

# Nucleic acid polymerases use a general acid for nucleotidyl transfer

Christian Castro<sup>1,4</sup>, Eric D Smidansky<sup>1,4</sup>, Jamie J Arnold<sup>1</sup>, Kenneth R Maksimchuk<sup>1</sup>, Ibrahim Moustafa<sup>1</sup>, Akira Uchida<sup>1</sup>, Matthias Götte<sup>2</sup>, William Konigsberg<sup>3</sup> & Craig E Cameron<sup>1</sup>

**Nucleic acid polymerases catalyze the formation of DNA or RNA from nucleoside-triphosphate precursors. Amino acid residues in the active site of polymerases are thought to contribute only indirectly to catalysis by serving as ligands for the two divalent cations that are required for activity or substrate binding. Two proton-transfer reactions are necessary for polymerase-catalyzed nucleotidyl transfer: deprotonation of the 3'-hydroxyl nucleophile and protonation of the pyrophosphate leaving group. Using model enzymes representing all four classes of nucleic acid polymerases, we show that the proton donor to pyrophosphate is an active-site amino acid residue. The use of general acid catalysis by polymerases extends the mechanism of nucleotidyl transfer beyond that of the well-established two-metal-ion mechanism. The existence of an active-site residue that regulates polymerase catalysis may permit manipulation of viral polymerase replication speed and/or fidelity for virus attenuation and vaccine development.**

Nucleic acid polymerases catalyze the formation of DNA or RNA from 2'-deoxyribonucleotides or ribonucleotides, respectively. Polymerases are therefore required for the reproduction, maintenance and expression of the genomes of all organisms, including viruses. Nucleotidyl transfer, the chemical reaction catalyzed by polymerases, is shown in **Figure 1**. Nucleophilic attack on the  $\alpha$ -phosphorous atom of the (2'-deoxy)ribonucleoside triphosphate by the primer 3'-hydroxyl leads to formation of a phosphodiester bond and release of pyrophosphate. All polymerases require two divalent cations, usually  $Mg^{2+}$ , for activity and use a two-metal-ion mechanism for nucleotidyl transfer<sup>1,2</sup> (**Fig. 1**). Metal A lowers the  $pK_a$  of the primer 3'-hydroxyl, thus facilitating deprotonation of this moiety for in-line nucleophilic attack<sup>3</sup>. Metal B orients the triphosphate for catalysis, stabilizes the negative charge that arises during formation of the pentavalent transition state and has been suggested to assist pyrophosphate release<sup>1,3</sup>. Although it is clear that deprotonation of the primer 3'-hydroxyl is required for catalysis<sup>4,5</sup>, only recently has it been suggested that protonation of the pyrophosphate leaving group occurs before its release from the enzyme<sup>4</sup>. The acceptor of the 3'-hydroxyl proton and donor of pyrophosphate proton are not known. It is now clear that chemistry can be at least partially rate limiting for nucleotide addition by all classes of nucleic acid polymerases<sup>4,6</sup>, as well as serving as a fidelity checkpoint<sup>6-8</sup>. Identification of the acceptor and donor for these proton-transfer reactions may inspire new mechanistic hypotheses for how catalytic efficiency can be tuned by the nature (correct versus incorrect) of the bound nucleotide.

Four classes of template-dependent nucleic acid polymerases exist: RNA-dependent RNA polymerase (RdRp); RNA/DNA-dependent DNA polymerase, the so-called reverse transcriptase (RT); DNA-dependent DNA polymerase (DdDp); and DNA-dependent RNA polymerase (DdRp). In this study, we have used the RdRp from poliovirus (PV), the RT from human immunodeficiency virus type 1 (HIV), the DdDp from bacteriophage RB69 and the DdRp from bacteriophage T7 as representatives of the four classes of polymerases because of the wealth of mechanistic and/or structural information available for these enzymes<sup>9-21</sup>. The objective of this study was to identify the proton donor to the pyrophosphate leaving group, as analysis of high-resolution structures of several polymerases poised for or undergoing catalysis implicated a basic amino acid residue rather than a water molecule as the proton donor. In addition, analysis of the pH dependence of nucleotide addition by PV RdRp was consistent with catalysis being dependent on a residue with a  $pK_a$  value of 10.5 (ref. 4).

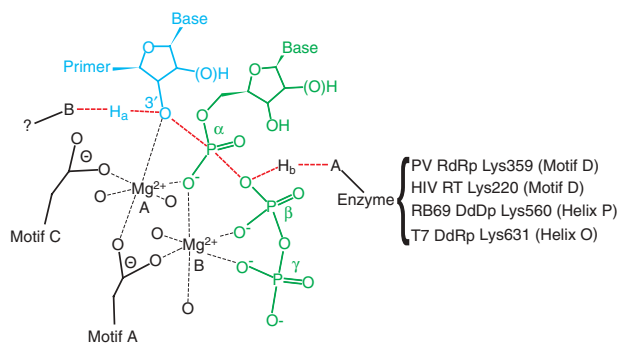
## RESULTS

### A general acid in nucleic acid polymerase catalysis

Analysis of the structural model for the RdRp from Norwalk virus (NV) in complex with primed template RNA and nucleotide showed that Lys374 of conserved structural motif D was located in the vicinity of the triphosphate moiety of the incoming nucleotide<sup>22</sup> (**Supplementary Figs. 1a,b** online). Structural motif D is conserved between RdRps and RTs and has no defined catalytic function<sup>23</sup>. Analysis of structure-based sequence alignments showed only one conserved

<sup>1</sup>Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, USA. <sup>2</sup>Department of Microbiology & Immunology, McGill University, Montreal, Quebec H3A 2B4, Canada. <sup>3</sup>Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, USA. <sup>4</sup>These authors contributed equally to this work. Correspondence should be addressed to C.E.C. (cec9@psu.edu).

Received 12 August 2008; accepted 26 November 2008; published online 18 January 2009; doi:10.1038/nsmb.1540



**Figure 1** Extending the two-metal-ion mechanism of nucleotidyl transfer to include general acid catalysis. Nucleoside triphosphate (green) enters the active site with a divalent cation ( $Mg^{2+}$ , metal B). This metal ion is coordinated by the phosphates of the nucleotide, an aspartate residue located in structural motif A of all polymerases, and probably water molecules (indicated as oxygen ligands to metal without specific designation). Metal B orients the triphosphate in the active site and may contribute to charge neutralization during catalysis. A second divalent cation binds ( $Mg^{2+}$ , metal A) that is coordinated by the 3'-hydroxyl of the primer terminus (cyan), the nucleotide  $\alpha$ -phosphate and aspartate residues of structural motifs A and C. Metal A lowers the  $pK_a$  of the 3'-hydroxyl, facilitating deprotonation and subsequent nucleophilic attack at physiological pH. As the transition state of nucleotidyl transfer is approached (indicated by dashed red lines), the primer 3'-hydroxyl proton,  $H_a$ , is transferred to an unidentified base (B), and we propose that the pyrophosphate leaving group is protonated ( $H_b$ ) by a basic amino acid on the enzyme. The positions of the candidate general acids of the model polymerases used in this study are shown.

lysine residue in motif D of RdRps and RTs, including the telomere RT, telomerase (TERT) (**Supplementary Fig. 1c**). These sequence alignments identified Lys359 of PV RdRp and Lys220 of HIV RT in conserved structural motif D as candidates for the proton donors in these systems (**Supplementary Fig. 1c**). Lys219 of HIV RT was ruled out as the putative general acid because of the absence of this residue in the RT from human immunodeficiency virus type 2 (HIV-2) (**Supplementary Fig. 1c**)—it is well established that the DNA polymerase activity of HIV-2 RT is on par with that observed for HIV RT<sup>24</sup>. Unlike Lys374 of NV RdRp, Lys359 of PV RdRp and Lys220 of HIV RT were not oriented in a position to interact readily with the triphosphate moiety of the incoming nucleotide (**Fig. 2**).

Structural motif D is one of the most dynamic elements of the palm subdomain of RdRps and RTs, with the position of this motif varying by as much as 6 Å when structures are compared (**Supplementary Fig. 1b**). Solution of the structure of HIV RT in complex with primed template DNA and nucleotide required cross-linking of the enzyme to DNA, which could have influenced the orientation of motif D<sup>14</sup>. This possibility is supported by a recently published model for the TERT elongation complex that positions the motif D lysine in a position to serve as a general acid<sup>25</sup> (**Fig. 2**). Therefore, only biochemical studies could be used to test the possibility that the conserved lysine in motif D functions as a general acid. Identification of candidates for the general acid in DNA-dependent DNA and RNA polymerases was more straightforward. Structural models show clearly that Lys560 in helix P of RB69 DdDp, a B family polymerase<sup>12</sup>, and Lys631 in helix O of T7 DdRp, an A family polymerase<sup>21</sup>, are positioned to serve as proton donors (**Fig. 2**) for these enzymes.

For each of the polymerases described above, we produced derivatives in which the candidate lysine proton donor was replaced with a leucine residue (**Supplementary Methods** online). We chose leucine

instead of alanine to prevent water from binding to the site and serving as a proton donor (**Supplementary Fig. 2** online). For all polymerases tested, the maximal rate constant for nucleotide incorporation ( $k_{pol}$ ) was reduced by 50-fold to 2,000-fold for the leucine derivative relative to the wild-type, lysine-containing enzyme (**Table 1**). This observation is consistent with a role for the lysine in the rate-limiting step for nucleotide incorporation. For all of the wild-type polymerases used here, chemistry is at least partially rate limiting<sup>4</sup>. With the exception of the leucine derivative of T7 DdRp, the  $k_{pol}$  values for each polymerase derive from experiments that included a saturating nucleotide concentration (five times the apparent dissociation constant ( $K_{d,app}$ )) (**Supplementary Methods**). In the case of the leucine derivative of T7 DdRp, the highest concentration of nucleotide attainable was two times the  $K_{d,app}$  (**Supplementary Methods**). This circumstance increases the error on the  $k_{pol}$  value measured; however, our ability to reach conclusions for the leucine derivative is not affected, because the observed 100-fold reduction in  $k_{pol}$  value for this derivative is much greater than the error of the measurement.

We did not observe a substantial difference in the  $K_{d,app}$  for the nucleotide substrate measured for the leucine derivatives of PV RdRp or HIV RT (**Table 1**). This observation is consistent with residues in the conserved, structural motif F of these enzymes functioning independently in triphosphate binding<sup>26</sup>. In contrast, the  $K_{d,app}$  value for nucleotide substrate measured for the leucine derivatives of RB69 DdDp and T7 DdRp increased 25-fold and 167-fold, respectively (**Table 1**). This observation is consistent with structural studies showing that these lysines interact with the nucleotide substrate<sup>18,20</sup> (**Fig. 2**).

Unlike with many other polymerases, the stability of PV RdRp on its nucleic acid primer template is not affected by pH; this allows us to interpret the pH dependence of the kinetics of nucleotide incorporation<sup>4</sup> (**Supplementary Methods**). Consistent with Lys359 of PV RdRp functioning as a proton donor during nucleotidyl transfer, the descending limb of the pH rate profile observed for wild-type PV RdRp<sup>4</sup> (**Fig. 3a** and **Supplementary Fig. 3** online) was lost in the profile for the K359L derivative (**Fig. 3a** and **Supplementary Fig. 3**). The ionization observed for the K359L derivative probably reflects deprotonation of the 3'-hydroxyl (**Fig. 3a**). Theoretical studies of nucleotidyl transfer by rat DNA polymerase- $\beta$  have suggested that the primer 3'-hydroxyl has a  $pK_a$  in the range of 8–9.5 (ref. 7). Because the  $pK_a$  values measured are kinetic  $pK_a$  values, the lack of equivalence in the  $pK_a$  value for the K359L derivative to any measured for the wild-type enzyme probably reflects a change in the rate-determining step for nucleotidyl transfer<sup>4</sup>. Only the chemistry should be rate limiting for the K359L derivative, whereas a conformational-change step and chemistry are equally rate limiting for the wild-type enzyme<sup>10</sup>.

### Loss of the general acid leads to single proton transfer

Theoretical studies of the free-energy landscape for phosphoryl transfer almost always show proton transfer in the transition state<sup>27</sup>, which would manifest experimentally as a solvent deuterium kinetic isotope effect<sup>4</sup>. The solvent deuterium kinetic isotope effect is defined as the ratio of  $k_{pol}$  values obtained when the reaction is performed in  $H_2O$  relative to values obtained when the reaction is performed in  $D_2O$ <sup>28</sup> (**Supplementary Methods**). The observation of a solvent deuterium kinetic isotope effect supports a proton-transfer reaction representing one of the microscopic steps reflected in the macroscopic rate constant,  $k_{pol}$ . Each wild-type polymerase showed a solvent deuterium kinetic isotope effect of 2–5 (**Table 1**). This observation is consistent with chemistry contributing to the rate-limiting step(s) reported by the nucleotide-incorporation assay, as observed for the

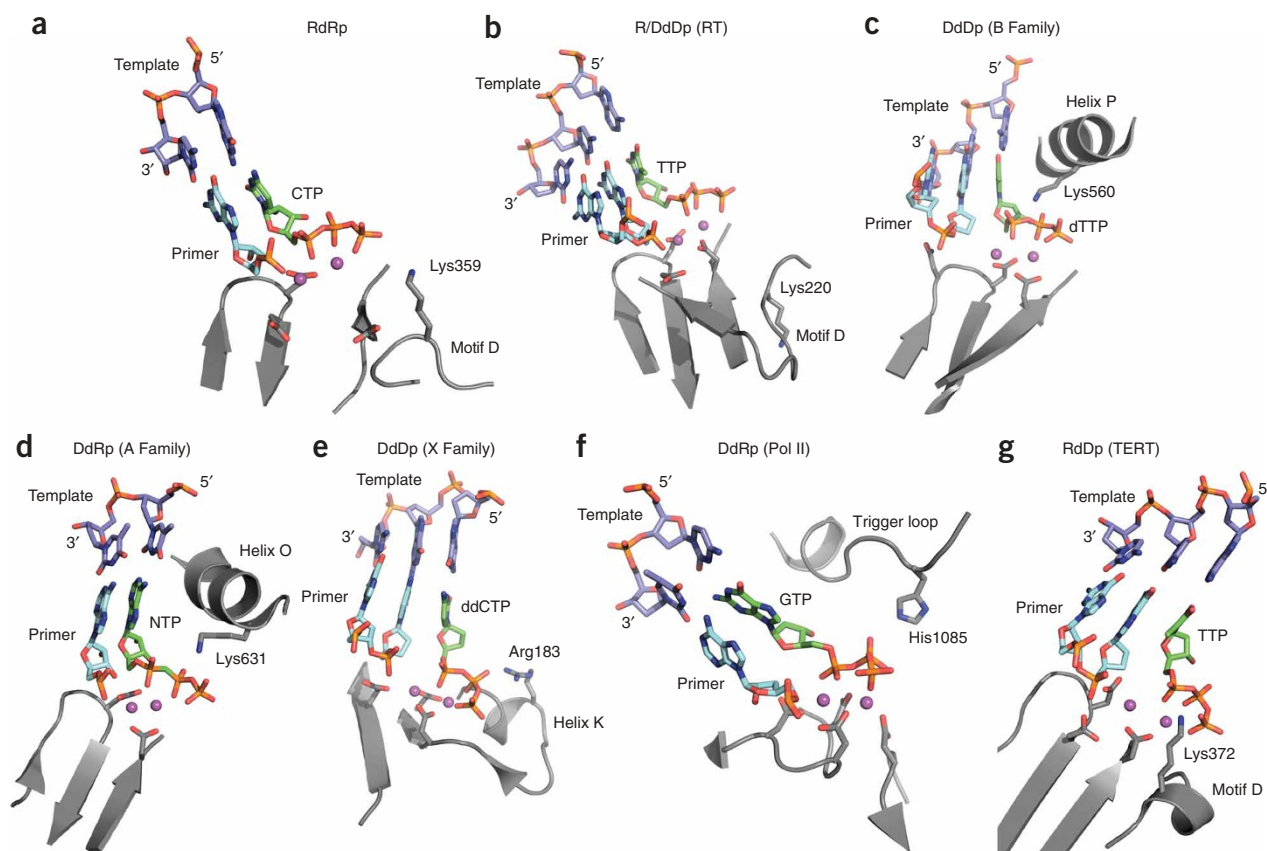
wild-type enzymes<sup>4</sup>. Notably, the solvent deuterium kinetic isotope effect measured for each leucine derivative was smaller than that observed for the corresponding wild-type enzyme (Table 1). This observation could be interpreted in one of two ways. First, it is possible that the extent to which chemistry limits nucleotide addition by the leucine derivatives is less than occurs for the wild-type enzymes. Second, it is possible that fewer protons are being transferred. This latter possibility would be consistent with the candidate lysine residues contributing one proton transfer during nucleotide incorporation.

To count the number of protons being transferred during the nucleotidyl transfer reaction, we performed a proton-inventory experiment<sup>4,29</sup> (Supplementary Methods). This experiment obtains  $k_{\text{pol}}$  values in the presence of different mole fractions of D<sub>2</sub>O ( $k_n$ ). The ratio of  $k_n / k_0$  ( $k_0$  is  $k_{\text{pol}}$  in H<sub>2</sub>O) is plotted as a function of the mole fraction of D<sub>2</sub>O ( $n$ ). If a single proton is transferred during the reaction, then the data should fit to a line. We carried out this experiment for all leucine derivatives except the T7 DdRp derivative, which could not be saturated with nucleotide. The data obtained for all of the leucine derivatives evaluated fit well to a straight line

(Fig. 3b), consistent with a single proton-transfer reaction occurring during nucleotidyl transfer for these enzymes. In contrast, the data for the corresponding wild-type enzymes failed to fall on a line defined by  $k_n / k_0$  values at  $n = 0$  and  $n = 1$ , consistent with a model in which more than one proton-transfer reaction occurs during nucleotidyl transfer (Fig. 3b). More extensive analysis of the proton-inventory experiment for the wild-type enzymes, including statistical analysis, was reported previously, convincingly demonstrating that two proton-transfer reactions occur in the transition state for nucleotidyl transfer by the wild-type enzymes used here<sup>4</sup>. These data provide compelling evidence that an active-site amino acid residue is used as a general acid to protonate the pyrophosphate leaving group and facilitate nucleotidyl transfer.

### General acid catalysis in the presence of Mn<sup>2+</sup>

All of the experiments described to this point were performed in the presence of Mg<sup>2+</sup>, because this divalent cation is considered to be the biologically relevant cofactor. In the presence of Mg<sup>2+</sup>, chemistry is only partially rate limiting for PV RdRp at pH 7.5 (ref. 10). In



**Figure 2** Interactions of NTP in the active sites of various polymerase families. In all ternary complexes, the NTP (green carbon atoms), primer and template (cyan and purple carbon atoms, respectively) and key active-site residues (gray carbon atoms) are shown as sticks; the metal ions are shown as magenta spheres. (a) The RdRp from PV (PDB 1RA6)<sup>19</sup> with the primer and template and CTP bound to the active site in the presence of two Mn<sup>2+</sup> ions. The primer and template, CTP and Mn<sup>2+</sup> ions were extracted from the structural complex of Norovirus (NV) polymerase (PDB 3BSO)<sup>22</sup> and modeled into the active site of PV RdRp. (b) HIV RT with TTP bound to the active site in the presence of two Mg<sup>2+</sup> ions (PDB 1RTD)<sup>14</sup>. (c) The RB69 DNA polymerase with the dTTP nucleotide bound in the presence of Ca<sup>2+</sup> ions (PDB 1IG9)<sup>12</sup>. (d) The T7 RNA polymerase with the modified nucleotide  $\alpha\beta$ -methylene-ATP bound in the presence of Mg<sup>2+</sup> (PDB 1S76)<sup>21</sup>. (e) DNA polymerase- $\beta$  with the ddCTP bound to the active site in the presence of Mg<sup>2+</sup> ions (PDB 1BPY)<sup>50</sup>. (f) Multisubunit RNA polymerase II from yeast (PDB 2E2H)<sup>31</sup> in complex with nucleic acids and GTP substrate in the presence of Mg<sup>2+</sup>. (g) The catalytic subunit of telomerase with TTP bound to the active site in the presence of two Mg<sup>2+</sup> ions (PDB 3DU5)<sup>25</sup>. Primer and template, TTP and Mg<sup>2+</sup> ions were extracted from the structural complex of HIV RT (PDB 1RTD)<sup>14</sup> and modeled into the active site of TERT as described previously<sup>25</sup>. Residues equivalent to Lys359 of PV RdRp predicted to function as general acids are labeled. The designations of the structural elements on which the general acid is located are also indicated explicitly.

**Table 1 Kinetic analysis of PV RdRp, HIV-1 RT, RB69 DdDp and T7 DdRp supports general acid catalysis in nucleotidyl transfer**

Parameter measured	PV RdRp		HIV RT		RB69 DdDp		T7 DdRp	
	WT	K359L	WT	K220L	WT	K560L	WT	K631L
$k_{\text{pol}}$ ( $\text{s}^{-1}$ )	50 ± 5	1 ± 0.1	60 ± 5	0.3 ± 0.1	200 ± 10	0.10 ± 0.01	60 ± 5	0.6 ± 0.1
$K_{\text{d,app}}$ ( $\mu\text{M}$ ) <sup>a</sup>	200 ± 20	700 ± 80	7 ± 1	5 ± 2	40 ± 5	1000 ± 100	300 ± 30	5.0 ± 1.0 × 10 <sup>4c</sup>
SDKIE <sup>b</sup>	3.0 ± 0.3	2.5 ± 0.3	2.2 ± 0.4	1.8 ± 0.4	4.2 ± 0.2	1.8 ± 0.2	5.2 ± 0.5	2.6 ± 0.5 <sup>c</sup>
PI <sup>d</sup>	2	1	2	1	2	1	2	n.d. <sup>e</sup>

<sup>a</sup> $K_{\text{d,app}}$  is for (d)ATP (Supplementary Methods). <sup>b</sup>SDKIE is the solvent deuterium kinetic isotope effect<sup>28</sup>, calculated as  $k_{\text{obs}}$  in H<sub>2</sub>O /  $k_{\text{obs}}$  in D<sub>2</sub>O at saturating (d)ATP concentration. <sup>c</sup> $K_{\text{d,app}}$ ,  $k_{\text{pol}}$  and SDKIE values listed for T7 K631L were obtained by using data collected with 80 mM ATP, a subsaturating concentration. <sup>d</sup>PI is the proton inventory<sup>29</sup>, calculated from a plot of  $k_n / k_0$  as a function of  $n$ . The data were fitted to a modified Gross-Butler equation for either a two-proton-transfer model (equation (1)) or a one-proton-transfer model (equation (2)). The value reported is the proton-transfer model that best fits the data. <sup>e</sup>n.d., not determined.

addition, the extent to which chemistry contributes to the observed rate constant for nucleotide addition changes as a function of pH<sup>4</sup>. In the presence of Mn<sup>2+</sup>, however, chemistry is the primary determinant of the rate constant for nucleotide addition by PV RdRp from pH 6.0 to pH 9.0 (ref. 4). Unfortunately, experiments cannot be performed above pH 9.0 in the presence of Mn<sup>2+</sup> owing to precipitation of the metal hydroxide.

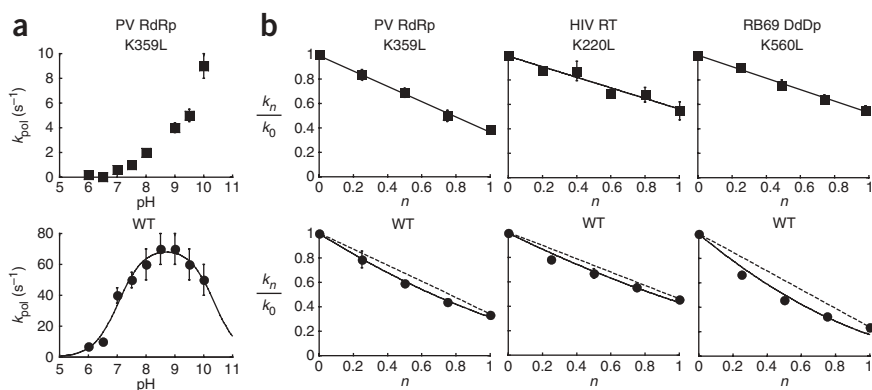
We evaluated the activity of the K359L derivative of PV RdRp in the presence of Mn<sup>2+</sup> (Supplementary Fig. 4 online). The observed rate constant for nucleotide incorporation was reduced to 1 s<sup>-1</sup>, ten-fold lower than that of the wild-type enzyme, with no substantial change in the apparent dissociation constant for nucleotide (Supplementary Fig. 4a). The phosphorothioate effect observed for the K359L derivative (6 ± 1) was essentially the same as that measured for wild-type enzyme (7 ± 1), consistent with chemistry remaining as the rate-limiting step (Supplementary Fig. 4a). The solvent deuterium kinetic isotope effect observed for the K359L derivative of PV RdRp was 3 ± 1—more than two-fold lower than the value of 7 ± 1 observed for the wild-type enzyme (Supplementary Fig. 4a). These data are consistent with the loss of a proton-transfer reaction, as observed in the presence of Mg<sup>2+</sup>. As expected, the pH dependence of the rate constant for nucleotide incorporation was now essentially identical within the experimentally accessible pH range (Supplementary Fig. 4a). These data are consistent with chemistry serving as the rate-limiting step for both the K359L derivative and the wild-type enzyme in the presence of Mn<sup>2+</sup>. The reduced rate constant for nucleotide

incorporation and the solvent deuterium kinetic isotope effect observed for the K359L derivative relative to the wild-type enzyme provide additional support for the role of Lys359 as a general acid. We conclude that PV RdRp, and probably the other polymerases, employ general acid catalysis, regardless of the divalent cation cofactor used.

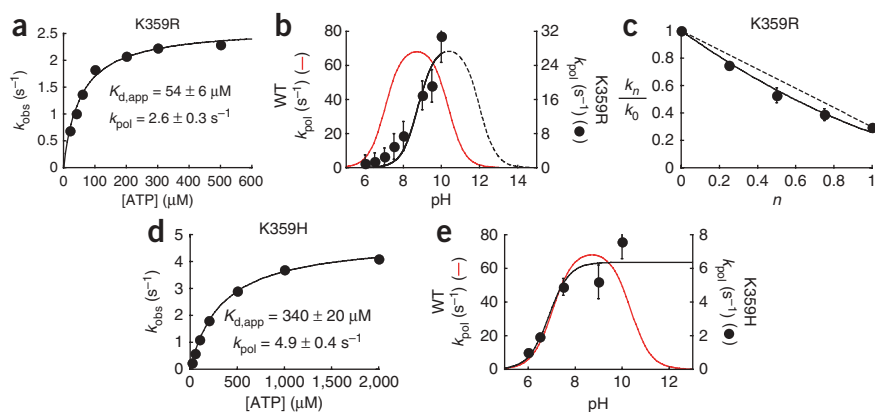
### Arginine and histidine substitute for lysine as the general acid

To test further the use of a general acid for nucleotidyl transfer, we produced a PV RdRp derivative in which Lys359 was substituted with arginine. In addition to testing our model for general acid catalysis, the ability of arginine to function at this position as a general acid would be consistent with the use of arginine as a general acid in X family polymerases such as DNA polymerase-β<sup>30</sup> (Fig. 2). A derivative containing an arginine at this position would be expected to function better than the K359L derivative but still worse than the wild-type enzyme, would have unique pH dependence relative to the wild-type enzyme and would show two proton-transfer reactions during nucleotidyl transfer. At pH 7.5, the K359L derivative bound nucleotide well and turned over ~three-fold faster than the K359L derivative and 20-fold slower than wild-type enzyme (Fig. 4a). At pH 10, the  $k_{\text{pol}}$  value for the K359L derivative had risen to within two-fold of the maximum value observed for wild-type enzyme at pH 8.5 and continued to rise (Fig. 4b). In contrast, the value for the wild-type enzyme was falling at pH 10 (Fig. 4b). This difference can be explained by the  $\Delta\text{p}K_{\text{a}}$  value of 2 between lysine and arginine. The descending limb of the pH rate profile for the K359L

**Figure 3** A conserved basic amino acid in the active site of multiple classes of polymerases participates directly in nucleotidyl transfer and functions as a general acid. (a) Values for  $k_{\text{pol}}$  plotted as a function of pH for K359L PV RdRp and wild-type (WT) PV RdRp. Kinetic data were obtained for AMP incorporation into S/S+1 (a symmetrical primer template substrate containing 2-aminopurine in the +1 templating position) using the stopped flow assay<sup>4</sup> in the presence of 5 mM MgCl<sub>2</sub> (Supplementary Methods). In the case of wild-type (WT) PV RdRp, the solid line shows the fit of the data to a model describing two ionizable groups (equation (3) in the Supplementary Methods), yielding  $\text{p}K_{\text{a}}$  values of 7.0 ± 0.1 and 10.5 ± 0.1 (ref. 4). Error bars indicate s.d. Plots of the  $\log(k_{\text{pol}})$  as a function of pH are provided in Supplementary Figure 3. (b) Proton-inventory plots for PV RdRp, HIV RT and RB69 DdDp, with either the leucine derivative or lysine-containing polymerase (Supplementary Methods). Values for  $k_n / k_0$  were plotted as a function of  $n$ — $n$  is the mole fraction of D<sub>2</sub>O,  $k_n$  is the observed rate constant for nucleotide incorporation in a reaction containing a given mole fraction of D<sub>2</sub>O and  $k_0$  is the observed rate constant for nucleotide incorporation in H<sub>2</sub>O. The solid lines represent the fit of the data to a one-proton-transfer model for the leucine derivative (filled squares) and to a two-proton-transfer model for the wild-type polymerase (filled circles) (Supplementary Methods). The dashed lines represent the predicted line for a one-proton-transfer model. Each data point represents the average of two or three independent experiments. Error bars indicate s.d. (<10% in all cases). Graphs for wild-type enzymes are adapted from ref. 4; more rigorous analysis of the two-proton model is reported therein.



**Figure 4** Altering nucleotidyl transfer kinetics by changing the amino acid that acts as the general acid. (a) Lys359 of PV RdRp was changed to arginine. Kinetic analysis of AMP incorporation by K359R PV RdRp at pH 7.5 fit to a hyperbola with a  $K_{d,app}$  for ATP of  $54 \pm 6 \mu\text{M}$  and a  $k_{pol}$  of  $2.6 \pm 0.1 \text{ s}^{-1}$ . (b) pH rate profile for K359R PV RdRp fit to a model describing a one-ionizable group, yielding a  $pK_a$  value of  $8.8 \pm 0.3$ . The dashed line shows the predicted curve for a second ionizable group with a  $pK_a$  of 12.0. The red line shows the curve of  $k_{pol}$  as a function of pH for wild-type (WT) PV RdRp with  $pK_a$  values of 7.0 and 10.5. Error bars indicate s.d. (c) Proton-inventory plot for K359R PV RdRp fit to a two-proton-transfer model. The dashed line represents the predicted line for a one-proton-transfer model. Each data point represents the average of two independent experiments. Error bars indicate s.d. ( $<10\%$ ). (d) Lys359 of PV RdRp was changed to histidine. Kinetic analysis of AMP incorporation by K359H PV RdRp at pH 7.5 fit to a hyperbola with a  $K_{d,app}$  for ATP of  $340 \pm 20 \mu\text{M}$  and a  $k_{pol}$  of  $4.9 \pm 0.4 \text{ s}^{-1}$ . (e) pH rate profile for K359H PV RdRp fit to a model describing a one-ionizable group, yielding a  $pK_a$  value of  $6.9 \pm 0.3$ . The red line shows the curve of  $k_{pol}$  as a function of pH for wild-type PV RdRp with  $pK_a$  values of 7.0 and 10.5. Error bars indicate the s.d.



derivative should appear at pH values greater than 11 (dashed line in Fig. 4b). Unfortunately, experiments cannot be performed at pH values higher than 10 because of the insolubility of metal hydroxide. Finally, the solvent deuterium kinetic isotope effect was  $3.4 \pm 0.1$ , essentially the same as that of the wild-type enzyme (Table 1), and proton-inventory data fit well to a two-proton model (Fig. 4c). These data strongly support the use of a general acid for nucleotidyl transfer and suggest that both lysine and arginine can serve this function.

Multisubunit RNA polymerases contain a histidine in the trigger loop that may serve as the general acid for these enzymes<sup>31</sup> (Fig. 2). Histidine can function at this position of the PV RdRp (Fig. 4d). We observed a ten-fold reduction in the  $k_{pol}$  value for the K359H derivative relative to the wild-type enzyme with only a modest (50%) increase in the  $K_{d,app}$  value for nucleotide at pH 7.5 (Fig. 4d). The  $pK_a$  value for the ascending limb of the pH rate profile for the K359H derivative was identical to that of the wild-type enzyme; the descending limb was absent, consistent with a  $pK_a$  value of 6–7 for histidine (Fig. 4e).

## DISCUSSION

We conclude that all four classes of nucleic acid polymerases use general acid catalysis for nucleotidyl transfer. Notably, the general acid is not absolutely essential but provides a 50-fold to 2,000-fold rate enhancement, depending upon the polymerase evaluated. The value of the  $pK_a$  for pyrophosphate in the polymerase active site is not known. The absence of an absolute requirement for a general acid in polymerase-catalyzed nucleotidyl transfer suggests that the value of the  $pK_a$  for the tetra-anionic form of pyrophosphate is not sufficiently high to preclude it from serving as a leaving group. The differences in rate enhancement conferred by the general acid for the different polymerases may reflect the extent to which the general acid also contributes to the overall neutralization of the negative charge in the active site. RdRps and RTs have evolved an independent structural motif (F) to bind the triphosphate<sup>26</sup>. In contrast, in other polymerase families the general acid is connected directly (helix O, helix K, trigger loop) or indirectly (helix P) to the structural elements that bind the triphosphate (Fig. 2). There is no doubt that neutralization of the negative charge that forms during the transition state will facilitate nucleotidyl transfer and probably also contributes to the reduced catalytic efficiency of the leucine derivatives<sup>12,32,33</sup>.

Protonation of the pyrophosphate leaving group in the mechanism of nucleotidyl transfer has been essentially ignored until recently. Most structures of polymerases poised for or undergoing catalysis reveal an interaction between a basic amino acid of the polymerase and the  $\beta$ -phosphate of the nucleotide substrate (Fig. 2), and this interaction has been interpreted as a nucleotide binding determinant. However, one study proposed that the interaction between the  $\beta$ -phosphate of the incoming nucleotide and His1085 (trigger loop) of yeast RNA polymerase II could link substrate recognition to catalysis if this histidine served as a general acid for protonation of the pyrophosphate leaving group<sup>31</sup>. Our recent studies of nucleotidyl transfer confirmed that protonation of the pyrophosphate leaving group does occur<sup>4</sup>, thus encouraging the identification of the proton donor.

Nature has used all three basic amino acids as the general acid (Fig. 2). The residue chosen may reflect the need to achieve a balance between the rate of polymerization and the fidelity of nucleotide incorporation. Enzymes that use a lysine as a general acid—for example, PV RdRp, HIV RT, RB69 DdDp and T7 DdRp—elongate nucleic acids with rate constants on the order of  $100 \text{ s}^{-1}$  (Table 1). In contrast, rat DNA polymerase- $\beta$  (which contains arginine as general acid) and yeast RNA polymerase II (histidine as a general acid) synthesize nucleic acid an order of magnitude more slowly at physiological pH<sup>8,30,34</sup>. The idea that a difference in the general acid leads to changes in elongation rate is supported by the reduced rate of RNA synthesis observed for the K359R and K359H derivatives of PV RdRp (Fig. 4). In addition, changing His1085 of yeast RNA polymerase II to tyrosine reduces the rate of catalysis by an order of magnitude without changing the observed affinity for the nucleotide substrate<sup>34</sup>. Our study would suggest that the reduced rate of catalysis for the H1085Y mutant at physiological pH is likely to reflect the increased  $pK_a$  value for the tyrosine hydroxyl proton relative to the histidine imino proton.

It has often been suggested that polymerase translocation occurs concomitantly with or after pyrophosphate release<sup>21,35</sup>. The best structural description of a polymerase undergoing catalysis was provided by work using a fragment of DNA polymerase I from *Bacillus stearothermophilus*, an A family polymerase<sup>36</sup>. In this system, movement of helix O is coupled to catalysis and translocation. Notably, the general acid of all polymerases seems to be associated with exceptionally dynamic structural elements (Supplementary Fig. 1 and Supplementary Fig. 5 online), and the movement of

these elements has been implicated in translocation<sup>12,18,21,23,31,34,36</sup>. It is possible that the protonation-deprotonation cycle of the general acid during catalysis toggles helix O (and its structural equivalents) between the closed (protonated) and open (deprotonated) states, thus driving the translocation reaction. Consistent with this possibility is the observation that changing the general acid to leucine leads to substantial reductions in processivity for all nucleic acid polymerases used in this study (**Supplementary Fig. 6** online).

The use of a general acid by polymerases for nucleotidyl transfer has important implications for reaction reversal, pyrophosphorolysis. Because pyrophosphate in solution should be protonated, deprotonation of pyrophosphate should not occur readily for enzymes that use lysine or arginine as a general acid, as these residues should be rapidly reprotonated by solvent after pyrophosphate release. In contrast, enzymes such as the multisubunit RNA polymerases that use a histidine residue should catalyze pyrophosphorolysis more efficiently, an observation that has been reported<sup>37–39</sup>. The presence of histidine in these polymerases may have been an early solution to deal with the problem of polymerase arrest caused by backtracking, misincorporation or template damage that is now dealt with by using editing factors<sup>40</sup>.

Finally, it is worth noting that, where studied, the element harboring the general acid has been shown to contribute not only to incorporation speed but also incorporation fidelity<sup>41–46</sup>. RNA virus pathogenesis and virulence are closely linked to replication speed and fidelity<sup>47,48</sup>. It has recently been shown that a PV mutant encoding an RdRp with increased fidelity is attenuated<sup>47,48</sup>, and this attenuated virus serves as an effective vaccine strain<sup>49</sup>. The identification of a single amino acid in the polymerase active site that can tune replication speed and perhaps modulate incorporation fidelity suggests the provocative hypothesis that a universal strategy for viral attenuation may exist that can be applied to the rational design of virus vaccine strains.

## METHODS

**Materials.** All general experimental materials, buffers, salts, and so on, were of the highest grade available from Sigma, Fisher or VWR. A complete list of any specialty reagents used is provided in the **Supplementary Methods**.

**Construction, expression and purification of polymerases and their derivatives.** We constructed derivatives of all polymerases: PV RdRp, RB69 DdDp, T7 DdRp and HIV RT by using standard recombinant DNA protocols, described in the **Supplementary Methods**. Briefly, we mutated DNA sequences by PCR. Forward and reverse primers (**Supplementary Table 1** online) used for amplification were selected based on the presence of unique restriction sites suitable for subcloning of the mutated DNA fragment into the expression plasmid for the wild-type polymerase.

We expressed all polymerases in *Escherichia coli* and purified them as described previously<sup>4</sup>. Briefly, we transformed cells with the appropriate plasmid and used these cells to produce an inoculum for large-scale growth. We induced gene expression during exponential growth by addition of isopropyl- $\beta$ -D-thiogalactopyranoside. We lysed the induced cells in appropriate buffers and we purified the enzymes to apparent homogeneity by using standard column chromatography resins and protocols. Details are provided in the **Supplementary Methods**.

**Nucleotide-incorporation experiments.** We performed nucleotide-incorporation experiments as described previously<sup>4</sup>. In general, we formed elongation complexes in the appropriate buffer by incubating polymerase with the appropriate primed template followed by rapid mixing with a nucleoside triphosphate substrate, generally ATP. Rapid mixing was performed by using either a chemical quench flow (CQF) or stopped flow instrument (both from KinTek). For the CQF experiments, we used <sup>32</sup>P-labeled primers. We monitored primer extension by phosphorimaging of polyacrylamide gels. For the stopped flow experiments, we used templates containing 2-aminopurine ribonucleoside

monophosphate on the 5' side of the templating nucleotide. Primer extension causes a fluorescence change that could be monitored using the stopped flow instrument. All details, including modifications, are provided in the **Supplementary Methods**.

## Solvent deuterium kinetic isotope effect and proton-inventory experiments.

We performed solvent deuterium kinetic isotope effect and proton-inventory experiments by monitoring pre-steady state nucleotide incorporation either using the CQF or stopped flow instruments. We prepared enzymes, substrates and buffers in 100% water or 100% D<sub>2</sub>O and then mixed them at the appropriate ratio to obtain 0%, 25%, 50%, 75% or 100% D<sub>2</sub>O. Deuterated glycerol was used in all solutions in D<sub>2</sub>O. The pD was used instead of pH for the solutions in D<sub>2</sub>O and was adjusted according to pD = pH + 0.4. Solvent deuterium kinetic isotope effect values were calculated as the ratio of  $k_{\text{pol}}$  values obtained in H<sub>2</sub>O divided by that obtained in D<sub>2</sub>O. Plots consisted of  $k_n / k_{\text{H}_2\text{O}}$  as a function of  $n$ , where  $k_n$  is the observed rate constant for nucleotide incorporation at a particular mole fraction of D<sub>2</sub>O,  $k_{\text{H}_2\text{O}}$  is the observed rate constant for nucleotide incorporation in H<sub>2</sub>O and  $n$  is the solvent mole fraction of D<sub>2</sub>O. Proton-inventory data were fit to the modified Gross-Butler equation for a two-proton-transfer model<sup>29</sup>:

$$\frac{k_n}{k_{\text{H}_2\text{O}}} = (1 - n + n \times \Phi 1)(1 - n + n \times \Phi 2) \quad (1)$$

or for a one-proton-transfer model<sup>29</sup>:

$$\frac{k_n}{k_{\text{H}_2\text{O}}} = (1 - n + n \times \Phi) \quad (2)$$

where  $k_n$  is the observed rate constant at the different percentages of D<sub>2</sub>O,  $k_{\text{H}_2\text{O}}$  is the observed rate constant in water,  $n$  is the mole fraction of D<sub>2</sub>O and  $\Phi$  is the inverse of the isotope effect for each ionizable group.

**Data analysis.** Observed rate constants ( $k_{\text{obs}}$ ) for nucleotidyl transfer at various concentrations of nucleotide were obtained by fitting product-versus-time data to an equation defining a single exponential. Values for  $K_{\text{d,app}}$  and  $k_{\text{pol}}$  were obtained by fitting  $k_{\text{obs}}$ -versus-[NTP] data to an equation defining a hyperbola. Data were fit by nonlinear regression using the program KaleidaGraph (Synergy Software). Specific equations used are provided in **Supplementary Methods**.

*Note: Supplementary information is available on the Nature Structural & Molecular Biology website.*

## ACKNOWLEDGMENTS

We thank S.J. Benkovic, P.C. Bevilacqua, J. Martin Bollinger, K. Murakami, K.D. Raney and J.C. Reese for comments on the manuscript. This study was supported by a grant (AI45818) from the US National Institutes of Health to C.E.C.

## AUTHOR CONTRIBUTIONS

C.E.C., J.J.A., C.C. and E.D.S. designed research; C.C., E.D.S., K.R.M., J.J.A., I.M. and A.U. performed research; M.G. and W.K. contributed new reagents and analytical tools; C.E.C., C.C., E.D.S. and J.J.A. analyzed data; C.E.C., E.D.S., C.C. and J.J.A. wrote the paper.

## COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/nsmb/>.

Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

- Steitz, T.A. A mechanism for all polymerases. *Nature* **391**, 231–232 (1998).
- Yang, W., Lee, J.Y. & Nowotny, M. Making and breaking nucleic acids: two-Mg<sup>2+</sup>-ion catalysis and substrate specificity. *Mol. Cell* **22**, 5–13 (2006).
- Fothergill, M., Goodman, M.F., Petruska, J. & Warshel, A. Structure-energy analysis of the role of metal ions in phosphodiester bond hydrolysis by DNA polymerase I. *J. Am. Chem. Soc.* **117**, 11619–11627 (1995).
- Castro, C. *et al.* Two proton transfers in the transition state for nucleotidyl transfer catalyzed by RNA- and DNA-dependent RNA and DNA polymerases. *Proc. Natl. Acad. Sci. USA* **104**, 4267–4272 (2007).

5. Florian, J., Goodman, M.F. & Warshel, A. Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *J. Am. Chem. Soc.* **125**, 8163–8177 (2003).
6. Showalter, A.K. & Tsai, M.D. A reexamination of the nucleotide incorporation fidelity of DNA polymerases. *Biochemistry* **41**, 10571–10576 (2002).
7. Florian, J., Goodman, M.F. & Warshel, A. Computer simulations of protein functions: searching for the molecular origin of the replication fidelity of DNA polymerases. *Proc. Natl. Acad. Sci. USA* **102**, 6819–6824 (2005).
8. Sucato, C.A. *et al.* DNA polymerase  $\beta$  fidelity: halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state. *Biochemistry* **47**, 870–879 (2008).
9. Anand, V.S. & Patel, S.S. Transient state kinetics of transcription elongation by T7 RNA polymerase. *J. Biol. Chem.* **281**, 35677–35685 (2006).
10. Arnold, J.J. & Cameron, C.E. Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of  $Mg^{2+}$ . *Biochemistry* **43**, 5126–5137 (2004).
11. Ferrer-Orta, C. *et al.* Sequential structures provide insights into the fidelity of RNA replication. *Proc. Natl. Acad. Sci. USA* **104**, 9463–9468 (2007).
12. Franklin, M.C., Wang, J. & Steitz, T.A. Structure of the replicating complex of a Pol  $\alpha$  family DNA polymerase. *Cell* **105**, 657–667 (2001).
13. Gohara, D.W., Arnold, J.J. & Cameron, C.E. Poliovirus RNA-dependent RNA polymerase (3Dpol): kinetic, thermodynamic, and structural analysis of ribonucleotide selection. *Biochemistry* **43**, 5149–5158 (2004).
14. Huang, H., Chopra, R., Verdine, G.L. & Harrison, S.C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669–1675 (1998).
15. Pomerantz, R.T., Temiakov, D., Anikin, M., Vassilyev, D.G. & McAllister, W.T. A mechanism of nucleotide misincorporation during transcription due to template-strand misalignment. *Mol. Cell* **24**, 245–255 (2006).
16. Sarafianos, S.G. *et al.* Structures of HIV-1 reverse transcriptase with pre- and post-translocation AZTMP-terminated DNA. *EMBO J.* **21**, 6614–6624 (2002).
17. Spence, R.A., Kati, W.M., Anderson, K.S. & Johnson, K.A. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Science* **267**, 988–993 (1995).
18. Temiakov, D. *et al.* Structural basis for substrate selection by T7 RNA polymerase. *Cell* **116**, 381–391 (2004).
19. Thompson, A.A. & Peersen, O.B. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J.* **23**, 3462–3471 (2004).
20. Yang, G., Franklin, M., Li, J., Lin, T.C. & Konigsberg, W. Correlation of the kinetics of finger domain mutants in RB69 DNA polymerase with its structure. *Biochemistry* **41**, 2526–2534 (2002).
21. Yin, Y.W. & Steitz, T.A. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **116**, 393–404 (2004).
22. Zamyatkin, D.F. *et al.* Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.* **283**, 7705–7712 (2008).
23. Canard, B., Chowdhury, K., Sarfati, R., Double, S. & Richardson, C.C. The motif D loop of human immunodeficiency virus type 1 reverse transcriptase is critical for nucleoside 5'-triphosphate selectivity. *J. Biol. Chem.* **274**, 35768–35776 (1999).
24. Hizi, A., Tal, R., Shaharabany, M. & Loya, S. Catalytic properties of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2. *J. Biol. Chem.* **266**, 6230–6239 (1991).
25. Gillis, A.J., Schuller, A.P. & Skordalakes, E. Structure of the *Tribolium castaneum* telomerase catalytic subunit TERT. *Nature* **455**, 633–637 (2008).
26. Ng, K.K., Arnold, J.J. & Cameron, C.E. Structure-function relationships among RNA-dependent RNA polymerases. *Curr. Top. Microbiol. Immunol.* **320**, 137–156 (2008).
27. Rosta, E., Kamerlin, S.C. & Warshel, A. On the interpretation of the observed linear free energy relationship in phosphate hydrolysis: a thorough computational study of phosphate diester hydrolysis in solution. *Biochemistry* **47**, 3725–3735 (2008).
28. Schowen, R.L. Mechanistic deductions from solvent isotope effects. *Progr. Phys. Org. Chem.* **9**, 275–332 (1972).
29. Venkatasubban, K.S. & Schowen, R.L. The proton inventory technique. *CRC Crit. Rev. Biochem.* **17**, 1–44 (1984).
30. Kraynov, V.S., Showalter, A.K., Liu, J., Zhong, X. & Tsai, M.D. DNA polymerase  $\beta$ : contributions of template-positioning and dNTP triphosphate-binding residues to catalysis and fidelity. *Biochemistry* **39**, 16008–16015 (2000).
31. Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D. & Kornberg, R.D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006).
32. Xiang, Y., Oelschlaeger, P., Florian, J., Goodman, M.F. & Warshel, A. Simulating the effect of DNA polymerase mutations on transition-state energetics and fidelity: evaluating amino acid group contribution and allosteric coupling for ionized residues in human Pol  $\beta$ . *Biochemistry* **45**, 7036–7048 (2006).
33. Yang, G., Lin, T., Karam, J. & Konigsberg, W.H. Steady-state kinetic characterization of RB69 DNA polymerase mutants that affect dNTP incorporation. *Biochemistry* **38**, 8094–8101 (1999).
34. Kaplan, C.D., Larsson, K.M. & Kornberg, R.D. The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by  $\alpha$ -amanitin. *Mol. Cell* **30**, 547–556 (2008).
35. Marchand, B. & Gotte, M. Site-specific footprinting reveals differences in the translocation status of HIV-1 reverse transcriptase. Implications for polymerase translocation and drug resistance. *J. Biol. Chem.* **278**, 35362–35372 (2003).
36. Johnson, S.J., Taylor, J.S. & Beese, L.S. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. USA* **100**, 3895–3900 (2003).
37. Erie, D.A., Yager, T.D. & von Hippel, P.H. The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 379–415 (1992).
38. Rudd, M.D., Izban, M.G. & Luse, D.S. The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proc. Natl. Acad. Sci. USA* **91**, 8057–8061 (1994).
39. Wang, D. & Hawley, D.K. Identification of a 3'  $\rightarrow$  5' exonuclease activity associated with human RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **90**, 843–847 (1993).
40. Erie, D.A., Hajiseyidjavad, O., Young, M.C. & von Hippel, P.H. Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science* **262**, 867–873 (1993).
41. Bebenek, A. *et al.* Dissecting the fidelity of bacteriophage RB69 DNA polymerase: site-specific modulation of fidelity by polymerase accessory proteins. *Genetics* **162**, 1003–1018 (2002).
42. Carroll, S.S., Cowart, M. & Benkovic, S.J. A mutant of DNA polymerase I (Klenow fragment) with reduced fidelity. *Biochemistry* **30**, 804–813 (1991).
43. Johnson, V.A. *et al.* Update of the drug resistance mutations in HIV-1: spring 2008. *Top. HIV Med.* **16**, 62–68 (2008).
44. Sousa, R. & Padilla, R. A mutant T7 RNA polymerase as a DNA polymerase. *EMBO J.* **14**, 4609–4621 (1995).
45. Suzuki, M., Yoshida, S., Adman, E.T., Blank, A. & Loeb, L.A. *Thermus aquaticus* DNA polymerase I mutants with altered fidelity. Interacting mutations in the O-helix. *J. Biol. Chem.* **275**, 32728–32735 (2000).
46. Zhang, H. *et al.* The L561A substitution in the nascent base-pair binding pocket of RB69 DNA polymerase reduces base discrimination. *Biochemistry* **45**, 2211–2220 (2006).
47. Arnold, J.J., Vignuzzi, M., Stone, J.K., Andino, R. & Cameron, C.E. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *J. Biol. Chem.* **280**, 25706–25716 (2005).
48. Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2006).
49. Vignuzzi, M., Wendt, E. & Andino, R. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.* **14**, 154–161 (2008).
50. Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J. & Pelletier, H. Crystal structures of human DNA polymerase  $\beta$  complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205–11215 (1997).