Identification of Structural Motifs of RNA Aptamers Interacting with the Stem-loop Sequence at the Gag-pol Junction of HIV-1 RNA

Eui-Su Shin, Soon Kang Hong,[+] and Bongrae Cho[*]

Department of Applied Chemistry, Cheongju University, Cheongju 360-764, Korea. [*]E-mail: brcho@cju.ac.kr
[+]Department of Fire Service Administration, Chodang University, Chonnam 534-701, Korea

To identify RNA motifs interacting with the stem-loop RNA which plays an important role in the synthesis of gag-pol fusion protein of HIV-1, a SELEX (Systematic Evolution of Ligands by Exponential Enrichment) was performed. RNA aptamers binding to the stem-loop sequence at the gag-pol junction of HIV-1 RNA were selected from a 30-mer random sequence library through 13 rounds of selection. Some of resulting RNA aptamers contained the consensus sequences AAGGGA and GUUGAU. The stem-loop RNA ligand interacted with synthetic RNA fragment containing two conserved sequence motifs AAGGGA and GUUGAU but did not bind to DNA oligonucleotides having the sequences identical to the RNA fragment. These results suggest that a specific tertiary interaction rather than Watson-Crick base-pairing is involved in the complex formation between RNA aptamer and the stem-loop RNA.

Key Words : HIV-1 RNA, Gag-pol junction, Stem-loop RNA, RNA aptamer, SELEX

Introduction

Human immunodeficiency virus type 1 (HIV-1), like other retroviruses, utilizes translational frameshift which allows the virus to maintain a fixed ratio of Gag protein to Gag-Pol fusion protein in infected cells. Maintenance of the Gag protein/Gag-Pol fusion protein ratio is important for the HIV-1 RNA dimerization and viral efficient assembly of infectious particles.[1,2] This conserved ratio results from a ribosomal frameshifting event which is dependent on downstream RNA stem-loop secondary structure in the genome of the HIV-1.[3,4]

The expression of genetic information in RNA can be regulated by designing DNA or RNA oligonucleotide complementary to target sequence. For example, antisense oligonucleotides were developed to specifically repress complementary target genes.[5,6] But down regulation of specific mRNA with antisense oligonucleotide needs to be studied further because mRNAs have diverse conformers. To overcome this structural problem, SELEX (Systematic Evolution of Ligands by Exponential Enrichment) was applied to isolate RNA aptamers which specifically bind to the hairpin RNA which plays an important role in the synthesis of gag-pol fusion protein of HIV-1 in this work.
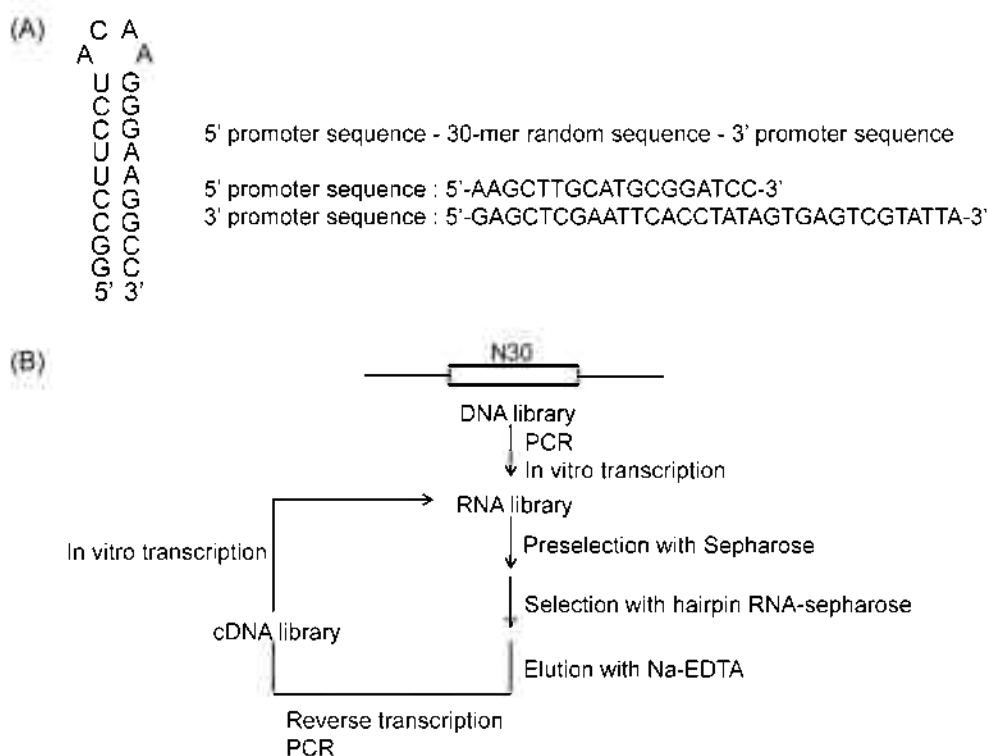
SELEX is a technique for isolating nucleic acid molecules (aptamers) with affinities for a target molecule from a random pool with a large number of sequences by iterative rounds of affinity selection and amplification. Target molecules include proteins, amino acids, nucleotides, antibiotics and RNA.[7-19] In this work, we isolated RNA aptamers which specifically bound to the stem-loop sequence of HIV-1 RNA from a 30-mer random RNA pool. The sequence analysis of these selected RNA aptamers showed that some of them contained conserved sequence motifs AAGGGA and GUU-GAU. The RNA oligonucleotides comprising two conserved sequence motifs were prepared to study the nature of their interaction with the stem-loop RNA in solution. The gel mobility shift assay and the enzymatic protection experiment showed that there was a tertiary interaction between the two RNA molecules.

Experimental Section

Construction of a random RNA library. The oligonucleotide library, 5'-AAGCTTGCATGCGGATCC-(N)30-GAGCTCGAATTCACCTATAGTGAGTCGTATTA-3', contained 30 randomized nucleotides flanked by primer binding site at 5' side and the promoter sequences at 3' side, respectively (Figure 1(A)). This library was amplified through five cycles of polymerase chain reaction (PCR). The library of 1 μg was transcribed *in vitro* with T7 RNA polymerase (RiboMAX™, Promega) in 20 μL of the reaction buffer. *In vitro* transcription products were purified by gel elution of the crush and soak method.[20] The gel-purified RNA was dissolved in 0.2 mL of the binding buffer (30 mM tris-acetate, pH 7.5, 60 mM magnesium acetate, 120 mM potassium acetate, and 120 mM ammonium acetate), incubated at 70 °C for 5 min, and slowly cooled to room temperature.

SELEX protocol. The affinity column for *in vitro* selection was prepared as described.[17] In brief, stem-loop RNA was purified by gel elution from *in vitro* transcribed products. The gel-purified stem-loop RNA was oxidized at the 3'-terminal sugar with NaIO₄ and then coupled to Sepharose-adipic acid hydrazide resin (Amersham Pharmacia Biotech). In order to minimize the enrichment of undesirable RNA species binding to the Sepharose resin itself, we pre-selected the RNA pool on the uncoupled Sepharose-adipic acid hydrazide resin. After being passed through the pre-column of Sepharose-adipic acid hydrazide resin, we then loaded the RNA pool onto the stem-loop RNA-attached affinity column.

(A)

```
      C A
    A   A
    U G
    C G
    C G
    U A
    U A
    C G
    C G
    G C
    G C
    5' 3'
```

5' promoter sequence - 30-mer random sequence - 3' promoter sequence

5' promoter sequence : 5'-AAGCTTGCATGCGGATCC-3'
3' promoter sequence : 5'-GAGCTCGAATTCACCTATAGTGAGTCGTATTA-3'

(B)

```
                                    ┌──────── N30 ────────┐
                                    │                     │
                                    └─────────────────────┘
                                         DNA library
                                            │ PCR
                                            ▼ In vitro transcription
                    ┌──────────────→   RNA library
                    │                       │ Preselection with Sepharose
   In vitro transcription │                 │
                    │                       │ Selection with hairpin RNA-sepharose
                    │                       │
               cDNA library                 │ Elution with Na-EDTA
                    │
           Reverse transcription
           PCR
```

Figure 1. (A) The proposed structure of the stem-loop sequence used for selection and a synthetic library containing 30 randomized nucleotides bound by constant T7 promoter and 3' primer. (B) Outline of the experimental strategy for SELEX. The RNA pool was prepared by *in vitro* transcription of the amplified DNA library carrying a randomized 30-mer DNA region. RNA aptamers binding to a stem-loop RNA were selected with the stem-loop RNA-attached affinity column chromatography. After the 13th round of selection, cDNA of each RNA aptamer was cloned into the pGEM-T Easy vector, and its sequence was analyzed.

We washed the column with binding buffer and then eluted the bound RNA with three column volumes of the elution buffer (25 mM Na-EDTA, pH 8.0). We recovered the selected RNAs by ethanol precipitation and reverse-transcribed it with an M-MLV reverse transcriptase (RT) (Bioneer) (or AMV reverse transcriptase, Promega) using a cDNA primer (5'-AAGCTTGCATGCGGATCC-3'). Then, the cDNAs were amplified by PCR with the cDNA primer and T7 primer (5'-TAATACGACTCACTATAGGTG-3'), and used for the next round of selection. After the 13th round of selection, the amplified cDNAs were cloned into the pGEM-T Easy vector (Promega), and their sequences were determined.

Gel mobility shift assay. Stem-loop RNA was internally labeled by *in vitro* transcription with $[\alpha\text{-}^{32}P]CTP$ (GE Healthcare). The labeled stem-loop RNA and RNA oligonucletides were separately heated in the binding buffer at 70°C for 5 min and slowly cooled to room temperature. The labeled RNA was mixed with 1 μL of RNA oligonucletides (500 pmol/μL) in a final volume of 10 μL and incubated at room temperature for 20 min. The complex was electrophoresed on a 15% non-denaturing polyacrylamide gel in the binding buffer at 40-90 V for 20 h at 4 °C.

Enzymatic protection experiment. RNA aptamer (or stem-loop RNA) was labeled at 5'-end with $[\gamma\text{-}^{32}P]ATP$ (GE Healthcare) and T4 polynucleotide kinase (New England BioLab), and renatured as above. The labeled RNA alone or mixed with the non-labeled cognate RNA was used for nuclease-cleavage reactions. The reaction volume was adjusted to 100 μL with the binding buffer, and incubated for 20 min at room temperature. Then, 0.2 U of S1 nuclease (Boehringer Manheim GmbH), 1 U of RNase T1 (Industrial Research Limited) or 0.1 U of RNase V1 (Pierce Molecular Biology) was added to the above mixture, and the reaction volume included an additional 1 mM $ZnCl_2$ for nuclease S1 cleavage. The reaction mixture was incubated for another 20 min at room temperature. The cleavage products were recovered by ethanol precipitation and electrophoresed on 15% polyacrylamide-7 M urea gels.

Results and Discussion

Selection and conserved sequence motifs of RNA aptamers. RNA molecules binding to the stem-loop RNA which plays an important role in the synthesis of gag-pol fusion protein of HIV-1, were selected from an RNA pool containing 30 randomized nucleotides. An RNA pool for SELEX was prepared by *in vitro* transcription using T7 RNA polymerase from the corresponding double-stranded DNA library of about 10^{14} independent sequences. Selection was performed with the stem-loop RNA-attached column which was prepared by coupling the 3'-terminal sugar of stem-loop RNA to the commercial Sepharose-adipic acid hydrazide resin. RNA library was passed through the affinity
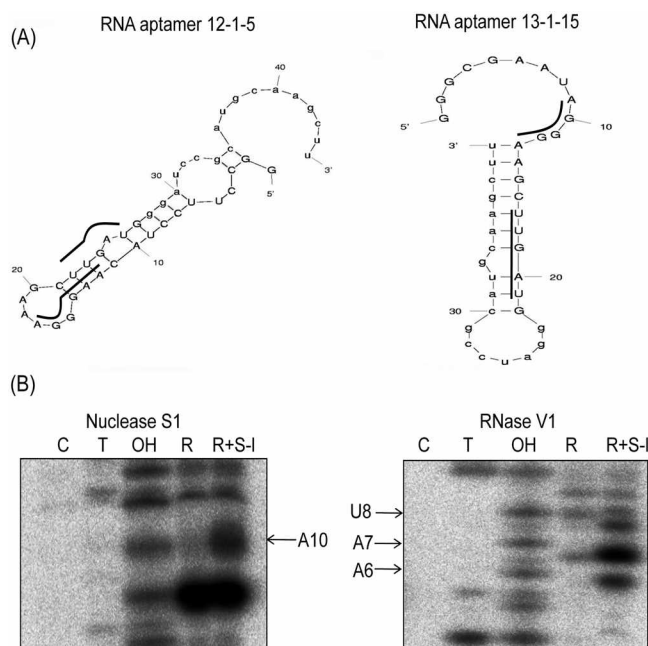
Structural Motifs of RNA Aptamers

Bull. Korean Chem. Soc. **2008**, Vol. 29, No. 11   2229

13-1-15   GGGCGAAU**AGGGA**AGC**UUGAU**

13-1-13   GGGCGGCCCNAAGC**GUUGAU**

13-1-3   GGCCUUCCUAC**AAGGGA**AUAA**GUUGAU**G

12-2-1   GGCCCGCGCAGGAAGC**UUGAUC**

12-1-9   GGCCAC**AGGGA**ANGUAAUACGACUCACUAUAG

12-1-7   **GGGA**UCCGCAUGCAAGCUUACCUCGAUG

12-1-5   GGCCUUCCUAC**AAGGGA**AAGC**UUGAU**G

12-1-4   GGCCUUCCUAC**AAGGGA**UAAGCUGAUC

12-1-1   GGCCAGGCAGAAGCUGCUG

Figure 2. Sequences of the randomized region in RNA aptamers. The consensus sequences are underlined in bold letters.

column and then RNA species bound to the column were eluted by reducing the ionic strength and chelating the Mg^{2+} with EDTA. The selected RNAs were recovered, reverse-transcribed, and then the transcribed cDNAs were amplified by PCR. A new pool of RNA, enriched in the stem-loop RNA-binding motifs, was prepared by transcription from the PCR-amplified DNA. This new RNA pool was used for the next round of selection and amplification. The stringency of the selection was given to lead to a more cohesive set of isolates by reducing the concentration of the stem-loop RNA in the affinity column as the number of selection cycle increased. After the 13th round of selection, the selected RNA aptamers were reverse-transcribed into cDNA, amplified by PCR and cloned into the pGEM-T easy vector for sequence analysis (Figure 1(B)). We searched the consensus sequences of the selected RNA aptamers with CLUSTAL W(1.83) multiple sequence alignment. Some of the selected RNA aptamers had the consensus sequences AAGGGA and GUUGAU (Figure 2).

Secondary structures of RNA aptamers. Optimal predicted secondary structures of RNA aptamers 12-1-5 and 13-1-15, including the constant sequences 5' and 3' to the 30 nt sequence, are shown in Figure 3(A). These structures were found by the MFOLD program accessed on the internet (www.bioinfo.rpi.edu/applications/mfold/old/rna). RNA aptamer 12-1-5 has three single-stranded regions and two double-stranded regions, a long double-stranded region of which has a bulge A25 as detected in RNA aptamer 13-1-3,[21-24] and RNA aptamer 13-1-15 has two single-stranded regions and one double-stranded region with a bulge G33. Conserved sequence motif, AAGGGA was extended in the single- and double-stranded region in the secondary structure models but conserved sequence motif, GUUGAU was in the double-stranded region having a bulge.

Protection of RNA aptamer by stem-loop RNA. Potential regions participating in the interaction between the stem-loop RNA and two RNA aptamers 12-1-5 and 13-1-15 were determined by footprinting using S1 nuclease (specific to single-stranded regions) and RNase V1 (specific to double-stranded regions) (Figure 3B). The residue A10 in RNA aptamer 12-1-5 became more susceptible to nuclease S1 in

Figure 3. (A) Predicted secondary structures of RNA aptamers 12-1-5 and 13-1-15. Primer sequences at the 5' and 3' ends of the transcript are indicated in lower case and unique portions of the molecules are indicated in capital letters. (B) Enzymatic probing of the interaction between the stem-loop RNA and RNA aptamers 12-1-5 and 13-1-15. RNA aptamers labeled at the 5'-end were partially digested with RNase V1 and nuclease S1 in the presence (R + S-I) or absence (R) of stem-loop RNA. Aliquots of enzymatic digestions were loaded onto a denaturing 15% polyacrylamide gel, along with the corresponding partial-alkaline hydrolysate (OH), the denaturing partial RNase T1 (T), and no digestion (C). The protected nucleotides are indicated by arrows.

the presence of a stem-loop RNA than the RNA aptamer alone, suggesting that this nucleotide was liberated from intramolecular interaction and became more accessible to nuclease S1 in solution upon forming a complex with a stem-loop RNA. The region A6AU8 in RNA aptamer 13-1-15 became more susceptible to RNase V1 in the presence of a stem-loop RNA than the RNA aptamer alone, suggesting that this region was stacked or in double-stranded region from intramolecular interaction and became more accessible to RNase V1 in solution upon forming a complex with a stem-loop RNA. However, it is rather surprising that two conserved sequence motifs, AGGGAA and UUGAU in two RNA aptamers were not protected by the stem-loop RNA from any nuclease digestions.

RNA fragments for the interaction with stem-loop RNA. In order to examine whether RNA oligonucleotides derived from two RNA aptamers 12-1-5 and 13-1-15 could bind to the stem-loop RNA, four species of RNA oligonucleotides were purchased from Bioneer. Two species are from RNA aptamer 12-1-5; one was 10-mer RNA oligonucleotide (CU-ACAAGGGA) containing the protected nucleotide A10 and one conserved sequence motif AGGGAA, and the other was 19-mer RNA oligonucleotide (CUACAAGGGAAAGCUU-GAU) containing the protected nucleotide A10 and two conserved sequence motifs AGGGAA and UUGAU. Two

(A)

(B)
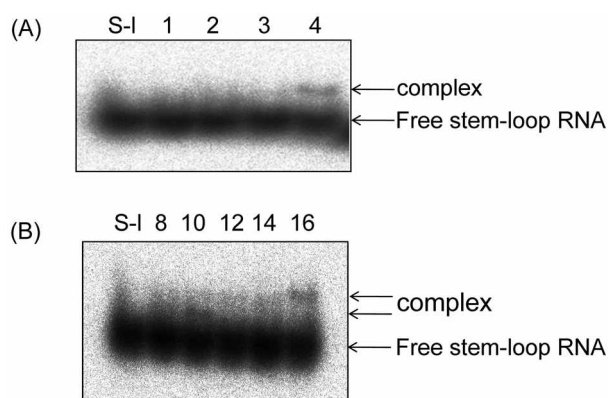
Figure 4. Characterization of binding activities of (A) four RNA oligonucleotides and (B) RNA oligonucleotides derived from the bound 16-mer RNA species for the stem-loop RNA. RNA oligonucleotides were added to the internally labeled stem-loop RNA, and the reaction mixtures were fractionated by 15% polyacrylamide gel electrophoresis under native conditions. The lane marked as (S-l) indicates stem-loop RNA alone without RNA oligonucleotides. (A) Lane 1, 10-mer RNA oligonucleotide (CUACAAGGGA); lane 2, 8-mer RNA oligonucleotide (AAUAGGGA); lane 3, 19-mer RNA oligonucleotide (CUACAAGGGAAAGCUUGAU); lane 4, 16-mer RNA oligonucleotide (AAUAGGGAAGCUUGAU). (B) The figures above each lane indicate the number of nucleotides composing RNA oilgonucleotides. Lane 8, 8-mer RNA oligonucleotides (AAUAGGGA); lane 10, 10-mer RNA oligonucleotides (AAUAGGGAAG); lane 12, 12-mer RNA oligonucleotides (AAUAGGGAAGCU); lane 14, 14-mer RNA oligonucleotides (AAUAGGGAAGCUUG); lane 16, 16-mer RNA oligonucleotides (AAUAGGGAAGCUUGAU).

other species are from RNA aptamer 13-1-15; one was 8-mer RNA oligonucleotide (AAUAGGGA) containing the protected region A6AU8 and one conserved sequence motif AGGGA, and the other was 16-mer RNA oligonucleotide (AAUAGGGAAGCUUGAU) containing the protected region A6AU8 and two conserved sequence motifs AGGG-AA and UUGAU. The interaction of these four RNA oligonucleotides with the stem-loop RNA was examined by gel mobility-shift assay (Fig. 4A). Only 16-mer RNA oligonucleotide derived from RNA aptamer 13-1-15 bound to the stem-loop RNA and other three RNA oligonucleotides did not. These data suggest that the protected region A6AU8 and two conserved sequence motifs AGGGAA and UUGAU in RNA aptamer 13-1-15 are important in the interaction between RNA aptamer and the stem-loop RNA.

To determine how many nucleotides of the bound 16-mer RNA oligonucleotide consist of minimal substrate necessary for binding to the stem-loop RNA, three species of 10-mer, 12-mer and 14-mer RNA oligonucleotides, deleted by every 2 nucleotides from the 3'-terminus of the bound 16-mer RNA were also purchased from Bioneer. Only 10-mer and 16-mer RNA oligonucleotides were found to bind to the stem-loop RNA to form a complex by the gel mobility-shift assay (Fig. 4B), suggesting that a minimal RNA substrate for the binding to the stem-loop RNA is composed of 10 nucleotides, and that 10-mer and 16-mer RNA are changed to different conformers by forming a complex with the stem-loop RNA. However, no interaction between the 8-mer, 12-

mer or 14-mer RNAs, and the stem-loop RNA was observed under the same condition. We don't know yet the reason, which further study is needed for.

Nature of the interaction between the RNA aptamer and the stem-loop RNA. To know whether the interaction between 10-mer RNA or 16-mer RNA and the stem-loop RNA occurs through Watson-Crick base-pairing, a gel mobility-shift assay was performed with 10-mer or 16-mer deoxyoligonucleotides which have identical sequences with the 10-mer or the 16-mer RNA oligonucleotides, respectively. 10-mer or 16-mer deoxyoligonucleotides did not bind to the stem-loop RNA (data not shown), implying that a specific tertiary interaction rather than normal Watson-Crick base-pairing is involved in the complex formation between 10-mer RNA or 16-mer RNA and the stem-loop RNA. However, the nature of this tertiary interaction remains to be demonstrated.

References

1. Park, J.; Morrow, C. D. *J. Virol.* **1991**, *65*, 5111.
2. Shehu-Xhilaga, M.; Crowe, S. M.; Mak, J. *J. Virol.* **2001**, *75*, 1834.
3. Parkin, N. T.; Chamorro, M.; Varmus, H. E. *J. Virol.* **1992**, *66*, 5147.
4. Telenti, A.; Martinez, R.; Munoz, M.; Bleiber, G.; Greub, G.; Sanglard, D.; Peters, S. *J. Virol.* **2002**, *76*, 7868.
5. Stein, C. A.; Krieg, A. M. *Applied Antisense Oligonucleotide Technology*; Stein, C. A., Krieg, A. M., Eds; Wiley: New York, 1998.
6. Crooke, S. T. *Annu. Rev. Med.* **2004**, *55*, 61.
7. Giver, L.; Bartel, D. P.; Zapp, M. L.; Green, M. R.; Ellington, A. D. *Gene* **1993**, *137*, 19.
8. Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550.
9. Burgstaller, P.; Kochoyan, M.; Famulok, M. *Nucleic Acids Res.* **1995**, *23*, 4769.
10. Binkley, J.; Allen, P.; Brown, D. M.; Green, L.; Tuerk, C.; Gold, L. *Nucleic Acids Res.* **1995**, *23*, 3198.
11. Lato, S. M.; Boles, A. R.; Ellington, A. D. *Chem. Biol.* **1995**, *2*, 291.
12. Lauhon, C. T.; Szostak, J. W. *J. Am. Chem. Soc.* **1995**, *117*, 1246.
13. Geiger, A.; Burgstaller, P.; von der Eltz, H.; Roeder, A.; Famulok, M. *Nucleic Acids Res.* **1996**, *24*, 1029.
14. Cho, B.; Taylor, D. C.; Nicholas, Jr. H. B.; Schmidt, F. J. *Bioorg. Med. Chem.* **1997**, *5*, 1107.
15. Haller, A. A.; Sarnow, P. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8521.
16. Wallace, S. T.; Schroeder, R. *RNA* **1998**, *4*, 112.
17. Ko, J.; Cho, B.; Ahn, J. K.; Lee, Y.; Park, I. *Bull. Korean Chem. Soc.* **1999**, *20*, 1335.
18. Ko, J.; Lee, Y.; Park, I.; Cho, B. *FEBS Lett.* **2001**, *508*, 300.
19. Cho, B. *Bull. Korean Chem. Soc.* **2005**, *26*, 2093.
20. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
21. Cho, B. *Bull. Korean Chem. Soc.* **2008**, *29*, 1063.
22. Cho, B. *Bull. Korean Chem. Soc.* **2008**, *29*, 2026.
23. Cho, B. *Bull. Korean Chem. Soc.* **2006**, *27*, 786.
24. Cho, B. *Bull. Korean Chem. Soc.* **2007**, *28*, 689.