COMMUNICATION

Synthesis of 2'-Deoxy-2'-fluoro-L-cytidine and Fluorinated L-Nucleic Acids for Structural Studies

Yuliya Dantsu,[a] Ying Zhang,[a] and Wen Zhang*[a,b]

[a] Dr. Yu. Dantsu, Dr. Y. Zhang, and Prof. Dr. W. Zhang
Department of Biochemistry and Molecular Biology
Indiana University School of Medicine
635 Barnhill Drive, Indianapolis, IN 46202, USA
E-mail: wz15@iu.edu
[b] Prof. Dr. W. Zhang
Melvin and Bren Simon Cancer Center
535 Barnhill Dr, Indianapolis, IN 46202, USA

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

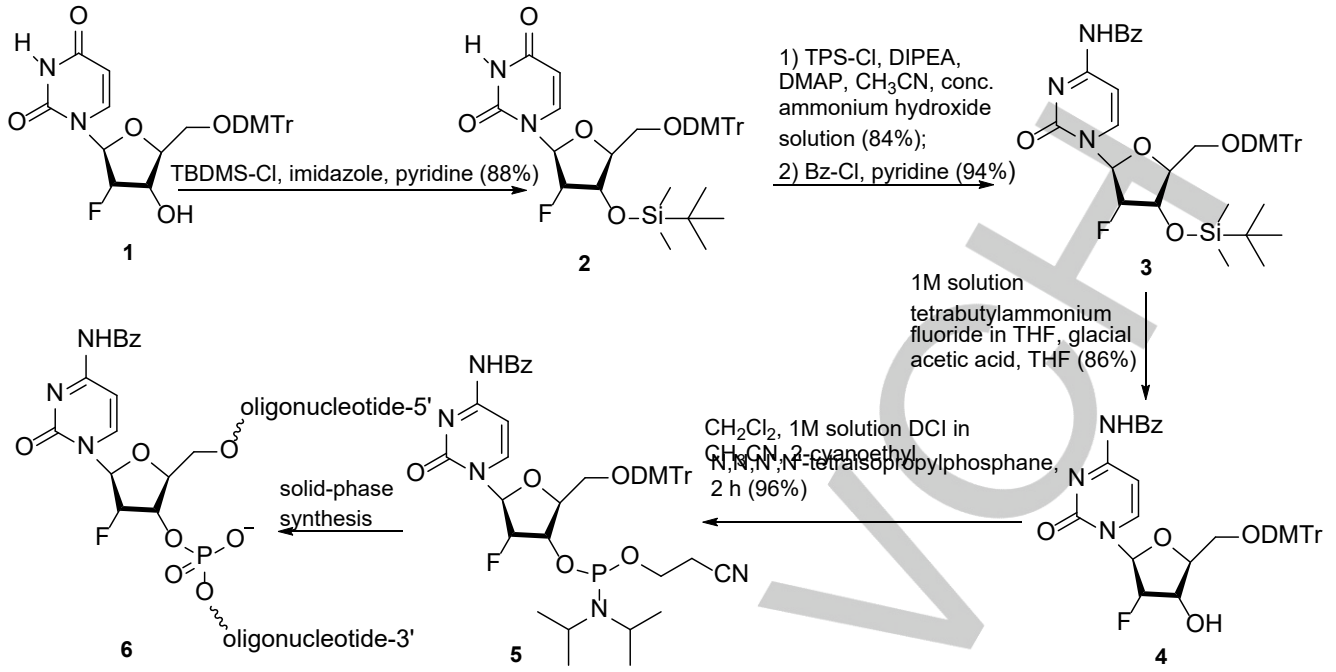
Abstract: The unique properties of fluorine atom on nucleic acid backbone can offer stricking functional and structural features. In order to extend the biological applications of L-type nucleic acid, we chemically incorporate the fluoro-modification into 2'-position of L-cytidine, and obtain a series of fluoro-modified L-DNAs/ L-RNAs. Our melting study indicates that the 2'-fluoro-modification does not disrupt the thermostabilities of wild-type L-nucleic acids. Consistently, our X-ray crystal structure reveals that fluoro-moiety cause no structural perturbation, and the fluoro-L-cytidine forms the Watson-Crick base pair with L-guanosine virtually identical to nonmodified L-type CG pair. This fluoro-modified cytidine provides a useful biochemical strategy to investigate L-nucleic acid as advanced molecular therapy.

Fluorination is one of the most common chemical derivatizations that have been widely used in pharmaceutical chemistry, due to the unique biophysical and biochemical properties of fluorine element, like small size and extreme electronegativity.[1] Particularly, replacement of H or OH group by fluorine element in nucleic acid results in the striking biochemical features. The small size of fluorine leads to minimal steric effect on nucleic acid functions, and the significant electronic effect of fluoride modification could result in the characteristic hydrogen bonding, as well as the alteration of nucleic acid sugar pucker and overall structure.[2] Additionally, fluorinated nucleic acids display the enhanced physiological and thermal stability to benefit their application as biomedical therapies.[3-4] The strategy of nucleic acid fluorination has been attracting extensive interest in the past several decades. The successful examples include the fluoride-modified nucleosides as anticancer, antiviral or antibacterial drug,[5] and the fluoride-modified oligonucleotides as therapeutic antisense,[6] siRNA[3] and aptamer[7] drugs. Plus, fluoro-derivatizations are widely used to probe the nucleic acid structures by ^{19}F NMR, due to the intense signal of ^{19}F nucleus and its sensitive chemical shift variation.[8] Hence, the ongoing development of novel fluorinated nucleic acid molecules is significant, in order to benefit the fields of innovative drug discovery and basic biochemical studies.

On the other hand, to address the challenge that nucleic acid therapy is susceptible to nuclease-mediated degradation in vivo,[9] a new technique of chemical modification has been pioneered, based on the nucleic acids composed of the entirely opposite handedness.[10-11] The non-naturally occurring L-type RNA/DNA is the enantiomer of native D-type RNA/DNA, and it forms versatile secondary and tertiary conformations in the mirror-image fashion. Owing to the unordinary chirality of L-nucleic acids, they are immune to the recognition of natural nuclease in human body. As a result, L-RNA/DNA molecules exhibit extraordinary stability and extended circulating time in vivo.[12] Unlike the classical siRNA or antisense therapies, L-nucleic acids cannot sequence-specifically bind to native disease gene through Watson-Crick base pairing.[13] The present investigation on L-nucleic acid drug mainly focuses on the selection of L-aptamers, which target the disease-relevant protein and RNA elements through 3D structural interactions.[14] The development of Systematic Evolution of Ligands by EXponential enrichment (SELEX) technique has enabled to isolate the structure-specific L-aptamers.[15] SELEX of functional L-nucleic acids has demonstrated outstanding potential in development of L-aptamer to specifically bind different molecular targets,[16-17] and even creation of a mirror-image polymerization system.[18-19] Besides, L-nucleic acid has also been applied to construct unique DNA nanostructure as advanced material and drug delivery device,[20] and the unique features of L-nanostructure could offer the superior stability to ensure the effective biodistribution in vivo.[21]

Considering the practical value of fluoro-modification in nucleic acid-related research fields, and the exceptional potential of L-nucleic acids as next-generation of molecular therapy, it is critical to further explore the fluorinated L-DNA/L-RNA to expand their biomedical applications. The fluorine-induced pseudo-hydrogen bonds and sugar conformational change are possible to engender the unique ligand-binding aptamers. The fluorine atom at 2'-position of L-(deoxy)ribose is likely to enhance the thermostability of L-duplex, which helps to keep the intact geometry of L-nanostructure in human body. Earlier, we chemically synthesized a mirror reflection of native 2'-F-uridine building block and, subsequently, investigated a systematic thermal stability and enzymatic assay of a series of 2'-fluoro-L-DNA/RNA oligonucleotides.[22] Our results revealed the enhancement of structural and thermal stabilities after fluoro-modification was incorporated into 2'-position of L-uridine.

This is the author's manuscript of the article published in final edited form as:
Dantsu, Y., Zhang, Y., & Zhang, W. (2021). Synthesis of 2′-Deoxy-2′-fluoro-L-cytidine and Fluorinated L-Nucleic Acids for Structural Studies. ChemistrySelect, 6(39), 10597–10600. https://doi.org/10.1002/slct.202103202



Scheme 1. Syntheses of 3′-O-(2-Cyanoethyl-N,N-diisopropyl- phosphoramidite)-4-N-benzoyl-5′-O-[bis(4-methoxyphenyl)phenylmethyl]-2′-deoxy-2′-fluoro-β-L-cytidine **5** and its corresponding oligonucleotides

Here we continue to investigate the chemically modified L-nucleic acids by introducing fluorine atom into 2'-positoin of L-cytidine and characterizing its functions by stability and structural studies. The synthetic route of the target building block, phosphoramidite **5**, is depicted in Scheme 1. The synthesis of 2′-deoxy-2′-fluoro-β-L-cytidine nucleoside **4** started from (4,4′-dimethoxytrityl)-2′-deoxy-2′-fluoro-β-L-uridine ***1,*** which was synthesized by following our previous report.[22] The following protection of 3' positions by tert-butyldimethylsilyl (TBDMS) group gave rise to the compound **2**. The carbonyl group at 4-position of uracil was then activated by triisopropylbenzenesulfonyl (TPS) group, and the derivatized L-uridine nucleoside analogue was finally converted into L-cytidine derivative **3** by amination reaction with concentrated ammonium hydroxide. After protection of the exocyclic amino group with benzoyl (Bz) group and deprotection of 3'-TBDMS by treating with 1M tetrabutylammonium fluoride, compound **4** was synthesized. The target phosphoramidite **5** was finally obtained in quantitative yield (98%). The fluoro-modified L-cytidine was then incorporated into different oligonucleotides for the downstream functional studies. All the synthetic L-oligonucleotides are purified by HPLC and characterized by MS analysis (Table 1 and Supporting Information).

The thermal stability of DNAs/RNAs is considered as an essential characteristic when applied as therapeutic agent in human body. Therefore, we perform the thermal denaturation study to determine whether the 2'-F moiety on L-cytidine has the potential stabilizing effect and the fluoro-modification is generating thermal perturbation. The intercalating fluorescence dye (Sybr green I) was exploited in melting temperature (T_m) measurement of various oligonucleotides. The Sybr green I binds only to double-stranded nucleic acids by intercalating between base pairs, thereby the change of fluorescence signal can be monitored during denaturing process. All the experiments were performed in the buffer of 750 mM NaCl, 10 mM $MgCl_2$, 50 mM Na_3PO_4, pH 6.8.

We first examined the T_m of L-DNA duplexes. The melting temperature of self-complimentary L-5′-dCGCGAATTCGCG-3′ (Table 1, Entry a) was measured to be 63.9 °C. When we introduced one fluoro-modified residue (Entry b), the T_m was slightly enhanced by 0.5 °C, to 64.4 °C. Besides, when 2 fluoro-modified residues were incorporated into the same sequence (Entry c), T_m could be improved by 0.7 °C, to 64.6 °C. These melting studies lend the evidence that, bringing in fluoro-modification to the 2'-position of sugar pucker of L-cytidine could give rise to a more thermal stable L-DNA. One hypothetical explanation is that, the introduced 2'-F moiety can rigidify the sugar conformation into L-3'-*endo,* which potentially strengthens the π-π stacking interaction in duplex. Plus, compared to the hydrogen atom at 2'-position of DNA, the introduced fluoride modification could induce additional pseudo hydrogen bonding of C-F···H-O-H to stabilize the duplex in water solution.

We then incorporated 2'-F into L-RNA backbone to test its effect. The self-complimentary L-RNA contained the same sequence as L-DNA, L-5′-CGCGAAUUCGCG-3′ (Entry d), and it displayed a T_m of 41.0 °C. After bringing in one fluoro-modification (Entry e), T_m remained almost identical to nonmodified L-RNA (40.9 °C). However, when two fluoro-modifications were present (Entry f), T_m dropped by ~0.7 °C to 40.3 °C. Different from of L-DNA, we didn't observe meaningful enhancement of 2'-fluorination to the thermal stability of L-RNA. What is more, multiple fluoride modifications could diminish the helical stability. Generally, 2'-fluorination switches the sugar pucker conformation because of its electronegativity rather than its hydrogen-like atomic size.[23] Therefore, compared to the 2'-fluorine atoms, the wild-type 2'-OH groups in L-RNA are probably more potent to configure the L-ribose to the 3'-*endo* conformation. Too many 2'-fluoride modifications in L-RNA may sacrifice the thermal stability to some

degree. And generally, L-RNA has a much lower T_m than L-DNA. This is likely due to that L-DNA has additional 4 consecutive 5-methyl groups on thymine nucleobases, which contribute significantly to the helical stability by the expanded stacking interaction.[24]

To seek the insights into the helical geometry of 2'-F-containing L-A and L-B form duplexes, we then carried on the circular dichroism studies to investigate the basic conformations of the synthetic oligonucleotides. The CD spectra of L-double strands and their D-formed counterparts (DNA b and b' and RNA e and e') were recorded. The native D-form 2'-F-modified DNA provided a conservative CD spectrum with amplitude bands: a positive band around 278 nm and a negative one around 252 nm. They are highly close to the characteristic wavelengths that a B-form DNA adopts. In contrast, its mirror-image L-form 2'-F-modified DNA showed the chiral inversion (negative around 278 nm and positive around 252 nm, Figure S24A). For the native 2'-F-modified RNA, the CD spectrum was characterized as the A-form like duplex, by a short negative band at 293 nm, an intense positive one at 262 nm. The corresponding chiral L-RNA displayed the inverted CD spectrum (Figure S24B). The CD characterization illustrates the synthetic fluoro-modified L-DNA and L-RNA are adopting the classical L-type B-form and A-form helical conformations, and the fluoro-modification generates no basic structural perturbation to the L-duplexes.

To further understand the structural effect 2'-F would produce to L-duplex, we carried out the crystallographic structural study of a self-complementary L-DNA, L-5′-dG$\underline{C_F}$GTACGC-3′. Compared to regular L-DNA, the crystallization of the fluoro- L-DNA was greatly facilitated with the assistance of 2'-modification to rigidify the conformation. The crystals could grow to the optimal size within 2 weeks and diffract to atomic resolution in crystallographic data collection. By molecular replacement, we successfully determined the crystal structure of this fluoro-L-DNA to 1.35 Å, using our previous published structure (PDB code 7KW4) as the searching model. The superimposition of this structure with the published L-DNA structure reveals that, these two structures have the highly identical left-handed helical structures (Figure 1A). The fluoro-L-cytidine forms Watson-Crick base pair with the L-guanosine, with the hydrogen bond lengths of 2.8-3.0 Å (Figure 1B). The fluoro-L-dC:L-dG base pair is also superimposable to the fluoro-L-dU:L-dA pair we observed previously (Figure 1C). The furanose ring of L-deoxycytidine is derivatized to the L-form 3′-*endo* conformation, and the 2′-fluoro atom is pointing to the minor groove of the L-helix. The pairing nucleotide of L-dG also has the 3′-*endo* conformation. This is different from our previous structure, that the L-dA had a 2′-*endo* conformation when pairing to fluoro-L-dU (Figure 1C). Moreover, some water molecules are observed to locate in the minor groove, close to the 2′-fluoro derivatization (Figure 1D). The potential fluorine-involved pseudo hydrogen bond and water-involved regular hydrogen bonds are likely to bridge the 2′-fluoro-L-deoxycytidine and the upstream L-deoxyguanosine, thereby structurally stabilizing the L-DNA duplex. The collected diffraction data, phasing and refinement statistics are included in the Supporting Information (Table S1 and S2).

In summary, we have successfully synthesized the 2'-fluoro-L-deoxycytidine phosphoramidite and incorporated the nucleoside into L-DNA and L-RNA for biophysical and structural studies. The experimental results reveal that the fluoride modification at sugar pucker of L-DNA can enhance the thermal stability, though it lightly diminishes the thermal stability when present in L-RNA. The crystal structure of fluoro-L-nucleic acid is highly similar to the corresponding nonmodified L-nucleic acids. We expect our results can encourage more in-depth investigation on chemically modified mirror-image nucleic acid, and eventually benefit the biomedical application of L-nucleic acid as molecular therapy. Moreover, the reported fluoro-modification at L-deoxyribose

Table 1. MS analyses and T_m measurements of synthetic L-oligonucleotides (n.d. = not determined)

Entry	Sequences	Isotopic mass *m/z* Measured (calcd.)	Duplex T_m (°C)
a	L-5′-dCGCGAATTCGCG-3′	$[M\text{-}3H]^{3-}$: 1213.8 (1213.9)	63.9±0.1
b	L-5′-dCGCGAATTC$_F$GCG-3′	$[M\text{-}3H]^{3-}$: 1219.8 (1219.9)	64.4±0.2
c	L-5′-dCGCGAATTC$_F$GC$_F$G-3′	$[M\text{-}3H]^{3-}$: 1225.6 (1225.9)	64.6±0.2
d	L-5′-rCGCGAAUUCGCG-3′	$[M\text{-}3H]^{3-}$: 1268.8 (1268.5)	41.0±0.4
e	L-5′-rCGCGAAUUC$_F$GCG-3′	$[M\text{-}3H]^{3-}$: 1269.5 (1269.2)	40.9±0.1
f	L-5′-rCGCGAAUUC$_F$GC$_F$G-3′	$[M\text{-}3H]^{3-}$: 1270.2 (1270.5)	40.3±0.1
b'	D-5′-dCGCGAATTC$_F$GCG-3′	$[M\text{-}2H]^{2-}$: 1830.8 (1830.8)	n.d.
e'	D-5′-rCGCGAAUUC$_F$GCG-3′	$[M\text{-}3H]^{3-}$: 1269.5 (1269.2)	n.d.

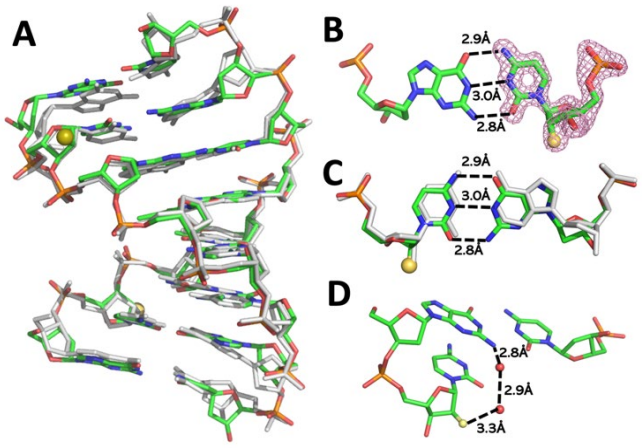


Figure 1. Fluoro-L-DNA crystal structure. A: Superimposed structures of 2'-F-dC- L-DNA (green stick, PDB 7MOO) and the corresponding 2'-F-dU-L-DNA (gray stick, PDB 7KW4). Fluorine atom is shown in gold. B: Local structure and electron density map of 2'-F-L-dC:L-dG base pair. The corresponding F_o–F_c omit map contoured at 1.5σ (pink mesh) indicates the ordered 2'-F-dC. C: Superimposed structures of 2'-F-L-dC:L-dG pair and 2'-F-L-dU:L-dA pair, which share the same location in the duplex. D: 2'-F is forming pseudo hydrogen bond with neighbouring water, to interact with upstream L-dG through hydrogen bonding network.

demonstrates a useful derivatization strategy to facilitate L-nucleic acid crystallization for structural studies.

Supporting Information Summary

All the general techniques for the organic and solid-phase syntheses of the target compounds, their physicochemical characterizations, including Nuclear Magnetic Resonance spectra, Liquid Chromatography Mass Spectrometry analyses, melting temperature and Circular Dichroism curves are detailed in the Supporting Information.

Acknowledgements

We thank the Zhang lab for helpful discussions, insightful commentary and careful revision of the manuscript. We thank Dr. L. Zeng and the Chemical Genomics Core at IUSM for LC-MS and NMR analysis. We thank Dr. Zdzislaw Wawrzak from the Life Sciences Collaborative Access Team beamline 21-ID-D at the Advanced Photon Source, Argonne National Laboratory (USA). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

Accession numbers

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data bank under accession code 7MOO.

Conflicts of interest

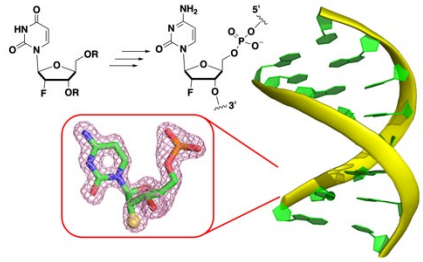
There are no conflicts to declare.

Keywords: crystal structure • fluoride modification • L-nucleic acid • nucleosides • oligonucleotides

[1] S. Purser, P. R. Moore, S. Swallow, V. Gouverneur, *Chem. Soc. Rev.* **2008**, *37*, 320-330.
[2] F. Guo, Q. Li, C. Zhou, *Org. Biomol. Chem.* **2017**, *15*, 9552-9565.
[3] E. L. Chernolovskaya, M. A. Zenkova, *Curr Opin Mol Ther.* **2010**, *12*, 158-167.
[4] P. S. Pallan, E. M. Greene, P. A. Jicman, R. K. Pandey, M. Manoharan, E. Rozners, M. Egli, *Nucleic Acids Res.* **2011**, *39*, 3482-3495.
[5] X.-L. Qiu, X.-H. Xu, F.-L. Qing, *Tetrahedron* **2010**, *66*, 789-843.
[6] A. M. Kawasaki, M. D. Casper, S. M. Freier, E. A. Lesnik, M. C. Zounes, L. L. Cummins, C. Gonzalez, P. D. Cook, *J Med Chem.* **1993**, *36*, 831-841.
[7] C. Kratschmer, M. Levy, *Nucleic Acid Ther.* **2017**, *27*, 335-344.
[8] L. G. Scott, M. Hennig, *Methods Enzymol.* **2016**, *566,* 59-87.
[9] J. Zhou, J. Rossi, *Nat Rev Drug Discov.* **2017**, *16,* 181-202.
[10] M. J. Damha, P. A. Giannaris, P. Marfey, L. S. Reid, *Tetrahedron Lett.* **1991**, *32*, 2573-2576.
[11] H. Urata, K. Shinohara, E. Ogura, Y. Ueda, M. Akagi, *J. Am. Chem. Soc.* **1991**, *113*, 8174-8175.
[12] K. P. Williams, X.-H. Liu, T. N. Schumacher, H. Y. Lin, D. A. Ausiello, P. S. Kim, D. P. Bartel, *Proc Natl Acad Sci U S A.* **1997**, *94*, 11285-11290.
[13] B. E. Young; N. Kundu; J. T. Sczepanski, *Chem. Eur. J.* **2019,** *25*, 7981-7990.
[14] S. Dey, J. T. Sczepanski, *Nucleic Acids Res.* **2020**, *48*, 1669-1680.
[15] J. C. Rosch, D. A. Balikov, F. Gong, E. S. Lippmann, *Eng. Rep.* **2020**, 2:e12089.
[16] J. T. Sczepanski, G. F. Joyce, *J. Am. Chem. Soc.* **2015**, *137*, 16032–16037.
[17] A. Vater, S. Klussmann, *Drug Discov Today* **2015**, *20*, 147-155.
[18] J. T. Sczepanski, G. F. Joyce, *Nature* **2014**, *515* , 440-442.
[19] K. F. Tjhung, J. T. Sczepanski, E. R. Murtfeldt, G. F. Joyce, *J. Am. Chem. Soc.* **2020**, *142*, 15331-15339.
[20] C. Lin, Y. Ke, Z. Li, J. H. Wang, Y. Liu, H. Yan, *Nano Lett.* **2009**, *9*, 433-436.
[21] H. B. D. Thai, K.-R. Kim, K. T. Hong, T. Voitsitskyi, J.-S. Lee, C. Mao, D.-R. Ahn, *ACS Cent Sci* **2020**, *6*, 2250-2258.
[22] Y. Dantsu, Y. Zhang, W. Zhang, *Org. Lett.* **2021**, *23*, 5007-5011.
[23] F. Wohlrab, T. Haertlé, T. Trichtinger, W. Guschlbauer, *Nucleic Acids Res.* **1978**, *5*, 4753-4760.
[24] S. Wärmländer, J. E. Sponer, J. Sponer, M. Leijon, *J Biol Chem* **2002**, *277*, 28491-28497.

Entry for the Table of Contents

Insert graphic for Table of Contents here.



There is increasing demand for the novel therapeutics with high stability and specificity. Here, we chemically optimized the mirror-image nucleic acids, by accomplishing the synthesis of 2'-F-modified L-Cytidine, as well as various oligonucleotides. The T_m measurement characterized the basic thermal stability, and the high-resolution crystal structure of L-DNA revealed the detailed structural features and demonstrated the structural stability.