Supporting Information

Single-molecule FRET reveals an innate fidelity checkpoint in DNA polymerase I

Svitlana Y. Berezhna[1*], Joshua P. Gill[1,2*], Rajan Lamichhane[1], David P. Millar[1].

(1) Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037.

(2) Present address: Life Technologies Inc., Carlsbad, CA.

(*) These authors contributed equally to the work.

Supporting Information consists of supplemental experimental procedures and seven figures combined in a single pdf file, which includes information on oligonucleotide labeling; effects of fluorophore labels on nucleotide incorporation activity; sample preparation for smFRET measurement; interconversion among different FRET states for L744C KF polymerase bound to DNA; histogram of smFRET efficiencies for binary L744C/L361A KF - DNA complexes; fluorescence intensity time traces and smFRET efficiency trajectories for L744C KF molecules binding to DNA primer/templates with dA in the template extension position in the presence of incorrect substrates dGTP and dCTP; histogram of smFRET efficiencies for L744C KF complexes in the presence of all four nucleotide substrates; fluorescence intensity time traces and smFRET efficiency trajectories for L744C KF polymerases binding to DNA molecules with dC in the template extension position; ES-TOF mass spectra of non-labeled and labeled KF derivatives.

Supporting Experimental Procedures.

Oligonucleotide labeling. Labeling of primer oligonucleotides with Alexa-Fluor 488 dyes was performed using primers that contained an internal modification with amino deoxythymidine dT C_6 at position 8 from the 3' end of the primer (Gene Link). Typically, 25 nmol of primer was incubated with a 5 to 8-fold excess of Alexa-Fluor 488 carboxylic acid, succinimidyl ester (Invitrogen) in 100 mM Na_2CO_3, pH 8.5 for 20 hours at room temperature. The excess free dye was removed by use of a Sephadex-G25 gel filtration column equilibrated with 10 mM TrisHCl, pH 7.5 and 1 mM EDTA. The labeled primer was separated from unlabeled primer by 20% denaturing PAGE. The extent of labeling, calculated from the optical absorbances at 490 and 260 nm, corresponding to dye and DNA peaks respectively, was typically 100% after purification (the dye absorbance at 260 nm was accounted for).

Effects of fluorophore labels on nucleotide incorporation activity. The effects of fluorophores introduced at specific labeling positions within KF or the DNA primer strand were assessed by a primer extension assay. Reactions were initiated by mixing 15 nM KF (labeled or unlabeled) with annealed primer-template duplex (10 μM primer, labeled or unlabeled, and 20 μM template) containing 100 mM of each dNTP. The reaction buffer contained 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, and 0.05 mg/mL BSA. The reaction was stopped by the addition of a formamide solution (containing 10 mM EDTA, pH 8.0) after times ranging from 1 to 30 min. All reactions were performed at 20 ^{o} C. The extension products were separated on a denaturing

sequencing polyacrylamide gel (20% polyacrylamide, 8 M urea, and 1xTBE), stained with SybrGold, and visualized under a UV lamp.

Sample preparation for smFRET measurements. The sample chamber was assembled using 1 mm thick quartz slides (TIRF Technologies) as an active surface and 0.17 mm thick microscope glass cover slips (Fisherbrand, 22 x 40 x 1 mm) as covers, sealed together by a single layer of a dual-sided adhesive tape (Scotch), according to the standard protocol used in the field[1]. To decrease non-specific binding of KF molecules to the surface, the quartz surface was pre-treated with a mixture (95:5) of PEG-silane (MW = 5000, Laysan Bio Inc.) and biotin-PEG-silane (MW = 5000, Laysan Bio Inc), using a protocol developed in our laboratory, followed by application of 0.2 mg/ml immuno pure streptavidin (Thermo Scientific). Subsequently, 20 µl of 100pM Alexa-Fluor 488 labeled biotinylated DNA was added to the sample chamber and allowed to incubate in the reaction buffer for 5 minutes for surface attachment. Unbound DNA was removed by flushing 200 µl of buffer through the sample chamber. The DNA-containing sample was examined on the TIRF microscope for determination of the DNA spot density, intensity and uniformity and for background signal measurement. Andor Solis software package was used in calibration data recording and processing. Typically, 100-200 well defined and approximately equally spaced fluorescence spots were detected within a single field of view. The laser intensity was slightly re-adjusted for each experiment to produce the same spot intensity in all experiments. Upon completion of DNA calibration measurements, a reaction mixture of 5-20 nM labeled KF and 1 mM dNTP (if present) was injected into the sample chamber through the side holes using custom-built

microfluidics and allowed to equilibrate for 5 minutes at room temperature. The reaction buffer contained 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 0.05 mg/mL BSA and 1 mM propyl gallate as an oxygen scavenger. Due to the small amount of KF storage buffer present, approximately 5% glycerol was also present in the reaction mixture.

References

(1) Roy, R.; Hohng, S.; Ha, T. *Nature Methods* **2008**, *5*, 507-516.

(2) Lam, W. C.; Van der Schans, E. J.; Joyce, C. M.; Millar, D. P. *Biochemistry* **1998**, *37*, 1513.

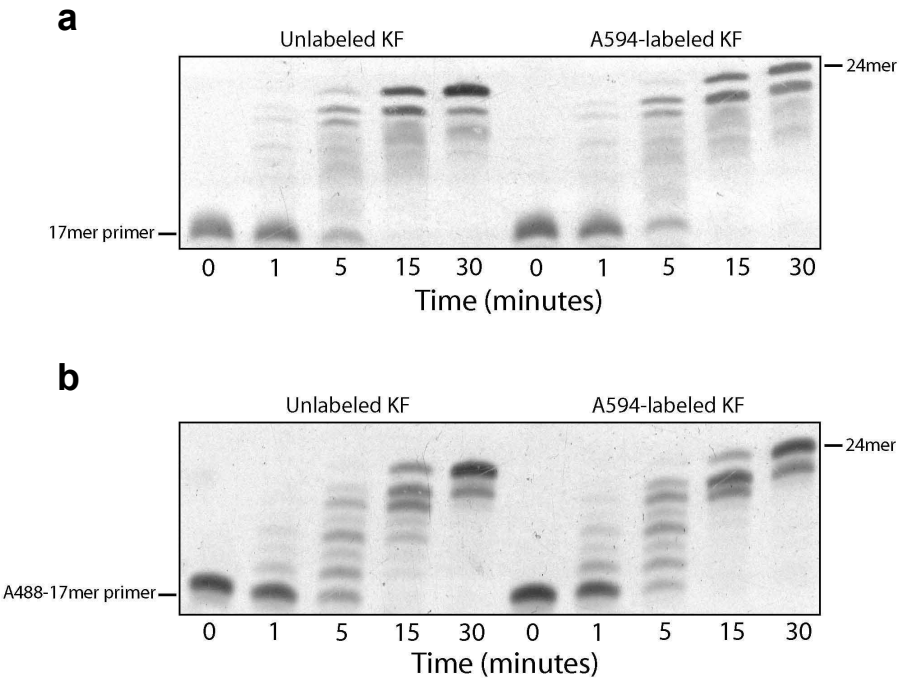


Figure S1. Effects of fluorescence labels on the enzymatic activity of KF polymerases. (a) Unlabeled or A594-labeled L744C KF enzymes (15 nM) were used to extend an unlabeled 17-mer primer (10 μM primer and 20 μM template). The similar extension time courses demonstrate no effect of the A594 dye on the enzymatic activity of labeled KF. (b) Unlabeled or A594-labeled KF were allowed to extend the A488-17mer primer, reagent concentrations are same as in (a). The similar extension time courses show that the A488 dye attached to the primer oligonucleotide has no significant effect on polymerase activity. Note that a very low polymerase concentration was deliberately used in the primer extension assays to ensure that the extension products would be laddered over a period of time, thereby revealing any effects arising from the dye labels on the DNA or KF.

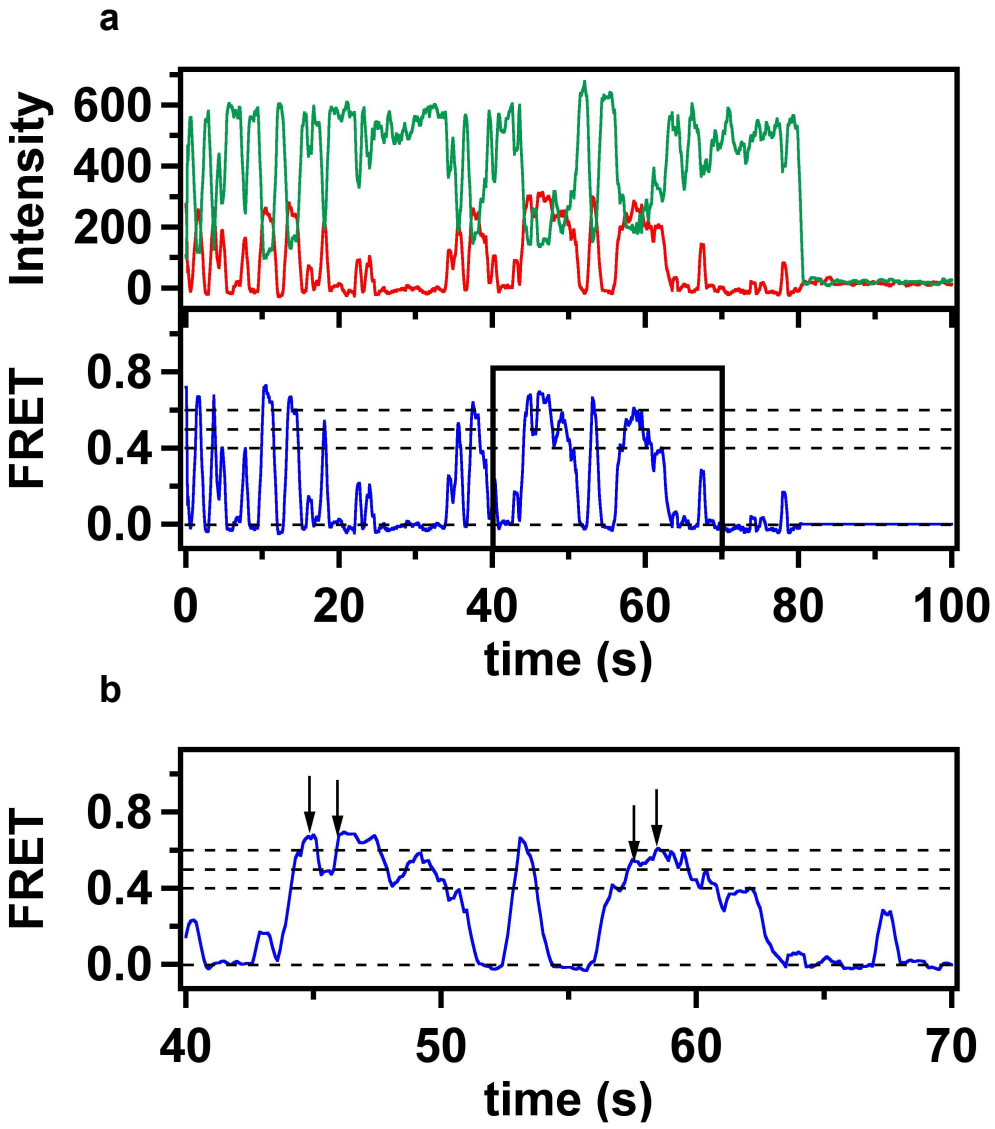


Figure S2. L744C KF polymerase bound to DNA displays interconversion between different FRET states. (a) While polymerase is bound to DNA, it is mostly observed in a single conformation, although some interconversion events can be captured during extended binding periods. (b) The enlarged fragment of the smFRET trajectory shows transitions occuring between the open (0.41 FRET), ajar (0.50 FRET) and closed (0.63 FRET) conformations, indicated by the arrows, demonstrating that these states are interconnected in a kinetic pathway. The guide lines positioned at 0, 0.4, 0.5 and 0.6 FRET efficiencies are for ease of visual inspection.

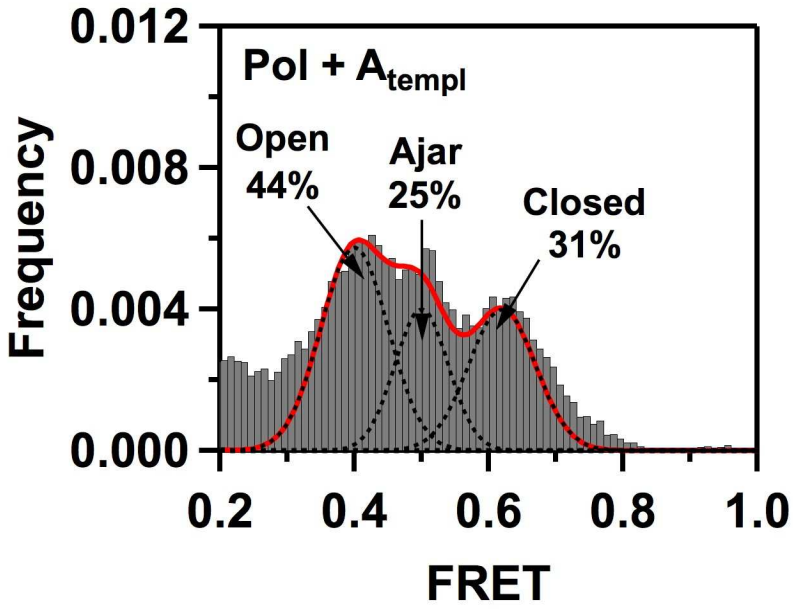


Figure S3. Histogram of smFRET efficiencies for binary L744C/L361A KF - DNA complexes. Binary Pol+A_{templ} complexes, formed by association of acceptor-labeled L744C/L361A mutant KF enzymes and donor-labeled DNA molecules with A in the template extension position. The L361A mutation is known to disrupt binding of DNA at the 3′-5′ exonuclease (exo) site[2]. The FRET efficiency histogram compiled from 149 individual trajectories reveals three bound states, with mean FRET efficiencies of 0.41, 0.50 and 0.63, as seen in corresponding L744C KF-DNA complexes (Fig. 2a, main text). However, the state at 0.9 FRET efficiency is absent, indicating that the peak represents a subpopulation of DNA primer/templates bound at the exo site. The histogram was fitted as described in the main text. Dashed black lines show individual Gaussian distributions with the red line corresponding to a sum of Gaussians. The percentage numbers on each curve indicate the fraction of complexes in each population.

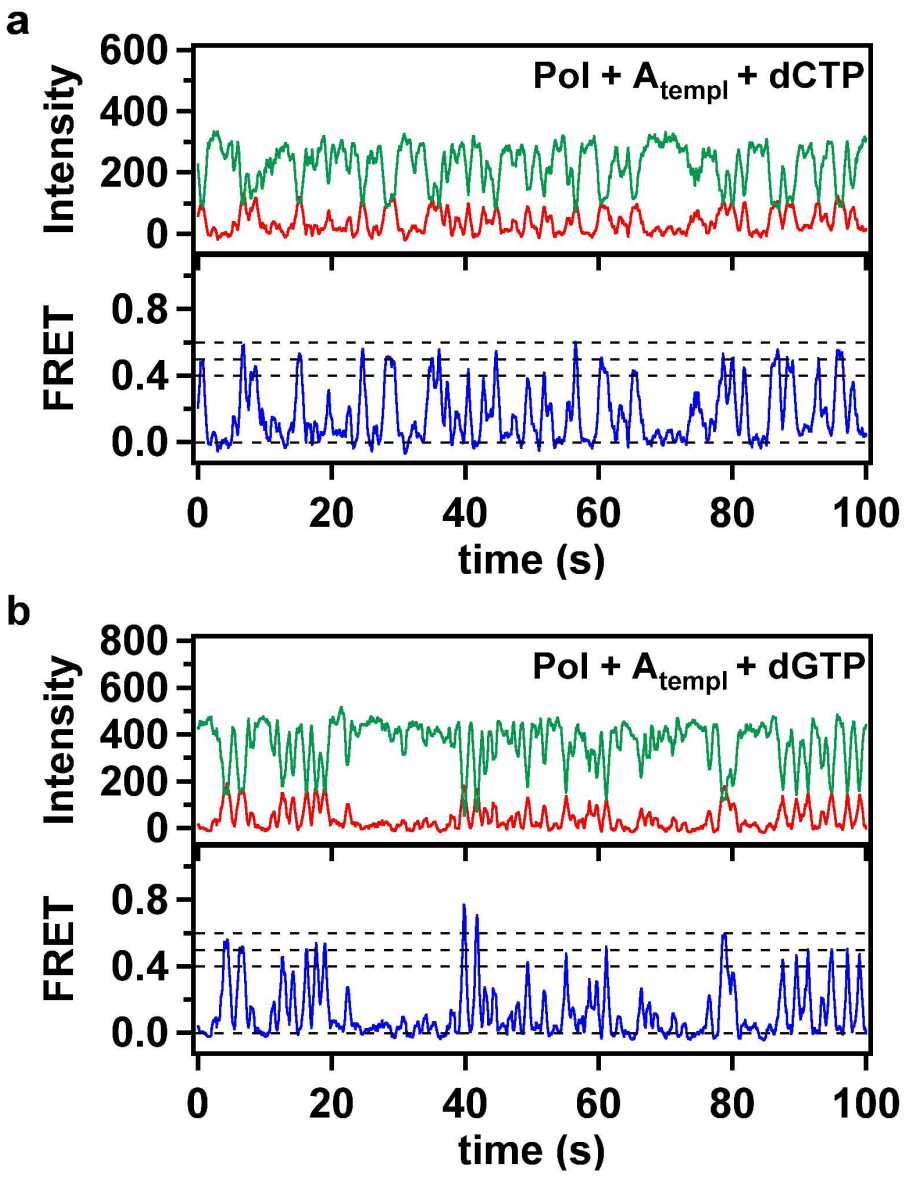


Figure S4. Fluorescence intensity time traces and smFRET efficiency trajectories for L744C KF molecules binding to DNA primer/templates with dA in the template extension position in the presence of incorrect substrates dGTP and dCTP. In the presence of only incorrect dCTP (a) or dGTP (b) nucleotides, polymerases most frequently populate short-lived 0.4 and 0.5 FRET states and dissociate without being able to form a stable ternary complex. The guide lines at 0, 0.4, 0.5 and 0.6 FRET efficiencies are added for ease of visual inspection.

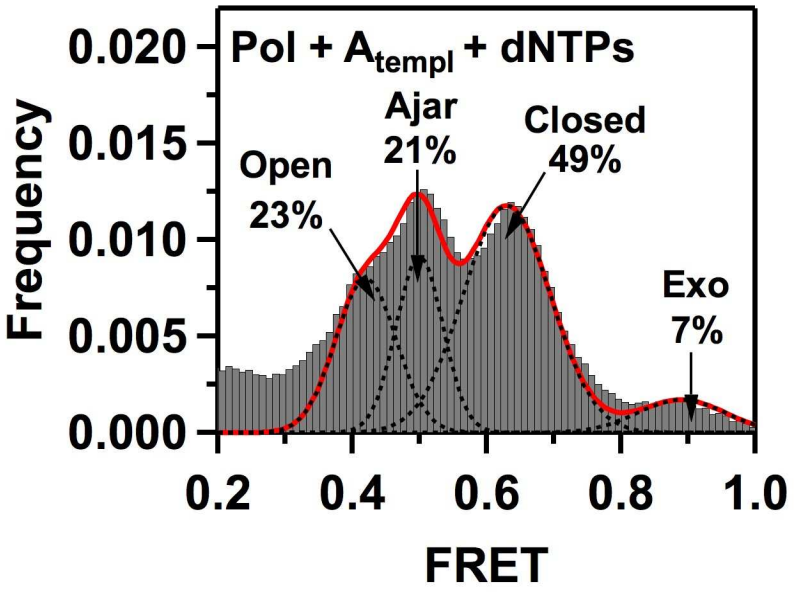


Figure S5. Histogram of smFRET efficiencies for L744C KF complexes in the presence of all four nucleotide substrates. Ternary Pol+A_{templ}+dNTPs complexes formed in the presence of 1 mM total mixed dNTPs (equal amounts of dTTP, dCTP, dGTP, dATP) show a redistribution of open (0.42 FRET), intermediate (0.50 FRET) and closed (0.63 FRET) states compared to either "only correct" (dTTP-dA) or "only incorrect" (dATP-dA) complexes (Fig. 2 in main text). The state at 0.9 FRET efficiency is also reduced compared to the dATP-dA complex (Fig. 2c, main text). Individual peaks in the histogram were fitted with Gaussian functions (black dashed lines). The solid red line corresponds to a sum of Gaussians. The percentage numbers on each peak indicate the fraction of complexes in each population. The histogram was compiled from 321 individual smFRET time trajectories.

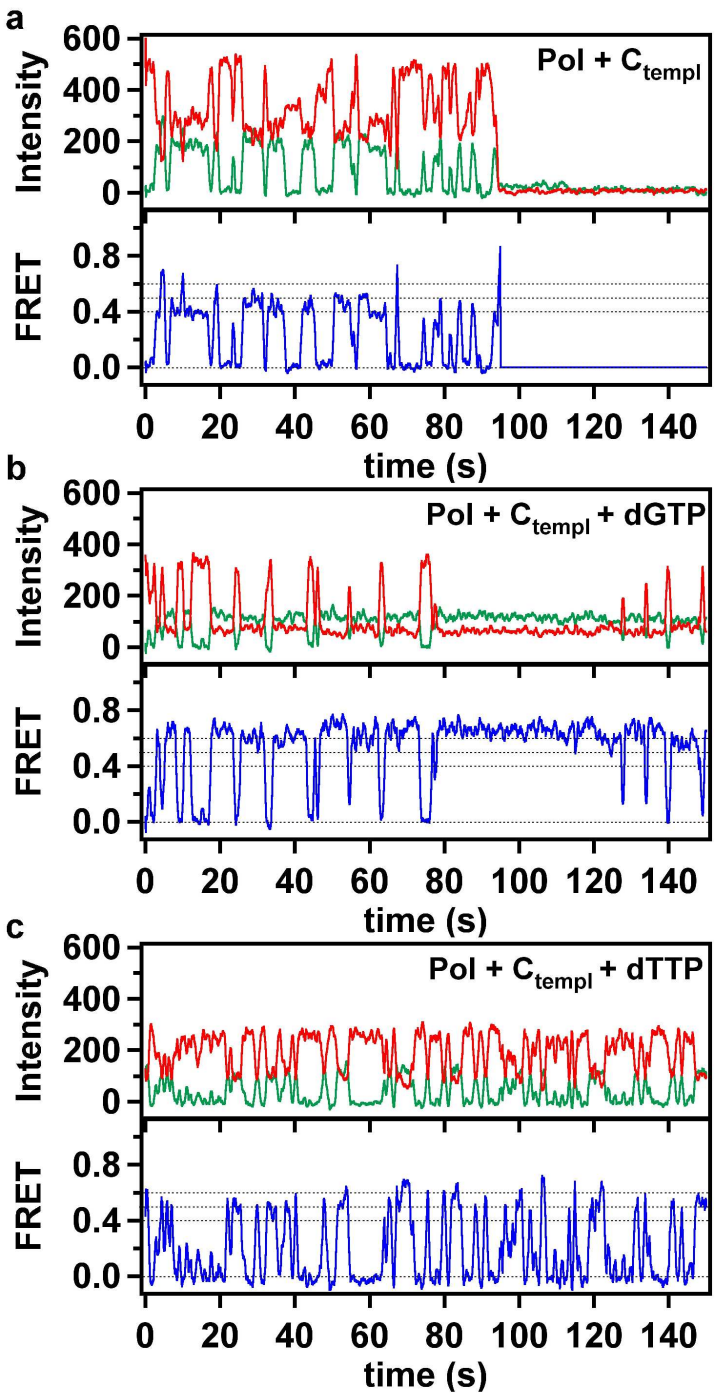
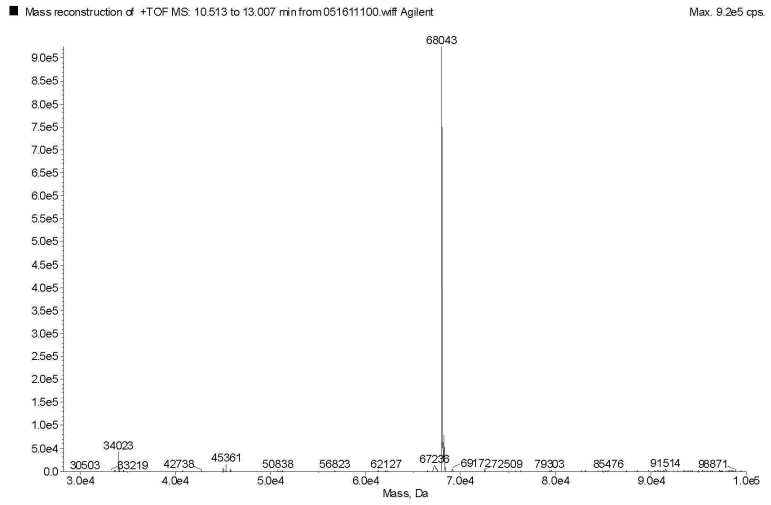


Figure S6. L744C KF polymerases binding to DNA molecules with dC in the template extension position. (a) Binary Pol+C_{templ} complexes with dC in the template extension position populate 0.41, 0.50 and 0.63 FRET states, similar to Pol+A_{templ} complexes (Fig. 2a, main text). In the case of extended binding events, a few inter-conversion transitions can be clearly resolved. (b) Ternary correct Pol+C_{templ}+dGTP complexes favor a relatively stable binding state with 0.64 FRET efficiency. (c) In the presence of only incorrect dCTP nucleotides, the polymerase frequently samples short-lived 0.4 and 0.5 FRET states. The guide lines at 0, 0.4, 0.5 and 0.6 FRET efficiencies are for ease of visual inspection.

a



b

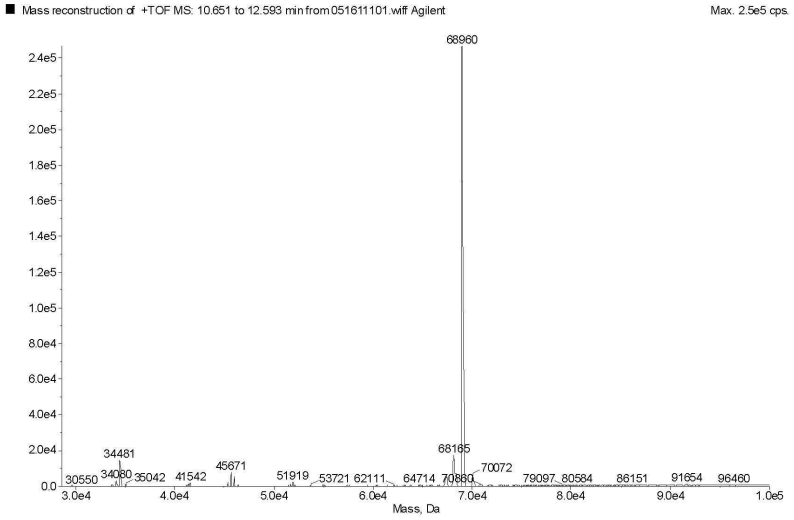


Figure S7. ES-TOF mass spectra of non-labeled and labeled KF derivatives. (a) Non-labeled L744C KF mutant has a single dominant peak at 68043 molecular weight, in good agreement with a calculated mass of 68030 for genotype D424A/L744C/C907S. (b) Alexa-Fluor 594 labeled L744C KF protein exhibits a single peak with molecular weight 68960, within instrumental error of the expected 68939 mass for Alexa-Fluor 594 labeled genotype D424A/L744C/C907S.