Selection of RNA-binding Peptides Containing an Arg-rich Motif

Naohiko Shimada, Reiko Iwase,† Tetsuji Yamaoka, and Akira Murakami*
Department of Polymer Science and Engineering, Kyoto Institute of Technology, Matsugasaki, Kyoto 606-8585
†Department of Bioscience, Teikyo University of Science and Technology, Yatsuzawa 2525, Umenohara, Kitatome-gun, Yamanashi 109-0193

(Received November 5, 2003; CL-031058)

In order to create RNA-binding peptides, we designed peptides based on the sequence information of the RNA-binding proteins and selected them using affinity chromatography. The five peptides out of 42 peptides were efficiently selected. The selected peptides containing Asn and Arg could recognize the RNA structure with the binding pocket such as the internal loop or G•U wobble pair.

Recently, the specific interaction between peptides and RNAs has been revealed. The BIV Tat peptide recognizes a bulge region of the TAR RNA[1] and the HIV Rev peptide recognizes an internal loop of the RRE RNA[2] with high affinities. It was also reported that the HIV Tat peptide derivative inhibited the gene expression *in vivo*.[3] Therefore, the RNA-binding peptides could be novel gene-regulation drugs. In order to design and create novel RNA-binding peptides, some of the cases including our previous report have adopted the sequence information of the RNA-binding proteins.[4] However, this methodology might be restricted to the known sequence information, which might limit further development in this research field. In this letter, we present a method to determine the novel combination of a peptide and a RNA motif. The method consists of two points. One is to design the peptide sequence. Tan and Frankel recently reported that there exist several RNA-binding proteins with similar sequences to the residue 34–40 (TRQARRN) Rev protein, and the Arg-rich motif formed the α-helical structure.[5] Based on this report, we assumed that $TRX_aX_bRRX_c$, where X represents an unconserved amino acid, could be the key motif of the RNA-binding peptides. That is, amino acids of X positions in the conserved sequence might be tightly correlated with the specificity of the recognition of the RNAs. To choose the amino acids for the X positions, we evaluated the occurrence frequency at the X positions in the RNA-binding proteins using the Protein Identification Resource (PIR). It was found that 198 sequences had the $TRX_aX_bRRX_c$ in 561 RNA-binding proteins. A, Q, and R were frequently found at position X_a, and G, L, N, and R at position of X_c. However, no amino acid was clearly designated in position of X_b. Based on these results, a randomized peptide, $TRX_aX_bRRX_c$, was finally designed as follows; X_a: R, A, or Q, X_b: R, A, Q, N, I, G, or T, and X_c: R or N. R and N for the X_c position were expected to possibly participate in the hydrogen bonding with the nucleobases of the RNA.[6,7] Furthermore, an RGG motif[8,9] (KRGGKRGGK) was added at both the *N*- and *C*-terminals in order to enhance the binding affinity and a tryptophan was added to the *N*-terminal as a fluorescent probe (Figure 1).

The second point of this method is related to the selection of the peptides. We adopted the affinity chromatography using a mixture of 16S- and 23S-rRNA from *E. coli*. The RNAs are comprised of many stem-loops, bulges, internal loops, which are reportedly the binding sites of the peptides. The affinity column was prepared by adding the mixture of rRNAs solution to a Sephadex A-25 (DEAE type). Almost all added rRNAs were anchored to the resin. The randomized peptides were then added to the column (RNA-column). As a control, Sephadex A-25 without rRNA (Control-column) was used. Almost all peptides were passed through the Control-column. On the other hand, about 20% of the peptides was retained in the RNA-column after washing with the buffer (10 mM Tris-HCl (pH = 7.5), 100 mM NaCl, 1 mM $MgCl_2$). It was found that the retained peptides were eluted using a buffer containing 0.2–1.0 M NaCl. The molecular weights of the peptides contained in the elution fraction were analyzed by MALDI–TOF MS. As shown in Figure 2, four major peaks were observed, while 15 major peaks were observed from the mixture of the randomized peptides. It was found that the four peaks corresponded to six peptides as summarized in Table 1. These results suggest that the RNA-column protocol was successful in selecting the rRNA-binding peptides from the randomized peptide pool. However, as the molecular weights of pep3 and pep4 are the same, it was not possible to assign the peptide based on the peak (c). Pep5 and pep6 are in the same situation (peak (d)). In order to evaluate the binding ability of the peptides to rRNA, the six peptides were individually synthesized, and the filter binding assay was carried out. As shown in Figure 3, five peptides (pep1–5) are clearly bound to rRNA with a higher affinity than pep6. These results indicat that pep6 was not included in the eluted peptides. That is, the sequences, QR (pep5) and RQ (pep6) in X_aX_b, in the peptides affected the large difference in their binding ability, and Q might be the key amino acid in the interaction. In fact, the peptides coded Q in X_b were not selected. Moreover, pep9, which encoded R at all X positions, was not selected. Both pep7 and pep8 that possess the same positive charges as pep3, pep4, and pep5 were not

NH_2-WKRGGKRGGK **$TRX_aX_bRRX_c$**KRGGKRGGK-$CONH_2$

Figure 1. Randomized peptide sequence. X_a: R, A, or Q, X_b: R, A,Q, N, I, G, or T, X_c: R or N.

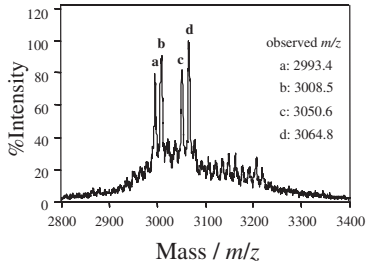

Figure 2. MALDI–TOF MS of 0.4–1.0 M NaCl elution fraction.

Table 1. Molecular weights of corresponding peptides with appearing four peaks by MALDI–TOF MS

Peptide	[a]Peak No.	Observed m/z	[b]Sequence	Calculated m/z
pep1	a	2993.6	TR**RG**RR**R**	2993.6
pep2	b	3008.5	TR**RN**RR**N**	3008.5
pep3	c	3050.6	TR**RN**RR**R**	3050.6
pep4	c	3050.6	TR**RR**RR**N**	3050.6
pep5	d	3064.8	TR**QR**RR**R**	3064.6
pep6	d	3064.8	TR**RQ**RR**R**	3064.6
pep7	-	-	TR**RT**RR**R**	3037.6
pep8	-	-	TR**RI**RR**R**	3049.7
pep9	-	-	TR**RR**RR**R**	3092.7

[a]Peak No. are indicated at the top of the peaks in Figure 2. [b]Bold text indicates the position of X. Pep7, pep8, and pep9 did not appeared in MALDI–TOF MS analysis.

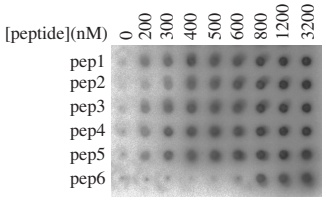

Figure 3. Filter binding assay of each peptides with rRNA. ^{32}P-3′-labeled rRNA: 6 nM, nitrocellulose filter: OPTITRAN BA-S 83 (Schleicher & Schuell), passed through the filter: room temperature, washing: 37 °C 500 μL buffer 2 times.

also selected. It was suggested that the peptides were not merely isolated by the coulombic interaction. In the five selected peptides, R was frequently found in the position of X_a, and R and N were found in the position of X_b. This result indicats that the selected peptides might interact with the nucleobases via hydrogen bonding.

In order to determine the binding sites of the selected peptides, the primer extension analysis using reverse transcriptase (RT) was performed. The RT reaction could be terminated by the RNA-binding molecules, and the resulting truncated cDNA could designate the binding sites. In order to suppress non-specific binding of the peptides to rRNA, a large molar excess of tRNA was used. Figure 4 shows the cDNA patterns in the presence of the peptides. In the absence of the peptide, some bands due to the tertiary structure of the rRNA or to the modified bases of the rRNA appeared. In the presence of pep1–5, additional bands appeared as indicated by the arrows. Pep6 did not produce any additional bands. The binding regions on the rRNAs determined by the length of the truncated cDNA are shown in Figures 4B and 4D. The common features of the binding sites were the presence of the G•U wobble pairs or apical loops. Furthermore, a G–C pair was juxtaposed with the G•U wobble pairs or with an apical loop of the binding sites. The RNA-binding molecules are often required for the binding pocket such as bulges, internal loops or single loops. Furthermore, RNA containing a G•U wobble pair and a loop could make the binding pocket for the RNA-binding molecules.[9–11] The G•U wobble pair of the binding sites for the selected peptides might provide the binding pocket for access of the peptides, and the peptides might recognize the G–C pair.

In conclusion, we succeeded in selecting the rRNA-binding peptides by affinity chromatography using the rRNA-anchoring

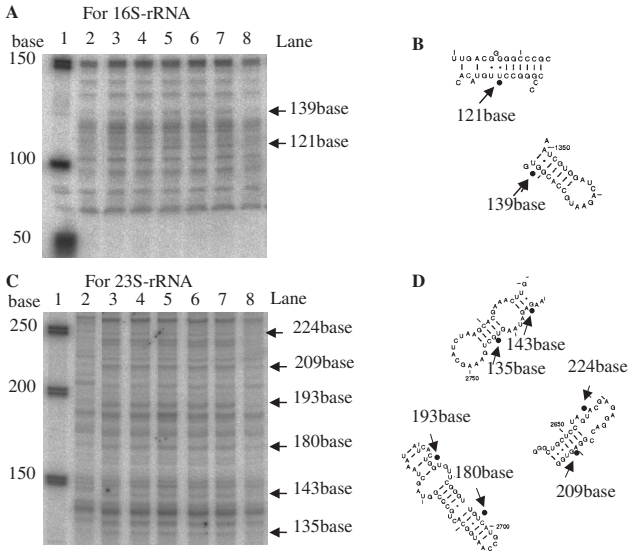

Figure 4. The primer extension assay and binding sites of the peptides for 16S-rRNA (A, B) and 23S-rRNA(C, D). ^{32}P-5′-labeled primer: for 16S-rRNA 5′-CGGTTACCTTGTTACG-3′ (complementary to the site 1496–1511) for 23S-rRNA 5′-GGTCTATTATGACCG-3′ (complementary to the site 2876–2891). RT reaction: [peptide] = 4 μM, [rRNA] = 13 nM, [tRNA] = 5 μM, 20 °C, 2 h, 100 units ReverTra Ace(TOYOBO), gel: 10% denaturing polyacrylamide gel, Lane1: DNA molecular weight marker XIII (Roche Diagnostics), Lane2: no peptide, Lane3–8: pep1, pep2, pep3, pep4, pep5, pep6, respectively. Filled circles: terminated sites.

column. Several peptides, which could bind to rRNA, were effectively selected. This result may lead to the design of new RNA-binding molecules. A detailed study to clarify the effect of the amino acid residues on the RNA recognition is now under way.

This work was partly supported by the Grant-in-Aid for Scientific Research (No. 14580609, AM, TY, RI), and Grants for Regional Science and Technology Promotion (AM) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1 J. D. Puglisi, L. Chen, S. Blanchard, and A. D. Frankel, *Science*, **270**, 1200 (1995).
2 J. L. Battiste, H. Mao, N. S. Rao, R. Tan, D. R. Muhandiram, L. E. Kay, A. D. Frankel, and J. R. Williamson, *Science*, **273**, 1547 (1996).
3 I. Hug, Y.-H. Ping, N. Tamilarasu, and T. M. Rana, *Biochemistry*, **38**, 5172 (1999).
4 N. Shimada, R. Iwase, T. Yamaoka, and A. Murakami, *Polym. J.*, **6**, 507 (2003).
5 R. Tan and A. D. Frankel, *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 5282 (1995).
6 C. Oubridge, N. Ito, P. R. Evans, C.-H. Teo, and K. Nagai, *Nature*, **372**, 432 (1994).
7 J. D. Puglisi, R. Tan, B. J. Calnan, A. D. Frankel, and J. R. Williamson, *Science*, **257**, 76 (1992).
8 M. Kiledjian and G. Dreyfuss, *EMBO J.*, **11**, 2655 (1992).
9 M. Frugier and P. Schimmel, *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 11291 (1997).
10 G. Varani and W. H. McClain, *EMBO Rep.*, **1**, 18 (2000).
11 M. Ruff, S. Krishnaswamy, M. Boeglin, A. Poterszman, A. Mitschler, A. Podjarny, B. Rees, J. C. Thierry, and D. Moras, *Science*, **252**, 1682 (1991).

Published on the web (Advance View) January 26, 2004; DOI 10.1246/cl.2004.212