Supporting Information for:

Nucleotide selection by the Y-family DNA polymerase Dpo4 involves template translocation and misalignment

Alfonso Brenlla,[1] Radoslaw P. Markiewicz,[1]

David Rueda,[*,1,2,3] and Louis J. Romano*[*,1]*

[1]*Department of Chemistry, Wayne State University, Detroit, MI 48202*

[2]*Department of Medicine, Section of Virology, Imperial College London, London W12 0NN, UK*

[3]*Single Molecule Imaging Group, MRC Clinical Sciences Center, Imperial College London, London W12 0NN, UK*

*To whom correspondence should be addressed.
E-mail: ljr@chem.wayne.edu or david.rueda@imperial.ac.uk
Phone: +1 313 577 2584 (L.J.R.) or +44 20 8383 1604 (D.R.).

Supplementary Table S1

DNA sequences used in this work.

Experiment	Name[1]	Sequence[2]
Single molecule	7Cy3 primer	5′-[Biotin]-GAT GAA AAT GAA-3′
	8Cy3 primer	5′-[Biotin]-GAT GAA AAT GAA G-3′
	9Cy3 primer	5′-[Biotin]-GAT GAA AAT GAA GG-3′
	10Cy3 primer	5′-[Biotin]-GAT GAA AAT GAA GGA-3′
	11Cy3 primer	5′-[Biotin]-GAT GAA AAT GAA GGA T-3′
	12Cy3 primer	5′-[Biotin]-GAT GAA AAT GAA GGA TA-3′
	Cy3-Template	3′-CTA CTT TTA CTT CCT ATC GAT AAT ACT C-5′
Extension assays	15-mer primer	5′-[Cy3]-GAT GAA AAT GAA GGA -3′
	28-mer short template	3′-CTA CTT TTA CTT CCT ATC GAT AAT ACT C-5′
	24-mer primer	5′-[Cy3]-GGA TTT GGA TGA AGG TGA AGC ATG
	33-mer long template	3′-CCT AAA CCT ACT TCC ACT TCG TAC CTA TAA TAC-5′

[1] Number indicates the number of nucleotides between primer terminus and Cy3 linked nucleotide.

[2] Cy3 is conjugated to blue, underlined thymine.

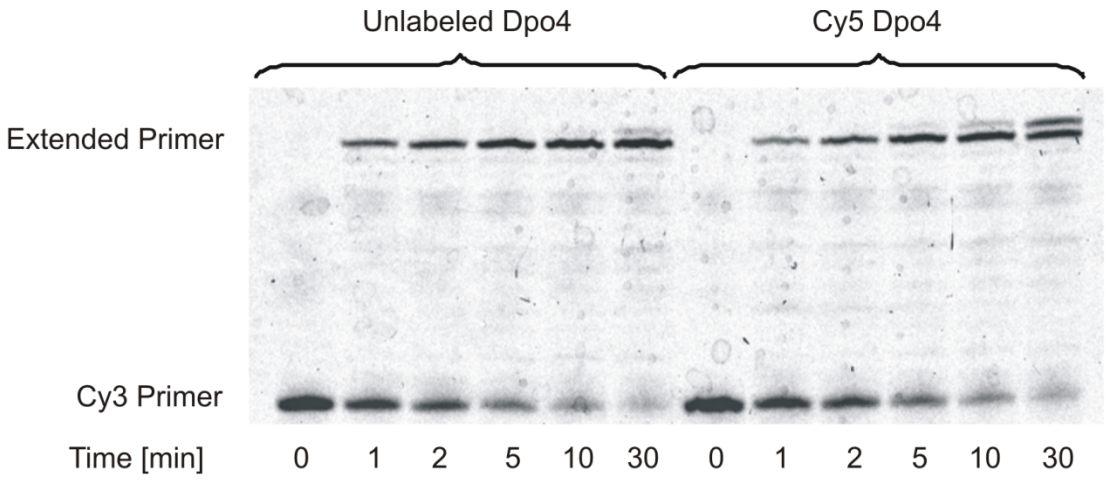


Figure S1. Cy5-labeled Dpo4 presents similar level of activity compared to unlabeled polymerase. Extension assays were carried out at 37^{o} C in buffer containing 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 100 µM dNTPs, 0.025 mg/mL BSA and 5 nM primer-template (15mer/28mer). Reactions were initiated by the addition of polymerase (final Dpo4 concentration 10 nM) and manually stopped with an equal volume of loading buffer (10 mM EDTA, 1 mg/mL bromophenol blue, in 10 mL of formamide). The samples were run on a 20 % denaturing polyacrylamide gel for ~16 hours at 800V. The gel was scanned on a Typhoon 9210 Variable Mode Imager (GE Healthcare).

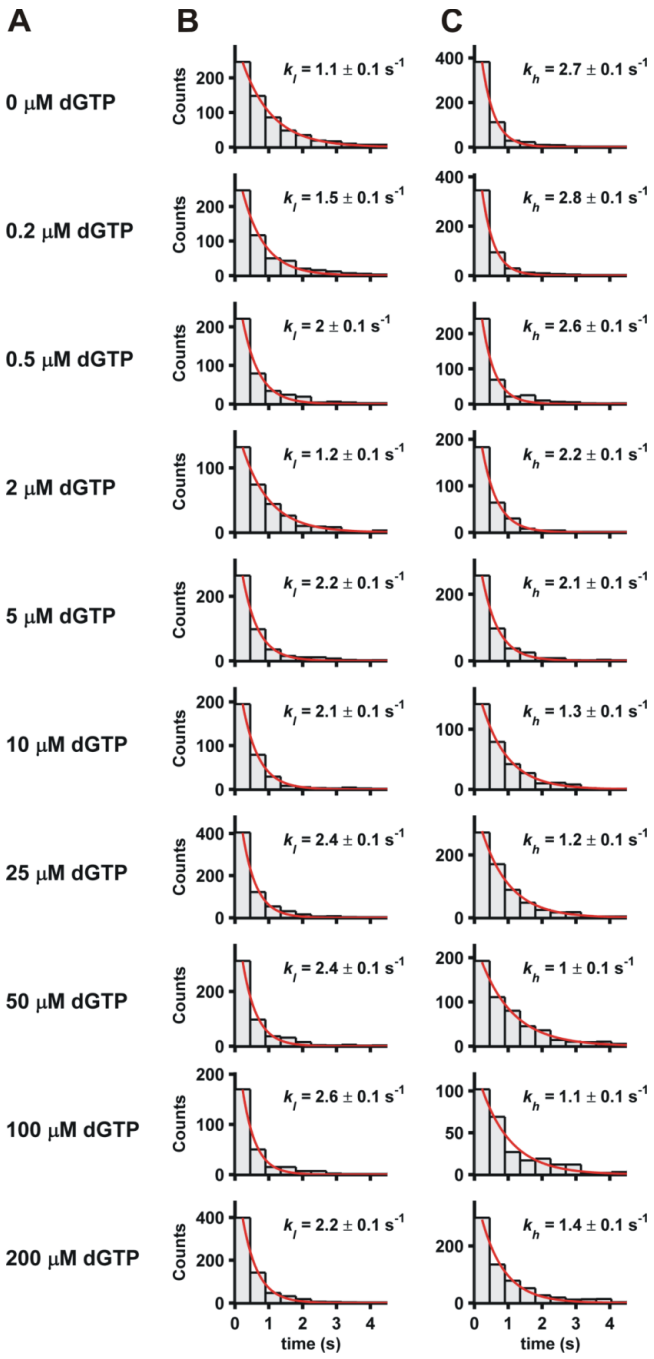


Figure S2. Dwell time distributions for t_h (see Figure 4A) and t_l were fit by monoexponential decay functions. *(A)* Concentration of dGTP used in each experiment. *(B)* Dwell time distribution for t_h and the fit to a monoexponential function. Kinetic rate constant k_l is shown on each graph. *(C)* Dwell time distribution for t_l and the fit to a monoexponential function. Kinetic rate constant k_h is shown on each graph.

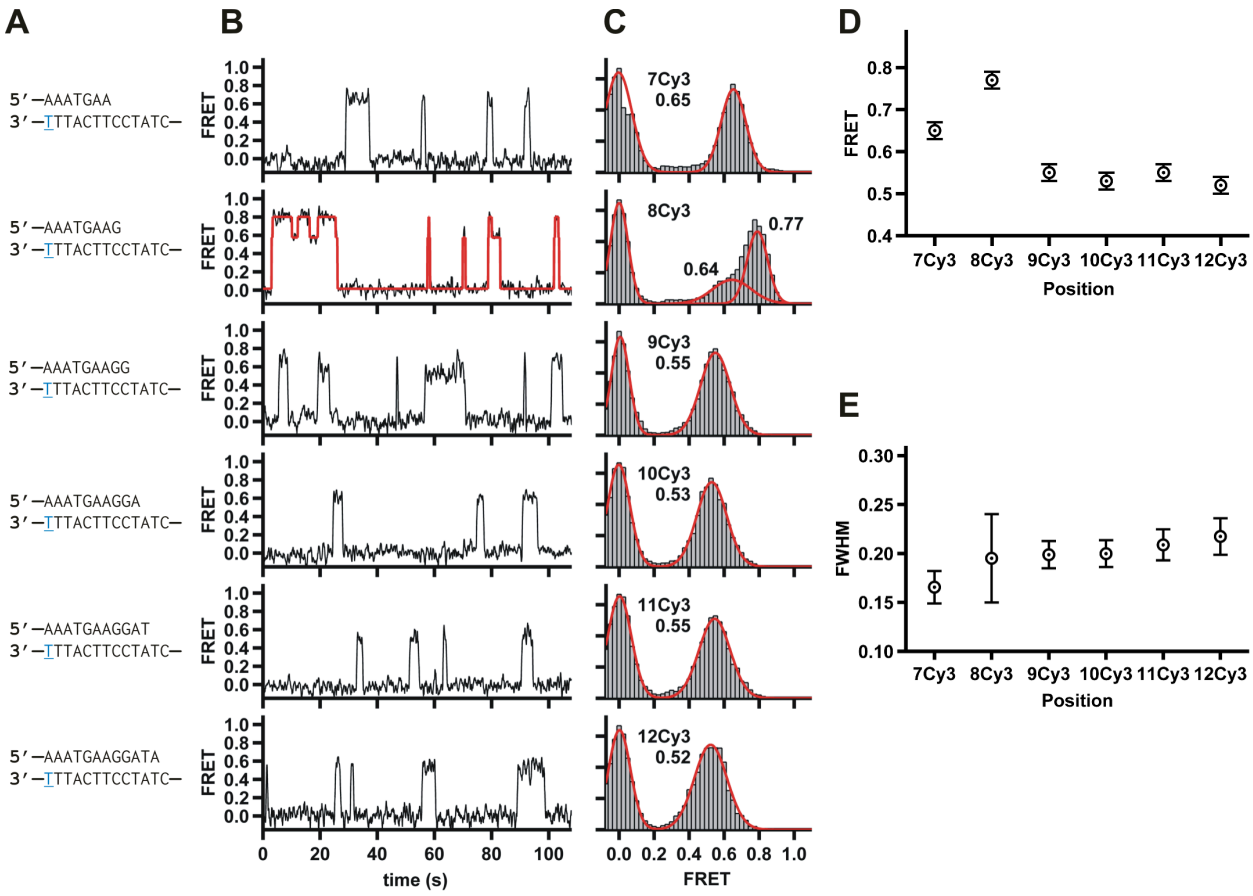


Figure S3. Dpo4 binding different DNA constructs in the presence of Mg. Experiments were carried out in the same buffer conditions as the experiments shown in **Figure 2**, except $CaCl_2$ was replaced with $MgCl_2$. *(A)* The sequence of each primer-template is shown. The same template was used in all DNA constructs, while the lengths of the DNA primers varied. The position of the conjugated Cy3 dye is marked with a blue, underlined thymine. *(B)* Characteristic FRET traces for Dpo4 binding to each duplex. The 8Cy3 trace shown was fit by a Hidden Markov model (red line). *(C)* SmFRET histograms for Dpo4 binding to each duplex with the Gaussian fits shown in red. 8Cy3 binary complex data was fit with two Gaussian functions peaking at 0.64 and 0.77. *(D)* SmFRET efficiencies from (C) plotted as a function of the number of nucleotides between the Cy3 and the primer-template terminus. *(E)* Full width at half maximum (FWHM) plotted as a function of the number of nucleotides between the Cy3 and the primer-template terminus. The FWHM for 8Cy3 was calculated by fitting the 8Cy3 binary complex data to a single Gaussian function. Errors in FRET are estimated to be ±0.02 and the errors for FWHM are the errors of the fit.

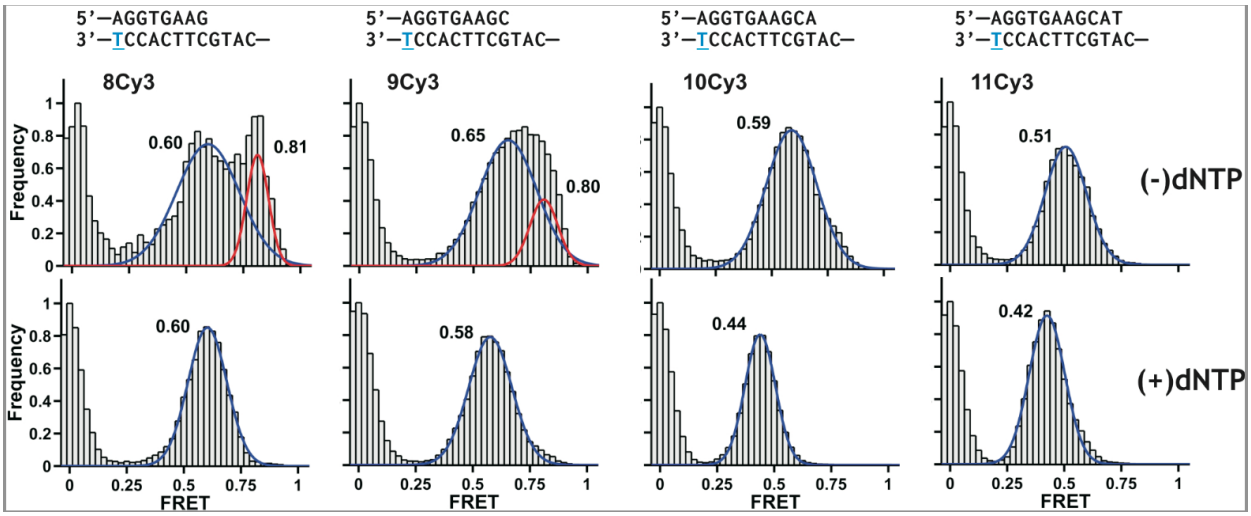


Figure S4. Dpo4 binding to a longer template with different sequence. The structure of the DNA duplex is shown on top. The Cy3 is conjugated to the underlined blue thymine. The top graphs represent the FRET distributions for the binary complex and their fits to two Gaussian functions (8Cy3 and 9Cy3) or one Gaussian function (10Cy3 and 11Cy3). The FRET distributions in the presence of the next correct nucleotide were fit by single Gaussian functions (bottom graphs).

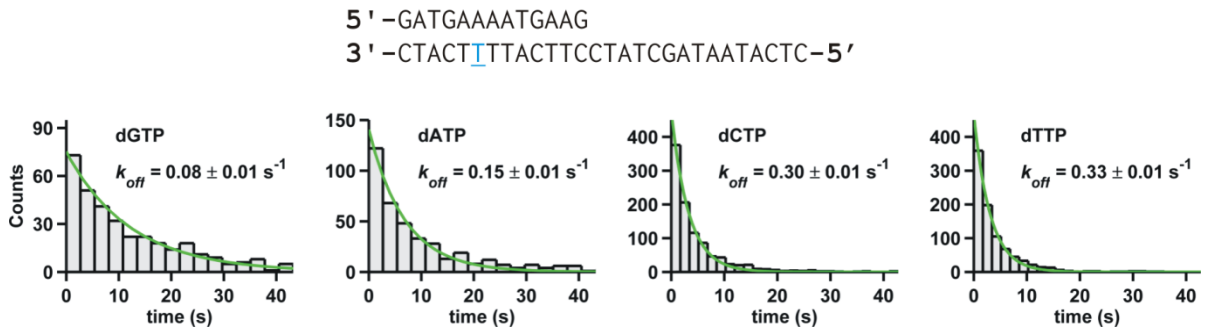


Figure S5. Incorrect nucleotides dCTP and dTTP destabilize the ternary complex, whereas dATP and dGTP induce a small stabilization of the ternary complex. Dwell time distribution (t_{on}, see Figure 1B) for the ternary complex with all four possible dNTPs, indicated at each graph. The concentration of dGTP (correct nucleotide) was 200 µM and the concentration of incorrect dNTPs was 1 mM. k_{off} was obtained from a fit to a monoexponential decay function. The primer-template sequence is shown above the plots.

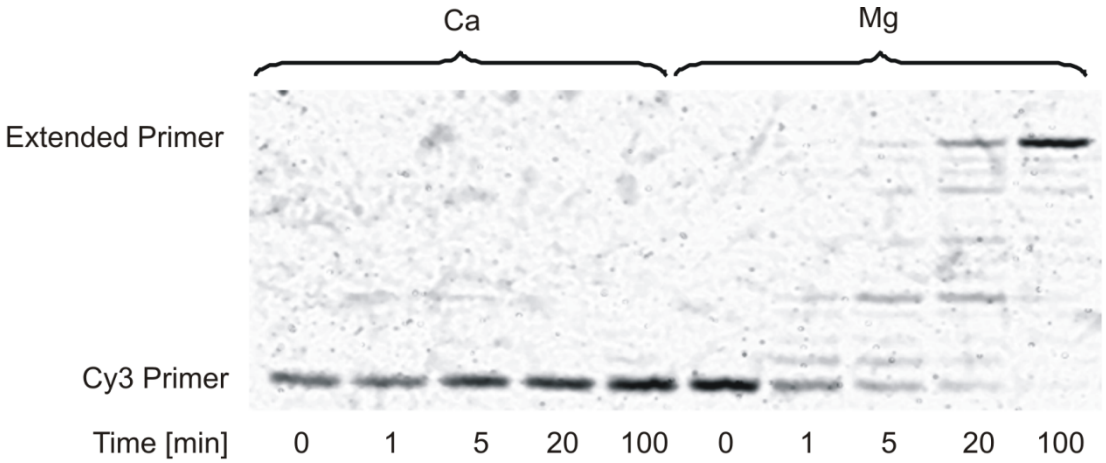
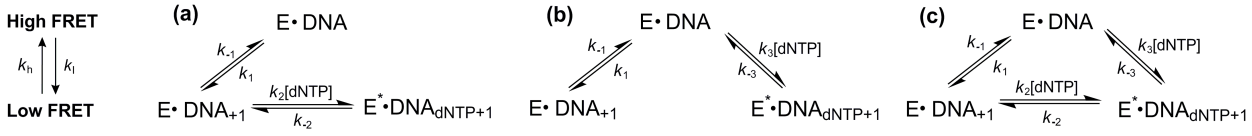


Figure S6. Ca inhibits Dpo4 activity. Reactions were carried out at 25^{o} C in buffer containing 50 mM Tris-HCl, pH 7.5, 3.5 mM $CaCl_2$ (left) or $MgCl_2$ (right), 100 µM dNTPs, 0.025 mg/mL BSA and 5 nM primer-template (15mer/28mer). Upon addition of polymerase (final concentration of Dpo4 10 nM), reactions were incubated for the times shown in the gel. Reactions were quenched and the gel scanned as described in **Figure S1**.

Dependence of k_h and k_l with dGTP concentration

Transition rates k_h and k_l (see Scheme S1) varied with dGTP concentration. We have identified two different conformations for Dpo4/DNA binary complex, but in principle we do not know which one(s) can bind dNTPs. If we assume only reversible processes, three possible mechanisms for dNTP binding arise. In the first two scenarios, **(a)** and **(b)**, only the translocated binary and ternary complex (insertion complex) respectively are able to bind dNTPs. In the third scenario **(c)**, both complexes can bind dNTPs.



Scheme S1. Three possible models for dNTP binding by the Dpo4/DNA binary complexes. E•DNA corresponds to the preinsertion binary complex with the terminal base pair bound at the active site. E•DNA$_{+1}$ is the insertion binary complex, while E*•DNAdNTP$_{+1}$ corresponds to the insertion ternary complex. For the 8Cy3 construct, the E•DNA$_{+1}$ and E*•DNAdNTP$_{+1}$ complexes FRET values are indistinguishable.

Equation 1 was used to fit the data shown in Figure 4C where K_{dGTP} is the affinity constant for dGTP, k_0 is the interconversion rate in the absence of dGTP and k_∞ is the interconversion rate at saturating concentrations of dGTP.

$$k_{obs} = k_0 + \left(k_\infty - k_0\right)\frac{[dGTP]}{K_d + [dGTP]} \qquad k_{obs} = k_0 + \left(k_\infty - k_0\right)\frac{[dGTP]}{K_{dGTP} + [dGTP]} \qquad \text{[eq. 1]}$$