J|A|C|S
COMMUNICATIONS

Published on Web 12/12/2001

Internal Derivatization of Oligonucleotides with Selenium for X-ray Crystallography Using MAD

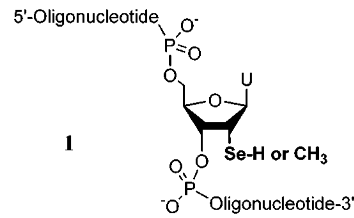
Quan Du,[†] Nicolas Carrasco,[†] Marianna Teplova,[‡] Christopher J. Wilds,[‡] Martin Egli,[‡] and Zhen Huang*[,†]

Department of Chemistry, Brooklyn College, and Program of Biochemistry and Chemistry, The Graduate School, The City University of New York, 2900 Bedford Avenue, Brooklyn, New York 11210, and Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235-1634

Received September 18, 2001

Determination of the three-dimensional structures of DNA oligonucleotides, DNA−drug complexes, ribozymes, and viral RNAs with high resolution is invaluable for gaining insights into their functions and mechanisms.[1−4] Several approaches, including heavy-atom soaking, cocrystallization, and halogen derivatization, have been used to label DNA and RNA for nucleic acid X-ray crystallography.[5] Heavy atom soaking and cocrystallization often prove not very successful for nucleic acid X-ray crystallography, and halogen derivatization is usually limited to short nucleotides. In protein X-ray crystallography, selenium is used to replace sulfur to mimic methionine,[6] and this selenomethionine derivatization method is widely used in phase and structure determination of proteins using multiwavelengh anomalous dispersion (MAD).[7] Indirect derivatization of RNAs using selenomethionine-labeled RNA-binding protein U1A for phase and structure determination has been successfully demonstrated, although RNA-binding protein U1A has to be prepared and appropriate positions for inserting the U1A-binding site have to be identified by constructing numerous ribozyme constructs and screening their complexes with the protein.[4,8] Therefore, direct labeling of nucleotides with selenium will largely simplify the derivatization effort and will facilitate nucleic acid X-ray crystallography.

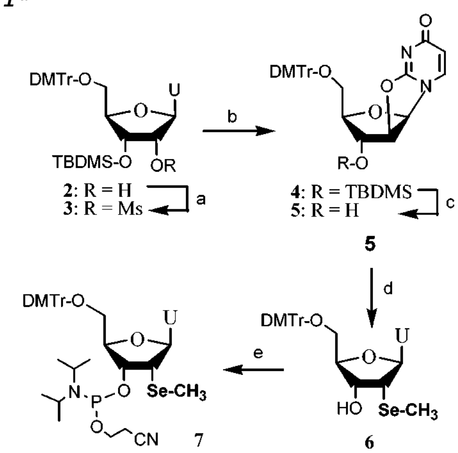
Oxygen atoms in nucleotides are the best choices for selenium selective substitution to mimic natural nucleotides. Recently, we have reported replacement of the 5′-oxygen of nucleosides with selenium and synthesis of oligonucleotides containing selenium at 5′-termini.[9] Demonstration of the stability of the selenium functionality and its compatibility with the solid-phase phosphoramidite chemistry prompted us to investigate the possibility of introducing selenium to other positions, especially internal positions. Here we describe the synthesis of oligonucleotides containing selenium at the 2′-α-position of uridine **1** (Figure 1), and reveal crystallization and structural studies of the selenium-containing oligodeoxyribonucleotides.



1

Figure 1.

Scheme 1 [a]



[a] a. MsCl/THF/TEA, 95% yield; b. toluene/ tetrahexylammonium hydrogen sulfate/ Na_2CO_3 (satd), 96% yield; c. $(Bu)_4N^+ F^-$, 95% yield; d. NaHSe, then CH_3I, or $NaSeCH_3$, 96%; e. $PCl(OCH_2CH_2CN)N(^iPr)_2$, 92% yield.

After mesylation of partially protected uridine **2** at the 2′-position (Scheme 1), the mesyl group was displaced by the uracil exo-2-oxygen in basic conditions.[10] A two-phase reaction system (toluene and aqueous Na_2CO_3), catalyzed by a phase-transfer catalyst, was developed to facilitate the nucleophilic substitution; anhydro-uridine **4** was formed in 96% yield. Since our experiments indicated that the bulky 3′-TBDMS group blocked selenide nucleophiles attacking at 2′ position from the α-face, this group was removed by the fluoride treatment. We also found that if NaHSe generated by reduction of selenium metal with $NaBH_4$[11] was used as the nucleophile to attack **5** at the 2′-position, an additional step was required to protect the resulting selenol from oxidation. When sodium methylselenide was used as the nucleophile to open the tricyclic ring of **5**, selenium−nucleoside **6** was obtained in 96% yield with methyl protection, which prevents oxidation of the selenium functionality. The selenide nucleophilic reactions were conducted in THF solution, which avoided the ring-opening at the 2-position, resulting in substitution at the base.[12] Compound **6** was analyzed by MS, ^{77}Se NMR, 2D-NMR, and NOE experiments to confirm the stereochemistry and the structure. Nucleoside **6** was finally converted to selenium-labeled phosphoramidite **7** in 92% yield by reaction with 2-cyanoethyl *N*,*N*-diisopropyl-chlorophosphoramidite.

Using the phosphoramidite **7**, DNA and RNA analogues containing selenium at the 2′-positions [DNA-octamer, 5′-G**U**$_{Se}$GTACAC;[13] DNA-decamer, 5′-GCGTAU$_{Se}$ACGC-3′;[14] RNA-hexamer, and 5′-r(CGU$_{Se}$AC)dG[3]5] were synthesized following standard solid-phase synthesis. The potential for scale-up was demonstrated by 10 μmol syntheses. As expected, the protected selenide functionality was

[†] The City University of New York.
[‡] Vanderbilt University.

10.1021/ja0171097 CCC: $22.00

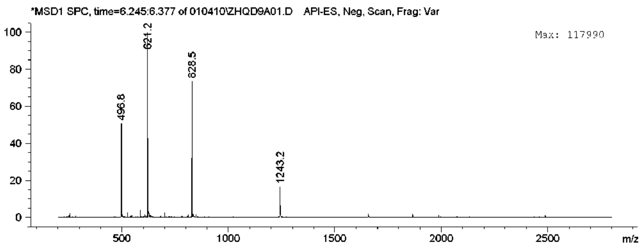


Figure 2. Electrospray MS of the octamer, 5′-G**U**$_{Se}$GTACAC, molecular formula $C_{78}H_{99}N_{30}O_{46}P_7Se$, MW = 2489.63 (including all isotopes). Measured (expected) *m*/*e*: $[M - 2H^+]^{2-}$ = 1243.2 (1243.8); $[M - 3H^+]^{3-}$ = 828.5 (828.8); $[M - 4H^+]^{4-}$ = 621.2 (621.4); $[M - 5H^+]^{5-}$ = 496.8 (496.9).

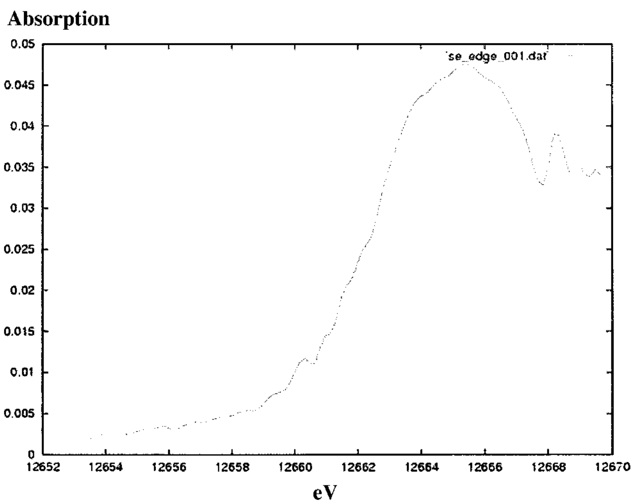


Figure 3. X-ray fluorescence spectrum of the decamer crystal. The theoretical value for the Se K edge is 12.6578 keV (0.9795 Å).

found stable in mild I_2 treatment (20 mM, 20 s) for the phosphite oxidation. The Se−oligonucleotides **1** with methyl protection were purified by HPLC, and the selenium functionality was confirmed by electrospray mass spectrometry. The MS spectrum of the octamer, shown as an example, is displayed in Figure 2, where a set of the nucleotide anions carrying negative charges from 2 to 5 were observed.

Crystallization conditions were screened, and diffraction quality crystals were identified. X-ray fluorescence spectra confirmed the presence of selenium in crystals (Figure 3). MAD data of the Se−decamer to 1.2 Å resolution were collected at the Advanced Photon Source, and the diffraction data were successfully phased on the basis of the selenium anomalous signal. Likewise, diffraction data of the octamer to 1.8 Å resolution were collected, and the structure of the octamer was determined by the molecular replacement technique. These X-ray structures confirmed the presence of the 2′-methylseleno group at the α-position of the uridine.

In both structures, the 2′-Me-Se-substituted furanoses display C3′-endo puckers, consistent with the A-form geometry of the unmodified decamer and octamer duplexes, which is adopted by RNA and A-form DNA. As previously established for 2′-O-methylated nucleotides and other 2′-O-modified ribonucleotide analogues,[14] the methyl groups of the methylseleno moieties are directed into the minor groove and the C3′−C2′−Se−Me torsion angles adopt an antiperiplanar conformation. Details of the structure determination and refinement results will be reported elsewhere.

In conclusion, we have developed a route for the synthesis of 2′-selenium uridine analogues and oligonucleotides containing selenium labels, and have demonstrated for the first time a new strategy to covalently derivatize nucleotides with selenium for phase and structure determination in X-ray crystallography. The 2′-α-postion selenium derivatization retains the native C3′-*endo* conformation of A-Form DNA and RNA molecules. As the solid-phase synthesis allows preparation of Se−RNA and Se−DNA in large scales, unlike the phosphoroselenonate-mediated autoligation of DNA strands,[16] this approach is suitable for RNA and A-Form DNA derivatization for X-ray crystallography. Selenium labels can also be incorporated into a large RNA molecule via ligation of a transcribed fragment and a synthetic fragment containing selenium labels. This derivatization method may serve as an alternative approach in phase and structure determination of RNA−protein and DNA−protein complexes by derivatizing RNA and DNA instead of proteins.

Acknowledgment. We thank Dr. Xiangpeng Kong at NYU for collecting diffraction data at Synchrotron Light Source at Brookhaven, Dr. Hsin Wang at the Staten Island College for assisting in high-field NMR data collection, and Dr. Soll and Dr. Edeninger at Hunter College for assisting in MS data collection and DNA synthesis. We are also grateful to Dr. Zdzislaw Wawrzak for help with data collection. This work was supported by PSC-CUNY Research Awards (69674-00-29 and 62392-00-31) and New Research Dimension Fund (to Z.H.), and NIH (GM-55237 to M.E.).

Supporting Information Available: HRMS, ^{1}H, ^{13}C, ^{77}Se NMR data and 2D-NMR spectra of the nucleoside analogues, mass spectra of the oligonucleotides, HPLC analysis, and crystallization (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Pley, H. W.; Flaherty, K. M.; Mckay, D. B. *Nature* **1994**, *372*, 68−74.
(2) Xiong, Y.; Sundaralingam, M. *Nucleic Acids Res.* **2000** *28*, 2171−2176.
(3) Hung, L.-W.; Holbrook, E. L.; Holbrook, S. R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5107−5112.
(4) Rupert, P. B.; Ferre-Damare, A. R. *Nature* **2001**, *410*, 780−786.
(5) For review, see: Holbrook, S. R.; Kim, S.-H. *Biopolymers* **1997**, *44*, 3−21.
(6) Yang, W.; Hendrickson, W. A.; Crouch, R. J.; Satow, Y. *Science* **1990**, *249*, 1398−1405.
(7) Obmolova, G.; Ban, C.; Hsieh, P.; Yang, W. *Nature* **2000**, *407*, 703−710.
(8) Ferre-D'Amare, A. R.; Zhou, Kaihong; Doudna, J. A. *Nature* **1998**, *395*, 567−574.
(9) Carrasco, N.; Ginsburg, D.; Du, Q.; Huang Z. *Nucleosides, Nucleotides Nucleic Acids* **2001**, *20*, 1723−1734.
(10) (a) Horwitz, J. P.; Chua, J.; Da Rooge, M.; Noel, M.; Klundt, I. L. *J. Org. Chem.* **1966**, *31*, 205−210. (b) Haraguchi, K.; Tnaka, H.; Maeda, H.; Itoh, Y.; Saito, S.; Miyasaka, T. *J. Org. Chem.* **1991**, 56, 5401−5408. (c) McGee, D. P. C.; Vanghn-Settle, A.; Vargeese, C. Zhai, Y. *J. Org. Chem*. **1996**, *61*, 781−785.
(11) Jakiwczyk, O. M.; Kristoff, E. M.; McPhee, D. *Synth. Commun.* **1993**, *23*, 195−199.
(12) Cosford, N. D. P.; Schinazi, R. F. *J. Org. Chem.* **1991**, *56*, 2161−2165.
(13) Jain, S. C.; Zon, G.; Sundaralingam, M. *Biochemistry* **1989**, *28*, 2360−2364.
(14) (a) Tereshko, V.; Portmann, S.; Tay, E. C.; Martin, P.; Natt, F.; Altmann, K.-H.; Egli, M. *Biochemistry* **1998**, *37*, 10626−34. (b) Lubini, P.; Zürcher, W.; Egli, M. *Chem. Biol.* **1994**, *1*, 39−45.
(15) Biswas, R.; Mitra, S. N.; Sundaralingam, M. *Acta Crystallogr., Sect. D* **1998**, *54*, 570−576.
(16) Xu, Y.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 9040−9041.

JA0171097