Supporting Information for:

Kinetic Mechanism of Translocation and dNTP Binding in Individual DNA Polymerase Complexes

Kate R. Lieberman, Joseph M. Dahl, Ai H. Mai, Ashley Cox,
Mark Akeson and Hongyun Wang

Supporting Information Text

Derivation of $1/\langle \Delta T_{post} \rangle$ at high [dGTP]

As shown in Figure 3B, the transition rates between the two post-translocation states are k_{on}[dGTP] and k_{off}. The time scale of relaxing to the equilibrium between two states is given by the reciporocal of the sum of the transition rates. At high dGTP conentration, k_{on}[dGTP] + k_{off} is large and as a result, the two post-translocation states are effectively in equilibrium. Thus, we treat the two post-translocation states as one composite state with the probability of each sub-state given by

$$p_{\text{E-DNA}} = k_{off} \;/\; (k_{on}[\text{dGTP}] + k_{off})$$

$$p_{\text{E-DNA-dNTP}} = k_{on}[\text{dGTP}] \;/\; (k_{on}[\text{dGTP}] + k_{off})$$

The overall transition rate from the composite post-translocation state to the pre-translocation state is

$$r_{\text{post→pre}} = r_4\, p_{\text{E-DNA-dNTP}} + r_2\, p_{\text{E-DNA}}$$

$$= (r_4\, k_{on}[\text{dGTP}] + r_2\, k_{off}) \;/\; (k_{on}[\text{dGTP}] + k_{off})$$

$$= r_4 + (r_2 - r_4)\, k_{off} \;/\; (k_{on}[\text{dGTP}] + k_{off})$$

$$= r_4 + (r_2 - r_4)\, K_d/[\text{dGTP}] - (r_2 - r_4)\, K_d^2/[\text{dGTP}]^2 + O(1/[\text{dGTP}]^3)$$

$$\approx r_4 + \alpha/[\text{dGTP}]$$

where $K_d = k_{off} \,/\, k_{on}$ and $\alpha = (r_2 - r_4)\, K_d$.

In the measured time traces of ionic current amplitude, the composite post-translocation state is detected as the lower amplitude state since the two post-translocation states yield the same amplitude level.

The mean dwell time of the composite post-translocation state satisfies

$$1/\langle \Delta T_{post} \rangle = r_{post \to pre} \approx r_4 + \alpha/[dGTP]$$

The three-state model and the method for estimating transition rates from measured time traces of amplitude

As shown in Figure 3, the 3 states are

State 1: pre-translocation state (upper amplitude)

State 2: post-translocation state (lower amplitude)

State 2B: post-translocation state with dGTP bound (lower amplitude)

Based on the analysis shown in Figure 4, we eliminated the direct transitions between the pre-translocation state and the dGTP bound post-translocation state (r_3[dGTP] and r_4 in Figure 3B). Thus, the 3-state model becomes the one shown in Figure 3C where the dGTP can only bind when the complex is in the post-translocation state. The 3 states are connected by 4 transition rates

r_1: transition rate from state 1 to state 2

r_2: transition rate from state 2 to state 1

k_{on}: first order rate constant of dGTP binding

k_{off}: rate of dGTP dissociating

At equilibrium, the probabilities of 3 states satisfy

$$p_1 + p_2 + p_{2B} = 1$$

$$\frac{p_{2B}}{p_2} = \frac{[dGTP]}{K_d}, \qquad K_d = \frac{k_{off}}{k_{on}}$$

$$\frac{p_1}{p_2} = \frac{r_2}{r_1}$$

Solving these equations, we obtain

$$p_1 = \frac{\frac{r_2}{r_1}}{1 + \frac{[dGTP]}{K_d} + \frac{r_2}{r_1}}$$

$$p_2 = \frac{1}{1 + \frac{[dGTP]}{K_d} + \frac{r_2}{r_1}}$$

$$p_{2B} = \frac{\frac{[dGTP]}{K_d}}{1 + \frac{[dGTP]}{K_d} + \frac{r_2}{r_1}}$$

Below we derive a method for estimating the 4 transition rates from data at individual values of voltage and [dGTP].

The two post-translocation states (states 2 and 2B) yield the same current amplitude. The measured time traces of amplitude show only two amplitude levels:

I_1: the true amplitude of the pre-translocation state (without noise)

I_2: the true amplitude of the two post-translocation states (without noise)

Note that I_1 is the upper amplitude and I_2 is the lower amplitude: $I_2 < I_1$.

Let

$S(t)$: state (1, 2, or 2B) of the complex atop of the pore at time t

$I(t)$: the *true* amplitude (without noise) at time t, corresponding to state $S(t)$.

$X(t)$: measured time trace of amplitude = $I(t)$ + noise

We have

$$I(t) = \begin{cases} I_1 , & S(t) = 1 \\ I_2 , & S(t) = 2 \text{ or 2B} \end{cases}$$

$$X(t) = I(t) + N(t)$$

Note that $S(t)$, $I(t)$ and $X(t)$ are all random processes.

We assume that the noise $N(t)$ has zero mean and that $N(t_1)$ is independent of $N(t_2)$.

We map $[I_2, I_1]$ to [-1, 1] and consider the scaled amplitude

$$Y(t) = \frac{2}{(I_1 - I_2)}\left(X(t) - \frac{I_1 + I_2}{2} \right)$$

The mean of $Y(t)$ has the theoretical expression

$$E[Y] = p_1 - (p_2 + p_{2B}) = \frac{\frac{r_2}{r_1} - \left(1 + \frac{[dGTP]}{K_d}\right)}{1 + \frac{[dGTP]}{K_d} + \frac{r_2}{r_1}} \qquad \text{(E01)}$$

This is the first equation for the unknown parameters.

To derive more equations, we consider the auto-correlation

$$R(t) \equiv E\left[Y(t_0)Y(t_0+t)\right].$$

Form the 3-state model shown in Figure 3C, we can show that the autocorrelation function has two properties:

1. $R(t)-\left(E[Y]\right)^2 = \left(1-\left(E[Y]\right)^2\right)\left[c_1\exp(-\lambda_1 t)+(1-c_1)\exp(-\lambda_2 t)\right]$

   where λ_1 and λ_2 are the eigenvalues of

$$\begin{pmatrix} (r_1+r_2) & k_{on}[dGTP] \\ r_2 & \left(k_{on}[dGTP]+k_{off}\right) \end{pmatrix}$$

   This property leads to two equations for the unknown parameters

$$\lambda_1+\lambda_2 = (r_1+r_2)+\left(k_{on}[dGTP]+k_{off}\right) \tag{E02}$$

$$\lambda_1\cdot\lambda_2 = (r_1+r_2)\left(k_{on}[dGTP]+k_{off}\right)-r_2 k_{on}[dGTP]$$

$$= r_1 k_{off}\left(\frac{r_2}{r_1}+\frac{[dGTP]}{K_d}+1\right) \tag{E03}$$

2. $R'(0) = -2\left(r_1 p_1 + r_2 p_2\right)$

   which gives us another equation for the unknown parameters:

$$\left(1-\left(E[Y]\right)^2\right)\left[c_1\lambda_1+(1-c_1)\lambda_2\right] = 2r_1\left(p_1+\frac{r_2}{r_1}p_2\right)$$

$$= \frac{4r_2}{1+\frac{[dGTP]}{K_d}+\frac{r_2}{r_1}} \tag{E04}$$

Thus, we obtain 4 equations for the 4 unknown parameters $\left(r_1, r_2, k_{on}, k_{off}\right)$:

$$\frac{\frac{r_2}{r_1}-\left(1+\frac{[dGTP]}{K_d}\right)}{1+\frac{[dGTP]}{K_d}+\frac{r_2}{r_1}} = E[Y]$$

$$\frac{4r_2}{1+\frac{[dGTP]}{K_d}+\frac{r_2}{r_1}} = \left(1-\left(E[Y]\right)^2\right)\left[c_1\lambda_1+(1-c_1)\lambda_2\right]$$

$$r_1 k_{off}\left(\frac{r_2}{r_1}+\frac{[dGTP]}{K_d}+1\right) = \lambda_1\cdot\lambda_2$$

$$\left(r_1 + r_2\right) + \left(k_{on}\left[dGTP\right] + k_{off}\right) = \lambda_1 + \lambda_2$$

In the above 4 equations, all quantities on the right hand side are calculated from data.

- $E[Y]$ is calculated directly from a time trace of amplitude $\{Y(t)\}$.
- $R(t)$ is calculated directly from a time trace of amplitude $\{Y(t)\}$.
- c_1, λ_1 and λ_2 are calculated by fitting measured values of $\{R(t)\}$ to the theoretical expression

$$R(t) - \left(E\left[Y\right]\right)^2 = \left(1 - \left(E\left[Y\right]\right)^2\right)\left[c_1 \exp\left(-\lambda_1 t\right) + \left(1 - c_1\right)\exp\left(-\lambda_2 t\right)\right].$$

The 4 unknown parameters are then solved from the 4 equations above. In this way, we can calculate a set of 4 parameters from each measured time trace of amplitude. At each individual voltage and [dGTP], we have 20 ~ 60 measured time traces. From multiple estimated sets of parameter values, we use the mean as a more accurate estimate and use the standard error as the error bar.

Table S1. Translocation and dNTP binding rates.

[dGTP]	Voltage	r_1 (s^{-1}) [a]	r_2 (s^{-1}) [b]	k_{on}($s^{-1}\mu M^{-1}$) [c]	k_{off}(s^{-1}) [d]
0 μM	140 mV	672.94 ± 33.68	1337.9 ± 24.34		
	150 mV	512.32 ± 9.81	1428.8 ± 19.31		
	160 mV	446.8 ± 5.63	1592.8 ± 51.44		
	170 mV	343.31 ± 6.81	1793.5 ± 15.94		
	180 mV	235.71 ± 4.4	1932.9 ± 25.06		
	190 mV	196.16 ± 14.13	1999.2 ± 27.28		
	200 mV	145.09 ± 10.65	2308.2 ± 101.39		
	210 mV	107.9 ± 3.73	2368.7 ± 63.09		
5 μM	140 mV	678.9 ± 18.1	1327 ± 112	12.99 ± 1.73	26.22 ± 2.94
	150 mV	538.1 ± 25.9	1534 ± 101	14.73 ± 1.83	27.84 ± 1.80
	160 mV	412.0 ± 21.8	1689 ± 112	15.31 ± 0.98	30.29 ± 2.03
	170 mV	324.5 ± 10.1	1765 ± 81	13.04 ± 0.96	30.13 ± 1.90
	180 mV	243.0 ± 12.8	2039 ± 128	15.75 ± 1.36	34.61 ± 2.41
	190 mV	181.5 ± 14.5	2118 ± 297	13.32 ± 1.15	30.01 ± 3.10
	200 mV	137.6 ± 12.3	2205 ± 181	13.12 ± 0.95	30.27 ± 1.99
	210 mV	109.7 ± 12.3	2406 ± 306	10.74 ± 1.63	27.42 ± 4.81
10 μM	140 mV	748.5 ± 77.1	1294 ± 264	15.87 ± 4.30	26.94 ± 4.71
	150 mV	575.2 ± 54.0	1477 ± 232	18.96 ± 2.78	28.20 ± 2.76
	160 mV	420.7 ± 27.7	1653 ± 181	19.19 ± 1.43	30.23 ± 1.54
	170 mV	342.2 ± 22.9	1801 ± 166	19.15 ± 1.46	30.09 ± 0.77
	180 mV	257.2 ± 10.4	1920 ± 75	20.94 ± 0.95	30.55 ± 0.65
	190 mV	202.0 ± 8.7	2100 ± 124	19.61 ± 1.17	32.19 ± 1.22
	200 mV	148.0 ± 12.5	2271 ± 318	18.74 ± 1.12	29.42 ± 1.22
	210 mV	100.4 ± 7.3	2194 ± 252	21.91 ± 1.19	35.52 ± 1.94

[dGTP]	Voltage	r_1 (s^{-1}) [a]	r_2 (s^{-1}) [b]	k_{on}($s^{-1}\mu M^{-1}$) [c]	k_{off}(s^{-1}) [d]
20 μM	140 mV	870.3 ± 120	1137 ± 487	10.20 ± 5.09	23.88 ± 4.26
	150 mV	680.0 ± 102	1303 ± 464	12.59 ± 3.95	24.07 ± 3.57
	160 mV	474.1 ± 61.9	1556 ± 334	14.57 ± 2.45	26.40 ± 1.79
	170 mV	366.8 ± 46.4	1703 ± 359	17.00 ± 1.63	29.88 ± 1.88
	180 mV	236.0 ± 15	1931 ± 173	18.81 ± 1.58	31.69 ± 0.81
	190 mV	200.2 ± 15.9	2098 ± 195	18.21 ± 0.62	29.78 ± 1.18
	200 mV	144.5 ± 12.2	2112 ± 271	17.25 ± 1.07	30.29 ± 1.09
	210 mV	102.7 ± 8.6	2349 ± 316	17.49 ± 1.15	32.15 ± 1.62
40 μM	180 mV	227.2 ± 21.8	1765 ± 318	18.01 ± 1.43	31.35 ± 1.02
	190 mV	171.5 ± 16.3	1940 ± 312	17.42 ± 1.17	31.67 ± 1.39
	200 mV	132.0 ± 9.0	2180 ± 310	19.24 ± 1.09	32.35 ± 1.55
	210 mV	87.11 ± 11.5	1985 ± 365	19.79 ± 1.54	35.30 ± 2.96

[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 dGTP association rate.
[d] The dGTP dissociation rate.
All values are reported with the standard error.