## Supporting Information:

# Time Course Analysis of Enzyme-Catalyzed DNA Polymerization 

Julius Rentergent, Max D. Driscoll, and Sam Hay *<br>Manchester Institute of Biotechnology, University of Manchester, Manchester M1 7DN, UK<br>E-mail: sam.hay@manchester.ac.uk

## Assay and initial conditions

For all assays, varying concentrations of DNA template-primer and dNTPs were prepared in 10 mm Tris $/ \mathrm{HCl}, 50 \mathrm{~mm} \mathrm{NaCl}, 10 \mathrm{~mm} \mathrm{MgCl}_{2}, 1 \mathrm{~mm}$ dithiothreitol, pH 8.0 or $\mathrm{pH}^{*}$ 8.0. The reaction was started by adding KF and the total reaction volume was 1 ml . At several time points, starting after 10 seconds, $100 \mu \mathrm{l}$ aliquots were removed and quenched by adding to $1.5 \mu \mathrm{l} \mathrm{PG}(3 \mu \mathrm{M}$ final concentration). Fluorescence measurements were made within 10 min of stopping the reaction.

The following list contains all concentrations of reactants for all experiments. KF dependence using T7pR80 template at $25^{\circ} \mathrm{C}$ : KF 0 to $200 \mathrm{~nm}, 50 \mathrm{~nm}, 500 \mu \mathrm{M}$ dNTP $(125 \mu \mathrm{~m}$ of each type). DNA dependence using T7pR80 template at $25^{\circ} \mathrm{C}$ : $5 \mathrm{~nm} \mathrm{KF}, 0$ to 25 nm DNA, $500 \mu \mathrm{M}$ dNTP. dNTP dependence using T7pR80 template at $25^{\circ} \mathrm{C}$, $\mathrm{pH} 8: 50 \mathrm{~nm} \mathrm{KF}, 100 \mathrm{~nm}$ DNA, 0 to $100 \mu \mathrm{M}$ dNTP. This experiment was repeated in $\mathrm{D}_{2} \mathrm{O}$, and with T 7 pR 30 template in $\mathrm{H}_{2} \mathrm{O}$ and in $\mathrm{D}_{2} \mathrm{O}$. DNA template length dependency was carried out using all four oligos at $25^{\circ} \mathrm{C}$ : 50 nm KF, 100 nm DNA, $500 \mu \mathrm{~m}$ dNTP. KF dependence using T7pR100 template was carried out at $5,10,15,20$ and $25^{\circ} \mathrm{C}$ : KF 0 to 1000 nM at 5,10 and $15^{\circ} \mathrm{C} ; 0$ to 500 nm KF at $20^{\circ} \mathrm{C} ; 0$ to 200 nm KF at $25^{\circ} \mathrm{C}, 50 \mathrm{~nm}$ DNA, $500 \mu \mathrm{M}$. Buffer was adjusted to produce pH 8 at all temperatures.

## Full results tables

Table S1. Full results of global nonlinear regression of Scheme 1 to the three data sets shown in Figure 2.3. $k_{1}, k_{-1}$ and $K_{\mathrm{d}}$ refer to the enzyme-template binding parameters, while $k_{\text {cat }}$ and $K_{\mathrm{m}}$ refer to the steady-state parameters of dNTP substrate.

| Parameter | Value $\pm$ Standard Error |
| :---: | :---: |
| Sum of Squares | $7.64 \times 10^{11}$ |
| $\Delta f /\left(\mathrm{cps} \mu \mathrm{M}{ }^{-1}\right)$ | $3.18 \times 10^{5} \pm 0.07 \times 10^{5}$ |
| $f_{0} /\left(\operatorname{cps~}_{\text {um }}{ }^{-1}\right)$ | $7.60 \times 10^{6} \pm 0.204 \times 10^{6}$ |
| $\mathrm{DNA}_{\text {mod, }}$ KF | $0.77 \pm 0.02$ |
| DNA ${ }_{\text {mod, DNA }}$ | $0.86 \pm 0.03$ |
| DNA ${ }_{\text {mod,dNTP }}$ | $0.72 \pm 0.02$ |
| $k_{1} /\left(\mu^{-1} \mathrm{~s}^{-1}\right)$ | $7.43 \pm 0.97$ |
| $k_{-1} / \mathrm{s}^{-1}$ | $0.074 \pm 0.013$ |
| $K_{\text {d }} / \mathrm{nM}$ | $9.96 \pm 1.79$ |
| $k_{\text {cat }} / \mathrm{s}^{-1}$ | $3.34 \pm 0.14$ |
| $K_{\mathrm{m}} / \mathrm{\mu M}$ | $0.51 \pm 0.04$ |
| $k_{\text {cat }} / K_{\mathrm{m}} /\left(\mu^{-1} \mathrm{~s}^{-1}\right)$ | $6.52 \pm 0.55$ |

Table S2. Full results for global fit to DNA length dependence data. The parameter $k_{- \text {end }}$ refers to the $k_{-1}$ value at the final $\mathrm{DNA}_{\mathrm{n}}$ position. The variation in $\Delta f$ for different templates is produced by different template properties, and by correlation with the initial template concentration, which is affected by both annealing efficiency and experimental error.

| Parameter | Value $\pm$ Standard Error |
| :---: | :---: |
| Sum of Squares | $1.61 \times 10^{11}$ |
| $\Delta f_{30} /\left(\operatorname{cps} \mu \mathrm{m}^{-1}\right)$ | $3.96 \times 10^{5} \pm 0.16 \times 10^{5}$ |
| $\Delta f_{50} /\left(\operatorname{cps~}_{\mu \mathrm{M}}{ }^{-1}\right)$ | $3.36 \times 10^{5} \pm 0.10 \times 10^{5}$ |
| $\Delta f_{80} /\left(\operatorname{cps~} \mathrm{mm}^{-1}\right)$ | $2.63 \times 10^{5} \pm 0.065 \times 10^{5}$ |
| $\Delta f_{100} /\left(\mathrm{cps}_{\mu \mathrm{M}}{ }^{-1}\right)$ | $3.01 \times 10^{5} \pm 0.060 \times 10^{5}$ |
| $f_{0} /\left(\operatorname{cps} \mu \mathrm{M}^{-1}\right)$ | $7.01 \times 10^{6} \pm 0.347 \times 10^{6}$ |
| $k_{- \text {end }} / \mathrm{s}^{-1}$ | $1.7 \pm 2.3$ |
| $k_{\text {cat }} / \mathrm{s}^{-1}$ | $2.5 \pm 0.1$ |

Table S3. Full results of temperature dependence regression. $\mathrm{KF}_{0}$ refers to the expected KF concentration, while $\mathrm{KF}_{\text {active }}$ labels the estimated active amount. $\Delta f$ was fixed at $3.2 \times 10^{5} \mathrm{cps}^{\mu \mathrm{M}}{ }^{-1}$ from calibration.

|  | Temperature $/{ }^{\circ} \mathrm{C}$ |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Parameter | 25 | 20 | 15 | 10 | 5 |
| SoS | $4.2 \times 10^{10}$ | $8.55 \times 10^{10}$ | $1.19 \times 10^{11}$ | $6.9 \times 10^{11}$ | $3.5 \times 10^{10}$ |
| $\mathrm{DNA}_{\text {mod }}$ | $0.82 \pm 0.01$ | $0.85 \pm 0.01$ | $0.85 \pm 0.01$ | $0.75 \pm 0.01$ | $0.82^{b}$ |
| $k_{1} /\left(\mu \mathrm{M}^{-1} \mathrm{~s}^{-1}\right)$ | $5.5 \pm 2.6$ | $44.5 \pm 24.9$ | $16.7 \pm 7.4$ | $1.53 \pm 0.29$ | $3.54 \pm 0.98$ |
| $K_{\mathrm{d}} / \mathrm{nM}$ | $20 \pm 7.9$ | $5.6 \pm 1.8$ | $-a$ | $-a$ | $-a$ |
| $k_{\text {cat }} / \mathrm{s}^{-1}$ | $5.6 \pm 0.36$ | $3.3 \pm 0.1$ | $1.40 \pm 0.02$ | $1.00 \pm 0.02$ | $0.43 \pm 0.01$ |
| $\mathrm{KF}_{\text {active }} / \mathrm{KF}_{0}$ | $0.95 \pm 0.17$ | $0.53 \pm 0.04$ | $0.41 \pm 0.02$ | $0.36 \pm 0.03$ | $0.27 \pm 0.01$ |

${ }^{a}$ not existent in irreversible model. ${ }^{b}$ fixed to average of previous four values.

Table S4. Full results of dNTP dependence data regression including solvent KIE $_{\text {obs }}$ values. $\mathrm{k}_{- \text {end }}$ refers to the $k_{-1}$ value at the final $\mathrm{DNA}_{\mathrm{n}}$ position.

| T7pR80 DNA template | Value in $\mathrm{H}_{2} \mathrm{O}^{a}$ | Value in $\mathrm{D}_{2} \mathrm{O}$ | KIE |
| :--- | :--- | :---: | :---: |
| $k_{\text {obs }}$ |  |  |  |
| $K_{\mathrm{m}} / \mathrm{s}^{-1}$ | $\mathrm{\mu M}$ | $3.34 \pm 0.14$ | $1.131 \pm 0.056$ |
| $k_{\text {cat }} / K_{\mathrm{m}} /\left(\mathrm{\mu M}^{-1} \mathrm{~s}^{-1}\right)$ | $0.51 \pm 0.04$ | $0.224 \pm 0.027$ | $2.28 \pm 0.19$ |
| Sum of Squares | $6.52 \pm 0.55$ | $5.06 \pm 0.73$ | $1.29 \pm 0.22$ |
| $\Delta f /\left(\operatorname{cps} \mu \mathrm{M}^{-1}\right)$ | $(3.18 \pm 0.07) \times 10^{5}$ | $\left(3.23 \times 10^{11}\right.$ |  |
| $f_{0} /\left(\operatorname{cps} \mu \mathrm{M}^{-1}\right)$ | $(7.60 \pm 0.20) \times 10^{6}$ | $(8.25 \pm 0.14) \times 10^{5}$ |  |

${ }^{a}$ from Table S1

| T7pR30 DNA template | Value in $\mathrm{H}_{2} \mathrm{O}$ | Value in $\mathrm{D}_{2} \mathrm{O}$ | KIE $_{\text {obs }}$ |
| :--- | :---: | :---: | :---: |
| $k_{\text {cat }} / \mathrm{s}^{-1}$ | $2.63 \pm 0.12$ | $0.834 \pm 0.019$ | $3.15 \pm 0.16$ |
| $K_{\mathrm{m}} / \mu \mathrm{M}$ | $0.858 \pm 0.067$ | $0.398 \pm 0.034$ | $2.16 \pm 0.25$ |
| $k_{\text {cat }} / K_{\mathrm{m}} /\left(\mathrm{\mu M}^{-1} \mathrm{~s}^{-1}\right)$ | $3.06 \pm 0.16$ | $2.10 \pm 0.16$ | $1.46 \pm 0.13$ |
| Sum of Squares | $8.43 \times 10^{11}$ | shared |  |
| $\Delta f /\left(\operatorname{cps~}_{\mathrm{cpm}}{ }^{-1}\right)$ | $(6.17 \pm 0.06) \times 10^{5}$ | shared |  |
| $f_{0} /\left(\mathrm{cps}_{\mathrm{cm}} \mathrm{M}^{-1}\right)$ | $(7.03 \pm 0.13) \times 10^{6}$ | shared |  |
| $k_{- \text {end }} / \mathrm{s}^{-1}$ | $0.43 \pm 0.34$ | shared |  |

## Sample time courses from reduced model



Figure S1. Reduced model approximation of a $100-\mathrm{mer}$. Simulated sample time course with different model lengths at 200 nm DNA polymerase, 100 nm DNA template and $500 \mu \mathrm{~m}$ dNTP.

## Analysis of pulse-chase/pulse-quench data



Figure S2. Time course analysis of pulse-chase/pulse-quench yield gap for KF. The data and model used were published previously in: Dahlberg, M. E., and Benkovic, S. J. (1991) Kinetic mechanism of DNA polymerase I(Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant. Biochemistry 30, 4835-4843. The data points from Figure 2 in the publication were extracted with: Rohatgi, A. WebPlotDigitizer at http://arohatgi.info/WebPlotDigitizer/.
The model shown in Scheme V in their publication (similar to the first two rows of Scheme 1 including the bracketed part) was implemented in Copasi together with the published parameter values and conditions for the experiment. The dependent variable of the pulse-quench time course (blue) is modelled as the sum of all $\mathrm{DNA}_{n+1}$ concentrations, while the dependent variable of the pulse-chase (red) time course is modelled as the sum of all $\mathrm{DNA}_{n+1}$ concentrations plus the concentration of $\mathrm{KF}^{*}$ $\mathrm{DNA}_{\mathrm{n}}$-dNTP. The model does not fit the data when the published parameters are used. The fit above was obtained with $k_{3}=182 \mathrm{~s}^{-1}\left(\right.$ up from $\left.50 \mathrm{~s}^{-1}\right)$ and $k_{5}=0.29 \mathrm{~s}^{-1}$ (down from $15 \mathrm{~s}^{-1}$ ). While these numbers are unlikely to reflect true rate constants in fact, a number of alternative parameters can be changed to give the same result it does suggest that one of the pre-chemistry step could be occuring at a similar rate as phosphoryl transfer $\left(150 \mathrm{~s}^{-1}\right)$. Nevertheless, one of the post-chemistry steps must be slow in order to describe these data.

## Mathematica code

This code can be used to produce the Scheme 1 model for a DNA template of arbitrary length.
$\mathrm{tf}=120 ;$
species $=P P ;$
reallength $=$.;
modellength $=$.
(* model with length adjustments in this cell ${ }^{*}$ )
poly[length_]:=\{
$E n z '[t]==$
Total[
Table $\left[k_{-1} * \mathrm{ET}_{i}[t]\right.$
$-k_{1} * T_{i}[t] * \operatorname{Enz}[t],\{i, 0$, length -1$\left.\}\right]$
]
$+k_{- \text {end }} * \mathrm{ET}_{\text {length }}[t]$
$-k_{\text {end }} * T_{\text {length }}[t] * \operatorname{Enz}[t]$,

Table $\left[T_{i}^{\prime}[t]==\right.$
$k_{-1} * \mathrm{ET}_{i}[t]$
$-k_{1} * T_{i}[t] * \operatorname{Enz}[t],\{i, 0$, length -1$\left.\}\right]$,

$$
\begin{aligned}
& T_{\text {length }}^{\prime}[t]== \\
& k_{- \text {end }} * \mathrm{ET}_{\text {length }}[t] \\
& -k_{\text {end }} * T_{\text {length }}[t] * \operatorname{Enz}[t]
\end{aligned}
$$

(* ET start ${ }^{*}$ )
$\mathrm{ET}_{0}^{\prime}[t]==$
$k_{1} * T_{0}[t] * \mathrm{Enz}[t]$
$-k_{-1} * \mathrm{ET}_{0}[t]$
$-\mathrm{kcat} * \mathrm{ET}_{0}[t] * 1 / 4 S[t] /(\mathrm{Km}+1 / 4 S[t]) *$ length $/$ reallength,
(*repeat*)
Table $\left[\mathrm{ET}_{i}^{\prime}[t]==\right.$
$k_{1} * T_{i}[t] * \operatorname{Enz}[t]$
$-k_{-1} * \mathrm{ET}_{i}[t]$
-kcat $* \mathrm{ET}_{i}[t] * 1 / 4 S[t] /(\mathrm{Km}+1 / 4 S[t]) *$ length/reallength
$+\mathrm{kcat} * \mathrm{ET}_{i-1}[t] * 1 / 4 S[t] /(\mathrm{Km}+1 / 4 S[t]) *$ length $/$ reallength,
$\{i, 1$, length -1$\}]$,
(*finish*)

$$
\begin{aligned}
& \mathrm{ET}_{\text {length }}^{\prime}[t]== \\
& k_{\text {end }} * T_{\text {length }}[t] * \mathrm{Enz}[t] \\
& -k_{- \text {end }} * \mathrm{ET}_{\text {length }}[t] \\
& + \text { kcat } * \mathrm{ET}_{\text {length }-1}[t] * 1 / 4 S[t] /(\mathrm{Km}+1 / 4 S[t]) * \text { length/reallength }
\end{aligned}
$$

$$
S^{\prime}[t]==
$$

$$
\operatorname{Sum}\left[-\mathrm{kcat} * \mathrm{ET}_{i}[t] * 1 / 4 S[t] /(\mathrm{Km}+1 / 4 S[t]),\{i, 0, \text { length }-1\}\right]
$$

$$
\operatorname{PP}^{\prime}[t]==
$$

$$
\operatorname{Sum}\left[\operatorname{kcat} * \mathrm{ET}_{i}[t] * 1 / 4 S[t] /(\mathrm{Km}+1 / 4 S[t]),\{i, 0, \text { length }-1\}\right],
$$

$\operatorname{Enz}[0]==\mathrm{enz} *$ Eerror,
$T_{0}[0]==$ dna $*$ Terror,
$S[0]==\mathrm{sub}$,
$\operatorname{PP}[0]==0$,
Table $\left[T_{i}[0]==0,\{i\right.$, length $\left.\}\right]$,
Table $\left[\mathrm{ET}_{i}[0]==0,\{i, 0\right.$, length $\left.\}\right]$
\}//Flatten;
solaccurate $=$ Table[
ParametricNDSolve[
poly[lengths],
species,
$\{t, 0 ., \mathrm{tf}\}$,
$\left\{k_{1}, k_{-1}, \mathrm{kcat}, \mathrm{Km}, k_{\text {end }}, k_{-\mathrm{end}}, \mathrm{enz}, \mathrm{dna}, \mathrm{sub}\right.$, Eerror, Terror $\}$,
MaxSteps $\rightarrow \infty$,
PrecisionGoal $\rightarrow 9$,
AccuracyGoal $\rightarrow 9$,
Method $\rightarrow$ "LSODA"],
\{lengths, $\{30,50,80,100\}\}$ ];

