the correlation times 2, 3, and 4 ns, and two different starting structures obtained from preliminary calculations using the program DIANA (Güntert et al., 1991). The number of MARDIGRAS calculations within RANDMARDI was set to 30, giving a total of 180 MARDIGRAS calculations for each of the five NOESY spectra. The peak integration errors used in RANDMARDI were obtained from a comparison of the volumes of the peaks above and below the diagonal whenever possible. From this comparison integration errors of approximately 10%, 20%, and 35% were derived for strong, medium, and weak signals, respectively. In cases where the volume could be obtained from only one side of the diagonal, the peak integration error was evaluated from the signal intensity using these errors. The upper and lower bound for each restraint was taken as the largest and smallest distance, respectively, calculated from the five NOESY spectra using the RANDMARDI procedure. No lower bound was used for restraints involving α-protons affected by the water suppression.

Pseudoatoms were treated in the following way. Methyl groups, nondistinguishable 2,6H and 3,5H pairs in phenylalanine and tyrosine rings, and nondistinguishable methylene pairs were treated as pseudoatoms by MARDIGRAS, with the resulting distance being to the geometrical center of the proton group. A three-site jump model for intra- and interresidue distances was used for modeling the internal motion of methyl groups, and a two-site jump model was used for methylene groups and aromatic rings (Liu et al.,

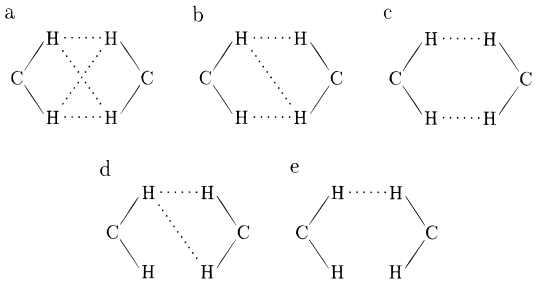

FIGURE 2: Five different situations encountered with NOEs (represented by dotted lines) between two nonstereospecifically assigned methylene groups. The restraints between the geometric centers were in arrangements a, b, and c the largest upper bound and the smallest lower bound. In arrangement d, the restraint was taken as the largest upper bound plus a pseudoatom correction for the group with NOEs to only one proton. A corresponding procedure was applied for the lower bound. In arrangement e, pseudoatom corrections for both groups were added to the upper bound and subtracted from the lower bound.

1992). Accordingly, no pseudoatom correction was applied in these cases. Nonstereospecifically assigned methylene protons with different chemical shifts were treated individually by MARDIGRAS. In cases where a NOE to only one of the methylene protons could be identified, the restraint was to the geometric center of the methylene group, and a pseudoatom correction factor (Wüthrich et al., 1983) was added to (subtracted from) to the upper (lower) bound. When there were NOEs between both protons of a methylene pair and another proton, the largest upper bound and the smallest lower bound were used as restraints to the geometric center. In cases where the NOEs were between two nonstereospecifically assigned methylene groups, distance restraints between the geometric centers were as shown in Figure 2. The lower bound was automatically reset to 1.8 Å, if the above procedure resulted in a smaller value. Redundant distance restraints were eliminated with the program DIANA prior to X-PLOR structure calculations.

A total of 39 J_{NNH^α} coupling constants were obtained from an analysis of the DQF-COSY spectrum as described elsewhere (Sørensen et al., 1996). In cases where $J_{NNH^\alpha} >$ 10 Hz, $J_{NNH^\alpha} >$ 9 Hz, $J_{NNH^\alpha} >$ 8 Hz, $J_{NNH^\alpha} <$ 6 Hz, and $J_{NNH^\alpha} <$ 5 Hz by at least 1 standard deviation, the applied restraints on the ϕ angles were $-140 < \phi < -100$, $-145 < \phi < -95$, $-160 < \phi < -80$, $-95 < \phi < -45$, and $-85 < \phi < -35$, respectively. The χ_1 angles were determined from a combination of intraresidual NH−H$^\alpha$ and NH−H$^\beta$ NOEs and $J_{H^\alpha H^\beta}$ coupling constants as outlined by Wagner et al. (1987). If one of the three staggered conformations was identified unambiguously, a χ_1 angle restraint of ±30° was applied. The observation of one small and one large $J_{H^\alpha H^\beta}$ coupling constant indicates a g^2t^3 or a t^2g^3 conformation. In cases where these conformations could not be distinguished, a χ_1 angle restraint of $-120° \pm 100°$ was applied.

(2) Structure Calculations. The structure calculations were carried out using the program X-PLOR 3.1 (Brünger, 1992). The starting point was the generation of substructures using the DG-SUB-EMBED protocol. The backbone carbon, nitrogen and hydrogen atoms and the β- and γ-carbons were included in the substructures. Subsequent to the generation of substructures, distance geometry simulated annealing (DGSA) calculations (Nilges et al., 1988) were carried out. The structures were first run through 1000 cycles of restrained energy minimization and then subjected to 4 ps of restrained molecular Verlet dynamics at 3000 K using time steps of 0.001 ps. The next step was a 4 ps cooling to 100 K. During the cooling stage the temperature was varied in steps of 50 K. Thus, 0.069 ps of restrained molecular Verlet dynamics were performed at each temperature using time steps of 0.001 ps. Subsequent to cooling, the structures were run through an additional 1000 cycles of restrained energy minimization. In the DGSA protocol, the van der Waals energies were represented by the simple repel function. During the cooling stage, the van der Waals interactions were increased by varying the force constant of the van der Waals repel function from 0.003 to 4 kcal·mol^{-1}·Å^{-4}.

Finally, the structures from the DGSA calculation were subjected to 3000 cycles of restrained energy minimization. In this final refinement, the full CHARMM energy (Brooks et al., 1983; Brünger, 1992) function was employed, including a Lennard−Jones potential function and an explicit hydrogen-bond energy term. The X-PLOR parameter file parmallh3x.pro was used. During all calculations the force constants used for the NOE and torsion angle restraints were 50 kcal·mol^{-1}·Å^{-2} and 200 kcal·mol^{-1}·rad^{-2}, respectively.

(e) Relaxation Data Analysis. The relaxation rates, R_1 and R_2, were found by a nonlinear least-squares fit of a single-exponential function to the peak volumes. For each residue, the ^{1}H−^{15}N NOE was calculated by dividing the volume of the peak in the spectrum recorded with ^{1}H saturation by the corresponding volume in the spectrum recorded without ^{1}H saturation. The ^{1}H−^{15}N peak volumes were determined using the program CROSSFIT (University of Copenhagen), based on a combination of linear prediction and least-squares analyses (Gesmar et al., 1994; Kristensen et al., 1996).

Relaxation Theory. The relaxation of the ^{15}N nuclei at high field strengths is primarily caused by the dipolar interaction with the directly attached proton and by the chemical shift anisotropy (CSA) mechanism. The ^{15}N relaxation parameters can be expressed using the values of the spectral density function J, evaluated at the five frequencies 0, ω_N, $\omega_H + \omega_N$, ω_H, and $\omega_H - \omega_N$ (Abragam, 1961):

$$R_1 = \frac{d^2}{4}[3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N) \quad (1)$$

$$R_2 = \frac{d^2}{8}[4J(0) + 3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{c^2}{6}[4J(0) + 3J(\omega_N)] \quad (2)$$

$$\mathrm{NOE} = + \frac{\gamma_H}{\gamma_N}\frac{1}{R_1}\frac{d^2}{4}[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)] \quad (3)$$

where $d^2 = (\mu_0 h/8\pi^2)^2 \gamma_H^2 \gamma_N^2 \langle 1/r_{NH}^3 \rangle^2$, $c^2 = \omega_N^2(\sigma_\parallel - \sigma_\perp)^2/3$, ω_H and ω_N are the Larmor frequencies of the ^{1}H and ^{15}N nuclei, while μ_0 is the permeability of vacuum, h is Planck's constant, γ_H and γ_N are the gyromagnetic ratios of the ^{1}H and ^{15}N nuclei, r_{NH} is the nitrogen−proton bond length, and $\sigma_\parallel$ and $\sigma_\perp$ are, respectively, the parallel and perpendicular components of the ^{15}N chemical shift tensor, assuming an axially symmetric chemical shift tensor. The value of the chemical shift anisotropy ($\sigma_\parallel - \sigma_\perp$) varies with the conformation of the polypeptide backbone and takes approximate values of −160 ppm (α-helix and turns) and −150 ppm (β-

sheet and random coil) (Shoji et al., 1990). For $r_{NH} = 1.02$ Å, $(\sigma_{\parallel} - \sigma_{\perp}) = -160$ ppm, and a magnetic field strength of 11.74 T the values $d^2 = 5.2 \times 10^9$ (rad/s)^2 and $c^2 = 8.6 \times 10^8$ (rad/s)^2 are obtained.

The spectral density function $J(\omega)$ is twice the cosine transform of the angular autocorrelation function (Lipari & Szabo, 1982a), which describes the fluctuations of each NH vector (Abragam, 1961). The spectral density function therefore gives the distribution of frequencies contained in the fluctuations of the NH vector and consequently contains information on the dynamical behavior of the protein backbone. For a protein in solution, $J(\omega)$ depends on both the overall tumbling of the macromolecule and the internal motions of the NH vector.

In addition to dipolar coupling and CSA relaxation mechanisms, R_2 can also be affected by chemical exchange processes. In this case an exchange term

$$R_2(\text{obs}) = R_2 + \pi\Delta\text{ex} \tag{4}$$

must be included. Here R_2 is given by eq 2 and Δex is the increase in ^{15}N line width due to chemical exchange.

Obviously, the three ^{15}N relaxation parameters are insufficient to determine the spectral density function at the five frequencies in eqs 1−3. Several solutions to this problem have been proposed, including analyses within the framework of a specific theoretical model of motion, *e.g.*, the "wobbling-in-a-cone" model (Woessner, 1962; Kinosita et al., 1977; Richarz et al., 1980). The most common approach is, however, the so-called "model-free" formalism of Lipari and Szabo (1982a,b), in which the overall tumbling of the molecule is described by a single correlation time, and the internal motions are described by two parameters, *i.e.*, a generalized order parameter and an effective internal correlation time. Alternatively, the "spectral density mapping" approach (Peng & Wagner, 1992a,b) can be applied. In this approach, six relaxation parameters are determined to evaluate the spectral density function at all five frequencies. There are, however, some problems in this approach that have led to the "reduced spectral density mapping" (Farrow et al., 1995a,b; Lefèvre et al., 1996) based on the ^{15}N R_1, R_2, and NOE relaxation data exclusively. Analyses of the ^{15}N relaxation data of α3(VI) using both the "model-free" approach and reduced spectral density mapping will be presented.

Calculation of the "Model-Free" Parameters. In the "model-free" approach the spectral density function $J(\omega)$ takes the form (Lipari & Szabo, 1982a)

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

where S^2 is an order parameter, τ_c is the correlation time for the overall tumbling of the molecule, and $1/\tau = (1/\tau_c + 1/\tau_i)$, where τ_i is the correlation time for the internal motion. For an NH vector completely fixed or completely free, S^2 takes the values 1 and 0, respectively. A physical interpretation of τ_i is more difficult and requires a specific motional model. The spectral density function in eq 5 reduces to

$$J(\omega) = \frac{2}{5}\left(\frac{S^2\tau_c}{1 + \omega^2{\tau_c}^2}\right) \tag{6}$$

in the limit $\tau_i \rightarrow 0$. Insertion of eq 6 into eq 3 shows that in this limit the heteronuclear NOE is independent of S^2 as well as of τ_i.

A more elaborate form of $J(\omega)$ was proposed by Clore et al. (1990) to account for data that could not be described by eq 5. In this description it is assumed that internal motions on two different time scales are present, resulting in the following form of $J(\omega)$

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1 + \omega^2{\tau_c}^2} + \frac{(S_f{}^2 - S^2)\tau}{1 + \omega^2\tau^2}\right] \tag{7}$$

where the order parameters of the fast and slow internal motions are called S_f^2 and S_s^2, respectively. The order parameter, S^2, is defined as $S_f^2 \cdot S_s^2$, and $1/\tau = (1/\tau_c + 1/\tau_{is})$, τ_{is} being the correlation time for the slow internal motion.

Determination of the Overall Correlation Time. The first step in a "model-free" analysis is the determination of the overall correlation time τ_c. The ratio R_1/R_2 is independent of both S^2 and τ_i when the internal motion is negligible (Kay et al., 1989), thus allowing a determination of τ_c on a residue-by-residue basis. In a first approximation we assume that internal motion is negligible for the NH vectors of residues involved in secondary structure elements. Accordingly, a total of 30 NH vectors [the residues Asp3(9)−Leu7(13), Ile18(24)−Asp24(30), Asn26(32)−Tyr35(41), residue Phe45(51), and the residues Gln48(54)−Val54(60)] were used in a preliminary determination of τ_c, giving a value of 2.95 ± 0.26 ns. In the next step, the 10 NH vectors that gave rise to a τ_c outside 1 standard deviation were discarded, and the overall correlation was determined from the relaxation data of the 20 remaining NH vectors by a least-squares procedure. In this procedure one overall correlation time and the individual order parameters and internal correlation times were optimized simultaneously using eqs 5 or 6. The optimization resulted in $\tau_c = 3.00$ ns, which was used in the following calculations.

The least-squares procedure used to determine τ_c consisted of a minimization through a grid-search of

$$\chi^2 = \sum_j \left[\frac{R_1(\text{c})_j - R_1(\text{o})_j}{\sigma(R_1)_j}\right]^2 + \left[\frac{R_2(\text{c})_j - R_2(\text{o})_j}{\sigma(R_2)_j}\right]^2 + \left[\frac{\text{NOE(c)}_j - \text{NOE(o)}_j}{\sigma(\text{NOE})_j}\right]^2 \tag{8}$$

where the sum is over the involved residues, and R_1(c), R_2(c), and NOE(c) are the values calculated from eqs 5 or 6. The σ values are the corresponding standard deviations.

Determination of the Parameters Describing the Internal Motions. Determination of S^2 and τ_i were done by optimizing these parameters in a least-squares procedure identical to the one described above. The optimization was done on a residue-by-residue basis, keeping the overall correlation time fixed at 3.00 ns. The spectral density functions in eqs 5 and 6 were used in the analysis of the ^{15}N backbone relaxation data of 31 and 20 residues, respectively. Note that only the R_1 and R_2 terms were included in the analysis when eq 6 was used, since the NOE in this case is independent of S^2 (*vide supra*).

Reduced Spectral Density Mapping Analysis. The reduced spectral density mapping was done using method 2 of Farrow

et al. (1995b). That is, $J(0.87\omega_H)$, $J(\omega_N)$, and $J(0)$ were determined from

$$J(0.87\omega_H) = \frac{4}{5d^2}\frac{\gamma_N}{\gamma_H}(\mathrm{NOE} - 1)R_1 \quad (9)$$

$$J(\omega_N) = \frac{R_1 - (7d^2/4)J(0.921\omega_H)}{(3d^2/4) + c^2} \quad (10)$$

$$J(0) = \frac{R_2 - (d^2/8)[3J(\omega_N) + 13J(0.955\omega_H)] - (c^2/2)J(\omega_N)}{d^2/2 + 2c^2/3} \quad (11)$$

where $J(0.921\omega_H)$ and $J(0.955\omega_H)$ were derived from $J(0.87\omega_H)$ through $J(\epsilon\omega_H) = (0.87/\epsilon)^2 J(0.87\omega_H)$, where $\epsilon =$ 0.921 or 0.955. In the case of residues Glu11(17), Gly12-(18), Gly37(43), and Lys44(50) no NOE could be measured, and $J(0.87\omega_H)$ was set to the average $J(0.87\omega_H)$ value obtained from the other residues.

Standard deviations were obtained from a Monte Carlo procedure in both the "model-free" approach and the reduced spectral density mapping. Thus, 30 R_1, R_2, and NOE data sets were generated by adding random errors to the measured parameters. The errors are normally distributed with the measured standard deviations. From the optimized parameters determined from the 30 data sets an average value and a standard deviation were derived for each parameter.

RESULTS

(a) Quality of the Calculated Structures. A total of 612 nontrivial NOE distance restraints were obtained. These restraints (Table 1 of the supporting information) are composed of 170 intraresidual restraints, 114 sequential restraints, and 328 intraresidual restraints. In addition, 24 ϕ and 23 χ_1 dihedral angle restraints (Table 2 of the supporting information) were included in the final calculations. The distribution of the NOEs is shown in Figure 3a.

Fifty structures were calculated using the procedure outlined in Experimental Procedures. The average values of the rms deviations of the 20 structures with the lowest total energy from their mean structure are 1.47 Å for all heavy atoms and 0.91 Å for all N, C^α, and C′ backbone atoms if residues Glu1(7), Thr13(19), Cys14(20), Gly36(42)−Gly39(45), and Val58(64) are excluded from the fitting. The corresponding rms deviations are 1.10 and 0.46 Å if only the β-sheet residues [Ile18(24)−Tyr35(41)] and the α-helix residues [Gln48(54)−Cys55(61)] are included in the fitting. The residual rms deviations from the mean structure are shown in Figure 3b. A best-fit superposition of the backbone of the 20 structures is shown in Figure 4.

The statistics in Table 1 show that the calculated structures are in good agreement with the experimental data. Furthermore, no residual distance restraint violation exceeded 0.3 Å in any of the 20 structures. Finally, the negative Lennard−Jones van der Waals energy of −249.433 kcal mol^{-1} indicates that the structures have good nonbonding contacts.

(b) Description of the Structure. Previously (Sørensen et al., 1996) the secondary structure was determined from NOEs and J_{NHH^α} coupling constants. The observation of a β-sheet spanning residues Ile18(24)−Tyr35(41) is confirmed by the backbone dihedral angles ϕ and ψ. Thus, residues Ile18(24)−Asp24(30) and Cys30(36)−Tyr35(41) all have ϕ angles in the range from −180° to −60° and ψ angles in the range

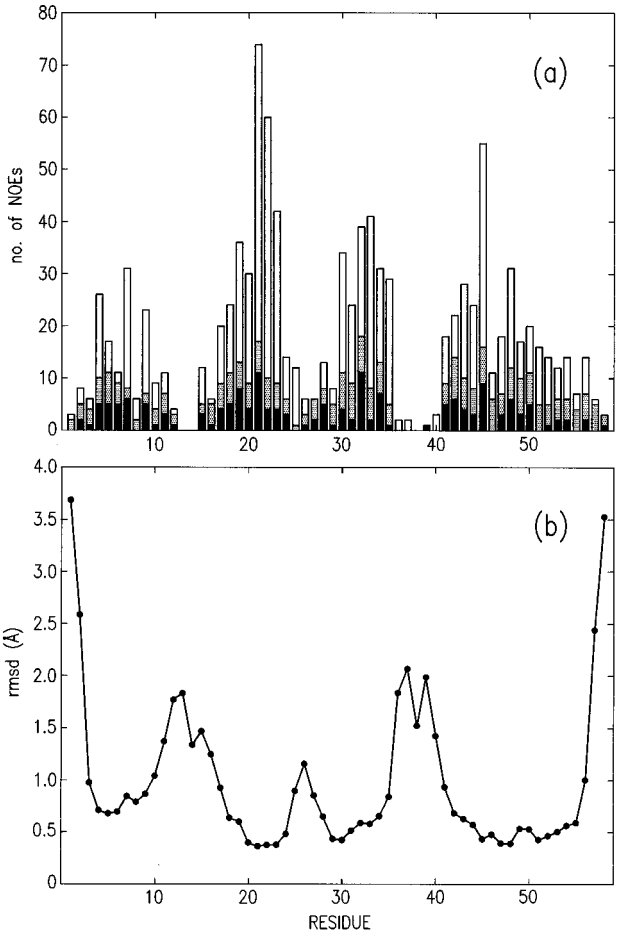

FIGURE 3: (a) Distribution of the NOEs used in the determination of the α3(VI) structure. All sequential and interresidual NOEs are represented twice, once for each of the involved residues. The intraresidual, sequential, and interresidual NOEs are indicated by a black bar, a dotted bar, and a empty bar, respectively. (b) Average N, C′, and C^α backbone rms deviations (in angstroms) from the mean structure for the 20 solution structures of α3(VI) with the lowest total energy.

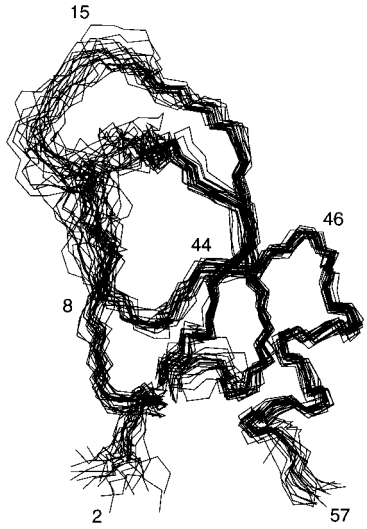

FIGURE 4: Superposition of the 20 solution structures of α3(VI) with the lowest total energy. Only the N, C′, and C^α backbone atoms for the residues from Thr2(8) to Gly12(18), from Arg15(21) to Tyr35(41), and from Gly40(46) to Pro57(63) are superimposed.

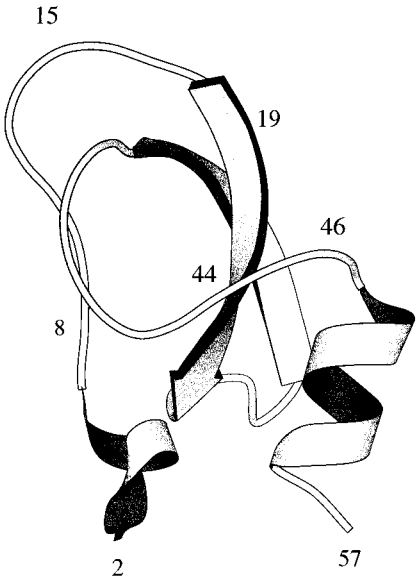

FIGURE 5: Schematic representation of the structure of α3(VI) highlighting the secondary structure. The figure was prepared using the program MOLSCRIPT (Kraulis, 1991).

Table 1: Structural Statistics

	mean	range
rms Deviations from NOE Restraints and from the Idealized Geometry Used within X-PLOR		
NOE (Å)	0.026	0.010
bond length (Å)	0.016	0.002
bond angles (deg)	3.430	0.204
improper dihedral angles (deg)	0.430	0.107
X-PLOR Potential Energies (kcal·mol^{-1})[a]		
total	−969.599	120.308
van der Waals	−249.433	35.111
NOE restraints	20.140	16.791
dihedral angle restraints	0.274	0.642

[a] The force constants of the NOE and torsion angle terms were 50 kcal·mol^{-1}·Å^{-2} and 200 kcal·mol^{-1}·rad^{-2}, respectively.

from 20° to 180° in the structure with the lowest total energy. The ψ angle of Ser29(35) is $-154.7° \pm 4.1°$ in the 20 structures and consequently falls outside the above range. It is therefore difficult to specify whether Ser29(35) is the last residue in the turn or the first residue in the second β-strand. In addition, Phe45(51) has ϕ and ψ angles in the above-mentioned range, which supports the observation (Sørensen et al., 1996) that this residue associates by hydrogen bonds to the central β-sheet to form a small third strand. Furthermore, the residues from Gln48(54) to Cys55(61) all have ϕ and ψ angles characteristic of an α-helix; that is, ϕ angles in the range from $-40°$ to $-180°$ and ψ angles in the range from $-10°$ to $-70°$. Accordingly, the backbone dihedral angles suggest that Cys55(61) is the last residue in the C-terminal α-helix. Finally, the observation of an irregular 3_{10}-helix from residues Thr2(8) to Leu7(13) is also supported by the ϕ and ψ dihedral angles. A schematic representation of the structure of α3(VI) highlighting the secondary structure is given in Figure 5.

Hydrogen Bonds. A total of 27 amide protons have been identified previously as slowly exchanging (Sørensen et al., 1996). It is therefore not surprising that the structure of α3(VI) obtained is characterized by numerous hydrogen bonds. Thus, 15 of the 27 slowly exchanging amide protons

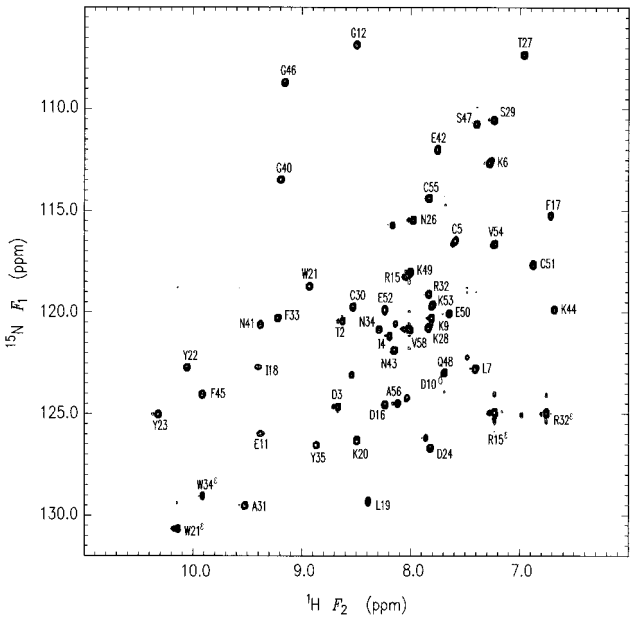

FIGURE 6: Contour plot of the first ^{1}H−^{15}N spectrum in the R_1 series. The assignments are marked by one-letter codes and residue numbers.

are involved in hydrogen bonds within the β-sheet or helical regions.

The amide protons of Tyr23(29), Lys28(34), Ser29(35), Asn43(49), Lys44(50), and Val58(64) are all involved in hydrogen bonds associated either with turns or with the tertiary structure of the protein. Five of these six hydrogen bonds can be identified immediately from the calculated structures. These are Tyr23(29) NH···δ OC Asn43(49), Lys28(34) NH···OC Asp24(30), Ser29(35) NH···γ OC Thr27(33), Asn43(49) NH···δ OC Asn41(47), and Lys44(50) NH···OC Glu42(48). In 11 or more of the selected structures these hydrogen bonds are characterized by the geometry [a N−C distance < 3.5 Å and a N−H−O bond angle in the range from 120° to 180° (Baker & Hubbard, 1984)] necessary for establishing a hydrogen bond. A hydrogen bond between Val58(64) NH and Ala56(62) CO was found in eight of the selected structures.

The final six amide protons [residues Thr2(8), Ile4(10), Leu19(25), Asn26(32), Arg32(38), and Gly36(42)], originally identified as slowly exchanging, are not involved in hydrogen bonds in more than, at most, six structures. This is in good agreement with the fact that all six amide protons only gave rise to very weak signals in the TOCSY spectrum used for the identification of slowly exchanging amide protons.

Finally, according to the definition above, three hydrogen bonds, Lys6(12) NH···OC Asp3(9), Thr27(33) NH···OC Asp24(30), and Cys55(61) NH···OC Cys51(57), were found in 19, 18, and 20 of the calculated structures, respectively, even though the amide protons were not characterized as slowly exchanging. Two of these hydrogen bonds are located in secondary structure elements [Lys6(12) NH···OC Asp3(9) in the N-terminal 3_{10}-helix and Cys55(61) NH···OC Cys51(57) in the C-terminal α-helix] while Thr27(33) NH···OC Asp24(30) is located in the turn.

(c) Relaxation Parameters. An almost complete ^{15}N resonance assignment for α3(VI) was obtained from a ^{1}H−^{15}N HSQC spectrum (Table 3 of the supporting information). Figure 6 shows a contour plot of the ^{1}H−^{15}N correlation experiment with the shortest delay in the R_1 series. The

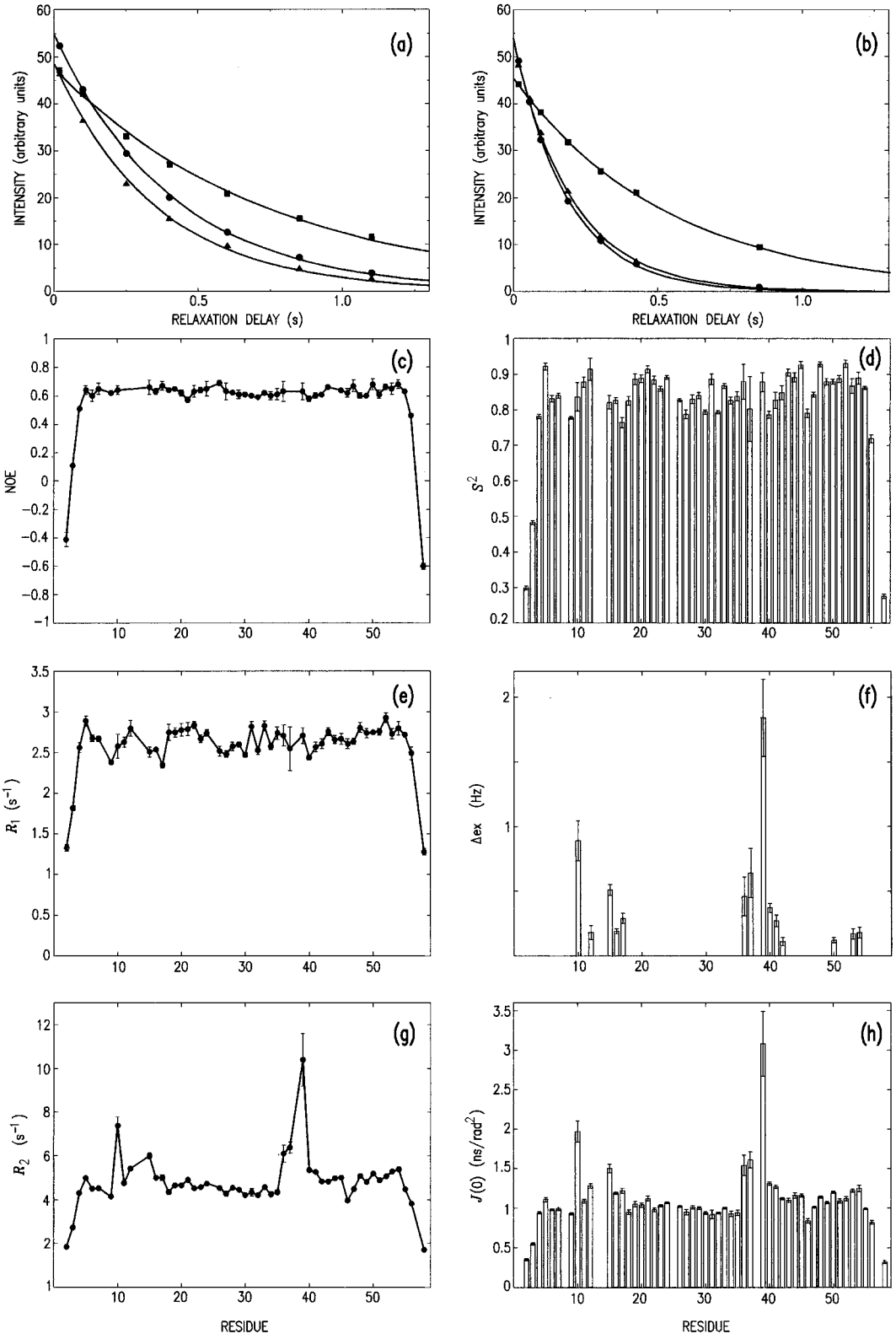

FIGURE 7: Plots of (a) the R_1 and (b) the R_2 relaxation curves for residues Thr2(8) (■), Trp21(27) (▲), and Gly40(46) (●) and of the measured ^{15}N relaxation parameters and the parameters obtained from the analysis (see text) as a function of residue number: (c) The ^{1}H−^{15}N heteronuclear NOE; (d) the order parameter, S^2; (e) the ^{15}N spin−lattice relaxation rate constant, R_1; (f) the exchange term, Δex; (g) the ^{15}N spin−spin relaxation rate constant, R_2; and (h) the value of the spectral density function at the frequency zero, $J(0)$.

quality of the relaxation data is illustrated by the relaxation curves shown in Figure 7a,b.

The calculated spin−lattice and spin−spin relaxation rate constants and the heteronuclear ^{1}H−^{15}N NOEs are plotted as a function of the amino acid sequence in Figure 7. The corresponding values are given in Table 4 of the supporting information. The experimental R_1 values of the backbone ^{15}N atoms range from 1.28 to 2.89 s^{-1}, whereas the corresponding R_2 values range from 1.72 to 10.4 s^{-1}. The average backbone ^{15}N R_2 value is 4.41 s^{-1} if the R_2 values larger than 5.2 s^{-1} are ignored [residues Asp10(16), Gly12(18), Arg15(21), Gly36(42), Gly37(43), Gly39(45), Gly40(46), Asn41(47), Lys53(59), and Val54(60)]. The backbone ^{15}N NOEs span the range from −0.60 to 0.69.

The "model-free" order parameters S^2 obtained from analysis of the relaxation data are plotted in Figure 7d. The backbone order parameters range from 0.28 to 0.93 with an average of 0.82 ± 0.13. An exchange term was included in the analysis of the relaxation data obtained for 14 residues. These residues all have $\chi^2 > 10$ and a R_2/R_1 ratio above average. The exchange term, Δex, is plotted against the amino acid sequence in Figure 7f.

Three nonterminal residues [Tyr22(28), Gly46(54), and Cys55(61)] have $\chi^2 > 10$ and a R_2/R_1 ratio significantly below average. This indicates the presence of two internal motions on different time scales (Clore et al., 1990). However, application of eq 7 in the analysis did not yield realistic dynamics parameters. A similar situation was encountered for Leu136 in *Bacillus subtilis* glucose permease IIa domain (Stone et al., 1992).

The calculated "model-free" parameters are listed in Table 5 of the supporting information. Figure 7h shows a plot of $J(0)$, the spectral density function at frequency 0, obtained from the reduced spectral density mapping analysis. The calculated values of the spectral density function at the frequencies 0, ω_N, and $0.87\omega_H$ are listed in Table 6 of the supporting information.

DISCUSSION

(a) Correlation between the Precision of the Calculated Structures and the Dynamics of α3(VI). The 20 structures of α3(VI) shown in Figure 4 and their rms deviations (Figure 3b) reveal that the secondary structure elements are well-defined in terms of convergence. However, the region around the Cys14(20)−Cys38(44) disulfide bond is poorly defined, which could indicate increased flexibility in this region. On the other hand, the poor definition primarily stems from the fact that only few structural restraints are observed and could, therefore, also be caused by spectral overlap and/or geometric factors. The cluster of glycines in positions 36(42), 37(43), 39(45), and 40(46) raises the question whether the small number of NOEs observed in the region around the Cys14(20)−Cys38(44) disulfide bond is simply a consequence of the lack of side chains or reflects a higher flexibility in this part of the molecule.

Isomerization of the Cys14−Cys38 disulfide bond in BPTI has been observed previously (Otting et al., 1993), suggesting that the poor definition of the Cys14(20)−Cys38(44) region of α3(VI) may be the result of a similar conformational flexibility. Furthermore, in a study of the complete C5 domain, Zweckstetter et al. (1996) found the loop region around the Cys14(20)−Cys38(44) disulfide bond to be more disordered than the rest of the 55-residue core. Although these observations strongly suggest that a higher conformational flexibility is indeed present, they cannot be taken as final proof. Recently, measurements of ^{15}N relaxation parameters have allowed a description of the backbone dynamics on time scales ranging from picoseconds to milliseconds [see, for example, the review by Peng and Wagner (1994)]. The ^{15}N relaxation data presented in Figure 7 therefore provide the additional information needed to establish whether or not poor structural definition reflects genuine flexibility.

Motions on the picosecond time scale are reflected in lower values of the heteronuclear $^1H-^{15}N$ NOE, whereas motions on the micro- to millisecond time scale result in larger R_2 values (Kay et al., 1989). In the "model-free" analysis, low values of S^2 correspond to motions on the pico- to nanosecond time scale, whereas the necessity of including an exchange term, Δex, in the analysis indicates motions on the micro- to millisecond time scale. Low values of $J(0)$ obtained from the reduced spectral density mapping indicates high-frequency motions. In contrast, motions on the micro- to millisecond time scale will, as reflected in R_2 values, result in larger values of $J(0)$ (see eq 11).

As mentioned above, the study of the complete C5 domain (Zweckstetter et al., 1996) indicated the presence of a Cys14(20)−Cys38(44) disulfide bond isomerization in the C-terminal Kunitz domain from human type VI collagen, similar to the one observed in BPTI (Otting et al., 1993). Further evidence for this isomerization is provided by the relaxation data presented here. Thus, the increased values of R_2 observed for residues Gly12(18), Arg15(21), Gly36(42), Gly37(43), and Gly39(45) (Figure 7g), the corresponding exchange terms (Figure 7f), and the larger $J(0)$ values (Figure 7h) all show that these residues are involved in motions on the micro- to millisecond time scale. Szyperski et al. (1993) found a correlation time of approximately 2 ms at 309 K for the Cys14−Cys38 disulfide bond isomerization process in BPTI. That is, the Cys14−Cys38 disulfide bond isomerization in BPTI and the motion affecting residues Gly12(18), Arg15(21), Gly36(42), Gly37(43), and Gly39(45) in α3(VI) occur on similar time scales. It should be noted that since no signals were observed in the $^1H-^{15}N$ spectra from residues Thr13(19), Cys14(20), and Cys38(44), no relaxation parameters could be determined for these residues. It seems likely that the signals from these residues are exchange-broadened beyond detection.

The large R_2 value observed for Asp10(16) is probably not a result of a Cys14(20)−Cys38(44) disulfide bond isomerization but is caused by interactions with internally bound water. Zweckstetter et al. (1996) and Arnoux et al. (1995) found two internal water molecules in a pocket between the residues from Pro8(14) to Asp10(16) and from Gly40(46) to Lys44(50). Exchange of internal water molecules with the bulk solvent has previously been seen to result in increased ^{15}N R_2 values (Clore et al., 1990).

In contrast to the above-mentioned correlation between a poorly defined region and protein mobility, the larger rms deviations observed for the residues in the β-turn [Asp24(30)−Ser29(35)] does *not* reflect increased conformational flexibility. This is evidenced by the fact that these residues have order parameters and $J(0)$ values (Figure 7d,h) similar to the well-defined β-sheet residues. The structural uncertainty of this region is therefore a result of geometric factors and spectral overlap, both of which reduce the number of observable NOEs, and illustrates the care one must take in making conclusions about protein mobility on the basis of structural uncertainty.

(b) Comparison of the α3(VI) Structure with Other Kunitz Domains. The structure of several Kunitz domains has been determined by X-ray crystallography and/or NMR, including the structure of BPTI (Berndt et al., 1992; Deisenhofer & Steigmann, 1975; Wlodawer et al., 1984, 1987), APPI (Heald et al., 1991; Hynes et al., 1990), dendrotoxin I (Lancelin et al., 1994), dendrotoxin K (Berndt et al., 1993), and a Kunitz-type proteinase inhibitor from the sea anemone *Stichodactyla helianthus* (Antuch et al., 1993). Recently, the X-ray structure of α3(VI) (Arnoux et al., 1995) and the NMR structure of the complete C5 domain (Zweckstetter et al.,

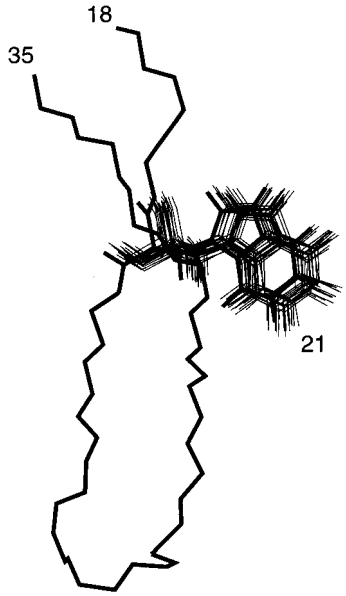

FIGURE 8: Position of the Trp21(27) ring in the 20 solution structures (thin lines) and the crystal structure (thick line). The N, C′, and C^{α} atoms of the residues from Thr2(8) to Gly12(18), from Arg15(21) to Tyr35(41), and from Gly40(46) to Ala56(62) in the solution structures were superimposed on the crystal structure.

1996) were published. A comparison between the NMR and X-ray structures of α3(VI) and between the NMR structures of α3(VI) and BPTI is given below.

Comparison of the NMR and X-ray Structures of α3(VI). Superposition of the well-defined part of the average α3(VI) NMR solution structure [the residues from Thr2(8) to Gly12(18), from Arg15(21) to Tyr35(41), and from Gly40(46) to Ala56(62)] with the X-ray crystal structure shows that the overall fold of the structures in the two phases is very similar. The rms deviation of the backbone C^{α}, N, and C′ atoms is 1.23 Å. Of particular interest in a comparison of the NMR and X-ray structures is the position of Trp21(27). In a previous paper (Sørensen et al., 1996), it was concluded that α3(VI) exists in two unequally populated conformations that differ in the orientation of the Trp21(27) ring. This conclusion was based partly on the expectation of similar positions of residue Trp21(27) in the solution and crystal structures of α3(VI). Figure 8 shows that the position of this residue is indeed the same in the solution and the crystal structures. Furthermore, it can be seen that the side chain of Trp21(27) is very well-defined in the NMR structures. This is a consequence of the large number of NOEs (74) involving Trp21(27) (Figure 3a), which is due to the position of Trp21(27) in the hydrophobic groove between the β-sheet and the C-terminal α-helix. Thus, a reorientation of the Trp21(27) ring is in excellent agreement with the observation of two sets of chemical shifts (Sørensen et al., 1996) for protons in both the β-sheet and the C-terminal α-helix, corresponding to the two conformations of Trp21(27). It should be noted that the orientation of the Trp21(27) ring shown in Figure 8 is the orientation in the major conformation. The two conformations are slowly exchanging on the NMR time scale at the temperature employed in the present study (303 K) and therefore give rise to separate signals. In principle, this allows the determination of both conformations. However, as a consequence of its low population (6.4%) no structural restraints could be obtained for the minor conformation.

Comparison of the NMR Structures of α3(VI) and BPTI. The overall folds of BPTI and α3(VI) are, not surprisingly, very similar. However, a superposition of the two structures indicates that the backbone of BPTI has a bend at position 40(46), whereas the backbone of α3(VI) is extended at this position. The extended structure of α3(VI) is the result of the Lys44(50) side chain, which is turned toward Glu42(48) and consequently "pushes" the backbone into the extended conformation at Gly40(46). The position of the Lys44 side chain in the pocket between the residues from Pro8(14) to Asp10(16) and from Gly40(46) to Lys44(50) may explain why only two internal water molecules are found in this pocket (Arnoux et al., 1995; Zweckstetter et al., 1996), whereas three internal water molecules are found in the similar pocket in BPTI (Otting & Wüthrich, 1989).

(c) Comparison of the Dynamics of α3(VI) and Other Kunitz Domains: (1) Comparison with the Complete C5 Domain. The NMR study of the complete C5 domain (Zweckstetter et al., 1996) showed that this 76-residue protein exists in several multiple conformations. Thus, three patches of residues with multiple conformations were identified: (1) one located in the central β-sheet, (2) one located around the Cys14(20)−Cys38(44) disulfide bond, and (3) one located in the vicinity of the Cys5(11)−Cys55(61) disulfide bond. In contrast, no evidence for the presence of multiple conformations in α3(VI) was found in the crystallographic data (Arnoux et al., 1995). As mentioned above, multiple conformations in the central β-sheet have also been observed for α3(VI) in solution (Sørensen et al., 1996). Furthermore, the relaxation data of α3(VI) show that the residues around the Cys14(20)−Cys38(44) disulfide bond are indeed conformationally flexible. It is therefore somewhat surprising that no evidence is found in the NMR data of α3(VI) for conformational heterogeneity of residues in the vicinity of the Cys5(11)−Cys55(61) disulfide bond; that is, only one signal was observed for the protons of Cys5(11), Lys6(12), Leu7(13), and Cys55(61) even at 285 K.

The different dynamical behavior observed for these residues in the complete C5 domain and in α3(VI), respectively, must be a consequence of the additional six N-terminal residues and 12 C-terminal residues in the complete C5 domain, as almost identical experimental conditions were applied in the NMR studies of the two proteins. It seems, therefore, as if the disordered N- and C-terminals in the complete C5 domain trigger or affect a dynamical process leading to the observation (Zweckstetter et al., 1996) of two signals with similar intensities for residues Cys5(11), Lys6(12), Leu7(13), and Cys55(61). It is, however, not clear whether the additional 18 residues of the two termini affect the exchange rate or the equilibrium constant of the dynamical process, or both. The fact that the residues Cys5(11), Lys6(12), Leu7(13), and Cys55(61) are all close to the Cys5(11)−Cys55(61) disulfide bond indicates that an isomerization of this disulfide bond may be the dynamical process responsible for the multiple conformations observed for these residues in the complete C5 domain (Zweckstetter et al., 1996).

(2) Comparison with BPTI. Figure 7h reveals that the values of $J(0)$ are in general larger for the residues from Glu42(48) to Val54(60) [average $J(0)$ value 1.13 ± 0.07 ns/rad^2) than for the β-sheet residues from Ile18(24) to Tyr35(41) (average $J(0)$ value 0.99 ± 0.06 ns/rad^2), which suggests that the dynamics of the two regions are different.

This is somewhat surprising, as tight interactions between the two regions are imposed by the Trp21(27)−Phe45(51) hydrogen bonds and the Cys30(36)−Cys51(57) disulfide bond. To determine whether this sequence-dependent difference in the $J(0)$ values is present in other Kunitz domain proteins, a reduced spectral density mapping was performed for BPTI using the R_1 and R_2 values obtained by Szyperski et al. (1993) [for all residues in the regions from Ile18 to Tyr35 and from Arg42 to Thr54 the $^1H-^{15}N$ NOEs were set to 0.43, which is the maximum theoretical value assuming the overall correlation time (2 ns) found by Szyperski et al. (1993)]. The average $J(0)$ values for the residues from Ile18 to Tyr35 and from Arg42 to Thr54 in BPTI are 1.23 ± 0.08 and 1.28 ± 0.12 ns/rad^2, respectively (residues Ala48 and Arg53 were excluded from the calculation, as no relaxation data are available for these residues). Clearly, in the case of BPTI there is no indication of significantly larger $J(0)$ values for the residues from Arg42 to Thr54 compared to the residues from Ile18 to Tyr35.

The different sequence dependences of $J(0)$ observed for α3(VI) and BPTI suggest that there is a difference in the relative dynamical behavior of the residues from Glu42(48) to Val54(60) compared to the residues from Ile18(24) to Tyr35(41) in the two molecules. The larger $J(0)$ values observed for residues Glu42(48) to Val54(60) in α3(VI) indicate that these residues are involved in a dynamical process on the micro- to millisecond time scale. However, the origin of this dynamical process (or processes) is not clear at present.

Finally, Figure 7h shows that Gly46(52) is characterized by a low $J(0)$ value (0.84), which indicates the presence of high-frequency motions. These high-frequency motions are not on the picosecond time scale, since the $^1H-^{15}N$ NOE of 0.62 observed for Gly46(52) (Table 4 of the supporting information) is very close to the upper limit of 0.64. Most likely the low $J(0)$ value observed for Gly46(52) is a result of a highly flexible glycine in a position where the Kunitz domain structure bends. Thus, there is no evidence for high-frequency motions in the relaxation data for the corresponding residue, Lys46, in BPTI (data not shown).

(d) Relation between the Structure and Dynamics of α3(VI) and Its Trypsin Inhibition. The crystal structure of α3(VI) revealed that the residues Phe17(23) and Thr13(19) have a dramatic influence on the structure of the reactive loop [the residues from Glu11(17) to Ile18(24)] compared to BPTI (Arnoux et al., 1995). Thus, in the crystal structure of α3(VI) the aromatic ring of Phe17(23) packs closely to the hydrophobic pocket consisting of Trp34(40) and Gly36(42). This is also observed in the NMR structures of the complete C5 domain (Zweckstetter et al., 1996) and α3(VI). As already discussed by Arnoux et al. (1995) and Zweckstetter et al. (1996), the presence of a phenylalanine in position 17(23) should have a negative effect on the binding to trypsin through (1) removal of the stabilizing hydrogen bonds in the BPTI/trypsin complex between Arg17 of BPTI and His40 of trypsin and through (2) steric hindrances between the aromatic ring and the residues Gln192 and Gly193 of trypsin.

Also, the presence of a threonine in position 13(19) is expected to have a negative influence on trypsin binding (Arnoux et al., 1995). The X-ray structure of α3(VI) indicates that Thr13(19) causes a rotation of the backbone orientation compared to BPTI, which has a proline in position 13. As a result of this rotation, two new internal hydrogen bonds [Thr13(19) NH···O Cys38(44) and Cys38(44) NH···O Thr13(19)] are established in the crystal structure of α3(VI). It was suggested (Arnoux et al., 1995) that the orientation of the Thr13(19) carbonyl oxygen toward Cys38(44) NH would reduce the inhibitory effect significantly by preventing the formation of a hydrogen bond between Thr13(19) O of α3(VI) and Gly216 NH of trypsin, *i.e.*, the hydrogen bond corresponding to the Gly216 NH···O Pro13 hydrogen bond observed in the BPTI/trypsin complex (Marquart et al., 1983; PDB access code 2PTC). However, trypsin binding studies show that substitution of Thr13(19) with Pro in the [D16(22)A, R15(21)K] mutant of α3(VI) only lowers K_D toward trypsin from 45 to 22 nM and that this reduction is only $^1/_4$ that obtained by substituting Arg15(21) with Lys in the D16(22)A mutant (K_D 335 nM → 45 nM). The NMR data presented here give some hints on the surprisingly small effect of substituting Thr13(19) with Pro. Thus, Thr13(19) is located in a poorly defined region with increased flexibility (*vide supra*). Accordingly, the orientation of Thr13(19) O is not restricted to the one observed in the crystal structure, allowing the Gly216 NH···O Thr13(19) hydrogen bond to be formed in the solution complex between trypsin and α3(VI).

CONCLUSIONS

The present study shows that the global fold of α3(VI) in solution is virtually identical to its global fold in the crystal phase and to the global folds of the well-defined core of the complete C5 domain and of BPTI. The backbone dynamics determined from the ^{15}N relaxation data provides further evidence of a Cys14(20)−Cys38(44) disulfide bond isomerization that occurs on the micro- to millisecond time scale in α3(VI). It is suggested that this dynamical process is the reason for the surprisingly small increase in trypsin affinity obtained by substituting Thr13(19) of α3(VI) with Pro, thus further emphasizing the importance of including protein dynamics in the discussion of protein function.

In contrast to the study of the complete C5 domain by Zweckstetter et al. (1996), the results obtained here show no evidence of multiple conformations of the residues in the vicinity of the Cys5(11)−Cys55(61) disulfide bond of α3(VI). This indicates that the additional six N-terminal residues and the 12 C-terminal residues in the complete C5 domain affect a dynamical process in the well-defined core of the complete C5 domain.

ACKNOWLEDGMENT

We thank Drs. Georg Ole Sørensen, Anne Marie Munk Jørgensen, and Søren M. Kristensen for computational assistance and advice and Mrs. Else Philipp and Mr. Jan Makropoulos for technical assistance. The fermentation, amino acid sequence analysis, mass spectrometry, and amino acid analysis were kindly done by Drs. Jørgen Petersen, Lars Thim, Per F. Nielsen, and Mogens Christensen, respectively. Dr. He Liu and Mr. Marco Tonelli are acknowledged for many valuable suggestions. The BPTI relaxation data were kindly provided by Professor Kurt Wüthrich and Dr. Thomas Szyperski.

SUPPORTING INFORMATION AVAILABLE

Six tables containing the distance and dihedral constraints, the ^{15}N chemical shifts, the ^{15}N R_1, R_2, and NOE values, the

results of the "model-free" analysis, and the spectral density mapping (21 pages). Ordering information is given on any current masthead page.

REFERENCES

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Chapter 8, Clarendon Press, Oxford, England.

Antuch, W., Berndt, K. D., Chávez, M. A., Delfin, J., & Wüthrich, K. (1993) *Eur. J. Biochem. 212*, 675−684.

Arnoux, B., Mérigeau, K., Saludjian, P., Norris, F., Norris, K., Bjørn, S., Olsen, O., Petersen, L., & Ducruix, A. (1995) *J. Mol. Biol. 246*, 609−617.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.

Berndt, K. D., Güntert, P., Orbons, L. P. M., & Wüthrich, K. (1992) *J. Mol. Biol. 227*, 757−775.

Berndt, K. D., Güntert, P., & Wüthrich, K. (1993) *J. Mol. Biol. 234*, 735−750.

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

Bodenhausen, G., Freeman, R., & Morris, G. A. (1976) *J. Magn. Reson. 23*, 171−175.

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

Borgias, B., & James, T. L. (1990) *J. Magn. Reson. 87*, 475−487.

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

Brünger, A. T. (1992) X-PLOR, version 3.1. A system for crystallography and NMR. *X-PLOR Manual*, Yale University, New Haven, CT.

Chu, M.-L., Zhang, R.-Z., Pan, T.-c., Stokes, D., Conway, D., Kuo, H.-J., Glanville, R., Mayer, U., Mann, K., Deutzmann, R., & Timpl, R. (1990) *EMBO J. 9*, 385−393.

Clore, G. M., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 7387−7401.

Deisenhofer, J., & Steigmann, W. (1975) *Acta Crystallogr. Sect. B 31*, 238−250.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. (1979) *Faraday Div. Chem. Soc. Symp. 13*, 49−55.

Dufton, M. J. (1985) *Eur. J. Biochem. 153*, 647−654.

Farrow, N. A., Zhang, O., Forman-Kay, J. D., & Kay, L. E. (1995a) *Biochemistry 34*, 868−878.

Farrow, N. A., Szabo, A., Torchia, D. A., & Kay, L. E. (1995b) *J. Biomol. NMR 6*, 153−162.

Fritz, H., & Wunderer, G. (1983) *Arzneim. Forsch./Drug Res. 33*, 479−494.

Gesmar, H., Nielsen, P. F., & Led, J. J. (1994) *J. Magn. Reson. 103B*, 10−18.

Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1985) *J. Am. Chem. Soc. 107*, 6394−6396.

Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1986) *J. Chem. Phys. 85*, 6837−6852.

Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.

Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.

Heald, S. L., Tilton, R. F., Jr., Hammond, L. J., Lee, A., Bayney, R. M., Kamarck, M. E., Ramabhadran, T. V., Dreyer, R. N., Davis, G., Unterbeck, A., & Tamburini, P. P. (1991) *Biochemistry 30* 10467−10478.

Hynes, T. R., Randal, M., Kennedy, L. A., Eigenbrot, C., & Kossiakoff, A. (1990) *Biochemistry 29*, 10018−10022.

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

Karplus, M., & McCammon, J. A. (1983) *Annu. Rev. Biochem. 52*, 263−300.

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.

Kinosita, K., Jr., Kawato, S., & Ikegami, A. (1977) *Biophys. J. 20*, 289−305.

Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S., & Ito, H. (1988) *Nature 331*, 530−532.

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

Kristensen, S. M., Sørensen, M. D., Gesmar, H., & Led, J. J. (1996) *J. Magn. Reson. 112B*, 193−196.

Lancelin, J.-M., Foray, M.-F., Poncin, M., Hollecker, M., & Marion, D. (1994) *Nat. Struct. Biol. 1*, 246−250.

Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593−626.

Lefèvre, J.-F., Dayie, K. T., Peng, J. W., & Wagner, G. (1996) *Biochemistry 35*, 2674−2686.

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

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

Liu, H., Thomas, P. D., & James, T. L. (1992) *J. Magn. Reson. 98*, 163−175.

Liu, H., Spielmann, P., Ulyanov, N. B., Wemmer, D. E., & James, T. L. (1995) *J. Biomol. NMR 6*, 390−402.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259−281.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.

Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) *Acta Crystallogr. B39*, 480−490.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.

Norris, K., Norris, F., Bjørn, S. E., Diers, I., & Petersen, L. C. (1990) *Biol. Chem. Hoppe-Seyler 371* (Suppl), 37−42.

Otting, G., & Wüthrich, K. (1989) *J. Am. Chem. Soc. 111*, 1871−1875.

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

Peng, J. W., & Wagner, G. (1992a) *J. Magn. Reson. 98*, 308−332.

Peng, J. W., & Wagner, G. (1992b) *Biochemistry 31*, 8571−8586.

Peng, J. W., & Wagner, G. (1994) *Methods Enzymol. 239*, 563−596.

Ponte, P., Gonzalez-DeWhitt, P., Schilling, J., Miller, J., Hsu, D., Greenberg, B., Davis, K., Wallace, W., Lieberburg, I., Fuller, F., & Cordell, B. (1988) *Nature 331*, 525−527.

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

Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250−254.

Richarz, R., Nagayama, K., & Wüthrich, K. (1980) *Biochemistry 19*, 5189−5196.

Shoji, A., Ozaki, T., Fujito, T., Deguchi, K., Ando, S., & Ando, I. (1990) *J. Am. Chem. Soc. 112*, 4693−4697.

Sørensen, M. D., Kristensen, S. M., & Led, J. J. (1995) *J. Magn. Reson. 107B*, 83−87.

Sørensen, M. D., Kristensen, S. M., Bjørn, S., Norris, K., Olsen, O., & Led, J. J. (1996) *J. Biomol. NMR 8*, 391−403.

Stone, M. J., Fairbrother, W. J., Palmer, A. G., Reizer, J., Saier, M. H., Jr., & Wright, P. E. (1992) *Biochemistry 31*, 4394−4406.

Szyperski, T., Luginbühl, P., Otting, G., Güntert, P., & Wüthrich, K. (1993) *J. Biomol. NMR 3*, 151−164.

Tanzi, R. E., McClatchey, A. I., Lamperti, E. D., Villa-Komaroff, L., Gusella, J. F., & Neve, R. L. (1988) *Nature 331*, 528−530.

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

Westaby, S. (1993) *Ann. Thorac. Surg. 55*, 1033−1041.

Wlodawer, A., Walter, J., Huber, H., & Sjölin, L. (1984) *J. Mol. Biol. 180*, 301−329.

Wlodawer, A., Nachman, J., Gilliland, G. L., Gallagher, W., & Woodward, C. (1987) *J. Mol. Biol. 198*, 469−480.

Woessner, D. E. (1962) *J. Chem. Phys. 36*, 1−4.

Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.

Zweckstetter, M., Czisch, M., Mayer, U., Chu, M.-L., Zinth, W., Timpl, R., & Holak, T. A. (1996) *Structure 4*, 195−209.

BI9705570