Supporting Information for:

Kinetic Mechanism at the Branchpoint Between the DNA Synthesis and Editing Pathways in Individual DNA Polymerase Complexes

Kate R. Lieberman, Joseph M. Dahl, and Hongyun Wang

Supporting Information Text

I. Dwell time distribution of the upper amplitude state for the 3 models in Figure 4

From mathematical point of view, the model in Figure 4c contains the models in Figure 4a and b as two special cases, respectively, when $r_5 = r_6 = 0$ and when $r_3 = r_4 = 0$. Thus, we study the model in Figure 4c.

To find the dwell time distribution, we analyze the problem of escaping from the upper amplitude state, which is a composite state consisting of state 1 and state 3 (Figure 4).

Let $P_1(t)$ = probability of state 1 at time t

$P_3(t)$ = probability of state 3 at time t

In the escaping problem, P_1 and P_3 are governed by the initial value problem

$$\begin{cases} \dfrac{d}{dt}\begin{pmatrix} P_1 \\ P_3 \end{pmatrix} = -A\begin{pmatrix} P_1 \\ P_3 \end{pmatrix}, \qquad A = \begin{pmatrix} (r_1 + r_3) & -r_4 \\ -r_3 & (r_4 + r_6) \end{pmatrix} \\ \begin{pmatrix} P_1(0) \\ P_3(0) \end{pmatrix} = \begin{pmatrix} \dfrac{r_2}{r_2 + r_5} \\ \dfrac{r_5}{r_2 + r_5} \end{pmatrix} \end{cases}$$

Let λ_1 and λ_2 denote the two eigenvalues of matrix A.

Both P_1 and P_3 are linear combinations of $\exp(-\lambda_1 t)$ and $\exp(-\lambda_2 t)$:

$$P_1(t) = a_1 \exp(-\lambda_1 t) + a_2 \exp(-\lambda_2 t)$$

$$P_1(t) = b_1 \exp(-\lambda_1 t) + b_2 \exp(-\lambda_2 t)$$

The probability of remaining in the upper amplitude state at time t is $P_1(t) + P_3(t)$.

The dwell time distribution of the upper amplitude state is

$$\begin{aligned} \rho_{\text{upper}}(t) &= -\frac{d}{dt}\left(P_1(t) + P_3(t)\right) \\ &= (a_1 + b_1)\lambda_1 \exp(-\lambda_1 t) + (a_2 + b_2)\lambda_2 \exp(-\lambda_2 t) \end{aligned}$$

In the escaping problem, we have $P_1(0) + P_3(0) = 1$, which gives us $(a_1 + b_1) + (a_2 + b_2) = 1$. As a result, we write the dwell time distribution as

$$\rho_{\text{upper}}(t) = c\,\lambda_1 \exp(-\lambda_1 t) + (1-c)\lambda_2 \exp(-\lambda_2 t)$$

Therefore, for all 3 models in Figure 4, the dwell time distribution of the upper amplitude state consists of 2 exponential modes.

II. Behavior of Q vs voltage for the 3 models in Figure 4

To distinguish which of the 3 models in Figure 4 is consistent with experimental observations, we consider quantity Q defined as

$$Q = \frac{\lambda_1 \lambda_2}{c\,\lambda_1 + (1-c)\lambda_2}$$

We first derive an analytical expression of Q for the general model in Figure 4c.

The numerator of Q is given by the determinant of the matrix

$$\lambda_1 \lambda_2 = \det \begin{pmatrix} (r_1 + r_3) & -r_4 \\ -r_3 & (r_4 + r_6) \end{pmatrix} = r_1 (r_4 + r_6) + r_3 r_6$$

Using the differential equation and the initial condition, we write the denominator of Q as

$$c\,\lambda_1 + (1-c)\lambda_2 = \rho_{\text{upper}}(0) = -\frac{d}{dt}\left(P_1(t) + P_3(t)\right)\Big|_{t=0}$$

$$= \begin{pmatrix} 1 & 1 \end{pmatrix} A \begin{pmatrix} P_1(0) \\ P_3(0) \end{pmatrix} = \frac{r_1 r_2 + r_6 r_5}{r_2 + r_5}$$

It follows that Q has the expression

$$Q = \frac{r_1 (r_4 + r_6) + r_3 r_6}{\dfrac{r_1 r_2 + r_6 r_5}{r_2 + r_5}}$$

Model in Figure 4a

Substituting $r_5 = r_6 = 0$ into the expression above, we have

$Q = r_4$ for model in Figure 4a

which is expected to be independent of the applied voltage.

Model in Figure 4b

Substituting $r_3 = r_4 = 0$ into the expression above, we have

$$Q = \frac{r_1 r_6}{\dfrac{r_1 r_2 + r_6 r_5}{r_2 + r_5}}$$

For the complexes formed between D12A/D66A and DNA1-OH, there are 2 separate time scales: 1) fast transitions between pre-translocation and post-translocation at the polymerase site, 2) slow transitions to and from the exonuclease site manifested as long stays at the exonuclease site and long waiting between

long stays (Figure 2b). Mathematically, for D12A/D66A with DNA1-OH, we have $(r_5, r_6) << (r_1, r_2)$. Consequently, Q becomes

$$Q \approx r_6 \qquad \text{for model in Figure 4b}$$

which increases as the applied voltage is reduced.

Model in Figure 4c

$$Q = \frac{r_1(r_4 + r_6) + r_3 r_6}{\frac{r_1 r_2 + r_6 r_5}{r_2 + r_5}}$$

For D12A/D66A with DNA1-OH, Figure 2bg shows long stays at the exonuclease site and long waiting between long stays. The 2 separate time scales inform us that

$$(r_3, r_4) << (r_1, r_2) \qquad \text{and} \qquad (r_5, r_6) << (r_1, r_2)$$

It follows that Q has the expression

$$Q \approx r_4 + r_6 \qquad \text{for model in Figure 4c}$$

which increases as the applied voltage is reduced.

In summary,

Model in Figure 4a: $Q = r_4$ which is independent of the voltage

Model in Figure 4b: $Q \approx r_6$ which increases as the voltage is reduced

Model in Figure 4c: $Q \approx r_4 + r_6$ which increases as the voltage is reduced

Q is calculated from ionic current time traces as follows: i) dwell time samples of the upper amplitude state are extracted from the time traces; ii) parameters λ_1, λ_2 and c are determined by fitting a model of 2 exponential modes to dwell time samples; iii) Q is calculated from the definition

$$Q = \frac{\lambda_1 \lambda_2}{c\lambda_1 + (1-c)\lambda_2}$$

Figure 5 shows Q vs voltage for D12A/D66A with DNA1-OH. It demonstrates clearly that Q is independent of the voltage. Therefore, we select the model in Figure 4a.

II. Extracting kinetic rates from dwell time samples

We determine kinetic rates (r_1, r_2, r_3, r_4) in the model shown in Figure 4a.

For the model in Figure 4a, the matrix is

$$A = \begin{pmatrix} (r_1 + r_3) & -r_4 \\ -r_3 & r_4 \end{pmatrix}$$

The two eigenvalues satisfy

$$c\lambda_1 + (1-c)\lambda_2 = \frac{r_1 r_2 + r_6 r_5}{r_2 + r_5} = r_1$$

$$\lambda_1\lambda_2 = \left(r_1 + r_3\right)r_4 - r_3r_4 = r_1r_4$$

$$\lambda_1 + \lambda_2 = r_1 + r_3 + r_4$$

From these 3 equations, we have

$$r_1 = c\,\lambda_1 + \left(1 - c\right)\lambda_2$$

$$r_3 = \frac{c\left(1 - c\right)\left(\lambda_1 - \lambda_2\right)^2}{c\,\lambda_1 + \left(1 - c\right)\lambda_2}$$

$$r_4 = \frac{\lambda_1\lambda_2}{c\,\lambda_1 + \left(1 - c\right)\lambda_2}$$

Kinetic rates r_1, r_3 and r_4 are calculated from intermediate parameters λ_1, λ_2 and c, which are determined by fitting a model of 2 exponential modes to the dwell time samples of the upper amplitude state.

Kinetic rate r_2 is calculated directly by fitting an exponential distribution to the dwell time samples of the lower amplitude state.

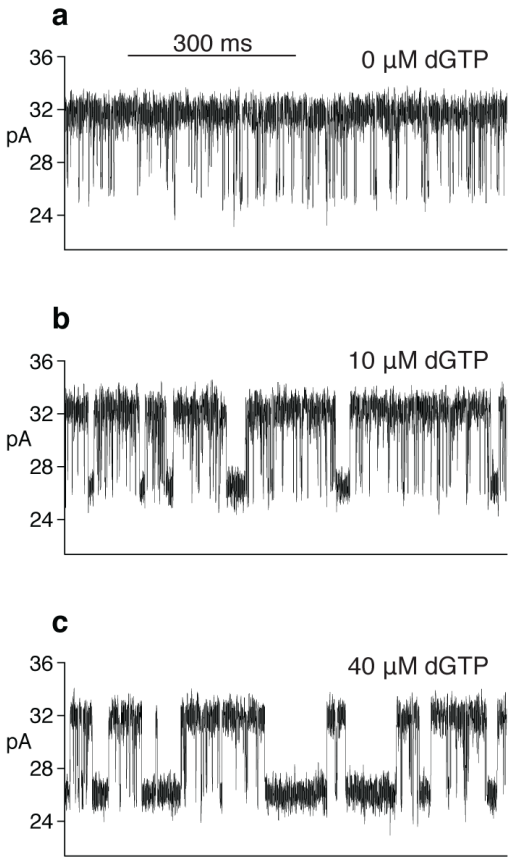

Figure S1: **Complementary dNTP binding stabilizes the post-translocation state in Φ29 DNAP complexes**. Ionic current traces for complexes formed between wild type Φ29 DNAP and the DNA1-H substrate, captured at 180 mV in the presence of (a) 0 µM (b) 10 µM or (c) 40 µM dGTP (complementary to the dCMP residue at n=0 of the template strand).

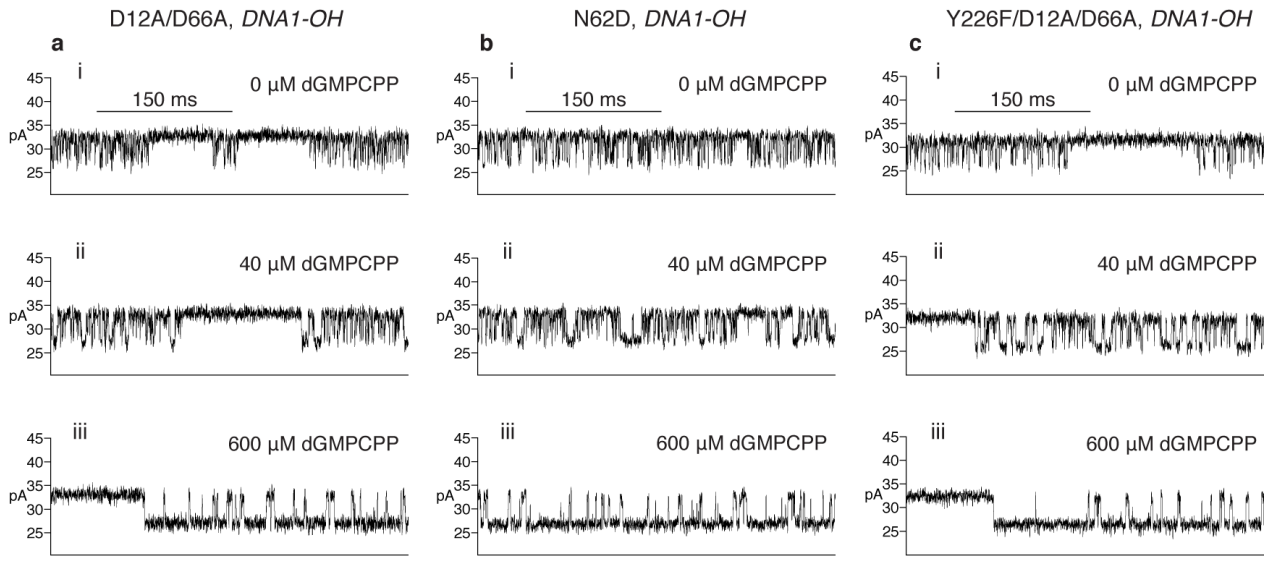

Figure S2: **Binding of a complementary, non-hydrolyzable dNTP analog stabilizes the post-translocation state in Φ29 DNAP complexes formed with DNA1-OH**. Ionic current traces for complexes formed between the (a) D12A/D66A, (b) N62D, or (c) Y226F/D12A/D66A mutants of Φ29 DNAP and the DNA1-OH substrate, captured at 180 mV in the presence of (i) 0 µM (ii) 40 µM or (iii) 600 µM deoxyguanosine-5′-[(alpha,beta)-methyleno]triphosphate (dGMPCPP; complementary to the dCMP residue at n=0 of the template strand).

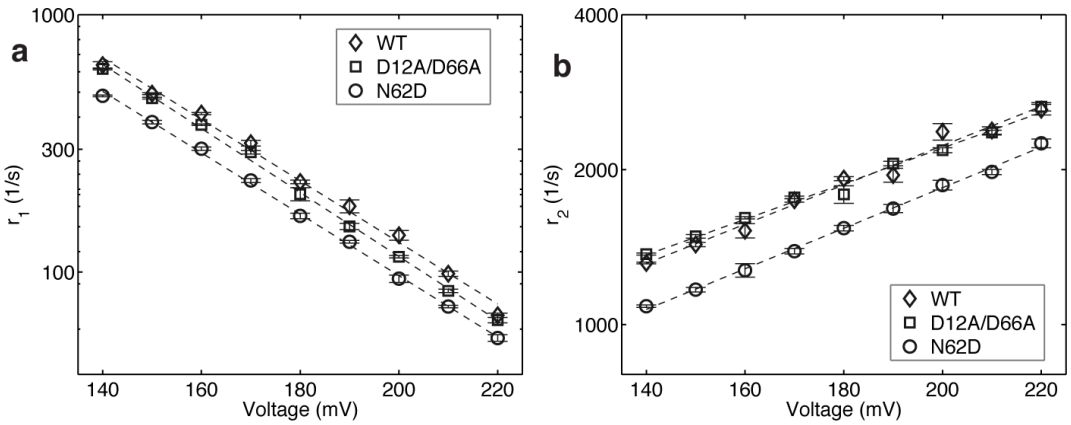

Figure S3: **Influence of the N62D mutation on the translocation rates in complexes formed with DNA1-H.** (a) Plots of log(r_1) vs. voltage and (b) log(r_2) vs. voltage for complexes formed beween DNA1-H and wild type Φ29 DNAP (diamonds), the D12A/D66A mutant (squares), or the N62D mutant (circles). Rates were determined using dwell time samples extracted from ionic current traces and a two-state model for the translocation step (Lieberman, K. R.; Dahl, J. M.; Mai, A. H.; Akeson, M.; Wang, H. *J Am Chem Soc* **2012**, *134*, 18816–18823).

Table S1. Translocation and primer strand transfer rates in complexes formed with DNA1-OH.

enzyme	Voltage	r_1 (s^{-1}) [a]	r_2 (s^{-1}) [b]	r_3 (s^{-1}) [c]	r_4 (s^{-1}) [d]
D12A/D66A	140 mV	1516.9 ± 38.83	1520.3 ± 23.95	11.11 ± 0.82	9.56 ± 0.83
	150 mV	1433.8 ± 11.66	1743.7 ± 8.05	12.15 ± 0.62	10.6 ± 0.62
	160 mV	1176.8 ± 17.89	2006.5 ± 33.0	13.19 ± 0.92	11.54 ± 3.07
	170 mV	1007.3 ± 20.71	2606.6 ± 39.21	11.98 ± 1.17	10.17 ± 0.91
	180 mV	861.07 ± 10.56	2935.3 ± 27.2	11.25 ± 0.48	10.8 ± 1.07
	190 mV	684.68 ± 21.87	3478.3 ± 52.93	10.47 ± 0.69	10.14 ± 3.55
	200 mV	506.64 ± 12.77	4616.7 ± 37.48	9.48 ± 0.89	9.3 ± 1.07
	210 mV	424.27 ± 7.37	5187.4 ± 76.88	10.93 ± 0.70	9.36 ± 0.70
	220 mV	329.48 ± 7.65	6044.9± 81.84	13.34 ± 1.16	12.84 ± 1.35
N62D	140 mV	1458.3 ± 78.11	698.18 ± 12.05	52.05 ± 31.2	371.54 ± 222.9
	150 mV	1297.9 ± 35.88	915.69 ± 21.84	35.92 ± 21.6	253.02 ± 151.8
	160 mV	1159 ± 22.02	1113 ± 14.11	27.45 ± 16.47	220.01 ± 123.9
	170 mV	995.87 ± 12.61	1399 ± 29.88	27.68 ± 11.57	228.53 ± 53.47
	180 mV	814.65 ± 15.41	1761.8 ± 47.62	30.7 ± 18.42	211.34 ± 101
	190 mV	696.38 ± 12.54	2081.7 ± 66.43	32.19 ± 19.31	201.84 ± 63.87
	200 mV	557.36 ± 8.47	2279.5 ± 37.8	32.41 ± 11.68	175.07 ± 19.13
	210 mV	525 ± 10.17	3000.3 ± 31.79	54.03 ± 8.65	174.52 ± 16.69
	220 mV	397.2 ± 16.06	3452.4 ± 81.2	75.79 ± 22.92	155.07 ± 27.65
Y226F/D12A/ D66A	140 mV	1107.5 ± 8.94	1197.5 ± 7.40	6.88 ± 0.62	15.27 ± 3.83
	150 mV	965.03 ± 30.50	1347.7 ± 32.54	6.72 ± 0.75	18.54 ± 3.40
	160 mV	769.5 ± 8.67	1561.6 ± 29.33	5.17 ± 0.47	11.8 ± 1.97
	170 mV	622.3 ± 6.83	1821.1 ± 9.93	4.87 ± 0.32	10.98 ± 1.55
	180 mV	450.38 ± 5.32	2077.1 ± 18.63	4.51 ± 0.33	11.4 ± 1.44
	190 mV	364.44 ± 8.91	2112.5 ± 38.13	4.65 ± 0.38	11.01 ± 0.97
	200 mV	286.62 ± 9.64	2279 ± 36.03	4.40 ± 0.36	10.06 ± 1.10
	210 mV	220.03 ± 2.0	2657.6 ± 33.04	4.74 ± 0.56	8.85 ± 0.93
	220 mV	154.27 ± 4.61	2939.1 ± 58.23	4.31 ± 0.43	10.28 ± 1.06

[a] The rate of transition from the pre-translocation to the post-translocation state.
[b] The rate of transition from the post-translocation to the pre-translocation state.
[c] The rate of primer strand transfer from the polymerase site pre-translocation state to the exonuclease site.
[d] The rate of primer strand transfer from the exonuclease site to the polymerase site pre-translocation state.
Rates were determined using dwell time samples extracted from ionic current traces and the three-state model in Figure 4a. All values are reported with the standard error.