BIOCHEMISTRY

 Volume 46, Number 7 February 20, 2007

Articles

Structural Approach to a Novel Tandem Repeat DNA-Binding Domain, STPR, by CD and NMR[†,‡]

Shin Saito,[§] Tomoyasu Aizawa,[§] Kyosuke Kawaguchi,[§] Takeshi Yamaki,[§] Daisuke Matsumoto,[§] Masakatsu Kamiya,[§] Yasuhiro Kumaki,[§] Mineyuki Mizuguchi,[||] Sigeharu Takiya,[⊥] Makoto Demura,*[,§] and Keiichi Kawano*[,§]

Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan, Faculty of Pharmaceutical Sciences, University of Toyama, Toyama 930-0194, Japan, and Division of Genome Dynamics, CRIS, Hokkaido University, Sapporo 060-0810, Japan

Received August 29, 2006; Revised Manuscript Received December 10, 2006

ABSTRACT: Fibroin-modulator-binding protein 1 (FMBP-1) is a factor that binds the transcriptional activation elements of the fibroin gene. It has a novel structure, consisting of four tandem repeats (R1−R4) of 23 amino acids each in the C-terminal half. This region is referred to as the STPR (score and three amino acid peptide repeat) domain and acts as a DNA-binding domain in FMBP-1. Interestingly, the homology among the four repeats is remarkably high. Here, we have determined the three-dimensional structures of the four repeats by NMR. All four repeat units have basically the same structure: a short α-helix in the N-terminal half maintained by a salt bridge and an N-capping box. CD studies showed that the full-length STPR domain was 31% helical in solution. This is explained by the connections among the four short helices that were determined separately by NMR. From the thermal-denaturation study, it can be deduced that these four helices in the full-length STPR domain moved flexibly with no interaction among them. However, the specific DNA caused a distinct increase, of up to 76%, in the α-helical content of the full-length STPR domain. This finding suggests that the binding of the full-length STPR domain to specific DNA causes an induced-fit conformational change that increases α-helicity; the poorly structured regions of the protein may form a regular secondary structure. Furthermore, the mutation analysis showed that the four repeats of the STPR domain raise the possibility of interaction with DNA in different ways.

The domestic silkworm *Bombyx mori* is a good candidate to be a host for the production of useful proteins (e.g., human collagen) on an industrial scale because its silk glands synthesize vast amounts of silk protein (*1*). The silk glands of *B. mori* develop during the embryonic stage and differentiate into three compartments: anterior, middle, and posterior silk glands. The posterior silk gland specially generates heavy-chain fibroin (H-fibroin, generally called

[†] This study was partially supported by the Program for the Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN), Japan and a grant for the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

[‡] The atomic coordinates (codes 1VD7, 1VD8, 1VD9, and 1VDA for R1, R2, R3, and R4, respectively) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/). The chemical shifts of R1, R2, R3, and R4 have been submitted to the BMRB under accession nos. 6156, 6157, 6154, and 6155, respectively.

* Corresponding author. Tel: +81-11-706-2770. Fax: +81-11-706-2770. E-mail: kawano@sci.hokudai.ac.jp (K.K.). Tel: +81-11-706-2771. Fax: +81-11-706-2771. E-mail: demura@sci.hokudai.ac.jp (M.D.).

[§] Graduate School of Science, Hokkaido University.

[||] University of Toyama.

[⊥] Division of Genome Dynamics, CRIS, Hokkaido University.

10.1021/bi061780c CCC: $37.00
Published on Web 01/24/2007

fibroin), light-chain fibroin (L-fibroin), and fibrohexamerin, which are assembled into an insoluble fibroin thread (*2*−*7*). However, the middle silk gland secretes several members of the sericin protein family that provide the thread with a sticky coating (*8*−*15*). This gland does not secrete fibroin protein (*7*, *16*). Interestingly, these silk proteins are synthesized and accumulate in the larval feeding stages of silkworm development but not in the moulting stages (*4*, *6*, *11*, *13*, *15*). These tissue- and spatial-specific gene expressions are this organism's most significant characteristics, and they are regulated by several transcriptional factors (*17*−*23*).

Fibroin-modulator-binding proteins 1, 2, and 3 (FMBP-1, -2, and -3)[1] were identified as specific binding factors that bind the AT-rich elements upstream from and in the intron of the fibroin gene (*23*). These elements were reported to enhance transcription from the fibroin core promoter, which covers only the TATA box and initiation region (*24*, *25*). A previous study found that FMBP-2 and FMBP-3 were the same proteins as two well-known transcriptional factors, the Bm fork head (Bm Fkh) and POU-M1, which had winged-helix-type and POU-homeodomain-type DNA-binding domains, respectively (*26*, *27*). However, the molecular and structural details of FMBP-1 remain unknown.

The FMBP-1 gene encodes a 218-amino-acid protein, which is divided into two regions according to its amino acid sequence (*28*). The N-terminal region, residues 1−98, shows a weak similarity to the AP2 domain, which is a DNA-binding domain of transcriptional factors in higher plants, such as APETALA2 (*29*), EREBP-1 (*30*), and RAV1 and RAV2 (*31*). However, the C-terminal half has a novel structure composed of four tandem repeats of 23 amino acids each (Figure 1). This region is referred to as the STPR (score and three amino acid peptide repeat) domain. A previous study indicated that the full-length STPR domain (R1−R4) specifically recognized the AT-rich sequence (ATNTWTNTA), whereas the truncation of even one repeat unit (R1−R3, R2−R4, and so on) almost completely eliminated the domain's binding ability (*28*). Therefore, the STPR domain acts as a DNA-binding domain in FMBP-1. To the best of our knowledge, the homology among the four repeat units of FMBP-1 is much higher (approximately 80%) than that of any transcription factors ever reported. In addition, this repeated structure, composed of 23 amino acids per repeat, is found in a wide spectrum of eukaryotes, including humans, mice, *Drosophila*, and *Caenorhabditis elegans*, according to various database searches (*28*). Thus, the STPR domain appears to be a general motif with important implications for the DNA-binding function, but it is not known exactly how the four repeat units bind DNA.

[1] Abbreviations: FMBP-1, fibroin-modulator-binding protein 1; STPR, score and three amino acid peptide repeat; CD, circular dichroism; NMR, nuclear magnetic resonance; Fmoc, *N*-(9-fluorenyl)-methoxycarbonyl; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid, GST, glutathione *S*-transferase; PBS, phosphate-buffered saline; DQF-COSY, double-quantum-filtered correlated spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; DANTE, delay alternating with mutation for tailored excitation; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt; EMSA, electrophoretic mobility shift assay; HSQC, heteronuclear single-quantum coherence spectroscopy; BMRB, Biological Magnetic Resonance Data Bank; SUMO, small ubiquitin-like modifier; PDB, Protein Data Bank.

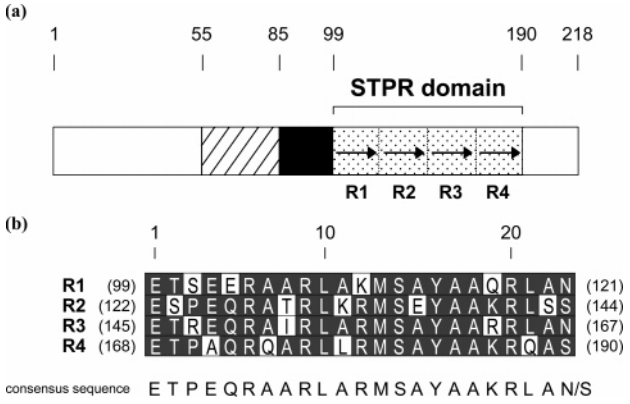

FIGURE 1: Sequences of the FMBP-1 STPR (score and three amino acid peptide repeat) domain. (a) Functional domains of FMBP-1. The STPR domain is a DNA-binding domain of FMBP-1. The box with diagonal lines and the black box show acidic and hyper-basic regions, respectively. (b) Sequence of the FMBP-1 STPR domain and homology among four repeats. The repeats were designated R1, R2, R3, and R4. The amino acids most frequently used among the four repeats are adopted as the amino acids of the consensus sequence. Amino acids are given in single-letter code, and conserved residues are shown as white letters against a black background.

Here, as a first step toward understanding the structural characteristics of the STPR domain, we used CD measurements to collect structural information about the four peptide fragments constituting that domain and the full-length STPR domain. We then determined the 3D structures of the peptides of the four repeat units in aqueous solution by NMR. Furthermore, to examine the effects of DNA on the secondary structure of the full-length STPR domain, the CD spectrum in the presence of the specific DNA was compared with that in the absence of DNA. Finally, to understand the correlation between structure and function, we prepared some mutants of the full-length STPR domain on the basis of the determined structures and evaluated their DNA-binding abilities. These results suggest that the full-length STPR domain partially forms four short α-helices in solution. However, the specific DNA caused a distinct increase, up to 76%, in the α-helical content of the full-length STPR domain, reminiscent of an induced-fit mechanism. This is the first article to describe the structural analysis of a novel DNA-binding domain, the STPR domain.

MATERIALS AND METHODS

Preparation of Peptides. For the CD and NMR experiments, the four 23-residue peptides that constitute the STPR domain (R1: NH_2-ETSEERAARLAKMSAYAAQRLAN-COOH; R2: NH_2-ESPEQRATRLKRMSEYAAKRLSS-COOH; R3: NH_2-ETREQRAIRLARMSAYAARRLAN-COOH; and R4: NH_2-ETPAQRQARLLRMSAYAAKRQAS-COOH) were purchased from Sigma Genosys (Hokkaido, Japan). These peptides were prepared using the solid-phase method with standard Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry. Purification was carried out by reverse-phase HPLC with a TSKgel ODS-80Ts column (TOSOH, Japan) to give final products of >95% purity. Each peptide was eluted with a 5−80% linear gradient of acetonitrile in 0.1% trifluoroacetic acid (TFA), and the eluant was monitored by UV absorbance at 225 nm. The peptides were characterized by mass spectrometry, which gave the expected mass peaks.

Preparation of the Full-Length STPR Domain for CD Spectral Measurements. To measure the CD spectrum, we constructed a protein comprising the full-length STPR domain. The preparation for the plasmid construction of this domain was detailed previously (*28*). The resultant pGEX-6P-1 plasmid, containing an insert that encoded a full-length STPR domain fragment, was introduced into the BL21 cell for overexpression of its product. *Escherichia coli* was cultured at 37 °C, and the expression of the GST-fusion protein was induced by the addition of 1 mM isopropyl-*β*-D-thiogalactopyranoside. After incubation for a further 4 h, the cells were harvested, washed, and resuspended in phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 3.2 mM KH_2PO_4). The cells were broken by sonication and the addition of 10% Triton X-100 to a final concentration of 1%. The supernatant obtained by centrifugation was mixed with glutathione−Sepharose (Amersham Biosciences, Piscataway, NJ) at 4 °C overnight. The Sepharose beads bound to fusion protein were washed with PBS. The sample was digested with PreScission Protease (Amersham Biosciences) for 16 h at 4 °C to remove the GST tag (the addition of 10 units PreScission Protease to 1 mg of protein), and the supernatant obtained by centrifugation was collected. Next, the sample solution was dialyzed against buffer A (50 mM sodium phosphate buffer at pH 7.0) and loaded on an SP sepharose column (Amersham Biosciences) equilibrated with buffer A. The protein was then eluted by a linear gradient from 0 to 1 M NaCl. The peak fractions were pooled, and the buffer was changed to 10 mM sodium phosphate buffer (pH 6.2). The sample was stored at −30 °C.

CD Spectroscopy. CD spectra were measured in each 23-residue peptide (R1, R2, R3, and R4) and in the full-length STPR domain, which were obtained by peptide synthesis and expression in *Escherichia coli*, respectively. CD experiments in the range of 200−250 nm were recorded on a Jasco J-725 spectropolarimeter using standard procedures. Measurements were taken using a 1 mm quartz cell at 25 °C, except for the thermal denaturation experiment. The concentrations of all four peptides and the full-length STPR domain were 40−60 and 15 μM, respectively, in a 10 mM sodium phosphate buffer (pH 6.2). The ellipticity values obtained for the buffer were subtracted from the experimental values obtained for the peptide or protein samples. The secondary structure contents of the peptides and proteins were calculated using the K2d algorithm (available at http://www.embl-heidelberg.de/~andrade/k2d.html) (*32*). CD data obtained at 25 °C were used as input for the program. Thermal denaturation curves were recorded in the temperature range 0−80 °C at 222 nm by a step−scan procedure at a heating rate of 25 °C/h.

Two-Dimensional ^1H NMR Spectroscopy and Structural Calculations. To determine the 3D structure by NMR, four peptides, each corresponding to a repeat of the STPR domain of FMBP-1 (R1, R2, R3, and R4), were targeted. The peptide concentration was 1.5−2.5 mM in the buffer (10 mM sodium phosphate buffer at pH 6.2) dissolved in either 99% D_2O or 90% H_2O/10% D_2O. In hydrogen ion titration analysis with NMR, the pH was downward-titrated by the stepwise addition of HCl up to pH 1.6. The NMR experiments were performed on Alpha 600 and ECA 600 spectrometers (JEOL, Tokyo, Japan). The majority of the NMR spectra were recorded at 20 °C, and some experiments were also recorded at 30 °C to resolve ambiguities. DQF-COSY spectra (*33*) were collected at 20 and 30 °C. TOCSY spectra (*34*) were obtained at 20 and 30 °C with a mixing time of 80 ms. NOESY spectra (*35*) were collected at 20 and 30 °C with mixing times of 150, 200, and 300 ms. The water signal was suppressed by the DANTE pulse (*36*). The size of each spectrum was 2048 complex points in the t2 dimension and 512 complex points in the t1. The proton chemical shifts were referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS). NMR spectra were assigned on an SGI O2 workstation using XEASY (*37*) and NMRPipe software (*38*).

NOE-derived interproton distance restraints were classified into three ranges: strong (1.8−2.7 Å), medium (1.8−3.5 Å), and weak (1.8−5.0 Å). An additional 0.5 Å was added at the upper boundary for NOEs involving methyl protons and methylene protons that were not assigned stereospecifically. Torsion angle restraints on the backbone ϕ angles were derived from $^3J_{HN\alpha}$ coupling constants from the high digital resolution DQF-COSY spectra. Backbone ϕ angles were restrained to $-60 \pm 30°$ for $^3J_{HN\alpha} < 6.0$ Hz. Structure calculations were performed using the distance geometry/simulated annealing method provided by CNS 1.1 software (*39*). The structure was determined using NOE-derived distance restraints and dihedral angle restraints. The force constants for the distance restraints were set to 50 kcal mol^{-1} Å^{-2} throughout all of the calculations. The final round of calculations began with 50 initial structures, and the best 15 structures were selected and analyzed with AQUA, PROCHECK-NMR (*40*), and MOLMOL (*41*). None of them had NOE violations of >0.2 Å.

CD Measurements for DNA Titration. To examine the conformational rearrangement of the STPR domain upon DNA binding, the 16-mer oligo DNA duplex (5′-GAATC-TATGTAAATAC-3′ and 5′-GTATTTACATAGATTC-3′) was used as the domain's recognition element. The underlined bases were essential for the DNA binding of FMBP-1 in a previous study (*28*). The concentrations of the full-length STPR domain and oligo DNA duplex were 15 and 20 μM, respectively, in 10 mM sodium phosphate buffer (pH 6.2). The spectrum of the full-length STPR domain in the presence of DNA was obtained by subtracting the spectrum of DNA from that of the mixture of STPR domain and DNA.

Preparation of Mutants for the Electrophoretic Mobility Shift Assay (EMSA). To evaluate the DNA-binding abilities by EMSA, we prepared some mutants of the full-length STPR domain as GST-fusion proteins. The plasmid of the full-length STPR domain used here had the same construction as that described in the Preparation of the Full-Length STPR Domain for CD Spectral Measurements section. We carried out site-directed mutagenesis using a QuikChange Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA). The processes for expression and purification were nearly the same as those described above, except that there was no cleavage of the GST tag. After the fusion protein was bound to the glutathione−Sepharose beads, the beads were washed with 50 mM Tris-HCl at pH 8.0. Finally, the bound protein was eluted with the elusion buffer (50 mM Tris-HCl at pH 9.5 and 5 mM glutathione), dialyzed against the NP40 buffer (50 mM Tris-HCl at pH 7.9, 12.5 mM $MgCl_2$, 100 mM NaCl,

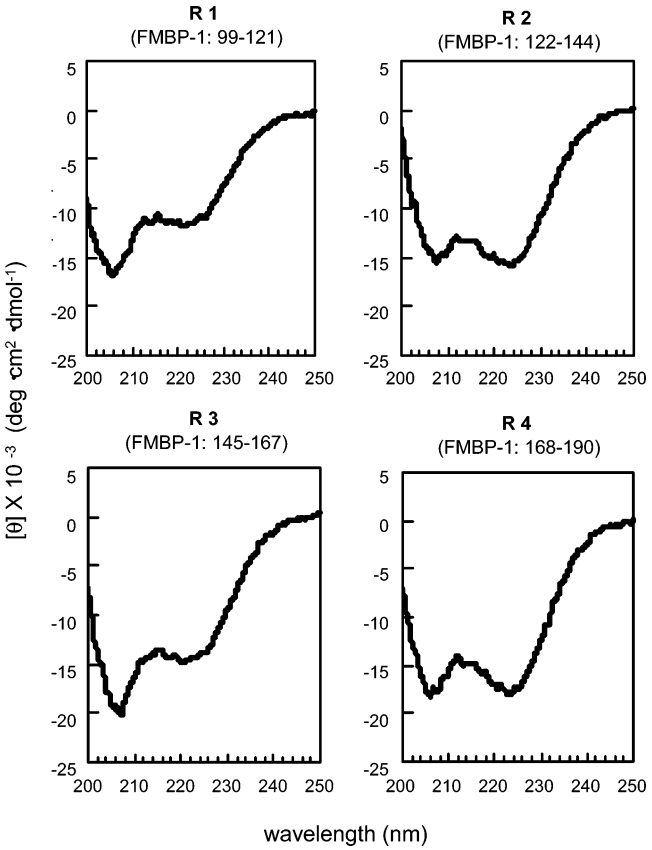

FIGURE 2: CD spectra of the individual repeat fragments of the STPR domain. The α-helix content of these fragments (R1, R2, R3, and R4) estimated by the K2D algorithm method (*32*) is 28, 37, 31, and 37%, respectively. The peptide concentration is 40−60 μM in 10 mM sodium phosphate buffer (pH 6.2). The vertical scale is normalized by the mole concentrations.

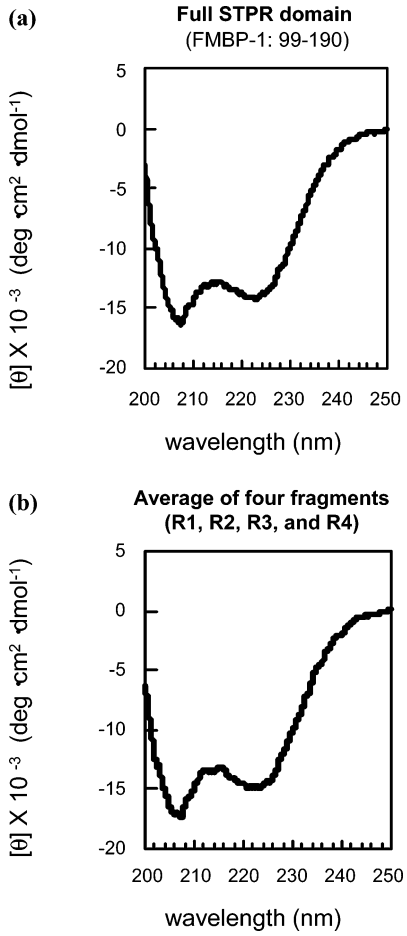

FIGURE 3: Comparison of the CD spectra between the full-length STPR domain and the repeat fragments. (a) CD spectrum of the full-length STPR domain (R1R2R3R4). The domain was expressed as a GST-fusion protein and then cleaved with site-specific protease. The α-helix content estimated by the K2D algorithm method (*32*) is 31%. The protein concentration is 15 μM in 10 mM sodium phosphate buffer (pH 6.2). The vertical scale is normalized by the mole concentrations. (b) Average CD spectrum of each repeat unit (R1, R2, R3, and R4). The mean of the ellipticity values was calculated from the four values obtained by measuring each fragment. The helical content was estimated to be 31%.

0.1 mM EDTA, 20% glycerol, 0.1% NP40, and 1 mM DTT), and stored at −80 °C.

EMSA. EMSAs were carried out as described by Takiya et al. (*28*). The oligonucleotide +290 probe (5′-AATTGATGAATCTATGTAAATACTGGGCAGACAATT-3′ and 5′-AATTGTCTGCCCAGTATTTACATAGATTCATCAATT-3′) was used as the STPR domain's recognition element. This probe is comprises the AT-rich sequences in the intron of the fibroin gene. The +290 probe (0.4−0.5 ng) was incubated with protein in 10 μL (20 ng) of reaction mixture (10 mM Tris-HCl at pH 7.9, 7.5 mM $MgCl_2$, 60 mM NaCl, 1 μg POLY(dI-dC), 20% glycerol, and 0.6% NP40) for 20 min on ice. The protein−probe complexes were separated on 7% polyacrylamide gels.

RESULTS AND DISCUSSION

Conformational Characteristics of Homologous Repeats. The four repeat units of the STPR domain have highly homologous sequences (Figure 1b). In addition, the full-length STPR domain can specifically recognize DNA (ATNTWTNTA) (*28*). In spite of these interesting characteristics, there have been no reports on this domain's tertiary structures. Therefore, we first divided the STPR domain into the four characteristic peptide fragments (R1, R2, R3, and R4) and collected information on their structures.

Figure 2 shows the CD spectra of the individual peptide fragments. As one can see, each peptide shows typical α-helical characteristics. In addition, the α-helical content of these fragments, as estimated by the K2d algorithm method (*32*), is 28−37%, and the helical contents among the individual peptides are similar. Interestingly, the helical orientation in R1 is more similar to that in R3 on the basis of the relative minima at 208 and 222 nm, whereas that in R2 appears more similar to the helical orientation in R4.

Next, to investigate the full-length STPR domain, we measured its CD spectrum. Figure 3a shows the CD spectrum of the full-length STPR domain. One can see that the domain also shows typical α-helical characteristics (helical content: 31%). Here, we calculated the mean ellipticity values among the four fragments (Figure 3b). Interestingly, the average CD spectrum (helical content: 31%) was similar to the CD spectrum of the full-length STPR domain. From these results, it may be deduced that each repeat unit by itself can form the original secondary structure, even though the full-length STPR domain is divided into four fragments.

Here, to test whether or not the four fragments act as independent structural units in the full-length STPR domain, we studied the melting behavior of the fragments and that of the full-length STPR domain. Figure 4 shows the

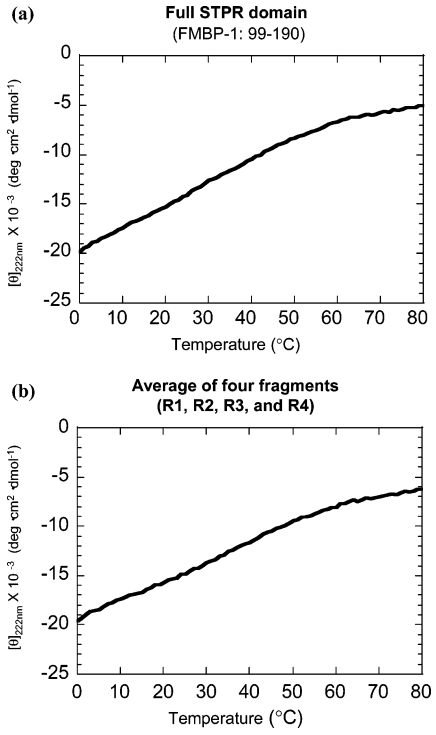

FIGURE 4: Unfolding profile of the full-length STPR domain and the repeat fragments. The ellipticity at 222 nm was monitored in the temperature range 0−80 °C by a step-scan procedure at a heating rate of 25 °C/h. (a) Thermal denaturation curve of the full-length STPR domain (R1R2R3R4). The protein concentration is 15 μM in 10 mM sodium phosphate buffer (pH 6.2). The vertical scale is normalized by the mole concentrations. (b) Average thermal denaturation curve of each repeat unit (R1, R2, R3, and R4). The peptide concentration is 40−60 μM in 10 mM sodium phosphate buffer (pH 6.2). The mean of the ellipticities at 222 nm was calculated from the four values obtained by measuring each fragment.

temperature dependence of the CD spectral ellipticities at 222 nm, which are influenced by helical content. There was no difference in thermal denaturation curves between the full-length domain (Figure 4a) and the average of the four fragments (Figure 4b). These data suggest that in the full-length STPR domain, each repeat most likely acts as an independent structural unit with no interaction among the four repeat units.

NMR Analyses of the Four STPR Repeats. In an attempt to understand the 3D structure of the full-length STPR domain, the ^{15}N-labeled recombinant protein, which includes the full-length STPR domain (R1−R4), was prepared in *E. coli*. Figure 5 shows the ^{1}H-^{15}N-HSQC NMR spectrum of the full-length STPR domain. We observed several downfield-shifted proton resonances at approximately 9.0 ppm. These data imply that the domain is partially structured, but the assignment and structure determination of the full-length domain were difficult because of the high overlap. From the CD measurements, the domain had the same helical content and the same melting behavior as the average of the four fragments (Figures 3 and 4). These data imply that the 3D structure of the full-length STPR domain is close to the architecture obtained by connecting the tertiary structures of the four peptide fragments. Thus, we prepared the four repeat fragments by peptide synthesis and determined their tertiary structures using NMR measurements.

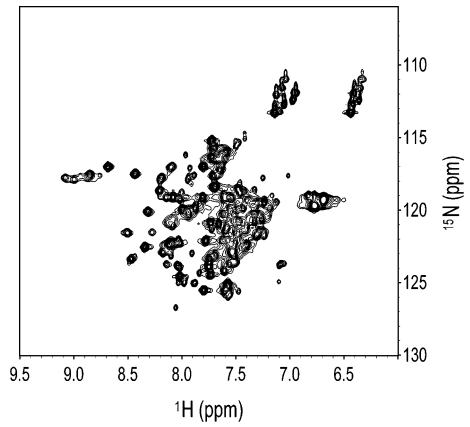

FIGURE 5: ^{1}H-^{15}N HSQC spectrum of the full-length STPR domain recorded at 10 °C. The protein concentration is 0.74 mM in 10 mM sodium phosphate buffer (pH 6.2) containing 60 mM NaCl.

The proton NMR signals were identified according to the standard procedure for the sequential assignment of a proton (*42*). The residues were sequentially assigned by first identifying spin systems through a combination of DQF-COSY spectra and TOCSY spectra. This procedure was followed by assigning the identified spin systems to particular residues in the peptide by the observation of sequential NH−NH (d_{NN}), C_αH−NH ($d_{\alpha N}$), and C_βH−NH ($d_{\beta N}$) NOEs. The fingerprint regions of the COSY spectra are shown in Figure 6. The NMR signals of the backbone protons without prolines and N-terminus glutamic acids were observed in each COSY spectrum. The proton NMR signals were almost completely assigned to four repeat units. All assigned chemical shifts were registered with the Biological Magnetic Resonance Data Bank (BMRB accession numbers: 6156, 6157, 6154, and 6155).

Elements of the secondary structures were deduced from the pattern of sequential and medium-range NOE connectivities. The sequential d_{NN} NOEs with strong intensities were observed in Ser-3−Met-13 of R1, Glu-4−Ala-7, and Arg-9−Arg-12 of R2, Glu-4−Leu-10 of R3, and Ala-4−Arg-12 of R4. The extensive unambiguous medium-range NOEs αN-(i,i+3), characteristic of α-helical conformations, were also detected at some N-terminal residues (e.g., Ser-3−Leu-10 in R1; Pro-3, Gln-5, and Thr-8 in R2; Arg-3−Leu-10 in R3; Ala-4−Gln-7, Arg-9−Leu-10, and Met-13 in R4). These results strongly suggest that the residues in the N-terminal half are involved in the formation of the α-helix.

The 3D structure of each repeat unit was calculated following the molecular dynamic protocol described in Materials and Methods. A final set of 125 NOE-derived distance constraints and 3 angle constraints in R1, 136 NOE-derived distance constraints and 2 angle constraints in R2, 182 NOE-derived distance constraints and 2 angle constraints in R3, and 133 NOE-derived distance constraints and no angle constraint in R4 were used as inputs for the distance geometry and simulated annealing calculations using the CNS 1.1 software. On the basis of low residual distance violations, 15 structures were selected. Figure 7 displays the 15 final structures superimposed for best fit over the backbone atoms of residues 4−9 on the mean coordinate structure as well as the ribbon diagrams with the lowest energy structures among the 15. All repeat units had a short α-helix in the N-terminal half (residues 3−10), but the C-terminal region was disor-

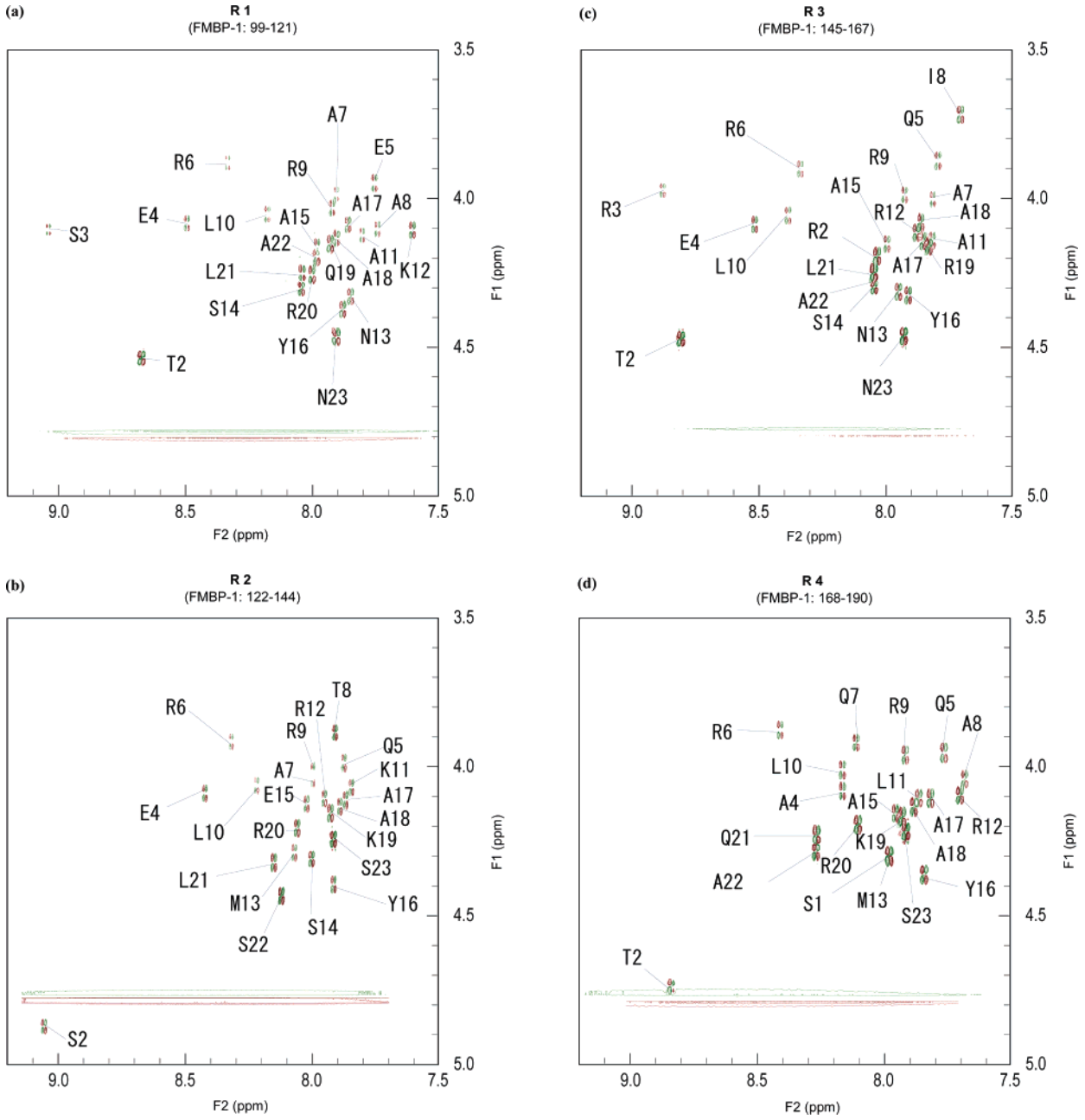

FIGURE 6: Fingerprint region from a COSY spectrum of the four repeat units of the FMBP-1 STPR domain: R1 (a), R2 (b), R3 (c), and R4 (d). Each spectrum was recorded at pH 6.2 in buffer (10 mM sodium phosphate buffer at pH 6.2) at 20 °C. Residue 3 is proline in R2 and R4. Amino acids are given in single-letter code.

dered. The regions corresponding to the helical structure (residues 3−10) were well-defined with pairwise rmsd's of 0.44, 0.32, 0.21, and 0.56 Å for the backbone atoms in R1, R2, R3, and R4, respectively. The CD spectrum of R1 was more similar to that of R3, whereas the CD spectrum of R2 appeared more similar to that of R4 (Figure 2). However, there was no evidence for these similarities in the NOE patterns or structural information. The difference among these CD spectra might be due to the primary sequence. The helical content estimated by the NMR study was 35% (of 23 residues, 8 residues were involved in the helical structure). These results were, thus, mostly consistent with the structural studies by CD measurements.

Determination of Helix-Stabilizing Factors in the STPR Repeat. From our results, it may be deduced that each repeat unit by itself can form an original secondary structure even if the full-length STPR domain is divided into four fragments. Thus far, several interactions have been reported to contribute to α-helix stability in short peptides (*43*). Here, we explored the possibility of several helix-stabilizing factors.

We first noted the well-conserved arginine residues at position 9 (denoted as Arg-9, the ninth residue from the N-terminus in each repeat fragment). In NMR studies, all of these arginines showed an abnormal downfield shift for an N_ϵ proton of 1−1.5 ppm at pH 6.2 compared with those of the other arginines. It is generally accepted that an abnormal NMR chemical shift is brought about by various interactions. To understand the cause of this abnormal shift, a pH titration experiment was performed. Figure 8 shows the pH dependences of chemical shifts for the four Arg N_ϵ protons of peptide R4 (Arg-6, -9, -12, and -20). Interestingly, the $N_\epsilon H$ proton resonance of Arg-9 shifted largely upfield

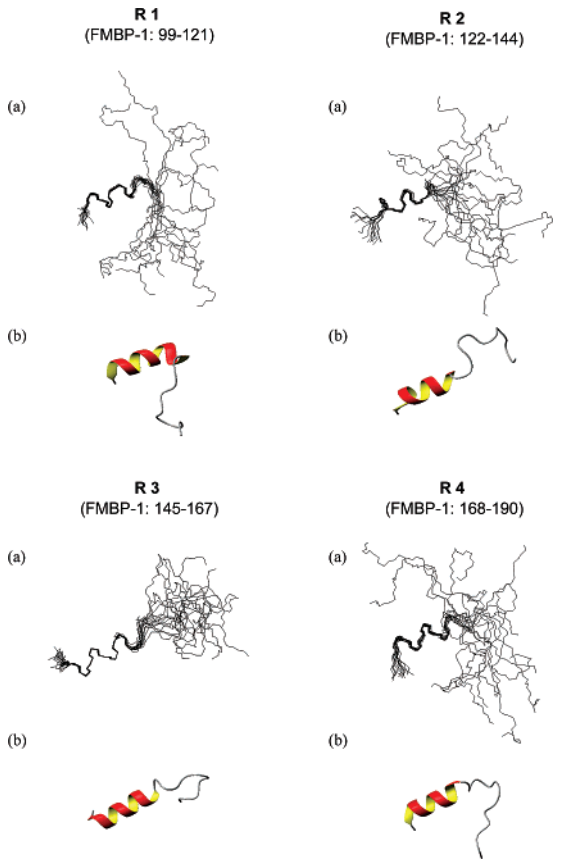

FIGURE 7: Three-dimensional structures of R1, R2, R3, and R4 of the FMBP-1 STPR domain (pdb code: 1VD7, 1VD8, 1VD9, and 1VDA, respectively). (a) View of the ensemble of 15 NMR structures of each repeat unit in the FMBP-1 STPR domain superimposed for best fit over the backbone atoms of residues 4−9 in the mean coordinate structure. (b) Ribbon diagram of each repeat unit with the lowest energy structure among the 15 final structures. These diagrams were generated using the program MOLMOL (*41*).

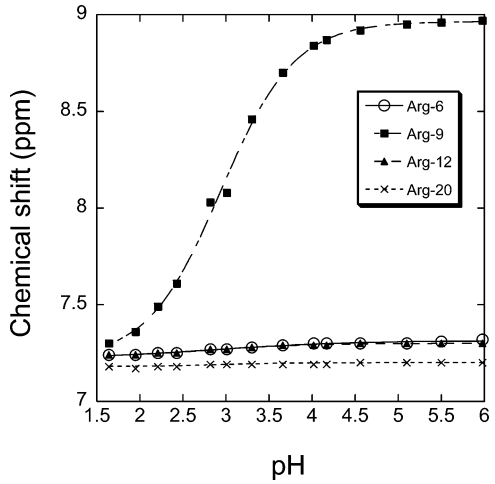

FIGURE 8: pH dependence of the chemical shift at 20 °C, in 60 mM NaCl, for the N_ϵ protons of Arg-6 (○), Arg-9 (■), Arg-12 (▲), and Arg-20 (×), which are included in R4. Titration data were fitted using the modified Henderson−Hasselbalch equation (*44*).

with decreasing pH in a sigmoidal fashion. This result suggests that the interaction that caused the abnormal downfield shifts in the side chain of Arg-9 disappeared at low pH. These titration data were fit using the modified Henderson−Hasselbalch equation: $\delta_{obs} = \delta_u + \Delta\delta/(1 + 10^{n(pK-pH)})$, where δ_u is the shift of the unprotonated arginine resonance, $\Delta\delta$ is the titration shift resulting from the protonation of a group with the p*K* value indicated in the equation, and *n* is the Hill coefficient (*44*). The calculated p*K* value of the sigmoid for Arg-9 $N_\epsilon H$ was 2.9 ($n = 1.0$). Thus, Arg-9 interacts with the side chain of an acidic amino acid. Here, we noted the side chain of Glu-1 as a candidate for the interaction partner of Arg-9 because Glu-1 is the only acidic amino acid in R4. The other repeat units (R1, R2, and R3) of the FMBP-1 STPR domain also have glutamic acid residues at position 1 (Figure 1b). This interaction is thought to be the $H-N_\epsilon\cdots O{=}C_\delta$ hydrogen-bonding interaction in the salt bridge between Glu-1 and Arg-9 (*45*). This kind of salt bridge interaction was reported in other proteins (*46*−*48*). For example, the C-peptide (residues 1−13) from RNase A (*47*) has a short helix stabilized by the salt bridge; it is formed by two residues, Glu-2 and Arg-10, with the same spacing as in STPR domain repeat units (Glu-1 and Arg-9).

Besides the salt bridge, another possible stabilizing factor for the helical structure could be the N-capping box, which is known as an initiator and stabilizer of α-helices in several peptides and proteins (*49*). The residues in the N-capping box are designated as

$$\ldots\ldots-N''-N'-N_{cap}-N_1-N_2-N_3-\ldots\ldots$$

where N_1 refers to the first helical residue, N_{cap} is a bridging residue belonging to both the loop and helix, and the primed residues refer to the loop regions flanking the helix. It is commonly assumed that two residues at the N_{cap} and N_3 positions are involved in the N-capping box, with Ser and Thr commonly found at the N_{cap} position and with Glu and Gln commonly found at N_3.

Here, we found the presence of the N-capping box by the characteristic ^{1}H NOE connectivities. As described above, the NMR study indicated that residues 3−10 were involved in the helical formation in each repeat unit so that Thr/Ser-2 and Gln/Glu-5 correspond to the N_{cap} and N_3 residues, respectively. Table 2 summarizes the NOE connectivities between the N_{cap} residue and the N_3 or N_2 residue. These NOE connectivities were compared with those of ShK toxin, which is an example of a peptide involved in the N-capping box (*50*). As shown in Table 2, we observed the characteristic NOE connectivities from Thr/Ser-2 protons to several protons, as observed in the ShK toxin. These results suggest that the N-capping box is likely to act as an initiator and stabilizer of the helix. From a comparison with the ShK toxin, it can be deduced that the N-capping box of the STPR repeats is also stabilized by the two hydrogen bonds from the backbone NH of the N_3 to the backbone oxygen and side-chain oxygen of N_{cap} (Figure 9).

From the COSY spectra of the STPR peptides (Figure 6), the N-terminal residues at the N_{cap} and N_1 positions were shifted so far downfield (N_{cap}, Thr/Ser-2; N_1, Ser/Arg-3). Particularly, the N_1 amide protons (Ser-3 in R1; Arg-3 in R3), which were invisible in R2 and R4 because of proline, exhibited larger downfield shifts. Calculations using NMR-derived structures indicated that C═O and C−N anisotropies made a large contribution to these chemical shifts (calculated with the SHIFTCALC program for chemical shift calculations available at http://www.nmr.group.shef.ac.uk/NMR/mainpage.html). Thus, the downfield shifts might be because of the capping interactions. Indeed, in the previous report

Table 1: Structural Statistics for the 15 Best Structures[a]

	R1 (FMBP-1:99-121)	R2 (FMBP-1:122-144)	R3 (FMBP-1:145-167)	R4 (FMBP-1:168-190)
structural restraints				
intraresidue	66	94	97	67
sequential	45	34	59	42
medium range	14	8	26	24
long range	0	0	0	0
Φ range	3	2	2	0
average potential energies (kcal mol^{-1})				
E_{total}	14.30 ± 0.27	21.15 ± 0.61	20.11 ± 0.75	23.54 ± 1.66
E_{bonds}	0.08 ± 0.01	0.53 ± 0.05	0.37 ± 0.06	0.76 ± 0.14
E_{angles}	13.01 ± 0.06	14.93 ± 0.18	14.43 ± 0.35	14.66 ± 0.32
E_{impr}	0.18 ± 0.01	0.21 ± 0.02	0.27 ± 0.06	0.23 ± 0.02
E_{VDW}	0.84 ± 0.16	1.97 ± 0.38	2.85 ± 0.82	2.17 ± 0.70
E_{NOE}	0.20 ± 0.08	3.50 ± 0.31	2.20 ± 0.64	5.73 ± 1.13
E_{cdih}	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
NOE violation >0.2 (Å)	0	0	0	0
dihedral angle violation >5(°)	0	0	0	0
Ramachandran plot (%)				
most favored region	65.1	54.3	55.9	57.0
additional allowed region	28.3	37.0	34.9	35.0
generously allowed region	4.1	6.0	7.9	5.7
disallowed region	2.5	2.7	1.3	2.3
pairwise rmsd (Å) (residues 3-10)				
backbone	0.44 ± 0.14	0.32 ± 0.16	0.21 ± 0.08	0.56 ± 0.22
all heavy atoms	1.74 ± 0.31	1.62 ± 0.28	1.64 ± 0.35	1.49 ± 0.21

[a] The data represent the structural statistics for the 15 energy-minimized structures obtained through the refinement protocol in CNS. E_{impr}, E_{VDW}, and E_{cdih} are improper dihedral angle energy, intramolecular van der Waals energy, and dihedral angle restraint energy, respectively.

Table 2: Summary of N-Cap NOE Connectivities Observed between the N_{cap} Residue and the N_2 or N_3 Residue

			FMBP-1 STPR domain				ShK[50],[c] (reference peptide)
			R1[a]	R2[a,b]	R3[a]	R4[a]	
N_{cap} residue	→ N_3 residue	N_2 residue					
NH	→ C_βH		+	+	+	+	+
NH	→ C_γH		+	+	+	+	+
C_βH	→	NH	+	+	+	+	+
C_γH	→	NH	+	*	+	+	+
C_βH	→ NH					+	+
C_γH	→ NH		+	*	+	+	+

[a] The plus signs indicate the NOE connectivities observed in both 150 and 300 ms mixing time NOESY spectra. [b] The asterisks denote that C_γH is absent. The N_{cap} residue is serine in R2. [c] The NOE connectivities were reported for the N-capping box by Lanigan et al. (*50*). The plus signs indicate the NOE connectivities observed in a 200 ms mixing time NOESY spectrum.

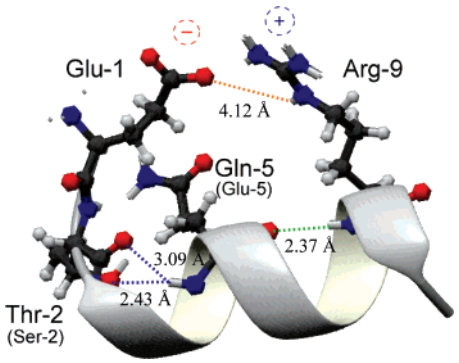

FIGURE 9: Schematic view of the intramolecular interactions of the STPR repeat. The intramolecular interactions that are predicted to be observed in R3 are shown as an example of the four STPR repeat units. The orange line represents the salt bridge, and the blue lines represent the hydrogen bonds of the N-capping box. The green line represents the hydrogen bond commonly observed in alpha helices.

on the N-capping box, two residues at the N_{cap} and N_1 positions showed a downfield shift for an amide proton, as the STPR peptides (*51*). In ShK toxin, the N_1 amide proton is also shifted downfield, although the N_{cap} amide proton has an upfield shift because of the ring current effect from the N_2 residue phenylalanine (*52*).

Interestingly, almost all of the repeat units from other eukaryotes contain four amino acids (Glu-1 and Arg-9, related to the salt bridge; and Thr/Ser-2 and Gln/Glu-5, related to the N-capping box) with the same spacing as that of the repeat unit of the FMBP-1 STPR domain (*28*). These observations support the hypothesis that the N-terminal half of each repeat in the STPR domain might include a well-defined helix and that each helix is stabilized by the salt bridge between Glu-1 and Arg-9 as well as by the N-capping box encompassing Thr/Ser-2 and Gln/Glu-5.

Conformational Rearrangement of the STPR Domain upon Binding DNA. The tertiary structural conformation of the full-length STPR domain was characterized from CD and NMR measurements. In solution, each repeat unit had a short α-helix in the N-terminal half (residues 3−10 in each unit), and the C-terminal region was flexible. In the full-length STPR domain, a total of four helices were connected with these flexible regions. As a result, each helix in the full-length STPR domain was predicted to flexibly move with no interaction among these helices.

However, there is no information about the conformation in the DNA-bound state. Here, to examine DNA's effects on the secondary structure of the full-length STPR domain, the CD spectrum was analyzed in the presence of specific DNA (Figure 10). Surprisingly, the presence of the specific DNA distinctly increased, by 76%, the α-helical content of the full-length STPR domain, whereas the α-helical content was 31% in the DNA-free state. This observation suggests a conformational rearrangement of the STPR domain upon

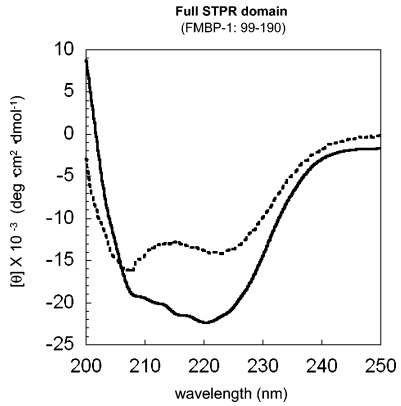

FIGURE 10: CD spectra of the full-length STPR domain in the absence (···) and presence of (—) the specific DNA duplex (5′-GAATCTATGTAAATAC-3′ and 5′-GTATTTACATAGATTC-3′). The domain was expressed as a GST-fusion protein and then was cleaved with site-specific protease. The concentrations of the full-length STPR domain and the oligo DNA were 15 and 20 μM, respectively, in 10 mM sodium phosphate buffer (pH 6.2). The spectrum of the full-length STPR domain in the presence of DNA was obtained by subtracting the spectrum of DNA from that of the mixture of STPR domain and DNA.

binding the AT-rich element of the fibroin gene, reminiscent of an induced-fit mechanism. This type of induced fit appears to be a common occurrence in the initiation of transcription, where poorly structured regions of the protein may form a regular secondary structure in the presence of DNA, enhancing protein−protein interactions and promoting an ordered nucleoprotein structure (*53*−*55*).

Likewise, the specific DNA was added to each repeat peptide fragment, but no significant effect was observed (data not shown). These results are in accord with our previous study, that is, the full-length STPR domain (R1−R4) specifically recognized the specific DNA sequence (ATNTWTNTA), whereas the truncation of even one repeat unit (R1−R3, R2−R4, and so on) almost completely eliminated its binding ability (*28*). It is thought that all four repeats are needed for the induced-fit mechanism in the case of the STPR domain because the four repeats might cooperatively behave in the presence of DNA.

Effects of Mutations on Binding to Specific DNA. In the DNA-free state, the present results suggest that the full-length STPR domain partially forms four short helices in solution. However, upon specific DNA binding, the full-length STPR domain showed a distinct increase in the α-helical content; the poorly structured regions of the protein are thought to form a regular secondary structure. Now we were interested in knowing how the four highly homologous repeats participate in DNA recognition. However, the NMR measurement of the full-length STPR domain in the presence of the DNA duplex failed to obtain a good ^{1}H-^{15}N-HSQC spectrum. Therefore, as a first step toward understanding the correlation between structure and function, we prepared four mutants of the full-length STPR domain on the basis of the determined structures. We constructed these mutants in order to destroy one salt bridge, as shown in Figure 11a ; Glu-1 in each repeat unit was replaced by Gln (denoted as M1, M2, M3, and M4). The DNA-binding ability of each mutant was compared with that of the wild type by the electrophoretic mobility shift assay (EMSA). Interestingly, the influence on DNA binding differed among the four mutants.

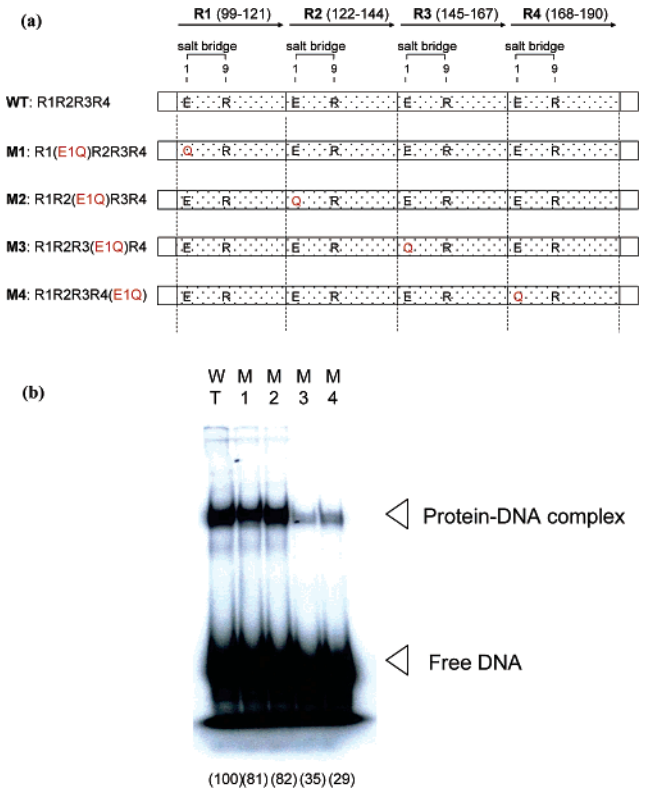

FIGURE 11: DNA binding activities of the wild type (WT) and four mutants (M1, M2, M3, and M4) of the STPR domain. (a) Sequence of the full STPR domain (R1R2R3R4) and its mutants (M1, M2, M3, and M4) used in EMSA analysis. The proteins shown in (a) were expressed as GST-fusion proteins. (b) Autoradiograph of polyacrylamide gels showing the EMSA analyses of STPR domain binding to DNA. The oligonucleotide +290 probe (5′-AATTGATGAATCTATGTAAATACTGGGCAGACAATT-3′ and 5′-AATTGTCTGCCCAGTATTTACATAGATTCATCAATT-3′) was used as the STPR domain's recognition element. The values in parentheses represent the relative intensities of the bands of the protein−DNA complex obtained by densitometric tracing.

The M1 and M2 mutants, in which glutamine residues replaced glutamic acid residues at position 1 of R1 or R2, possessed almost the same (in excess of 80%) DNA-binding abilities as the wild type (Figure 11b). However, the binding abilities of the M3 and M4 mutants were markedly reduced. These results suggest that each repeat has a different role in the interaction with DNA, even though the four repeat units have highly homologous sequences and similar helical structures.

DNA-binding domains consisting of four helices have been seen in several proteins (*56*−*59*). For example, PIAS1, which functions as an E3-type SUMO, has a four-helix bundle with an up−down−extended loop−down−up topology (*56*). In this example, the contributions to DNA binding differ among the four helices: the regions included in the α2- and α3-helices are contact sites for DNA. In the already reported four-helix bundle proteins, the homology among the four helices is not very high; therefore, it is not surprising that the DNA-binding abilities differ among the four helices. In the STPR domain, however, the roles of the four repeat units in DNA-binding abilities apparently differed despite their highly homologous sequences and the similarity of their helical structures. The orientations of the α-helices might be important for specific DNA binding.

According to the results of the NMR analysis, all four repeat units have basically the same structure: a short α-helix in the N-terminal half maintained by a salt bridge and an N-capping box. In the DNA-free state, this short helix derived from each unit was connected with flexible loops. As a result, each helix in the full-length STPR domain was predicted to flexibly move, with no interaction among these helices. Furthermore, the binding of the STPR domain to a specific DNA target sequence caused a conformational change that increased α-helicity. Because there have been no reports about DNA-binding domains composed of highly homologous repeats such as the STPR domain, the determination of the solution structure of a specific DNA complex will provide information about the novel DNA-binding motif. In addition, this repeated structure, composed of 23 amino acids per repeat, is found in a wide spectrum of eukaryotes, including humans. The 3D structure observed here may help to elucidate how this group of proteins binds to DNA and regulates transcription.

ACKNOWLEDGMENT

We thank the staff of the High-Resolution NMR Laboratory, Graduate School of Science, Hokkaido University, for the NMR measurements.

REFERENCES

1. Tomita, M., Munetsuna, H., Sato, T., Adachi, T., Hino, R., Hayashi, M., Shimizu, K., Nakamura, N., Tamura, T., and Yoshizato, K. (2003) Transgenic silkworms produce recombinant human type III procollagen in cocoons, *Nat. Biotechnol. 21*, 52−56.
2. Suzuki, Y., and Brown, D. D. (1972) Isolation and identification of the messenger RNA for silk fibroin from Bombyx mori, *J. Mol. Biol. 63*, 409−429.
3. Suzuki, Y., and Suzuki, E. (1974) Quantitative measurements of fibroin messenger RNA synthesis in the posterior silk gland of normal and mutant Bombyx mori, *J. Mol. Biol. 88*, 393−407.
4. Suzuki, Y., and Giza, P. E. (1976) Accentuated expression of silk fibroin genes in vivo and in vitro, *J. Mol. Biol. 107*, 183−206.
5. Kimura, K., Oyama, F., Ueda, H., Mizuno, S., and Shimura, K. (1985) Molecular cloning of the fibroin light chain complementary DNA and its use in the study of the expression of the light chain gene in the posterior silk gland of Bombyx mori, *Experientia 41*, 1167−1171.
6. Couble, P., Moine, A., Garel, A., and Prudhomme, J. C. (1983) Developmental variations of a nonfibroin mRNA of Bombyx mori silkgland, encoding for a low-molecular-weight silk protein, *Dev. Biol.* (San Diego, CA, U.S.) *97*, 398−407.
7. Maekawa, H., and Suzuki, Y. (1980) Repeated turn-off and turn-on of fibroin gene transcription during silk gland development of Bombyx mori, *Dev. Biol.* (San Diego, CA, U.S.) *78*, 394−406.
8. Sprague, K. U. (1975) The Bombyx mori silk proteins: characterization of large polypeptides, *Biochemistry 14*, 925−931.
9. Gamo, T., Inokuchi, T., and Lauferc, H. (1977) Polypeptides of fibroin and sericin secreted from the different sections of the silk gland in *Bombyx mori*, *Insect Biochem. 7*, 285−293.
10. Okamoto, H., Ishikawa, E., and Suzuki, Y. (1982) Structural analysis of sericin genes. Homologies with fibroin gene in the 5′ flanking nucleotide sequences, *J. Biol. Chem. 257*, 15192−15199.
11. Ishikawa, E., and Suzuki, Y. (1985) Tissue- and stage-specific expression of sericin genes in the middle silk gland of Bombyx mori, *Dev., Growth Differ. 27*, 73−82.
12. Couble, P., Michaille, J. J., Garel, A., Couble, M. L., and Prudhomme, J. C. (1987) Developmental switches of sericin mRNA splicing in individual cells of Bombyx mori silkgland, *Dev. Biol.* (San Diego, CA, U.S.) *124*, 431−440.
13. Obara, T., and Suzuki, Y. (1988) Temporal and spatial control of silk gene transcription analyzed by nuclear run-on assays, *Dev. Biol.* (San Diego, CA, U.S.) *127*, 384−391.
14. Michaille, J. J., Garel, A., and Prudhomme, J. C. (1990) Cloning and characterization of the highly polymorphic Ser2 gene of Bombyx mori, *Gene 86*, 177−184.
15. Matsunami, K., Kokubo, H., Ohno, K., and Suzuki, Y. (1998) Expression pattern analysis of SGF-3/POU-M1 in relation to sericin-1 gene expression in the silk gland, *Dev., Growth Differ. 40*, 591−597.
16. Tsuda, M., Ohshima, Y., and Suzuki, Y. (1979) Assumed initiation site of fibroin gene transcription, *Proc. Natl. Acad. Sci. U.S.A. 76*, 4872−4876.
17. Matsuno, K., Hui, C. C., Takiya, S., Suzuki, T., Ueno, K., and Suzuki, Y. (1989) Transcription signals and protein binding sites for sericin gene transcription in vitro, *J. Biol. Chem. 264*, 18707−18713.
18. Hui, C. C., Matsuno, K., and Suzuki, Y. (1990) Fibroin gene promoter contains a cluster of homeodomain binding sites that interact with three silk gland factors, *J. Mol. Biol. 213*, 651−670.
19. Suzuki, T., Matsuno, K., Takiya, S., Ohno, K., Ueno, K., and Suzuki, Y. (1991) Purification and characterization of an enhancer-binding protein of the fibroin gene. I. Complete purification of fibroin factor 1, *J. Biol. Chem. 266*, 16935−16941.
20. Suzuki, T., Takiya, S., Matsuno, K., Ohno, K., Ueno, K., and Suzuki, Y. (1991) Purification and characterization of an enhancer-binding protein of the fibroin gene. II. Functional analyses of fibroin factor 1, *J. Biol. Chem. 266*, 16942−16947.
21. Durand, B., Drevet, J., and Couble, P. (1992) P25 gene regulation in Bombyx mori silk gland: two promoter-binding factors have distinct tissue and developmental specificities, *Mol. Cell. Biol. 12*, 5768−5777.
22. Horard, B., Julien, E., Nony, P., Garel, A., and Couble, P. (1997) Differential binding of the Bombyx silk gland-specific factor SGFB to its target DNA sequence drives posterior-cell-restricted expression, *Mol. Cell. Biol. 17*, 1572−1579.
23. Takiya, S., Kokubo, H., and Suzuki, Y. (1997) Transcriptional regulatory elements in the upstream and intron of the fibroin gene bind three specific factors POU-M1, Bm Fkh and FMBP-1, *Biochem. J. 321*, 645−653.
24. Takiya, S., Hui, C. C., and Suzuki, Y. (1990) A contribution of the core-promoter and its surrounding regions to the preferential transcription of the fibroin gene in posterior silk gland extracts, *EMBO J. 9*, 489−496.
25. Hui, C. C., and Suzuki, Y. (1989) Enhancement of transcription from the Ad2 major late promoter by upstream elements of the fibroin- and sericin-1-encoding genes in silk gland extracts, *Gene 85*, 403−411.
26. Fukuta, M., Matsuno, K., Hui, C. C., Nagata, T., Takiya, S., Xu, P. X., Ueno, K., and Suzuki, Y. (1993) Molecular cloning of a POU domain-containing factor involved in the regulation of the Bombyx sericin-1 gene, *J. Biol. Chem. 268*, 19471−19475.
27. Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T., and Suzuki, Y. (1995) Silk gland factor-1 involved in the regulation of Bombyx sericin-1 gene contains fork head motif, *J. Biol. Chem. 270*, 9340−9346.
28. Takiya, S., Ishikawa, T., Ohtsuka, K., Nishita, Y., and Suzuki, Y. (2005) Fibroin-modulator-binding protein-1 (FMBP-1) contains a novel DNA-binding domain, repeats of the score and three amino acid peptide (STP), conserved from Caenorhabditis elegans to humans, *Nucleic Acids Res. 33*, 786−795.
29. Jofuku, K. D., den Boer, B. G., Van Montagu, M., and Okamuro, J. K. (1994) Control of Arabidopsis flower and seed development by the homeotic gene APETALA2, *Plant Cell 6*, 1211−1225.
30. Ohme-Takagi, M., and Shinshi, H. (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element, *Plant Cell 7*, 173−182.
31. Kagaya, Y., Ohmiya, K., and Hattori, T. (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants, *Nucleic Acids Res. 27*, 470−478.
32. Andrade, M. A., Chacón, P., Merelo, J. J., and Morán, F. (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network, *Protein Eng. 6*, 383−390.
33. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) Improved spectral resolution in COSY 1H NMR spectra of proteins via double quantum filtering, *Biochem. Biophys. Res. Commun. 117*, 479−485.

34. Braunschweiler, L., and Ernst, R. (1983) Coherence transfer by isotopic mixing: Application to protein correlation spectroscory, *J. Magn. Reson. 53*, 521−528.
35. Macura, S., Huang, Y., Suter, D., and Ernst, R. (1981) Two-dimensional chemical-exchange and cross-relaxation spectroscopy of coupled nuclear spins, *J. Magn. Reson. 43*, 259−281.
36. Zuiderweg, E., Hallenga, K., and Olejniczak, E. (1986) Improvement of 2D NOE spectra of biomolecules in H_2O solution by coherent suppression of the solvent resonance, *J. Magn. Reson. 70*, 336−343.
37. Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological molecules, *J. Biomol. NMR 5*, 1−10.
38. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
39. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
40. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
41. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graph. 14*, 51−55.
42. Wüthrich, K. (1986) Sequence-Specific Resonance Assignments in Proteins, in *NMR of Proteins and Nucleic Acids*, pp 130−161, John Wiley & Sons, New York.
43. Munoz, V., and Serrano, L. (1995) Helix design, prediction and stability, *Curr. Opin. Biotechnol. 6*, 382−386.
44. Gao, G., DeRose, E. F., Kirby, T. W., and London, R. E. (2006) NMR determination of lysine pK_a values in the Pol λ lyase domain: mechanistic implications, *Biochemistry 45*, 1785−1794.
45. Liu, A., Hu, W., Majumdar, A., Rosen, M. K., and Patel, D. J. (2000) NMR detection of side chain-side chain hydrogen bonding interactions in 13C/15N-labeled proteins, *J. Biomol. NMR 17*, 305−310.
46. Bosshard, H. R., Marti, D. N., and Jelesarov, I. (2004) Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings, *J. Mol. Recognit. 17*, 1−16.
47. Fairman, R., Shoemaker, K. R., York, E. J., Stewart, J. M., and Baldwin, R. L. (1990) The Glu 2−·Arg 10+ side-chain interaction in the C-peptide helix of ribonuclease A, *Biophys. Chem. 37*, 107−119.
48. Wu, Z., Li, X., de Leeuw, E., Ericksen, B., and Lu, W. (2005) Why is the Arg5-Glu13 salt bridge conserved in mammalian alpha-defensins? *J. Biol. Chem. 280*, 43039−43047.
49. Aurora, R., and Rose, G. D. (1998) Helix capping, *Protein Sci. 7*, 21−38.
50. Lanigan, M. D., Tudor, J. E., Pennington, M. W., and Norton, R. S. (2001) A helical capping motif in ShK toxin and its role in helix stabilization, *Biopolymers 58*, 422−436.
51. Lyu, P. C., Wemmer, D. E., Zhou, H. X., Pinker, R. J., and Kallenbach, N. R. (1993) Capping interactions in isolated alpha helices: position-dependent substitution effects and structure of a serine-capped peptide helix, *Biochemistry 32*, 421−425.
52. Tudor, J. E., Pennington, M. W., and Norton, R. S. (1998) Ionisation behaviour and solution properties of the potassium-channel blocker ShK toxin, *Eur. J. Biochem. 251*, 133−141.
53. Spolar, R. S., and Record, M. T. (1994) Coupling of local folding to site-specific binding of proteins to DNA, *Science 263*, 777−784.
54. Johnson, N. P., Lindstrom, J., Baase, W. A., and von Hippel, P. H. (1994) Double-stranded DNA templates can induce alpha-helical conformation in peptides containing lysine and alanine: functional implications for leucine zipper and helix-loop-helix transcription factors, *Proc. Natl. Acad. Sci. U.S.A. 91*, 4840−4844.
55. Werner, M. H., Gronenborn, A. M., and Clore, G. M. (1996) Intercalation, DNA kinking, and the control of transcription, *Science 271*, 778−784.
56. Okubo, S., Hara, F., Tsuchida, Y., Shimotakahara, S., Suzuki, S., Hatanaka, H., Yokoyama, S., Tanaka, H., Yasuda, H., and Shindo, H. (2004) NMR structure of the N-terminal domain of SUMO ligase PIAS1 and its interaction with tumor suppressor p53 and A/T-rich DNA oligomers, *J. Biol. Chem. 279*, 31455−31461.
57. Assa-Munt, N., Mortishire-Smith, R. J., Aurora, R., Herr, W., and Wright, P. E. (1993) The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage lambda repressor DNA-binding domain, *Cell 73*, 193−205.
58. Da, G., Lenkart, J., Zhao, K., Shiekhattar, R., Cairns, B. R., and Marmorstein, R. (2006) Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes, *Proc. Natl. Acad. Sci. U.S.A. 103*, 2057−2062.
59. Clubb, R. T., Schumacher, S., Mizuuchi, K., Gronenborn, A. M., and Clore, G. M. (1997) Solution structure of the I gamma subdomain of the mu end DNA-binding domain of phage mu transposase, *J. Mol. Biol. 273*, 19−25.

BI061780C