

## A Scalable Synthesis of $\alpha$ -L-Threose Nucleic Acid Monomers

Sujay P. Sau, Nour Eddine Fahmi, Ien-Yu. Liao, Saikat Bala, and John C. Chaput\*

<sup>†</sup>Department of Pharmaceutical Sciences, University of California, Irvine, California 92697, United States

Supporting Information

ABSTRACT: Recent advances in polymerase engineering have made it possible to copy information back and forth between DNA and artificial genetic polymers composed of TNA ( $\alpha$ -L-threofuranosyl-(3',2') nucleic acid). This property, coupled with enhanced nuclease stability relative to natural DNA and RNA, warrants further investigation into the structural and functional properties of TNA as

an artificial genetic polymer for synthetic biology. Here, we report a highly optimized chemical synthesis protocol for constructing multigram quantities of TNA nucleosides that can be readily converted to nucleoside 2'-phosphoramidites or 3'triphosphates for solid-phase and polymerase-mediated synthesis, respectively. The synthetic protocol involves 10 chemical transformations with three crystallization steps and a single chromatographic purification, which results in an overall yield of 16-23% depending on the identity of the nucleoside (A, C, G, T).

#### INTRODUCTION

TNA ( $\alpha$ -L-threofuranosyl-(3',2') nucleic acid) is an artificial genetic polymer in which the natural ribose sugar found in RNA has been replaced by the tetrose sugar of  $\alpha$ -L-threose. In contrast to natural DNA and RNA, which have a six-atom backbone repeat unit connected by 3',5'-phosphodiester linkages, TNA has a backbone periodicity of five atoms (or bonds) with phosphodiester linkages occurring at the 2' and 3' vicinal positions of the threose sugar. Despite this major structural difference, TNA is capable of forming stable antiparallel Watson-Crick duplexes and shows efficient crosspairing with complementary strands of DNA and RNA.<sup>1-3</sup> The ability to transfer genetic information between TNA and RNA, coupled with the chemical simplicity of threose relative to ribose, has fueled interest in TNA as a possible progenitor of RNA during a hypothetical period in life's history known as the RNA world. 4-6 TNA is also being explored as a source of nuclease resistant affinity reagents (aptamers) and catalysts for synthetic biology and molecular medicine, as the constitutional structure of TNA is stable under biological conditions. Ongoing efforts in both of these areas have inspired researchers to engineer polymerases that can "transcribe" DNA into TNA and other polymerases that can "reverse transcribe" TNA back into 12 By including a selective amplification step in the replication cycle, methods are now being developed to isolate functional TNA molecules by in vitro selection.

Motivated by a desire to explore the structural and functional properties of TNA by in vitro selection and atomic-level structure determination studies by solution NMR and X-ray crystallography, we have systematically evaluated the chemical synthesis of TNA monomers with the goal of devising a synthetic pathway that is amenable to large-scale (multigram) synthesis. A new chemical synthesis strategy was necessary because the original approach (Scheme 1) first described by Eschenmoser and colleagues suffered from a number of

Scheme 1. Original Strategy Developed to Synthesize α-L-Threose Nucleosides<sup>15</sup>

shortcomings that included low overall yield (2-6%), numerous silica gel purification steps, and poor regioselectivity during nucleoside tritylation. While the two DMT regioisomers (3' and 2') 5a and 5b can be chromatographically resolved, only trace amounts of the 2'-isomer is obtained in the case of adenosine and cytosine, which necessitates an additional series of protection-deprotection steps to synthesize the nucleotide 3'-triphosphates of adenosine and cytosine.<sup>12</sup>

Here, we describe a highly optimized chemical synthesis strategy for generating TNA nucleoside precursors that can be readily converted to 2-phosphoramidites for solid-phase synthesis or 3'-triphosphates for polymerase-mediated primer-extension reactions. Our approach involves a total of 10 chemical transformations with three crystallization and a single chromatographic purification steps and results in an overall yield of 16-23% depending on the identity of the nucleoside (A, C, G, T). We demonstrate that this new strategy can be

Received: December 8, 2015 Published: February 19, 2016

2302

<sup>&</sup>lt;sup>‡</sup>The Biodesign Institute, Arizona State University, Tempe, Arizona 85287-5301, United States

The Journal of Organic Chemistry

used to produce multigram quantities of TNA monomers required to explore the structural and functional properties of TNA polymers.

## ■ RESULTS AND DISCUSSION

The synthesis of L-threonolactone (2) from L-ascorbic acid (vitamin C) has been reported previously (Scheme 1). Accordingly, L-ascorbic acid undergoes oxidative cleavage in the presence of hydrogen peroxide and CaCO<sub>3</sub> to produce the calcium salt of L-threonate (6). Compound 6 is then converted to the desired L-threonolactone (2) using a Dowex cation exchange resin to promote acidification and intramolecular lactonization. However, due to the limited solubility of 6, both steps require large volumes of water that must be removed prior to cyclization and crystallization.

We began by examining the crystallization of 6 from the resulting aqueous reaction mixture. Literature methods require reducing the volume of the aqueous solution prior to crystallization with methanol; however, this is a tedious process, especially when the reaction is performed on a large scale. In an attempt to avoid this step, we examined possible conditions for obtaining pure 6 without volumetric reduction under reduced pressure. Through minimal optimization, we found that the addition of 2 volumetric equivalents of methanol directly to the reaction mixture allowed 6 to precipitate as a white crystalline material in 85% yield. This subtle change to the protocol resulted in a yield that compared favorably against previous reports  $(65-79\%)^{15,17,19}$  and allowed large-scale reactions to be performed with substantially higher throughput.

Next, we probed the conversion of 6 to 2 (Scheme 2). While previous literature methods invoke the use of a Dowex cation

# Scheme 2. Synthesis of Universal Glycosyl Donor for TNA Nucleosides <sup>418</sup>

HO 
$$\stackrel{\text{HO}}{\longrightarrow}$$
 HO  $\stackrel{\text{HO}}{\longrightarrow}$  HO  $\stackrel{\text{O}}{\longrightarrow}$  HO  $\stackrel{\text{O}}$ 

"Reagents and conditions: (a) (i) CaCO<sub>3</sub>, 30% aq. H<sub>2</sub>O<sub>2</sub>, H<sub>2</sub>O,18 h, 0 °C to rt; (ii) active charcoal, 70 °C, 2 h, 85%; (b) oxalic acid, *para*toluenesulfonic acid (cat), CH<sub>3</sub>CN, 2 h, reflux, 93%; (c) benzoyl chloride, 1:10 pyridine—CH<sub>2</sub>Cl<sub>2</sub>, 0.5 h, 0 °C, 64%; (d) *tert*-butyldiphenylchlorosilane, imidazole, DMAP (cat), CH<sub>2</sub>Cl<sub>2</sub>, 18 h, 0 °C to rt; (e) 1 M DIBAL-H in toluene, dry 1,2-dimethoxyethane, 0.5 h, -78 °C; (f) Ac<sub>2</sub>O–DMAP (5 equiv, 1.5 equiv) in CH<sub>2</sub>Cl<sub>2</sub>, 2 h, -78 °C to rt, 95% from 7.

exchange resin to generate threonic acid from **6**, such methodology is cost prohibitive for routine large-scale ( $\geq$ 100 g) synthesis. We, therefore, sought an alternative method for removing calcium from the reaction mixture and promoting the subsequent *in situ* cyclization to the threonolactone. For this purpose, we examined several calcium-precipitating reagents such as hydrofluoric acid, sulfuric acid, and phosphoric

acid. Of these, dilute sulfuric acid produced the desired compound in crystalline form with yields that are comparable to what has been achieved previously using Dowex. However, because of the hygroscopic nature of L-threonolactone (2), crystallization failed whenever the compound was not properly dried.

As an alternative to dilute  $H_2SO_4$ , we tested oxalic acid as an in situ calcium exchange reagent. We found that refluxing a heterogeneous mixture of 6 in acetonitrile with 1 equiv of oxalic acid and a catalytic amount of para-toluenesulfonic acid produced the desired L-threonolactone (2) and calcium oxalate as an insoluble precipitate that could be easily separated by filtration. Using this approach, pure 2 is obtained as a white solid in 93% yield after coevaporation with ethyl acetate. In addition to eliminating the need for an expensive cation exchange resin, oxalic acid provided a streamlined method for obtaining large quantities of L-threonolactone (2) without the strict need for rigorous drying of a dilute aqueous solution to an anhydrous state.

A major limitation of the original synthetic route developed by Eschenmoser and colleagues was the low regioselectivity observed when  $\alpha$ -L-threofuranosyl nucleosides 4 are tritylated with DMT-Cl (Scheme 1). In most cases, tritylation produced a mixture of 2' and 3' regioisomers that could be separated by careful silica gel chromatography to generate pure 5a and 5b. In principle, this synthetic strategy would have been ideal if 5a and 5b were obtained in equal amounts; however, in practice, the tritylation of adenosine and cytosine  $\alpha$ -L-threofuranosyl nucleosides yielded only trace amounts of the 2'-DMT isomer (5b). Thus, in order to obtain the precursor compounds required to synthesize TNA 3'-nucleoside triphosphates, a cumbersome strategy of protection and deprotection was developed to convert 5a into 5b. 14

In an effort to generate antiviral compounds based on the structural framework of TNA nucleosides, Herdewijn and colleagues developed a regioselective strategy that defined the 2' and 3' hydroxyl positions early in the synthetic pathway. 18,20 Accordingly, the authors reported that 2 could be selectively converted to the 2-O-benzoyl L-threonolactone derivative (7) by the addition of benzoyl chloride to a solution of 2 and imidazole in acetonitrile after 8 h of stirring from −5 °C to room temperature (Scheme 2). Although this method proved unsatisfactory in our hands, selective benzoylation of 2 was achieved with the addition of 1 equiv of benzoyl chloride to 2 in 1:10 pyridine-dichloromethane solution after stirring for 30 min at 0 °C. Following the reaction, we further discovered that 7 could be precipitated as a pure white powder in 64% yield after stirring overnight at 4 °C from a mixture of dichloromethane and hexanes.

In analogy to previous work on threofuranosyl nucleosides, <sup>18</sup> compound 7 was converted to the universal glycosyl donor 8 following three successive chemical transformations. Accordingly, 7 was silylated at the 3-hydroxy position with *tert*-butyldiphenylchlorosilane, followed by DIBAL-H reduction of the lactone to the lactol, and finally acetylation at the anomeric position to give 8 in 96% yield from 7. Filtration of 8 through a short-pad of silica gel proved highly facile and efficient (see <sup>1</sup>H NMR in the Supporting Information) due to the strong polarity differences between 8 and the nonwashable reagents.

From 8, we constructed a complete set of TNA nucleosides for thymidine (T, 9c), cytidine (C, 9b), adenosine (A, 9a), and guanosine (G, 9d) using a Vorbrüggen glycosylation (Scheme 3) that involved heating the glycosyl donor and desired

The Journal of Organic Chemistry

Scheme 3. Synthesis of TNA Phosphoramidite Precursors<sup>a</sup>

"Reagents and conditions: (a) thymine, *N,O*-bis(trimethylsilyl)-acetamide (BSA), trimethylsilyltriflate (TMSOTf), CH<sub>3</sub>CN, 2 h, 60 °C; (b) *N*<sup>4</sup>-benzoylcytosine, BSA, TMSOTf, CH<sub>3</sub>CN, 2 h, 60 °C; (c) *N*<sup>6</sup>-benzoyladenine, BSA, TMSOTf, toluene, 2.5 h, 95 °C; (d) (i) *N*<sup>2</sup>-acetyl-*O*<sup>6</sup>-diphenylcarbamoyl-guanine, BSA, 1,2-dichloroethane, 0.5 h 70 °C (ii) TMSOTf, CH<sub>3</sub>CN, toluene, 1.5 h, 70 °C; (e) 1 M TBAF in THF, 1 h, 0 °C, 64% for 9a, 70% for 9b, 43.5% for 9c, 69% for 9d; (f) DMT-Cl, DMAP, pyridine, 18 h, 70 °C; (g) 1 M NaOH, THF–MeOH, 20 min, 0 °C, 71% for 10a, 65% for 10b, 63% for 10c; (h) DMT-Cl, AgOTf, 2,4,6-trimethylpyridine, dry CH<sub>2</sub>Cl<sub>2</sub> 18 h, 70 °C, 52% for 10d.

nucleobase (in protected form) in the presence of trimethylsilyl trifluoromethane-sulfonate (TMSOTf). After workup, the 3'tert-butyldiphenylsilyl protecting group was removed following treatment with tetrabutylammonium fluoride for 1 h at 0 °C. Glycosylation and desilylation proceeded smoothly for thymine,  $N^4$ -benzoyl cytosine, and  $N^2$ -acetyl- $O^6$ -diphenylcarbamoyl guanine to give 9a, 9b, and 9d, respectively, in 64-70% yield after crystallization. However, glycosylation with N6benzoyl adenine resulted in a 3:2 mixture of  $N^9$ - and  $N^7$ regioisomers. Surprisingly, similar isomeric mixtures were obtained when SnCl<sub>4</sub> was substituted for TMSOTf, which gave exclusively the  $N^9$ -regioisomer for glycosylation with 1,2,3tribenzoyl threose. 15 Although the  $N^9$ - and  $N^7$ -regioisomers are separable by silica gel chromatography, we found that the crude isomeric mixture could be converted to the thermodynamically favored  $N^9$ -isomer by heating in dry toluene in the presence of 1 equiv of TMSOTf for 1 h at 80 °C. Further optimization showed that glycosylation in toluene at 95 °C provided the desired N<sup>9</sup>-regioisomer as the major product. Crystallization

after desilylation gave pure  $N^9$ -adenosine nucleoside (9c) in 43.5% yield.

Compounds 9a-d are key intermediates in the divergent synthesis of TNA nucleoside-3'-triphosphates and 2'-phosphoramidites. For L-threofuranosyl nucleoside 3'-triphosphates, compounds 9a-d can be phosphorylated using the standard Ludwig and Eckstein method, followed by treatment with concentrated NH4OH to remove the sugar and nucleobase protecting groups. 14 For L-threofuranosyl nucleoside 2'phosphoramidites, compounds 9a-d are tritylated with DMT-Cl and treated with 1 M NaOH for 20 min at 0 °C to remove the 2'-benzoyl group to obtain compounds 10a-d (Scheme 3).<sup>17</sup> While the tritylation of 9a-c proceeded efficiently under standard conditions, tritylation of 9d required rigorous azeotropic removal of water and the use of AgOTf as the catalyst. Removal of the O<sup>6</sup>-diphenylcarbamoyl (DPC) group with 90% TFA or glacial acetic acid allowed tritylation to occur under standard conditions, suggesting that the bulky DPC group limits addition of the DMT by steric hindrance.<sup>21,3</sup>

In our laboratory, we routinely synthesize TNA nucleosides 9a-d using the above-mentioned methodology. In a typical synthesis run, 125 g of 1 is converted into 80 g of 7 with an overall yield of 49%. In contrast to previously described methods, modifications to the purification protocol of 6 and 2 makes it possible to generate large quantities of highly pure 7 in just 5 days. The conversion of 7 to 8 is generally performed with 16.5 g of 7 to obtain 35 g of 8 with an overall yield of 96% in 24 h. The N-glycosylation of 8, followed by desilylation, affords nucleosides 9a-d in 64-70% yield, with the noted exception of 9c, which is obtained in slightly lower yield (43.5%). We have found that the synthesis of 9a-d from 7 can be performed on scales as large as 50 g, which is a dramatic improvement that will help accelerate the synthesis and characterization of TNA polymers.

In summary, we provide a scalable and highly optimized synthesis protocol for constructing L-threofuranosyl nucleosides that are immediate precursors of TNA triphosphates and phosphoramidite monomers. Several key challenges have been resolved and purification steps were minimized to increase the yield and throughput of key intermediates. The conversion of vitamin C to L-threonolactone was optimized to minimize the cost and labor required for generating a key intermediate in the synthesis pathway. The universal glycosyl donor 8 was prepared from 7 in three chemical steps, followed by filtration through a short pad of silica gel. The complete set of TNA nucleosides (9a-d) were synthesized in two reactions from 8 and purified by crystallization. Final tritylation, followed by debenzoylation, provided the desired compounds (10a-d) after silica gel column chromatography. Together, these changes make it possible to synthesize multigram-scale quantities of TNA monomers that can be used to synthesize TNA polymers.

## **■ EXPERIMENTAL SECTION**

**General Methods.** Except as otherwise noted, all nonaqueous reactions were carried out in oven-dried glassware under a balloon pressure of argon or nitrogen. Reagents were commercially available and used as received; anhydrous solvents were purchased as the highest grade. Reactions were monitored by thin-layer chromatography using 0.25 mm Silicycle or EM silica gel 60 F<sub>2.54</sub> plates. Column chromatography was performed using Silicycle 40–60 mesh silica gel. Yields are reported as isolated yields of spectroscopically pure compounds.  $^1\mathrm{H}$  and  $^{13}\mathrm{C}$  NMR spectra were obtained using 400 and 500 MHz NMR spectrometers. Chemical shifts are reported in parts per million (ppm,  $\delta$ ) referenced to the  $^1\mathrm{H}$  resonance of TMS.  $^{13}\mathrm{C}$ 

spectra are referenced to the residual  $^{13}$ C resonance of the solvent (DMSO- $d_6$ , 39.52 ppm, CD<sub>3</sub>OD, 49.00 ppm). Splitting patterns are designated as follows: s, singlet; br, broad; d, doublet; dd, doublet of doublets: t, triplet: q, quartet: m, multiplet.

doublets; t, triplet; q, quartet; m, multiplet. **Calcium-L-threonate** (6). To a cold  $(0-5 \, ^{\circ}\text{C})$  solution containing 125 g (0.71 mol) of L-ascorbic acid (1) in 1 L of H<sub>2</sub>O was slowly added 125 g (1.25 mol) of CaCO<sub>3</sub> with a spatula over 30 min (evolution of CO<sub>2</sub> was observed). To the resulting heterogeneous slurry was added 250 mL of 30% aq. H<sub>2</sub>O<sub>2</sub> dropwise over a period of ca. 1 h with stirring at 0-5 °C. The reaction mixture was allowed to warm to r.t and stirred overnight. The heterogeneous slurry was filtered, and the filter cake was washed with two-100 mL portions of H<sub>2</sub>O. The filtrate was treated with 25 g of activated carbon Darco G-60 and then heated to 50 °C, until peroxide was no longer detected using Merckoquant 1001-1 peroxide test strips. The hot suspension was filtered, and the solid material was washed with two-50 mL portions of H<sub>2</sub>O. The washings and the filtrate were combined and crystallized by the addition of 2 volume equivalents of methanol while stirring for 16 h at 4 °C. The solid material was filtered, washed with two-50 mL portions of MeOH, and dried under high vacuum at 40 °C. Calcium Lthreonate monohydrate (6) was obtained as a white solid: yield 99.4 g (85%); <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  3.60 (dd, 1H, I = 11.6, 7.6 Hz),

3.66 (dd, 1H, J=11.6, 5.2 Hz), 3.95 (m, 1H) and 4.03 (m, 1H). L-Threonolactone (2). To a suspension of calcium L-threonate (6) (99.4 g, 0.30 mol) in dry acetonitrile (500 mL) were added anhydrous oxalic acid (28.8 g, 0.32 mol) and para-toluenesulfonic acid monohydrate (1.0 g). The heterogeneous mixture was stirred at reflux for 3 h. The hot mixture was allowed to cool to room temperature and filtered. The filter cake was washed with 50 mL of acetonitrile, and the combined filtrate was evaporated under reduced pressure to produce a colorless syrup. The residue was dissolved in 100 mL of EtOAc and evaporated to dryness to give L-threonolactone (2) as a white solid: yield 66.5 g (93%); silica gel TLC  $R_{\rm f}$  0.36 (EtOAc),  $^{1}$ H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  3.93 (dd, 1H, J=8.8, 7.2 Hz), 4.20 (d, 1H, J=7.2Hz), 4.29 (dd, 1H, J=14.0, 6.8 Hz) and 4.41 (dd, 1H, J=9.8, 6.8 Hz), 4.82 (s, 2H).

2-O-Benzoyl-L-threonolactone (7).<sup>18</sup> To a cold (0-5 °C) solution containing 66.5 g (0.56 mol) of L-threonolactone (2) in 1.2 L of CH<sub>2</sub>Cl<sub>2</sub> and 135 mL of anhydrous pyridine under argon was added benzoyl chloride (72.0 mL, 0.62 mol) dropwise. The mixture was stirred under argon for 30 min at 4 °C. The crude reaction mixture was then sequentially washed with 1 N HCl (3 × 400 mL) and brine (200 mL). The combined aqueous layer was back extracted with two-100 mL portions of  $CH_2C\overline{l}_2$ . To the combined organic layers was added 2 volume equivalents of hexane over 1 h, and the solution was left stirring for 16 h at 4 °C. The product was collected by filtration and dried under vacuum. 2-O-Benzoy-L-threonolactone (7) was obtained as a white solid: yield 80.1 g (64%); silica gel TLC R<sub>f</sub> 0.30 (1:1 hexanes-EtOAc); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  3.81 (brs, 1H), 4.16 (dd, 1H, J = 9.2, 7.6 Hz), 4.64 (dd, 1H, J = 9.2, 8.0 Hz), 4.73 (q, 1H, J = 7.6 Hz), 5.41 (d, 1H, J = 6.8 Hz), 7.48 (t, 2H, J = 8.0 Hz), 7.64 (t, 1H, J = 8.0 Hz) and 8.10 (d, 2H, J = 7.6 Hz).

1-O-Acetyl-2-O-benzoyl-3-O-tert-butyldiphenylsilyl-1-threofuranose (8). <sup>18</sup> To a cold (0–5 °C) solution containing 16.5 g (74.3 mmol) of 2-O-benzoyl-1-threonolactone (7), 60 mg of DMAP and 10.2 g (150 mmol) of imidazole in 160 mL of  $\rm CH_2Cl_2$  was added dropwise tert-butyldiphenylchlorosilane (20 mL, 75.0 mmol). The reaction mixture was warmed to r.t. and stirred under argon for 16 h. The solvent was evaporated under reduced pressure, and the residue was dissolved in 300 mL of hexane. The organic phase was sequentially washed with 1 N HCl (100 mL), H<sub>2</sub>O (100 mL), and brine (100 mL). The organic layer was dried over MgSO<sub>4</sub> and evaporated to give 2-O-benzoyl-3-O-tert-butyldiphenylsilyl-1-threonolactone as a colorless syrup (35 g). The crude material was used directly without further purification; silica gel TLC ( $R_{\rm f}$  0.50 in 3:1 hexanes—EtOAc).

To a cold  $(-78\,^{\circ}\text{C}$ , acetone/dry ice) solution containing 35.0 g of crude 2-*O*-benzoyl-3-*O*-tert-butyldiphenylsilyl-L-threonolactone in 150 mL of 1,2-dimethoxyethane was added a 1 M solution of DIBAL-H in toluene (100 mL, 100 mmol) dropwise over a 10 min period. The mixture was stirred under argon for an additional 20 min at  $-78\,^{\circ}\text{C}$ .

TLC showed the reaction to be complete (R<sub>f</sub> 0.35 in 3:1 hexanes-EtOAc). A premade solution containing Ac<sub>2</sub>O (367 mmol) and DMAP (115 mmol) in  $CH_2Cl_2$  ( $Ac_2O/DMAP/CH_2Cl_2 = 35 \text{ mL}/14.0$ g/40 mL) was then added dropwise at -78 °C. After 10 min, the reaction mixture was removed from the dry ice bath and allowed to stir for 1.5 h while warming to r.t. The mixture was then diluted with 200 mL of hexanes and poured into a cold stirring 1 N aq. HCl solution (200 mL). The organic layer was separated and washed with H<sub>2</sub>O (100 mL), sat. aq. NaHCO3 (100 mL), and brine (100 mL). The organic layer was dried over MgSO<sub>4</sub> and evaporated to give 42 g of crude product as a yellow syrup. The product was suspended in 25 mL of 10% EtOAc in hexanes, passed through a short pad of silica (50 g), and washed out with 10% EtOAc in hexanes (200 mL). An anomeric mixture of 1-O-acetyl-2-O-benzoyl-3-O-tert-butyldiphenylsilyl-L-threofuranose (8) was obtained as a colorless syrup: yield 36.0 g (96%); silica gel TLC R<sub>f</sub> 0.50 (3:1 hexanes-EtOAc).

**1-(2'-O-Benzoyl-α-1-threofuranosyl)thymine (9a).** To a solution containing 34.0 g (67.4 mmol) of 1-*O*-acetyl-2-*O*-benzoyl-3-*Otert*-butyldiphenylsilyl-L-threofuranose (8) and 8.9 g (70.5 mmol) of thymine in 140 mL of anhydrous acetonitrile was added 45 mL (145 mmol) of *N*,*O*-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 60 °C. TMSOTf (20.0 mL, 108 mmol) was added dropwise, and stirring was continued for another 2 h at 60 °C, after which time TLC analysis (1:1 hexanes—EtOAc) showed the reaction to be complete. The mixture was cooled to r.t, diluted with 300 mL of EtOAc, and poured into 200 mL of cold sat. aq. NaHCO<sub>3</sub> solution with stirring. The organic layer was separated and washed with  $\rm H_2O$  (100 mL) and brine (100 mL), dried over MgSO<sub>4</sub>, and concentrated under reduced pressure. The crude nucleoside was obtained as a white foam (41 g) and was used directly without further purification; silica gel TLC  $\rm R_f$  0.37 (1:1 hexanes—EtOAc).

To a cold (0-5 °C) solution containing 41 g of crude 1-(2'-Obenzoyl-3'-O-tert-butyldiphenylsilyl- $\alpha$ -L-threofuranosyl)thymine in THF (250 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 70.0 mL, 70.0 mmol), and the mixture was stirred for 1 h at 0-5 °C. The solvent was evaporated under reduced pressure, and the residue was dissolved in 300 mL of EtOAc. The organic layer was separated and washed twice with H<sub>2</sub>O (100 mL) and brine (100 mL), dried over MgSO<sub>4</sub>, and concentrated under reduced pressure to give 20 g of the crude nucleoside as a yellow syrup. The syrup was dissolved in 50 mL of CH<sub>2</sub>Cl<sub>2</sub> and crystallized by the addition of 50 mL of hexanes. 1-(2'-O-Benzoyl-α-L-threofuranosyl)thymine (9a) was obtained as a white solid: yield 14.44 g (64%); silica gel TLC R<sub>f</sub> 0.35 (1:2 hexanes-EtOAc); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  1.90 (s, 3H), 4.05 (brs, 1H), 4.20 (dd, 1H I = 4.4, 10.0 Hz), 4.29 (d, 1H, J = 10.0 Hz), 4.51 (brs, 1H), 5.43 (s, 1H), 5.87 (d, 1H, J = 2.0Hz), 7.36 (s, 1H), 7.47 (t, 2H, J = 7.6 Hz), 7.61 (t, 1H, J = 7.2 Hz), 8.02 (d, 2H, J = 8.2 Hz) and 8.90 (brs, 1H); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD):  $\delta$  12.5, 74.7, 76.5, 83.9, 91.4, 111.2, 129.7, 130.4, 130.8, 134.8, 138.7, 152.3, 166.4, 166.6; HRMS (ESI-TOF): [M + Na]<sup>+</sup> calcd for C<sub>16</sub>H<sub>16</sub>N<sub>2</sub>O<sub>6</sub>Na, 355.0906; found, 355.0910.

 $N^4$ -Benzoyl-1-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)cytosine (9b). To a solution containing 33.5 g (66.3 mmol) of 1-O-acetyl-2-Obenzoyl-3-O-tert-butyldiphenylsilyl-L-threofuranose (8) and 15.0 g (70.0 mmol) of  $N^4$ -benzoylcytosine in 150 mL of anhydrous acetonitrile was added 40.0 mL (154 mmol) of N,O-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 60 °C. TMSOTf (38.0 mL, 199 mmol) was added dropwise, and stirring was continued at 60 °C for another 2 h, after which time TLC (1:1 hexanes-EtOAc) showed the reaction to be complete. The mixture was cooled to room temperature, diluted with 200 mL of EtOAc, and poured into 200 mL of sat. aq. NaHCO3 solution with stirring. The white suspension was filtered over Celite, and the organic layer was separated and washed with H<sub>2</sub>O (150 mL) and brine (150 mL), dried over MgSO<sub>4</sub>, and concentrated under reduced pressure. The crude nucleoside was obtained as a white foam (36.0 g) and was used directly without further purification; silica gel TLC R<sub>f</sub> 0.75 (EtOAc).

To a cold (0–5 °C) solution containing 36.0 g of crude  $N^4$ -benzoyl-1-(2'-O-benzoyl-3'-O-tert-butyldiphenylsilyl- $\alpha$ -L-threofuranosyl)-cytosine in THF (200 mL) was added dropwise tetrabutylammonium

fluoride (1 M solution in THF, 70.0 mL, 70.0 mmol), and the mixture was stirred at 0–5 °C for 1 h. The solvent was evaporated under diminished pressure, and the residue was dissolved in 500 mL of EtOAc. The organic layer was washed with 1 N aq. HCl (150 mL) and H<sub>2</sub>O (150 mL), and then was stirred with brine (150 mL) at r.t. to precipitate the product.  $N^4$ -Benzoyl-1-(2′-O-benzoyl- $\alpha$ -L-threofuranosyl)cytosine (9b) was obtained as a white solid: yield 19.4 g (70%); silica gel TLC  $R_f$  0.40 (1:2 hexanes–EtOAc); <sup>1</sup>H NMR (400 MHz, DMSO- $d_6$ )  $\delta$  4.3 (m, 3H), 5.39 (s, 1H), 5.85 (d, 1H J = 2.8 Hz), 5.99 (s, 1H), 7.41 (d, 1H, J = 7.2 Hz), 7.50–7.64 (m, 5H), 7.72 (t, 1H, J = 7.2 Hz), 8.03 (t, 4H, J = 7.6 Hz), 8.22 (d, 1H, J = 7.6 Hz), and 11.27 (s, 1H); <sup>13</sup>C NMR (125 MHz, DMSO- $d_6$ ):  $\delta$  72.4, 76.6, 81.5, 90.9, 95.7, 128.5, 128.9, 129.5, 132.8, 133.2, 133.9, 145.7, 154.5, 163.4, 164.5, 167.4; HRMS (ESI-TOF):  $[M + Na]^+$  calcd for  $C_{22}H_{19}N_3O_6Na$ , 444.1172; found, 444.1180.

 $N^6$ -Benzoyl-9-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)adenine (9c). A mixture of 40.5 g (80.0 mmol) of 1-O-acetyl-2-O-benzoyl-3-O-tert-butyldiphenylsilyl-L-threofuranose (8) and 19.1 g (80.0 mmol) of N<sup>6</sup>-benzoyladenine was coevaporated with 50 mL of anhydrous toluene and then suspended in 160 mL of anhydrous toluene. To the suspension was added 42.0 mL (163.8 mmol) of N,O-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 95 °C. Once all the suspension was dissolved, TMSOTf (22.2 mL, 118.8 mmol) was added dropwise, and stirring was continued at 95 °C for another 2.5 h, after which TLC (1:1 hexanes-EtOAc) showed complete consumption of the starting material and more polar major products corresponding to the  $N^9$  glycosides. The mixture was cooled to room temperature, then poured into a stirring mixture of 150 g of ice and 150 mL of sat. aq. NaHCO3 solution and stirred for 30 min. The organic layer was separated and washed with H<sub>2</sub>O (50 mL) and brine (50 mL), dried over MgSO<sub>4</sub>, and concentrated under reduced pressure. The crude nucleoside was obtained as a yellow foam: yield 54.3 g.

To a cold  $(0-5 \, ^{\circ}\text{C})$  solution (ice-bath) containing 54.3 g of  $N^{6}$ benzoyl-9-(2'-O-benzoyl-3'-O-tert-butyldiphenylsilyl-α-L-threofuranosyl)adenine in THF (160 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 80.0 mL, 80.0 mmol), and the mixture was stirred at 0-5 °C for 30 min. The solvent was evaporated under diminished pressure, and the residue was dissolved in 200 mL of EtOAc. The organic layer was washed with H<sub>2</sub>O (50 mL) and brine (50 mL). The combined aqueous layer was back-extracted with EtOAc (2 × 50 mL). The combined organic layer was dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The resulting crude material was dissolved in 150 mL of 1% methanol in CH<sub>2</sub>Cl<sub>2</sub> and crystallized by slow addition of 120 mL of hexanes. N<sup>6</sup>-Benzoyl-9- $(2'-O-benzoyl-\alpha-L-threofuranosyl)$ adenine (9c) was obtained as a white solid: yield 15.5 g (43.5%); silica gel TLC  $R_{\rm f}$  0.35 (EtOAc); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  4.30 (dd, 1H J = 4.0, 10.0 Hz), 4.39 (d, 1H, J = 10.0 Hz), 4.67 (d, 1H, J = 3.2 Hz), 5.63 (s, 1H), 6.07 (d, 1H, J = 1.6 Hz), 7.56 (m, 6H), 8.04 (t, 4H, J = 8.1 Hz), 8.27(s, 1H), and 8.81 (s, 1H).

 $N^2$ -Acetyl-O<sup>6</sup>-diphenylcarbamoyl-9-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)guanine (9d). To a solution containing 8.45 g (26.1 mmol) of  $N^2$ -acetyl- $O^6$ -diphenylcarbamoyl guanine<sup>23</sup> in 200 mL of a mixture of anhydrous dichloroethane and toluene (1:2 v/v) was added 15.8 mL (64.6 mmol) of N,O-bis(trimethylsilyl)acetamide, and the mixture was stirred for 30 min at 70 °C. The solvent was removed under reduced pressure, and the residue was dissolved in 55 mL of anhydrous toluene. 1-O-Acetyl-2-O-benzoyl-3-O-tert-butyldiphenylsilyl-L-threofuranose (8) (11 g, 21.8 mmol) in 65 mL of anhydrous toluene was added dropwise using a canula, and the mixture was heated to 70 °C. TMSOTf (8.5 mL, 45.9 mmol) was then added dropwise, and the mixture was stirred at 70 °C for another 1.5 h, after which TLC (1:1 hexanes-EtOAc) showed complete consumption of (8). The mixture was cooled to room temperature, diluted with 300 mL of EtOAc, and poured into 150 mL of sat. aq. NaHCO3 solution with stirring, resulting in a purple suspension. The suspension was filtered, and the organic layer was separated and washed with H2O (150 mL) and brine (150 mL), dried over MgSO<sub>4</sub>, and concentrated under reduced pressure. The crude nucleoside was obtained as a

purple foam (16.12 g) and was used directly without further purification; silica gel TLC  $R_f$  0.7 (1:2 hexanes—EtOAc).

To a cold  $(0-5 \, ^{\circ}\text{C})$  solution containing 16.12 g of crude  $N^{2}$ -acetyl- $O^6$ -diphenylcarbamoyl-9-(2'-O-benzoyl-3'-O-tert-butyldiphenylsilyl- $\alpha$ -L-threofuranosyl)guanine in THF (50 mL) was added dropwise tetrabutylammonium fluoride (1 M solution in THF, 19.0 mL, 19.0 mmol), and the mixture was stirred at 0-5 °C. After 15 min, TLC was checked and another 9 mL of TBAF was added, and the mixture was stirred for 45 min. The solvent was evaporated under diminished pressure, and the residue was dissolved in 250 mL of EtOAc. The organic layer washed with 1 N aq. HCl (50 mL), H<sub>2</sub>O (50 mL), and brine (50 mL), dried over MgSO<sub>4</sub>, and concentrated under reduced pressure. The resulting crude product was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (30 mL) and crystallized by addition of an equal amount of hexanes.  $N^2$ -Acetyl- $O^6$ -diphenylcarbamoyl-9-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)guanine (9d) was obtained as a white solid: yield 8.9 g (69%); silica gel TLC R<sub>f</sub> 0.25 (1:2 hexanes-EtOAc); <sup>1</sup>H NMR (400 MHz, DMSO-d6)  $\delta$  2.20 (s, 3H), 4.32 (dq, 2H, J = 3.6, 7.6 Hz), 4.60 (m, 1H), 5.78 (t, 1H, J = 2.0 Hz), 5.96 (d, 1H, J = 4.0 Hz), 6.31 (d, 1H, J = 1.6 Hz), 7.33 (t, 2H, J = 6.0 Hz), 7.43–7.60 (m, 10H), 7.72 (t, 1H, J = 6.0 Hz), 8.03 (dd, 1H, J = 0.8, 6.4 Hz), 8.60 (s, 1H), and 10.81 (s, 1H);  $^{13}$ C NMR (125 MHz, DMSO- $d_6$ ):  $\delta$  24.6, 73.2, 74.8, 82.3, 87.9, 120.1, 128.7, 128.9, 129.5, 129.6, 134.0, 141.6, 144.3, 150.1, 152.3, 154.2, 155.3, 165.0, 169.2; HRMS (ESI-TOF): [M + Na]+ calcd for C<sub>31</sub>H<sub>26</sub>N<sub>6</sub>O<sub>7</sub>Na, 617.1761; found, 617.1750.

1-(3'-O-Dimethoxytrityl-α-L-threofuranosyl)thymine (10a). <sup>15</sup> A mixture containing 3.2 g (9.6 mmol) of 1-(2'-O-benzoyl-α-L-threofuranosyl)thymine (9a), 4.88 g (14.4 mmol, 1.5 equiv) of DMT-Cl, and 100 mg of DMAP (0.82 mmol) was coevaporated twice with 20.0 mL of anhydrous pyridine. The mixture was dissolved in 40 mL of anhydrous pyridine and stirred under argon at 80 °C for 16 h. The solvent was evaporated under diminished pressure, and the residue was partitioned between EtOAc (50 mL) and H<sub>2</sub>O (50 mL). The organic layer was washed with brine (50 mL), dried (MgSO<sub>4</sub>), and evaporated.

The residue was dissolved in a mixture of 38 mL of THF and 30.5 mL of MeOH, and then cooled to 0 °C. To the cold solution was added 15.0 mL of ice-cold 1 N aq. NaOH while stirring. After 30 min, the reaction was quenched by addition of 100 mL of water. The volatile solvent was removed under diminished pressure, and the residual aqueous solution was extracted with EtOAc (3 × 50 mL). The combined organic layer was dried (MgSO<sub>4</sub>) and evaporated. The residue was purified on a silica gel column (50 g silica bed), eluting with 50–100% EtOAc in hexanes containing 2% Et<sub>3</sub>N in steps of 5% increase in EtOAc for every 200 mL. 1-(3'-O-Dimethoxytrityl- $\alpha$ -L-threofuranosyl)thymine (10a) was obtained as a white foam: yield 3.6 g (71%);  $^{1}$ H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  1.82 (s, 3H), 3.30 (d, 1H, J = 9.6 Hz), 3.67 (s, 3H), 3.71 (s, 3H), 3.91 (d, 1H, J = 13.6 Hz), 4.19 (s, 1H), 4.23 (s, 1H), 5.72 (s, 1H), 6.79 (m, 4H), 7.12–7.38 (m, 9H), 7.44 (s, 1H) and 10.92 (s, 1H).

 $N^4$ -Benzoyl-1-(3'-O-dimethoxytrityl- $\alpha$ -L-threofuranosyl)-cytosine (10b). <sup>15</sup> A mixture of 3.18 g (7.7 mmol) of  $N^4$ -benzoyl-1-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)cytosine (9b), 3.84 g (11.3 mmol, 1.5 equiv) of DMT-Cl, and 120 mg of DMAP (0.99 mmol, 0.1 equiv) was coevaporated twice with 15.0 mL of anhydrous pyridine. The residue was suspended in 50 mL of anhydrous pyridine and stirred under argon for 16 h at 80 °C. The solvent was evaporated under diminished pressure, and the residue was partitioned between EtOAc (50 mL) and H<sub>2</sub>O (50 mL). The organic layer was washed with brine (50 mL), dried (MgSO<sub>4</sub>), and evaporated.

The residue was dissolved in a mixture of 38 mL of THF and 30.5 mL of MeOH, and then cooled to 0 °C. To the cold solution was added 7.7 mL of ice-cold 1 N aq. NaOH while stirring. After 10 min, another 7.7 mL of ice-cold 1 N aq. NaOH was added. Ten minutes after the second addition, the reaction was quenched by addition of 40 mL of 10% NH<sub>4</sub>Cl in water. The volatile solvent was removed under diminished pressure, and the residue was diluted by addition of 50 mL of EtOAc. The organic layer was separated, washed with H<sub>2</sub>O (50 mL) and brine (50 mL), dried (MgSO<sub>4</sub>), and evaporated. The residue was purified on a silica gel column (50 g silica bed), eluting with 70–100% EtOAc in hexane containing 1% Et<sub>3</sub>N in steps of 5% increase in EtOAc

for every 200 mL.  $N^4$ -Benzoyl-1-(3'-O-dimethoxytrityl- $\alpha$ -L-threofuranosyl)cytosine (10b) was obtained as a white foam: yield 3.12 g (65%);  $^1$ H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  3.32 (d, 1H, J = 10.0 Hz), 3.71 (d, 1H, J = 4.0 Hz), 3.74 (s, 6H), 4.25 (s, 1H), 4.33 (s, 1H), 5.70 (s, 1H), 6.79 (m, 4H), 6.75–7.35 (m, 9H), 7.47–7.67 (m, 4H), 7.94 (d, 2H, J = 8.0 Hz), 8.05 (d, 1H, J = 8.0 Hz), and 9.10 (brs, 1H).

 $N^6$ -Benzoyl-9-(3'-O-dimethoxytrityl- $\alpha$ -L-threofuranosyl-adenine (10c). <sup>15</sup> A mixture of 3.39 g (7.6 mmol) of  $N^6$ -benzoyl-9-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)adenine (9c), 3.86 g (11.4 mmol, 1.5 equiv) of DMT-Cl, and 93 mg (0.76 mmol, 0.1 equiv) of DMAP was coevaporated twice with 7.0 mL of anhydrous pyridine. The residue was dissolved in 35 mL of anhydrous pyridine, and the mixture was stirred at 75 °C under argon for 16 h. The solvent was evaporated under diminished pressure, and the residue was partitioned between EtOAc (50 mL) and  $H_2O$  (50 mL). The organic layer was washed with brine (50 mL), dried (MgSO<sub>4</sub>), and evaporated.

The residue was dissolved in a mixture of 38 mL of THF and 30.5 mL of MeOH, and then cooled to 0 °C. To the cold stirred solution was added 7.6 mL of ice-cold 1 N aq. NaOH solution. After 10 min, another 7.6 mL of ice-cold 1 N aq. NaOH solution was added. Ten minutes after the second addition, the reaction was quenched with 40 mL of 10% aq. NH<sub>4</sub>Cl. The volatile solvent was removed under diminished pressure, and the aqueous residue was diluted with 50 mL of EtOAc. The organic layer was separated, washed with H<sub>2</sub>O (50 mL) and brine (50 mL), dried (MgSO<sub>4</sub>), and evaporated. The residue was purified on a silica gel column (70 g silica bed), eluting with 30-80% EtOAc in hexanes containing 1% Et<sub>3</sub>N. Stepwise elution was performed with 10% increases in EtOAc for each 200 mL volume.  $N^6$ -Benzoyl-9-(3'-O-dimethoxytrityl- $\alpha$ -L-threofuranosyl)adenine (10c) was obtained as a white foam: yield 3.6 g (63%); <sup>1</sup>H NMR (400 MHz, DMSO- $d_6$ ) 3.45 (dd, 1H, J = 5.2, 9.6 Hz), 3.59 (dd, 1H, J = 3.2, 9.6 Hz), 3.73 (s, 6H), 4.17 (m, 1H), 4.50 (m, 1H), 5.91 (d, 1H, J = 2.8Hz), 5.93 (d, 1H, J = 4.8 Hz), 6.86 (m, 4H), 7.13-7.33 (m, 9H), 7.56(t, 2H, J = 7.6 Hz), 7.65 (t, 1H, J = 7.2 Hz), 8.06 (d, 2H, J = 7.2 Hz),8.56 (s, 1H), 8.75 (s, 1H), 11.22 (brs, 1H).

 $N^2$ -Acetyl- $O^6$ -diphenylcarbamoyl-9-(3'-O-dimethoxytrityl- $\alpha$ -L-threofuranosyl)guanine (10d).  $^{15}$  5.2 g (8.74 mmol) of  $N^2$ -acetyl- $O^6$ -diphenylcarbamoyl-9-(2'-O-benzoyl- $\alpha$ -L-threofuranosyl)guanine (9d) was coevaporated with 20 mL of anhydrous pyridine under reduced pressure, and the solid residue was dried under vacuum for 18 h. To the residue was added 5.92 g (17.48 mmol, 2 equiv) of DMT-Cl, anhydrous dichloromethane (60 mL), 2.3 mL of 2,4,6-trimethylpyridine (17.48 mmol, 2 equiv), and 675 mg (2.62 mmol, 0.3 equiv) of silver triflate. The mixture was stirred under nitrogen for 18 h. TLC (3:1 hexanes-EtOAc) showed the reaction to be complete. The solvent was removed under diminished pressure, and the residue was dissolved in 150 mL of dichloromethane. The organic layer was washed with 0.2 (N) HCl (300 mL), brine (60 mL), dried (MgSO<sub>4</sub>), and evaporated. The residue was purified on a silica gel column (4  $\times$ 15 cm), eluting with 100 mL of 99:1 DCM-Et<sub>3</sub>N to 300 mL of 97:2:1 DCM-MeOH-Et<sub>3</sub>N. The product was obtained as a pale yellow powder: yield 5.8 g (74.3%).

The compound (3.0 g, 3.34 mmol) was dissolved in 45 mL of THF-MeOH-water (5:4:1) and cooled to 0-5 °C in a water-ice bath. Ice-cold 1 N aq. NaOH (2.1 mL) was added dropwise to the mixture, and the mixture was stirred for 10-15 min. TLC (hexane-EtOAc 2:1) showed that the reaction was not complete. Another 2.1 mL of cold 1 N aq. NaOH was added dropwise, and stirring continued for another 15-20 min. The reaction was quenched with 15 mL of 10% aq. NH<sub>4</sub>Cl, and the crude product was extracted with EtOAc (2  $\times$ 50 mL). The combined organic layer was dried over MgSO<sub>4</sub> and evaporated under reduced pressure. The residue was purified on a silica gel column (4 × 15 cm) using a gradient elution with 100 mL of (85:14:1) hexane-EtOAc-Et<sub>3</sub>N, followed by 100-300 mL of (75:24:1) hexane-EtOAc-Et<sub>2</sub>N and (65:34:1) hexane-EtOAc-Et<sub>3</sub>N.  $N^2$ -Acetyl- $O^6$ -diphenylcarbamoyl-9-(3'-O-dimethoxytrityl- $\alpha$ -Lthreofuranosyl)guanine (10d) was obtained as a pale yellow powder: yield 1.9 g (71%); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  2.17 (s, 3H), 2.73 (s, 1H), 3.42 (s, 2H), 3.71 (s, 6H), 4.33 (m, 1H), 4.45 (s, 1H), 5.83

(d, 1H, *J* = 2.4 Hz), 6.78 (m, 4H), 7.12–7.45 (m, 19H), 8.19 (s, 1H), 8.81 (s, 1H).

## ASSOCIATED CONTENT

## **S** Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.joc.5b02768.

NMR data for compounds 2, 6, 7, 8, 9a-d, and 10a-d (PDF)

#### AUTHOR INFORMATION

## **Corresponding Author**

\*Tel.: 949-824-8149. E-mail: jchaput@uci.edu.

#### Notac

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

We wish to thank members of the Chaput lab and J. Heemstra for helpful comments and suggestions. This work was supported by the DARPA Folded Non-Natural Polymers with Biological Function Fold F(x) Program under award number N66001-14-2-4054 and a grant from the National Science Foundation grants (1304583).

#### REFERENCES

- (1) Schoning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. Science 2000, 290, 1347.
- (2) Yang, Y.-W.; Zhang, S.; McCullum, E. O.; Chaput, J. C. J. Mol. Evol. 2007, 65, 289.
- (3) Ebert, M.-O.; Mang, C.; Krishnamurthy, R.; Eschenmoser, A.; Jaun, B. J. Am. Chem. Soc. 2008, 130, 15105.
- (4) Orgel, L. E. Science 2000, 290, 1306.
- (5) Joyce, G. F. Nature 2002, 418, 214.
- (6) Chaput, J. C.; Yu, H.; Zhang, S. Chem. Biol. 2012, 19, 1360.
- (7) Chaput, J. C.; Szostak, J. W. J. Am. Chem. Soc. 2003, 125, 9274.
- (8) Chaput, J. C.; Ichida, J. K.; Szostak, J. W. J. Am. Chem. Soc. 2003, 125, 856.
- (9) Kempeneers, V.; Vastmans, K.; Rozenski, J.; Herdewijn, P. Nucleic Acids Res. 2003, 31, 6221.
- (10) Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C. J. Am. Chem. Soc. 2005, 127, 7427.
- (11) Yu, H.; Zhang, S.; Dunn, M.; Chaput, J. C. J. Am. Chem. Soc. 2013, 135, 3583.
- (12) Dunn, M. R.; Larsen, A. C.; Fahmi, N.; Zahurancik, W. J.; Meyers, M.; Suo, Z.; Chaput, J. C. J. Am. Chem. Soc. 2015, 137, 4014.
- (13) Yu, H.; Zhang, S.; Chaput, J. C. Nat. Chem. 2012, 4, 183.
- (14) Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. Org. Lett. 2005, 7, 1485.
- (15) Schoning, K.-U.; Scholz, P.; Wu, X.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. Helv. Chim. Acta 2002, 85, 4111.
- (16) Zhang, S.; Yu, H.; Chaput, J. C. Curr. Protoc. Nucleic Acid Chem. 2013, 52, 4.54.1.
- (17) Zhang, S.; Chaput, J. C. Curr. Protoc. Nucleic Acid Chem. 2012, 4,
- (18) Dumbre, S. G.; Jang, M.-Y.; Herdewijn, P. J. Org. Chem. 2013, 78, 7137.
- (19) Wei, C. C.; De Bernardo, S.; Tengi, J. P.; Borgese, J.; Weigele, M. J. Org. Chem. 1985, 50, 3462.
- (20) Wu, T.; Froeyen, M.; Kempeneers, V.; Pannecouque, C.; Wang, J.; Busson, R.; De Clercq, E.; Herdewijn, P. J. Am. Chem. Soc. 2005, 127, 5056.
- (21) Milecki, J.; Foldesi, A.; Fischer, A.; Adamiak, R. W.; Chattopadhyaya, J. J. Labelled Compd. Radiopharm. 2001, 44, 763.
- (22) Jang, M.-Y.; Froeyen, M.; Herdewijn, P. Tetrahedron Lett. 2013, 54, 6084.
- (23) Zou, R.; Robins, M. J. Can. J. Chem. 1987, 65, 1436.