Supporting Information

RB69 DNA Polymerase Mutants with Expanded Nascent Base-Pair Binding Pockets are Highly Efficient but have Reduced Base Selectivity†

Hong Zhang[‡§], Jeff Beckman[‡§], Jimin Wang[‡], William Konigsberg*[,‡]

[§] *These authors contributed equally to this paper.*

[‡] *Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA.*

* Corresponding Author: william.konigsberg@yale.edu

Averaging the $K_{d,app}$ and k_{pol} data for correct and incorrect dNTPs. The wt exhibits only small differences in $K_{d,app}$ and in the maximum incorporation rate (k_{pol}) for correct dNTPs (Table 1 of the main text). When the $K_{d,app}$ values were averaged for all four correct dNTPs, the resulting value was 57 μM with a standard deviation (σ) of about 20% from the mean. The average k_{pol} value was 210 s^{-1} (σ of 20%). The average incorporation efficiency, $k_{pol}/K_{d,app}$, was 4.1 $s^{-1}\mu M^{-1}$ (σ of 40%).

All four mutant RB69 pols had similar kinetic parameters for correct incorporation as wt (Table 1). They had an average $K_{d,app}$ of 46 μM (σ of 30%) and an average k_{pol} of 210 s^{-1} (σ of 30%) (assuming 300 s^{-1} for each k_{pol} involving the triple mutant). These values are comparable to those of the wt (Table 1 of the main text), suggesting that there is likely to be a common mechanism with similar kinetic parameters for the incorporation of all four correct nucleotides by wt RB69 pol and all of its NBP mutants. This was supported by the stopped-flow fluorescence data mentioned in the main text of the paper.

The kinetic parameters for incorporation of incorrect nucleotides by the wt varied over a wide range. For example, the $K_{d,app}$ values for incorporation of incorrect dNMPs varied from 730 μM for the dGTP/dT pair to > 2,000 μM for several dNTP/dN mispairs. The k_{pol} values among the 12 mispairs, those that could be calculated, varied by 40 fold, ranging from 0.003 s^{-1} for dGTP/dA to 0.13 s^{-1} for dATP/dA. The wt catalytic efficiencies, calculated from the $k_{pol}/K_{d,app}$ ratios, varied from $3x10^{-7}$ $s^{-1}\mu M^{-1}$ for dCTP/dC to $1.7x10^{-4}$ $s^{-1}\mu M^{-1}$ for dCTP/dA. This 1,000-fold range in misincorporation efficiency strongly suggests that the wt has distinct features that restrict incorporation of certain mispairs. To circumvent these constraints, incorrect dNMPs are likely to be incorporated by nucleotidyl transfer pathways that differ in subtle ways from those of correct dNMPs as proposed for T7 DNA polymerase.

When differences in the kinetic parameters for incorporation of incorrect dNMPs were examined using the single, double, and triple mutants, the greatest variation was seen in k_{pol} values which

increased progressively from the single to the triple mutant. For example, the dTTP/dG mispair had a k_{pol} value of 0.05 s^{-1} with the L561A single mutant, a value of 1.7 s^{-1} with the L561A/Y567A double mutant, and a k_{pol} of 8.7 s^{-1} with the L561A/S565G/Y567A triple mutant (Table 1). Comparison of the k_{pol} values for other mispairs with the single, double, and triple mutants showed the same trend (Table 1). However, the variations in the $K_{d,app}$ values for incorrect dNTPs with the different mutants were much less pronounced, having a range of only 6 fold, falling between 200 and 1200 μM. The dTTP/dG mispair is the notable exception where an average $K_{d,app}$ of 45 μM was obtained, a value close to that of the correct dCTP/dG pair when using wt (70).

In contrast to the averaged base discrimination exhibited by the NBP mutants, the base selectivity varies considerably among the different sets of mispairs. In the main text of the paper this was shown in a bar graph that encompassed all of the data. A simpler representation of a subset of these results is shown here in Fig. S1.

Correlation between increased NBP volume, ΔV, and the change in free energy, ΔΔG, for nucleotidyl transfer. We investigated whether there was a correlation between the change in NBP volume among NBP mutants and the change in free energy for incorporating a correct versus an incorrect dNMP. With respect to wt RB69 pol, the change in NBP volume ΔV is 52, 156, 171, 327, and 379 $Å^3$ for the S565G, L561A, Y567A, L561A/Y567A, and L561A/S565G/Y567A

Figure S1. Relating incorporation Efficiencies ($k_{pol}/K_{d,app}$) for Correct and Incorrect Bases by Wild Type RB69 pol and Four Mutants as Estimated from Pre-Steady-State Chemical-Quench. Note the different scales on the Y axes.

mutants, respectively, using the surface residue assumption (*1-3*). From our kinetic data (Table 1), plots of log (k_{pol}) or log (k_{pol}/K_{dapp}) for all 12 possible mispairs versus ΔV revealed a steady increase in value as ΔV increased. The graph of log (K_{dapp}) versus ΔV however showed only a small increase (data not shown), consistent with the fact that K_{dapp} values did not change substantially as the number of substitutions increased (with few exceptions).

These data were used to calculate the effects of the increased NBP volume on the overall K_{dapp} and transition state free energy barriers for nucleotidyl transfer. For each enzyme, the two components were calculated by averaging the free energy changes obtained from dividing the kinetic parameters K_{dapp} or k_{pol} (the latter to relate transition states) for incorporation of each of the four correctly paired dNTPs by each of the three incorrect dNTP pairs for a total of twelve values and then fitting these values to the equations: $\Delta\Delta G_T = \Sigma[RT\ln(K_{d,app,correct}/K_{d,app,incorrect})]/12$ and $\Delta\Delta G_K = \Sigma[RT\ln(k_{pol,correct}/k_{pol,incorrect})]/12$. When $\Delta\Delta G_T$, $\Delta\Delta G_K$, and $\Delta\Delta G = \Sigma[RT\ln[(k_{pol}/K_{d,app,correct})/(k_{pol}/K_{d,app,incorrect})]]/12$ were plotted against ΔV, going from the wt to the triple mutant, the relationships were linear (Fig. S2). The slopes of ΔΔG and $\Delta\Delta G_K$ versus ΔV were similar, – 0.011 and – 0.009 kcal/mole/Å^3 with least squares line-fitting correlation coefficients of 0.99 and 0.98, respectively, whereas the slope of $\Delta\Delta G_T$ was relatively shallow (–0.0016 kcal/mole/Å^3) with a poor line-fit correlation coefficient of 0.69 (Fig. S2). The correlation for the $\Delta\Delta G_T$ plot was relatively poor because binding of a dNTP includes a complex set of

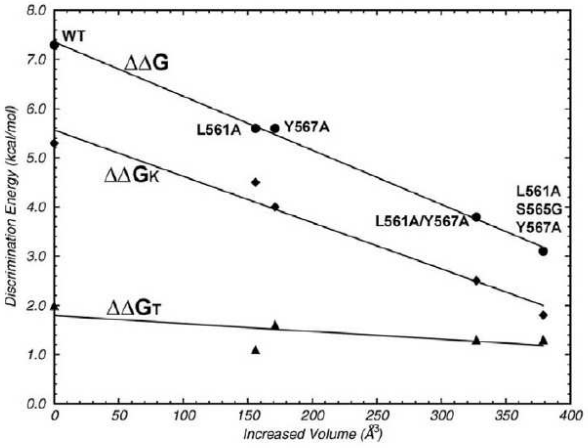

Figure S2. Plots of free energy change (calculated as an average of all base pairing combinations, see text for details) versus increased NBP volume. $\Delta\Delta G_K$ represents change in the transition-state free energy required to form product (calculated from k_{pol} values), $\Delta\Delta G_T$ represents change in free energy upon binding of a dNTP to form an active ternary complex (calculated from $K_{d,app}$ values), and ΔΔG is the total free energy change (calculated from $k_{pol}/K_{d,app}$ values).

interactions primarily involving shape recognition, altered hydrogen bonding between mispaired bases, and nascent base stacking of the incoming dNTP with its 5' neighbor in the primer strand.

Overall, the graphs show that substitution of Y567, S565, and/or L561 in the wt enzyme removed features that increased catalytic energy barriers for misincorporation. Substitution of Y567 or L561 with Ala lowered energy barriers for incorporation of incorrect dNMPs by an average of 1.7 kcal/mole, an additional 1.8 kcal/mole reduction after converting single mutants to the double mutant (L561A/Y567A), and a further reduction of 0.7 kcal/mole after conversion of the double to the triple mutant (Fig. S2).

The triple mutant is unique in exhibiting two phases during a single turnover progress curve. As mentioned in the main text, product formation catalyzed by the triple mutant fit best to a double exponential when the assays were carried out under single turnover conditions (Fig. 3 of the main text). This was the case for all the correct dNTPs as well as for some other dNTP/dN mispairs. The deviation from a single exponential for incorporation of other incorrect dNMPs with slower k_{pol} values ($< 5\ s^{-1}$) was less pronounced.

When 20-fold excess unlabeled primer-template (P/T) was included with the incoming dNTP prior to mixing with the enzyme:^{32}P-labeled P/T in the rapid chemical quench instrument, the resulting progress curves fit a single exponential for all the assays involving the triple mutant (Fig. 3B,C of the main text). Neither the wt, nor the single or double mutants required double exponentials to fit the experimental data points even when unlabeled primer-template was omitted. We rationalize the results with the triple mutant as follows. When the triple mutant encounters the P/T, which is in limiting concentrations, the enzyme binds a percentage of the P/T in a productive fashion so that, upon introduction of an incoming correct dNTP, product formation ensues. However, a portion of the primer-template (close to 50%) binds to the enzyme to give a "dead end" binary complex. The enzyme has to

dissociate from the DNA to allow another enzyme molecule to bind the DNA productively. The rate of the second phase of the exponential is about 5 s^{-1}, which is comparable to the rate of dissociation of the enzyme from the P/T and is rate limiting. This rate is comparable to that observed with T4 pol where the P/T:enzyme dissociation rate was measured directly (*4*). We do not know why this happens only with the triple mutant but not with any of the other NBP mutants. We can only infer that the extra space in the NBP of the triple mutant allows the primer-template to bind in a way that does not lead to product formation even if an incoming dNTP is present. Another possibility is that a small fraction of the primer-template binds to the triple mutant in the exo rather than in the pol mode and that switching from the exo to the pol mode occurs at the rate of ~ 5 s^{-1}, the rate found for exo to pol switching with T4 pol (*4*). Arguing against this is the excess DNA trap experiment where the inclusion of excess unlabeled P/T trapped all free enzyme before it could rebind the ^{32}P-labeled P/T. If transfer of the primer-template from the exo to the pol mode occurred without dissociation of the enzyme:^{32}P-labeled primer-template complex, then the inclusion of unlabeled P/T would not have affected the progress curve for product formation.

REFERENCES

1. Gerstein, M., Tsai, J., and Levitt, M. (1995) The volume of atoms on the protein surface: calculated from simulation, using Voronoi polyhedra, *J. Mol. Biol. 249*, 955-966.
2. Richards, F. M. (1974) The interpretation of protein structures: total volume, group volume distributions and packing density, *J. Mol. Biol. 82*, 1-14.
3. Chothia, C. (1976) The nature of the accessible and buried surfaces in proteins, *J. Mol. Biol. 105*, 1-12.

4. Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., and Benkovic, S. J. (1992) Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4, *Biochemistry 31*, 10984-10994