Synthesis of a 2′-Se-thymidine Phosphoramidite and Its Incorporation into Oligonucleotides for Crystal Structure Study
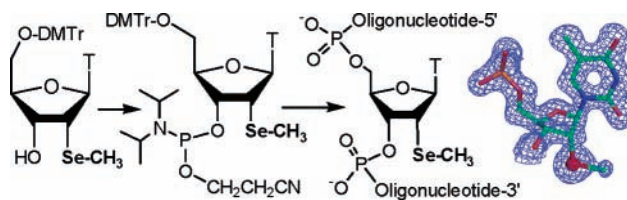
Jia Sheng, Jiansheng Jiang, Jozef Salon, and Zhen Huang*

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

huang@gsu.edu

ABSTRACT

To investigate nucleic acids with selenium derivatization for crystallography, we report the first synthesis of 2′-methylseleno-thymidine phosphoramidite and its incorporation into DNAs and RNAs by solid-phase synthesis with over 99% coupling yield. The d(GT$_{Se}$GTACAC)$_2$ crystal structure was also determined at 1.40 Å resolution using Se phasing, revealing that this Se derivatization did not cause significant structure perturbation, consistent with our UV melting study. In addition, we observed that the Se modification largely facilitated the crystallization.

Nucleic acids play key roles in genetic information storage, expression, and translation and have attracted intense research attention for a long time.[1−4] X-ray crystallography has been widely used to study the structure and function of nucleic acid molecules.[5,6] To facilitate crystal structure determination of nucleic acids, especially using multiwavelength anomalous dispersion (MAD) or single-wavelength anomalous dispersion (SAD) phasing,[7] Huang, Egli, and co-workers were the first to develop the selenium derivatization strategy.[8−13] So far, the 2′-Se derivatization has been introduced to U and C by us and to A and G by other groups.[9−15] Using the Se

derivatization and phasing, research laboratories have already reported several crystal structures of DNA and RNA molecules,[10,14−16] including the homo-DNA structure, the phasing of which otherwise proved extremely challenging.[16]

Natural tRNAs and rRNAs are heavily modified, and over 100 different modifications have been observed.[17] Ribo-

(1) Nougayrède, J.-P.; Homburg, S.; Taieb, F.; Boury, M.; Brzuszkiewicz, E.; Gottschalk, G.; Buchrieser, C.; Hacker, J.; Dobrindt, U.; Oswald, E. *Science* **2006**, *313*, 848.

(2) Adams, P. L.; Stahley, M. R.; Kosek, A. B.; Wang, J.; Strobel, S. A. *Nature* **2004**, *430*, 45.

(3) Benoff, B.; Yang, H.; Lawson, C.; Parkinson, G.; Liu, J.; Blatter, E.; Ebright, Y. W.; Berman, H. M.; Ebright, R. H. *Science* **2002**, *297*, 1562.

(4) Watson, J. D.; Crick, G. H. C. *Nature* **1953**, *171*, 737.

(5) Lima, L. M. T. R.; Cordeiro, Y.; Tinoco, L. W.; Marques, A. F.; Oliveira, C. L. P.; Sampath, S.; Kodali, R.; Choi, G.; Foguel, D.; Torriani, I.; Caughey, B.; Silva, J. L. *Biochemistry* **2006**, *45*, 9180.

(6) (a) Ferre-D'Amare, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567. (b) Egli, M. *Curr. Opin. Chem. Biol.* **2004**, *8*, 580. (c) Holbrook, S. R.; Kim, S. H. *Biopolymers* **1997**, *44*, 3.

(7) (a) Yang, W.; Hendrickson, W. A.; Crouch, R. J.; Satow, Y. *Science* **1990**, *249*, 1398−1405. (b) Hendrickson, W. A.; Ogata, C. M. *Methods Enzymol.* **1997**, *276*, 494. (c) Liu, D.; Lepore, B. W.; Petsko, G. A.; Thomas, P. W.; Stone, E. M.; Fast, W.; Ringe, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11882.

(8) Carrasco, N.; Ginsburg, D.; Du, Q.; Huang, Z. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 1723.

(9) Du, Q.; Carrasco, N.; Teplova, M.; Wilds, C. J.; Egli, M.; Huang, Z. *J. Am. Chem. Soc.* **2002**, *124*, 24.

(10) (a) Teplova, M.; Wilds, C. J.; Wawrzak, Z.; Tereshko, V.; Du, Q.; Carrasco, N.; Huang, Z.; Egli, M. *Biochimie* **2002**, *84*, 849. (b) Wilds, C. J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M. *J. Am. Chem. Soc.* **2002**, *124*, 14910.

(11) (a) Buzin, Y.; Carrasco, N.; Huang, Z. *Org. Lett.* **2004**, *6*, 1099. (b) Carrasco, N.; Huang, Z. *J. Am. Chem. Soc.* **2004**, *126*, 448. (c) Carrasco, N.; Buzin, Y.; Tyson, E.; Halpert, E.; Huang, Z. *Nucleic Acids Res.* **2004**, *32*, 1638. (d) Salon, J.; Chen, G. X.; Portilla, Y.; Germann, M. W.; Huang, Z. *Org. Lett.* **2005**, *7*, 5645. (e) Logan, G.; Igunbor, C.; Chen, G.-X.; Davis, H.; Simon, A.; Salon, J.; Huang, Z. *Synlett* **2006**, *10*, 1554.

(12) Carrasco, N.; Williams, J. C.; Brandt, G.; Wang, S.; Huang, Z. *Angew. Chem., Int. Ed.* **2006**, *45*, 94.

thymidine (rT) is a common modification in mature tRNAs and rRNAs. Though the ribothymidine is the invariable modification at position 54 of tRNAs of nearly all bacteria and enkarya,[17b] its function in the TΨC loop has not been well understood. Because our previous study indicated that the Se derivatization at the 2′-positions of the nucleotides does not cause significant perturbation of RNAs,[9−11] the Se incorporation at the 2′-position of ribothymidine in tRNAs and rRNAs will help to obtain insight into the important rT modification. Furthermore, the Se incorporation at the same 2′-position of thymidine will help to study the structure and function of DNA molecules. Therefore, there is an urgent need to synthesize the 2′-Se-thymidine derivative, as an analogue of both ribothymidine and thymidine for the RNA and DNA studies. We report here the first synthesis of the 2′-methylseleno-thymidine derivative and the corresponding phosphoramidite, its incorporation into DNAs and RNAs by solid-phase synthesis, and the UV melting and X-ray crystal structure studies of the Se-derivatized DNAs and RNAs.

The synthesis of the 2′-methylseleno-ribothymidine derivative (**4**) started from the glycosylation of the protected D-ribose (**1**) with thymine under the Vorburggen conditions.[18] Intermediate **3** was synthesized by minor modifications of the literature procedures.[19] In brief, due to the 2′-neighboring group effect, this glycosylation generated exclusively the protected β-ribothymidine derivative, followed by the NaOMe treatment to give fully deprotected ribothymidine (**2**) in 88% yield over two steps. After the intramolecular cyclization of **2** in DMF, the 2,2′-anhydrothymidine was selectively protected with the DMTr group at the 5′-position to give **3** in 87% yield in two steps (Scheme 1). In the Se functionality incorporation step, we initially encountered problems of low yield and long reaction time in the Se introduction into anhydroribothymidine **3**. We attempted to increase the reaction yield by protecting the 3′-hydroxyl group and using different reducing reagents, such as LiAlH$_4$ and DIBAH, to
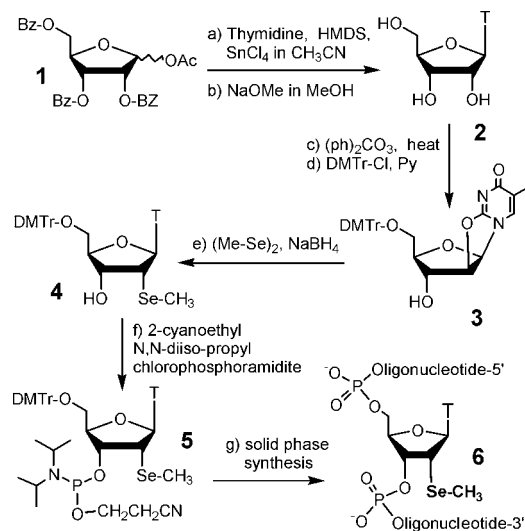
form the methyl selenide nucleophile via the reduction of dimethyl diselenide. Unfortunately, we observed the nucleobase reduction and the elimination reaction. These problems were later overcome by using a weaker reducing reagent (NaBH$_4$) at a higher temperature (50 °C) in EtOH, which gave 82% yield in 3 h. Finally, the 2′-Se-ribothymidine (**4**) was converted to the corresponding phosphoramidite (**5**) in 90% yield.[11a,c]

Similar to solid-phase oligonucleotide synthesis using the thymidine phosphoramidite,[9,11a,c] this novel 2′-Se-ribothymidine phosphoramidite was incorporated into several DNA and RNA oligonucleotides under the standard conditions in the presence of 5-BMP as the coupling reagent. Analogous to our previous report, the 2′-Se functionality on rT was stable under the mild I$_2$ treatment (20 mM, 20 s).[9,11c] The synthesized oligonucleotides were cleaved off the solid support and fully deprotected with concentrated ammonia at 55 °C overnight, followed by reversed-phase HPLC purification twice: DMTr-on and DMTr-off purification. A typical HPLC profile of the Se-DNAs with the DMTr groups is shown in Figure 1A, which is almost identical to that of the corresponding native DNA (data not shown). The HPLC
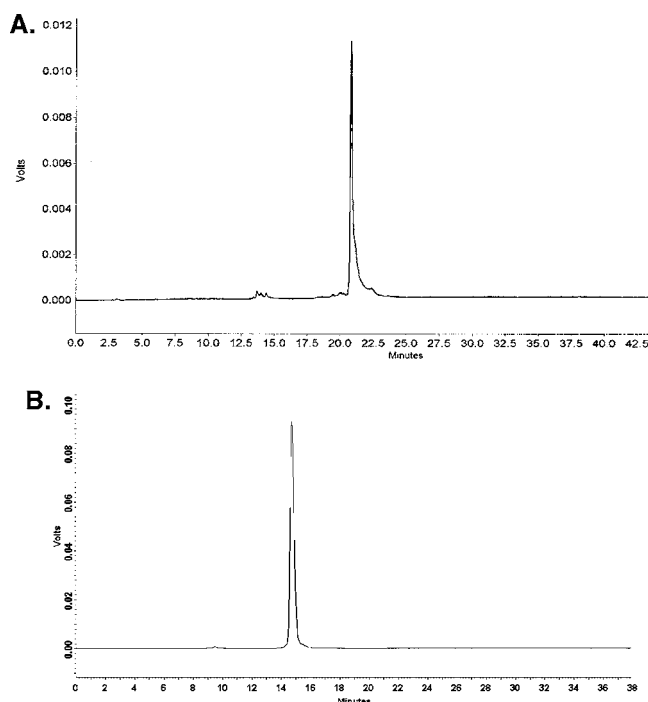

Figure 1. Reversed-phase HPLC analysis and purification of the Se-DNAs. (A) The HPLC analysis profile of the crude DMTr-on Se-DNA (5′-DMTr-GT$_{Se}$GTACAC-3′) after cleavage from the solid support and deprotection. The HPLC analysis was performed on a Zorbax SB-C18 column (4.6 × 250 mm) with a linear gradient from buffer A (20 mM triethylammonium acetate, pH 7.1) to 70% buffer B (50% acetonitrile, 20 mM triethylammonium acetate, pH 7.1) in 20 min; the Se-DNA retention time was 20.6 min. (B) The HPLC purification profile of the DMTr-off Se-DNA (5′-GT$_{Se}$GTACAC-3′). The sample was eluted on a Zorbax SB-C18 column (21 × 250 mm; flow rate of 6 mL/min) with the same gradient and buffers; its retention time was 14.8 min.

Scheme 1. Synthesis of 5′-DMTr-2′-MeSe-thymidine Phosphoramidite (**5**) and Oligonucleotides Containing This Se Derivatization (**6**)

analysis of the native and modified DNA octamers indicated that the Se-modified and nonmodified phosphoramidites had the same incorporation efficiency, revealing a high coupling yield for the Se-T phosphoramidite (99%) as the total formation of the short abortive oligonucleotides was less than 5%. A representative HPLC purification profile of the DMTr-off Se-DNAs is shown in Figure 1B. All of the purified Se-oligonucleotides were also confirmed by MALDI-TOF MS analysis. A typical MS spectrum of the Se-oligonucleotides is shown in Figure 2; all MS data of Se-DNAs and Se-RNAs are collected in Table 1.
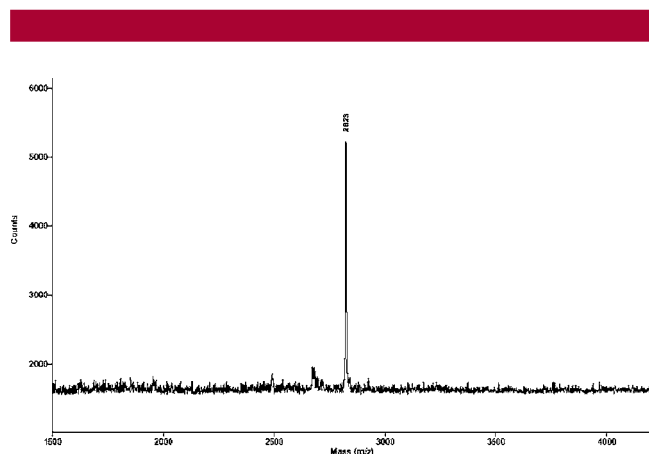

Figure 2. MALDI-TOF MS analysis of the DMTr-off Se-DNA (ATGGT$_{Se}$GCTC). Molecular formula: $C_{89}H_{114}N_{32}O_{54}P_8Se$. FW: 2822.8. Found $[M + H]^+$: 2823. Calcd $[M + H]^+$: 2823.8.

To compare the thermostability of the Se-oligonucleotides with their native ones, the UV melting experiment was carried out. The samples (2 μM DNA or RNA duplexes or hairpin loop), dissolved in the buffer of 50 mM NaCl, 10 mM NaH$_2$PO$_4$−Na$_2$HPO$_4$ (pH 6.5), 0.1 mM EDTA, and 10 mM MgCl$_2$, were heated to 85 °C and allowed to cool to room temperature slowly. The oligonucleotide samples were then stored at 5 °C for 2 h before the UV melting measurements were carried out. All of the denaturing curves

Table 1. MALDI-TOF MS Data of Se-oligonucleotides

entry m/z	Se-oligonucleotides	measured (calcd)
a	Se-DNA-1 (5′-GT$_{Se}$GTACAC−3′) $C_{79}H_{101}N_{30}O_{46}P_7Se$: FW 2502.5	$[M + H]^+$: 2503 (2503.5)
b	Se-DNA-2 (5′-ATGGT$_{Se}$GCTC-3′) $C_{89}H_{114}N_{32}O_{54}P_8Se$: FW 2822.8	$[M + H]^+$: 2823 (2823.8)
c	Se-DNA-3 (5′-GGAT$_{Se}$GGGCG-3′) $C_{91}H_{112}N_{42}O_{51}P_8Se$: FW 2936.9	$[M + H]^+$: 2938 (2937.9)
d	Se-DNA-4 (5′-GCGT$_{Se}$ATACGC-3′) $C_{98}H_{125}N_{38}O_{58}P_9Se$: FW 3120.4	$[M + H]^+$: 3121 (3121.4)
e	Se-RNA-1 (5′-GGT$_{Se}$AUUGCGGUACC-3′) $C_{135}H_{169}N_{52}O_{97}P_{13}Se$: FW 4553.7	$[M + H]^+$: 4555 (4554.7)
f	Se-RNA-2 (5′-UGAGCT$_{Se}$UCGGCUC-3′) $C_{124}H_{157}N_{45}O_{91}P_{12}Se$: FW 4184.5	$[M + H]^+$: 4186 (4185.5)
g	Se-RNA-3 (5′-CUGUGT$_{Se}$UCGAUCCACAG-3′) $C_{162}H_{204}N_{60}O_{118}P_{16}Se$: FW 5454.2	$[M + H]^+$: 5455 (5455.2)

were collected by a Cary 300 UV−visible spectrophotometer with a temperature controller at a heating rate of 0.5 °C/min. The data were analyzed in accordance with the convention of Puglisi and Tinoco.[20] The melting temperatures and typical curves are presented in Table 2 and the

Table 2. UV Melting Temperatures of the 2′-Se-RNAs and -DNAs

entry	DNA/RNA pairs	melting temp (°C, native)
	RNA	
a	5′-GGT$_{Se}$AUUGCGGUACC-3′ 3′-CCA--UAACGCCAUGG-5′	53.5 (54.2)
b	5′-UGAGCT$_{Se}$UCGGCUC-3′ 3′-ACUCGA--AGCCGAG-5′	47.2 (47.8)
c	5′-CUGUGT$_{Se}$UCGAUCCACAG-3′ (the tRNA TΨC hairpin loop)	51.7 (52.4)
	DNA	
d	5′-ATGGT$_{Se}$GCTC-3′ 3′-TACCA--CGAG-5′	42.9 (42.6)
e	5′-GGAT$_{Se}$GGGCG-3′ 3′-CCTA--CCCGC-5′	40.1 (40.2)

Supporting Information, respectively. There were no significant differences between the melting temperatures of the 2′-Se-oligonucleotide duplexes and the native ones. In addition, the Se-modified and native tRNA TΨC loops have almost the same melting temperatures. Therefore, our UV thermostability study indicates that this selenium modification

(13) Brandt, G.; Carrasco, N.; Huang, Z. *Biochemistry* **2006**, *45*, 8972.
(14) (a) Hobartner, C.; Micura, R. *J. Am. Chem. Soc.* **2004**, *126*, 1141. (b) Hobartner, C.; Rieder, R.; Kreutz, C.; Puffer, B.; Lang, K.; Polonskaia, A.; Serganov, A.; Micura, R. *J. Am. Chem. Soc.* **2005**, *127*, 12035.
(15) (a) Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Hobartner, C.; Polonskaia, A.; Phan, A. T.; Wombacher, R.; Micura, R.; Dauter, Z.; Jaschke, A.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2005**, *12*, 218. (b) Moroder, H.; Kreutz, C.; Lang, K.; Serganov, A.; Micura, R. *J. Am. Chem. Soc.* **2006**, *128*, 9909.
(16) Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A. *J. Am. Chem. Soc.* **2006**, *128*, 10847.
(17) (a) Becker, H. F.; Motorin, Y.; Florentz, C.; Giege, R.; Grosjean, H. *Nucleic Acids Res.* **1998**, *26*, 3991. (b) Sprinzl, M.; Vassilenko, K. S. *Nucleic Acid Res.* **2005**, *33*, D139. (c) McCloskey, J. A.; Rozenski, J. *Nucleic Acids Res.* **2005**, *33*, D135.
(18) (a) Vorbruggen, H.; Bennua, B. *Chem. Ber.* **1981**, *114*, 1279. (b) Huang, Z.; Benner, S. A. *J. Org. Chem.* **2002**, *67*, 3996.
(19) (a) Mansuri, M. M.; Starrett, J. E.; Wos, J. A.; Tortolani, D. R.; Brodfuehrer, P. R.; Howell, H. G.; Martin, J. C. *J. Org. Chem.* **1989**, *54*, 4780. (b) Saito, Y.; Zevaco, T. A.; Agrofoglio, L. A. *Tetrahedron* **2002**, *58*, 9593. (c) Shi, J.; Du, J.; Ma, T.; Pankiewicz, K. W.; Patterson, S. E.; Tharnish, P. M.; McBrayer, T. R.; Stuyver, L. J.; Otto, M. J.; Chu, C. K.; Schinazi, R. F.; Watanabe, K. A. *Bioorg. Med. Chem.* **2005**, *13*, 1641.

(20) Puglisi, J. D.; Tinoco, I., Jr. *Methods Enzymol.* **1989**, *180*, 304.

does not significantly alter the duplex and loop structures, which is consistent with our previous UV melting results of the 2′-Se-DNA and RNA oligonucleotides.[11a,c] As the UV melting study cannot provide detailed information on the structure change, NMR structural analysis will be needed for further characterization.

In our crystallization study, we found that the Se-DNA (GT$_{Se}$GTACAC, self-complementary) crystallized in 2−3 weeks in the native buffer, where the native DNA (GTG-TACAC) crystallized over 2 months.[21,22] When the Nucleic Acid Mini Screen kit (Hampton Research, with 24 diversified crystallization buffers) was screened, to our pleasant surprise, we found that the Se-DNA crystallized overnight in 20 out of 24 buffers (buffers no. 5−24). In contrast, the native DNA did not crystallize at all over many weeks in these Hampton buffers. It was also reported[15b] that the two RNAs (12-mer and 16-mer) containing the 2′-Se-guanosines crystallized, respectively, in 17 and 33 buffers, whereas their corresponding native RNAs crystallized in 15 and 24 buffers, individually. Though the crystallization conditions of these Se-RNAs expanded slightly, it appears that these two Se-RNAs, which crystallized mostly as thin needle crystals, behaved similarly to many Se-Met-derivatized proteins that required the fine tuning of their native crystallization conditions. By contrast, the majority of these Se-DNA crystals grown under the new buffers was larger than the native DNA crystals grown under the native buffer conditions, and these Se-DNA crystals also diffracted well. Our results reveal that the 2′-Se derivatization of thymidine facilitated the crystal packing and growth and appeared to help expand the crystallization conditions.

Finally, the structure of the Se-DNA crystal, grown under the buffer no. 7, was determined at 1.40 Å resolution via the Se SAD phasing and refinement. We found that the determined Se-DNA structure (Figure 3) is superimposable to the native structure (2.0 Å resolution) with the same space group,[21] indicating that this Se derivatization does not alter the native structure significantly. In addition, this higher structure resolution allows observation of ordered 43 water molecules and other structure features. The Se-derivatized structure reveals that the 2′-Se-furanose displays the 3′-endo sugar pucker, which is consistent with the sugar pucker of A-form DNAs and RNAs, and the 2′-methylseleno group is placed in the minor groove of the duplex. Data collection, phasing, and refinement statistics of the determined Se-DNA structure (2HC7) are listed in the Supporting Information.

In conclusion, we have successfully synthesized the 2′-Se-thymidine phosphoramidite (5) as an analogue of both thymidine and ribothymidine phosphoramidites for the solid-phase Se-oligonucleotide synthesis. Subsequently, we have incorporated it into DNAs and RNAs with over 99%
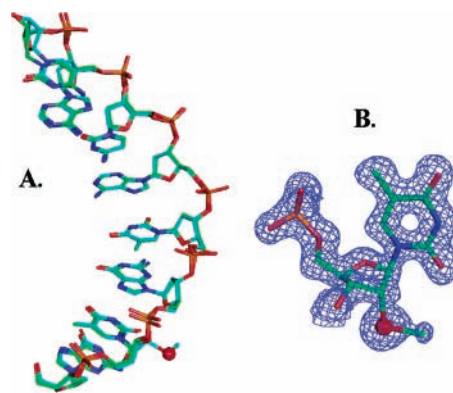

Figure 3. DNA crystal structures (Se in red). (A) Superimposed structure comparison of the native DNA (GTGTACAC, 1DNS in green) and the Se-derivatized DNA (GT$_{Se}$GTACAC, 2HC7 in cyan). (B) The structure model and the electron density map of 2′-Se-ribothymidine in the structure.

coupling yield. Our thermal denaturing study of the Se-modified DNAs and RNAs also reveals that this Se derivatization did not significantly change the stability of the structures of the DNA and RNA duplexes and the tRNA TΨC loop. Furthermore, our X-ray crystal structure study with a Se-modified DNA octamer (GT$_{Se}$GTACAC) shows that the Se-derivatized DNA structure (1.40 Å resolution) is virtually identical to the native structure (2.0 Å resolution) with the same space group. More importantly, we observed that this Se modification dramatically expanded the crystallization conditions of the Se-DNA and largely facilitated crystal growth while retaining the high diffraction quality. In addition to DNA investigations, this 2′-Se-thymidine can also serve as a ribothymidine analogue for the structure and function studies of tRNAs and rRNAs. As the Se derivatization can facilitate the phase determination and crystallization, it will significantly impact the structure and function studies of nucleic acids and their protein complexes using X-ray crystallography. In addition, we are in the process of further investigating the mechanism of the crystal growth facilitated by the Se derivatization.

Supporting Information Available: Experimental procedures, ^{1}H and ^{13}C NMR, UV melting curves, and HRMS analytical data. This material is available free of charge via the Internet at http://pubs.acs.org.

OL062937W

(21) Jain, S.; Zon, G.; Sundaralingam, M. *Biochemistry* **1989**, *28*, 2360.
(22) Thota, N.; Li, X. H.; Bingman, C.; Sudaralingam, M. *Acta Cryst.* **1993**, *D49*, 282.