

## Design of RNA-Binding Oligopeptides Based on Information of RNA-Binding Protein

Naohiko SHIMADA, Reiko IWASE, Tetsuji YAMAOKA, and Akira MURAKAMI<sup>†</sup>

*Department of Polymer Science and Engineering, Kyoto Institute of Technology, Matsugasaki,  
Kyoto 606–8585, Japan*

(Received January 24, 2003; Accepted April 2, 2003)

**ABSTRACT:** In this study, we demonstrated the methodology to design RNA-recognizing oligopeptides. The sequences of the oligopeptides were selected from the information of polypyrimidine tract binding protein (PTB). The oligoRNA sequence was chosen from the hepatitis C virus internal ribosome entry site (HCV IRES) region, where PTB is supposed to interact. The interaction between the oligopeptides and oligoRNA was examined by UV, fluorescence, and circular dichroism (CD) spectroscopy. Oligopeptides which can interact with oligoRNA in a specific manner were explored.

**KEY WORDS** Oligopeptide / Polypyrimidine Tract Binding Protein (PTB) / RNA-Binding Oligopeptide / Fluorescence /

In the pathways of the life cycles of HCV and human immunodeficiency virus (HIV), it has been revealed that some proteins may play important roles in the expression of gene functions. A typical example is observed in the polypyrimidine tract binding protein (PTB or hnRNP I), which is a 57kDa protein. PTB was identified as a nuclear localized splicing factor that specifically binds to the polypyrimidine tract within the pre-mRNA introns with a high affinity ( $K_d \approx 10^{-8}$  M).<sup>1,2</sup> Furthermore, PTB was identified as one of the additional factors in translation initiation. PTB stimulates the IRES (internal ribosomal entry site)-dependent initiation of translation in some virus systems.<sup>3,4</sup> In ribosomes, various ribosomal proteins interact with ribosomal RNA. Recently, X-ray crystallography and nuclear magnetic resonance (NMR) studies have revealed the tertiary structures of the bacterial ribosome-rRNA complex.<sup>5,6</sup> Information from the protein-RNA interaction implies that there might exist certain oligopeptide motifs that can recognize RNA in a sequence specific manner.

Based on the implication, the trials to design oligopeptides that can recognize RNA have been developed in this decade.<sup>7,8</sup> The efforts are of interest not only in understanding the detailed mechanisms of the PTB-RNA interaction, but also in designing novel RNA-recognizing molecules. Especially, the latter may propose a novel concept for the development of new drugs. That is, such oligopeptides may act as competitive reagents to PTB in the regulation of the RNA functions. Similar attempts have been reported by Rana *et al.*<sup>9</sup> They have intensively studied the design of oligopeptides that recognize the trans-activation-

responsive sequence in HIV. This study showed the feasibility of oligopeptides to recognize the tertiary folded structure of RNA. In spite of these situations, the method to design such oligopeptides has not been well developed. As to DNA-binding proteins, details of the interaction have been analyzed with angstrom resolution. The reason why studies of RNA-binding proteins remain undeveloped might be attributed to the diverse tertiary structure of RNA. Only recently, Steitz's group reported their pioneering works using an X-ray crystallographic technique.<sup>5</sup> These studies revealed the tertiary structure of the large subunit of a bacterial ribosome and a large amount of the related information have been accumulated to date. Nevertheless, the difficulty in developing novel RNA-binding peptides still remains.

In this study, we describe the trial to design RNA-recognizing oligopeptides. Not like other trials, we dared not to adopt combinatorial chemistry to search for such oligopeptides. Instead, our design strategy is based on information from the system between PTB and mRNA in HCV. The sequence of the oligopeptide was selected from that of the PTB and it was modified by adding other functional peptide sequences. The oligoRNA sequence was also derived from the HCV IRES region, and some substitution of the bases was also done. Some interesting features are disclosed.

### EXPERIMENTAL

*Sample Preparations for CD and Fluorescence Measurements.* The oligopeptides and oligoRNAs with the designated concentrations in 20 mM phosphate

<sup>†</sup>To whom correspondence should be addressed (E-mail: akiram@ipc.kit.ac.jp).

buffer, 0.1 M NaCl (pH 7.0) were incubated at 25°C for 30 min prior to mixing. The mixing solutions were allowed to stand in the cuvette at 25°C for 5 min.

Reversed phase HPLC was done by using LC-10, SPD-6AV, SLC-6A (Shimadzu Corp., Japan) equipped with a reversed phase column (Capcel Pak; C18, 3.9 mm × 300 mm, 4.5 mm × 150 mm, Shiseido Inc. Japan) using acetonitrile (AN) containing 0.1% trifluoroacetic acid (TFA) for the oligopeptide at an AN linear gradient from 5% to 65% for 30 min or 0.1 M triethylammonium acetate (pH 7.0) for the oligoRNA at a linear gradient from 5% to 20% AN for 15 min.

The oligoRNAs were synthesized using an ABI 391 DNA synthesizer (Applied Biosystems, Foster, CA) with rA<sup>bz</sup>, rC<sup>bz</sup>, rG<sup>ibu</sup>, and rU β-cyanoethyl-*N,N*-diisopropylaminophosphoramidites protected with 5'-*O*-DMTr and 2'-*O*-TBDMS (Millipore, Bedford, MA). The cleavage from the controlled pore glass support and the deprotection of the base protecting groups were achieved *in situ* by adding 3:1 (v/v) 28% ammonium hydroxide/EtOH at 55°C for 14 h. The filtrate was evaporated to dryness. Deprotection of TBDMS was done in 1 M tetrabutylammonium fluoride in tetrahydrofuran (THF) (Sigma-Aldrich Corp., St. Louis, MO) for 24 h at room temperature. The solution was desalted using a Sephadex G-25 (Amersham Pharmacia Biotech, Uppsala, Sweden) and was subjected to reversed phase HPLC for the final purification. The purity of oligoRNAs was confirmed by 10% denaturing polyacrylamide gel electrophoresis.

Oligopeptides were synthesized on a 9050 Plus Pepsynthesizer (Millipore, Bedford, MA) using Fmoc amino acids (Calbiochem-Novabiochem Corp., San Diego, CA) and *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (PE Biosystems, Warrington, UK) as the activating reagent. The PAL-PEG-PS resin (PE Biosystems, Warrington, UK) was used to retain the carboxy-terminal amide residue. The coupling and capping times were 60 min and 5 min, respectively. The oligopeptides were cleaved and deprotected with 90% TFA/5% thioanisole/3% ethanedithiol/2% *m*-cresole for 2 h at room temperature. After the resin was removed, the filtrate was evaporated to *ca.*, 20% of the volume and was then added to ethyl ether for precipitation of the oligopeptides. The oligopeptides were purified by reversed phase high performance liquid chromatography (HPLC). The Concentration of an oligopeptide was determined based on the absorbance of an intrinsic tryptophan at 280 nm using  $\epsilon = 5690 \text{ M}^{-1}\text{cm}^{-1}$ .<sup>10</sup>

The UV melting profiles were obtained using a Hitachi 2000A spectrophotometer equipped with a programmable thermal controller (Hitachi. Corp., Japan).

The temperature was raised at the rate of 1.0°C min<sup>-1</sup>.

Fluorescence spectra were obtained using a Shimadzu 5300PC fluorescence spectrophotometer (Shimadzu Corp., Japan) equipped with a thermal controller (Emission spectrum:  $\lambda_{\text{exc.}} = 342 \text{ nm}$ , Slit; 5 nm (exc.), 10 nm (em.)). When the titration experiment was carried out by varying the oligoRNA concentration, the Raman light scattering was monitored as an inner control for excitation efficacy to quantitatively evaluate the fluorescence.

The circular dichroism spectra were obtained using a JASCO J-720 CD spectrophotometer (Nihon Bunko Corp., Japan) equipped having a thermal controller with a bandwidth of 1.0 nm. The  $\alpha$ -helicity was calculated for the chain length dependence of helicity by using the mean residual ellipticity at 222 nm.<sup>11</sup> Other detailed conditions for the measurements are described in the figure captions.

**Fluorescence Titrations.** The observed relative fluorescence intensity ( $F_{\text{obs}}$ ) is shown in this study as eq 1.

$$F_{\text{obs}} = F_{\text{free}}([P]/[P]_0) + F_{\text{bound}}([PR]/[P]_0) \quad (1)$$

where  $[P]_0$ ,  $[P]$ , and  $[PR]$  are the concentrations of the total oligopeptide, free oligopeptide and oligopeptide-oligoRNA complex, respectively.  $F_{\text{free}}$  and  $F_{\text{bound}}$  are the relative fluorescence intensities of the free oligopeptide and complex oligopeptide, respectively. The dissociation constant of the oligopeptide-oligoRNA ( $K_d$ ) is defined in terms of the equilibrium molar concentrations of the reactant and product as in eq 2.

$$[PR]/[P]_0 = n[R]/(K_d + [R]) \quad (2)$$

Here,  $n$  and  $[R]$  are the number of binding sites and free oligoRNA, respectively. Eq 2 allows the solution for  $[PR]$ , because each oligopeptide species may exhibit the extent of fluorescence as shown in eq 3.

$$F_{\text{obs}} = (F_{\text{bound}}/2[P]_0)(K_d + n[P]_0 + [R]_0 - \sqrt{\{(K_d + n[P]_0 + [R]_0)^2 - 4n[R]_0[P]_0\}}) \quad (3)$$

Each  $K_d$  was determined by fitting the  $F_{\text{obs}}$  value to eq 3 using a nonlinear least-squares global fitting program (IgorPro 2.03, WaveMetrics Inc.). The fitting of the curve was evaluated as the best-fit parameter to be approximately  $n = 1$ .

## RESULTS AND DISCUSSION

### Design of Oligopeptides

The general concepts to design an oligopeptide that can recognize RNA in a specific manner are of interest in designing novel RNA acting reagents, *i.e.*, RNA regulation molecules. The concepts of the design can be

**Table I.** Sequences of oligopeptides used in this study

oligopeptides	Sequence
peptide-I	RVKILFNKKE NALVQMADGN QAQLAMSHLN GW
peptide-II	<u>PKKKRKV</u> RVKILFNKKE NALVQMADGN QAQLAMSHLN GW
peptide-III	<u>CAAAA</u> <u>KVKIEFNKKE</u> <u>NAEVQMADNG</u> <u>QAAAA</u> W  .....     .....

Each peptide is derived from PTB at residues 366–396. Peptide-II and -III are modified as described in the text. The modified points are indicated as underlined texts. The dashed lines indicate points assuming a salt bridge.

classified into two categories. The one is to choose the sequence from information about RNA-binding proteins.<sup>8</sup> In the case that the target RNA binding domains of RNA-binding protein have been specified, this concept can be adopted to design the peptide sequences. At the same time, it is essential that some amino acids such as Asn, Gln, Lys, and Arg should be contained in peptide sequences. These amino acids appear with high frequency in the sequence of the RNA-binding domains of the RNA-binding protein. The other concept is to choose the sequence from a random oligopeptide pool made by the combinatorial synthesis.<sup>12, 13</sup> This method has been widely adopted to select the RNA-binding peptides. This study concerned with the former category.

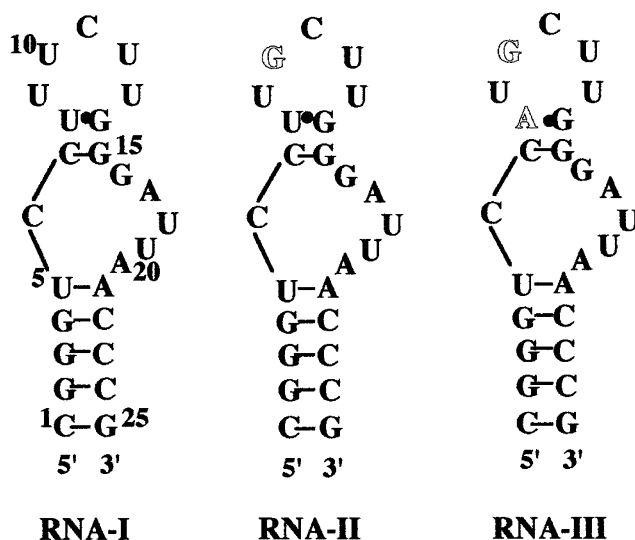
The oligopeptide sequence was selected from the PTB of the HCV IRES binding protein. The PTB consists of four RNA recognition motifs, RRM-1~RRM-4. RRM is a highly conserved sequence in various RNA-binding proteins.<sup>14, 15</sup> Especially, RRM-3 and RRM-4 are reportedly very important for RNA binding.<sup>16–18</sup> Furthermore, based on the mutation analysis, it was revealed that the sequence 366–396 of RRM-3 significantly contributed to the RNA recognition.<sup>16</sup> We selected and synthesized the oligopeptide (peptide-I) to which a tryptophan was added at the *C*-terminal as shown in Table I. The tryptophan was adopted as an inner fluorescent probe. Based on peptide-I, two other peptides were designed. Peptide-II has the PKKKRKV sequence at the *N*-terminal of peptide-I. The addition of the basic oligopeptide sequence might facilitate the interaction with RNA mainly due to the electrostatic interaction between the negatively charged RNA and the oligopeptide. Furthermore, as PKKKRKV is known as a nuclear localization signal, peptide-II could be transferred to nuclei by this signal and then it might interact RNAs. From another point of view, peptide-III was designed as follows. Peptide-I contains seven amino acids with amide residues and four amino acids with basic ones. It is regarded that these residues are responsible for recognizing RNA. Especially, seventh, eleventh (N), and fifteenth (Q) amino acids from *N*-terminal of

peptide-I locate at the (i) and (i+4) positions (Table I). If peptide-I forms an  $\alpha$ -helix, these amide residues could locate in-line outside the helix and might favorably interact the RNA. A similar concept has been reported in designing DNA-binding peptides.<sup>29</sup> Therefore, it is expected that such manipulations of the peptide sequence might facilitate  $\alpha$ -helix formation, and might enforce the ability to interact with RNA. Accordingly, this design included two aspects at one time. The one was the addition of an oligoalanine tract to both the *N*- and *C*-termini of peptide-I. The addition might facilitate oligopeptides to form an  $\alpha$ -helix in aqueous media. The other is the alteration of the amino acids of peptide-I to facilitate it to form salt bridges.<sup>22</sup> That is, 390G and 391N were altered to 390N and 391G, respectively. The *N*-terminal was acetylated and the *C*-terminal was amidated in order to avoid any helix-destabilizing effects.

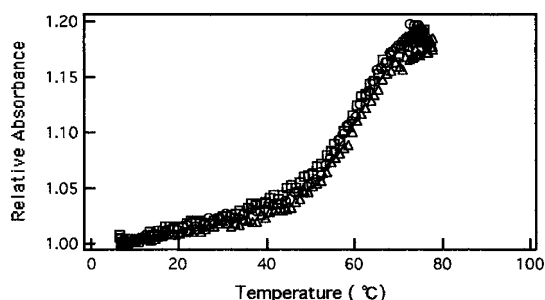
In 35% trifluoroethanol, the circular dichroism spectra of peptide-II showed a 33%  $\alpha$ -helicity. Under physiological conditions, peptide-I, -II, and -III alone showed smaller  $\alpha$ -helicities (6.5, 8.6, and 7.8%, respectively) (Figure 4a). These results may be explained by the large flexibility of the peptides alone in aqueous media.

#### Design of OligoRNA

The oligoRNAs (25-mer) were designed based on the IRES sequence of the HCV mRNA whose secondary structure was proposed.<sup>23</sup> Based on the report that PTB is bound to the 178–202 nt region of the IRES of mRNA HCV,<sup>24</sup> three oligoRNAs were synthesized as shown in Figure 1. RNA-I was from the wild type sequence. RNA-II was designed by the substitution of U10 of RNA-I to G10. RNA-III was designed by the further substitution of U8 of RNA-II to A8. Figure 2 shows the UV melting profiles of the oligoRNAs. They showed similar melting profiles with a single transition temperature, suggesting that oligoRNAs hold the hairpin structure at 25 °C under physiological conditions. The CD spectra of the oligoRNAs indicated that these oligoRNAs hold the typical A-form structure (Figure 4b). Based on these results, it can be assumed that RNA-



**Figure 1.** Sequences and assumed secondary structures of the oligoRNAs. The oligoRNAs were designed based on the HCV IRES RNA (position at 178–202 nt). Mutated points are indicated as outlined text.

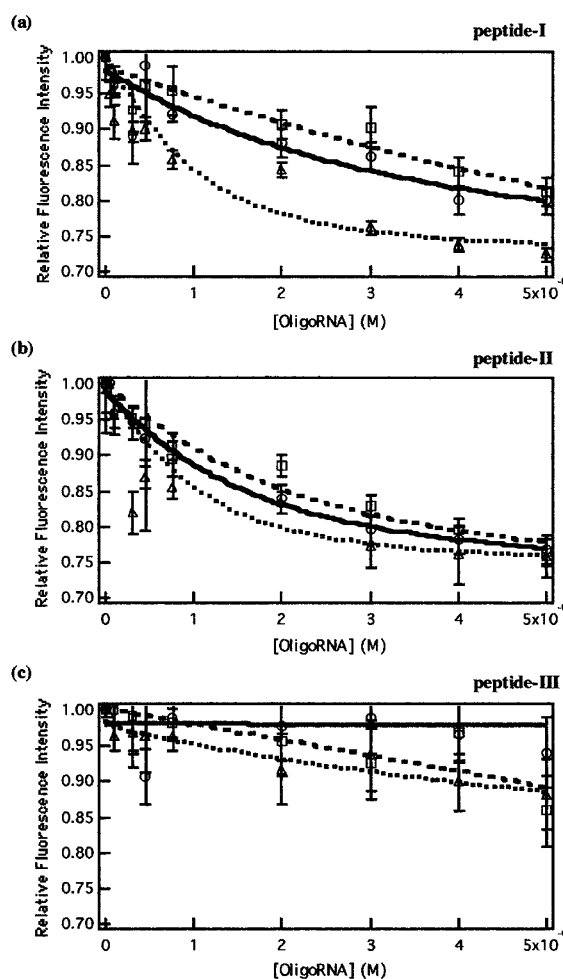


**Figure 2.** Melting profiles for RNA-I (○), RNA-II (□), and RNA-III (△) in 20 mM phosphate buffer (pH 7.0) containing 0.1 M NaCl. The absorbance was monitored at 270 nm: [oligoRNA], 5  $\mu$ M.

I, -II, and -III have a similar hairpin-type 3D-structure with a similar thermal stability.

#### Interaction between Oligopeptide and OligoRNA

It has been reported that the fluorescence intensity of the internal fluorescent residue, Trp, was quenched when the peptide interacts with nucleic acids.<sup>25–28</sup> In our case, the fluorescence intensities of peptide-I to -III decreased as the oligoRNA concentration increased as shown in Figure 3, though the extents of the decrease were varied from 5% to 25%. Dissociation constant ( $K_d$ ) of these systems was estimated from the results according to eq 3 and the  $K_d$ s are summarized in Table II. Peptide-I interacted with RNA-I with a  $K_d = 3.0 \mu$ M. Substitution of U10 of RNA-I to G10 to give RNA-II, which destroyed the pyrimidine tract, significantly increased the  $K_d$  with peptide-I to 13  $\mu$ M, suggesting that U10 might appreciably participate in the interaction. On the other hand, the further substitution of U8 of RNA-II to A8 to give RNA-III, decreased the  $K_d$  with peptide-I to 1.8  $\mu$ M. These results suggest that



**Figure 3.** The fluorescence titration curves for peptide-I (a), peptide-II (b), and peptide-III (c) with RNA-I (—○—), RNA-II (---□---) and RNA-III (···△···) at 25 °C in 20 mM phosphate buffer (pH 7.0) containing 0.1 M NaCl ( $\lambda_{exc}$ , 296 nm,  $\lambda_{em}$ , 355 nm). The data points represent the average of three experiments: [oligopeptide], 1  $\mu$ M.

**Table II.** Dissociation constants ( $K_d$ ) for oligopeptide-oligoRNA binding

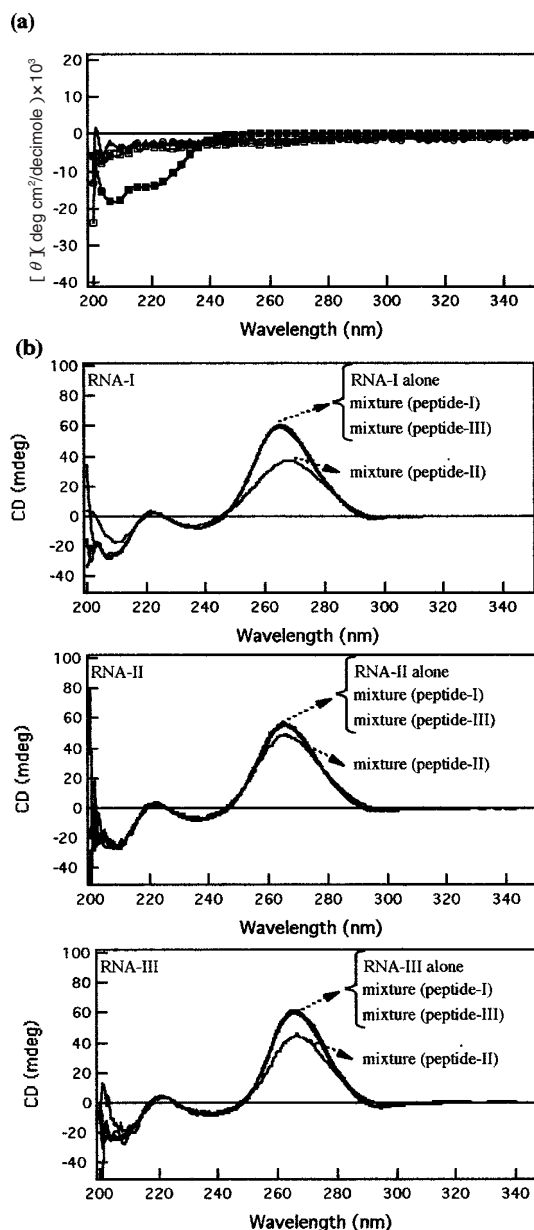
oligoRNAs	$K_d(\mu\text{M})$		
	peptide-I	peptide-II	peptide-III
RNA-I	3.0	1.0	> 20
RNA-II	13	1.5	> 20
RNA-III	1.8	0.4	> 20

Dissociation constants were calculated using eq 3 as described in the EXPERIMENTAL section.

peptide-I interacted with oligoRNA in a specific manner, and that U10 of RNA-I plays an important role in the interaction. In order to evaluate the conformation of peptide, the CD difference spectra were obtained. The CD spectra were almost identical to that of peptide-I alone (data not shown), suggesting that peptide-I did not appreciably form any particular tertiary structure in the presence of the oligoRNA. The CD spectrum of the oligoRNA was unchanged in the presence of peptides-I (Figure 4b), suggesting that the oligoRNA structure maintained the hairpin structure in the interaction with peptide-I. Therefore, it seems that peptide-I might interact with the oligoRNAs by adapting itself to the RNA structure. Curry *et al.* recently reported that RRM-3 formed a certain tertiary structure,  $\beta$ - $\alpha$ - $\beta$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$ , in the absence of RNA.<sup>16</sup> This suggests that the interaction mode of the oligopeptides might differ from that of RNA-binding proteins.

The effect of the modification of peptide-I was examined. Peptide-II interacted more preferably with all oligoRNAs than peptide-I. The basic region is assumed to interact with the RNA without specificity because the mode of interaction is electrostatic. Therefore, the selectivity of peptide-II for each oligoRNA was decreased than peptide-I. Peptide-III did not appreciably interact with the RNAs. The addition of oligoalanine did not provide peptide-I with an  $\alpha$ -helicity (Figure 4a) and the effect of the design to make the peptide form the salt bridges could not be clearly evaluated ( $K_d > 20 \mu\text{M}$ ). The detailed mechanism of the selectivity of the interaction has not been revealed. The CD spectra of these series showed distinguished features compared with the peptide-I series. In all cases, the Cotton effects decreased to 60–70% of those of peptide-I and the oligo RNA mixtures. These results showed that the tertiary structures of the RNA were significantly deformed by the interaction with peptide-II (Figure 4b). It is likely that the basic motif of peptide-II facilitated the binding ability of peptide-II to RNAs and that the binding deformed the RNA structures.

In conclusion, the extraction of the oligopeptide sequences from the RNA-binding proteins was quite successful for developing RNA-binding oligopeptides. Further modification changed the binding ability and



**Figure 4.** (a) CD spectra of peptide-I (○), peptide-II (□), and peptide-III (△) in 20 mM phosphate buffer (pH 7.0) containing 0.1 M NaCl at 25 °C. Filled squares represent peptide-II in same buffer containing 35% TFE: [oligopeptide], 1  $\mu\text{M}$ . (b) CD spectra of RNA-I, RNA-II, and RNA-III in the absence and in the presence of oligopeptides at 25 °C in 20 mM phosphate buffer containing 0.1 M NaCl: [oligopeptide], 1  $\mu\text{M}$ ; [oligoRNA], 5  $\mu\text{M}$ .

the recognition specificity of the sequence and the tertiary structure of the oligoRNA. In general, the methods to develop RNA-acting peptides have not been established. Most successful trials, though the cases were not many, have come from information about the natural RNA-binding proteins. It might be rational to first adopt the method we studied here and then to adopt the combinatorial procedure to further optimize the sequence of the oligopeptide.

*Acknowledgment.* This research was partially supported by a 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, Science, Sports and Culture of Japan.

## REFERENCES

1. A. Ghetti, S. Piñol-Roma, W. M. Michael, C. Morandi, and G. Dreyfuss, *Nucleic Acids Res.*, **20**, 3671 (1992).
2. G. W. Witherell, A. Gil, and E. Wimmer, *Biochemistry*, **32**, 8268 (1993).
3. A. Kaminski, M. T. Howell, and R. J. Jackson, *EMBO J.*, **9**, 3753 (1990).
4. A. Kaminski, S. L. Hunt, J. G. Patton, and R. J. Jackson, *RNA*, **1**, 924 (1995).
5. N. Ban, P. Nissen, J. Hansen, P. B. Moore, and T. A. Steitz, *Science*, **289**, 905 (2000).
6. S. C. Agalarov, G. S. Prasad, P. M. Funke, C. D. Stout, and J. R. Williamson, *Science*, **288**, 107 (2000).
7. A.-M. Baker and D. E. Draper, *J. Biol. Chem.*, **270**, 22939 (1995).
8. M. Schwemmler, M. Gorch, M. Bader, T. F. Sarre, and K. Hilse, *FEBS Lett.*, **251**, 117 (1989).
9. I. Huq, Y.-H. Ping, N. Tamilarasu, and T. M. Rana, *Biochemistry*, **38**, 5172 (1999).
10. H. Edelhoch, *Biochemistry*, **6**, 1978 (1967).
11. N. D. Lazo and D. T. Downing, *Biochemistry*, **36**, 2559 (1997).
12. M. Frugier and P. Schimmel, *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 11291 (1997).
13. J. E. Barrick, T. T. Takahashi, J. Ten, T. Xia, and R. W. Roberts, *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 11291 (2001).
14. K. Nagai, C. Oubridge, T. H. Jesse, J. Li, and P. R. Evans, *Nature*, **348**, 515 (1990).
15. D. W. Hoffman, C. C. Qury, B. L. Golden, S. W. White, and J. D. Keene, *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 2495 (1991).
16. M. R. Conte, T. Grune, J. Ghuman, G. Kell, A. Ladas, S. Matthews, and S. Curry, *EMBO J.*, **19**, 3132 (2000).
17. Y. L. Oh, B. Hahm, Y. K. Kim, H. K. Lee, J. W. Lee, O. K. Song, K. Tsukiyama-Kohara, M. Kohara, A. Nomoto, and S. K. Jang, *Biochem. J.*, **331**, 169 (1998).
18. I. Pérez, J. G. McAfee, and J. G. Patton, *Biochemistry*, **36**, 11881 (1997).
19. V. I. Bruskov, *Mol. Biol. (Moscow)*, **9**, 304 (1975).
20. 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).
21. C. Oubridge, N. Ito, P. R. Evans, C.-H. Teo, and K. Nagai, *Nature*, **372**, 432 (1994).
22. S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 8898 (1987).
23. K. Tsukiyama-Kohara, N. Iizuka, M. Kohara, and A. Nomoto, *J. Virol.*, **66**, 1476 (1992).
24. N. Ali and A. Siddiqui, *J. Virol.*, **69**, 6367 (1995).
25. S. Padmanabhan, W. Zhang, M. W. Capp, C. F. Anderson, and M. T. Record Jr., *Biochemistry*, **36**, 5193 (1997).
26. F. Burn, J.-J. Toulemé, and C. Héleène, *Biochemistry*, **14**, 558 (1975).
27. W. Zhang, J. P. Bond, C. F. Anderson, T. M. Lohman, and M. T. Record, Jr., *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 2511 (1996).
28. M. C. Morris, P. Vidal, L. Chaloin, F. Heitz, and G. Divita, *Nucleic Acids Res.*, **25**, 2730 (1997).
29. B. Y. Wu, B. L. Gaffney, R. A. Jones, and J. W. Taylor, *Peptides*, **105**, 265 (1996).