J. Biochem. **136**, 29–37 (2004)
DOI: 10.1093/jb/mvh091

Side Chain–Side Chain Interactions of Arginine with Tyrosine and Aspartic Acid in Arg/Gly/Tyr–Rich Domains within Plant Glycine-Rich RNA Binding Proteins

Yasuhiro Kumaki[1], Katsutoshi Nitta[2], Kunio Hikichi[3], Takeshi Matsumoto[3] and Norio Matsushima[3,*]

[1]High-Resolution NMR Laboratory and [2]Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810; and [3]School of Health Sciences, Sapporo Medical University, Sapporo 060-8556

Received March 12, 2004; accepted April 17, 2004

Plant glycine-rich RNA-binding proteins (GRRBPs) contain a glycine-rich region at the C-terminus whose structure is quite unknown. The C-terminal glycine-rich part is interposed with arginine and tyrosine (arginine/glycine/tyrosine (RGY)–rich domain). Comparative sequence analysis of forty-one GRRBPs revealed that the RGY-rich domain contains multiple repeats of Tyr-$(Xaa)_h$-$(Arg)_k$-$(Xaa)_l$, where Xaa is mainly Gly, "k" is 1 or 2, and "h" and "l" range from 0 to 10. Two peptides, 1 ($G^1G^2Y^3G^4G^5G^6R^7R^8D^9G^{10}$) and 2 ($G^1G^2R^3R^4D^5G^6G^7Y^8G^9G^{10}$), corresponding to sections of the RGY-rich domain in *Zea mays* RAB15, were selected for CD and NMR experiments. The CD spectra indicate a unique, positive band near 228 nm in both peptides that has been ascribed to tyrosine residues in ordered structures. The pH titration by NMR revealed that a side chain-side chain interaction, presumably an $H\text{-}N^{\varepsilon}\cdots O{=}C^{\gamma}$ hydrogen bonding interaction in the salt bridge, occurs between Arg (*i*) and Asp (*i* + 2). 1D GOESY experiments indicated the presence of NOE between the aromatic side chain proton and the arginine side chain proton in the two peptides suggesting strongly that the Arg (*i*) aromatic side chain interacts directly with the Tyr (*i* ± 4 or *i* ± 5) side chain.

Key words: glycine-rich RNA binding protein, arginine/glycine/tyrosine–rich domain, arginine-tyrosine side chain interactions, arginine-aspartic acid side chain interactions.

Abbreviations: GRRBP, glycine-rich RNA-binding protein; RGY-rich domain, arginine-, glycine-, and tyrosine-rich domain; RRM, RNA recognition motif; *Zm*RAB15, RAB15 from *Zea mays*; GOESY, Gradient enhanced NOE spectroscopy; GROESY, Gradient enhanced ROESY.

Plant glycine-rich RNA binding proteins (GRRBPs) are small, approximately 15–17 kDa, and comprise two very diverse and characteristic regions (Fig. 1a) (*1*, *2*). The RNA recognition motif (RRM) of ~90 amino acids occupies the N-terminal part of the protein and is highly homologous among different proteins. The C-terminal part is highly variable in length (~20 to ~80 amino acids) and is extremely rich in glycine residues interposed with arginine and tyrosine; we term this region the arginine/glycine/tyrosine (RGY)–rich domain (Fig. 1b). The RGY-rich domain contains repeats of the RGG-box, which has been identified as an RNA-binding motif in other proteins (*3*). This protein family has been extended to include representatives from mammals and cyanobacteria (*4–6*). Their function remains unknown, although the finding that maize and tobacco proteins accumulate in the nucleolus suggests a role in pre-rRNA processing (*7*, *8*).

The three-dimensional structure of the RRM domain within the U1A protein has been determined, revealing a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ folding topology (*9*). In contrast, the structure of the C-terminal RGY-rich domain is quite unknown. The RGY-rich domain is characterized by quasi-repetitive amino acid sequences configured in tandem, as described later herein (Fig. 1b). The RGY-rich domain is highly flexible because of the numerous glycines. The high flexibility and repetitiveness in the RGY-rich domain appear to contribute to the difficulty in determining the three-dimensional structure by conventional NMR and X-ray analyses.

Repetitive RGY-rich domains occur widely in many proteins. They are present in the prions that are responsible for the transmissible spongiform encephalopathies (*10*). The RGY-rich domain in prions consists of four tandem repeats of PHGGGWGQ. NMR experiments revealed that the PHGGGWGQ segment adopts a loop conformation with histidine (*i*) and tryptophan (*i* + 4) in close proximity (*11*, *12*). A cation-π interaction appears to occur between histidine and tryptophan. Thus, aromatic side chain interactions that may form a loop conformation are expected to occur in RGY-rich domains.

The purpose of the present study is to investigate the structure of the C-terminal RGY-rich domains within GRRBPs. First, comparative sequence analysis was performed on the quasi-repetitive RGY-rich domains in forty-one plant GRRBPs. The results were also utilized to select peptide sequences to be used in CD and NMR

*To whom correspondence should be addressed Tel: +81-11-611-2111, Fax: +81-11-613-7134, E-mail: matusima@sapmed.ac.jp

Fig. 1. **Schematic diagram of glycine-rich RNA-binding protein (GRRBP) (a) and multiple alignment of the amino acid sequences of nine representative proteins: *Zm*RAB15, *Cr*GRP-2, *Cr*GRP-1, *Sb*GRP, *St*GRP, ScGRP-1, *Ng*GRP, *Dc*GRP, and *Ns*GRP-1b (b).** The amino acid sequences of these GRRBPs were taken from sequence data banks (SWISS-PROT, PRF, PIR or Genbank): *Zm*RAB15(P10979), *Cr*GRP-2(AF200323), *Cr*GRP-1(AF200321), *Sb*GRP(AF310215) *St*GRP(Z49197), *Dc*GRP(Q03878), ScGRP-1(T10479), *Ng*GRP(2415521A) and *Ns*GRP-1b(D16205). The quasi-repetitive, arginine-, glycine-, and tyrosine-rich domain (RGY-rich domain) is solid, while the RNA recognition motif (RRM) is hatched. The amino acid sequences of the multiple alignments are indicated only for the RGY-rich domain. The sequence RR(E/D) in the RGY-rich domains is shaded and aromatic residues Tyr or Phe, Trp are indicated by bold type. Abbreviations: species: *Cr*, *Catharanthus roseus*; *Dc*, *Daucus carota*; *Ng*, *Nicotiana glutinosa*; *Ns*, *Nicotiana sylvestris*; *Sb*, *Sorghum bicolar*; *Sc*, *Solanum commersonii*; *Zm*, *Zea mays*.

experiments. Secondly, these experiments were performed on two short peptides corresponding to sections of the RGY-rich domain within GRRBPs. NOE signals were obtained in a GOESY experiment that detects small signals of interest (*13*). We identified side chain–side chain interactions between Arg (i) and Asp (i + 2), and between Arg (i) and Tyr ($i \pm 4$ or $i \pm 5$).

MATERIALS AND METHODS

Amino Acid Sequence—The amino acid sequences of multiple alignments of RGY-rich domains were extracted from forty-one GRRBPs from twenty-four plant pecies: GRP-7 and GRP-8 from *Arabidopsis thaliana*; GRP-10 from *Brassica napus*; GRP-1, GRP-2 and GRP-3 from *Catharanthus roseus*; GRP from *Citrus unshiu*; GRP from *Daucus carota*; two GRPs from *Euphorbia esula*; GRP from *Glycine max*; two GRPs from *Hordeum vulgare*; GRP from *Lycopersicon esculentum*; GRP-1 from *Medicago sativa*; GRP from *Nicotiana glutinosa*; GRP-1a, GRP-1b, and GRP-1c from *Nicotiana sylvestris*; GRP-1 and GRP-2 from *Oryza sativa*; GRP from *Picea glauca*; GRP-1 from *Pelargonium x hortorum*; GRP1 from *Physcomitrella patens*, GRP-1 and GRP-2 from *Ricinus communis*; GRP from *Rumex obtusifolius*; GRP1A and GRP2A from *Sinapis alba*; GRP, GRP-1 and GRP-2 from *Sorghum bicolar*; GRP-S1 and GRP-S2 from *Sorghum vulgare*; GRP-1 from *Solanum commersonii*; GRP from *Solanum tuberosum*; GRP-1 from *Triticum aestivum*; and RAB15 and three GRPs from *Zea mays*.

Peptides Samples—We designed the following four peptides; **1**, H-$G^1G^2Y^3G^4G^5G^6R^7R^8D^9G^{10}$-OH; **2**, H-$G^1G^2R^3R^4D^5G^6G^7Y^8G^9G^{10}$-OH; **3**, H-$G^1G^2F^3G^4G^5G^6R^7R^8D^9G^{10}$-OH; **4**, H-$G^1G^2R^3R^4D^5G^6G^7F^8G^9G^{10}$-OH. The four peptides were purchased from SIGMA Genosys Japan K.K. (Ishikari, Japan). All peptides were chemically synthesized using solid-phase Fmoc chemistry on an Applied Biosystems Japan Ltd. Pioneer peptide synthesizer. Synthetic peptides were purified by reverse-phase HPLC using a TSKgel ODS-80Ts column (Tosoh Coop., Japan).

Circular Dichroism (CD)—CD spectra were recorded on a JASCO J-725 spectropolarimeter (Jasco, Japan) at peptide concentrations of 0.21–0.36 mM, pH 4.3–4.4 and pH 6.1–6.3, and 278K. Typically a cell with a 0.2 cm pathlength was used to obtain spectra between 190 and 260 nm with sampling points every 0.5 nm. Direct CD measurements (θ, in mdeg) were converted to molar ellipicity, [θ] using the relationship [θ] = θ/cl, where c is the concentration and l is the pathlength. Molar ellipticity [θ] is in units deg cm^2 dmol^{-1}.

NMR Experiments—All NMR spectra were measured on JEOL ALPHA 500 spectrometers. For the measurements, the peptides were dissolved in H_2O/D_2O (90%/10%) or D_2O (99.96%). The pH values were adjusted to 4.3 for H_2O with acetic acid-d_4 and KOD and the pD values were adjusted to 4.3 and 6.2 for D_2O. Peptide concentrations were approximately 12.0 mM. ^{1}H-NMR spectra for 12 mM **1** and **2**, pH 4.3, 278 K were compared with those for a sufficiently diluted solution of 0.3 mM with no significant difference observed. DSS was added to the solutions as an internal reference (0.0 ppm).

Averaged diffusion coefficients of the six 12 mM solutions at pH 4.3 for peptides **1**, **2**, **3**, and **4** and pH 6.2 for peptides **1** and **2** were determined by the stimulated echo method (*14*) at 278K. The diffusion coefficients of the 12 mM peptide solutions were also determined as a function of diffusion time ranging from 11 ms to 176 ms. The

results of the two experiments were used to exclude the possibility of intermolecular NOEs, as discussed later.

Sequence specific assignments of proton resonances for **1** and **2** were obtained by 2D NOESY (mixing time = 100, 175, 250 ms) (*15*, *16*), TOCSY (mixing time = 100 ms) (*17*), ROESY (mixing time = 250 ms) (*18*), and DQF-COSY (*19*) experiments at 278 K. The 2D NMR spectra were measured in the phase-sensitive mode using TPPI-State for quadrature detection in f_1. The water signal was suppressed by presaturation or by the WATERGATE method in TOCSY and NOESY (*20*), and by coherence selection with pulsed field gradients in DQF-COSY experiments. Temperature coefficients of the chemical shifts of amide proton resonances were estimated from the 2D TOCSY spectra obtained at 278 K and 303 K. The $^3J_{NHC}\alpha_H$ coupling constants were derived from the absorptive and dispersive components of f_2 active splitting of the DQF-COSY spectra (f_2 digital resolution: 0.82 Hz) at 278K, which effectively eliminates errors due to linewidth (*21*).

The measurements of GOESY (mixing time = 100, 175, 250 ms) incorporating selective refocusing and pulsed field gradient enhancement (*13*) and GROESY (mixing time = 250 ms) (*22*), which facilitate the detection of small NOE signals of interest, were performed for **1** and **2** in D_2O solution. The $C^\delta H$ and $C^\varepsilon H$ protons of the Tyr side chain were selected as target protons. The use of D_2O can eliminate the signals of exchangeable protons whose chemical shifts are close to those of the $C^\delta H$ and $C^\varepsilon H$ protons. A RE-BURP pulse of 80 ms duration (*23*) was used as a selective refocusing π pulse. For **3** and **4**, GROESY measurements (mixing time = 250 ms) were performed using a RE-BURP pulse of 16.3 ms duration. The carrier frequency of these pulses was positioned at the center of the aromatic signal region. The number of scans for all of the GOESY and GROESY measurements was 47,104.

The ^1H chemical shifts of arginine and aspartic acid were measured as a function of pH. The chemical shifts of Arg $N^\varepsilon H$ were determined in 90% H_2O/10% D_2O at 278 K, and those of Asp $C^\alpha H$ were determined in 100% D_2O at 278 K (not deuterium corrected).

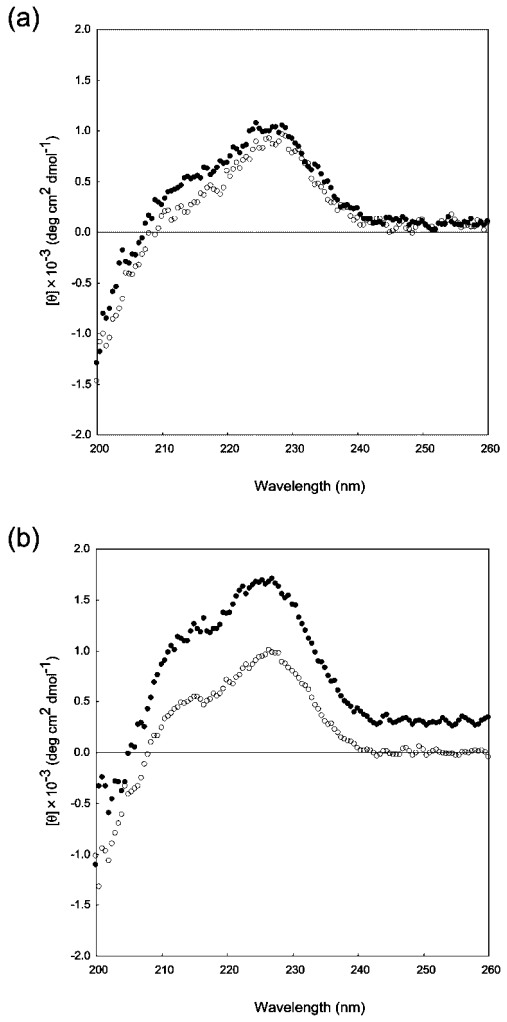

Fig. 2. **CD spectra of 1 (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) (a) and 2 (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) (b).** (a) 0.21 mM at 298K (no buffer); (b) 0.36 mM at 298K (no buffer). open circles, at pH.4.1 for **1** and pH 4.3 for **2**; solid circles, at pH 6.1 for **1** and pH 6.3 for **2**.

RESULTS

Sequence Analysis of the RGY-Rich Domains—Multiple alignments between RGY-rich domains were performed by a key residue of the aromatic amino acids (Fig. 1b). This analysis revealed that the RGY-rich domains with a large variety in length are represented by tandem repeats consisting of Tyr-$(Xaa)_h$-$(Arg)_k$-$(Xaa)_l$ or Tyr-$(Xaa)_m$ containing no Arg residues, where Tyr is sometimes replaced by Phe or Trp, Xaa is mainly Gly but includes polar residues such as Glu, Asp, Asn, Ser, Ala, Gln, and rarely other residues, and where "k" is 1 or 2, and "h" or "l" and "m" are highly variable, ranging from 1 to 10, and 2 to 8, respectively.

The segments of Tyr-$(Xaa)_h$-$(Arg)_k$-$(Xaa)_l$ are frequently described by Tyr-$(Gly)_h$-Arg-Arg-Asp/Glu-$(Gly)_l$. They include characteristic sequences such as YGG(R/Q)RE, YGGGRR(E/D), YGGGGRRE, RR(E/D/N)GGY, RREGGGY, RREGGGGGY, RREGGGGGGY, and RREGGGGGGGY. In particular, YGGGRR(E/D) and RR(E/D/N)GGY appear with the highest frequency. Thus, for the first NMR study of the RGY-rich domains we selected two segments containing RRDGGY or YGGGRRD. The two segments are **1** (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) and **2** (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) corresponding to 103–112 and 107–116, respectively, of *Z. mays* RAB15, which contains 157 amino acids (Fig. 1b) (*24*). These two segments are also seen in *C. roseus* GRP2 and GRP3 (*25*).

CD Spectra—The far UV-CD spectra of peptides **1** and **2** are shown in Fig. 2. The main feature is a fairly intense peak near 228 nm in both peptides and a peak near 215 nm in **2**. Although a 229-nm extremum is not usually observed, bands near 230 nm have been observed in the spectra of some proteins, including the phage fd gene 5 protein (Ffg5p) (*26–29*). The origin of these bands has been attributed to aromatic groups. For Ffg5p, the large band has been confirmed to be due primarily to the L_a bands of the five tyrosyl residues (*28*). Thus, the 228-nm CD bands observed here are attributed to the tyrosine residues in ordered structures (Tyr3 in **1** and Tyr8 in **2**).

NMR Resonance Assignments—The sequential spectral assignments for the two peptides, **1** and **2**, were com-

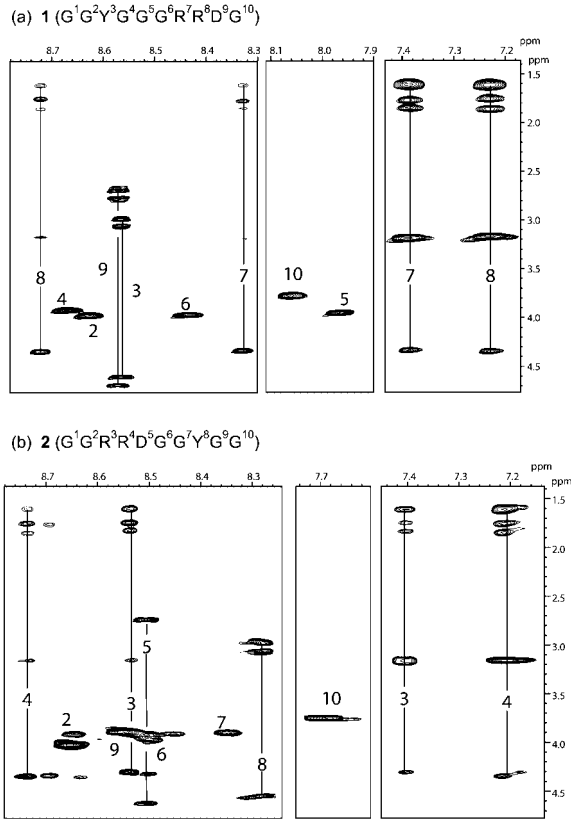

Fig. 3. **Aliphatic to the backbone amide regions and the side chain amino region of the arginine of the 2D-TOCSY spectra of two peptides, 1 (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) (a) and 2 (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) (b).**

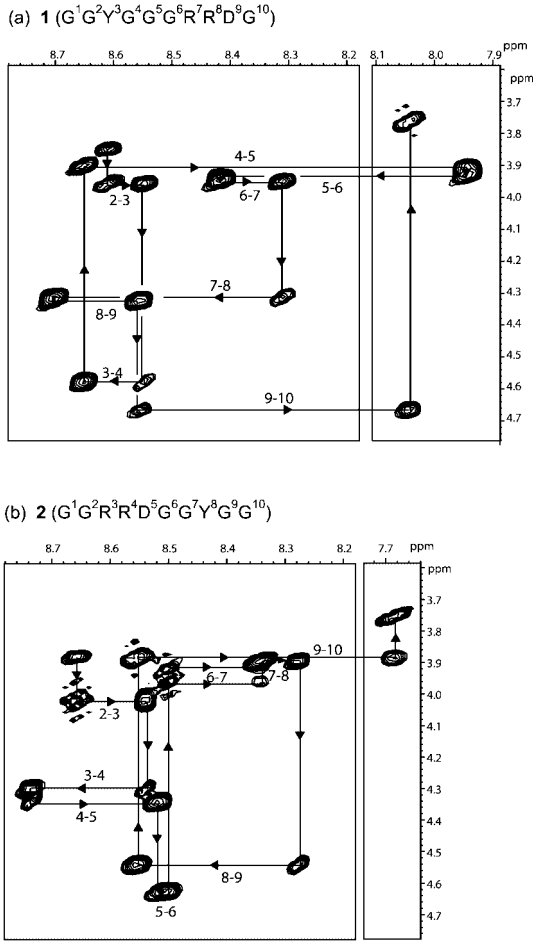

Fig. 4. **Finger-print regions of the 2D-NOESY spectra of 1 (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) (a) and 2 (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) (b).**

pleted by TOCSY and NOESY experiments in H_2O/D_2O (90%/10%), at 278K, pH 4.3 (Figs. 3 and 4, and Table 1). Minor components are not observed in the NMR spectra.

Nuclear Overhauser Effect (NOE) Connectivities between Arg and Tyr—Figure 5 shows the 1D NOE spectra for **1** and **2** in D_2O solution at 278K, pD 4.3 which were obtained by GOESY with selective refocusing of the C$^\varepsilon$H and C$^\delta$H protons of Tyr side chain (mixing time, 175 ms). As expected, medium-range NOE connectivities were detected: NOE between the Tyr3 C$^\varepsilon$H proton and the C$^\delta$H proton of Arg7 and/or Arg8 [d$\varepsilon\delta$ (Tyr3, Arg7/8) for abbreviation] and d$\varepsilon\alpha$ (Tyr3, Arg7/8) for **1** and d$\delta\varepsilon$ (Arg3/4, Tyr8), d$\varepsilon\alpha$ (Tyr8, Gly10) and d$\delta\alpha$ (Tyr8, Gly10) for **2**. The other NOEs observed by GOESY are described in Fig. 5. The corresponding NOEs were also observed at pD 6.3.

All NOEs obtained by the 1D GOESY experiment were confirmed by 1D GROESY (*22*). The NOEs between the arginine side chain and the aromatic side chain observed for **1** and **2** were also observed for **3** and **4**, which are characterized by the Tyr to Phe substitution in **1** and **2** (data not shown). Thus, the present NOE data for the four peptides demonstrate that the side chain of Arg (i) is in close proximity with the aromatic side chain of Tyr/Phe ($i \pm 4$ or $i \pm 5$).

Chemical Shifts of Arg Residues—The chemical shift of the N$^\varepsilon$H proton of Arg7 clearly differs from the signal of Arg8 in **1** (Fig. 6). Similarly, in **2** the N$^\varepsilon$H proton signal of Arg3 clearly differs from that of Arg4.

Merutka *et al.* (*30*) have reported proton chemical shifts of a series of disordered linear peptides (H-G^{1}G^{2}X^{3}G^{4}G^5-OH, with X being one of the 20 naturally occurring amino acids). The chemical shifts were obtained in 90% H_2O/10% D_2O at 277.2K, pH 5.0, which is close to the conditions used in this study. Therefore, the chemical shift of the Arg residue in the linear pentapeptide was compared with those of the two Arg residues in **1** and **2**. As shown in Fig. 3, the signals of the N$^\varepsilon$H proton of Arg7 but not Arg8 in **1** occur approximately 0.11 ppm downfield from those of the Arg3 residue in the GGRGG segment (*30*). Similarly, the signals of the N$^\varepsilon$H proton of Arg3 but not Arg4 in **2** occur approximately 0.15 ppm downfield from those of the Arg residue in the GGRGG segment (Fig. 5) (*30*).

The signals of the amide proton (NH) of Arg7 in **1** and of Arg3 in **2** occur approximately 0.37 ppm and 0.14 ppm upfield, respectively, from those of the Arg residue in the GGRGG segment (Fig. 6) (*30*).

*Temperature Coefficients and Chemical Shifts of Amide Proton Resonances of Gly5 in **1** and Gly10 in **2***—The extent to which an amide proton either participates in hydrogen bonding or is exposed to solvent is reflected in the

Table 1. **^{1}H-NMR chemical shifts, vicinal coupling constants, and temperature coefficients of amide proton chemical shifts for two peptides (1 and 2).**

Residue	Chemical shift[a]						$^3J_{NHC\alpha H}$[b]	$\Delta\delta/\Delta T$[c]
	NH	$C^\alpha H$	$C^\beta H$	$C^\gamma H$	$C^\delta H$	$C^\varepsilon H/N^\varepsilon H$		
1 (H-$G^1G^2Y^3G^4G^5G^6R^7R^8D^9G^{10}$-OH)								
Gly1		3.86						
Gly2	8.61	3.97						5.49
Tyr3	8.55	4.59	2.98/3.05		7.14	6.84	7.8	5.71
Gly4	8.65	3.92						7.77
Gly5	7.95	3.94						2.00
Gly6	8.42	3.97						5.54
Arg7	8.32	4.33	1.77/1.85	1.61	3.18	7.37	7.8	6.29
Arg8	8.71	4.35	1.75/1.86	1.61	3.17	7.21	6.5	7.54
Asp9	8.56	4.69	2.68/2.76				7.7	8.46
Gly10	8.05	3.77						6.46
2 (H-$G^1G^2R^3R^4D^5G^6G^7Y^8G^9G^{10}$-OH)								
Gly1		3.89						
Gly2	8.67	4.03						4.63
Arg3	8.55	4.31	1.75/1.83	1.62	3.17	7.41	6.4	7.32
Arg4	8.75	4.36	1.77/1.85	1.61	3.17	7.22	6.4	7.49
Asp5	8.52	4.63	2.75				7.3	6.28
Gly6	8.51	3.93/3.98						6.80
Gly7	8.35	3.90						4.80
Tyr8	8.29	4.55	2.98/3.06		7.15	6.84	7.3	7.09
Gly9	8.56	3.90						6.84
Gly10	7.69	3.76						3.54

[a]Chemical shifts for each amino acid at 278 K. Chemical shifts are reported in parts per million. [b] $^3J_{NHC\alpha H}$ is three-bond scalar coupling between the backbone amide proton and the α proton. $^3J_{NHC\alpha H}$ is reported in Hz. [c]$\Delta\delta/\Delta T$ is the temperature coefficient of the amide proton chemical shifts. Coefficients measured between 278 and 298K are included (ppb/K).

temperature coefficient of the amide proton chemical shift ($-\Delta\delta/\Delta T$). An amide proton showing $-\Delta\delta/\Delta T$ below 5.0 ppb/K is sequestrated significantly from water, a state that is not consistent with an unordered structure in which amide protons are completely exposed to water.

Peptides **1** and **2** have the GGYGG segment at the N-terminus and C-terminus, respectively. The Gly residues, Gly5 in **1** and Gly10 in **2**, that are located at the *i*+2 position when tyrosine occupies the *i* position, show a remarkable reduction in amide proton temperature coefficients (Table 1). Gly5 in **1** and Gly10 in **2** also show significant deviations in their NH chemical shifts (Table 1). These results are consistent with those for Gly5 in H-$G^1G^2X^3G^4G^5$-OH containing the aromatic residues Tyr, Phe, and Trp (*30*).

pH Titration of Arg and Asp Residues by NMR—Fig. 7a shows the chemical shifts versus pH of the $N^\varepsilon H$ protons of Arg7 and Arg8 and the $C^\alpha H$ proton of Asp9 in **1**. Similarly, the pH titration of the protons of Arg3, Arg4 and Asp 5 in **2** is shown in Fig. 7b. Surprisingly, the $N^\varepsilon H$ proton resonances of Arg7 in **1** and Arg3 in **2** shift largely upfield with decreasing pH in a sigmoidal fashion and then reach the chemical shifts of a random coil below about pH 3, while those of Arg8 $N^\varepsilon H$ in **1** and Arg4 $N^\varepsilon H$ in **2** hardly change (Fig. 7a). In **1**, the midpoint of the sigmoid for Arg7 $N^\varepsilon H$ (3.7) is consistent with 3.8 for Asp9 $C^\alpha H$. The same result was also obtained for Arg3 $N^\varepsilon H$ and Asp5 $C^\alpha H$ in **2** (Fig. 7b). The observed midpoints for Asp9 in **1** and Asp5 in **2** correspond to the pK_a value for ionization of the side chain of aspartic acid. Thus, it is concluded that the side chain-side chain interaction occurs between Arg (*i*) and Asp (*i* + 2) in both peptides.

DISCUSSION

Intramolecular and Intermolecular Interactions—The NMR data including the NOEs were obtained for 12 mM peptide solutions. The observed side chain-side chain interactions, such as Arg-Asp and Tyr-Arg interactions, may be intermolecular interactions. However, the possibility is very low based on the following observations: (*1*) the far-UV CD spectra were observed in very dilute solutions; (*2*) average diffusion coefficients of the 12 mM peptide solutions ($1.51 \pm 0.05 \times 10^{-10}$ m^2 s^{-1}) were close to 1.59×10^{-10} m^2 s^{-1} of the 0.3 mM peptide solution at pH 4.3 for **1**; (*3*) the NMR spectra showed no minor components (Figs. 3–5); and (*4*) the diffusion coefficients of the 12 mM solutions were independent of diffusion time.

The unusual, positive 228 nm extremum in the CD spectra indicates that the tyrosine residues in **1** and **2** participate in ordered structures even in dilute solutions. The results (*2*) indicate that the monomer is dominant in 12 mM solution. Even if dimers exist, the population will be small. Results (*3*) and (*4*) indicate that even if a small population of dimers exists, the exchange rate between dimers and monomers is too fast to exhibit intermolecular cross relaxation. Moreover, it has been reported that intermolecular interactions occur between Ac-KKKKK-NH_2 and Ac-DDXDD-NH_2, in which X is Tyr and Phe (*31*). Despite the order of 10^3 M^{-1} for the association constants, NOEs between intermolecular residues have not

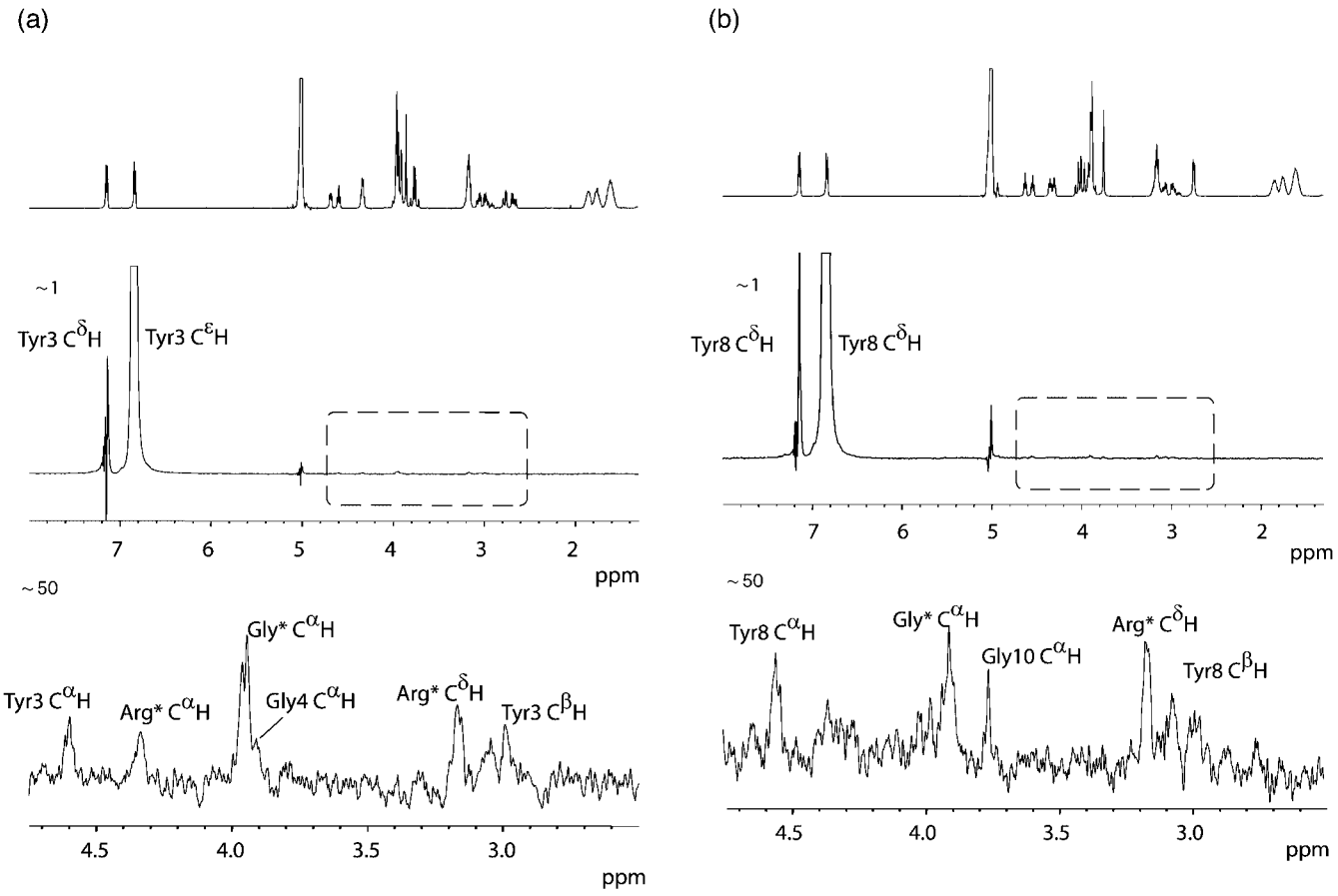

Fig. 5. **The 1D GOESY spectra of** 1 **(H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) (a) and** 2 **(H- G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) (b).** The top and bottom of the figure show the 1D GOESY spectra that are expansions of the region enclosed by the rectangle. The GOESY spectra were measured at 278 K using a 175ms mixing time in D_2O solution at 278 K, pD 4.3. Arg* indicates Arg7 and/or Arg8 in **1** and Arg3 and/or Arg4 in **2**, and Gly* indicate Gly2, Gly5, and/or Gly6 in **1** and Gly1, Gly6, Gly7, and/or Gly9 in **2** whose spectra show overlapping resonances. The following NOEs are detected in the figure: for **1**, dεδ (Tyr3, Arg*), dεα (Tyr3, Arg*), dεα (Tyr3, Gly*), dδα (Tyr3, Gly*), dεα(Tyr3, Gly4), dαε (Tyr3, Tyr3), and dβε (Tyr3, Tyr3), and, for **2**, dδε (Arg*, Tyr8), dεα (Tyr8, Gly10), dεα (Tyr8, Gly*), dαε (Tyr3, Tyr3), and dβε (Tyr3, Tyr3).

been observed (*31*). This strongly suggests that the Arg-Tyr interactions are intramolecular interactions, although the possibility of intermolecular interactions is not rigorously excluded.

The medium-range NOEs coming from the Arg-Tyr side chain interactions were not observed by 2D NOESY. The upfield shifts of Arg side chain protons, which are expected in Arg-Tyr interactions due to cation-π interactions (*31*), were also not observed. Thus, the population of folded molecules having Arg-Tyr interactions and Arg-Asp interactions is considered to be low.

Side Chain–Side Chain Interactions between Arg and Asp—The pH titration by NMR revealed that side chain–side chain interactions occur between Arg (*i*) and Asp (*i* + 2) at pH 4.3–4.4. The interaction is presumably the H-N$^{\varepsilon}$···O=C$^{\gamma}$ hydrogen bonding interaction in the Arg (*i*)-Asp (*i* + 2) salt bridge. Hydrogen bond formation induces a downfield shift of the N$^{\varepsilon}$H proton of arginine (*32*). In fact, such downfield shifts were observed for Arg7 in **1** and for Arg 3 in **2** (Fig. 6).

Interestingly, IDITIS™ data analyses have revealed that, in several intramolecular contacts, the interacting aspartate and arginine residues are very close in the sequence and frequently have an $i \rightarrow i + 2$ relationship (*33*). This relationship is consistent with the Arg-Arg-Asp segment in **1** and **2**.

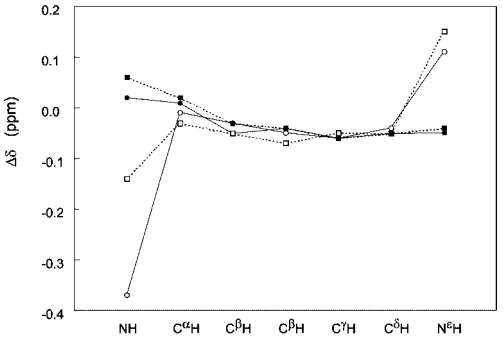

Fig. 6. **Deviations of the side-chain chemical shifts of Arg7 and Arg8 in 1 (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) and Arg3 and Arg4 in 2 (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) from those of the GGRGG segment (*25*).** The chemical shifts for **1** and **2** were obtained in 90% H_2O/10% D_2O at 278 K, pH 4.3, and for the GGRGG segment in 90% H_2O/10% D_2O at 277.2 K, pH 5.0. open circles, Arg7 in **1**; solid circles, Arg8 in **1**; open squares, Arg3 in **2**; solid squares, Arg4 in **2**.

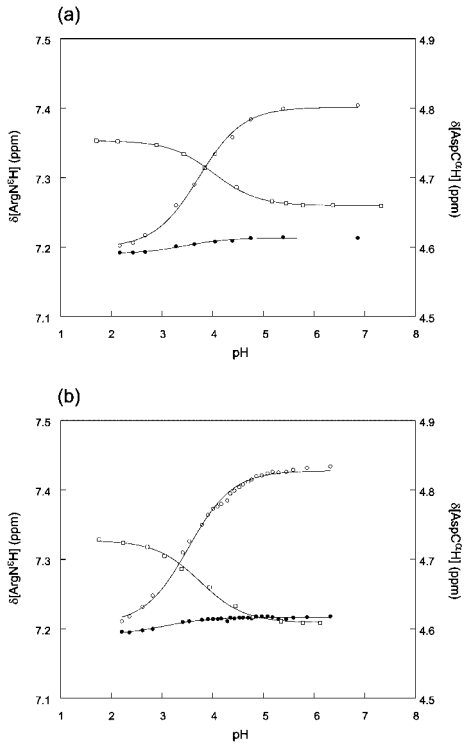

Fig. 7. **Plots of the ^{1}H chemical shifts of arginine and aspartic acid versus pH.** (a) Arg7 N$^{\varepsilon}$H (open circles), Arg8 N$^{\varepsilon}$H (solid circles) and Asp9 C$^{\alpha}$H (open squares) in **1** (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH). (b) Arg3 N$^{\varepsilon}$H (open circles), Arg4 N$^{\varepsilon}$H (solic circles) and Asp5 C$^{\alpha}$H (open squares) in **2** (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH). The chemical shifts of Arg N$^{\varepsilon}$H were determined in 90% H_2O/10% D_2O at 278 K, and those of Asp C$^{\alpha}$H were determined in 100% D_2O at 278K (not deuterium corrected).

Side Chain–Side Chain Interactions between Arg and Tyr—The NOE data indicated that the side chain–side chain interactions between Arg and Tyr are interrupted by 3 to 4 glycine residues. Aromatic side chains are known to participate in a variety of relatively strong and weak interactions including aromatic-aromatic interactions (*34–39*), cation-π interactions (*40–42*), and amino-aromatic interactions (*43–46*). Cation-π interactions have been observed for short peptides such as P^{1}P^{2}K^{3}Y^{4}D^{5}K^6 (*47*) and K^{1}C^{2}K^{3}K^{4}K^{5}G^{6}D^{7}D^{8}F^{9}C^{10}D^{11} (*31*) in NMR experiments. Since the arginine residues were positively charged under the experimental conditions used, the observed Arg-Tyr side chain interactions are ascribed to cation-π interactions.

Interaction between the Aromatic Side Chain Proton and Backbone Amide Proton—Gly5 of **1** and Gly10 of **2** showed a remarkable reduction in their NH proton temperature coefficients. Gly10 of **2** and Gly5 of **1** showed deviations in the NH chemical shift (Table 1) that concur well with those for Gly5 of G^{1}G^{2}X^{3}G^{4}G^5 and for Gly3 of X^{1}A^{2}G^3 containing the aromatic residues Tyr, Phe, and Trp (*30*, *48*, *49*). These observations are ascribed to aromatic (*i*) –backbone NH (*i* + 2) interactions (*48*).

It is concluded that, in addition to cation-π interactions (*37–39*), amino-aromatic interactions (*43–46*) appear to contribute to the formation of Arg-Tyr side chain interactions.

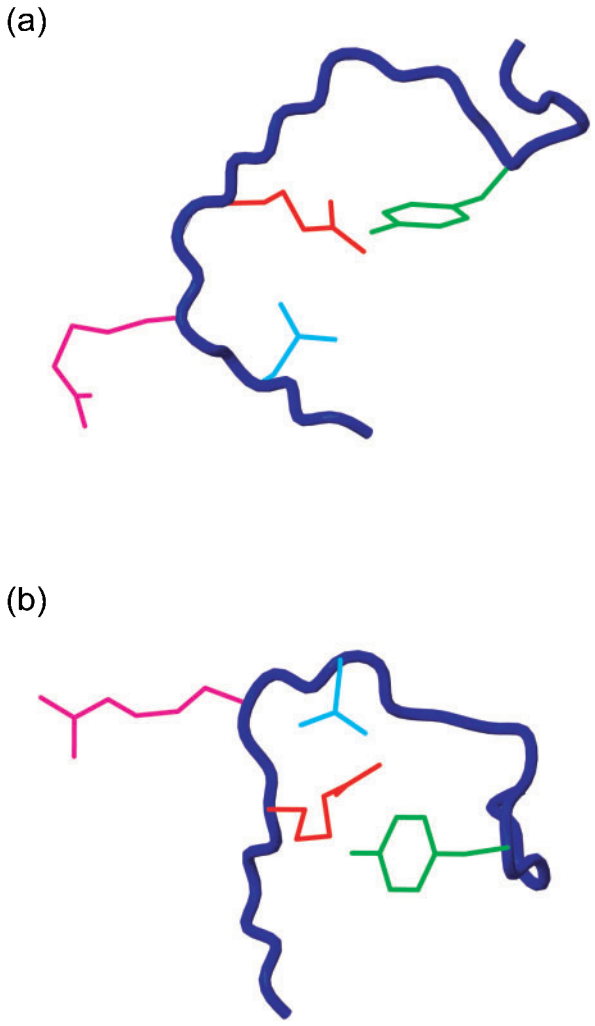

Fig. 8. **Plausible structures of 1 (H-G^{1}G^{2}Y^{3}G^{4}G^{5}G^{6}R^{7}R^{8}D^{9}G^{10}-OH) (a) and 2 (H-G^{1}G^{2}R^{3}R^{4}D^{5}G^{6}G^{7}Y^{8}G^{9}G^{10}-OH) (b).** Blue shows the backbone chain represented by the ribbon. Red/cyan, green and light blue show the side chains of Arg, Tyr and Asp, respectively.

Loop Search by the PDB File—The side chain–side chain interactions between Arg and Tyr suggest that the two peptides, **1** and **2**, can adopt a loop conformation, as seen for PHGGGWGQ in prions (*11*, *12*). The common feature among these peptide sequences is the presence of an aromatic residue and a positively charged residue separated by an intervening glycine-rich region. A loop, which is characterized by an aromatic side chain and a positively charged side chain in close proximity, is expected to be present in the known structures of proteins. We searched for this loop using the peptide sequences (Y/F/W)GGx(R/K) and (R/K)xxGG(Y/F/W) (where "x" is any amino acid) by PDB (http://www.cbrc.jp/papia/papiaJ.html). Loops were found in five proteins with segments consisting of KKNGGY (1g8m), KVDGGY (1dfx), KVSGGY (1pdk and 1qpx), YGGSR (1qgb) and WGGTK (1agg).

Plausible models for **1** and **2** based on the observed NMR data are shown in Fig. 8. The side chain–side chain interactions between Arg and Asp or Tyr contribute to the formation of the loop structures in these models.

Implications for the Function of RGY-Rich Domains—In the RZ-1 protein, which is a GRRBP, the C-terminal RGY-rich domain, in addition to the RRM, is essential for RNA-binding activity (*50*). Although the RZ-1 RGY-rich domain shows low similarity to the RAB15 RGY-rich domain studied here, it appears to be represented by Tyr-$(Xaa)_h$-$(Arg)_k$-$(Xaa)_l$, as seen in RAB15. Moreover, the arginine and tyrosine residues in the sex-lethal protein interact directly with ribonucleic acid of RNA (*51*). Thus, it is possible to speculate that the loop structures of the RAB15 RGY-rich domain contribute to RNA-binding activity.

CONCLUSION

Sequence analysis of plant GRRBPs revealed that the RGY-rich domains contain multiple repeats of Tyr-$(Xaa)_h$-$(Arg)_k$-$(Xaa)_l$, where Xaa is mainly Gly, "k" is 1 or 2, and "h" and "l" range from 0 to 10. The CD study indicated that the tyrosine residues in the two peptides, **1** and **2**, are in ordered structures. The NMR study indicated that side chain-side chain interactions occur between Arg (i) and Asp ($i + 2$) and between Arg and Tyr interrupted by 3 to 4 residue glycine-rich regions. We speculate that these interactions form the loop structures in the RGY-rich domains.

We thank Masakatsu Kamiya for help with the CD experiment. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan (to N.M.).

REFERENCES

1. Alba, M.M. and Pagès, M. (1998) Plant proteins containing the RNA-recoginition motif. *Trends Plant Sci.* **3**, 15–21
2. Sachetto-Martins, G., Franco, L.O., and de Oliveira, D.E. (2000) Plant glycine-rich proteins: a family or just proteins with a common motif? *Biochim. Biophys. Acta* **1492**, 1–14
3. Burd, C.G. and Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615–621
4. Nishiyama, H., Higashitsuji, H., Yokoi, H., Itoh, K., Danno, S., Matsuda, T., and Fujita, J. (1997) Cloning and characterization of human CIRP (cold-inducible RNA-binding protein) cDNA and chromosomal assignment of the gene. *Gene* **204**, 115–120
5. Derry, J.M., Kerns, J.A., and Francke, U. (1995) RBM3, a novel human gene in Xp11.23 with a putative RNA-binding domain. *Hum. Mol. Genet.* **4**, 2307–2311
6. Sato, N. (1994) A cold-regulated cyanobacterial gene cluster encodes RNA-binding protein and ribosomal protein S21. *Plant Mol. Biol.* **24**, 819–823
7. Alba, M.M., Culianezmacia, F.A., Goday, A., Freire, M.A., Nadal, B., and Pagès, M. (1994) The maize RNA-binding protein, MA16, is a nucleolar protein located in the dense fibrillar component. *Plant J.* **6**, 825–834
8. Moriguchi, K., Sugita, M., and Sugiura, M. (1997) Structure and subcellular localization of a small RNA-binding protein from tobacco. *Plant J.* **12**, 215–221
9. Nagai, K., Oubridge, C., Ito, N., Avis, J., and Evans, P. (1995) The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. *Trends Biochem. Sci.* **20**, 235–240
10. Prusiner, S.B. (1998) Prions. *Proc. Natl Acad. Sci. USA* **95**, 13363–13383
11. Yoshida, H., Matsushima, N., Kumaki, Y., Nakata, M., and Hikichi, K. (2000) NMR studies of model peptides of PHGGGWGQ repeats within the N-terminus of prion proteins: a loop conformation with histidine and tryptophan in close proximity. *J. Biochem.* (*Tokyo*) **128**, 271–281
12. Zahn, R. (2003) The octapeptide repeats in mammalian prion protein constitute a pH-dependent folding and aggregation site. *J. Mol. Biol.* **334**, 477–488
13. Stonehouse, J., Adell, P., Keeler, J., and Shaka, A.J. (1994) Ultrahigh-quality NOE spectra. *J. Amer. Chem. Soc.* **116**, 6037–6038.
14. Tanner, J.E. (1990) Use of the stimulated echo in NMR diffusion studies. *J. Chem. Phys.* **52**, 2523–2526
15. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553
16. Macura, S. and Ernst, R.R. (1980) Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy. *Mol. Phys.* **41**, 95–117
17. Braunschweiler, L. and Ernst, R.R. (1983) Coherence transfer by isotropic mixing—application to proton correlation spectroscopy. *J. Magn. Reson.* **53**, 521–528
18. Davis, D.G. and Bax, A. (1985) Practical aspects of two-dimensional transverse NOE spectroscopy. *J. Magn. Reson.* **63**, 207–213
19. Piantini, U., Sørensen, O.W., and Ernst, R.R. (1982) Multiple quantum filters for elucidating NMR coupling networks. *J. Amer. Chem. Soc.* **104**, 6800–6801
20. Piotto, M., Saudek, V., and Sklenár, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665
21. Kim, Y. and Prestegard, J.H. (1989) Measurement of vicinal couplings from cross peaks in COSY spectra. *J. Magn. Reson.* **84**, 9–13
22. Adell, P., Parella, T., Sanchez-Ferrando, F., and Virgili, A. (1995) Clean selective spin-locking spectra using pulsed field gradients. *J. Magn. Reson.* **B108**, 77–80
23. Geen, H. and Freeman, R. (1991) Band-selective radiofrequency pulses. *J. Magn. Reson.* **93**, 93–141
24. Gomez, J., Sanchez-Martinez, D., Stiefel, V., Rigau, J., Puigdomenech, P., and Pagès, M. (1988) A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. *Nature* **334**, 262–264
25. Veau, B., Oudin, A., Courtois, M., Chenieux, J.C., Hamdi, S., Rideau, M., and Clastre, M. (2000) Cloning of two cDNAs encoding crGRP2 and crGRP3 (accession Nos.AF200323 and AF200322), the first members of the RRM-GRP family in Catharanthus roseus (PGR00–049). *Plant Physiol.* **122**, 1459
26. Fasman, G.D., Foster, R.J., and Beychok, S. (1966) The conformational transition associated with the activation of chymotrypsinogen to chymotrypsin. *J. Mol. Biol.* **19**, 240–253
27. Beychok, S., Armstrong, J.M., Lindblow, C., and Edsall, J.T. (1966) Optil rotatory dispersion and circular dichroism of human carbonic anhydrases B and C. *J. Biol. Chem.* **241**, 5150–5160
28. Day, L.A. (1973) Circular dichroism and ultraviolet absorption of a deoxyribonucleic acid binding protein of filamentous bacteriophage. *Biochemistry* **12**, 5329–5339
29. Ettinger, M.J. (1974) Spectral properties of "non-blue" cupric copper in proteins. circular dichroism and optical spectra of galactosidase. *Biochemistry* **13**, 1242–1247
30. Merutka, G., Dyson, H.J., and Wright, P.E. (1995) Random coil' ^1H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J. Biomol. NMR* **5**, 14–24
31. Pletneva, E.V., Laederach, A.T., Fulton, D.B., and Kostic, N.M. (2001) The role of cation-π interactions in biomolecular association. Design of peptides favoring interactions between cationic and aromatic amino acid side chains. *J. Amer. Chem. Soc.* **123**, 6232–6245
32. 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**, 305–10

33. Mitchell , J.B., Thornton, J.M., Singh, J. and Price, S.L. (1992) Towards an understanding of the arginine-aspartate interaction. *J Mol Biol*. **226**, 251–262
34. Burley, S.K. and Petsko, G.A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23–28
35. Singh, J. and Thornton, J.M. (1985) The interaction between phenylalanine rings in proteins. *FEBS Lett.* **191**, 1–6
36. Burley, S.K. and Petsko, G.A. (1988) Weakly polar interactions in proteins. *Adv. Protein Chem*. **39**, 125–189
37. Samanta, U., Pal, D., and Chakrabarti, P. (1999) Packing of aromatic rings against tryptophan residues in proteins. *Acta Crystallogr. Sect. D* **8**, 1421–1427
38. Chelli, R., Gervasio, F.L., Procacci, P., and Schettino, V. (2002) Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *J. Amer. Chem. Soc.* **124**, 6133–6143
39. Hunter, C.A., Singh, J., and Thornton, J.M. (1991) Pi-pi interactions: the geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol*. **218**, 837–846
40. Scrutton, N.S. and Raine, A.R. (1996) Cation-pi bonding and amino-aromatic interactions in the biomolecular recognition of substituted ammonium ligands. *Biochem. J.* **319**, 1–8
41. Gallivan, J.P. and Dougherty, D.A. (1999) Cation-pi interactions in structural biology. *Proc. Natl Acad. Sci. USA* **96**, 9459–9464
42. Shi, Z., Olson, C.A., and Kallenbach, N.R. (2002) Cation-pi interaction in model alpha-helical peptides. *J. Amer. Chem. Soc.* **124**, 3284–3291
43. Burley, S.K. and Petsko, G.A. (1986) Amino-aromatic interactions in proteins. *FEBS Lett.* **203**, 139–143
44. Mitchell, J.B., Nandi, C.L., McDonald, I.K., Thornton, J.M., and Price, S.L. (1994) Amino/aromatic interactions in proteins is the evidence stacked against hydrogen bonding? *J. Mol. Biol.* **239**, 315–331
45. Tóth, G., Watts, C.R., Murphy, R.F., and Lovas, S. (2001). Significance of aromatic-backbone amide interactions in protein structure. *Proteins* **43**, 373–81.
46. Worth, G.A. and Wade, R.C. (1995). The aromatic-(i + 2) amine interaction in peptides. *J. Phys. Chem.* **99**, 17473–17482
47. Burghardt, T.P., Juranic, N., Macura, S., and Ajtai, K. (2002) Cation-pi interaction in a folded polypeptide. *Biopolymers* **63**, 261–272
48. Kemmink, J. and Creighton, T.E. (1993) Local conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor and their roles in folding. *J. Mol. Biol.* **234**, 861–878
49. Dyson, H.J., Merutka, G., Waltho, J.P., Lerner, R., and Wright P.E. (1992) Folding of peptide fragments comprising the complete sequence of proteins: Models for initiation of protein folding. I. Myohemerythrin. *J. Mol. Biol.* **226**, 795–817
50. Hanano, S., Sugita, M., and Sugiura, M. (1996) Structure and expression of the tobacco nuclear gene encoding RNA-binding protein RZ-1: The existence of an intron in the 3′-untranslated region. *DNA Res.* **3**, 65–71
51. Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y. and Yokoyama, S. (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398**, 579–585