Supplementary Information

Evaluating the Rate and Substrate Specificity of Laboratory Evolved XNA Polymerases

Ali Nikoomanzar, Matthew R. Dunn, and John C. Chaput*

Departments of Pharmaceutical Sciences, Chemistry, and
Molecular Biology and Biochemistry
University of California, Irvine, CA 92697-3958. United States

This PDF file includes:

Methods

General information

DNA and RNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa). FANA and TNA oligonucleotides were synthesized on an ABI3400 DNA synthesizer using chemical synthesis reagents purchased from Glen Research (Sterling, Virginia). FANA phosphoramidites were purchased from Glen Research (Sterling, Virginia). FANA triphosphates were purchased from Metkinen Chemistry (Kuopio, Finland). TNA triphosphates and phosphoramidites were obtained by chemical synthesis as described previously[1,2]. Oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis, electroeluted, ethanol precipitated, and quantified by UV absorbance. Thermopol buffer was purchased from New England Biolabs (Ipswitch, MA). All chemical reagents, including NTPs and dNTPs were purchased from Sigma Aldrich (St. Louis, Missouri). Fluorescent dyes were obtained from the following vendors: SYTO 9 and SYBR Green (Thermo Fisher Scientific, Waltham, Massachusetts), LC Green Plus (BioFire Defense, Salt Lake City, Utah), Eva Green (Biotium, Fremont, California), and BOXTO (TATAA Biocenter, Gothenburg, Sweden). CleanAmp dNTP mix was purchased from TriLink Biotechnologies (San Diego, California). PCR plates and optical film were purchased from Applied Biosystems (Foster City, California). Nickel and heparin affinity columns were purchased from GE Healthcare (Little Chalfont, United Kingdom).

Polymerase expression and purification

Engineered polymerases were expressed and purified as previously described[3]. Purified enzymes are shown in Supplementary Figure 1. Briefly, XL1-blue *E. coli* cells carrying custom pGDR11 polymerase expression plasmids were streaked onto LB-ampicillin (100 µg/mL) agar plates and grown overnight at 37°C. A single isolate was used to inoculate 50 mL of LB-ampicillin (100 µg/mL) liquid medium. The starter culture was grown to confluency overnight at 37°C with shaking at 225 rpm. The starter culture (1:100 v/v) was used to inoculate 1 L of LB-ampicillin (100 µg/mL) liquid medium and grown at 37°C with shaking at 225 rpm. At OD_{600} = 0.6, the expression culture was removed and cooled to 15°C. The culture was then induced with IPTG at a final concentration of 0.5 mM and incubated overnight at 15°C with shaking at 225 rpm. Cells were harvested by centrifugation for 10 min at 5,000 rpm and 4°C. The cell

pellet was resuspended in 40 mL buffer [10 mM Tris pH 8.0, 500 mM KCl, 10% glycerol] and lysed by sonication on ice following a 2 sec on and 2 sec off protocol for 2 min and repeated once after the lysate cooled for 5 min. Native *E. coli* proteins were denatured by heating the sample for 60 min at 80°C and then immediately cooling the lysate for 30 min on ice at 4°C. The lysate was clarified by centrifugation for 10 min at 20,000 rpm and 4°C. Nucleic acids were precipitated by adding 10% (v/v) polyethyleneimine (PEI) to a final concentration of 0.5%, mixing by inversion, incubating for 30 min on ice at 4°C, and centrifuging for 30 min at 20,000 rpm and 4°C. The supernatant was transferred to a fresh centrifuge tube and PEI was removed and the target polymerase precipitated by adding 60% (w/v) ammonium sulfate, mixing by inversion, incubating for 30 min at 4°C, and then centrifuging for 30 min at 20,000 rpm and 4°C. Protein pellets were suspended in 4°C buffer [10 mM tris pH 8.0, 50 mM KCl, 10% glycerol]. Particulate was removed by centrifuging for 10 min at 20,000 rpm and 4°C. Recombinant polymerases were purified by heparin affinity chromatography with step elutions of 100, 250, 500, and 1000 mM KCl. Fractions corresponding to protein of the correct size were verified by SDS PAGE, combined, quantified by UV absorbance at 280 nm, and stored at 4°C. Purified proteins were verified by SDS PAGE (see Figure S1).

Screening fluorescent dyes for a dose-dependent linear response

Five high-resolution melting dyes (SYTO 9, Eva Green, SYBR Green I, BOXTO, and LC Green Plus) were screened for a linear dose-dependent response against a set of max standards (SI Table 1). Each max standard was serially diluted in water to obtain six 10x stock solutions that spanned a concentration range of 25 µM to 780 nM. The dye-binding assay was performed in 96-well format. Each measurement (10 µL) contained 1 µL of max standard, 1 µL of 10x Thermopol buffer [200 mM Tris-HCl, 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 20 mM $MgSO_4$, and 1% Triton X-100, pH 8.8], 2 µL of fluorescent dye (see Tables S2-3), 1 µL of nucleoside triphosphate and 5 µL of nuclease-free water. Each fluorescent dye was evaluated at four different concentrations 0.5x, 1x, 2x, and 4x, where x denotes the manufacturers recommended concentration. In some cases, the manufacturer provides the precise dye concentration while other manufacturers provide the concentration in units of x-fold. Wells that contained the TNA max standard were tested in the presence of 1 mM $MnCl_2$ by substituting 1 µL of water for 1 µL of a 10 mM $MnCl_2$ solution. The plate was sealed with optical film and placed into an Applied Biosystems Quantstudio

6 Flex quantitative real-time PCR instrument. The max standard was denatured by heating for 2 min at 95°C and annealed at 55°C. After equilibration (~60 s), the fluorescence intensity was recorded and the data was fit to a linear equation using Microsoft Excel. The linear range and R^2 values were calculated from the maximum linear region observed for a line plotted between the minimum and maximum duplex concentrations (78 nM to 2.5 uM; 32-fold). The signal-to-noise ratio was obtained from the minimum and maximum fluorescence values observed for the linear region of each plot.

Baseline measurements

Measurements of baseline fluorescence for the min and max standards (SI Table 1) were determined in 96-well format. Each measurement (10 µL) contained 1 µM of either the min or max standard, 1x Thermopol buffer, 100 µM of each nucleoside triphosphate, the corresponding fluorescent dye poised at the optimal concentration (see below), and polymerase buffer minus polymerase [1 mM Tris pH 8.0, 50 mM KCl, 1% glycerol]. Baseline measurements were performed in triplicate by aliquoting 10 µL from a 5x mastermix of the reactions into 3 separate wells and taking the average of all fluorescent values to generate a single minimum and maximum fluorescence baseline.

Steady-state measurements

Steady-state kinetic measurements were performed in 96-well format. Each measurement (10 µL) contained 1 µM of the self-priming 30mer HP template, referred to as the min standard (SI Table 1), 1x Thermopol buffer, 100 µM of each nucleoside triphosphate, the corresponding fluorescent dye poised at the optimal concentration (see below), and the natural or engineered polymerase. The following dyes and dye concentrations were used: DNA/DNA (5 µM Syto 9), DNA/TNA (4x Eva Green), DNA/RNA (2x LC Green Plus), DNA/FANA (2x LC Green Plus). TNA reactions were supplemented with 1mM $MnCl_2$. Polymerase reactions were initiated by adding 1 µL of a 10x polymerase solution. The 10x polymerase solutions were prepared as a 2-fold dilution series in buffer [10 mM Tris pH 8.0, 500 mM KCl, 10% glycerol] that spanned a polymerase concentration range of 100 nM to 6.25 nM. The PCR plate was sealed with optical film and immediately placed inside the real-time PCR instrument. Reactions were denatured for 2 min at 95°C and extended for 1 h at 55°C with fluorescence intensity measurements

collected at 6 s intervals. The first 18 sec of data were excluded to eliminate possible artifacts caused by temperature equilibration. Fluorescence data for each reaction was normalized by subtracting baseline fluorescence and dividing by the difference between min and max fluorescence values. Fluorescence was converted to nucleotides per polymerase using the conversion factor ($F_{max} = C_n*n*C_p^{-1}$, where F_{max} is the maximum fluorescence for the max standard, C_n is the concentration of the nucleic acid, n is the number of incorporated nucleotides, and C_P is the concentration of polymerase). Nucleotides per polymerase were plotted over time and the initial velocities were determined by performing a linear regression on the initial data. Steady-state kinetic values were determined by calculating the slope of the linear trajectory obtained for the polymerase concentration that yielded the longest linear range under saturating (>100-fold) primer-template conditions. The reported values derive from a representative plot taken from at least three independent replicates.

Reaction condition optimization

Kinetic measurements were performed in 96-well format to evaluate the extension efficiency of TNA synthesis under a full factorial analysis of reaction conditions. Each measurement (10 µL) contained a final concentration of 1 µM primer-template complex (SI Table 1, L11 library, PBS8 primer), 100 µM tNTPs, 4x Eva Green fluorescent dye, 1mM $MnCl_2$, 1 µM polymerase, and a titrated concentration of KCl [0-21 mM], $MgCl_2$ [0.5-4 mM], or Tris-HCl pH [7.5-9.0]. Reactions were initiated as described previously for the steady-state measurements, denatured for 2 min at 95°C, and extended for 2 h at 50°C, 55°C and 60°C for a total of 288 unique reaction conditions. Fluorescence intensity measurements were collected at 15 s intervals. The first 15 sec of data were excluded to eliminate possible artifacts caused by temperature equilibration. Polymerase activity values were reported as a heat map based on initial velocity and change in fluorescence.

Table S1. Oligonucleotide sequences. Modifications are written following IDT nomenclature. N represents a degenerate position with variable proportions of A, T, G, and C nucleobases. Underlined lower case sequences on the max standard were constructed of DNA, RNA, TNA, and FANA to produce four separate version of the max standard.

Name	**DNA sequence (5' -> 3')**
Min standard (30mer HP)	TCTCTATAGTGAGTCGTATAGGTGGTATCCGAAAGGATACCACC
Max standard	TCTCTATAGTGAGTCGTATAGGTGGTATCCGAAAGGATACCACCtatacgactcactatagaga
L11 library	GGATCGTCAGTGCATTGAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGTGGTATCCCCAAGGGGAC N=A:T:C:G=1:1:0.3:0.3
PBS8 20mer	GTCCCCTTGGGGATACCACC

Table S2. **Dye screening and optimization for DNA:DNA and DNA:TNA duplexes.** Dyes were chosen by selecting for the highest linear-fit (R^2) upon reaching a cut-off of a 32-fold linear range (LR) and a signal-to-noise (SNR) greater than 15-fold. Optimal dye concentrations are highlighted in green.

Duplex	Dye	Concentration	Linear Range	Signal-to-Noise	R^2
DNA:DNA	Eva Green	2x	16	19.11	0.9804
		1x	16	10.26	0.9412
		0.5x	4	4.12	0.9772
		0.25x	-	-	-
	LC Green Plus	2x	32	32.86	0.9903
		1x	32	45.67	0.9853
		0.5x	32	16.03	0.9169
		0.25x	32	13.13	0.9310
	SYBR Green I	2x	32	114.58	0.9841
		1x	16	36.46	0.9758
		0.5x	8	5.65	0.9999
		0.25x	-	-	-
	BOXTO	10 µM	32	33.97	0.9743
		5 µM	16	13.65	0.9749
		2.5 µM	-	-	-
		1.25 µM	-	-	-
	SYTO 9	10 µM	32	31.92	0.9923
		5 µM	32	40.99	0.9970
		2.5 µM	32	17.96	0.9770
		1.25 µM	32	13.54	0.9689
DNA:TNA	Eva Green	4x	32	32.52	0.9921
		2x	32	24.64	0.9918
		1x	32	17.24	0.9437
		0.5x	32	9	0.9463
	LC Green Plus	4x	32	23.88	0.9795
		2x	32	27.04	0.9945
		1x	32	17.28	0.9793
		0.5x	32	16.37	0.9990
	SYBR Green I	4x	32	11.74	0.9721
		2x	8	4.54	0.9869
		1x	16	5.2	0.9407
		0.5x	32	6.31	0.9203
	BOXTO	20 µM	4	2.74	0.8168
		10 µM	8	5.1	0.8920
		5 µM	8	2.63	0.9098
		2.5 µM	8	2.48	0.9771
	SYTO 9	20 µM	32	8.53	0.9793
		10 µM	32	5.68	0.9517
		5 µM	32	4.27	0.9228
		2.5 µM	32	4.29	0.9497

Table S3. **Dye screening and optimization for DNA:RNA and DNA:FANA duplexes.** Dyes were chosen by selecting for the highest linear-fit (R^2) upon reaching a cut-off of a 32-fold linear range (LR) and a signal-to-noise (SNR) greater than 15-fold. Optimal dye concentrations are highlighted in green.

Duplex	Dye	Concentration	Linear Range	Signal-to-Noise	R^2
DNA:RNA	Eva Green	4x	8	5.65	0.9880
		2x	8	5.57	0.9868
		1x	4	3	0.9854
		0.5x	4	2.68	0.9620
	LC Green Plus	2x	32	20.86	0.9975
		1x	32	17.89	0.9778
		0.5x	32	17.98	0.9882
		0.25x	32	18.85	0.9803
	SYBR Green I	4x	8	6.17	0.9840
		2x	4	3.35	0.9941
		1x	4	2.69	0.9766
		0.5x	2	1.37	0.9350
	BOXTO	20 μM	16	3.16	0.9051
		10 μM	32	7.03	0.8256
		5 μM	32	8.72	0.8683
		2.5 μM	8	4.9	0.9530
	SYTO 9	20 μM	32	16.15	0.9983
		10 μM	32	15.06	0.9946
		5 μM	32	12.55	0.9864
		2.5 μM	32	15.29	0.9900
DNA:FANA	Eva Green	4x	8	6.45	0.9922
		2x	4	3.84	0.9973
		1x	4	2.83	0.9811
		0.5x	2	1.62	0.9821
	LC Green Plus	2x	32	18.34	0.9920
		1x	32	14.4	0.9720
		0.5x	32	13.12	0.9508
		0.25x	32	11.15	0.9454
	SYBR Green I	4x	4	3.2	0.9939
		2x	2	1.7	0.9899
		1x	2	1.16	0.8495
		0.5x	2	1.19	0.8653
	BOXTO	20 μM	32	14.18	0.9433
		10 μM	32	2.25	0.8456
		5 μM	32	3.05	0.7939
		2.5 μM	2	1.22	0.8820
	SYTO 9	20 μM	32	15.99	0.9881
		10 μM	32	14.16	0.9780
		5 μM	32	14	0.9782
		2.5 μM	32	12.87	0.9739

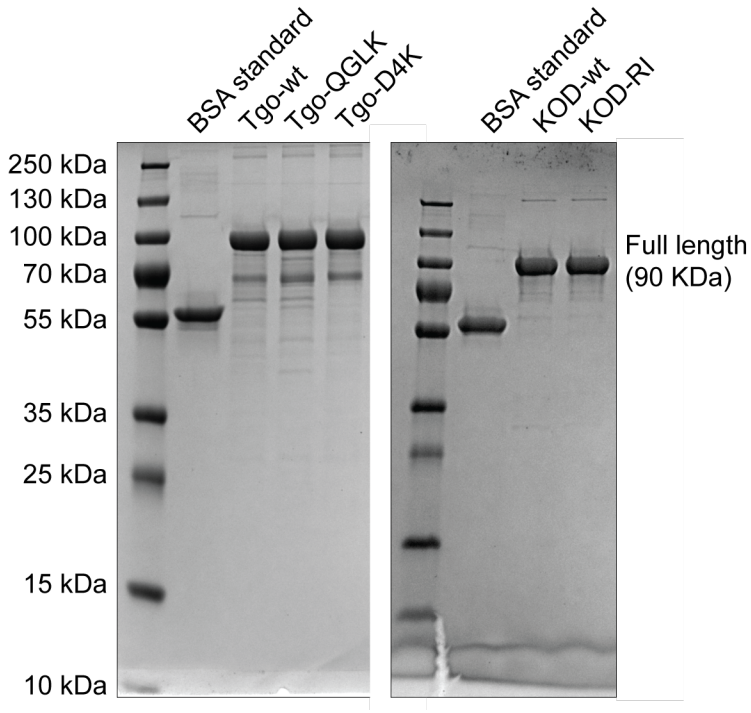

Figure S1. Validation of purity and concentration of purified polymerases. Polymerases were expressed in XL1-blue *E. coli* cells and purified by heparin affinity chromatography. The concentration of each polymerase was normalized by UV_{280} absorbance. Polymerase concentration and purity was further analyzed by SDS-PAGE.

References:

(1) Sau, S. P.; Fahmi, N. E.; Liao, J.-Y.; Bala, S.; Chaput, J. C. *J. Org. Chem.* **2016**, *81*, 2302-2307.

(2) Sau, S. P.; Chaput, J. C. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 3271-3273.

(3) Nikoomanzar, A.; Dunn, M. R.; Chaput, J. C. *Curr. Protoc. Nucleic Acid Chem.* **2017**, *69*, 4.75.