ORGANIC LETTERS

2013
Vol. 15, No. 15
3934–3937

Synthesis of 6-Se-Guanosine RNAs for Structural Study

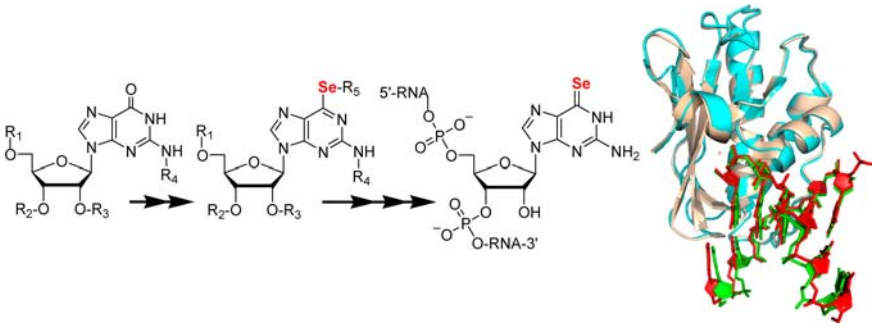
Jozef Salon, Jianhua Gan, Rob Abdur, Hehua Liu, and Zhen Huang*

Department of Chemistry, Georgia State University, Atlanta, Georgia 30303, United States

huang@gsu.edu

Received June 14, 2013

ABSTRACT



6-Se-guanosine phosphoramidite and RNAs have been synthesized by selenium substitution of the 6-oxygen atom, and it is revealed that the Se-derivatization is relatively stable and that bulge and wobble structures can better accommodate a large Se atom than a duplex. This Se-modification is useful in the structural study of RNAs and their protein complexes.

RNA secondary structures and motifs help to diversify structures and functions of noncoding RNAs (ncRNA). Bulges and U/G wobble pairs are frequent structural motifs in ncRNAs.[1,2] Functions of these motifs are revealed by structural study, such as X-ray crystallography, one of the most powerful tools for structure determination of biomacromolecules and complexes.[3,4] Structures of proteins and protein−nucleic acid complexes are commonly determined by selenium-derivatized (i.e., selenomethionyl) proteins via multiwavelength or single-wavelength anomalous diffraction (MAD or SAD) phasing.[5−7] Inspired by the Se-derivatized protein approach, our research group has pioneered and developed selenium-derivatized nucleic acids (SeNA) for the X-ray crystal structure study of nucleic acids and protein−nucleic acid complexes.[8−10]

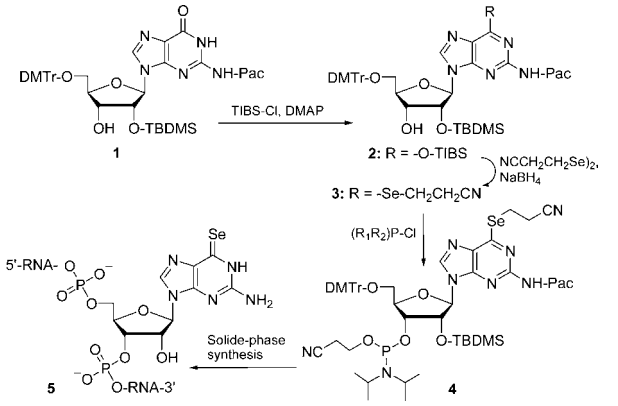
To facilitate the RNA structure study, the Se-atom-specific modification has been developed by us. Herein we report the first synthesis of the 6-Se-guanosine (^{Se}G) phosphoramidite and ^{Se}G-containing RNAs. Since ncRNAs often contain bulges and U/G wobble pairs in the duplex regions,[1,6] the stability of the G/C pairs close to bulge and wobble-pair structures has been investigated, by using the selenium atomic probe. The UV-melting study indicates that, in the presence of the bulge and wobble structures, the native and Se-modified G/C structures have similar stability. This Se atomic probe can be used in studying RNA secondary structures. Moreover, the ^{Se}G-RNA was used as the transient Se-modification to enhance crystallization, although the Se-derivatization is normally used for phase determination. Since crystallization is a major challenge in X-ray crystallography, a Se-RNA/DNA/protein complex was used to explore the crystallization and crystal quality. Particularly, we intended to investigate the crystallization property of RNA transiently modified with

(1) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678–1685.
(2) Macrae, I. J.; Zhou, K.; Li, F.; Repic, A.; Brooks, A. N.; Cande, W. Z.; Adams, P. D.; Doudna, J. A. *Science* **2006**, *311*, 195–198.
(3) Liu, Q.; Dahmane, T.; Zhang, Z.; Assur, Z.; Brasch, J.; Shapiro, L.; Mancia, F.; Hendrickson, W. A. *Science* **2012**, *336*, 1033–1037.
(4) Nakanishi, K.; Weinberg, D. E.; Bartel, D. P.; Patel, D. J. *Nature* **2012**, *486*, 368–374.
(5) Hendrickson, W. A. *Trends Biochem. Sci.* **2000**, *25*, 637–643.
(6) Ferre-D'Amare, A. R.; Zhou, K. H.; Doudna, J. A. *Nature* **1998**, *395*, 567–574.
(7) Schirle, N. T.; MacRae, I. J. *Science* **2012**, *336*, 1037–1040.
(8) Carrasco, N.; Ginsburg, D.; Du, Q.; Huang, Z. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 1723–1734.
(9) Lin, L.; Sheng, J.; Huang, Z. *Chem Soc Rev* **2011**, *40*, 4591–4602.
(10) Sun, H.; Sheng, J.; Hassan, A. E.; Jiang, S.; Gan, J.; Huang, Z. *Nucleic Acids Res.* **2012**, *40*, 5171–5179.

10.1021/ol401698n
Published on Web 07/16/2013

the Se-functionality. Surprisingly, the transient Se-RNA complexed with DNA and RNase H has offered high-quality crystals and resulted in high-resolution structure determination.

Scheme 1. Synthesis of ^{Se}G-RNAs (**5**)



Despite the synthesis of the 6-selenoguanine and its derivatives five decades ago,[11] synthesis of RNA containing the Se-guanine has not been accomplished because of the synthetic challenge, even though 6-S-purines have been introduced into nucleic acids.[12] Recently our laboratory has successfully developed a novel strategy to incorporate the selenium functionality to the 6-position of deoxyguanosine in DNA.[13] This successful strategy has encouraged us to introduce the selenium functionality to the 6-position of guanosine in RNA. Thus the synthesis of the 6-selenoguanosine (^{Se}G) phosphoramidite and ^{Se}G-RNAs is developed and reported.

Our development of the 2-cyanoethyl-seleno protection and deprotection for the Se-nucleobase nucleic acids[10,14] has encouraged us to protect the 6-Se-functionality on guanosine with the 2-cyanoethyl-seleno group. In addition, the 2-cyanoethyl protecting group can be removed under ultramild conditions (0.05 M K_2CO_3 in methanol). Since strong basic conditions can cause deselenization, phenoxyacetyl (Pac) was used as the protecting group for the 2-NH_2 of this 6-Se-guanosine phosphoramidite (**4**), which can also be removed under the ultramild conditions. Our synthesis (Scheme 1) started from the partially protected guanosine derivative (**1**). This commercial compound was activated at the 6-position of the guanine with a selective sulfonylation. To avoid protection of the 3′-hydroxyl group, 2,4,6-triisopropylbenzenesulfonyl chloride (TIBS-Cl) was explored.[13] Without purification of the TIBS-activated intermediate (**2**), the protected 6-Se-guanosine derivative (**3**) was obtained (82% yield in two steps) by the substitution of the sulfonyl-activating group with sodium 2-cyanoethylselenide generated by the reduction of di-(2-cyanoethyl) diselenide with $NaBH_4$. In a satisfactory yield, the 6-Se-guanosine derivative (**3**) was converted to the Se-phosphoramidite (**4**).[10,13] The characterization of these Se-nucleoside derivatives is presented in Figures S1−S10 in the Supporting Information (SI).

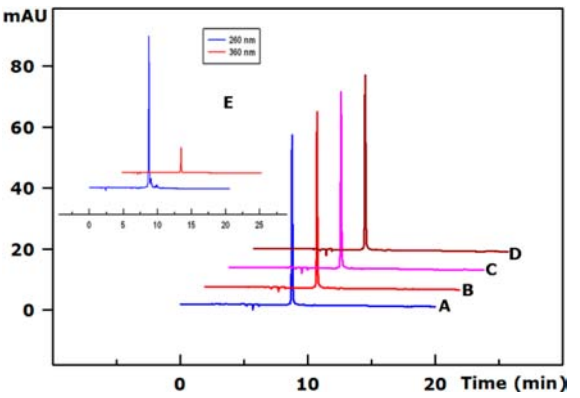


Figure 1. HPLC analysis and thermostability study. The purified RNA samples [^{Se}G-RNA: r(5′-G-^{Se}G-UAUUGCGGUACC-3′); 3 μM each] were heated at 60 °C in 100 mM sodium phosphate buffer (pH 7.6) for 1, 2, and 3 h. (A) RP-HPLC analysis at 360 nm, no heat; (B) 1-h heating (360 nm); (C) 2-h heating (360 nm); (D) 3-h heating (360 nm); (E) 3-h heating (260 nm) in blue line; 3-h heating (360 nm) in red line. HPLC conditions: Welchrom XB-C18 column (4.6 mm × 250 mm); flow rate, 1 mL/min; column temperature, 25 °C; gradient, 5 to 35% B in 10 min. Buffer A: 10 mM TEAAc (pH 7.4). Buffer B: 10 mM TEAAc (pH 7.4) in 60% acetonitrile dissolved in water (6:4 in v/v).

Finally, the ^{Se}G-phosphoramidite was incorporated into RNAs by solid-phase synthesis. The ^{Se}G-RNAs were synthesized using **4**, the ultramild CE phosphoramidites (A, C, and G),[10] and BTT activator. Since strong basic conditions (such as NH_3-H_2O) for the deprotection cause the deselenization of ^{Se}G, the K_2CO_3 deprotection conditions were used to remove all ultramild protecting groups.[10,13] If RNAs have any guanosine nucleotides, phenoxyacetic anhydride (Pac_2O, instead of Ac_2O) is used in the capping step to avoid the guanosine acetylation, which is difficult to remove under K_2CO_3 treatment. After the K_2CO_3 deprotection and fluoride treatment to remove TBDMS groups, the synthesized ^{Se}G-RNAs (**5**) were purified by HPLC twice (with and without the 5′-DMTr group). Interestingly, the deprotected Se-RNAs are yellow colored, which is the first observation of the colored RNAs after a single atom replacement, while natural RNAs are colorless. In addition, to measure the coupling efficiency of the ^{Se}G-phosphoramidite (**4**), 5′-DMTr-^{Se}GG dinucleotide was synthesized, analyzed by RP-HPLC, and compared with the synthesis and analysis of the native 5′-DMTr-GG, which indicated a high coupling yield (over 96%). The purified Se-RNAs were confirmed by HPLC and MS (Table 1). The typical HPLC, MS, and UV profiles of the ^{Se}G-RNAs are shown in Figures 1, S11, and S12.

(11) Mautner, H. G.; Chu, S. H.; Jaffe, J. J.; Sartorelli, A. C. *J. Med. Chem.* **1963**, *6*, 36–39.
(12) Coleman, R. S.; Arthur, J. C.; McCary, J. L. *Tetrahedron* **1997**, *53*, 11191–11202.
(13) Salon, J.; Jiang, J.; Sheng, J.; Gerlits, O. O.; Huang, Z. *Nucleic Acids Res.* **2008**, *36*, 7009–7018.
(14) Salon, J.; Sheng, J.; Jiang, J.; Chen, G.; Caton-Williams, J.; Huang, Z. *J. Am. Chem. Soc.* **2007**, *129*, 4862–4863.

Table 1. MALDI-TOF MS Data of the ^{Se}G-RNAs

entry	Se-RNAs	measured (calcd) m/z[a]
1	5′-UUC-^{Se}G-CG-3′	$[M + H]^+$: 1915
	$C_{56}H_{71}N_{20}P_5O_{41}Se$: FW 1914.1	(1915.1)
2	5′-CCG CGC-^{Se}G-CG dG-3′	$[M + H]^+$: 3237
	$C_{95}H_{121}N_{40}P_9O_{66}Se$: FW 3236.4	(3237.4)
3	5′-UAC UAA C-^{Se}G-U AGU A-3′	$[M + H]^+$: 4173
	$C_{124}H_{153}N_{49}P_{12}O_{87}Se$: FW 4171.5	(4172.5)
4	5′-UAC UAA CGU A-^{Se}G-U A-3′	$[M + H]^+$: 4173
	$C_{124}H_{153}N_{49}P_{12}O_{87}Se$: FW 4171.5	(4172.5)
5	5′-G-^{Se}G-U AUU GCG GUA CC-3′	$[M + H]^+$: 4526
	$C_{133}H_{165}N_{52}P_{13}O_{97}Se$: FW 4524.5	(4525.5)
6	5′-GGU AUU-^{Se}G-CG GUA CC-3′	$[M + H]^+$: 4526
	$C_{133}H_{165}N_{52}P_{13}O_{97}Se$: FW 4524.5	(4525.5)
7	5′-G-^{Se}G-$U_{2'SeMe}$-AUUGCGGUACC$_{2'OMe}$-3′	$[M + H]^+$: 4602
	$C_{135}H_{169}N_{52}P_{13}O_{95}Se_2$: FW 4600.7	(4601.7)
8	5′-G-^{Se}GUA-$U_{2'SeMe}$-UGCGGUACC$_{2'OMe}$-3′	$[M + H]^+$: 4602
	$C_{135}H_{169}N_{52}P_{13}O_{95}Se_2$: FW 4600.7	(4601.7)

[a] 3-Hydroxypicolinic acid (HPA) matrix was used.

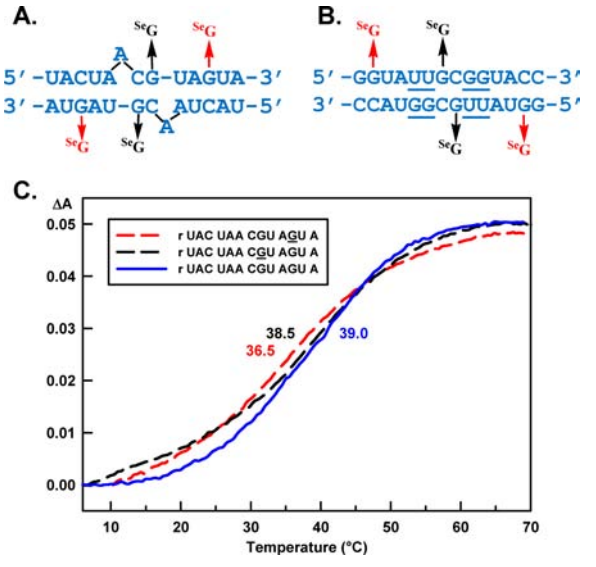


Figure 2. RNA secondary structures and normalized UV-melting curves of RNA **1**-**6** in Table 2. (A) Secondary structures of RNAs: **1** (native in blue), **2** (with black ^{Se}G), and **3** (with red ^{Se}G). (B) Secondary structures of RNAs: **4** (native in blue), **5** (with black ^{Se}G), and **6** (with red ^{Se}G). The underlined sequences are U/G wobble pairs. (C) Blue curve: melting curve of the native RNA containing two adenosine bulges (**1**; T_m = 39.0 °C). Black curve: melting curve of **2** (T_m = 38.5 °C). Red curve: melting curve of **3** (T_m = 36.5 °C). The melting temperatures of the duplexes (2.5 μM) were measured in 10 mM phosphate buffer (pH 6.5) containing 50 mM NaCl. The absorbance of the samples was monitored at 260 nm, and the duplexes were headed from 5 to 70 °C with a rate of 1 °C/min.

Moreover, the HPLC analysis indicated that this Se-functionality in RNA is relatively stable (Figure 1). The ^{Se}G-RNA (5′-G-^{Se}G-UAUUGCGGUACC-3′) was

Table 2. UV-Melting Temperature of the ^{Se}G-RNAs

entry	RNAs	T_m (°C) (ΔT_m per Se)
1	(5′-UACUAAC---G-UA---G-UA-3′)$_2$	39.0 ± 0.1
2	(5′-UACUAAC-^{Se}G-UA---G-UA-3′)$_2$	38.5 ± 0.2 (0.3)
3	(5′-UACUAAC---G-UA-^{Se}G-UA-3′)$_2$	36.5 ± 0.3 (1.3)
4	(5′-G---G-UAUU-G---CGGUACC-3′)$_2$	54.0 ± 0.1
5	(5′-G---G-UAUU-G^{Se}-CGGUACC-3′)$_2$	47.2 ± 0.2 (3.4)
6	(5′-G-^{Se}G-UAUU-G---CGGUACC-3′)$_2$	40.1 ± 0.2 (6.9)

heated at 60 °C for 3 h, and no significant deselenization of the Se-RNA was observed. The UV-melting temperatures of the ^{Se}G-RNAs were measured (T_m, Table 2) to examine the impact of ^{Se}G on the G/C pairs close to various secondary structures (Figure 2A and 2B), such as the bulge, U/G wobble pair, and duplex. Three typical melting curves are shown in Figure 2C. The melting temperatures indicate that the ^{Se}G/C pairs next to the two A-bulges do not cause significant destabilization (0.3 °C per Se modification) of the bulged structure (Figure 2A), while ^{Se}G/C in the duplex structure causes more destabilization (1.3 °C per Se) in the same sequence. In this RNA crystal structure, these two A-nucleotides flip out,[15] and the space generated by the bulges can better accommodate the large Se atoms. Thus, these two Se atoms barely cause any structural destabilization (0.3 °C per Se), which is consistent with the crystal structure study. Similarlly, the ^{Se}G/C close to the U/G wobble pair (Figure 2B) causes less destabilization (3.4 °C per Se) than that in a duplex (6.9 °C per Se). Significant destabilization of RNA duplex by the ^{Se}G/C pair has also been observed in a DNA duplex (decreased by up to 11 °C per Se).[13] These results suggest that the Se-accommodation is dependent on the locations and secondary structures of nucleic acids. Thus, the selenium atomic probe can be utilized to study the secondary structures of RNAs and RNA−protein complexes.

Furthermore, we have found that the selenium derivatization can help reduce the local conformational flexibility of the modified molecules and duplexes.[17,18] Thus we hypothesized that the Se heavy atom may reduce molecular dynamics, thereby facilitating molecular packing and crystallization. Thus, in addition to using Se-derivatization for phase determination, Se-modification was used for crystal growth. To demonstrate the proof of principle, the crystallization of the RNA/DNA/RNase H complex was examined with the Se-derivatized RNA. Initially, crystals were yellow and generated by the yellow ^{Se}G-RNA. The crystal color can be conveniently used to indicate whether RNA is present in a crystal. As shown in Figure 1, the ^{Se}G-RNA is relatively stable. However, during the crystallization over weeks in buffer, the Se-modification was naturally

(15) Berglund, J. A.; Rosbash, M.; Schultz, S. C. *RNA* **2001**, *7*, 682–691.
(16) Nowotny, M.; Yang, W. *EMBO J.* **2006**, *25*, 1924–1933.
(17) Jiang, J.; Sheng, J.; Carrasco, N.; Huang, Z. *Nucleic Acids Res.* **2007**, *35*, 477–485.
(18) Sheng, J.; Zhang, W.; Hassan, A. E.; Gan, J.; Soares, A. S.; Geng, S.; Ren, Y.; Huang, Z. *Nucleic Acids Res.* **2012**, *40*, 8111–8118.

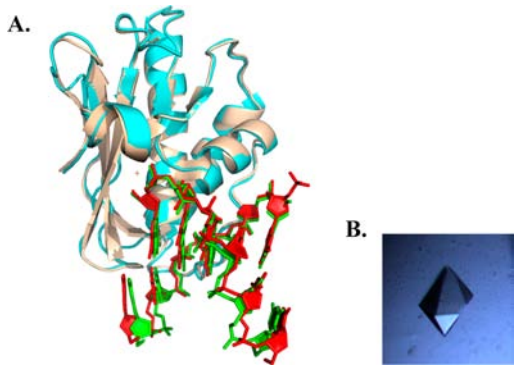


Figure 3. Structure comparison and crystal picture of the RNA/DNA/RNase H complexes. (A) The Se-transient native structure (resolution: 1.60 Å; PDB: 3ULD) is in cyan (protein) and red (RNA/DNA duplex), while the direct native structure (resolution: 2.70 Å; PDB: 2G8U) is in gray (protein) and green (RNA/DNA duplex).[16] The sequences of the RNA (5′-UCGACA-3′) and DNA (5′-ATGTCG-3′) are the same in these two structures. (B) Picture of the crystal ($0.2 \times 0.2 \times 0.2\ mm^3$; slightly yellow) of the transient Se-RNA/DNA/RNase H complex.

removed (by hydrolysis) from the Se-modified RNA. In addition to the time, the crystallization pH (6.0; SI) probably also played a critical role in the deselenization. Due to the deselenization, the crystal's yellow color gradually became colorless. The replacement of the selenium atoms by oxygen generated the native crystals eventually. The loss of selenium was confirmed by MS analysis of the crystals of the Se-RNA [5′-UC(^{Se}G)ACA-3′] complexed with DNA (5′-ATGTCG-p-3′, one-base overhang at 5′-end) and RNase H. Surprisingly, the transient Se-RNA complexed with DNA and RNase H has offered high quality crystals. The crystal structure was determined by the molecular replacement, and the structure assisted by the transient Se-modification offered high resolution (PDB ID: 1.60 Å; PDB: 3ULD), while the native crystals grown with the corresponding native RNA complexed with DNA and RNase H did not offer high resolution (only 2.70 Å). These two native structures from directly and transiently grown crystals are virtually identical (Figure 3). The high resolution offered by the transient Se-derivatization can probably be attributed to the reduced molecular dynamics and better molecular packing assisted by the Se heavy atom initially. This Se-modification may have great potential in investigating structures of RNAs and protein−RNA complexes.

In summary, the 6-Se-guanosine (^{Se}G) phosphoramidite and ^{Se}G-derivatized RNAs have been synthesized. The Se-modification is relatively stable, and the bulge and wobble structures can better accommodate the large Se atom than a duplex structure. This Se atomic probe can be used in studying RNA secondary structures, and the yellow color of the ^{Se}G-RNA is useful for protein−RNA cocrystallization visualization. Moreover, we have demonstrated the proof of principle that this novel Se-derivatization can be used for transiently Se-assisted molecular packing and crystallization for high-resolution structure determination. Our exciting discoveries will open a new research avenue in structure and function studies of ncRNAs. This Se probe also offers a useful tool for the structural study of RNAs as well as their protein complexes.

Acknowledgment. We would like to thank the beamline staff at Advanced Light Source beamline 8.2.1, in Lawrence Berkeley National Laboratory, for their help in the data collection. Plasmids expressing RNase H proteins were kindly given to us by Dr. Wei Yang at the NIH. This work was financially supported by the NIH (R01GM095881) and the Georgia Cancer Coalition (GCC) Distinguished Cancer Clinicians and Scientists program.

Supporting Information Available. Experimental procedures, ^{1}H and ^{31}P NMR, HRMS and MALDI-TOF MS analytical data, HPLC profiles. These materials are available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.