Received: 9 April 2020 | Revised: 19 May 2020 | Accepted: 22 May 2020

DOI: 10.1002/bip.23388

FULL PAPER

Biopolymers WILEY

Synthesis and polymerase recognition of a pyrrolocytidine TNA triphosphate

Hui Mei[1,2] | Yajun Wang[1] | Eric J. Yik[1] | John C. Chaput[1]

[1]Departments of Pharmaceutical Sciences, Chemistry, Molecular Biology and Biochemistry, University of California, Irvine, CA, USA

[2]Shenzhen Key Laboratory of Synthetic Genomics, Guangdong Provincial Key Laboratory of Synthetic Genomics, CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Science, Shenzhen, China

Correspondence
John C. Chaput, Departments of Pharmaceutical Sciences, Chemistry, Molecular Biology and Biochemistry, University of California, Irvine, CA 92697-3958.
Email: jchaput@uci.edu

Funding information
National Science Foundation

Abstract

Synthetic genetics is an area of synthetic biology that aims to extend the properties of heredity and evolution to artificial genetic polymers, commonly known as xeno-nucleic acids or XNAs. In addition to establishing polymerases that are able to convert genetic information back and forth between DNA and XNA, efforts are underway to construct XNAs with expanded chemical functionality. α-L-Threose nucleic acid (TNA), a type of XNA that is recalcitrant to nuclease digestion and amenable to Darwinian evolution, provides a model system for developing XNAs with functional groups that are not present in natural DNA and RNA. Here, we describe the synthesis and polymerase activity of a cytidine TNA triphosphate analog (6-phenyl-pyrrolocytosine, tC^{p}TP) that maintains Watson-Crick base pairing with guanine. Polymerase-mediated primer extension assays show that tC^{p}TP is an efficient substrate for Kod-RI, a DNA-dependent TNA polymerase developed to explore the functional properties of TNA by *in vitro* selection. Fidelity studies reveal that a cycle of TNA synthesis and reverse transcription occurs with 99.9% overall fidelity when tC^{p}TP and 7-deaza-tGTP are present as TNA substrates. This result expands the toolkit of TNA building blocks available for *in vitro* selection.

KEYWORDS

synthetic biology, synthetic genetics, unnatural nucleobase, xeno-nucleic acids

1 | INTRODUCTION

The emerging field of synthetic genetics aims to explore the structural and functional properties of xeno-nucleic acid polymers, commonly referred to as XNAs, by *in vitro* selection.[1–3] To date, five different types of XNA polymers with nonribose sugars [1,5-anhydrohexitol nucleic acid (HNA), arabino nucleic acid (ANA), 2′-fluoroarabino nucleic acid (FANA), cyclohexenyl nucleic acid (CeNA), and α-L-threose nucleic acid (TNA)] have undergone iterative rounds of selection and amplification to produce the first examples of XNA aptamers and catalysts.[4–11] These results, which are only just beginning, have attracted significant attention due to the unique physicochemical properties of XNA relative to their natural counterparts.[12] Most XNAs are resistant to nuclease digestion, adopt altered helical geometries, and exhibit increased levels of chemical stability.[13–15] These features make XNA aptamers and catalysts ideal molecules for emerging applications in synthetic biology and biotechnology.[16]

In vitro selection is a powerful method for generating nucleic acid molecules with desired functional properties.[17] Moving beyond the natural bases of adenine (A), cytosine (C), guanine (G), and thymine (T) requires chemical synthesis to generate new XNA building blocks and polymerase-mediated replication systems that can recognize these molecules as substrates for XNA synthesis. Whereas an extensive amount of work has been done to augment the functionality of natural DNA and RNA,[18–22] similar efforts are only just beginning for XNAs. Some of the first examples include work by Holliger and co-workers to evaluate the fidelity of an *in vitro* selected RNA polymerase ribozyme,[23] and our recent synthesis of TNA polymers (Figure 1A) bearing a tricyclic

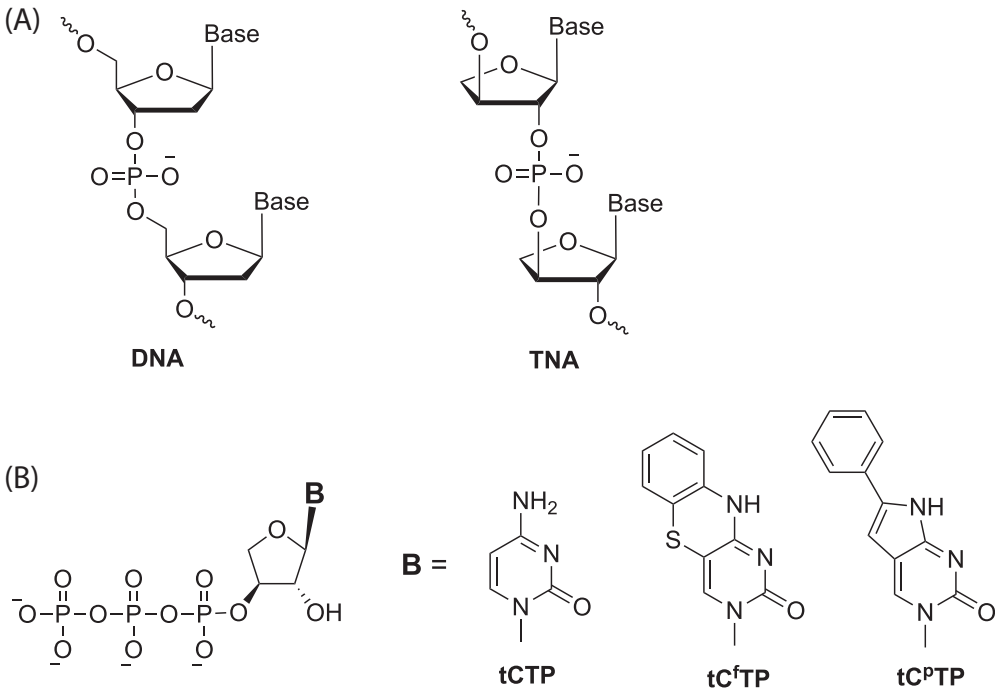


FIGURE 1 Molecular structure of TNA and TNA triphosphates. A, Constitutional structure for the linearized backbone of DNA (left) and α-L-threofuranosyl-(3′-2′) nucleic acid, TNA (right). B, Molecular structures of TNA cytidine triphosphates with natural and modified bases

phenothiazine ring system as a cytidine base (tC^f),[24] 7-deaza-guanine analogs,[25] and 5-alkynyl uracil analogs.[26] Another recent example by Hollenstein reveals the potential for side chain compatibility for 5-ethynyl uracil analogs of FANA.[27]

Previous studies using nucleic acids with expanded chemical functionalities have yielded DNA aptamers with enhanced functional activity.[28] In particular, hydrophobic functionalities were found to strongly promote the successful isolation of aptamers that bind with high affinity to their protein targets, as nicely exemplified by SOMAmers.[29] Related studies have shown the ability for diversity-enhancing functional groups to increase the likelihood of encoding shorter sequences and occupying nonoverlapping epitopes more frequently.[20] Finally, a systematic evaluation of chemically diverse functional groups has led to the understanding that nonpolar side chains outperform other chemical repertoires in head-to-head competition assays involving selection and replication.[30] These findings suggest that similar modifications could improve the quality of XNA aptamers, like TNA, that are being explored as next-generation protein affinity reagents.

Here, we report the chemical synthesis and polymerase recognition properties of a TNA nucleoside triphosphate of pyrrolocytosine (tC^{P}TP, Figure 1B), a cytosine analog bearing a hydrophobic phenylpyrrolo functionality. We show that tC^{P}TP is recognized by a laboratory-evolved TNA polymerase known as Kod-RI. Fidelity studies reveal that a cycle of TNA synthesis and reverse transcription occurs with 99.9% overall fidelity when tC^{P}TP and 7-deaza-tGTP are present together in the transcription mixture. These results suggest tC^{P}TP could be used to explore new regions of chemical space that are not accessible to TNA polymers with natural bases.

2 | MATERIALS AND METHODS

2.1 | General methods and materials

All nonaqueous reactions were performed using oven-dried glassware under inert gas atmosphere. All chemicals and solvents were of laboratory grade as obtained from commercial suppliers and were used without further purification. Thin-layer chromatography (TLC) was performed on TLC aluminum sheets covered with silica gel 60 F254 (Sigma-Aldrich, St. Louis, Missouri). Flash column chromatography (FC): SiliCycle 40-60 mesh silica gel (SiliCycle Inc., Quebec City, Canada). Yields are reported as isolated yields of pure compounds. UV quantification data are analyzed on NanoDrop 2000c using Beer's Law. ^{1}H, ^{13}C, and ^{31}P NMR spectra were obtained using Bruker DRX400 and Bruker DRX500 NMR spectrometers (Bruker, Billerica, Massachusetts). δ values in ppm relative to Me4Si or corresponding deuterium solvents as internal standard (^{1}H and ^{13}C). ^{31}P NMR values are reported in ppm relative to an external standard of 85% H3PO4. Reversed-phase HPLC was carried out on a C18 reverse-phase 250 × 9.4 mm HPLC column (Thermo Scientific) using a mobile phase of 0.1 M triethylammonium acetate buffer (pH 7.0)/acetonitrile. α-L-Threofuranosyl nucleoside 3′-monophosphate (8), 3′-phosphoro (2-methyl) imidazolide (9), and 3′-triphosphate (3) were analyzed by analytical HPLC with a reverse-phase column (C18 150 × 4.6 mm, 5 μm particle size, Thermo Scientific). ThermoPol buffer was purchased from New England Biolabs (Ipswich, Massachusetts). DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa). Recombinant Kod-RI polymerase was expressed and purified from *E coli* as previously described.[31] TNA triphosphates bearing natural bases were synthesized as previously described.[32]

2.2 | Chemical synthesis

2.2.1 | 1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-6-phenyl-furano[2,3-d]pyrimidin-2(3H)-one (5)

To a solution of 1-(2′-*O*-benzoyl-α-L-threofuranosyl)-5-iodo-uracil 4 (2.1 g, 4.73 mmol), Pd(PPh3)4 (0.56 g, 0.48 mmol) and CuI (0.18 g, 0.94 mmol) in anhydrous DMF (15 mL) were added anhydrous Et3N (1.3 mL, 9.4 mmol) and phenylacetylene (1.2 mL, 10.9 mmol). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 16 hours. The solvent was diluted with ethyl acetate (80 mL), washed with saturated Na_2EDTA aqueous solution (80 mL × 6) and brine (100 mL), dried over $MgSO_4$, and concentrated under reduced pressure to give a black residue. The residue was purified by flash chromatography (silica gel, CH_2Cl_2/MeOH, 30:1) to give crude product 2.0 g. To the crude product (2.0 g) in methanol (15 mL) was added Et_3N (6 mL) and CuI (191 mg, 1.0 mmol). The reaction mixture was heated at 50°C for 4 hours, and then methanol was removed under reduced pressure. The residue was washed with saturated Na_2EDTA solution and H_2O. Compound **5** was obtained as a yellow amorphous solid (1.52 g, 77%). TLC (hexane/EtOAc, 1:1): R_f = 0.21. ^{1}H NMR (500 MHz, DMSO-d_6): δ 8.67 (s, 1H), 8.05 (d, J = 7.5 Hz, 2H), 7.84 (d, J = 7.5 Hz, 2H), 7.73 (t, J = 7.5 Hz, 1H), 7.59 (t, J = 7.5 Hz, 2H), 7.52 (t, J = 7.5 Hz, 2H), 7.47 (t, J = 7.5 Hz, 1H), 7.32 (s, 1H), 6.08 (s, 1H), 5.80 (s, 1H), 5.42 (s, 1H), 4.42 (d, J = 9.5 Hz, 1H), 4.36-4.34 (m, 2H). ^{13}C NMR (126 MHz, DMSO-d_6): δ 171.4, 164.4, 153.9, 153.6, 139.3, 133.9, 129.6, 129.5, 129.1, 128.9, 128.3, 124.6, 106.5, 99.7, 91.9, 81.4, 77.0, 72.4. HRMS (ESI-TOF) calcd for $C_{23}H_{18}N_2O_6Na$ $[M + Na]^+$ 441.1063; observed 441.1058.

2.2.2 | 1-(α-L-Threofuranosyl)-6-phenyl-pyrrolo[2,3-d]pyrimidin-2(3H)-one (6)

Compound **5** (77 mg, 0.18 mmol) was dissolved in sat. NH_3-MeOH (2 mL) and NH_4OH (4 mL), and transferred to a sealed tube. The reaction mixture was stirred at 50°C for 16 hours. The solvent was evaporated, and the residue was purified by flash chromatography (silica gel, CH_2Cl_2/MeOH, 10:1) to afford free nucleoside **6** as a yellow solid (46 mg, 80%). TLC (CH_2Cl_2/MeOH, 9:1): R_f = 0.19. ^{1}H NMR (500 MHz, CD_3OD): δ 8.53-8.51 (m, 1H), 7.76-7.37 (m, 5H), 6.71-6.70 (m, 1H), 5.91-5.90 (m, 1H), 4.39-4.19 (m, 4H). ^{13}C NMR (126 MHz, CD_3OD): δ 160.6, 141.8, 138.7, 132.0, 130.1, 129.7, 126.3, 111.3, 98.4, 96.6, 82.0, 78.3, 76.7. HRMS (ESI-TOF) calcd for $C_{16}H_{15}N_3O_4Na$ $[M + Na]^+$ 336.0960; observed 336.0967.

2.2.3 | 1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-6-phenyl-furano[2,3-d]pyrimidin-2(3H)-one 3′-*O*-bis(2-cyanoethyl)-phosphotriester (7)

To a suspension of **5** (0.42 g, 1.0 mmol) and DMAP (24 mg, 0.20 mmol) in CH_2Cl_2 (10 mL) was added *N*,*N*-diisopropylethylamine (DIPEA) (261 μL, 1.5 mmol), followed by the addition of 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (318 μL, 1.5 mmol). After stirring for 20 minutes at room temperature, the solution was diluted with CH_2Cl_2 (40 mL) and extracted with saturated aqueous $NaHCO_3$ (40 mL). The organic layer was washed with brine, dried over Na_2SO_4, and evaporated to dryness for the next step. To a stirring solution of the residue in acetonitrile (10 mL) was added 3-hydroxypropionitrile (164 μL, 2.4 mmol), followed by a solution of 0.45 M tetrazole in acetonitrile (5.3 mL, 2.4 mmol). After being stirred for 1 hour at room temperature, H_2O_2 (30% in H_2O) (255 μL, 2.5 mmol) was added to the solution. After being stirred at room temperature for 20 minutes, the solution was diluted with CH_2Cl_2 (50 mL) and washed with brine. The organic layer was dried over Na_2SO_4 and evaporated. The residue was purified by flash chromatography (silica gel, CH_2Cl_2/acetone, 5:1) to afford **7** (0.26 g, 43%) as a white foam. TLC (CH_2Cl_2/acetone, 4:1): R_f = 0.34. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.32 (s, 1H), 8.06 (d, J = 7.5 Hz, 2H), 7.78 (d, J = 7.5 Hz, 2H), 7.63 (t, J = 7.5 Hz, 1H), 7.50-7.39 (m, 4H), 6.87 (s, 1H), 6.31 (s, 1H), 5.71 (s, 1H), 5.08-5.06 (m, 1H), 4.71 (d, J = 11.5 Hz, 1H), 4.49 (d, J = 11.5 Hz, 1H), 4.37-4.30 (m, 3H), 4.25-4.22 (m, 1H), 2.84-2.71 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.4, 165.2, 156.2, 154.9, 135.8, 134.2, 130.2, 130.0, 129.1, 128.8, 128.5, 125.2, 116.6, 116.5, 108.5, 98.1, 92.3, 80.4 (d, $J_{C,P}$ = 8.1 Hz), 79.2 (d, $J_{C,P}$ = 6.5 Hz), 75.2 (d, $J_{C,P}$ = 6.5 Hz), 63.2 (d, $J_{C,P}$ = 6.5 Hz), 63.0 (d, $J_{C,P}$ = 6.5 Hz), 19.8 (t, $J_{C,P}$ = 9.7 Hz). ^{31}P NMR (162 MHz, $CDCl_3$): δ − 2.36. HRMS (ESI-TOF) calcd for $C_{29}H_{25}N_4O_9PNa$ $[M + Na]^+$ 627.1257; observed 627.1255.

2.2.4 | 1-(α-L-Threofuranosyl)-6-phenyl-pyrrolo[2,3-d]pyrimidin-2(3H)-one 3′-*O*-monophosphate (8)

In a sealed tube, compound **7** (100 mg, 0.165 mmol) was combined with sat. NH_3-MeOH (2 mL) and NH_4OH (4 mL), and stirred at 45°C for 16 hours. The solution was cooled down to room temperature, diluted with water, and washed with CH_2Cl_2 (3 × 40 mL). The aqueous layer was lyophilized to afford the product as an ammonium salt. The precipitate was collected by centrifugation at 4400 rpm at room temperature for 15 minutes, and the resulting pellet was washed twice with 20 mL of acetone and dried under high vacuum. The nucleoside 3′-monophosphate **8** was obtained as a yellow solid in quantitative yield (70 mg, 100%). ^{1}H NMR (500 MHz, D_2O): δ 8.32 (s, 1H), 7.33-7.31 (m, 5H), 6.43 (s, 1H), 5.87 (s, 1H), 4.67-4.63 (m, 2H), 4.53-4.49 (m, 2H). ^{13}C NMR (126 MHz, D_2O): δ 157.9, 155.7, 140.5, 138.2, 129.9, 129.4, 129.3, 129.0, 125.3, 111.1, 98.2, 94.7, 80.1, 78.5, 76.8. ^{31}P NMR (162 MHz, D_2O): δ 3.25. HRMS (ESI-TOF) calcd for $C_{16}H_{15}N_3O_7P$ $[M - H]^-$ 392.0648; observed 392.0654.

2.2.5 | 1-(α-L-Threofuranosyl)-6-phenyl-pyrrolo[2,3-d]pyrimidin-2(3H)-one 3′-*O*-phosphor-(2-methyl) imidazolide (9)

To a stirring solution containing nucleoside 3′-monophosphate **8** (55 mg, 0.13 mmol) and 2-methylimidazole (150 mg, 1.83 mmol) in anhydrous DMSO (0.7 mL) and DMF (0.7 mL) was added triethylamine

(0.15 mL), triphenylphosphine (150 mg, 0.57 mmol), and 2,2′-dipyridyldisulfide (180 mg, 0.83 mmol). The reaction was stirred under N_2 protection for 3 hours at room temperature with monitoring by analytical HPLC. After consumption of the starting material, the product was precipitated by dropwise addition of the reaction mixture to a stirring solution containing 60 mL of acetone, 50 mL of diethyl ether, 3 mL of triethylamine, and 0.6 mL of saturated $NaClO_4$ in acetone. The precipitant was collected by centrifugation at 4400 rpm for 15 minutes at room temperature. The pellet was washed twice with 20 mL of washing solution (acetone/diethyl ether 1:1) and dried under high vacuum to afford **9** as a sodium salt (white solid, 56 mg, 91%). It was used for the next step directly without further purification. ^{31}P NMR (162 MHz, DMSO-d_6): δ – 10.06. HRMS (ESI-TOF) calcd for $C_{20}H_{20}N_5O_6PNa$ [M + Na]$^+$ 480.1049; observed 480.1036.

2.2.6 | 1-(α-L-Threofuranosyl)-6-phenyl-pyrrolo [2,3-d]pyrimidin-2(*3H*)-one 3′-*O*-triphosphate (3)

To an anhydrous DMF (2 mL) solution of compound **9** (74 mg, 154 μmol) was added tributylamine (76 μL, 0.32 mmol) and tributylammonium pyrophosphate (176 mg, 0.32 mmol), and the reaction was stirred at room temperature for 6 hours under N_2 protection with monitoring by analytical HPLC. After the reaction was finished, the reaction mixture was added dropwise to a stirring solution containing 20 mL of acetone and 2 mL of saturated $NaClO_4$ in acetone. The precipitate was collected by centrifugation at 4400 rpm for 15 minutes at room temperature and dried under vacuum for 1 hour. The crude precipitate was dissolved in 2 mL of 0.1 M triethylammonium acetate buffer and purified by semi-preparative HPLC. Fractions containing triphosphates were collected, concentrated, pH adjusted by triethylamine to 8.0, and lyophilized to afford the product as a triethylammonium salt. The solid product was resuspended in 1 mL of methanol and was added dropwise to a solution containing 20 mL of acetone and 1 mL of saturated $NaClO_4$ in acetone. The solution was centrifuged at 4400 rpm for 15 minutes at room temperature. The supernatant was discarded, and the pellet was washed with 30 mL acetone and dried under vacuum. The desired triphosphate **3** was obtained as sodium salt in a white solid form. The yield was measured by UV absorbance at 260 nM using NanoDrop 2000c, assuming an extinction coefficient of 21.0 mM^{-1} cm^{-1} and found to be 38% (58.5 μmol). ^{1}H NMR (400 MHz, D_2O): δ 8.38 (s, 1H), 7.58 (d, J = 5.2 Hz, 2H), 7.40-7.35 (m, 3H), 6.66 (s, 1H), 5.88 (s, 1H), 4.91-4.89 (m, 1H), 4.65 (d, J = 7.2 Hz, 1H), 4.54-4.50 (m, 2H). ^{31}P NMR (162 MHz, D_2O): δ – 5.62 (d, J = 13.0 Hz), –12.50 (d, J = 14.6 Hz), –21.58 (t, J = 13.0 Hz). HRMS (ESI-TOF) calcd for $C_{16}H_{15}N_3O_{13}P_3Na_4$ [M -3H + 4Na]$^+$ 641.9409; observed 641.9408.

2.3 | Primer extension assay on 3.17 DNA template

Primer-extension reactions were performed in a 10 μL reaction using 10 pmol of IR800-PBS2 primer and 10 pmol of 3.17 DNA template. Each reaction contained the DNA primer and template, 1X ThermoPol buffer [20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8], 1 μM Kod-RI, 100 μM of each tNTP, and 1 mM $MnCl_2$. Reactions were incubated for 1 hour at 55°C, quenched with stop buffer (8 M urea, 45 mM EDTA), and analyzed by 20% denaturing urea PAGE.

2.4 | Fidelity of TNA replication

The fidelity assay was performed as previously described.[25]

3 | RESULTS

3.1 | Synthesis of α-L-threofuranosyl-6-phenylpyrrolocytidine nucleoside

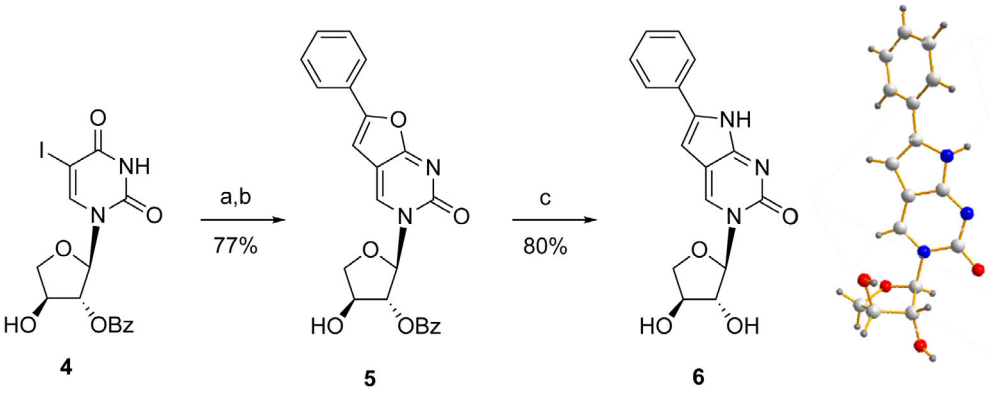
Our synthetic strategy (Scheme 1) started from α-L-threofuranosyl 5-iodouracil **4**.[26] A palladium and Cu(I) catalyzed Sonogashia cross-coupling reaction of **4** with phenylacetylene, followed by Cu(I) catalyzed cyclization in methanol at 50°C for 4 hours produced the desired furanouridine **5**. The free α-L-threofuranose phenylpyrrolocytosine nucleoside **6** was obtained by treating **5** with ammonium hydroxide to remove the 2′-*O*-benzoyl protecting group and convert the furan to the corresponding pyrrole. The molecular structure of **6** was confirmed by small molecule X-ray diffraction, which showed the desired nucleobase and correct arrangement of hydroxyl groups on the furanose sugar.

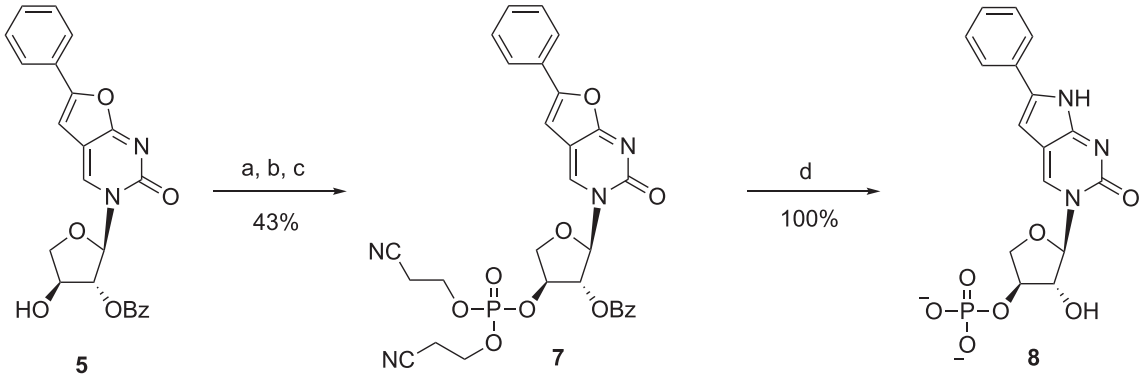
3.2 | Synthesis of TNA 3′-triphosphate

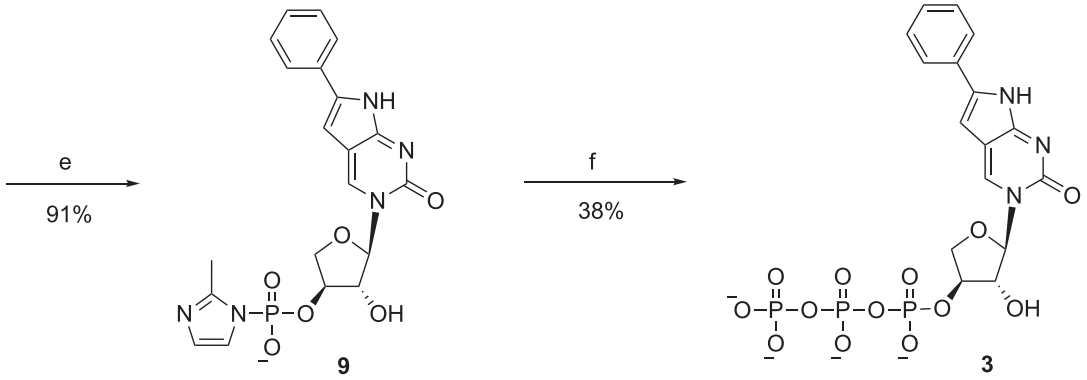
TNA triphosphates are generally more challenging to synthesize than DNA and RNA triphosphates because phosphorylation occurs at the more sterically encumbered 3′-hydroxyl group rather than the 5′-hydroxyl group found in DNA and RNA nucleosides. To overcome this problem, we followed a synthetic strategy that has proven successful for other TNA triphosphates.[32] In this approach, the free 3′-hydroxyl group is converted to an activated monophosphate that is treated with inorganic pyrophosphate to produce the desired nucleoside triphosphate, which is then purified by HPLC chromatography.

Initial attempts at constructing the pyrrolocytidine nucleoside 3′-monophosphate (tC^{p}MP) **8** utilized a dibenzyl phosphoramidite reagent as the phosphorus source. However, subsequent deprotection by hydrogenation led to unwanted opening of the bicyclic ring system. To overcome this problem, we revised our synthetic strategy (Scheme 2) to obtain the desired TNA nucleoside monophosphate **8**. Accordingly, nucleoside **5** was treated with 2-cyanoethoxy-*N*,*N*,-diisopropylchlorophosphine in the presence of Hünig's base and DMAP to obtain the 3′-*O*-phosphoramidite, which was subsequently converted to the corresponding trialkyl phosphite by exchanging the *N*,*N*-diisopropyl group with a base-labile cyanoethoxy group. The trialkyl phosphite was then oxidized with H_2O_2 to afford the

SCHEME 1 Reagents and conditions. A, Phenylacetylene, $Pd(PPh_3)_4$, CuI, triethylamine, DMF, r.t., 16 hours; B, CuI, triethylamine, MeOH, 50°C, 4 hours; C, NH_3-MeOH-H_2O, 55°C, 16 hours







SCHEME 2 Reagents and conditions. A, $NC(CH_2)_2OP(Cl)N(i\text{-}Pr)_2$, $(i\text{-}Pr)_2EtN$, DMAP, CH_2Cl_2, r.t., 20 minutes; B, 3-hydroxypropionitrile, tetrazole, CH_3CN, r.t., 1 hour; C, H_2O_2, CH_3CN, r.t., 20 minutes; D, NH_3-MeOH-H_2O, 45°C, 16 hours; E, 2-methylimidazole, PPh_3, 2,2′-dipyridyl disulfide, triethylamine, DMF-DMSO, r.t., 3 hours; F, tributylammonium pyrophosphate, tributylamine, anhydrous DMF, r.t., 6 hours

phosphate triester **7**, which was purified by silica gel chromatography. The desired TNA nucleoside 3′-monophosphate **8** was obtained as an ammonium salt following deprotection with 30% aqueous ammonium hydroxide. Next, TNA monophosphate **8** was converted to an imidazole-activated monophosphate **9** using a Mitsunobu reaction. Subsequent displacement of the imidazole group with pyrophosphate resulted in the desired tC^{P}TP **3**, which was obtained as sodium salt following HPLC purification, lyophilization, and precipitation with sodium perchlorate. The purity of the nucleoside 3′-monophosphates (**8**), 3′-phosphoro(2-methyl)-imidazolide (**9**) and 3′-triphosphate (**3**) were confirmed by analytical HPLC.

3.3 | TNA synthesis with modified TNA cytosine triphosphates

We evaluated the pyrrolocytidine TNA triphosphate **3** as a substrate for TNA synthesis using the TNA polymerase Kod-RI.[33] In this assay, a 5′- IR800-labeled DNA primer was annealed to a DNA template and incubated with Kod-RI for 1 hour at 55°C in the presence of chemically synthesized tNTP substrates.[34] As a comparison, three different primer extension reactions were performed in which the tNTP mixture contained tCTP **1**, tC^{f}TP **2**, or tC^{P}TP **3** (Figure 2). Analysis of the reaction products by denaturing polyacrylamide gel electrophoresis

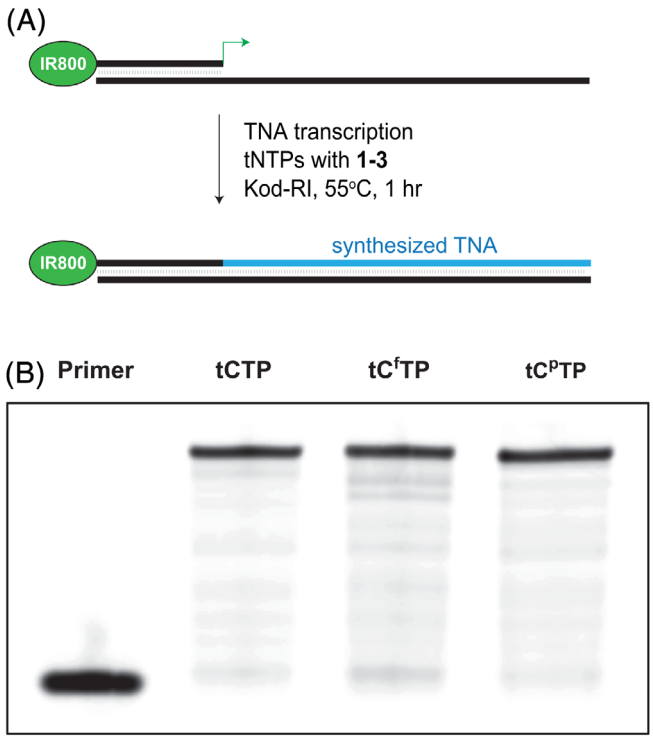


FIGURE 2 Polymerase-mediated synthesis of TNA with cytosine triphosphates (**1-3**). A, Schematic representation of primer extension assay. B, Denaturing PAGE analysis of TNA synthesis

(PAGE) revealed the formation of full-length products in all cases, as evidenced by complete conversion of the starting primer into the full-length TNA product. Relative to the unmodified tCTP **1** and modified tC^{f}TP **2**, tC^{P}TP **3** showed comparable yields of full-length product, indicating **3** is a suitable substrate for Kod-RI TNA polymerase.

3.4 | Fidelity of TNA replication with tC^{P}TP 3

To ensure that Kod-RI faithfully incorporates tC^{P}TP residues during TNA synthesis, we measured the fidelity of TNA replication (DNA → TNA → DNA) by sequencing the cDNA product obtained from a cycle of TNA synthesis and reverse transcription. This assay is operationally different than the fidelity of a single-nucleotide incorporation event, as it measures the accuracy by which a DNA template is copied into TNA, purified, and then copied back into DNA. Several controls were implemented to ensure that sequencing results represented the true fidelity of TNA synthesis. The most important of which included the use of a DNA primer that carried a double nucleotide mismatched that resulted in a TT → AA transversion in the replicated product but would remain a mismatch in a contaminating DNA template.

We sequenced >1000 nucleotide positions from the cDNA isolated from a cycle of TNA transcription and reverse transcription. Analysis of the data yielded an overall replication fidelity of 98.4% (error rate of 1.7×10^{-2}), which is comparable to unmodified tCTP and tC^{f}TP (Table 1).[35] Since the majority of the mutations were G → C transversion caused by G-G mispairing during TNA synthesis, we sought to improve the fidelity of TNA synthesis by substituting tGTP in the transcription mixture for the base analog 7-deaza-tGTP.[25] 7-deaza-tGTP is known to suppress the formation of the G-G mispairing during TNA synthesis by preventing the formation of a G-G Hoogsteen base-pair in the enzyme active site.[25] When 7-deaza-tGTP was present along with tC^{P}TP, the fidelity of TNA replication increased to 99.9% with an error rate of 1×10^{-3}. This level of fidelity is among the highest for XNA synthesis and reverse transcription, which often exhibit error rates in the range of 4.3×10^{-3} to 5.3×10^{-2}.[24]

TABLE 1 Fidelity of TNA Synthesis using tNTPs with modified bases[a]

	tC^{f}TP[24]	tC^{P}TP	(tC^{P}TP + 7dG)
Transitions	10.9	2.7	–
Transversions	5.4	14.5[b]	–
Insertions	–	–	–
Deletions	–	–	0.95
Error rate	1.6×10^{-2}	1.7×10^{-2}	0.1×10^{-2}
Fidelity (%)	98.4	98.3	99.9

[a]Fidelity measurements of TNA synthesis were performed using tC^{P}TP supplemented with three tNTPs with unmodified bases, or tC^{P}TP, 7-deaza-tGTP and two other unmodified tNTPs. Values are normalized to 1000 sequenced bases.
[b]Most of the transversions were G to C mutations in the DNA template.

4 | DISCUSSION

The field of synthetic genetics aims to establish the tools and techniques required to explore the structural and functional properties of alternative genetic polymers with backbone structures that are distinct from those found in nature.[1–3] One important goal of this area of research is to establish a new generation of aptamers and catalysts with physicochemical properties not found in natural DNA and RNA. For example, many XNAs have backbone structures that are resistant, or in some cases, recalcitrant to nuclease digestion or exhibit improved stability against high acid or basic conditions.[13,15] This feature makes them strong candidates for diagnostic and therapeutic applications that warrant high biological or chemical stability.[16]

More recently, efforts have begun to expand the chemical diversity of XNA polymers beyond the set of bases found in nature. Previously, we have explored the unnatural bases of diaminopurine and 7-deazaguanine as possible strategies for improving the efficiency and fidelity of TNA synthesis.[25,36] These initial approaches allowed us to improve the fidelity of TNA synthesis. Moving forward, we are working to augment the nuclease-resistant properties of TNA with expanded chemical functionality toward the development of biostable TNA molecules harboring protein-like properties by using extra chemical moieties attached to nucleobases. Indeed, similar approaches have proven successful for generating DNA aptamers with slow off-rates to protein targets as potential aptamer-based platform to identify biomarkers from clinically-relevant samples.[18–22] As target neutralizing reagents or ligands for identifying biomarkers, chemically expanded

TNA, as well as other XNA aptamers, would be more compatible to nuclease-intense physiological conditions, relative to their natural counterparts.

5 | CONCLUSIONS

In summary, we synthesized and evaluated the enzymatic properties of a TNA triphosphate bearing 6-phenyl-pyrrolocytidine, abbreviated tC^{p}TP. Pyrrolocytidine is a modified nucleobase with enhanced hydrophobicity and strong fluorescence that is highly sensitive toward environmental perturbations, which expands the chemical space of artificial genetic polymers by introducing novel photophysical properties into a nucleic acid system with a synthetic, nonnatural backbone structure. Enzymatic incorporation studies indicate that the tC^{p} triphosphate is an efficient substrate for an engineered TNA polymerase. Last, we found that TNA templates synthesized with tC^{p}TP and 7-deaza-tGTP exhibit 99.9% overall fidelity, which is among the highest for XNA replication systems. Together, these results provide a platform for synthesizing TNA libraries with enhanced hydrophobic and fluorescent properties.

ACKNOWLEDGMENT

We wish to thank members of the Chaput lab for helpful discussions and suggestions. This work was supported by the National Science Foundation (MCB: 1946312).

CONFLICT OF INTEREST

The authors declare no competing interest.

REFERENCES

[1] J. C. Chaput, H. Yu, S. Zhang, *Chem. Biol.* **2012**, *19*, 1360.
[2] G. Houlihan, S. Arangundy-Franklin, P. Holliger, *Acc. Chem. Res.* **2017**, *50*, 1079.
[3] V. B. Pinheiro, D. Loakes, P. Holliger, *Bioessays* **2012**, *35*, 113.
[4] V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S. Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn, P. Holliger, *Science* **2012**, *336*, 341.
[5] A. I. Taylor, V. B. Pinheiro, M. J. Smola, A. S. Morgunov, S. Peak-Chew, C. Cozens, K. M. Weeks, P. Herdewijn, P. Holliger, *Nature* **2015**, *518*, 427.
[6] H. Yu, S. Zhang, J. C. Chaput, *Nat. Chem.* **2012**, *4*, 183.
[7] I. Alves Ferreira-Bravo, C. Cozens, P. Holliger, J. J. DeStefano, *Nucleic Acids Res.* **2015**, *43*, 9587.
[8] Y. Wang, A. K. Ngor, A. Nikoomanzar, J. C. Chaput, *Nat. Commun.* **2018**, *9*, 5067.
[9] E. Eremeeva, A. Fikatas, L. Margamuljana, M. Abramov, D. Schols, E. Groaz, P. Herdewijn, *Nucleic Acids Res.* **2019**, *47*, 4927.
[10] A. E. Rangel, Z. Chen, T. M. Ayele, J. M. Heemstra, *Nucleic Acids Res.* **2018**, *46*, 8057.
[11] M. R. Dunn, C. M. McCloskey, P. Buckley, K. Rhea, J. C. Chaput, *J. Am. Chem. Soc.* **2020**, *142*, 7721.
[12] G. F. Joyce, *Science* **2012**, *336*, 307.
[13] M. C. Culbertson, K. W. Temburnikar, S. P. Sau, J.-Y. Liao, S. Bala, J. C. Chaput, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 2418.
[14] I. Anosova, E. A. Kowal, M. R. Dunn, J. C. Chaput, W. D. Van Horn, M. Egli, *Nucleic Acids Res.* **2016**, *44*, 1007.
[15] J. K. Watts, A. Katolik, J. Viladoms, M. J. Damha, *Org. Biomol. Chem.* **2009**, *7*, 1904.
[16] M. R. Dunn, R. M. Jimenez, J. C. Chaput, *Nat. Rev. Chem.* **2017**, *1*, 0076.
[17] D. S. Wilson, J. W. Szostak, *Annu. Rev. Biochem.* **1999**, *68*, 611.
[18] M. Hollenstein, C. J. Hipolito, C. H. Lam, D. M. Perrin, *Nucleic Acids Res.* **2009**, *37*, 1638.
[19] J. D. Vaught, C. Bock, J. Carter, T. Fitzwater, M. Otis, D. Schneider, J. Rolando, S. Waugh, S. K. Wilcox, B. E. Eaton, *J. Am. Chem. Soc.* **2010**, *132*, 4141.
[20] B. N. Gawande, J. C. Rohloff, J. D. Carter, I. von Carlowitz, C. Zhang, D. J. Schneider, N. Janjic, *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 2898.
[21] K. Sefah, Z. Yang, K. M. Bradley, S. Hoshika, E. Jimenez, L. Zhang, G. Zhu, S. Shanker, F. Yu, D. Turek, W. Tan, S. A. Benner, *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 1449.
[22] Y. Wang, E. Liu, C. H. Lam, D. M. Perrin, *Chem. Sci.* **2018**, *9*, 1813.
[23] J. Attwater, S. Tagami, M. Kimoto, K. Butler, E. T. Kool, J. Wengel, P. Herdewijn, I. Hirao, P. Holliger, *Chem. Sci.* **2013**, *4*, 2804.
[24] H. Mei, C. Shi, R. M. Jimenez, Y. Wang, M. Kardouh, J. C. Chaput, *Nucleic Acids Res.* **2017**, *45*, 5629.
[25] H. Mei, J. Y. Liao, R. M. Jimenez, Y. Wang, S. Bala, C. McCloskey, C. Switzer, J. C. Chaput, *J. Am. Chem. Soc.* **2018**, *140*, 5706.
[26] H. Mei, J. C. Chaput, *Chem. Commun. (Camb)* **2018**, *54*, 1237.
[27] F. Levi-Acobas, A. Katolik, P. Rothlisberger, T. Cokelaer, I. Sarac, M. J. Damha, C. J. Leumann, M. Hollenstein, *Org. Biomol. Chem.* **2019**, *17*, 8083.
[28] Y. Brundo, D. R. Liu, *Chem. Biol.* **2009**, *16*, 265.
[29] L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T. R. Keeney, N. Kim, T. H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S. K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R. M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D. I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J. D. Vaught, M. Vrkljan, J. J. Walker, M. Watrobka, S. Waugh, A. Weiss, S. K. Wilcox, A. Wolfson, S. K. Wolk, C. Zhang, D. Zichi, *Plos One* **2010**, *5*, e15004. https://doi.org/10.1371/journal.pone.0015004.
[30] P. A. Lichtor, Z. Chen, N. H. Elowe, J. C. Chen, D. R. Liu, *Nat. Chem. Biol.* **2019**, *15*, 419.
[31] Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C., *Curr. Protoc. Nucleic Acid Chem.* **2017**, *69*(1), 4-75. https://doi.org/10.1002/cpnc.33.
[32] S. Bala, J. Y. Liao, H. Mei, J. C. Chaput, *J. Org. Chem.* **2017**, *82*, 5910.
[33] M. R. Dunn, C. Otto, K. E. Fenton, J. C. Chaput, *ACS Chem. Biol.* **2016**, *11*, 1210.
[34] S. P. Sau, J. C. Chaput, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 3271.
[35] H. Yu, S. Zhang, M. R. Dunn, J. C. Chaput, *J. Am. Chem. Soc.* **2013**, *135*, 3583.
[36] J. C. Chaput, J. W. Szostak, *J. Am. Chem. Soc.* **2003**, *125*, 9274.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Mei H, Wang Y, Yik EJ, Chaput JC. Synthesis and polymerase recognition of a pyrrolocytidine TNA triphosphate. *Biopolymers*. 2020;e23388. https://doi.org/10.1002/bip.23388