Supplementary Information

A Novel Mechanism of Enhanced Transcription Activity and Fidelity for Influenza A Viral RNA-dependent RNA Polymerase

Xinzhou Xu\textsuperscript{a,c}, Lu Zhang\textsuperscript{b,\textasteriskcentered}, Julie Tung Sem Chu\textsuperscript{c}, Yuqing Wang\textsuperscript{a,c}, Alex Wing Hong Chin\textsuperscript{c,j}, Tin Hang Chong\textsuperscript{a,d}, Zixi Dai\textsuperscript{d}, Leo Lit Man Poon\textsuperscript{c,j}, Peter Pak-Hang Cheung\textsuperscript{a,d,g,h,\textasteriskcentered}, Xuhui Huang\textsuperscript{a,d,\textasteriskcentered}

\textsuperscript{a}The Hong Kong University of Science and Technology-Shenzhen Research Institute, Hi-Tech Park, Nanshan, Shenzhen 518057, China
\textsuperscript{b}State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, Fujian 350002, China
\textsuperscript{c}School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China
\textsuperscript{d}Department of Chemistry, State Key Laboratory of Molecular Neuroscience, The Hong Kong University of Science and Technology, Kowloon, Hong Kong, China
\textsuperscript{e}Bioengineering Graduate Program, Department of Biological and Chemical Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
\textsuperscript{f}University of Chinese Academy of Sciences, Beijing 100049, China
\textsuperscript{g}Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China.
\textsuperscript{h}Li Ka Shing Institute of Health Sciences, Li Ka Shing Medical Sciences Building, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China
\textsuperscript{i}HKU-Pasteur Research Pole, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China
\textsuperscript{j}Centre for Immunity and Infection, Hong Kong Science Park, Hong Kong, China

*To whom correspondence may be addressed. Email: luzhang@fjirsm.ac.cn or ppcheung@cuhk.edu.hk or xuhuihuang@ust.hk
Section 1 Supplementary Text

1.1 Cell-based mini-genome luciferase assay
The pcDNA3-WSN-PB1 (wildtype and mutants), PB2, PA, and NP genes were used to express the recombinant IAV polymerase in the HEK-293T cells. The reporter gene plasmid pPolI-NS-Luc plasmid (pBZ81A36) contains the negative-sense influenza vRNA-like firefly luciferase gene flanked by human Pol I and IAV polymerase 3’ promoter and 5’ terminator. pCMV-Gaussian luciferase plasmid (NEB) was added to serve as transfection control. The plasmids mix were prepared in a ratio of PB1: PB2: PA: NP: NS-Luc: Gaussian = 1: 1: 1: 2: 1 :1 into JetPrime buffer (Polyplus) according to the transfection instruction. 4 × 10^4 cells in 100 μl culture medium were seeded in each well of a 96 well plate and transfected for 48 hours. For the negative control, pcDNA-PB1 was omitted. After 48 hours of incubation at 37 degrees Celsius, the cells were lysed according to the Dual Luciferase Reporter Gene Assay Kit (Beyotime) protocol. Briefly, The first reading was obtained from firefly luciferase and quenched, and the second reading was obtained from gaussian luciferase. Cell lysates were transferred to a 96-well white plate and the signal was measured on the Varioskan LUX Multimode Microplate Reader (Thermofisher) in luminescence mode. The relative polymerase activity was represented as Firefly luciferase intensity / Gaussian luciferase intensity and normalized to the intensity of wild-type polymerase (100%) (Fig. 2A and 3A). The Gaussian luciferase intensity was normalized to the intensity of polymerase without the PB1 gene (Fig. S3).

1.2 Influenza A virus polymerase complex purification
The TAP tag on the PB2 subunit consisting of a TEV protease cleavage site and a protein A domain was used as tandem affinity tags in the purification steps according to previous works (1, 2). Prior to plasmid transfection, the HEK-293T cells were passaged at least three times before seeded in 10 ml to a 10 cm dish at 70% confluence. Cells were transfected with 1.6 μg of each of PB1 (wildtype and mutants), PA, PB2-TAP plasmids by 10 µl of JetPrime (Polyplus) at 90% confluence in the next day according to the instruction. Cells were collected after 48 hours of incubation and gently washed in 500 μl ice-cold PBS buffer for three times. Cells were suspended in lysis buffer [50 mM Tris-HCl pH 8.0, 200 mM NaCl, 25% glycerol, 0.5% Igepal CA-630, 1 mM dithiothreitol (DTT), 1 × complete EDTA-free protease inhibitor cocktail (PIC, Roche)] and incubated on a
rotating shaker at 4 °C for 1 hour to release the protein content from cells. A white silk-like precipitate will become visible in the lysate, indicating the end of the lysis process. The lysate was centrifuged at 18,000 × g for 15 minutes, and the supernatant was diluted four times in binding buffer [50 mM Tris-HCl pH 8.0, 200 mM NaCl] together with 50 μl pre-washed IgG Sepharose 6 Fast Flow beads (GE Healthcare, 17-0969-01) on a rotator at 4 °C for 2 hours. The mixer was then centrifuged at 500 ×g for 1 min and washed in wash buffer [10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 10% glycerol, 0.1% Igepal CA-630] for three times to remove contaminants. Bound IAV polymerase complex on the IgG beads were then resuspended in 300 μl cleavage buffer [10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 10% glycerol, 0.1% Igepal CA-630, 1 mM DTT, 1 × PIC] and released by 5 μl AcTEV protease (Thermofisher 12575015) on a rotator at 4 °C overnight. The mixture was then centrifuged at 17,000 ×g for 5 minutes, and the supernatant containing IAV polymerase complex was collected, aliquoted in 50 μl size, flash freeze in liquid nitrogen, and then stored in -80 °C freezer until use.

To quantify the amount of polymerases, we used Bovine Serum Albumin (BSA) as the mass standard to estimate the yield of purified wild-type and mutant IAV RdRp complexes in SDS-PAGE (Fig. S4A). In particular, we prepared the BSA standard by diluting the commercial BSA standard solution (2 mg/mL, Thermofisher) into 30 ng/μl, 20 ng/μl, 10 ng/μl, 5 ng/μl. We loaded 6 μl of each BSA standard in parallel with 10 μl of purified RdRp complexes in 8 % SDS-PAGE. The gel was stained in InstantBlue (Abcam), and the yield was estimated from the standard curve fitted from various concentrations of BSA. The mass of wild-type and mutant RdRp complex is estimated between 5 to 7 ng/μl.

To verify if mutations made on PB1 affect the correct expression of PB1 subunits and the formation of polymerase complexes, we have conducted the western blot analysis on PB1, PB2, and PA subunits of the purified wild-type and mutant IAV RdRps. The results suggested that all three subunits from wild-type and mutant IAV polymerases were expressed correctly at the expected molecular weight (Fig. S4C). We also stained the SDS-PAGE to visualize the expression and integrity of all three subunits after the purification. As shown in Fig. S4A and S4B, the PB1, PB2-TAP, PA subunits were resolved and visualized in 8 % SDS-PAGE as well as the co-purified Hsp70, similar to a previous protocol publication (1). Altogether, our results from western blot and
gel stain indicate that the introduced K235 and R239 mutant systems had well expressed all polymerase subunits as well as maintained the RdRp complex integrity.

1.3 Capping and radioactive labeling of RNA primer
Synthetic 5’ diphosphate RNA of 11nt (5’-ppGAAUACUCAAG-3’) (Chemgenes) was 5’ capped with a radiolabeled cap-0 and a cap-1 structure to initiate the in vitro transcription as described previously (2). Briefly, a 20 μl reaction was setup one hour at 37 °C with 1 μM 11-nt pp-RNA primer, 0.25 μM [α-32P] GTP (3,000 Ci mmol⁻¹, Perkin-Elmer BLU006H250UC), 0.8 mM S-adenosylmethionine, 2.5 U/μl 2'-O-methyltransferase (NEB), 0.5 U/μl Vaccinia Capping System (NEB) and 2U/μl RNase Inhibitor (NEB). The Reaction mixture was column purified (NEB), eluted with 20 μl RNase-free water, and stored at -20 °C. The radioactive intensity was assessed by placing the 1.5 ml tube near the Geiger meter detector to around 1000 counts per second (cps) per μl.

1.4 Capped RNA initiated in vitro transcription assay
The radioactive capped RNA initiated in vitro transcription of the IAV polymerase was set up in a 10 μl reaction. The mixture contains 5 mM Tris-HCl pH 8, 5 mM MgCl₂, 75 mM NaCl, 5% Glycerol, 0.05% Igepal CA-630, 1 mM DTT, 1.6 μM 30nt 3’ RNA template (5’-GCAUUG UCGCACAUCAGUACCUGCUUUCGCU-3’), 1.6 μM 5’ RNA promoter (5’-AGUAGAAACAAGGU-3’), 2 U/μl RNase Inhibitor (NEB M0314), 50 cps of radioactive capped RNA, 1 mM ATP, 1 mM CTP, 1 mM UTP, 1 mM GTP and 30 ng purified IAV polymerase. The reaction mixtures were incubated at 30 °C for 1 hour, denatured in an equal amount of RNA sample buffer [90% formamide, 10 mM EDTA, 0.01% bromophenol blue, and 0.01% xylene cyanol] at 95 °C for 5 mins and chilled on ice for 5 mins. The RNA sample was then separated on a 20% urea-PAGE in 1 × TBE buffer under 200 V for 2 to 3 hours on Mini-PROTEAN Tetra cell (Biorad). The gel was then placed against Molecular Dynamics Storage Phosphor Screen (Kodak) overnight, and the screen was scanned by Sapphire Biomolecular Imager (Azure Biosystems). The gel image was analysed in Image Lab (Biorad).

To ensure that the proper amount of primer was used to best visualize the product completion in the in vitro transcription assay, we conducted an in vitro transcription experiment by gradually...
increasing the concentration of radioactively labelled primers (Fig. S5). The remaining conditions were kept constant to ensure that the only variable is the primer concentration. When the primer concentration is less than 50 cps, the extension percentages remain at greater than 31%. However, when the primer concentration exceeds 50 cps, the extension percentage drops to 20.0 % and 10.6 %, indicating that an excessive amount of primer would decrease the extension percentage.

1.5 in vitro single-turnover NTP incorporation

The method utilizes single-turnover incorporation of one incoming NTP at various concentrations in a time-course manner, which was previously described in the kinetics characterization of yeast RNA polymerase II (3). Briefly, the radioactive capped RNA initiated in vitro transcription of the IAV polymerase was set up in a 10 μl reaction. The mixture contains 5 mM Tris-HCl pH 8, 5 mM MgCl₂, 75 mM NaCl, 5% Glycerol, 0.05% Igepal CA-630, 1 mM DTT, 1.6 μM 14-nt 3’ RNA template (5’-GGCCUGCUUUUGCU-3’), 1.6 μM 5’ RNA promoter, 2 U/ μl RNase Inhibitor (NEB), 50 cps of radioactive capped RNA and 30 ng purified influenza A virus polymerase. After pre-incubating the reaction mixtures at 30 ℃ for 30 min, various amounts of CTP or UTP were added to the mixture to achieve the final concentration of 10 μM, 50 μM, 100 μM, 200 μM, 500 μM, 1000 μM. For correct pairing of CTP, the reactions were started for 0.5 sec, 2sec, 5 sec, 10 sec, 30 sec, 60 sec, and rapidly stopped by immersing the tubes in liquid nitrogen and denatured in an equal amount of RNA sample buffer [90% formamide, 10 mM EDTA, 0.01% bromophenol blue and 0.01% xylene cyanol] at 95℃ for 5 mins and chilled on ice for 5 mins. For wobble pairing of UTP, all settings are identical except the reaction was extended to 5 sec, 15 sec, 30 sec, 60 sec, 120 sec, 300 sec, 600 sec. The RNA sample was then separated on a 20% urea-PAGE in 1 × TBE buffer under 8 watts for 6 to 8 hours on Protean II xi cell (Biorad). The gel was then placed against Molecular Dynamics Storage Phosphor Screen (Kodak) overnight, and the screen was scanned by Sapphire Biomolecular Imager (Azure Biosystems). The gel image was analysed in Image Lab (Biorad).

The IAV genome's template sequence starts with 3'-UCGUUUU or 3'-UCGCUUU, which are highly conserved and crucial for transcription efficiency and structural stability of RNA-bound RdRp (Fig. S7A) (4). Furthermore, to assess the fidelity of IAV transcription in vitro, we employed an 11nt radioisotope-labelled capped primer terminating in AG-3', which mimics the IAV RdRp's
unique cap-snatching mechanism in infected host cells (5). To effectively commence transcription, the primer's ending AG'-3 will couple and prime with the 3'-UC in the 3' template's priming region. Thus, fidelity may be quantified by comparing the NTP incorporation to the G at the template's third position (we use single-turnover NTP integration to assess fidelity, see Fig. S7B). Therefore, in our setup, only CTP and UTP quantification (but not ATP and GTP) is feasible as CTP and UTP will pair with the 3rd position G but not with the subsequent Us in the highly conserved 3'-UCCGUUU template sequence and our adopted AG-3' primer (Fig. S7B). This would yield a single extension product (Fig. S7C and S7D), and thus satisfy the fidelity measurement condition. On the contrary, if the IAV RdRp is soaked in ATP or GTP, the ATP and GTP will couple with the following Us in the template strand despite the mismatch in the initial incorporation of ATP:G or GTP:G. This results in uncertainty in RNA extension (Fig. S7E and S7F) and makes it difficult to conduct single-turnover NTP incorporation to measure fidelity.

It is worthy to note that there should be multiple extension products with the AG transcription assays as stated in the previous literature (6-8). The multiple extension products are mainly due to different priming schemes, including a primary C initiated product by AG-3’ priming with 3’-UC, a secondary G initiated product by G-3’ priming with 3’-U, and a few realignment products by realigning the primer after an initial addition of several nucleotides (as illustrated in Fig. S8A). As a result, multiple bands can be observed as reported in the previous studies using a 14-nt template (7). However, in this study, we used a 30-nt template which is originated from the 3’ end of the PA segment and observed only one compact band, because the multiple extension products from the 30-nt template (e.g., products with a length of 39-nt and 40-nt) could not be resolved clearly in the Urea-PAGE (Fig. S8B). Therefore, we believe that the discrepancy between the multiple extension products observed in the previous literature and 1-nt extension product captured in the current study is mainly attributed to the resolving ability of the Urea-PAGE gel weakens when template length increases from 14-nt to 30-nt. To further verify the different resolving ability of Urea-PAGE gel towards the 14-nt and 30-nt template, we performed the transcription assay on a 30-nt template in parallel with a 14-nt template. Indeed, consistent with previous studies, we observed multiple bands from 14-nt template extension while only one compact band from the 30-nt template extension (Fig. S8B). Besides, we found that our extension products from 30-nt template migrate slower than expected even in denaturing environment of 8 M urea. We believe it
is due to the strong hybridization formed by product RNA and the complimentary template, as suggested by a recent paper (9).

The above-discussed multiple products shall not impact our single-turnover NTP incorporation assay (Fig. S6), as we conducted it by limiting the NTP pool to ensure only a single NTP extension product occurred (Fig. S8C). In particular, only CTP or UTP was used for the single-turnover assay, which pairs with the G on the template following the initial 3'-UC priming site. Due to the absence of GTP in the NTP pool, the secondary G initiated products will not be extended beyond C on the template when the primer's ending G pairs with the template's initial 3'-U. (Fig. S8D). Therefore, we only observed 1-nt extension in Fig. S6.

1.6 Data fitting
Nonlinear-regression data fitting was performed using GraphPad Prism 8. The quantified time and concentration-dependent product formation for different NTP substrates are fitted into a biphasic exponential equation (I) which consequently yields two phases correspond to the fast \( k_{\text{fast}} \) and slow \( k_{\text{slow}} \) observed rates for the incorporation. The substrate concentration dependence of the fast phase is then fitted into the Michaelis-Menten equation (II) to obtain the \( k_{\text{pol}} \) and \( K_{d,\text{app}} \) for each NTP substrate. The specificity constant was determined by the ratio of \( k_{\text{pol}} \) and \( K_{d,\text{app}} \) (III). The discrimination score is calculated by the different specificities from CTP and UTP in the equation (IV).

\[
Y = Y_{\text{max}} \times \text{Fast}\% \times (1 - e^{-k_{\text{fast}} t}) + Y_{\text{max}} \times \text{Slow}\% \times (1 - e^{-k_{\text{slow}} t})
\]  

\[
k_{\text{fast}} = \frac{k_{\text{pol}}[S]}{K_{d,\text{app}} + [S]}
\]  

\[
\text{Specificity} = \frac{k_{\text{cat}}}{K_m} = \frac{k_{\text{pol}}}{K_{d,\text{app}}}
\]  

\[
\text{Discrimination} = \frac{\text{Specificity}_{\text{correct}}}{\text{Specificity}_{\text{incorrect}}} = \frac{k_{\text{pol,CTP}}/K_{d,\text{app,CTP}}}{k_{\text{pol,UTP}}/K_{d,\text{app,UTP}}}
\]

1.7 Fast and slow phases of NTP substrate incorporation
Single turn-over in vitro transcription was performed under various NTP concentrations. The observed reaction rate \( (k_{\text{obs}}) \) at each substrate concentration was obtained through the non-linear
regression fitting of product formation over time (Fig. S6). The fitting is not optimal when the data are fitted into one phase exponential equation, as the end of the fitted curve does not even out after the initial rapid incorporation. The obtained fast \((k_{\text{fast}})\) and slow \((k_{\text{slow}})\) observed rates in the biphasic exponential equation represented the co-existence of two populations of IAV RdRPs in the conformational equilibrium. \(k_{\text{fast}}\) represents the fast incorporation of NTP substrate in the reaction contributed by the poised RdRps, while \(k_{\text{slow}}\) is counted for the catalysis of suboptimal RdRps that requires a rate-limiting step of associated isomerization (e.g., translocation and active site closure). A similar two-phase kinetics has also been observed for the nucleotide incorporation kinetics of yeast RNA polymerase II (3).

The \(k_{\text{fast}}\) from different NTP substrate concentrations was then fitted into the Michaelis Menten equation, where the maximum rate of NTP incorporation \((k_{\text{pol}})\) and apparent dissociation constant \((K_{\text{d,app}})\) can be calculated. \(k_{\text{pol}}\) represents the reaction rate at enormous substrate concentration while \(K_{\text{d,app}}\) is a result of NTP loading and binding to the active site. Previous works in characterizing RNA polymerases provided a simplified model to determine the enzyme kinetic parameters (10-12). These assumptions include: the equilibrium of NTP loading and binding is a rapid single-step event, \(k_{\text{pol}}\) is the overall rate-limiting step. Such simplification is reasonable as the catalyst rate constant \((k_{\text{cat}})\) and Michaelis constant \((K_{m})\) in the multistep reactions cannot be well resolved due to the experimental limitations (3). Thus, the specificity \((k_{\text{cat}}/K_{m})\) constant, as a measurement of enzyme efficiency for the catalysis of one specific NTP substrate, can be alternatively derived from \(k_{\text{pol}}/K_{\text{d,app}}\), and it holds true for both cognate and noncognate NTP base pairing. By computing the specificity of cognate over noncognate NTP substrate, it will subsequently derive the discrimination score for cognate over noncognate NTP, which serves as an enzymatic explanation for the fidelity of NTP incorporation.

### 1.8 Exploration of the fitting parameters on the percentage of the fast phase

During the fitting of the biphasic exponential equation, the percentage of fast phase \((\text{fast}\%)\) describes the percentage of IAV RdRP population that is at optimal conformation for catalyzing substrates, while the percentage of slow phase \((\text{slow}\% = 100\% - \text{fast}\%)\) represents the suboptimal population in the reaction. To determine the fast\% for both wildtype and K235R RdRP in CTP and UTP base pairing, we calculated the sum of squared errors (SSE) for each base pairing scenario.
for wildtype and K235R and chose the corresponding fast% based on the minimal medium and variation in SSE (Fig. S9A). The fast% for CTP substrate is 58% for wildtype and 53% for K235R RdRP, and the fast% for UTP substrate is 31% for wildtype and 34% for K235R RdRP. The fast% is considerably higher for cognate NTP than noncognate NTP, as the reaction of incorporating cognate NTP is much favored than noncognate NTP for both wildtype and K235R RdRPs. Nevertheless, we also performed a spectrum scanning to survey if the higher fidelity of K235R mutant depends on the value of fast%. We plotted a 3D diagram for the discrimination (Z-axis) of wildtype and K235R RdRP against the fast% of CTP (X-axis) and UTP (Y-axis) (Fig. S9B). And it is evident that the plane representing the discrimination of K235R mutant is always above that of wildtype, indicating the conclusion of the higher fidelity of K235R mutant is valid regardless of the fast%.

1.9 MD simulations of IAV RdRp

The cryo-EM structure of the IAV elongation complex (PDBID: 6SZV (9)) was used as the structural basis to construct the IAV RdRp with the template-product duplex. Missing protein residues were modelled using modeller9.21 (13). To match with the experimental sequence, the UTP analog present in the cryo-EM structure was replaced with a CTP, and its base-paired template nucleotide was also mutated to G by Coot (14). The protonation states of histidine residues were predicted using propka3.0 module (15) in the pdbpqr2.2.1 (16) package, and manual examination was performed to ensure that the coordination with the magnesium ion can be formed as well as the hydrogen bond network can be maintained. The RdRp complex was placed in a dodecahedron box, the size of which was defined to ensure that the box edges at least 12 Å away from the complex surface. The complex was solvated with TIP3P water molecules (17), and sufficient counter ions were added to render the whole system neutral.

Amber99sb-ildn force field (18) was adopted to simulate protein and nucleotides. The default parameters for 3’-cytosine nucleotide were used for the base and sugar of the CTP, while the parameters for the triphosphate tail were taken from those developed by Meagher et al. (19). To reduce the instability caused by the excessive electrostatic repulsion between the two neighboring magnesium ions, dummy atom model (20, 21) was used for simulating the magnesium ions in the PA and PB1 domains.
First, the system was energy minimized for 10,000 steps. Afterwards, position restraint simulation under NVT ensemble (T=303K) was performed for 200 ps with a force constant of 10 \( \text{kJ} \times \text{mol}^{-1} \times \text{Å}^{-2} \) on all the heavy atoms of the complex, followed by another 500 ps position restraint simulation under NPT ensemble (T = 303 K, P = 1 bar). The last configuration after the position restrained simulations was used for performing five parallel 100 ns production simulations under NVT (T=303 K) ensemble, with the temperature gradually increases from 50K to 303K in the first 2 ns. In the production simulations, V-rescale thermostat (22) was applied with the coupling time constant of 0.1 ps. The long-range electrostatic interactions beyond the cut-off at 12 Å were treated with the Particle-Mesh Ewald method (23). Lennard-Jones interactions were smoothly switched off from 10 Å to 12 Å. The neighbor list was updated every 10 steps. An integration time step of 2.0 fs was used and the LINCS algorithm (24) was applied to constrain all the bonds. We saved the snapshots every 20 ps. All simulations were performed with Gromacs 5.0 (25).

Similar procedures as described above were adopted to construct and equilibrate the systems containing K235 or R239 mutants as well as those with UTP binding. Five parallel 100ns simulations with random seeds were performed for each mutant.

1.10 Structural analysis

The MD conformations in the first 20 ns of the 100 simulations were removed for the subsequent structural analysis. To measure the distance between NTP and K235, R239 or their mutant (Fig. 2F, Fig. 5C-5D, Fig. S11 and Fig. S12), we used the two oxygen atoms attached to the \( \alpha \), \( \beta \) or \( \gamma \)-phosphate atom of NTP to calculate their minimum distance to the atoms in the corresponding protein residue. The nitrogen atom (“NZ”) in the side chain of lysine (K), the nitrogen atoms (“NH1” and “NH2”) in the side chain of arginine (R) and the \( \beta \)-carbon atom of the alanine (A) were used in the distance calculations.

1.11 Investigation of K235R mutation on the active site configuration

First, we calculate the effect of the K235R mutation on the critical distances for CTP incorporation catalysis (Fig. S12). Firstly, we computed the distance between the two magnesium ions (Fig. S12A), which has been suggested to be less than 4.0 Å for the catalysis to occur (26, 27). We
discovered that the K235R mutant does not perturb this distance, which is 2.1 Å in both wildtype and K235R mutant systems (Fig. S12B). Secondly, we calculated the distance between the Pα atom in the incoming CTP and the O3' atom in the nascent strand's 3'-terminal nucleotide (Fig. S12A). This distance also needs to be smaller than 4.0 Å for the catalysis to occur (28, 29). Both the wildtype and K235R mutant systems exhibit an averaged distance of 3.5 Å (Fig. S12B), indicating that CTP can be efficiently incorporated in both systems. The above results indicate that the K235R mutant does not affect the critical distances for the catalysis, which is consistent with the experimental observation that both WT and K235R can efficiently synthesize RNAs.

Next, we examined the effect of the K235R mutant on the orientation of the R239 relative to the Pβ atom of CTP, which also plays an important role in the catalysis of the phosphoryl transfer reaction. As shown in Fig. S16B and S16E, the K235R mutant does not alter the relative position of R239's side chain to the Pβ atom of CTP. These results indicate that the mutation of K235R has negligible impact on the orientation of R239. However, we found that the K235R mutant makes itself approach closer to the Pβ atom than the wildtype K235 (Fig. S16B and S16E), and this observation is also consistent with the distance analysis in Fig. 5C.

Finally, we investigated the effect of the K235R mutation on the interactions between positively charged residues (R239 and K235/K235R) and the CTP triphosphate moiety. In particular, we discovered that the side-chain of R239 is in close contact (3.5 Å) with the Pα group of the incoming CTP in the wildtype system (black curve in Fig. S13A), which may serve to stabilize the transition state during the formation of the phosphodiester bond. The K235R mutant well maintains this interaction, with the mean distance equal to 3.5 Å (red curve in Fig. S13A). Interestingly, we discovered that the K235R mutant form new interactions between K235R and the Pα & Pγ groups of CTP (red curves in Fig. S13B and S13C), and these interactions are absent in the wildtype system (black curve in Fig. S13B and S13C). These newly formed interactions between K235R/R239 and the triphosphate moiety of CTP well explain the synergistic effects of K235R and R239 to facilitate the NTP incorporation.
1.12 Comparison between the K235R mutation in IAV RdRp and K65R mutation in HIV-1 RT

In HIV-1 reverse transcription (RT), the side chain of K65 is in contact with the substrate's triphosphate moiety at a minimum distance of 2.4 Å (see vdW representations of K65 and NTP in Fig. S14A). In IAV RdRp, by contrast, the side chain of K235 is located at a distance of 3.7 Å or larger from the substrate's triphosphate moiety (Fig. S14C). As a result, we hypothesize that HIV-1 RT has a more compact active site. To investigate the effect of the Lys-to-Arg mutation on the space between the substrate and the respective protein residue in HIV-1 RT, we also generated structural model by mutating the Lysine residue while retaining its orientation to the substrate. We discovered that the K65R mutation brings it closer to two phosphate atoms of substrate at distances of 2.4 Å and 2.7 Å, respectively (Fig. S14B). The Lys-to-Arg mutation in IAV RdRp, on the other hand, allows the residue to be at least 3.0 Å away from the substrate (Fig. S14D). These findings suggest that the K65R mutation in HIV-1 RT could render the active site too crowded and may significantly reduce the active site's stability. This could also explain why the K65R mutation impairs transcription activity.

1.13 Calculations of relative binding free energy

We calculated the relative binding free energy (ΔΔG) between the K235R mutant and wildtype RdRp for both cognate CTP and non-cognate UTP. To perform these calculations, we applied the free energy perturbation method (30, 31) based on explicit solvent MD simulations. Based on the thermodynamic cycle as shown in Fig. S15A, we can obtain difference in the binding free energy (ΔΔG=ΔG4−ΔG3, see Fig. S15A) of a particular NTP between K235R (ΔG4) and wildtype (ΔG3) by calculating the free energy changes due to the K235R mutation in the NTP-unbound (ΔG1) and NTP-bound state (ΔG2) separately. The Hamiltonian of the system is coupled to a parameter λ, which was used to drive the system from state wildtype (λ=0) to state K235R mutant (λ=1). In our calculations, we chose 26 λ values (λ= 0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.42, 0.44, 0.46, 0.48, 0.50, 0.52, 0.54, 0.56, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1), conducted 40-ns simulations at each λ, and finally obtained the binding free energy using the Multistate Bennett Acceptance Ratio method (32, 33). In particular, for each λ, 10,000-steps energy minimization was first performed, followed by 200 ps position restraint simulations (force constant
of 10 kJ×mol⁻¹×Å⁻² on all the heavy atoms) under NVT (T=303 K) ensemble and another 200ps under NPT (T=303 K and P=1 bar) ensemble. The last conformation of the position restraint simulation was used to seed one 40 ns simulation under NVT ensemble (T=303 K). The convergence of the relative binding free energy has been validated by every 5ns and the results become plateau after ~30 ns (Fig. S15C). Consistent with the tendency as observed for the $K_{d,app}$ of cognate CTP versus non-cognate UTP upon mutation, our results clearly indicate the less favourable binding free energy of UTP compared to CTP (the difference is ~3.4 kJ/mol, see Fig. S15B).

We noticed that our relative binding free energy suggests that K235R improves the binding of CTP ($\Delta\Delta G=-5.2±0.5$ kJ/mol) and UTP ($\Delta\Delta G=-1.8±0.7$ kJ/mol) to the active site (Fig. S15B), while the experimentally measured $K_{d,app}$ suggests a higher dissociation constant for K235R (Table S1). We anticipate that this difference is due to the contributions of the NTP loading. In particular, the longer sidechain of Lys-to-Arg mutation in K235R may introduce steric effect to hinder the NTP loading (leading to a larger value of $K_{d,app}$), as K235R lies at the exit of the NTP loading channel before entering the active site. This inhibitory effect would exist for all NTP regardless of the base, and thus lead to an increase in $K_{d,app}$ for both CTP and UTP at a similar magnitude.

1.14 Geometric clustering to select representative configurations

To select the representative configurations, we applied the K-Center clustering algorithm (34, 35) to divide the MD configurations into five groups based on the geometry similarity of the heavy atoms in NTP and two positively charged protein residues (resids 235 and 239). Notably, we found that the distributions of the K235R-UTP system contain two clusters with significant populations (see these two representative conformations in Fig. S17D). Interestingly, one of these configurations (Right panel of Fig. S17D with population 27%) displays a relatively larger distance between K235R and $\beta$-phosphate of the UTP, possibly due to larger fluctuations of the phosphate tail induced by the relatively unstable wobble base pair. In contrast, all the other three systems (WT-CTP, WT-UTP, and K235R-CTP) are dominated by a single cluster (see the representative conformation for each system in Fig. S17A-C). These observations indicate that the conformational flexibility is most pronounced in the K235R mutant RdRp with non-cognate UTP, which is consistent with the observations in Fig. S16.
1.15 NGS sequencing of wildtype and PB1-K235R recombinant live viruses

We generated the wild-type and PB1-K235R mutant A/WSN/1933(H1N1) recombinant virus from eight pHW2000 plasmids encoding each of the IAV genome segments, including either the wild-type or mutant K235 PB1 gene. The titres of passage 1 (P1) viruses were determined using the plaque assay, and the P2 viruses were passaged by inoculating P1 viruses into MDCK cells at a multiplicity of infection (MOI) of 0.001. In 10 cm culture dishes, P3 viruses were multiplied by inoculating MDCK cells with P2 viruses at an MOI of 0.001 and subjected to plaque assays to quantify the viral titres (Fig. 6A). 10 mL of media containing P3 wild-type and K235R mutant viruses were collected and centrifuged at 500 × g for 5 minutes to eliminate floating cell debris. Amicon Ultra-15 100 K was used to concentrate the supernatant medium to 500 µl at 1500 × g, 4 °C for 30 minutes. The residual genomic DNA and RNA impurities were digested for 30 minutes on ice using Micrococcal Nuclease (NEB). NucleoZOL (MACHEREY-NAGEL) was used to extract viral RNA from the medium, and the Monarch RNA Cleanup Kit (NEB) was used to purify the viral RNA according to the instructions. The Ct value for 1 µl of purified RNA was determined to be between 18 and 20 using the Luna One-Step RT-qPCR Kit (NEB) in conjunction with the 3' consensus RT-PCR primer (MBTuni-12, 5'-ACGCGTGATCAGCAAAAGCAGG) and the 5' consensus PCR primer (MBTuni-13, 5'-ACGCGTGATCAGTAGAACAAGG) as described in the published protocol (36). To amplify the viral RNA, a 5 µl sample was reverse transcribed using ProtoScript II Reverse Transcriptase (NEB) using the MBTuni-12 primer and amplified for 22 cycles using Q5 Hot Start High-Fidelity DNA Polymerase (NEB) using the MBTuni-12 and MBTuni-13 primers. Monarch PCR & DNA Cleanup Kit was used to purify the PCR products (NEB). The NGS collection was constructed using 2 ng DNA from two separate cultures of P3 wild-type and K235R mutant IAV mediated by the Tn5-tagmentation method (37). To summarize, the viral genome PCR results were processed with Tn5 transposase, resulting in major products of tagged dsDNA fragments of around 300 bp. Twelve cycles of PCR were used to repair and index the dsDNA fragment, which was then fed into the MGI DNBSEQ-G400 platform.

The sequencing reads were first trimmed in paired-end mode with cutadapt (38) using default parameter and mapped into reference IAV genome (GenBank Source Sequence Accession: LC333182 to LC333189) using bwa (39). Reads with Indel and soft-clipped were removed. Mapped reads (sam file) were sorted and converted to bam files by samtools (40) and the Mpileup
files were generated by samtools with the following criteria: 1) Sequencing base with quality scores over 30, 2) Mapping quality over 40. The number of mutation sites was counted by screening the mpileup with the following criteria: 1) Error rate $\geq 1\%$, 2) Depth $\geq \frac{1000}{\text{Error rate}(%)}$.

We analysed and compared wild-type and K235R viral genome from two biological replicates. The average sequencing depths of both wild-type and K235R are $\sim10^5$ per replicate (Fig. 6B). The number of SNV sites in wildtype live virus is $19.25 \pm 0.50$ (Mean $\pm$ SD), whereas the number of SNV sites in K235R is $12 \pm 2.94$ (Mean $\pm$ SD). Our results suggest that K235R is indeed a high-fidelity mutant at the virus level. Our findings imply that K235R is really a virus-level high-fidelity mutant (Fig. 6C and 6D).
Section 2 Supplementary Table and Figures:

|       | $k_{pol}$ (min$^{-1}$) | $K_{d,app}$ (μM) | $k_{pol}/K_{d,app}$ (μM$^{-1}$min$^{-1}$) | Discrimination | Fold change |
|-------|-------------------------|------------------|------------------------------------------|----------------|-------------|
| **CTP** |                         |                  |                                          |                |             |
| WT    | 219 ± 30                | 126 ± 56         | 1.7 ± 0.8                                | 41.9 ± 22.1    | 1           |
| K235R | 620 ± 89                | 186 ± 77         | 3.3 ± 1.5                                | 192.4 ± 94.3   | 4.6 ± 3.3   |
| **UTP** |                         |                  |                                          |                |             |
| WT    | 6.8 ± 0.5               | 163 ± 39         | 0.04 ± 0.010                             |                |             |
| K235R | 7.8 ± 0.7               | 449 ± 88         | 0.02 ± 0.004                             |                |             |

Table. S1. Table summary of kinetic parameters for CTP and UTP incorporation against guanosine template.
Fig. S1. The structure diagrams of RNA virus RdRps and sequence alignment of viruses from the Orthomyxoviridae family. (A) The relative location of Arginine and Lysine residues on motif F to the incoming NTP in the catalytic center is structurally conserved among different types of RNA viruses. The K235 and R239 in IAV RdRp (PDB: 6SZV), the K174 and R182 in Norwalk virus RdRp (PDB: 3BSO), the R268 and R270 in Φ6 RdRp (PDB: 1HI0), the K65 and R72 in HIV-1 RT (PDB: 1RTD), the K167 and R174 in poliovirus RdRp (PDB: 3OLA), the K908 and R914 in Qβ RdRp (PDB: 3AVT) are shown together with the incoming NTP. The motif F was rendered in light yellow. (B) sequence alignment of IAV, IBV, ICV, IDV, and Salmon isavirus on the motif F of PB1 subunits. The equivalent K235 (green) and R239 (blue) in IAV are well conserved in other viruses of the Orthomyxoviridae family.
Fig. S2. The Isoleucine residue on motif F may play important roles in template-product RNA duplex translocation. (A) Schematic representative showing the I241 on the motif F of IAV RdRp and (B) I176 on the motif F of Enterovirus 71 (EV71) RdRp. The I176 in EV71 RdRp was proposed to be one of the critical residues in template-product RNA duplex translocation. The location of the isoleucine residue of IAV and EV71 RdRps shows a high similarity in the elongation complex. The RNA products, template, incoming CTP, and isoleucine residues are in red, cyan, orange, and magenta, respectively.
Fig. S3. Dual-luciferase minigenome assay on R239 and K235 mutant IAV RdRps. The relative Gaussian luciferase activity from wild-type and mutant PB1 was normalized to the ΔPB1 group (100%). Error bars are ±SD, n = 3, one-way ANOVA with Dunnett’s multiple comparison test, ****P < 0.0001.
Fig. S4. Gel stain and western blot of purified wild-type and mutant IAV RdRps. (A) Purified wild-type (WT), R239, and K235 mutant polymerases were resolved by 8% SDS-PAGE and stained by Coomassie blue stain (Abcam, ab119211). BSA standards with known mass from 30 ng, 60 ng, 120 ng to 180 ng were prepared and loaded parallel with the wild-type and mutant polymerases. The mass of wild-type and mutant polymerases were estimated by BSA standard curve to between 5 to 7 ng/µl. (B) The gel stain of wild-type and K235R mutant IAV polymerase. Bands represent PB2-TAP, PA subunits were labelled. The co-purified Hsp70 is also indicated. (C) An equal volume of purified wild-type (WT) and mutant polymerases were resolved by 10% SDS-PAGE and incubated with either anti-PB1 (Thermo Fisher, PA5-34914), anti-PB2 (Thermo Fisher, PA5-32220), or anti-PA (Thermo Fisher, PA5-31315) antibody.
Fig. S5. *In vitro* transcription with a series dilution of radioisotope labelled primer. Series dilution of \([\alpha-32P]\) GTP labelled capped primer with a concentration of 200 counts per seconds (cps), 100 cps, 50 cps, 25 cps, and 12.5 cps were used to initiate the *in vitro* transcription of a 30-nt 3'-vRNA template in an equal amount of wild-type polymerases at 30 °C for 1 hour. The relative extension percentage was calculated by dividing the intensity of the extension product over the total intensity from the extension product and unused primer of that lane. The extension percentages remain the same until the primer cps higher than 50 cps, indicating an excessive primer were used in the reaction.
Fig. S6. Time and concentration-dependent single-turnover CTP and UTP incorporation on guanosine template. (A) Gel image of the single-turnover of CTP:G and UTP:G incorporation assay on purified wild-type (WT) and K235R mutant polymerases. The in vitro incorporation was initiated by mixing polymerase-template complexes with 1000 µM, 500 µM, 200 µM, 100 µM, 50 µM, 10 µM CTP or UTP. The CTP:G incorporation was stopped at 0.5s, 2s, 5s, 10s, 30s, 60s, while the UTP:G incorporation was stopped at 5s, 15s, 30s, 60s, 120s, 300s, 600s. The reaction was resolved on 20% denaturing urea polyacrylamide gel. The upper band represents the incorporated nucleotide, and the lower band represents the unincorporated primer. (B) Two-phase exponential fitting of the incorporation products. The quantified and normalized incorporation products were fitted into a biphasic exponential curve and plotted over time.
Fig. S7. Schematic representations of *in vitro* transcription of all pairing scenarios in single-turnover NTP incorporation assay. (A) The IAV genome's highly conserved 5' and 3' UTR sequences form a unique corkscrew structure within the IAV RdRp, which is required to maintain polymerase activity. (B) The guanine (G) of the third base from the template's 3' end serves as the fidelity checkpoint following transcription initiation. (C) Schematic representation of the incorporation of a single nucleotide from CTP and UTP pairing with G. (D) A typical 20% denaturing gel of a single-turnover CTP incorporation assay demonstrates that only one round of CTP addition occurs. (E) Schematic representation of the possible incorporation of multiple nucleotides from GTP and ATP against Us following the initial mispairing with G. (F) A typical 20% denaturing gel of a single-turnover GTP incorporation assay demonstrates that GTP will form wobble pairings with the following Us: Miss-pairing of G(template), resulting in +2 and +3 nt extension products.
Fig. S8. **Comparison of multiple in vitro transcription schemes.** (A) The capped primer terminating in AG will generate multiple extension products, including those initiated from C and G in the third and second positions on the 3' template, respectively, as well as the re-alignment products. (B) A 14-nt or 30-nt 3' template was used to initiate the in vitro transcription assay of wild-type RdRp. For 1 hour, the reactions were carried out in parallel at 30 °C. In denaturing gel, transcription products were resolved. (C) Illustration of a single-turnover NTP incorporation assay used to determine NTP's specificity. The NTP pairing with G in the third position of the template's 3' end will be evaluated. (D) Only CTP or UTP is used to initiate the single-turnover assay. The wobble pairing priming route between the capped primer's ending G and the template's initial C will not proceed because no GTP is present in the reaction to pair with the template's coming C.
Fig. S9. SSE calculation and percentage of fast phase determination. (A) Bar charts represent the SSE against the percentage of the fast phase. (B) The 3D diagram for the discrimination (Z-axis) of wild-type and K235R RdRp against the percentage of the fast phase of CTP (X-axis) and UTP (Y-axis).
**A Primer labeling**

+ Capping enzyme  
+ Radioactive GTP  

\[ 5'\text{-ppGAAUACUAG-3'} \xrightarrow{\text{+Capping enzyme}} 5'\text{-m7G-GAAUACUAG-3'} \]

**B In vitro transcription**

- **Full transcription**
  
  \[ \text{m7G-GAAUACUAG} \xrightarrow{+\text{NTP}} \text{+ Template: UCGUUUCGUCCAG} \]

- **C-G correct pairing**
  
  \[ \text{m7G-GAAUACUAG} \xrightarrow{+\text{CTP}} \text{+ Template: UCGUUUCGUCCGG} \]

- **U-G wobble pairing**
  
  \[ \text{m7G-GAAUACUAG} \xrightarrow{+\text{UTP}} \text{+ Template: UCGUUUCGUCCGG} \]

**Fig. S10. RNA primer capping and in vitro transcription assay reaction scheme.** (A) The diagram shows the radioactive [α-32P] GTP capping of the 11nt pp-RNA primer by the vaccinia capping enzyme. The right panel shows the capped RNA as a 12nt band resolved on 20% Urea-PAGE RNA gel. (B) Different in vitro transcription assay schemes. The two vertical lines on the 3’ terminal AG of the RNA primer represent the initial priming sites. The arrows represent the direction of transcription.
Fig. S11. The critical distance between $P_\beta$ and 235/239 residue in wildtype, R239K, and R239A mutant RdRps. The x-axis represents the distance between the $P_\beta$ group and the terminal N atom of the side chain of K235. The y-axis represents the distance between the $P_\beta$ group and the terminal atom of the side chain of R239 or R239K/A (see supplementary section 1.10 for details).
Fig. S12. Investigation of the impact of K235R mutant on the critical distance for catalysis. (A) Schematic representative showing the MgA-MgB and O3’-Pα distances; (B) The distances calculated using the MD conformations of wildtype and K235R mutant systems. The means and standard deviations were estimated by combining all the five trajectories of each system after removing the first 20ns in each trajectory.
Fig. S13. Examination over the influence of K235R mutant on the distance from the side chain of R239/K235/K235R to the triphosphate moiety of CTP. (A) Histogram of the distance between R239 and $P_\alpha$ group of CTP in the wildtype (in black) and K235R mutant (in red) systems. (B) Histogram of the distance between residue 235 and $P_\alpha$ group of CTP in the wildtype (in black) and K235R mutant (in red) systems. (C) Similar to (B) but $P_\gamma$ group of CTP was used for the calculation. The histogram was computed by using the five trajectories with the conformations from the first 20 ns removed. See supplementary section 1.10 for details of distance calculations.
Fig. S14. Comparison of the active site configuration in IAV RdRp with that in HIV-I RT.
(A) The relative position of K235 (shown in yellow) and substrate in HIV-I RT. The two oxygen atoms attached with $P_{\alpha}$ and $P_{\gamma}$ atoms of the substrate are shown in red spheres. The nitrogen atom in the side chain of K65, which is closest to the triphosphate moiety of the substrate is shown in the yellow sphere. (B) The same as (A) but with K65R mutant. (C) The relative position of K235 (shown in yellow) and substrate in IAV RdRp. The similar molecular representation as used in (A) is applied. (D) The same as (C) but with K235R mutant.
Fig. S15. Investigation of the impact of K235R mutant on the binding free energy of NTP. (A) The thermodynamic cycle was designed to calculate how K235R mutant affects the binding free energy of NTP. (B) Comparison of the relative binding free energy of cognate CTP with that of non-cognate UTP. (C) Validation of the convergence of the relative binding free energy for CTP and UTP upon K235R mutation.
Fig. S16. The impacts of the K235R mutant on the relative orientation of R239 and K235/K235R side-chain to the Pβ atom of NTP. (A) The definition of the coordinate system using the base of NTP (shown in orange). The Pβ atom of NTP, the nitrogen atom of R239 and K235 used for plotting the relative orientation are shown in orange, green, and blue spheres, respectively. (B) The x and y coordinates of Pβ atom of CTP in the wildtype IAV RdRp, the nitrogen atom of R239 and K235 are shown in red, orange, and blue contours. For R239, the NH1 or NH2 atom, which is nearer to the Pβ atom is plotted. (C) The same as (B) but for wildtype IAV RdRp with UTP binding at the active site. (D) The same as (A) but for K235R mutant. (E) The same as (B) but for K235R mutant IAV RdRp with CTP binding at the active site. (F) The same as (C) but for K235R mutant IAV RdRp with UTP binding at the active site. In (B)-(C) and (E)-(F), the density of the contour plots was estimated by combining the conformations from all the five 100ns trajectories after removing the data in the first 20ns.
Fig. S17. Representative configurations of the NTP and two positively charged residues in the MD simulations. The NTP is shown in orange. R239 and K235/K235R are displayed, with dashed lines connecting between their nitrogen atoms and the oxygen atoms attached to the Pβ atom of NTP. The populations of the respective group where the representative configuration belongs to are labelled.
Fig. S18. NTP analog incorporation on wild-type and K235R IAV polymerases. 500 µM of 3’dCTP and 2’F-dCTP were mixed with wild-type (WT) and K235R polymerases. The reaction was incubated for one hour and subjected to 20% denaturing urea polyacrylamide gel. The +1 site represents the incorporated NTP analog into the primer.
References:

1. Velthuis, te, A.J.W., Long, J.S. and Barclay, W.S. (2018) Assays to Measure the Activity of Influenza Virus Polymerase. In Influenza Virus, Methods and Protocols. Springer New York, New York, NY, Vol. 1836, pp. 343–374.

2. Velthuis, te, A.J.W., Robb, N.C., Kapanidis, A.N. and Fodor, E. (2016) The role of the priming loop in influenza A virus RNA synthesis. Nat. Microbiol, 1, 16029.

3. Kellinger, M.W., Song, C.-X., Chong, J., Lu, X.-Y., He, C. and Wang, D. (2012) 5-formylcytosine and 5-carboxylyctosine reduce the rate and substrate specificity of RNA polymerase II transcription. Nat. Struct. Mol. Biol., 19, 1–4.

4. Tomescu, A.I., Robb, N.C., Hengrung, N., Fodor, E. and Kapanidis, A.N. (2014) Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. Proc. Natl. Acad. Sci. U.S.A., 111, E3335–42.

5. De Vlugt, C., Sikora, D. and Pelchat, M. (2018) Insight into Influenza: A Virus Cap-Snatching. Viruses, 10, 641.

6. Koppstein, D., Ashour, J. and Bartel, D.P. (2015) Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation. Nucleic Acids Res., 43, 5052–5064.

7. Oymans, J. and Velthuis, te, A.J.W. (2018) A Mechanism for Priming and Realignment during Influenza A Virus Replication. Journal of Virology, 92.

8. De Vlugt, C., Sikora, D., Rocheleau, L. and Pelchat, M. (2019) Priming and realignment by the influenza a virus RdRp is dependent on the length of the host primers and the extent of base pairing to viral RNA. Virology, 536, 91–100.

9. Wandzik, J.M., Kouba, T., Karuppasamy, M., Pflug, A., Drncová, P., Provazník, J., Azevedo, N. and Cusack, S. (2020) A Structure-Based Model for the Complete Transcription Cycle of Influenza Polymerase. Cell, 10.1016/j.cell.2020.03.061.

10. Kireeva, M.L., Nedialkov, Y.A., Cremona, G.H., Purto, Y.A., Lubkowska, L., Malagon, F., Burton, Z.F., Strathern, J.N. and Kashlev, M. (2008) Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. Mol. Cell, 30, 557–566.

11. Joyce, C.M. and Benkovic, S.J. (2004) DNA polymerase fidelity: kinetics, structure, and checkpoints. Biochemistry, 43, 14317–14324.

12. Johnson, K.A. (1993) Conformational coupling in DNA polymerase fidelity. Annu Rev Biochem, 62, 685–713.

13. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys, 79, 926–935.
14. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O. and Shaw, D.E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, **78**, 1950–1958.

15. Meagher, K.L., Redman, L.T. and Carlson, H.A. (2003) Development of polyphosphate parameters for use with the AMBER force field. *J Comput Chem*, **24**, 1016–1025.

16. Duarte, F., Bauer, P., Barrozo, A., Amrein, B.A., Purg, M., Aqvist, J. and Kamerlin, S.C.L. (2014) Force field independent metal parameters using a nonbonded dummy model. *J Phys Chem B*, **118**, 4351–4362.

17. Jiang, Y., Zhang, H., Feng, W. and Tan, T. (2015) Refined Dummy Atom Model of Mg(2+) by Simple Parameter Screening Strategy with Revised Experimental Solvation Free Energy. *J Chem Inf Model*, **55**, 2575–2586.

18. Webb, B. and Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci*, **86**, 2.9.1–2.9.37.

19. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallographica Section D*, **66**, 486–501.

20. Olsson, M.H.M., Søndergaard, C.R., Rostkowski, M. and Jensen, J.H. (2010) PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p KaPredictions. *J. Chem. Theory Comput.*, **7**, 525–537.

21. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A. and Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **32**, W665–7.

22. Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J Chem Phys*, **126**, 014101.

23. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G. (1995) A smooth particle mesh Ewald method. *J Chem Phys*, **103**, 8577–8593.

24. Hess, B., Bekker, H., Berendsen, H.J.C. and Fraaije, J.G.E.M. (1997) LINCS: A linear constraint solver for molecular simulations. *J Comput Chem*, **18**, 1463–1472.

25. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B. and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1-2**, 19–25.

26. Chaudret, R., Piquemal, J.-P. and Andrés Cisneros, G. (2011) Correlation between electron localization and metal ion mutagenicity in DNA synthesis from QM/MM calculations. *Phys. Chem. Chem. Phys.*, **13**, 11239–11247.
27. Gleghorn, M.L., Davydova, E.K., Basu, R., Rothman-Denes, L.B. and Murakami, K.S. (2011) X-ray crystal structures elucidate the nucleotidyl transfer reaction of transcript initiation using two nucleotides. Proc Natl Acad Sci USA, 108, 3566–3571.

28. Radhakrishnan, R. (2007) Coupling of fast and slow modes in the reaction pathway of the minimal hammerhead ribozyme cleavage. Biophys J, 93, 2391–2399.

29. Sgrignani, J. and Magistrato, A. (2012) The Structural Role of Mg2+ Ions in a Class I RNA Polymerase Ribozyme: A Molecular Simulation Study. J Phys Chem B, 116, 2259–2268.

30. Zwanzig, R.W. (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. J Chem Phys, 22, 1420–1426.

31. Jorgensen, W.L. and Ravimohan, C. (1985) Monte Carlo simulation of differences in free energies of hydration. J Chem Phys, 83, 3050–3054.

32. Shirts, M.R. and Chodera, J.D. (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys, 129, 124105.

33. Klimovich, P.V., Shirts, M.R. and Mobley, D.L. (2015) Guidelines for the analysis of free energy calculations. J Comput Aided Mol Des, 29, 397–411.

34. Gonzalez, T.F. (1985) Clustering to minimize the maximum intercluster distance. Theoretical Computer Science, 38, 293–306.

35. Beauchamp, K.A., Bowman, G.R., Lane, T.J., Maibaum, L., Haque, I.S. and Pande, V.S. (2011) MSMBuild2: Modeling Conformational Dynamics on the Picosecond to Millisecond Scale. J. Chem. Theory Comput., 7, 3412–3419.

36. Zhou, B., Donnelly, M.E., Scholes, D.T., St George, K., Hatta, M., Kawaoka, Y. and Wentworth, D.E. (2009) Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses. Journal of Virology, 83, 10309–10313.

37. Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S. and Sandberg, R. (2014) Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc, 9, 171–181.

38. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, 17, 3–12.

39. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25, 1754–1760.

40. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Subgroup, 1.G.P.D.P. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25, 2078–2079.