ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

Functional Comparison of Laboratory-Evolved XNA Polymerases for Synthetic Biology

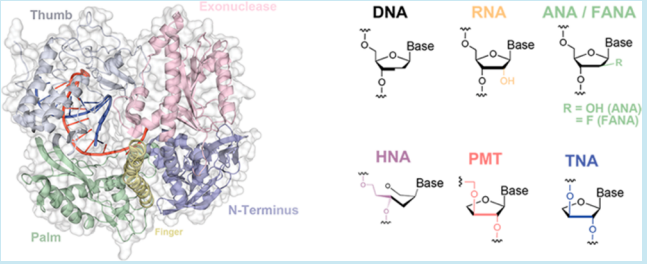
Esau Medina,[⊥] Eric J. Yik,[⊥] Piet Herdewijn, and John C. Chaput*

Cite This: *ACS Synth. Biol.* 2021, 10, 1429–1437

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Artificial genetic polymers (XNAs) have enormous potential as new materials for synthetic biology, biotechnology, and molecular medicine; yet, very little is known about the biochemical properties of XNA polymerases that have been developed to synthesize and reverse-transcribe XNA polymers. Here, we compare the substrate specificity, thermal stability, reverse transcriptase activity, and fidelity of laboratory-evolved polymerases that were established to synthesize RNA, 2′-fluoroarabino nucleic acid (FANA), arabino nucleic acid (ANA), hexitol nucleic acid (HNA), threose nucleic acid (TNA), and phosphonomethylthreosyl nucleic acid (PMT). We find that the mutations acquired to facilitate XNA synthesis increase the tolerance of the enzymes for sugar-modified substrates with some sacrifice to protein-folding stability. Bst DNA polymerase was found to have weak reverse transcriptase activity on ANA and uncontrolled reverse transcriptase activity on HNA, differing from its known recognition of FANA and TNA templates. These data benchmark the activity of current XNA polymerases and provide opportunities for generating new polymerase variants that function with greater activity and substrate specificity.



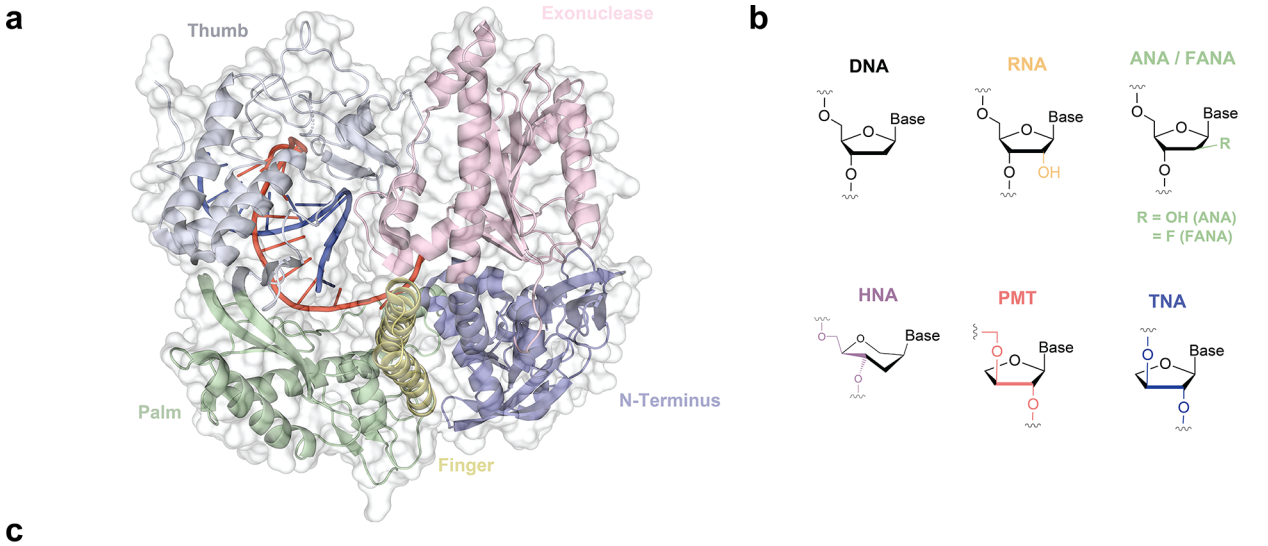
INTRODUCTION

Polymerase engineering technologies have made it possible to synthesize synthetic genetic polymers (XNAs) with backbone structures that are distinct from those found in nature.[1] These enzymes, along with other polymerases that copy XNA templates back into DNA,[2,3] have enabled researchers to explore new regions of chemical space by Darwinian evolution.[4,5] To date, molecular evolution experiments had been reported for five different classes of XNA polymers, including (1′,5′-anhydrohexitol nucleic acid (HNA),[6,7] arabino nucleic acid (ANA),[6,8] 2′-fluoroarabino nucleic acid (FANA),[6,9−11] cyclohexenyl nucleic acid (CeNA),[6] and α-L-threofuranosyl nucleic acid (TNA).[12−16] Similar polymerase engineering efforts have also enabled the replication of locked nucleic acid (LNA) polymers with promising results on the evolution of mixed-backbone LNA-modified RNA polymers.[17] Together, these results demonstrate that the well-established concepts of heredity and evolution are no longer limited to DNA and RNA but instead represent a general property of any genetic system that is capable of replication by faithfully copying parent and daughter strands.

Although the isolation of XNA molecules with catalytic and ligand binding activity represents an important landmark in the field of molecular evolution, researchers are also interested in expanding the use of XNAs to other applications in synthetic biology. One monumental challenge involves establishing genetically engineered cells that can maintain a synthetic XNA chromosome in actively dividing cells, thereby creating a genetic firewall between the biological information on the cell and the information introduced for synthetic biology purposes.[18] Though a daunting task, researchers have already shown that DNA plasmids carrying small HNA segments can restore thymidylate synthase activity in *E. coli*, demonstrating that limited stretches of XNA can be recognized by the natural polymerase machinery of the cell.[19] A fundamentally different approach for using XNAs in synthetic biology involves extending the concept of DNA information storage to synthetic genetic polymers with backbone structures that are recalcitrant to nuclease digestion, thereby preventing the accidental loss of information through unintended nuclease exposure. Toward this goal, TNA was evaluated as a biologically stable soft material for low-energy, high-density information storage using engineered polymerases that enable the writing and reading of digital information in TNA polymers.[20]

Recognizing that nucleic acid synthesis is the cornerstone of synthetic biology, the success of current and future XNA applications will ultimately hinge on the ability to encode and decode information in synthetic genetic polymers. Benchmarking the activities of existing XNA polymerases is an important

Received: February 3, 2021
Published: May 24, 2021



c

Polymerase	Substrate	Mutations
Tgo-QGLK	RNA	V93Q, D141A, E143A, Y409G, A485L, E664K
Tgo-D4K	ANA/FANA	V93Q, D141A, E143A, L403P, A485L, K569H, P657T, E658Q, T676I, Y663H, E664K, D669A, K671N
Tgo-6G12/I521L	HNA	V93Q, D141A, E143A, A485L, I521L, V589A, E609K, I610M, K659Q, E664Q, Q665P, R668K, D669A, T676R, A681S, K671N, K674R, L704P, E730G
Tgo-EPFLH	RNA	V94Q, D141A, E143A, H147E, L403P, L408F, A485L, I521L, E664H
Kod-RSGA	TNA	D141A, E143A, A485R, N491S, R606G, T723A

Figure 1. Engineered XNA polymerases. (a) Crystal structure of Kod-RI TNA polymerase (PDB: 5VU8) bound to a DNA primer−template duplex showing domains and subdomains by color. (b) The chemical structure of each XNA system showing the modified sugar moiety. (c) The amino acid mutations identified by directed evolution that facilitate RNA, ANA, FANA, HNA, PMT, and TNA synthesis. Mutations are colored by domain.

first step in this process, as it will help gauge the success of future directed evolution efforts aimed at isolating new variants with improved primer-extension efficiency, substrate specificity, and fidelity. Here, we describe a biochemical investigation into the substrate specificity, thermal stability, reverse transcriptase activity, and fidelity of XNA polymerases that were established to synthesize RNA, 2′-fluoroarabino nucleic acid (FANA), arabino nucleic acid (ANA), hexitol nucleic acid (HNA), threose nucleic acid (TNA), and phosphonomethylthreosyl nucleic acid (PMT). We find that the mutations acquired to facilitate XNA synthesis function by increasing the tolerance of the enzyme for sugar-modified substrates, implying that additional mutations are needed to complete the transition from generalists to specialists. Half of the engineered polymerases show signs of diminished activity after an extended thermal challenge at 90 °C, indicating that some of the acquired mutations have a negative effect on protein-folding stability. Bst DNA polymerase was found to have weak reverse transcriptase activity on ANA and showed signs of uncontrolled reverse transcriptase activity on HNA. These properties differ considerably from the known activity of *Geobacillus stearothermophilus* (Bst) DNA polymerase on FANA and TNA templates, where reverse transcription is both controlled and reasonably efficient. Together, these data benchmark the activity of XNA polymerases that have been discovered to replicate synthetic genetic polymers with novel backbone structures and provide rich opportunities for discovering the next generation of XNA polymerases.

RESULTS

We have previously described the discovery and utilization of Kod-RSGA and Tgo-QGLK, two B-family DNA polymerases isolated from the hyperthermophilic archaeal species *Thermococcus kodakarensis* (Kod) and *Thermococcus gorgonarius* (Tgo) that have been engineered by directed evolution to synthesize TNA and RNA oligonucleotides, respectively.[21,22] To generate the full set of XNA polymerases required to complete the study, we synthesized genes for the FANA and ANA polymerase Tgo-D4K,[6] the HNA polymerases Tgo 6G12 and Tgo-6G12/I521L,[6,8] and the PMT polymerase Tgo-EPLFH.[23] The set of six engineered polymerases carries up to 17 amino acid mutations that are primarily located in the thumb and finger regions of the catalytic domain (Figures 1 and S1). In addition, each engineered polymerase also carries the 3′,5′-exonuclease stalling (exo-) mutations D141A and E143A to inhibit proofreading activity during XNA synthesis. Genes for the new polymerases were constructed by Gibson assembly from synthetic DNA oligonucleotides and sequence-confirmed. All six XNA polymerases along with exonuclease deficient (exo-) versions of four wild-type polymerases *Thermococcus* sp. 9°N (9N), *Pyrococcus* sp. Deep vent (DV), Tgo, and Kod,[22] were expressed and purified from *E. coli* lysate using standard protocols that have been described previously (Figure S2).[24]

Assessing the success of any polymerase engineering effort involves comparing the activity of the evolved variant to the starting wild-type enzyme. We therefore began by evaluating the ability for four naturally occurring archaeal DNA polymerases (9N, DV, Tgo, and Kod) commonly used in DNA polymerase engineering efforts to extend a DNA

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

primer−template duplex with chemically synthesized XNA triphosphates (xNTPs).[1] In each case, the polymerases were challenged to extend a DNA primer with 20 XNA nucleotides by copying the DNA template into a complementary XNA strand (Figure 2). The reactions were performed using both

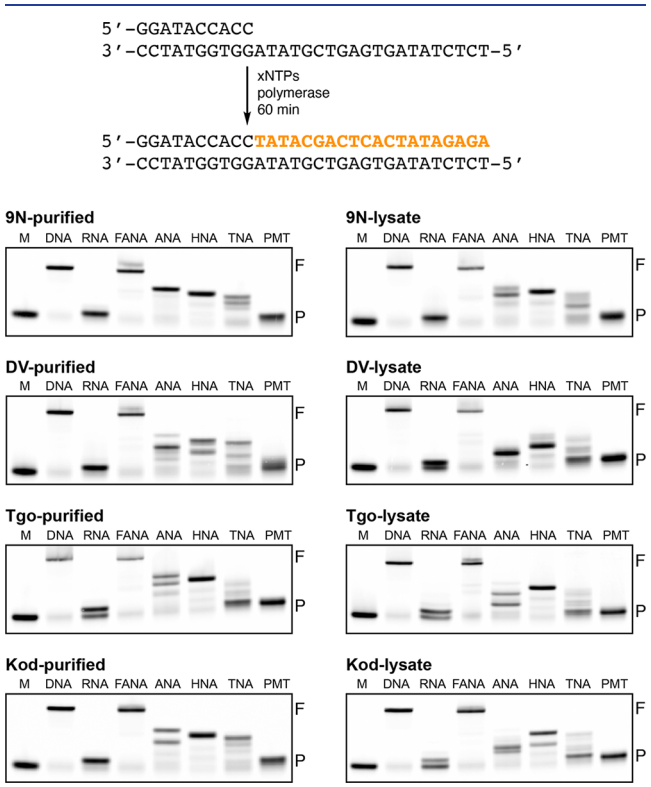


Figure 2. Substrate specificity of naturally occurring archaeal B-family DNA polymerases. Sequence of the DNA primer−template complex (black) with the XNA product shown in bold orange. Purified polymerase (left) and heat-treated clarified lysate (right) were compared in side-by-side reactions. Primer-extension reactions were performed by incubating an 5′ IR680-labeled primer−template duplex (500 nM) with xNTPs (100 μM each) and enzyme (500 nM) in 1× ThermoPol buffer for 60 min at 55 °C. Reactions were analyzed by denaturing polyacrylamide gel electrophoresis with fluorescent imaging on a LiCOR imager. P, primer; F, full-length product.

heat-treated lysate and purified enzyme to determine if any significant differences exist between purified and unpurified enzyme. Following an incubation period of 1 h at 55 °C, the reactions were quenched with formamide, and the primer-extension products were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE).

The results from our substrate specificity assay are shown in Figure 2. Several observations are immediately apparent from the set of gel images. First, it is clear that heat-treated lysate yields extension patterns that are nearly identical to the purified enzyme, demonstrating that heat-treated lysate is a viable approach for routine polymerase screening. Second, none of the polymerases recognize RNA substrates due to the strong steric gate activity of natural polymerases.[25] Third, each of the naturally occurring archaeal DNA polymerases is able to copy the DNA template into a full-length FANA product, which is consistent with the notion that FANA is a close structural analogue of DNA.[26] Fourth, ANA, HNA, and TNA are poor substrates for naturally occurring archaeal DNA polymerases, as evidenced by the appearance of truncated bands on each of the gels. This observation supports the theory that DNA polymerases have multiple gatekeeper activities; one of which can detect changes in the duplex geometry as modified nucleotides are incorporated onto the 3′ end of the DNA primer.[27] Finally, none of the polymerases recognize PMT as a substrate for XNA synthesis, implying that XNAs with both sugar and phosphate modifications are strong candidates for orthogonal genetic systems.[28]

Next, we evaluated the substrate specificity of purified versions of XNA polymerases that have been evolved in the laboratory for specific DNA-dependent XNA synthesis activities.[1,5] At the time of our study, LNA polymerases were not yet known, so the evaluation was restricted to DNA, RNA, FANA, ANA, HNA, TNA, and PMT. Similar to our analysis of wild-type DNA polymerases, a systematic primer-extension analysis was performed by separately incubating the six engineered polymerase variants with each xNTP mixture for 1 h at 55 °C in standard polymerase buffer devoid of manganese ions.

Analysis of the gel images shown in Figure 3 reveals an interesting arrangement of activity. While each XNA polymer-

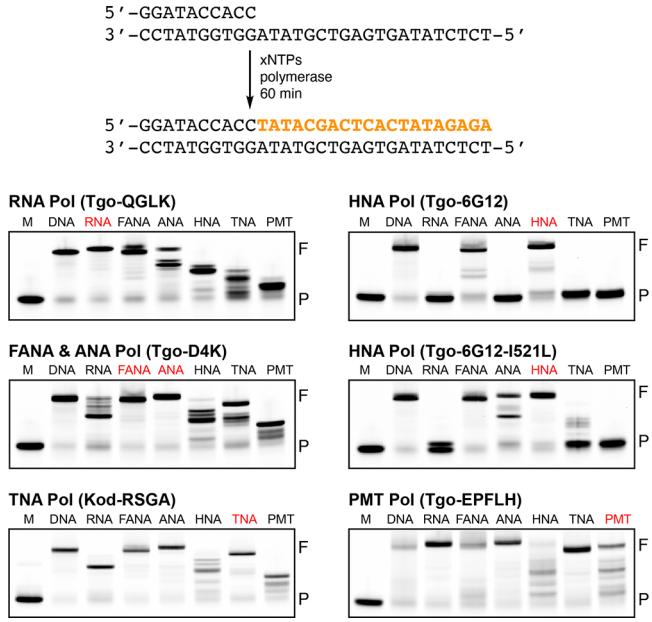


Figure 3. Substrate specificity of laboratory-evolved XNA polymerases. Sequence of the DNA primer−template complex (black) with the XNA product provided in bold orange. Cognate substrates are indicated in red above each gel. Primer-extension reactions were performed by incubating a 5′ IR680-labeled primer−template duplex (500 nM) with xNTPs (100 μM each) and enzyme (500 nM) in 1× ThermoPol buffer for 60 min at 55 °C. Reactions were analyzed by denaturing polyacrylamide gel electrophoresis with fluorescent imaging on a LiCOR imager. P, primer; F, full-length product.

ase is capable of synthesizing its cognate XNA substrate, Tgo-EPFLH did show lower than expected activity for PMT. This is due to the absence of manganese ions in the reaction mixture *versus* the presence of these ions in the original report of the PMT polymerase, as validated in separate primer-extension reactions performed in the presence of 1 mM $MnCl_2$ (Figure S3).[23] Manganese ions were excluded from the current study so that polymerase activities could be compared under standard magnesium buffer conditions typical of most polymerase studies and to avoid unwanted mutations.[1]

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

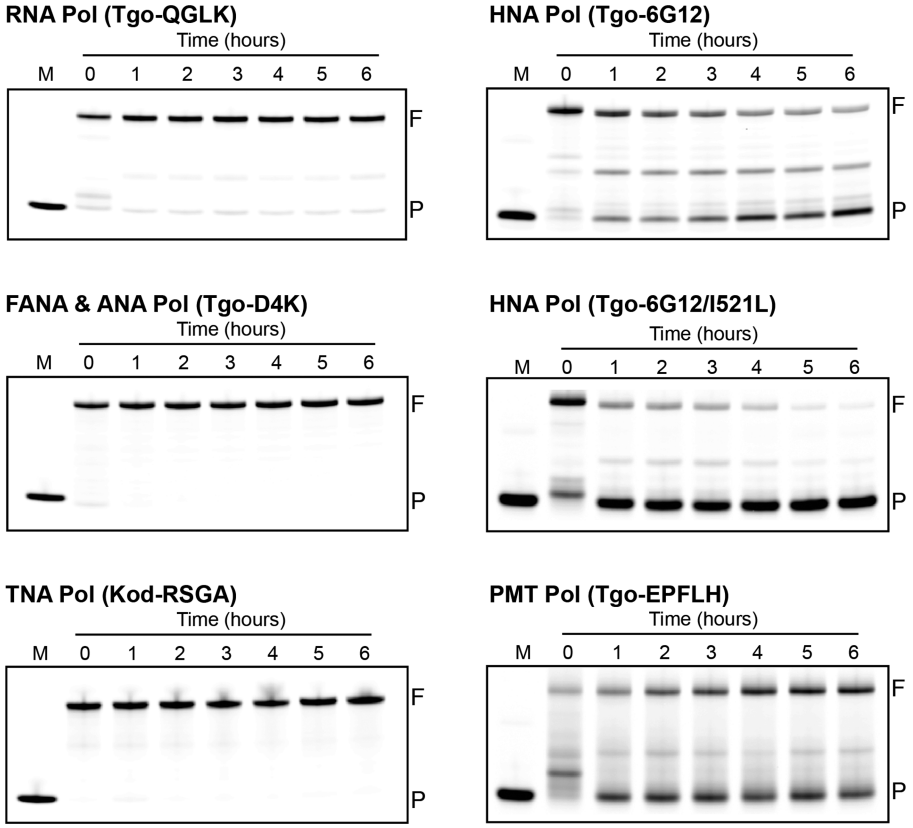


Figure 4. Thermostability of laboratory-evolved XNA polymerases. XNA polymerases were thermally challenged by heating the enzyme at 90 °C for 1−6 h prior to functional analysis in a DNA synthesis reaction. At designated time points, functional activity was assessed by incubating the enzyme with a 5′ IR680-labeled primer−template duplex (500 nM) and dNTPs (100 μM each) for 60 min at 55 °C. Reactions were analyzed by denaturing polyacrylamide gel electrophoresis with fluorescent imaging on a LiCOR imager. P, primer; F, full-length product.

Additionally, all of the engineered polymerases continue to recognize DNA substrates as well as other noncognate XNA substrates, indicating that the gain-of-function activity observed for the evolved XNA polymerases is due to a broadening of substrate specificity rather than a convergence on new active site configurations that function with higher specificity. The HNA polymerase Tgo-6G12 is the most specific polymerase, having activity for HNA, DNA, and FANA but no other XNA substrates. The other XNA polymerases exhibit promiscuous activity, especially Kod-RSGA and Tgo-EPFLH, both of which synthesize full-length products for XNAs other than their cognate TNA and PMT substrates, respectively. For example, Kod-RSGA shows strong activity for ANA, while Tgo-EPFLH exhibits activity for RNA, ANA, and TNA.

Intrigued by the strong cross-reactivity of Kod-RSGA and Tgo-EPFLH, we performed a time course analysis to compare the efficiency of primer extension of ANA and TNA using cognate and noncognate enzyme pairs. Our analysis reveals that Kod-RSGA is capable of synthesizing the full-length ANA product, albeit ∼4-fold slower than the cognate D4K polymerase previously evolved for ANA (Figure S4).[6] The results are somewhat different for TNA, as Tgo-EPFLH was only able to generate trace amounts of the full-length TNA product after 3 h, while the cognate Kod-RSGA polymerase converted all of the primer to full-length TNA product (Figure S4). The reduced activity of Tgo-EPFLH relative to the data observed in the comprehensive screen (Figure 3) is likely due to the longer length of the template strand, which increased the unpaired region from 20 to 70 nucleotides as a more stringent test of XNA synthesis activity.

Since XNA polymerases have thus far primarily been used to isolate XNA aptamers and catalysts by *in vitro* selection, we chose to evaluate the ability of each XNA polymerase to synthesize a naïve XNA library. In this experiment, each XNA polymerase was challenged to extend a 20-mer DNA primer annealed to an 80-mer DNA template containing a central random region of 40 contiguous nucleotide positions. Analysis of the primer-extension assays by denaturing PAGE indicates that all but one of the engineered polymerases are able to copy the DNA library into XNA (Figure S5). The only exception was Tgo-EPFLH, which produced a smear of truncated bands, indicating that this polymerase is not sufficiently active to copy DNA pools into PMT.

Tgo-EPFLH was previously reported to be highly selective against dNTP substrates, which is also consistent with the results observed in our polymerase activity screen shown in Figure 3. However, we were concerned that reduced formation of the DNA product observed under single-end-point detection conditions (1 h at 55 °C) could be due to continued DNA synthesis that proceeds *via* an untemplated primer-extension mechanism. To address this concern, we performed a time course analysis of Tgo-EPFLH mediated DNA synthesis in the presence and absence of the complementary template. The results from this assay (Figure S6) show that Tgo-EPFLH is prone to significant terminal transferase activity, whereby dNTP substrates are randomly added to the DNA primer. Although all of the XNA polymerases exhibit modest levels of

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

Table 1. Aggregate Replication Fidelity[a]

XNA	polymerase	reverse transcriptase	indel error ($\times 10^{-3}$)	misincorporation error ($\times 10^{-3}$)	fidelity (%)	ref.
FANA	Tgo-D4K	RT521K	5.0	9.45	98.6	1
FANA	Tgo-WT	Bst	0.83	0.83	99.8	2
ANA	Tgo-D4K	RT521K	5.8	7.66	98.8	1
ANA	Tgo-D4K	Bst	0.9	1.80	99.7	
HNA	Tgo-6G12	RT521	19.7	6.80	97.4	1
HNA	Tgo-6G12/I521L	RT521	n/a	3.00	99.7	3
TNA	Kod-RSGA	Bst	3.7	9.50	99.1	4

[a]n/a = not available. Bold entries denote new fidelity results obtained using Bst DNA polymerase for XNA reverse transcription. Nonbold entries denote comparison data from previous studies using the XNA reverse transcriptases RT521, RT521K, or Bst.

template-free DNA synthesis, the problem is more acute for Tgo-EPFLH (Figure S7). This observation, which to our knowledge has not been discussed previously, may imply that the acquired mutations are reducing the dependency of the evolved polymerases for the template.

Next, we evaluate the effect of the gain-of-function mutations discovered by directed evolution on the thermal stability of the engineered polymerase scaffolds. Reduced thermal stability is a common problem observed among engineered proteins generated by directed evolution but is rarely discussed among engineered polymerases developed for synthetic biology.[1] For this assay, the set of engineered polymerases was thermally challenged by preheating the enzymes for 1−6 h at 90 °C in polymerase buffer. At designated time points, the enzyme was cooled to room temperature, combined with the primer−template duplex and dNTP substrates, and incubated for 1 h at 55 °C to test for enzymatic activity *via* DNA synthesis in a primer-extension assay. Analysis of the resulting gels (Figure 4) reveals a range of DNA synthesis activity after the thermal challenge. The RNA, FANA/ANA, and TNA polymerases, Tgo-QGLK, Tgo-D4K, and Kod-RSGA, respectively, show no changes in activity even after 6 h of preincubation at 90 °C, demonstrating that the acquired mutations found in these polymerases do not appear to effect the folding stability of the protein. By contrast, the HNA polymerases Tgo-6G12 and Tgo-6G12/I521L as well as the PMT polymerase Tgo-EPFLH exhibit a diminished capacity for DNA synthesis, indicating that some of the acquired mutations have an unintended negative effect on protein-folding stability.

We have previously shown that a naturally occurring DNA polymerase I member isolated from thermophilic bacterial species *Geobacillus stearothermophilus* (Bst) is able to complete the XNA replication cycle by copying TNA and FANA oligonucleotides back into DNA (Table 1).[2,10] X-ray crystal structures of the post-translocated product of DNA synthesis on XNA strands composed entirely of TNA or FANA suggest that structural plasticity within the active site allows the enzyme to function as a XNA-dependent DNA polymerase.[29] This phenomenon was extensively studied in a recent data archiving project, in which Bst was found to faithfully recover 22 349 bytes of digital information stored in 7451 unique TNA oligonucleotides.[20]

Encouraged by these results, we wished to explore the ability of Bst DNA polymerase to reverse-transcribe other synthetic congeners of XNA. In particular, we focused on ANA and HNA oligonucleotides, which currently require the engineered polymerases RT521K and RT521 to effect the conversion of ANA and HNA information back into DNA (Table 1), respectively.[3] As with TNA and FANA previously,[2,10] we evaluated the ability for Bst DNA polymerase to synthesize DNA on ANA and HNA templates, respectively. Analysis of the resulting gels (Figure S8) indicate that Bst is capable of generating only modest amounts (<5%) of full-length DNA product on ANA. The problem is worse for HNA, with truncated adducts observed after 2 h and overextended products observed at the 3 h incubation time. We postulate that the overextension may be due to strand switching, but further studies are needed to confirm this hypothesis.

Given the ability for Bst DNA polymerase to synthesize small amounts of full-length DNA product on the ANA template, we decided to measure the fidelity of ANA replication using an assay (Table 1 and Figure S9) that involves a complete replication cycle (DNA → XNA → DNA).[2,10] This assay calculates the aggregate fidelity of replication, which is operationally different than the more restricted view of fidelity as single-nucleotide incorporation events. To ensure that the fidelity values represent the true fidelity of the replication cycle, we use a TT mismatch in the primer region that results in a double-AA transversion in strands that are correctly replicated as well as no reverse transcriptase and no template controls in the PCR reaction (Figure S10). Alignment of the sequences from more than 1000 nucleotide incorporation events (Figures S11) yields a replication fidelity of 99.7% for ANA replication, which is among the highest observed for known XNA replication systems.

DISCUSSION

Recent advances in the field of polymerase engineering have given rise to a new generation of enzymes that can recognize a wide range of nucleotide substrates with modifications made to the sugar, phosphate, and nucleobase moieties.[1,30,31] The use of these enzymes in practical applications aimed at generating biologically stable versions of receptor and catalysts provides new opportunities for diagnostic and therapeutic applications as well as informational systems for data archiving. Although the progress in this area has been inspiring, especially the ability to synthesize wholly unnatural sequence-defined synthetic genetic polymers, the limitation of current enzymes threatens to slow future synthetic biology advances that require precision synthesis of XNA polymers. Thus, an important goal in the field of polymerase engineering is to discover new polymerase variants with biochemical activities that more closely approximate their natural counterparts.

One of the more significant weaknesses uncovered in the current study is the propensity for engineered polymerases to function with broad substrate specificity, or in the case of FANA, activities that are already inherent to naturally occurring DNA polymerases. This finding implies that most

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

polymerase engineering technologies have succeeded by converting highly efficient DNA polymerase "specialists" into weakly active XNA polymerase "generalists" that exhibit activity for their cognate XNA substrate while retaining promiscuous activity for substrates that derive from other classes of nucleic acid molecules. Determining the reasons for low substrate specificity among laboratory-evolved XNA polymerases could provide insight into important questions in polymerase engineering. Some of the more critical questions include, what is the best approach for identifying the determinants of substrate specificity? To what extent can directed evolution remodel the enzyme active site for specific xNTP recognition? And how does substrate specificity affect enzyme catalysis? Finding the answers to these questions will ultimately lead to improvements in enzyme engineering capabilities that will in turn expand the growing field of synthetic genetics to the broader scientific community by providing enzymes that are more efficient and reliable for downstream users wishing to advance the goals of synthetic genetics.

Given that polymerases follow a complicated, largely unknown catalytic pathway, directed evolution will undoubtedly play an important role in identifying amino acid residues that confer increased activity and substrate specificity on the polymerase scaffold. The challenge, however, is to determine which library designs are best suited to address this problem. One approach would be to systematically explore each amino acid position in the polymerase scaffold by programmed allelic mutagenesis, a method that combines deep mutational scanning with next-generation high-throughput DNA sequencing.[32] From such analyses, second-generation libraries could then be designed that explore beneficial sites in various combinations by saturation mutagenesis. This approach is similar to the path taken to identify the TNA polymerase Kod-RSGA.[21] It is also representative of the approach that was originally used to identify the Therminator mutation found in all known XNA polymerases and thought to be a key determinant of substrate specificity.[33] Another approach is to explore new regions of sequence space through the recombination of homologous enzyme scaffolds.[34] Shuffling approaches have proven effective at addressing problems involving substrate specificity, such as the conversion of a galactosidase enzyme into a fucosidase.[35] By analogy, it would seem that a similar approach could be beneficial for fine-tuning the substrate specificity of XNA polymerases.

Coupled to the need for better polymerase libraries is an equally important need for more structural information on current and future XNA polymerases. At present, there exists a paucity of structural information available in the protein databank on laboratory-evolved polymerases that have been developed to copy genetic information back and forth between DNA and XNA. In an effort to address this problem, we have previously reported five X-ray crystal structures for an engineered TNA polymerase that capture the process of template recognition, substrate binding, catalysis, and translocation.[36] More recently, Delarue and colleagues have reported the apo and post-translocated structures for the HNA polymerase Tgo-6G12.[37] Although these structures provide new insights into the mechanism of XNA synthesis, more structures are needed to understand the local and global changes responsible for new gain-of-function activities as well as the underlying biochemical role of individual mutations. Here, it will be particularly important to correlate biochemical information with structural interpretations on how adaptive mutations enhance substrate specificity while eliminating the harmful effects of untemplated XNA synthesis.

In summary, we have compared the substrate specificity, thermal stability, reverse transcriptase activity, and fidelity of laboratory-evolved polymerases that were established to synthesize RNA and five different examples of XNA polymers. Together, these data benchmark the activity of current XNA polymerases and provide opportunities for generating new polymerase variants that function with greater activity and substrate specificity.

MATERIALS AND METHODS

Reagents. DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa). Threose nucleic acid (TNA) triphosphates were obtained by chemical synthesis as described previously.[38,39] 2′-Fluoroarabino nucleic acid (FANA) and arabinonucleic acid (ANA) triphosphates were purchased from TriLink Biotechnologies (San Diego, CA). 1,5-Anhydrohexitol nucleic acid (HNA) and 3′-2′ phosphonomethyl-threosyl nucleic acid (PMT) triphosphates were obtained by chemical synthesis as previously described.[23,40] ThermoPol buffer, Q5 site-directed mutagenesis kit, KLD, Gibson assembly were purchased from New England Biolabs (Ipswich, MA). A DNA clean-up kit was purchased from Zymo (Irvine, CA). Ammonium persulfate (APS) was purchased from Sigma-Aldrich (St. Louis, Missouri). A TOPO-TA cloning kit, deoxynucleotide triphosphates (dNTPs), ethylenediaminetetraacetic acid (EDTA), urea, acrylamide, and bis-acrylamide were purchased from Thermofisher Scientific (Waltham, Massachusetts). Tetramethyl-ethylenediamine (TEMED) was purchased from Bio-Rad (Hercules, California). Heparin affinity columns were purchased from GE Healthcare (Little Chalfont, United Kingdom).

Plasmid Synthesis. The DNA sequence (2319 bp) for *Thermococcus gorgonarius* wild-type (Tgo-WT) exo$^-$ DNA polymerase was obtained from GenBank (KP682507.1). The Tgo-WT exo$^-$ construct encodes 773 amino acids, including the exonuclease silencing mutations (D141A and E143A). For simplicity, the Tgo-V93Q/A485L variant was used as the starting point for constructing the XNA polymerase variants, as these mutations are found in each of the engineered polymerases. Tgo-V93Q/A485L was generated by site-directed mutagenesis (SDM) from Tgo-WT exo$^-$. The DNA primers for SDM were designed to contain mismatched nucleotides near the center of the forward primer to prevent annealing error. The reverse primer was designed to be flush to the forward primer, creating a blunt end. PCR (50 μL reaction volume) was performed using Q5 DNA polymerase with the following cycling conditions: melting at 95 °C for 2 min, 30× amplification cycle (95 °C for 30 s, 65 °C for 45 s, 72 °C for 5 min), and a final extension of 72 °C for 3 min. PCR reactions were analyzed by 1% agarose gel electrophoresis with ethidium bromide staining to confirm amplicon length and purity. Upon confirmation, 1 μL of the PCR product was treated with kinase, ligase, and Dpn1 (NEB KLD) for 20 min at room temperature per the manufacturer's protocol. KLD-treated samples were transformed into DH5-α *E. coli* per the manufacturer's protocol and plated onto 100 ng/μL of ampicillin-containing agar plates and grown overnight at 37 °C. Single colonies were picked from the plate, grown in liquid LB media, miniprepped, and sent out for Sanger sequencing. Trace alignments from Sanger sequencing were performed

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

using CLC Main software (Qiagen). PCR conditions and all subsequent steps were performed the same, unless clearly stated. Tgo-6G12 was generated by Gibson assembly with a gBlock containing the mutations: V589A, E609 K, I610M, K659Q, E664Q, Q665P, R668 K, D669Q, K671H, K674R, T676R, A681S, L704P, and E730G. The gBlock was designed to contain overlaps with the pGDR11 Tgo(exo-)-V93Q/A485L plasmid. Primers for the plasmid were designed with a 15−40 nucleotide overlap in length and with a melting temperature greater than 48 °C for Gibson assembly. The annealing step for PCR was changed to a temperature gradient ranging from 58 to 73 °C. PCR reactions were analyzed by agarose gel electrophoresis with ethidium bromide staining. The reactions were pooled and cleaned with a Zymo DNA clean and concentrator kit per manufacturer's protocol with two additional wash steps. Gibson assembly was performed with purified DNA and gBlock. Upon sequence verification of Tgo-6G12, Tgo-6G12/I521L was generated from the Tgo-6G12 plasmid by SDM at position I521L. Tgo-EPFLH was generated from the pGDR11 Tgo (exo-)-V93Q plasmid by SDM at H147E, I521L, E664H and Gibson assembly with a gBlock, containing mutations: L403P, L408F, and A485L. Tgo-QGLK and Tgo-D4K were generated as previously described.[41]

Recombinant Polymerase Expression and Purification. Expression and purification protocols were performed as previously described.[24] In brief, XL1-Blue *E. coli* strain carrying a pGDR11 vector encoding the polymerase of interest was grown and expressed in liquid LB media with carbenicillin. Cells were lysed and centrifuged, and the supernatant was then collected and treated with 10% (v/v) PEI for 15 min. Treated lysate was subsequently centrifuged (20 000 rpm, 4 °C, 20 min) to remove precipitated nucleic acids. Ammonium sulfate precipitation was performed, and the protein pellet was then resuspended in equilibration buffer (10 mM Tris pH 8.0, 50 mM NaCl, 10% glycerol). For primer-extension assays using lysate, the resuspended protein pellet was enough to perform the assay. However, most polymerases were purified on a heparin affinity column using a concentration gradient (5, 25, 50, 75, and 100% buffer B) made by mixing buffer A [10 mM Tris, pH 8.0, 50 mM NaCl, 10% glycerol] and buffer B [10 mM Tris, pH 8.0, 1 M NaCl, 10% glycerol] using an automated ÄKTA FPLC protein purification system.

Primer-Extension Assay. Primer-extension assays were performed in 20 μL reaction volumes containing 500 nM IR-680-labeled DNA primer annealed to an equivalent amount of DNA template (or library), 500 nM polymerase, 1× 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] and a 100 μM concentration of each xNTP. The DNA primer was annealed to the template by heating in 1× ThermoPol buffer for 5 min at 95 °C and stepwise cooled for 1 min at 70 °C, 1 min at 55 °C, 1 min at 35 °C, and 1 min at 4 °C. The reactions were initiated by adding the xNTPs and polymerase and incubating the solution for 1 h at 55 °C. DNA and FANA synthesis reactions were performed for 15 min, and XNA library reactions were incubated for 2 h. The reactions were quenched with stop buffer [95% formamide, 25 mM EDTA, pH 8.0] and analyzed by denaturing polyacrylamide gel electrophoresis with gel imaging performed using a LI-COR Odyssey CLx imager.

Thermostability Assay. Aliquots of purified polymerase (100 μL volumes, 10 μM polymerase) were heated for 1−6 h in storage buffer [10 mM Tris, pH 8.0, 500 mM NaCl, 10% glycerol] on a heat block at 90 °C. To prevent evaporation, the solutions were topped with mineral oil (40 μL volumes). At specified times, aliquots (10 μL volumes) were removed and kept at 24 °C until the 6 h time point was collected. Immediately following the thermal challenge, the polymerases were assayed for activity in a DNA synthesis reaction as described above.

Fidelity Assay. Fidelity measurements were performed using a DNA primer (1 μM PBS8_extra) containing a double-nucleotide mismatch (TT-TT) was annealed to the DNA template to form the primer−template duplex (1 μM). The primer-extension reaction was performed in 1× ThermoPol buffer [20 mM Tris-HCl, 10 mM $(NH4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8, NEB] supplemented with either hNTPs or aNTPs (100 μM of each xNTP) and the corresponding HNA or ANA polymerase. The reactions were incubated for 3 h at 55 °C. The fully extended product was purified by denaturing PAGE and reverse-transcribed into cDNA using 1 μM Bst-BF DNA polymerase and a corresponding DNA primer. The cDNA was amplified by PCR, ligated into a TOPO vector, and cloned into *E. coli* DH5α competent cells. Individual colonies were grown in liquid media and sequenced using the M13R primer by Retrogen, San Diego, CA. DNA sequences were aligned with the template used and analyzed using MEGA7 software.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00048.

Methods, Table S1, and Figures S1−S11 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

Authors

Esau Medina − *Departments of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Eric J. Yik − *Departments of Pharmaceutical Sciences, University of California, Irvine, California 92697-3958, United States*

Piet Herdewijn − *KU Leuven, Rega Institute for Medical Research, 3000 Leuven, Belgium;* orcid.org/0000-0003-3589-8503

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.1c00048

Author Contributions

⊥E.M. and E.J.Y. are cofirst authors.

Author Contributions

J.C.C. and P.H. conceived of the project and designed the experiments. E.M. and E.J.Y. performed all of the experiments. E.M., E.J.Y., and J.C.C. wrote the manuscript. All authors reviewed and commented on the manuscript.

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We wish to thank members of the Chaput lab for their helpful comments on the manuscript. This work was supported by a grant from the National Science Foundation (MCB: 1946312) to J.C.C. and a grant from the Belgian FWO (G085321N) to P.H. E.M. was supported by the NIH IMSD training grant (GM055246). We wish to thank C. Liu and M. Abramov for assistance with the chemical synthesis of PMT and HNA triphosphates.

■ REFERENCES

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

(2) Dunn, M. R., and Chaput, J. C. (2016) Reverse transcription of threose nucleic acid by a naturally occurring DNA polymerase. *ChemBioChem 17*, 1804−1808.

(3) Houlihan, G., Arangundy-Franklin, S., Porebski, B. T., Subramanian, N., Taylor, A. I., and Holliger, P. (2020) Discovery and evolution of RNA and XNA reverse transcriptase function and fidelity. *Nat. Chem. 12*, 683−690.

(4) Chaput, J. C., Yu, H., and Zhang, S. (2012) The emerging world of synthetic genetics. *Chem. Biol. 19*, 1360−1371.

(5) Taylor, A. I., Houlihan, G., and Holliger, P. (2019) Beyond DNA and RNA: The Expanding Toolbox of Synthetic Genetics. *Cold Spring Harbor Perspect. Biol. 11*, a032490.

(6) 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., and Holliger, P. (2012) Synthetic genetic polymers capable of heredity and evolution. *Science 336*, 341−344.

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

(8) Taylor, A. I., Pinheiro, V. B., Smola, M. J., Morgunov, A. S., Peak-Chew, S., Cozens, C., Weeks, K. M., Herdewijn, P., and Holliger, P. (2015) Catalysts from synthetic genetic polymers. *Nature 518*, 427−430.

(9) Alves Ferreira-Bravo, I., Cozens, C., Holliger, P., and DeStefano, J. J. (2015) Selection of 2′-deoxy-2′-fluoroarabinonucleotide (FANA) aptamers that bind HIV-1 reverse transcriptase with picomolar affinity. *Nucleic Acids Res. 43*, 9587−9599.

(10) Wang, Y., Ngor, A. K., Nikoomanzar, A., and Chaput, J. C. (2018) Evolution of a General RNA-Cleaving FANA Enzyme. *Nat. Commun. 9*, 5067.

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

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

(13) Dunn, M. R., McCloskey, C. M., Buckley, P., Rhea, K., and Chaput, J. C. (2020) Generating Biologically Stable TNA Aptamers that Function with High Affinity and Thermal Stability. *J. Am. Chem. Soc. 142*, 7721−7724.

(14) Mei, H., Liao, J.-Y., Jimenez, R. M., Wang, Y., Bala, S., McCloskey, C., Switzer, C., and Chaput, J. C. (2018) Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *J. Am. Chem. Soc. 140*, 5706−5713.

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

(16) Zhang, L., and Chaput, J. C. (2020) In Vitro Selection of an ATP-Binding TNA Aptamer. *Molecules 25*, 4194.

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

(18) Herdewijn, P., and Marliere, P. (2009) Toward safe genetically modified organisms through the chemical diversification of nucleic acids. *Chem. Biodiversity 6*, 791−808.

(19) Pochet, S., Kaminski, P. A., Van Aerschot, A., Herdewijn, P., and Marliere, P. (2003) Replication of hexitol oligonucleotides as a prelude to the propagation of a third type of nucleic acid in vivo. *C. R. Biol. 326*, 1175−84.

(20) Yang, K., McCloskey, C. M., and Chaput, J. C. (2020) Reading and Writing Digital Information in TNA. *ACS Synth. Biol. 9*, 2936−2942.

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

(22) Dunn, M. R., Otto, C., Fenton, K. E., and Chaput, J. C. (2016) Improving Polymerase Activity with Unnatural Substrates by Sampling Mutations in Homologous Protein Architectures. *ACS Chem. Biol. 11*, 1210−9.

(23) Liu, C., Cozens, C., Jaziri, F., Rozenski, J., Marechal, A., Dumbre, S., Pezo, V., Marliere, P., Pinheiro, V. B., Groaz, E., and Herdewijn, P. (2018) Phosphonomethyl Oligonucleotides as Backbone-Modified Artificial Genetic Polymers. *J. Am. Chem. Soc. 140*, 6690−6699.

(24) Nikoomanzar, A., Dunn, M. R., and Chaput, J. C. (2017) Engineered polymerases with altered substrate specificity: expression and purification. *Curr. Protoc. Nucleic Acid Chem. 69*, 1.

(25) Brown, J. A., and Suo, Z. (2011) Unlocking the sugar ″steric gate″ of DNA polymerases. *Biochemistry 50*, 1135−42.

(26) Anosova, I., Kowal, E. A., Dunn, M. R., Chaput, J. C., Van Horn, W. D., and Egli, M. (2016) The structural diversity of artificial genetic polymers. *Nucleic Acids Res. 44*, 1007−1021.

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

(28) Chaput, J. C., Herdewijn, P., and Hollenstein, M. (2020) Orthogonal Genetic Systems. *ChemBioChem 21*, 1408−1411.

(29) Jackson, L. N., Chim, N., Shi, C., and Chaput, J. C. (2019) Crystal structures of a natural DNA polymerase that functions as an XNA reverse transcriptase. *Nucleic Acids Res. 47*, 6973−6983.

(30) Chen, T., and Romesberg, F. E. (2014) Directed polymerase evolution. *FEBS Lett. 588*, 219−229.

(31) Houlihan, G., Arangundy-Franklin, S., and Holliger, P. (2017) Engineering and application of polymerases for synthetic genetics. *Curr. Opin. Biotechnol. 48*, 168−179.

(32) Araya, C. L., and Fowler, D. M. (2011) Deep mutational scanning: assessing protein function on a massive scale. *Trends Biotechnol. 29*, 435−442.

(33) 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., and Jack, W. E. (2019) Therminator DNA Polymerase: Modified Nucleotides and Unnatural Substrates. *Front. Mol. Biosci. 6*, 28.

(34) Crameri, A., Raillard, S.-A., Bermudez, E., and Stemmer, W. P. C. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature 391*, 288−291.

(35) Zhang, J. H., Dawes, G., and Stemmer, W. P. C. (1997) Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. U. S. A. 94*, 4504−4509.

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

(37) Samson, C., Legrand, P., Tekpinar, M., Rozenski, J., Abramov, M., Holliger, P., Pinheiro, V. B., Herdewijn, P., and Delarue, M. (2020) Structural Studies of HNA Substrate Specificity in Mutants of

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437

an Archaeal DNA Polymerase Obtained by Directed Evolution. *Biomolecules 10*, 1647.
(38) Sau, S. P., Fahmi, N. E., Liao, J.-Y., Bala, S., and Chaput, J. C. (2016) A scalable synthesis of α-L-threose nucleic acid monomers. *J. Org. Chem. 81*, 2302−2307.
(39) Liao, J.-Y., Bala, S., Ngor, A. K., Yik, E. J., and Chaput, J. C. (2019) P(V) Reagents for the Scalable Synthesis of Natural and Modified Nucleoside Triphosphates. *J. Am. Chem. Soc. 141*, 13286−13289.
(40) Vastmans, K., Pochet, S., Peys, A., Kerremans, L., Van Aerschot, A., Hendrix, C., Marliere, P., and Herdewijn, P. (2000) Enzymatic incorporation in DNA of 1,5-anhydrohexitol nucleotides. *Biochemistry 39*, 12757−65.
(41) Nikoomanzar, A., Dunn, M. R., and Chaput, J. C. (2017) Evaluating the rate and substrate specificity of laboratory evolved XNA polymerases. *Anal. Chem. 89*, 12622−12625.

https://doi.org/10.1021/acssynbio.1c00048
ACS Synth. Biol. 2021, 10, 1429−1437