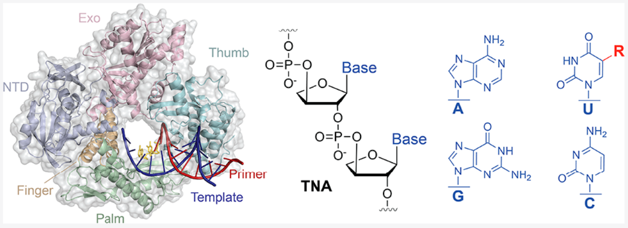
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JACS Article

Synthesis and Polymerase Recognition of Threose Nucleic Acid Triphosphates Equipped with Diverse Chemical Functionalities

Qingfeng Li, Victoria A. Maola, Nicholas Chim, Javeena Hussain, Adriana Lozoya-Colinas, and John C. Chaput*

Cite This: *J. Am. Chem. Soc.* 2021, 143, 17761−17768 Read Online



ABSTRACT: Expanding the chemical space of evolvable non-natural genetic polymers (XNAs) to include functional groups that enhance protein target binding affinity offers a promising route to therapeutic aptamers with high biological stability. Here we describe the chemical synthesis and polymerase recognition of 10 chemically diverse functional groups introduced at the C-5 position of α-L-threofuranosyl uridine nucleoside triphosphate (tUTP). We show that the set of tUTP substrates is universally recognized by the laboratory-evolved polymerase Kod-RSGA. Insights into the mechanism of TNA synthesis were obtained from a high-resolution X-ray crystal structure of the postcatalytic complex bound to the primer−template duplex. A structural analysis reveals a large cavity in the enzyme active site that can accommodate the side chain of C-5-modified tUTP substrates. Our findings expand the chemical space of evolvable nucleic acid systems by providing a synthetic route to artificial genetic polymers that are uniformly modified with diversity-enhancing functional groups.

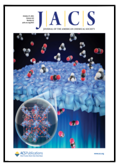
INTRODUCTION

The dominance of proteins over nucleic acids as scaffolds for biological receptors and catalysts is generally explained by the presence of 20 chemically diverse amino acids in comparison to the four chemically similar nucleobases of adenine (A), cytosine (C), guanine (G), and thymine (T) or uracil (U).[1] Despite this weakness, aptamers can more easily evolve in comparison to antibodies, making them a promising class of molecules for diagnostic and therapeutic applications that require target-specific binding via a programmable nucleic acid scaffold.[2−4] In recent years, the biological stability of nucleic acid aptamers has improved by the development of engineered forms of naturally occurring DNA polymerases that allow for the evolution of xeno-nucleic acid (XNA) aptamers comprised entirely of non-natural backbone architectures.[5] To date, aptamers have been discovered for the XNA systems of 2′-fluoroarabino nucleic acid (FANA),[6,7] hexitol nucleic acid (HNA),[8,9] and threose nucleic acid (TNA, Figure 1)[10−13] as well as a mixed-backbone system consisting of 2′-*O*-methoxy RNA and locked nucleic acid (LNA).[14] Although substitution of the ribose sugar for another sugar or sugar-like moiety can substantially improve the pharmacokinetic properties of the oligonucleotide,[13,15] these polymers often utilize the same nucleobases as those found in the natural genetic polymers of DNA and RNA.[16] This level of chemical similarity, consisting primarily of highly polar amino and carbonyl groups, is thought to limit the targetability of aptamers to a subset of the human proteome.[17]

Augmenting DNA scaffolds with diversity-enhancing functional groups is an effective route for generating aptamers that function with improved target binding affinity and specificity.[18] The simplest approach involves using click chemistry to introduce new a chemical functionality onto alkyne-modified nucleotides that have been incorporated into a DNA oligonucleotide.[19] Krauss and Mayer pioneered this strategy to evolve base-modified aptamers against targets that are challenging for DNA libraries having only natural bases.[20,21] Soh and colleagues extended this strategy to a particle display format that allows for greater control over the selection parameters.[22] Though it is amenable to a broad range of functional groups, including bulky substituents that maybe difficult to prepare enzymatically, the click chemistry approach generates a heterogeneous mixture of base-modified oligonucleotides due to subquantitative coupling of the alkynyl group to its corresponding azido-modified side chain.

Preparing uniformly modified nucleic acid libraries demands synthetic strategies that introduce new chemical functionalities into the oligonucleotide during the synthesis step of library construction. One approach uses a template-directed ligase-mediated assembly path for constructing highly functionalized

Received: August 16, 2021
Published: October 12, 2021



https://doi.org/10.1021/jacs.1c08649

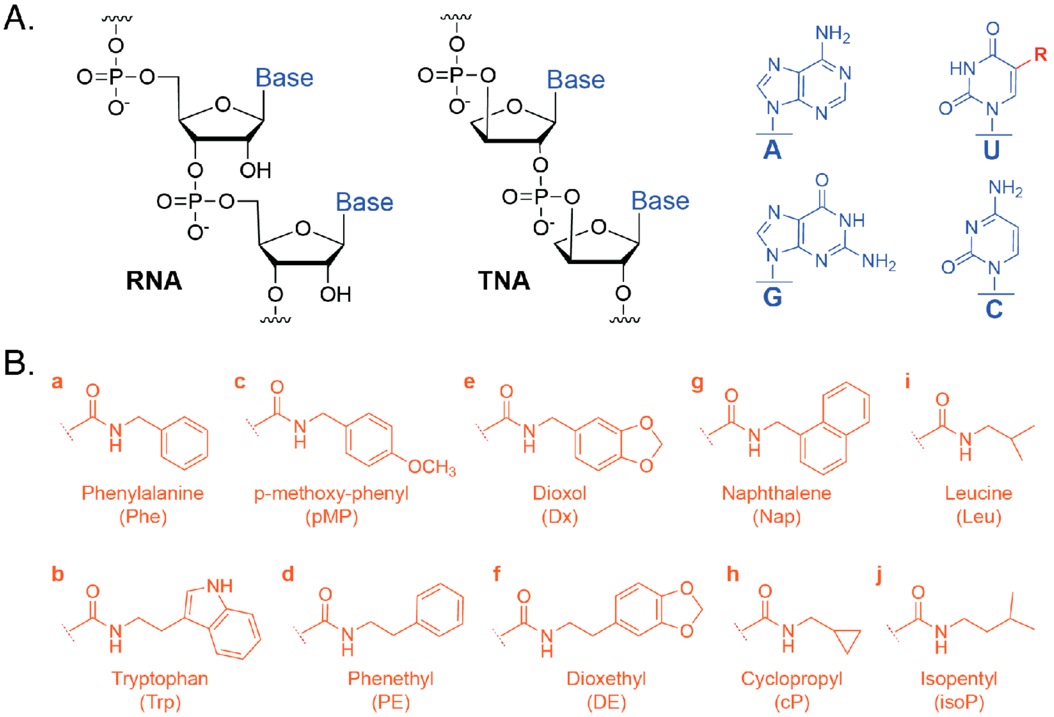


Figure 1. Chemically modified TNA triphosphates. (A) Constitutional structure for the linearized backbone of ribose nucleic acid (RNA) and 3′,2′-α-L-threofuranosyl nucleic acid (TNA) with the chemical structure of the nucleobases shown on the right. The location of the side chain modifications is denoted by an R group (red) at the C-5 position of the uracil nucleobase. (B) The chemical structure of the side chains prepared as C-5-modified tUTP derivatives with the amide linkage attached to the C-5 carbon atom of uracil.

nucleic acid polymers (HFNAP) that mimic proteins by containing diverse sets of chemical functionalities within the same oligonucleotide.[23−26] HFNAPs have been selected to bind several therapeutic targets, including proprotein convertase subtilisin/kexin type 9 (PCSK9), interleukin-6, and thrombin.[23,24] The polymerase-mediated synthesis of chemically modified nucleoside triphosphates offers a different approach to assembling uniformly modified oligonucleotides by incorporating base-modified residues directly into the growing strand. This strategy has been broadly explored as a general route to DNA aptamers with expanded chemical and physical properties.[27−32]

We have previously developed an *in vitro* selection system that allows for the rapid isolation of TNA aptamers by covalently linking freshly synthesized TNA strands to their encoding DNA template.[10] The DNA-display strategy avoids the need for a separate TNA reverse transcription step, as functional aptamers that remain bound to the target are recovered by amplifying their encoding DNA sequence using the polymerase chain reaction. In a proof-of-principle demonstration, TNA aptamers isolated from just three rounds of a stringent, affinity-based selection were found to bind HIV reverse transcriptase with solution binding affinity constants (K_D) of 0.4−4.0 nM and remain functional in the presence of biological nucleases.[11] This finding provides strong motivation for pursuing the development of functionally enhanced TNA aptamers as biologically stable affinity reagents for diagnostic and therapeutic applications.

Here, we describe the chemical synthesis and polymerase recognition of 10 chemically diverse functional groups introduced at the C-5 position of α-L-threofuranosyl uridine nucleoside triphosphate (tUTP, Figure 1). The stereochemistry of the modified nucleosides was unambiguously determined by solving the small-molecule X-ray crystal structures of two representative members. We further show that the set of tUTP substrates is universally recognized by the engineered polymerase Kod-RSGA, which is currently the best example of a DNA-dependent TNA polymerase.[33] We obtained insight into the mechanism of TNA synthesis by solving a high-resolution X-ray crystal structure of the postcatalytic complex of Kod-RSGA bound to the primer−template duplex. The crystal structure reveals a large cavity in the major groove region of the primer−template binding site that can accommodate the side chain of C-5-modified tUTP substrates. Together, our findings support the growing field of synthetic genetics by providing a synthetic route to TNA polymers that are uniformly equipped with new chemical functionality.

■ RESULTS

We began by chemically synthesizing a set of 10 C-5 modified tUTP substrates (Figure 1) following a palladium cross-catalyzed carboxyamidation reaction that couples 5-iodo-1-(2′-*O*-benzoyl-α-L-threofuranosyl)uracil to aryl or alkyl amines via *in situ* formation of an amide linkage.[34] The side chains analyzed in this study included the aryl ring systems of phenylalanine (Phe), tryptophan (Trp), *p*-methoxyphenyl (pMP), phenethyl (PE), dioxol (Dx), dioxethyl (DE), and naphthalene (Nap), while the alkyl groups included the leucine (Leu), cyclopropyl (cP), and isopentyl (isoP) side chains (Figure 1). Although other conjugation chemistries are available for linking functional groups to uracil nucleosides,[21] we viewed this strategy as a highly versatile route to uniformly modified TNA oligonucleotides. In addition to broad chemical reactivity, the resulting amide linkage benefits from reduced hydrophobicity relative to carbon linkers and limits the

https://doi.org/10.1021/jacs.1c08649
J. Am. Chem. Soc. 2021, 143, 17761−17768

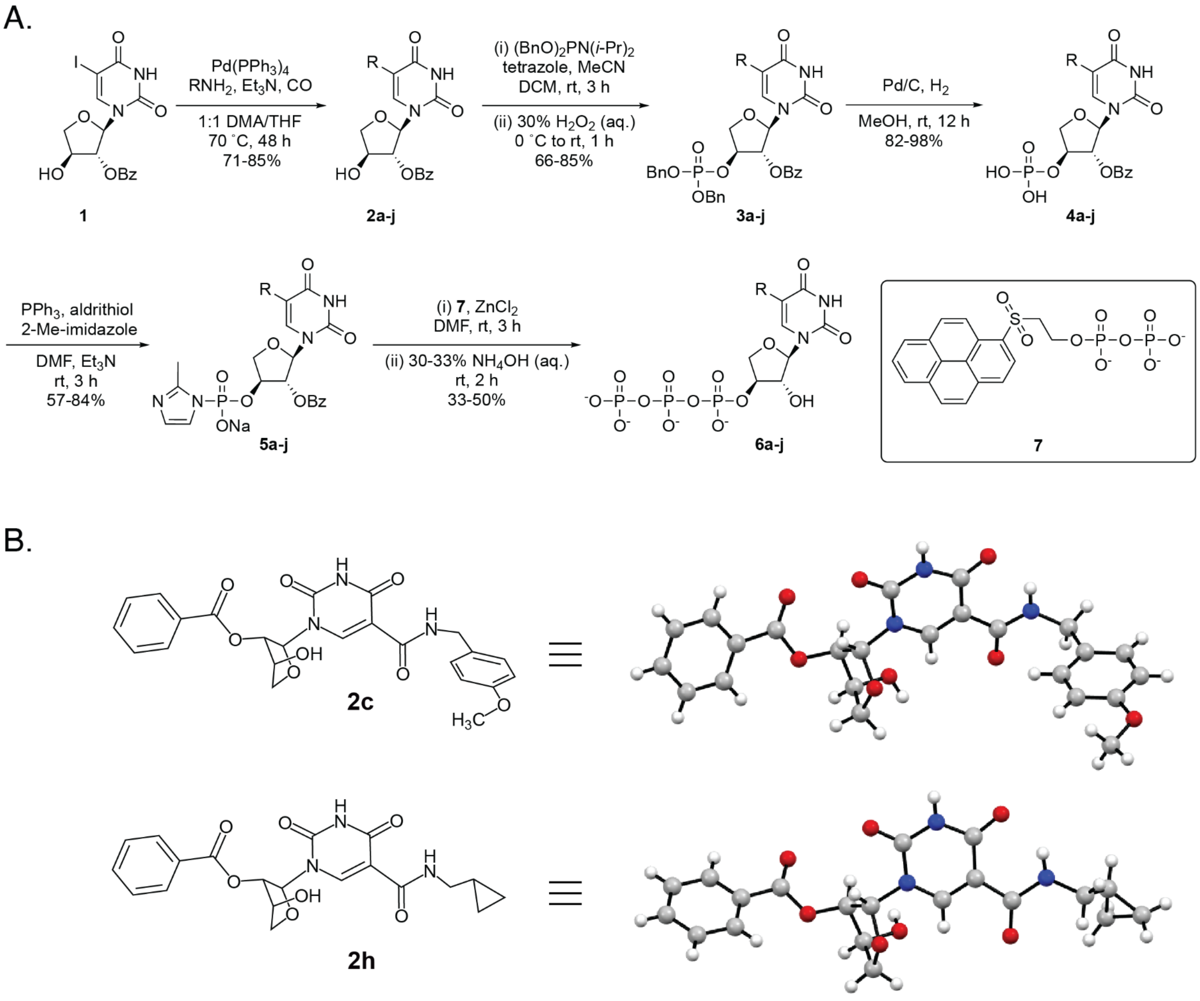


Figure 2. Chemical synthesis and structural characterization of TNA monomers. (A) Synthetic strategy used to prepare a set of 10 chemically distinct C-5 modified tUTP derivatives. Analogues a−j correspond to the 10 unique side chains provided in Figure 1B. (B) Small-molecule X-ray crystal structures of the C-5-modified uridine nucleosides prepared for *p*-methoxyphenyl (pMP, **2c**) and cyclopropyl (cP, **2h**).

number of rotatable bonds between the functional group and nucleobase.

The starting 5-iodo-1-(2′-*O*-benzoyl-*α*-L-threofuranosyl)-uracil nucleoside **1** was prepared in eight steps from L-ascorbic acid using a Vorbrüggen reaction to conjugate 5-iodouracil to an orthogonally protected threose sugar (Scheme S1).[35] Subsequent conjugation of the side chain to the nucleobase was accomplished in separate reactions by heating the desired aryl or alkyl amine with the 5-iodo-TNA nucleoside **1** in the presence of 10 mol % of $Pd(PPh_3)_4$ and CO (Figure 2A). The stereochemistry of the modified nucleosides was unambiguously determined by solving small-molecule X-ray crystal structures for the *p*-methoxyphenyl (**2c**) and cyclopropyl (**2h**) analogues. As illustrated in Figure 2B, both side chains are attached to the uracil base via an amide linkage at the C-5 position with proper stereochemistry observed at the C1′, C2′, and C3′ sugar positions (Figures S1 and S2 and Tables S1−S3). After purification by silica gel chromatography, the C-5-modified uracil nucleosides were converted to their corresponding nucleoside-3′-triphosphates **6** by transforming the 3′-hydroxyl group into the activated nucleoside 3′-monophosphate **5** that was coupled to pyrene pyrophosphate **7** to produce a fully protected TNA nucleoside-3′-triphosphate.[36] Following purification, the triphosphate precursor was deprotected for 1 h at 24 °C in concentrated ammonium hydroxide and precipitated as a sodium salt to afford the *α*-L-threofuranosyl uracil nucleoside triphosphates **6** carrying the desired side chain at the C-5 position (see the Supporting Information for synthetic protocols and compound characterization). This study also required the preparation of all four TNA triphosphates with the natural bases of adenine (tATP), cytosine (tCTP), thymine (tTTP), and guanine (tGTP), which were each synthesized in 14 steps from L-ascorbic acid as described previously.[36,37]

We confirmed that the set of base-modified TNA triphosphates were viable substrates for TNA synthesis by evaluating their incorporation into a TNA oligonucleotide library using a primer extension assay. Accordingly, a 5′-IR-labeled DNA primer annealed to a DNA library containing a central random region was incubated with a mixture of chemically synthesized tNTPs and Kod-RSGA for 2 h at 55 °C. In these reactions, the tNTP mixture contained either standard bases only or a tNTP solution in which the tTTP substrate was individually replaced with one of the 10 base-modified tUTP substrates. An analysis of the primer extension reactions by denaturing polyacrylamide gel electrophoresis indicates that the primer was extended to a full-length product in all cases, as evidenced by the presence of a discrete slower-moving band

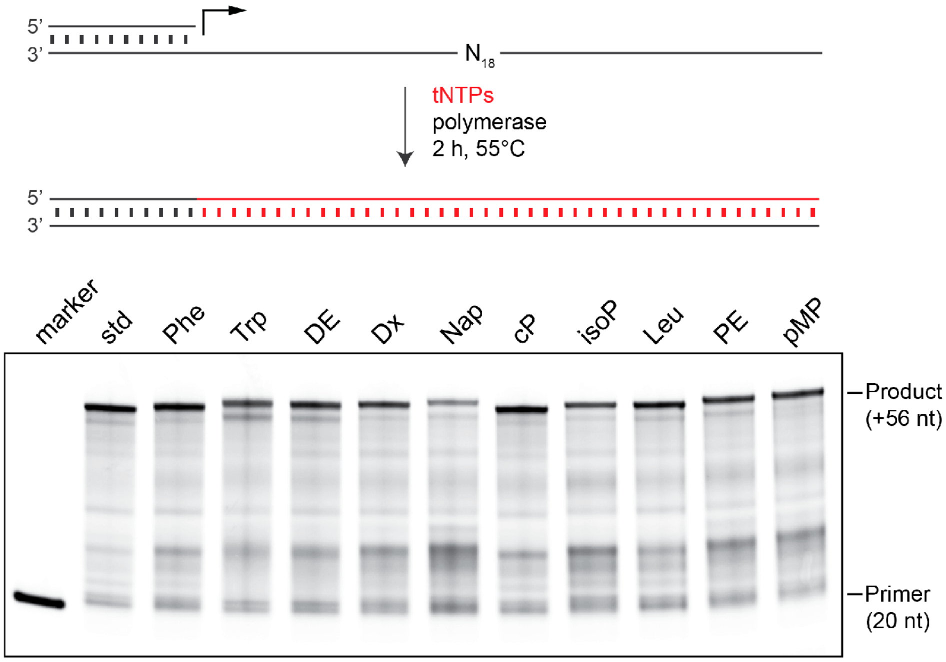


Figure 3. Polymerase-mediated synthesis of chemically modified TNA oligonucleotides having diverse sequence composition. A 5′-IR680-labeled DNA primer annealed to a DNA library was extended with TNA using a laboratory-evolved TNA polymerase. In the cartoon image, DNA is depicted in black and TNA is shown in red. The reaction conditions included 1 μM primer/template duplex, 100 μM tNTPs, 1 μM Kod-RSGA, and 1X ThermoPol buffer with heating for 2 h at 55 °C. Products of the primer-extension reaction were analyzed by denaturing polyacrylamide gel electrophoresis. The primer only and the standard base reaction denote the markers for the starting primer and fully elongated product strands. Abbreviations: standard (std), phenylalanine (Phe), tryptophan (Trp), dioxyethyl (DE), dioxol (Dx), naphthalene (Nap), cyclopropyl (cP), isopentyl (isoP), leucine (Leu), phenethyl (PE), and *p*-methoxyphenyl (pMP).

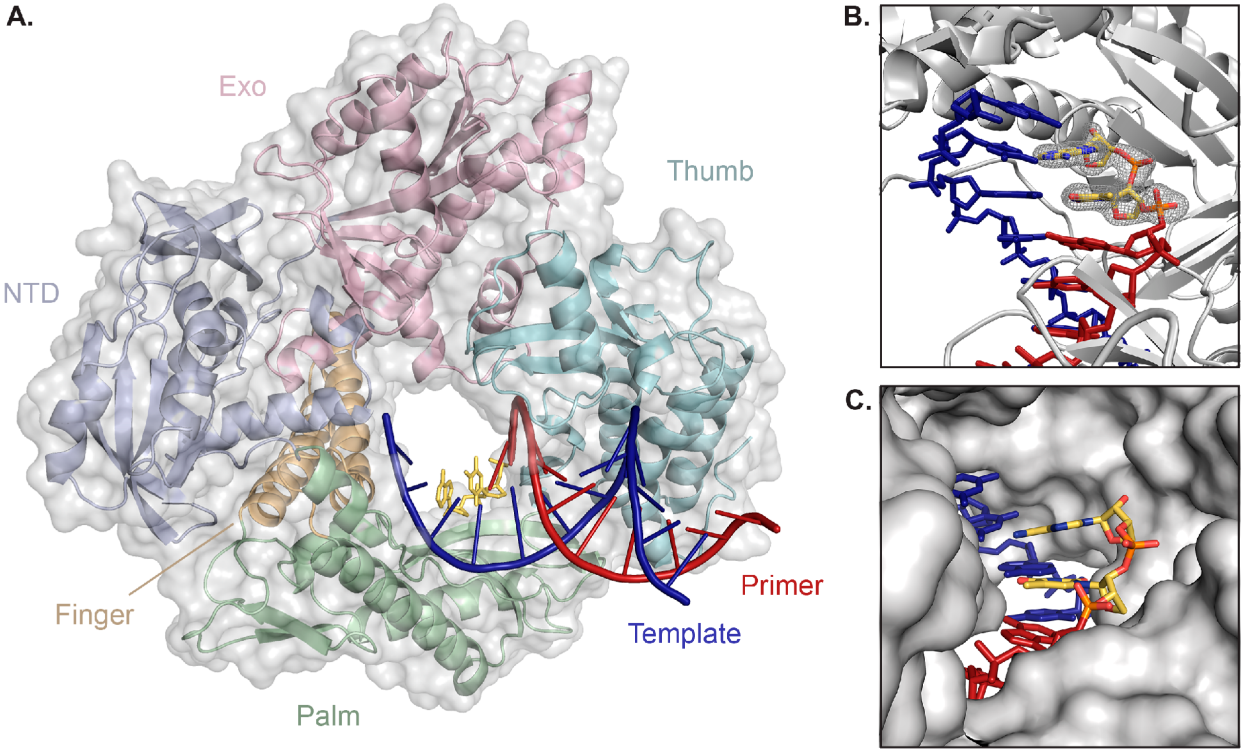


Figure 4. Crystal structure of Kod-RSGA TNA polymerase. (A) Anatomy of engineered TNA polymerase Kod-RSGA colored by domain. The DNA primer−template duplex is shown in red and blue, respectively, sitting in the groove between the palm and thumb subdomains. Two α-L-threofuranosyl nucleic acid (TNA) additions (tT and tA) are shown in yellow. (B) The active site of Kod-RSGA, showing Watson−Crick base pairing between the TNA and DNA nucleotides. A polder map of the tT and tA additions is shown in gray and contoured at 5.0σ. (C) Surface representation of the enzyme-active site in the binary conformation overlooking the cavity in the major groove of the TNA/DNA heteroduplex.

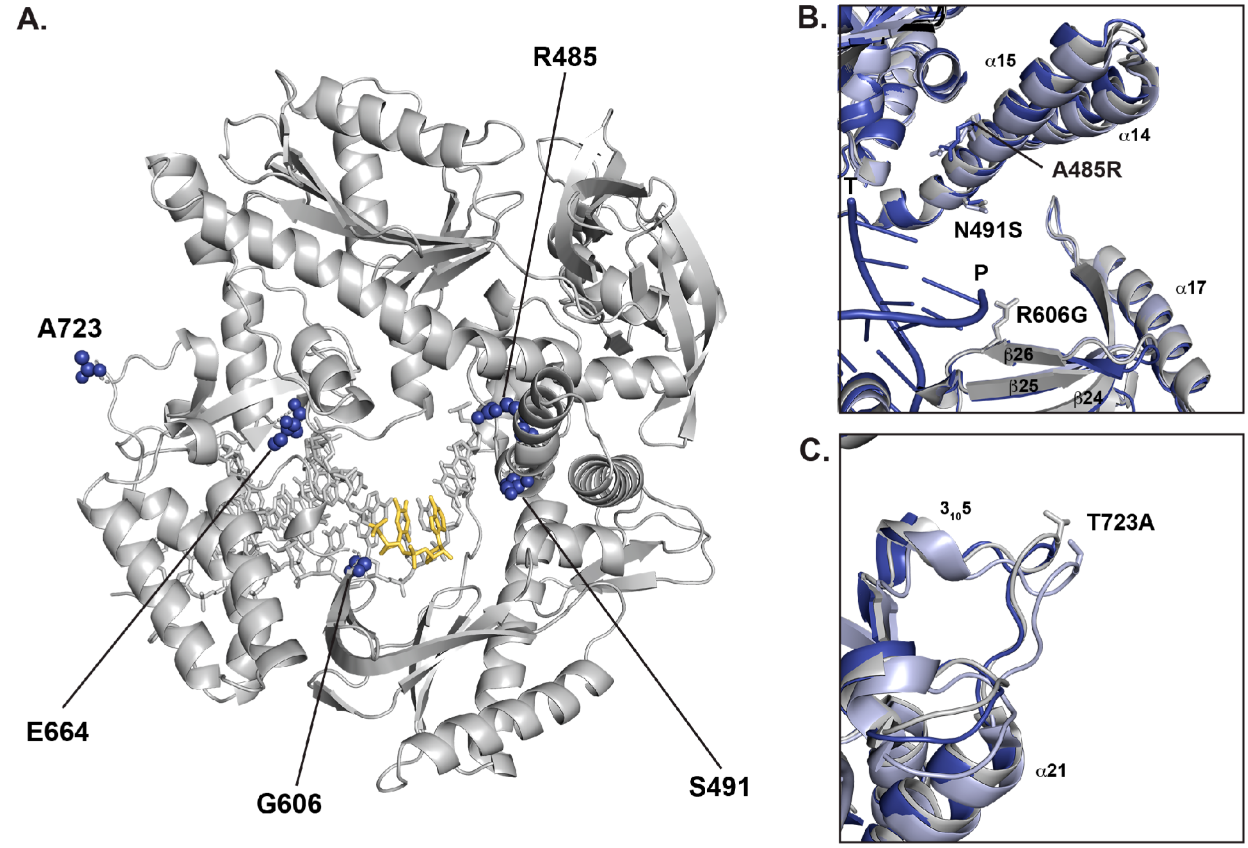


Figure 5. Kod-RSGA gain-of-function mutations. (A) Binary structure of Kod-RSGA in complex with a DNA duplex (light gray) with two TNA additions (yellow) to the 3′-end of the DNA primer. Locations of mutated residues in Kod-RSGA are shown in blue: A485R, N491S, R606G, and T723A, including the reversion mutation I664E from Kod-RI. (B) Mutated residues of Kod-RSGA finger (A485R, N491S) and thumb (R606G) subdomains shown in blue, superimposed with Kod-RI (light blue; PDB: 5VU9) and Kod WT (gray; PDB: 4K8Z). (C) Solvent-exposed loop region harboring the T723A mutation of the thumb, superimposed with Kod-RI and Kod-WT.

that comigrates with the positive control containing only natural bases (Figure 3). Although slight differences in primer extension efficiency are observed between the different TNA substrates, for example, Nap and isoP produce less full-length material in comparison to the other modified tNTPs, it may be possible to improve the incorporation of these substrates by optimizing the reaction conditions through increased enzyme concentration or reaction time.

Structural insights into Kod-RSGA were obtained by crystallizing the postcatalytic product of a primer extension reaction that added two TNA residues onto the 3′-end of a DNA primer−template duplex. Kod-RSGA was recently identified as a highly efficient TNA polymerase from a programmed allelic mutagenesis assay that comprehensively screened all possible single-point mutations in the catalytic domain (~400 aa) of the wild-type DNA polymerase isolated from the hyperthermophilic archaeal species *Thermococcus kodakarensis* (Kod).[38] Kod-RSGA functions with a catalytic rate of ~10 nt/min, which is ~20-fold faster than that of Kod-RI,[39] the only other TNA polymerase for which a crystal structure is available.[40] Diffraction-quality crystals were identified from a screening of ~600 conditions in a hanging-drop format using a Mosquito crystallization robot and subsequently optimized to 0.1 M $Mg(OAc)_2$, 0.1 M MOPS (pH 7.0), and 10% PEG 8000. The best crystal resolved to a diffraction limit of 2.1 Å (Table S4) and was solved by a molecular replacement using Kod-RI as the search model (PDB: 5VU9) for the Kod-RSGA structure.

Kod-RSGA adopts a disk-shaped architecture (Figure 4A) that contains an N-terminal domain (NTD), an exonuclease domain (Exo), and a catalytic domain consisting of the palm, finger, and thumb subdomains. As with Kod-RI and the parent wild-type enzyme, the primer−template duplex is bound in a deep groove defined by the palm and thumb subdomains.[40,41] The finger subdomain lies in an open conformation, ready to accept an incoming tNTP substrate. The DNA duplex is primarily recognized by contacts to the phosphodiester backbone, with a smaller number of contacts to the sugar and base moieties as previously described.[40,41] All of the sugar and nucleobase contacts occur through the minor groove, which is consistent with the tendency for B-family DNA polymerases to accept modified nucleotides bearing functional groups at the C-5 pyrimidine and C-7 deazapurine positions.[42] A polder map contoured at 5σ (Figure 4B) clearly shows that the newly incorporated tT and tA nucleotides at positions 12 and 13 of the primer strand adopt Watson−Crick base pairs with their complementary nucleotides in the DNA template. An analysis of the protein surface reveals a large opening near the enzyme active site (Figure 4C) with the N7 atom of the newly incorporated tA nucleobase pointing toward the center of the cavity which is accessible to bulk solution.

Kod-RSGA differs from wild-type Kod by the presence of the mutations A485R, N491S, R606G, and T723A, which confer TNA synthesis activity on the natural enzyme (Figure 5), as well as the 3′-5′-exonuclease silencing mutations D141A and E143A. Mutations A485R and N491S are located in the finger subdomain, while mutations R606G and T723A are located in the palm subdomain. The A485R mutation is known to be a strong determinant of sugar recognition, allowing natural polymerases to recognize a wide variety of sugar-modified substrates.[43] The change in enzymatic activity is thought to be due to altered dynamics between the open and closed states of the fingers, which increase the occupancy of the closed conformation necessary for catalysis.[43] Mutations N491S and R606G are both located in the enzyme active site (Figure 5B), with 491S situated directly over the nucleobase of the incoming tNTP, as observed for dNTP in the wild-type Kod structure.[41] The 606G mutation faces the primer—template duplex, adjacent to the expected nucleotide binding site. Both mutations, N491S and R606G, substitute bulky side chains for smaller, more flexible groups that may help position the tNTP substrate in the polymerase active site for chemical bond formation. Finally, mutation T723A is located on a large solvent-exposed loop on the surface of the protein, ~45 Å from the enzyme active site. We postulate that this mutation may play a role in the long-range dynamics required for efficient primer-extension activity.

■ DISCUSSION

Equipping xeno-nucleic acid polymers, such as TNA, with diversity-enhancing functional groups poses a considerable synthetic challenge, as these systems are not commercially available and can only be obtained by chemical synthesis. For example, a complete set of all four TNA substrates (tATP, tCTP, tTTP, and tGTP) requires 56 chemical transformations to convert L-ascorbic acid into the desired nucleoside triphosphates.[36,37] Synthesis of the 10 C-5 modified tUTP analogues described in the current study necessitates an additional 150 chemical transformations (15 steps each), with the critical step being the palladium-catalyzed cross-coupling reaction required to functionalize the uracil nucleobase with a side chain. Further complicating the problem is the known tendency for the 3′-hydroxyl position of α-L-threofuranosyl nucleosides to be substantially less reactive than the 5′-hydroxyl position of most DNA and RNA analogues.[44] The 3′-hydroxyl position of TNA is more sterically encumbered in comparison to the 5′-hydroxyl position of DNA due to its close proximity to the nucleobase and suffers from the fact that it is a secondary alcohol, which is less nucleophilic than a primary alcohol. Despite these challenges, the current methodology proved exceptionally versatile, allowing for a broad range of chemically diverse side chains to be installed on the tUTP substrate.

Recognizing that our synthetic approach demanded the production of reasonably large quantities (50−100 mg) of C-5 modified tUTP substrates for use as substrates during the *in vitro* selection of functionally enhanced TNA aptamers, we chose to pursue a convergent phosphorylation strategy that avoids the need for purification by high-performance liquid chromatography (HPLC).[36] In our experience, the separation of highly polar nucleoside triphosphates from polar side products using HPLC is a tedious process that limits the scale of nucleoside triphosphate synthesis to a few tens of milligrams per week, which was unsuitable for the current application.[45] In contrast to more conventional nucleoside triphosphate reagents,[46−48] we chose to couple an organic P(V) pyrene pyrophosphate reagent to the protected form of the imidazole-activated nucleoside monophosphates of C-5-modified tUMPs.[36] The product of this reaction is a fully protected nucleoside triphosphate derivative that is sufficiently hydrophobic so that it can be purified by silica gel chromatography.[36] Following purification, the triphosphate intermediate is deprotected with concentrated ammonium hydroxide (33% aqueous) and precipitated as the sodium salt using standard conditions. This synthetic approach was highly effective at generating the desired tUTP analogues in reasonable yield, purity, and scale.

The next synthetic challenge that we faced was the synthesis of base-modified TNA libraries by primer extension. To our surprise, Kod-RSGA was able to recognize each of the modified substrates without the need for any optimization of the reaction conditions. In this case, the TNA polymerase was able to extend a DNA primer with 36 contiguous TNA residues by copying a library of DNA templates into TNA. A comparison of the various primer-extension reactions by denaturing PAGE reveals that the modified uridine nucleotides are incorporated into the TNA library with the same high level of efficiency as for the TNA thymidine nucleotide. The only modified substrates that exhibited slightly reduced levels of full-length product were the naphthalene and isopentyl analogues, which may be physically constrained in the enzyme active site by the methylene linker connecting the naphthalene ring system to the amide moiety at the uracil C-5 position. However, future studies are needed to fully elucidate the relationship between the carbon linker and the incorporation efficiency of bulky side chains.

We gained structural insights into the mechanism of C-5-modified tUTP substrate incorporation into TNA oligonucleotides by solving the X-ray crystal structure of Kod-RSGA to high resolution. Our crystal structure captured the postcatalytic product of a primer extension reaction that added two TNA residues onto the 3′-end of a DNA primer. Consistent with the known proclivity of archaeal B-family DNA polymerases to accept a broader range of base-modified substrates in comparison to bacterial A-family DNA polymerases,[49] our structure revealed a wide, solvent-exposed channel in the enzyme active site that opens to the major groove of the DNA substrate. Similar observations have also been made for the wild-type Kod DNA polymerase.[41,50] The pocket is large enough to accommodate a wide assortment of aryl and aliphatic side chains found at the C-5 position of the modified tUTP substrates, which would occupy the major groove following their incorporation into the growing TNA strand.

The crystal structure of Kod-RSGA also revealed the location and orientation of the gain-of-function mutations required for efficient TNA synthesis. Although it is difficult to interpret the functional role of these amino acid changes in the absence of the tNTP substrate, we speculate that these changes confer TNA synthesis activity on the wild-type enzyme through a combination of factors that optimize substrate recognition in the enzyme active site and allow for more favorable short- and long-range dynamic interactions. For example, mutation N491S is situated at a position that would contact the nucleobase of the incoming tNTP substrate, while the R606G mutation faces the minor groove of the primer—template duplex, just below the expected nucleotide binding site. The locations of these mutations suggest that they may help coordinate the tNTP substrate in the polymerase active site for chemical bond formation. Additionally, these mutations substitute bulky side chains for smaller, more flexible residues that could play a role in short-range dynamics. In contrast, the A485R and T723A mutations are located outside the active site, suggesting that they may be responsible for interactions involving the opening and closing of the finger subdomain, which is required for chemical bond formation. For example, these mutations may lower the fidelity of the ajar conformation observed in A-family DNA polymerases, which is thought to be a critical checkpoint in the recognition of incoming nucleoside triphosphates.[51] Alternatively, they could relax the specificity of other checkpoints in the DNA synthesis mechanism, such as the geometry of the duplex following the incorporation of several modified nucleotides.[52]

CONCLUSION

In summary, we report the synthesis and polymerase recognition of TNA triphosphates equipped with diverse chemical functionalities. We have found that, despite significant chemical differences, the modified side chains are efficiently incorporated into TNA libraries. The ability to construct functionally enhanced TNA libraries by polymerase extension suggests that it should be possible to explore new regions of chemical space by *in vitro* selection. Such experiments could one day narrow the gap between aptamers and antibodies as affinity reagents in diagnostic and therapeutic applications.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c08649.

> Methods and supplementary figures, schemes, and tables, compound characterization with ^{1}H, ^{13}C, and ^{31}P NMR spectra, and synthetic methods used to prepare the 10 tUTP analogues (PDF)

Accession Codes

CCDC 2096319 and 2102461 contain the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, or by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033. Coordinates and structure factors for the crystal structure of Kod-RSGA have been deposited in the PDB with the accession code 7RSU.

AUTHOR INFORMATION

Corresponding Author

John C. Chaput − *Department of Pharmaceutical Sciences, Department of Chemistry, Department of Molecular Biology and Biochemistry, and Department of Chemical and Biomolecular Engineering, University of California, Irvine, California 92697-3958, United States*; orcid.org/0000-0003-1393-135X; Phone: 949-824-8149; Email: jchaput@uci.edu

Authors

Qingfeng Li − *Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Victoria A. Maola − *Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Nicholas Chim − *Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Javeena Hussain − *Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Adriana Lozoya-Colinas − *Department of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.1c08649

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We wish to thank members of the Chaput laboratory for helpful comments and suggestions on the manuscript. This work was supported by a sponsored research project from X, The Moonshot Factory and by a grant from the National Science Foundation (MCB: 1946312).

REFERENCES

(1) Wilson, D. S.; Szostak, J. W. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* **1999**, *68*, 611−647.

(2) Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. Analysis of aptamer discovery and technology. *Nat. Rev. Chem.* **2017**, *1*, 0076.

(3) Bittker, J. A.; Phillips, K. J.; Liu, D. R. Recent advances in the in vitro evolution of nucleic acids. *Curr. Opin. Chem. Biol.* **2002**, *6*, 367−374.

(4) Mayer, G. The chemical biology of aptamers. *Angew. Chem., Int. Ed.* **2009**, *48*, 2672−2689.

(5) Nikoomanzar, A.; Chim, N.; Yik, E. J.; Chaput, J. C. Engineering polymerases for applications in synthetic biology. *Q. Rev. Biophys.* **2020**, *53*, e8.

(6) Alves Ferreira-Bravo, I.; Cozens, C.; Holliger, P.; DeStefano, J. J. Selection of 2′-deoxy-2′-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Res.* **2015**, *43*, gkv1057.

(7) Rose, K. M.; Alves Ferreira-Bravo, I.; Li, M.; Craigie, R.; Ditzler, M. A.; Holliger, P.; DeStefano, J. J. Selection of 2′-Deoxy-2′-Fluoroarabino Nucleic Acid (FANA) Aptamers That Bind HIV-1 Integrase with Picomolar Affinity. *ACS Chem. Biol.* **2019**, *14*, 2166−2175.

(8) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S. Y.; McLaughlin, S. H.; Herdewijn, P.; Holliger, P. Synthetic genetic polymers capable of heredity and evolution. *Science* **2012**, *336*, 341−344.

(9) Eremeeva, E.; Fikatas, A.; Margamuljana, L.; Abramov, M.; Schols, D.; Groaz, E.; Herdewijn, P. Highly stable hexitol based XNA aptamers targeting the vascular endothelial growth factor. *Nucleic Acids Res.* **2019**, *47*, 4927−4939.

(10) Yu, H.; Zhang, S.; Chaput, J. C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **2012**, *4*, 183−187.

(11) Dunn, M. R.; McCloskey, C. M.; Buckley, P.; Rhea, K.; Chaput, J. C. Generating biologically stable TNA aptamers that function with high affinity and thermal stability. *J. Am. Chem. Soc.* **2020**, *142*, 7721−7724.

(12) Rangel, A. E.; Chen, Z.; Ayele, T. M.; Heemstra, J. M. In vitro selection of an XNA aptamer capable of small-molecule recognition. *Nucleic Acids Res.* **2018**, *46*, 8057−8068.

(13) Li, X.; Li, Z.; Yu, H. Selection of threose nucleic acid aptamers to block PD-1/PD-L1 interaction for cancer immunotherapy. *Chem. Commun.* **2020**, *56*, 14653−14656.

(14) Hoshino, H.; Kasahara, Y.; Kuwahara, M.; Obika, S. DNA Polymerase Variants with High Processivity and Accuracy for Encoding and Decoding Locked Nucleic Acid Sequences. *J. Am. Chem. Soc.* **2020**, *142*, 21530−21537.

(15) Culbertson, M. C.; Temburnikar, K. W.; Sau, S. P.; Liao, J.-Y.; Bala, S.; Chaput, J. C. Evaluating TNA stability under simulated physiological conditions. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 2418−2421.

(16) Chaput, J. C.; Herdewijn, P. What Is XNA? *Angew. Chem., Int. Ed.* **2019**, *58*, 11570−11572.

(17) Gold, L. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS One* **2010**, *5*, e15004.

(18) Brudno, Y.; Liu, D. R. Recent progress toward the templated synthesis and directed evolution of sequence-defined synthetic polymers. *Chem. Biol.* **2009**, *16*, 265−276.

(19) Pfeiffer, F.; Rosenthal, M.; Siegl, J.; Ewers, J.; Mayer, G. Customised nucleic acid libraries for enhanced aptamer selection and performance. *Curr. Opin. Biotechnol.* **2017**, *48*, 111−118.

(20) Temme, J. S.; MacPherson, I. S.; DeCourcey, J. F.; Krauss, I. High temperature SELMA: evolution of DNA-supported oligomannose clusters which are tightly recognized by HIV bnAB 2G12. *J. Am. Chem. Soc.* **2014**, *136*, 1726−1729.

(21) Tolle, F.; Brandle, G. M.; Natzner, D.; Mayer, G. A versatile approach towards nucleobase-modified aptamers. *Angew. Chem., Int. Ed.* **2015**, *54*, 10971−10974.

(22) Gordon, C. K. L.; Wu, D.; Pusuluri, A.; Feagin, T. A.; Csordas, A. T.; Eisenstein, M. S.; Hawker, C. J.; Niu, J.; Soh, H. T. Click-Particle Display for Base-Modified Aptamer Discovery. *ACS Chem. Biol.* **2019**, *14*, 2652−2662.

(23) Kong, D.; Yeung, W.; Hili, R. In Vitro Selection of Diversely Functionalized Aptamers. *J. Am. Chem. Soc.* **2017**, *139*, 13977−13980.

(24) Chen, Z.; Lichtor, P. A.; Berliner, A. P.; Chen, J. C.; Liu, D. R. Evolution of sequence-defined highly functionalized nucleic acid polymers. *Nat. Chem.* **2018**, *10*, 420−427.

(25) Brudno, Y.; Birnbaum, M. E.; Kleiner, R. E.; Liu, D. R. An in vitro translation, selection and amplification system for peptide nucleic acids. *Nat. Chem. Biol.* **2010**, *6*, 148−155.

(26) Niu, J.; Hili, R.; Liu, D. R. Enzyme-free translation of DNA into sequence-defined synthetic polymers structurally unrelated to nucleic acids. *Nat. Chem.* **2013**, *5*, 282−292.

(27) Vaish, N. K.; Larralde, R.; Fraley, A. W.; Szostak, J. W.; McLaughlin, L. W. A novel, modification-dependent ATP-binding aptamer selected from an RNA library incorporating a cationic functionality. *Biochemistry* **2003**, *42*, 8842−51.

(28) Battersby, T. R.; Ang, D. N.; Burgstaller, P.; Jurczyk, S. C.; Bowser, M. T.; Buchanan, D. D.; Kennedy, R. T.; Benner, S. A. Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog. *J. Am. Chem. Soc.* **1999**, *121*, 9781−9789.

(29) Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. A self-cleaving DNA enzyme modified with amines, guanidines and imidazoles operates independently of divalent metal cations (M2+). *Nucleic Acids Res.* **2009**, *37*, 1638−49.

(30) Wang, Y.; Liu, E.; Lam, C. H.; Perrin, D. M. A densely modified M(2+)-independent DNAzyme that cleaves RNA efficiently with multiple catalytic turnover. *Chem. Sci.* **2018**, *9*, 1813−1821.

(31) Cheung, Y. W.; Rothlisberger, P.; Mechaly, A. E.; Weber, P.; Levi-Acobas, F.; Lo, Y.; Wong, A. W. C.; Kinghorn, A. B.; Haouz, A.; Savage, G. P.; Hollenstein, M.; Tanner, J. A. Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 16790−16798.

(32) Gawande, B. N.; Rohloff, J. C.; Carter, J. D.; von Carlowitz, I.; Zhang, C.; Schneider, D. J.; Janjic, N. Selection of DNA aptamers with two modified bases. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 2898−2903.

(33) Medina, E.; Yik, E. J.; Herdewijn, P.; Chaput, J. C. Functional Comparison of Laboratory-Evolved XNA Polymerases for Synthetic Biology. *ACS Synth. Biol.* **2021**, *10*, 1429−1437.

(34) Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. Expanding the chemistry of DNA for in vitro selection. *J. Am. Chem. Soc.* **2010**, *132*, 4141−51.

(35) Mei, H.; Chaput, J. C. Expanding the chemical diversity of TNA with tUTP derivatives that are substrates for a TNA polymerase. *Chem. Commun.* **2018**, *54*, 1237−1240.

(36) Liao, J.-Y.; Bala, S.; Ngor, A. K.; Yik, E. J.; Chaput, J. C. P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates. *J. Am. Chem. Soc.* **2019**, *141*, 13286−13289.

(37) Sau, S. P.; Fahmi, N. E.; Liao, J.-Y.; Bala, S.; Chaput, J. C. A scalable synthesis of α-L-threose nucleic acid monomers. *J. Org. Chem.* **2016**, *81*, 2302−2307.

(38) Nikoomanzar, A.; Vallejo, D.; Yik, E. J.; Chaput, J. C. Programmed allelic mutagenesis of a DNA polymerase with single amino acid resolution. *ACS Synth. Biol.* **2020**, *9*, 1873−1881.

(39) Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C. Evaluating the rate and substrate specificity of laboratory evolved XNA polymerases. *Anal. Chem.* **2017**, *89*, 12622−12625.

(40) Chim, N.; Shi, C.; Sau, S. P.; Nikoomanzar, A.; Chaput, J. C. Structural basis for TNA synthesis by an engineered TNA polymerase. *Nat. Commun.* **2017**, *8*, 1810.

(41) Kropp, H. M.; Betz, K.; Wirth, J.; Diederichs, K.; Marx, A. Crystal structures of ternary complexes of archaeal B-family DNA polymerases. *PLoS One* **2017**, *12*, e0188005.

(42) Hottin, A.; Marx, A. Structural Insights into the Processing of Nucleobase-Modified Nucleotides by DNA Polymerases. *Acc. Chem. Res.* **2016**, *49*, 418−27.

(43) Gardner, A. F.; Jackson, K. M.; Boyle, M. M.; Buss, J. A.; Potapov, V.; Gehring, A. M.; Zatopek, K. M.; Correa, I. R., Jr.; Ong, J. L.; Jack, W. E. Therminator DNA Polymerase: Modified Nucleotides and Unnatural Substrates. *Front. Mol. Biosci.* **2019**, *6*, 28.

(44) Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. Synthesis of a-L-threofuranosyl nucleoside triphosphates (tNTPs). *Org. Lett.* **2005**, *7*, 1485−1487.

(45) Bala, S.; Liao, J.-Y.; Mei, H.; Chaput, J. C. Synthesis of α-L-Threofuranosyl Nucleoside 3′-Monophosphates, 3′-Phosphoro(2-Methyl)imidazolides, and 3′-Triphosphates. *J. Org. Chem.* **2017**, *82*, 5910−5916.

(46) Hollenstein, M. Nucleoside triphosphates–building blocks for the modification of nucleic acids. *Molecules* **2012**, *17*, 13569−91.

(47) Kore, A. R.; Srinivasan, B. Recent Advances in the Synthesis of Nucleoside Triphosphates. *Curr. Org. Synth.* **2014**, *10*, 903−934.

(48) Roy, B.; Depaix, A.; Perigaud, C.; Peyrottes, S. Recent Trends in Nucleotide Synthesis. *Chem. Rev.* **2016**, *116*, 7854−97.

(49) Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. Systematic characterization of 2 ′-deoxynucleoside-5 ′-triphosphate analogs as substrates for DNA polymerases by polymerase chain reaction and kinetic studies on enzymatic production of modified DNA. *Nucleic Acids Res.* **2006**, *34*, 5383−5394.

(50) Kropp, H. M.; Diederichs, K.; Marx, A. The Structure of an Archaeal B-Family DNA Polymerase in Complex with a Chemically Modified Nucleotide. *Angew. Chem., Int. Ed.* **2019**, *58*, 5457−5461.

(51) Wu, E. Y.; Beese, L. S. The structure of a high fidelity DNA polymerase bound to a mismatched nucleotide reveals an ″ajar″ intermediate conformation in the nucleotide selection mechanism. *J. Biol. Chem.* **2011**, *286*, 19758−67.

(52) Miller, H.; Grollman, A. P. Kinetics of DNA polymerase I (Klenow fragment exo-) activity on damaged DNA templates: effect of proximal and distal template damage on DNA synthesis. *Biochemistry* **1997**, *36*, 15336−42.