A novel mechanism of enhanced transcription activity and fidelity for influenza A viral RNA-dependent RNA polymerase

Xinzhou Xu[1,5], Lu Zhang[2,6,*], Julie Tung Sem Chu[3], Yuqing Wang[1,5],
Alex Wing Hong Chin[3,10], Tin Hang Chong[1,4], Zixi Dai[4], Leo Lit Man Poon[3,9,10],
Peter Pak-Hang Cheung[1,4,7,8,*] and Xuhui Huang [1,4,*]

[1]The Hong Kong University of Science and Technology-Shenzhen Research Institute, Hi-Tech Park, Nanshan, Shenzhen 518057, China, [2]State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, Fujian 350002, China, [3]School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China, [4]Department of Chemistry, State Key Laboratory of Molecular Neuroscience, The Hong Kong University of Science and Technology, Kowloon, Hong Kong, China, [5]Bioengineering Graduate Program, Department of Biological and Chemical Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China, [6]University of Chinese Academy of Sciences, Beijing 100049, China, [7]Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong, China, [8]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, [9]HKU-Pasteur Research Pole, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China and [10]Centre for Immunity and Infection, Hong Kong Science Park, Hong Kong, China

ABSTRACT

During RNA elongation, the influenza A viral (IAV) RNA-dependent RNA polymerase (RdRp) residues in the active site interact with the triphosphate moiety of nucleoside triphosphate (NTP) for catalysis. The molecular mechanisms by which they control the rate and fidelity of NTP incorporation remain elusive. Here, we demonstrated through enzymology, virology and computational approaches that the R239 and K235 in the PB1 subunit of RdRp are critical to controlling the activity and fidelity of transcription. Contrary to common beliefs that high-fidelity RdRp variants exert a slower incorporation rate, we discovered a first-of-its-kind, single lysine-to-arginine mutation on K235 exhibited enhanced fidelity and activity compared with wild-type. In particular, we employed a single-turnover NTP incorporation assay for the first time on IAV RdRp to show that K235R mutant RdRp possessed a 1.9-fold increase in the transcription activity of the cognate NTP and a 4.6-fold increase in fidelity compared to wild-type. Our all-atom molecular dynamics simulations further elucidated that the higher activity is attributed to the shorter distance between K235R and the triphosphate moiety of NTP compared with wild-type. These results provide novel insights into NTP incorporation and fidelity control mechanisms, which lay the foundation for the rational design of IAV vaccine and antiviral targets.

INTRODUCTION

Pandemic and seasonal influenza A virus (IAV) poses a significant threat to the avian, swine and human populations. It belongs to the negative-sense, segmented RNA virus from the *Orthomyxoviridae* family that comprises the genera *Influenzavirus* (A, B, C, D), *Thogotovirus*, *Quaranjavirus* and *Isavirus*. The RNA-dependent RNA polymerase (RdRp) is the central machinery for all transcription and genome replication processes of viruses and is implicated in evolution, transmission, pathogenesis and host range restriction ([1]). The IAV RdRp is a processive nucleic acid polymerase, which performs template-directed biosynthesis of long nucleotide acid chains using nucleoside triphosphates (NTPs) as substrates. The catalytic core of the IAV trimeric RdRp contains the pre-A (also known as F) and A-

E motifs. Structure-based alignment with RNA- and DNA-dependent RNA polymerases, including viruses from *Picornaviridae*, *Flaviviridae*, *Caliciviridae*, *Leviviridae*, *Cystoviridae* and *Retroviridae* families, shows a well-conserved configuration in Motif F, which has been speculated to play essential roles in controlling the rate and fidelity of viral transcription (2,3). During each nucleotide addition cycle, a phosphodiester bond is formed between the terminal 3′-hydroxyl group of the RNA product strand and the α-phosphate group of incoming NTP. Acid-base reaction involving the β-phosphate group is one of the essential steps in the catalysis, and positively charged residues in the motif F near the β-phosphate are expected to be critical for the phosphoryl transfer reaction (4,5). Interestingly, these viral RNA polymerases adopt highly conserved functional dual residues in the motif F to interact with the triphosphate moiety of the NTP substrates, which consists of a primary arginine residue located in the vicinity of the α-, β-phosphate moiety, a secondary arginine or lysine residue close to the β-, γ-phosphate moiety (Figure 1 and Supplementary Figure S1A). Previous studies have suggested two main functions of these dual residues based on structural or theoretical analyses. They may either probe and align with the triphosphate moiety to form the hydrogen bond and facilitate the NTP binding (6) or act as a proton donor for the pyrophosphate leaving group (7,8).

Due to the similarities in the active site architecture of IAV RdRp with other RdRps (Supplementary Figure S1A), it is therefore expected that IAV RdRp adopts a similar mechanism of employing positively charged residues to facilitate catalysis. The RdRp of the IAV possesses highly conserved R239 and K235 in Motif F that is present in all viruses classified to the *Orthomyxoviridae* family, including influenza A, B, C and D viruses (Figure 1A, B and Supplementary Figure S1). From a recently resolved bat IAV elongation complex with an incoming nonhydrolyzable UTP analog substrate (9), these two residues, namely R239 and K235, are closely located to the triphosphate moiety of the NTP in the elongating RdRp with both side chains pointing towards the β-phosphate group (Figure 1A). The relative positions are suggestive of their significances in NTP incorporation. However, the exact functions of these two positively charged residues are yet to be elucidated for IAV and most viral RdRps. Previous work established that point mutations in the active site residues usually resulted in the loss of polymerase activity. As a result, quantifying fidelity on those mutant RdRps is difficult. For example, K65 in HIV-1 retrotranscriptase (RT), K159 in coxsackievirus B3 (CVB3) and K155 in Hepatitis C Virus (HCV) RdRps are residues that are similar to K235 in IAV RdRp. Previously published researches have demonstrated that the K65R, K159R and K155A mutations significantly reduced the polymerase activity (10–12). Similarly, R174 in Poliovirus (PV) RdRp and R158 in Hepatitis C virus (HCV) RdRp structurally corresponds to R239 in IAV RdRp, and their mutations such as R174K and R158A also significantly reduced polymerase activity (13,14). Identifying the critical roles of these residues is of paramount importance to understand the mechanism of transcription of influenza viruses and the evolutionary pathway of viruses in the *Orthomyxoviridae* family, as well as the development of novel drugs

and vaccines that target NTP incorporation efficiency and fidelity.

With the absence of proofreading mechanisms in IAV RdRp, the incorporation of the cognate against noncognate NTP is solely dictated by the stability of the base-pairing between the template and incoming NTP tightly networked by the residues in the active site. The fidelity control mechanisms employed by different viruses may differ due to structural differences in RdRp and the presence or absence of exonuclease domains (15). Although no direct investigation of the kinetics of IAV RdRp's nucleotide addition cycle has been conducted, we hypothesize that the rate-limiting step in the nucleotide addition cycle of IAV RdRp is generally contributed by active site closure and phosphoryl transfer based on the study of poliovirus RdRp, which shares structural similarities with IAV RdRp in the catalytic region (16). Similar rate-limiting steps have been proposed for RNA Polymerase II in yeast (17). Motif F may play important roles in rate-limiting steps as well as fidelity control in the IAV RdRp nucleotide addition cycle. Recent research on influenza polymerase has revealed that the motif F is required for NTP binding and stabilization (15). Our single-turnover NTP incorporation assays and MD simulations further revealed that the positively charged residues R239 and K235 on motif F are required for substrate stabilization at the active site and can be used to manipulate fidelity. Additionally, we noticed the side chain of I241 on motif F near the base moiety of the + 1 template and thus may affect the translocation dynamics of IAV RdRp (Supplementary Figure S2). Consistently, a recent publication has suggested that I176 on the motif F of Enterovirus 71 (EV71) RdRp, which is corresponding to I241 in IAV RdRp, may be a critical residue in the template-product RNA duplex translocation (18).

The fidelity of transcription, which is dependent on the incorporation rates for the cognate over noncognate NTP, was estimated to be $\sim 10^{-3}$ to 10^{-4} errors per base pair in IAV (19). This mutation rate is equivalent to one mutation per newly synthesized virus. IAV RdRp mutation during viral replication can give rise to the dynamic changes on influenza virus proteins that allow rapid acquisition of drug resistance, evasion of host immune responses and new host adaptation (20). However, IAV RdRp must balance evolutionary fitness and the fidelity of transcription to ensure faithful virus replication. This balance has been experimentally exploited for clinical applications, where fidelity-variant viruses generated by point mutations of the RdRp are excellent candidates for the attenuated vaccine design (21,22). However, all the discovered high-fidelity IAV mutants so far have collaterally coupled with compromised enzyme activity (21,23), posing a challenge in the development of vaccines against emerging pandemic diseases. For example, we previously discovered the high fidelity V43I mutation in PB1 subunits of IAV RdRp with a slower rate for incorporating GTP analogs (21,23). A recent study also reported an IAV RdRp with R229K point mutation in PB1, which exhibited elevated fidelity and resistance to favipiravir (24). Unfortunately, the latter example involves a mutant polymerase with severe compromise in enzyme activity in the absence of other compensatory mutations. This common drawback hinders mutant viruses from being
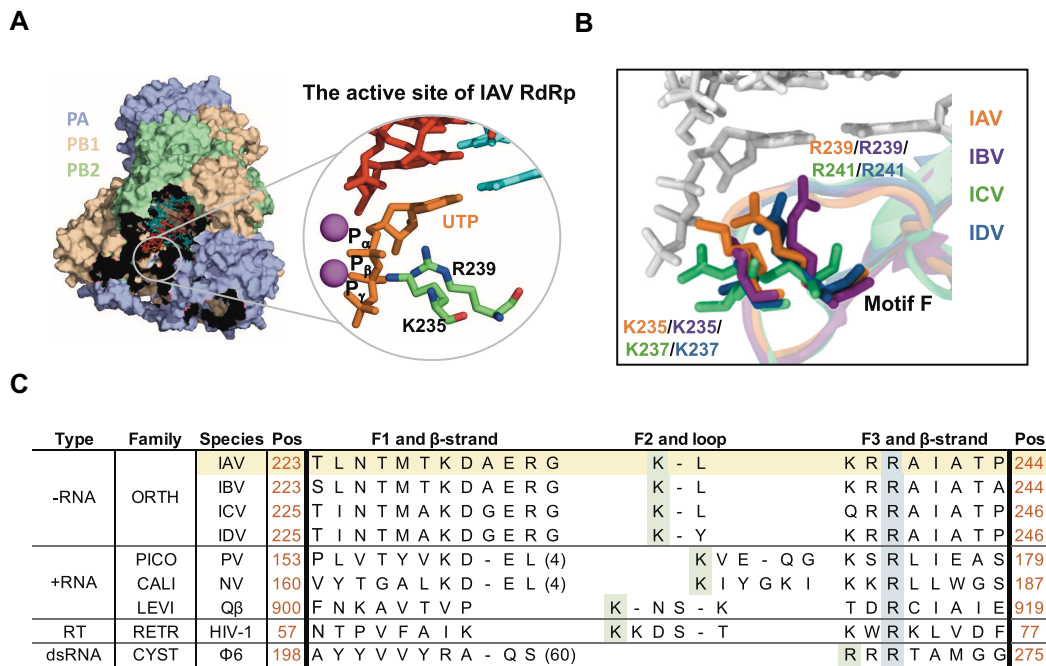
Figure 1. Diagrams showing the cartoon structure of IAV polymerase active site and structure-based alignment on the motif F of various RNA viral RdRps. (**A**) The elongation complex of bat IAV RdRp with the magnified active site. The left structure diagram exposed the active site of the elongation complex (PDB: 6SZV). The PA, PB1, PB2 subunits of the complex were rendered in blue, yellow and green, respectively. The right diagram magnified the active site of the left diagram. The RNA product and template were in red and cyan. The incoming UTP was in orange. The two magnesium ions were in magenta. Two positively charged amino acids, Arg 239 and Lys 235 of PB1 subunit, were pointing to the phosphate tail of the UTP. (**B**) The structural alignment on the active site of influenza A (IAV), B (IBV), C (ICV) and D (IDV) viral RdRps. The RdRp structures of IBV (PDB: 5MSG), ICV (PDB: 5D98), IDV (PDB: 6KV5) were aligned to the bat IAV RdRp elongation complex (PDB: 6SZV). The K235/K237 and R239/R241 on Motif F of IAV, IBV, ICV, IDV were represented in orange, purple, green and blue, respectively. The white background shows the incoming CTP pairing with the template. (**C**) A structure-based alignment on Motif F of various RNA virus RdRps. An arginine residue (highlighted in blue) in the submotif F3 and a lysine residue (highlighted in green) in the submotif F2 are structurally conserved among different types of RNA viruses referring to IAV (highlighted in yellow). The aligned residues in the same column represent their stereo locations are <4 Å. The alignment includes four negative-stranded RNA viruses: Orthomyxoviridae/influenza virus A, B, C, D (IAV, IBV, ICV, IDV); three positive-stranded RNA viruses: Picornaviridae/poliovirus (PICO/PV), Caliciviridae/Norwalk virus (CALI/NV), Leviviridae/bacteriophage Qβ (LEVI/Qβ); one retrovirus (RT): Retroviridae/human immunodeficiency virus-1 (RETR/HIV-1); one double-stranded RNA viruses: Cystoviridae/bacteriophage Φ6 (CYST/Φ6). Dash (-) represents the missing residue in the alignment.

exploited for vaccine development, as a high viral replication rate must be maintained for manufacturing attenuated vaccines. Hence, it is of great significance to determine if K235 and R239 in IAV RdRp can play a role in controlling the fidelity of transcription by fine-tuning the incorporation rates of the cognate and noncognate NTP.

In the present work, we adopted a multi-level approach to provide significant insights into the transcription mechanism by assessing the roles of two critical residues, the highly conserved Arg and Lys in the IAV RdRp from the *Orthomyxoviridae* family of viruses. We showed by cell-based mini-genome, *in vitro* transcription assays and enzyme kinetics assays that R239 and K235 are critical for maintaining transcriptional activity, and they exhibit a very low tolerance for mutations. Surprisingly, we discovered for the first time that a polymerase with the K235R mutation could exert both increased efficiency for cognate NTP incorporation and enhanced fidelity of transcription. In particular, we performed *in vitro* single-turnover NTP incorporation assay to profile the enzyme kinetics properties in cognate and noncognate base pairings under an assumption that there are no unexpected mismatch products extended in the reaction. Intriguingly, we found that the higher fidelity of the K235R mutant is mainly contributed by the significantly

higher activity for the cognate NTP but reduced activity for the noncognate NTP. K235R mutant drastically increased the incorporation rate (k_{pol}) while only slightly altered the dissociation constant $K_{d,app}$ for the cognate NTP. We also conducted all-atom molecular dynamics (MD) simulations to provide an atomic explanation for the higher incorporation rate of K235R mutant RdRp being due to a shorter distance to the triphosphate moiety. Our finding provides novel insights into the unique function of the two positively charged residues in the active site of the IAV polymerase. The K235R mutant RdRp serves as a solid foundation for the rational design of live attenuated vaccines by manipulating transcriptional fidelity via structure-based strategies. To design a rational live attenuated vaccine, it is necessary for the mutant RdRp to have comparable or higher replication activity while exhibiting variable fidelity (22). On the one hand, mutant viruses that retain or even enhance replication activity are ideal candidates for vaccine development due to their ability to replicate efficiently inside host cells. On the other hand, mutational variations in fidelity are useful for the attenuated vaccine as they always disrupt the balance between accurate transcription and efficient adaptation that wild-type viruses originally maintained. Both high- and low-fidelity mutations are being explored as

potential vaccine strategies. However, RNA virus strains that contained high-fidelity polymerase are attenuated *in vivo* via the inability to adapt instead of via another mechanism of generating lethal mutants as in the case for low-fidelity variants (15). The biological outcomes in the virus population are more stable and predictable. Thus, mutant viruses with comparable or high activity, while also with high fidelity would make excellent vaccine candidates.

MATERIALS AND METHODS

Cell culture

Madin-Darby canine kidney (MDCK) cells were maintained in Dulbecco's modified Eagle's medium (DMEM), supplemented with high glucose, GlutaMAX™, pyruvate (Gibco 10569010), 10% Fetal Bovine Serum (FBS, Gibco 10270106) and 1% penicillin-streptomycin (P/S, Gibco 15140122). HEK-293T cells were maintained in Advanced DMEM/F-12 (Gibco 12634010) supplemented with 2% FBS and 0.2% P/S and 4 mM L-Glutamine (Gibco 25030081). Cells were trypsinized and subcultured by TrypLE™ Express Enzyme (Gibco 12605028) once they reached full confluence. All cultures were maintained at 37°C and 5% CO_2. The virus infection medium was formulated with DMEM medium, supplemented with high glucose, GlutaMAX™, pyruvate, 1% P/S and 1 μg/ml TPCK-trypsin without FBS.

Plasmids

The pcDNA3-WSN-PB2-TAP plasmid was a generous gift from Prof. Aartjan te Velthuis from the University of Cambridge. K235 and R239 mutations were introduced by site-direct mutagenesis kit (NEB E0554S) according to the protocol. Briefly, the wild-type pcDNA3-WSN-PB1 and pPolI-WSN-PB1 plasmids were kindly provided by George Brownlee from University of Oxford amplified by primers containing mutant bases. The PCR products were ligated and transformed into DH5α competent cells (NEB C2987) followed by miniprep and Sanger sequencing. The mutation plasmids were further checked for correct expression in HEK-293T cells by western blot against PB1 antibody (Thermofisher PA5-34914).

Virus

Recombinant influenza viruses were rescued using a previously described reverse genetics system (25). The wild-type and mutant recombinant influenza A/WSN/1933(H1N1) viruses were rescued by co-transfecting the eight pHW2000-WSN plasmids into a 1:1 MDCK and HEK293T cell mix. For one well of the 6-well dish, 80% of confluence cells in the virus infection medium were transfected with 250 ng of each pPolI-WSN plasmids mixed with 6 μl TransIT-LT1(Mirus MIR 2300) in 200 μl Opti-MEM (Gibco 51985034). The culture media were collected and spun down after 48 h. The clear supernatants were flash-frozen in liquid nitrogen and stored at -80°C.

TCID$_{50}$ assay

Confluent MDCK cells seeded in 96-well plates were washed twice with PBS and infected with H1N1 WSN wild-type and the K235R mutant viruses. The viruses were 5-fold serial diluted from 10^{-2} in the virus infection medium, and 100 μl of the diluted viruses were added to each well of the plate. After 3 days of incubation, the cells were washed with PBS and fixed with 70% ethanol for 30 min and stained with 2% crystal violet solution (BBI A600331-0025) for 10 min. The detection limit is $10^{2.048}$ TCID$_{50}$ per ml.

Virus growth kinetics

The growth kinetics curves were determined for H1N1 WSN wild-type and the K235R mutant viruses. MDCK cells seeded in 24-well plates were infected with viruses at the MOI of 0.01 supplemented with 1 μg/ml TPCK-trypsin in DMEM medium without FBS. The culture media were collected and spun down after 3, 6, 9, 24, 48 and 72 h of infection. The virus titer was then determined by TCID$_{50}$ assay.

Plaque assay

Confluent MDCK cells seeded in six-well plates were washed twice with PBS and infected with H1N1 WSN wild-type and the K235R mutant viruses. The viruses were serial diluted in the virus infection medium, and 1000 μl of the diluted viruses were added to each well of the plate. After 1 h of inoculation, the virus infection medium was removed, and the cells were washed once with PBS and covered with agar overlay in the virus infection medium at 37°C and 5% CO_2. After 3 days, the agar overlay was removed, and the cells were fixed with 70% ethanol for 30 min and stained with 2% crystal violet solution (BBI A600331-0025) for 10 min. The stained plates were imaged and the plaques were counted.

Cell-based mini-genome luciferase assay

The pcDNA3-WSN-PB1 (wild-type 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. A total of 4×10^4 cells in 100 μl culture medium were seeded in each well of a 96-well plate and transfected for 48 h. For the negative control, pcDNA-PB1 was omitted. After 48 h of incubation at 37°C, the cells were lysed and the signal from two luciferases were measured according to the Dual Luciferase Reporter Gene Assay Kit (Beyotime RG089) protocol. The Gaussian luciferase intensity was normalized to the intensity of polymerase without the PB1 gene (Supplementary Figure S3). See Supplementary Section S1.1 for details.

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 (1,26). 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 (wild-type 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 h 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 h 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 min, 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 h. 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 min, 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. The yield of IAV polymerases were quantified by BSA standard from known concentration in SDS-PAGE (Supplementary Figure S4A). The mass of wild-type and mutant RdRp complex is estimated between 5 and 7 ng/μl. The western blot against PB2, PB1 and PA was also performed to ensure that the mutations introduced in PB1 would not affect the IAV polymerase complex formation (Supplementary Figure S4B). See Supplementary Section S1.2 for details.

Capping and radioactive labeling of RNA primer

Synthetic 5′ diphosphate RNA of 11 nt (5′-ppGAAUAC UCAAG-3′) (Chemgenes) was 5′ capped with a radiolabeled cap-0 and a cap-1 structure to initiate the *in vitro* transcription as described previously (26). Briefly, a 20 μl reaction was setup 1 h at 37°C with 1 μM 11-nt pp-RNA primer, 0.25 μM [α-32P]GTP (3000 Ci mmol⁻¹, Perkin-Elmer BLU006H250UC), 0.8 mM S-adenosylmethionine, 2.5 U/μl 2′-O-methyltransferase (NEB M0366), 0.5 U/μl Vaccinia Capping System (NEB M2080) and 2U/μl RNase Inhibitor (NEB M0314). The Reaction mixture was column purified (NEB T2030) and stored in -20°C. See Supplementary Section S1.3 for details.

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 30 nt 3′ RNA template (5′-GCAUUG UCGCAAUCAGUACCUGCUUUCGCU-3′), 1.6 μM 5′ RNA promoter (5′-AGUAGAAACAAGGUA-3′), 2 U/μl RNase Inhibitor (NEB M0314), 50 c.p.s 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 h, 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 min and chilled on ice for 5 min. The extended radioactive capped RNA products were resolved in 20% urea-PAGE. The gel image was scanned and analyzed on Image Lab (Biorad). See Supplementary Figure S5 and Supplementary Section S1.4 for details.

In vitro single-turnover NTP incorporation and data fitting

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 (27). 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′-GGCCUGCUUUUG CU-3′), 1.6 μM 5′ RNA promoter, 2 U/ μl RNase Inhibitor (NEB M0314), 50 c.p.s of radioactive capped RNA and 30 ng purified influenza A virus polymerase. After pre-incubating the reaction mixtures at 30°C for 30 min, various amounts of CTP or UTP were added to the mixture to achieve the final concentration of 10, 50, 100, 200, 500 and 1000 μM. For correct pairing of CTP, the reactions were started for 0.5, 2, 5, 10, 30, 60 s, 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°C for 5 min and chilled on ice for 5 min. For wobble pairing of UTP, all settings are identical except the reaction was extended to 5, 15, 30, 60, 120, 300 and 600 s. The incorporation products were separated in 20% urea-PAGE. The gel image was scanned and analyzed for the nonlinear regression fitting. In short, the quantified time and concentration-dependent product formation for different NTP substrates are fitted into a biphasic exponential equation which consequently yields two phases correspond to the fast (k_{fast}) and slow (k_{slow}) observed rates for the incorporation. The substrate concentration dependence of the fast phase is then fitted into the Michaelis–Menten equation to obtain the k_{pol} and $K_{d,app}$ for each NTP substrate. The specificity constant was determined by the ratio of k_{pol} and $K_{d,app}$ and the discrimination score is calculated by the different specificities from CTP and UTP in the equation. See Supplementary Figures S6–S9, and Supplementary Sections S1.5–1.8 for details.

Structural-based alignment on the motif F

The structures of motif F together with the incoming NTP of the influenza A virus RdRp [PDB: 6SZV (9)], the Norwalk virus RdRp [PDB: 3BSO (28)], the bacteriophage Φ6 RdRp [PDB: 1HI0 (29)], the human immunodeficiency virus-1 RT [PDB: 1RTD (30)], the poliovirus RdRp [PDB: 3OLA (7)] and the bacteriophage Qβ RdRp [PDB: 3AVT (31)] were extracted and aligned in mTM-align online server (32,33). The aligned residues within 4 Å were plotted in the same column. The lysine and arginine residues in other RNA viruses equivalent to the K235 and R239 in IAV were annotated in the Pymol (34) together with the NTP substrates.

Sequence alignment

The PB1 subunit sequence of IAV (accession: AGI54945.1), IBV (accession: QIQ42332.1), ICV (accession: BAV18492.1), IDV (accession: YP_009449556.1) and Salmon isavirus (accession: YP_145804.1) were obtained from NCBI (35,36). The sequences were aligned in MUSCLE (37).

MD simulations and structural analysis of IAV RdRp

We constructed the model of IAV RdRp with the template-product duplex and NTP bound at the active site based on the cryo-EM structure of the IAV RdRp elongation complex (PDB: 6SZV (9)). In our model, the active site adopts the 'closed' configuration with MgA ion coordinated with D308, D445 and D446. When constructing the model with UTP occupying the active site, we determined the orientation of the UTP by comparing it to the template guanidine nucleotide in the crystal structure of an RNA duplex containing a U:G wobble pair (PDBID: 433D (38)). For the K235 and R239 mutants, the corresponding residue was mutated using Pymol. To equilibrate each system for the MD simulations, 10 000-steps energy minimization was first performed, followed by 200 ps position restraint NVT (T = 303 K) simulation and another 500 ps position restraint NPT (P = 1 bar, T = 303 K) simulation with a force constant of 10 kJ × mol^{-1} × Å^{-2} on all the heavy atoms of the complex. Afterward, five parallel 100 ns production simulations under NVT (T = 303 K) ensemble were performed with different random seeds, where the first 2 ns was used to increase the temperature from 50 to 303 K gradually. After removing the first 20 ns of each trajectory, the MD conformations from the remaining part of the trajectory were used for the subsequent distance analysis. In the simulations, we used Amber99sb-ildn force field (39) to simulate protein and nucleotides. Parameters of the triphosphate moiety of NTP substrate and the magnesium ions were taken from previous works (40–42). See Supplementary Sections S1.9 and 1.10 for details.

NGS sequencing of wild-type 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 and the viruses were amplified to passage 3. Media containing P3 wild-type and K235R mutant viruses were collected and concentrated. The residual genomic DNA and RNA impurities were digested using Micrococcal Nuclease (NEB M0247), and the viral RNA was extracted by NucleoZOL (MACHEREY-NAGEL) and purified by the Monarch RNA Cleanup Kit (NEB T2030). The Ct value for 1 μl of purified RNA was determined by Luna One-Step RT-qPCR Kit (NEB E3005) in conjunction with the 3′ consensus RT-PCR primer (MBTuni-12, 5′-ACGCGTGATC AGCAAAAGCAGG) and the 5′ consensus PCR primer (MBTuni-13, 5′-ACGCGTGATCAGTAGAAACAAGG) as described in the published protocol (43). The viral RNA was then reverse-transcribed from 5 μl sample by Proto-Script II Reverse Transcriptase (NEB M0368) using the MBTuni-12 primer and amplified for 22 cycles using Q5 Hot Start High-Fidelity DNA Polymerase (NEB M0493) using the MBTuni-12 and MBTuni-13 primers. 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 (44), which was then fed into the MGI DNBSEQ-G400 platform. The sequencing reads were first trimmed in paired-end mode with cutadapt (45) using default parameter and mapped into reference IAV genome (GenBank Source Sequence Accession: LC333182 to LC333189) using bwa (46). Reads with Indel and soft-clipped were removed. Mapped reads (sam file) were sorted and converted to bam files by samtools (47) and the Mpileup files were generated by samtools with the following criteria: (i) sequencing base with quality scores over 30 and (ii) mapping quality over 40. The single-nucleotide variants (SNVs) were screened and reported by screening the mpileup with the following criteria: (i) Error rate ≥ 1% and (ii) Depth ≥ $\frac{1000}{\text{Error rate(\%)}}$. See Supplementary Section S1.15 for details.

RESULTS AND DISCUSSION

R239 is essential for maintaining the transcription activity of IAV RdRp

Given the primary role of the β-phosphate group for NTP addition, we aim to identify critical residues participating in the NTP incorporation via stabilizing the β-phosphate group (Figure 1A). We first conducted a structure-based alignment comparison by overlaying the X-ray or Cryo-EM structures of the motif F on the polymerases across different families of viruses, including poliovirus from *Picornaviridae* (PV) (7), Japanese encephalitis virus from *Flaviviridae* (JEV) (48), Norwalk virus from *Caliciviridae* (NV) (28), Escherichia virus Qβ from *Leviviridae* (31), Pseudomonas virus Φ6 from *Cystoviridae* (49), human immunodeficiency virus-1 from *Retroviridae* (HIV-1) (30), influenza virus A (9), B (50), C (51) and D (52) (Figure 1B–C and Supplementary Figure S1A). Our 3D structure-based alignment identified an R239 in Motif F of the IAV polymerase that shared a highly conserved position in 3D space across different families of viruses. This residue corresponds to the R174 in poliovirus, R474 in JEV, R182 Norwalk virus, R914

in Qβ, R270 in Φ6, R74 in HIV-1 (Figure 1C and Supplementary Figure S1A), which is consistent with previous observations based on sequence alignment (2,53). This R239 in IAV RdRp is located within interactive distance to the β-phosphate of the NTP substrate in the enclosed cavity of Motif F (Figure 1A). Its terminal guanidinium moiety is probing and aligning with the oxygen atoms of the β-phosphate of incoming NTP, which suggests that R239 may participate in catalysis by interacting with the β-phosphate group (Figures 1A and 2E). As R239 carries positive charges, we expect the charge could be the key to its function. To evaluate this hypothesis, we performed site-directed mutagenesis on R239 to assess the precise function of this residue. As expected, mutations of R239 to any uncharged residues (such as serine, glutamine, glycine, alanine and leucine) lead to complete abolishment of polymerase activity in the luciferase mini-genome assay (Figure 2A and Supplementary Section S1.1 for details). Similar results were also observed from *in vitro* transcription assays using purified wild-type and mutant polymerase complexes to extend an 11-nt capped primer RNA (see Figure 2B,C, Supplementary Figures S4 and S10, and Supplementary Sections S1.2–1.4 for details). These results indicate that the positive charge is essential for maintaining the function of this residue. However, we observed that even though the positive charge was preserved by mutating arginine to lysine, the polymerase activity was not retained, with only 5% *in vitro* activity compared to wild-type polymerase (Figure 2B,C). We also attempted to rescue recombinant viruses carrying the R239K mutation at the RdRp, but they were not viable (Figure 2D). All these results suggest that R239 is an invariant amino acid residue in PB1.

To determine how R239 mutant disrupts the transcription activity of IAV RdRp, we performed MD simulations of wild-type and mutant RdRps (five 100-ns independent simulations for each system, see Materials and Methods section for details). For the wild-type RdRp, we measured the minimum distance between the nitrogen atoms of the guanidinium moiety in R239 and the oxygen atoms attached to the β-phosphate group of incoming NTP (Figure 2E). The distance distribution is evenly centered at ∼4.5 Å (Figure 2F), a distance that allows its participation in the phosphoryl transfer reaction (54). However, when mutating R239 to lysine, the distance significantly fluctuates, leading to only a negligible fraction of conformations that remain catalytically active, while a considerable portion of inactive conformations with a large separation (>6 Å) from the β-phosphate group of NTP (Figure 2F). This is consistent with the experimental observations that the transcription activity of R239K RdRp is decreased orders of magnitudes compared to the wild-type RdRp. Furthermore, MD simulations by substituting R239 with alanine demonstrate that the distance between R239A and NTP is extended to ∼8 Å (Figure 2F), implicating that the point mutation completely abolishes its role in facilitating catalysis. This result also matches with the observed eliminated transcription activity in our experimental assays.

Altogether, both our *in vitro* transcription assays and MD simulations confirm that R239 is essential for maintaining transcription activity. The positive charge and an appropriate size of the side chain ensure the R239 is in close proximity to the β-phosphate group of incoming NTP and could facilitate catalysis. Our finding is consistent with a previous work demonstrating that the alanine substitutions on the positively charged residues in Motif F of IAV RdRp could lead to various degrees of loss in polymerase activity, such as the R239A mutation that led to a complete loss of enzyme activity (55). It is also worth noting that the distance between the β-phosphate group of incoming NTP and the secondary residue K235 of the functional dual residues in the active site is not perturbed upon the mutation on R239 (Supplementary Figure S11), which suggests that the role R239 plays in the reaction cannot be compensated by K235. These two residues may participate in the reaction with each being indispensable. This is the first observation in the viral polymerase of a negative-strand RNA virus.

K235R mutation enhances the transcription activity of IAV RdRp

Sequence alignments that mainly focus on the residues with functional conservation or sequence similarity may not be able to reveal critical residue networks that are structurally conserved (56–58). Based on the 3D structural-based alignment of Motif F in various viral RdRps (Figure 1B, C and Supplementary Figure S1A) complementing the sequence alignment, we identified the presence of functional dual residues that are prevalent in most RNA virus polymerases. The primary residue is conserved to be arginine, while the secondary residue can be arginine or lysine in different families of viruses. This feature is well exemplified by the R174 and K167 in PV, the R182 and K174 in NV, the R474 and K463 in JEV, the R270 and R268 in Φ6, the R914 and K908 in Qβ, and the R72 and K65 in HIV-1 (Figure 1C and Supplementary Figure S1A). The RdRps of influenza A, B, C, D viruses and salmon isavirus in the *Orthomyxoviridae* family also share highly conserved functional dual residues, as in R239 and K235 for IAV (Figure 1B and Supplementary Figure S1B). Hence, K235 provides a valuable opportunity to understand how NTP incorporation rates and fidelity are being fine-tuned and gives insight into the divergent evolution of influenza RdRps.

To reveal how K235 may contribute to NTP incorporation, we mutated its side chains to evaluate its effect on transcription activity. In the mini-genome assay, mutations that remove the positive charge of the side chains (mutating lysine to alanine, leucine, glutamic acid, asparagine or glutamine) resulted in nearly complete loss of polymerase activity (Figure 3A). Strikingly, we discovered that K235R mutation by preserving the positive charge while modifying the size of the side chain results in up to as much as two-fold enhanced polymerase activity over wild-type RdRp. This is of tremendous functional implications because mutations in the highly conserved catalytic region have always led to the severe defect of polymerase activity (13,59,60). To further confirm this finding, we performed a capped-dependent *in vitro* transcription assay to measure the transcription efficiency of wild-type and K235 mutant RdRps. Consistent with the results from the mini-genome assay (Figure 3A), mutations of K235 to amino acids with neutral side-chains abolished RdRp activity, while K235R mutant possessed a 2-fold increase in transcription efficiency
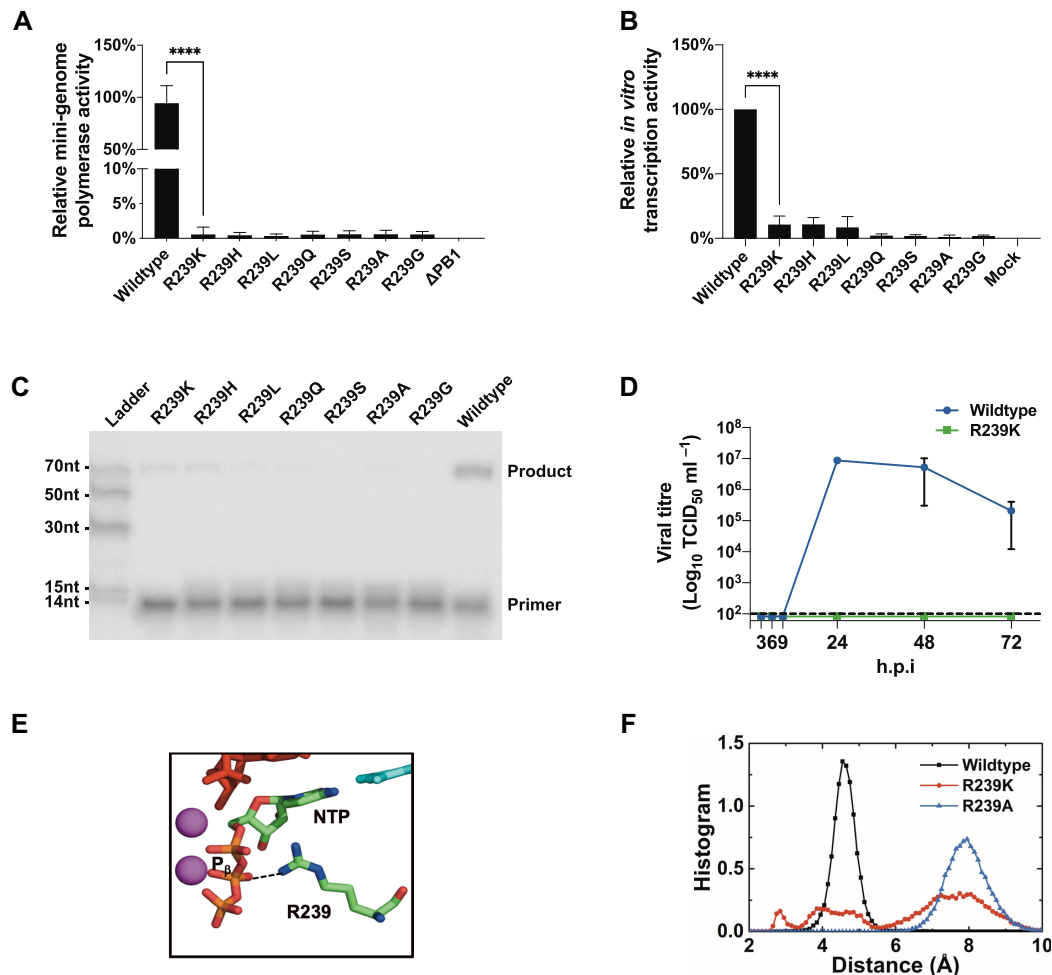
Figure 2. Arg 239 is critical to maintaining the normal function of the IAV polymerase. (**A**) Cell-based dual-luciferase minigenome assay of WT and various R239 mutants. HEK-293T cells were co-transfected with PA, PB2, NP, NS-firefly luciferase, Gaussian luciferase and PB1 plasmids coding one of the WT or R239 mutant PB1 gene. The relative polymerase activity was represented as firefly luciferase intensity/Gaussian luciferase intensity and normalized to WT (100%). Error bars represent standard deviation ($n = 3$) followed by one-way ANOVA with Dunnett's multiple comparison test, ****$P < 0.0001$. (**B**) *in vitro* transcription assay of purified WT and R239 mutant polymerases. The [α-32P] GTP labeled 11nt capped RNA was extended on a 30 nt RNA template. The band intensity from R239 mutants was quantified and normalized to WT. Error bars represent standard deviation ($n = 3$) followed by one-way ANOVA with Dunnett's multiple comparison test, ****$P < 0.0001$. (**C**) The denaturing urea polyacrylamide gel image of the *in vitro* transcription assay of purified WT and R239 mutant polymerases. The reactions were incubated for 1 h at 30°C. Unincorporated primer and full product were indicated in the graph. (**D**) WT and R239K live virus rescue and growth kinetics measurement. WT and R239K recombinant viruses were rescued by co-transfecting cells with the eight plasmids coding each gene of A/WSN/1933 strain and inoculated to MDCK cells at MOI = 0.01. The culture media were collected 3, 6, 9, 24, 48 and 72 h after infection. The TCID$_{50}$ titer concentration was measured and plotted against hours post infection (h.p.i) for WT (blue) and R239K (green). R239K mutant cannot be rescued by recombinant virus from plasmids. The dashed line represents the detection limit of the TCID$_{50}$ assay. (**E**) The diagram showing the Arg 239 in the catalytic center is located near the incoming NTP, and the terminal amide group of the R239 side chain is pointing to the β-phosphate group of the NTP. The black dashed line connecting the terminal nitrogen atom of the residue 239 and the oxygen atom of the β-phosphate group represented the distance to measure in the MD simulations. (**F**) Histogram of the distance measured in (**E**).

to synthesize the full product with all four NTPs (Figure 3B and C). The recombinant viruses carrying the K235R mutation also show no compromise on the viral replication (Figure 3D). These results demonstrate that the K235R mutant can significantly promote transcription activity and maintain normal viral growth kinetics compared to wild-type.

K235R mutation renders IAV RdRp higher fidelity

Transcription rate and fidelity are highly interrelated, both of which are sensitive to mutations of the RdRp. For example, the previously identified high-fidelity V43I and K229R mutants show accurate incorporation at the cost of de-

creased transcription rate (21,23,24). Therefore, it is of surprise that a mutation of K235, a critical residue in the active site, can significantly enhance polymerase activity. Due to the interplay between the transcription activity and fidelity, it would be interesting to determine if K235 could also play a role in controlling transcription fidelity. However, existing studies did not explicitly measure the time and concentration-dependent incorporation rate of cognate and noncognate NTP separately, while the conclusion of high- or low-fidelity was solely drawn by the incorporation efficiency of the miss-pairing NTP or mutagenic NTP analogs. Hence, we developed a kinetics assay to address this challenge by directly quantifying the transcription fidelity.
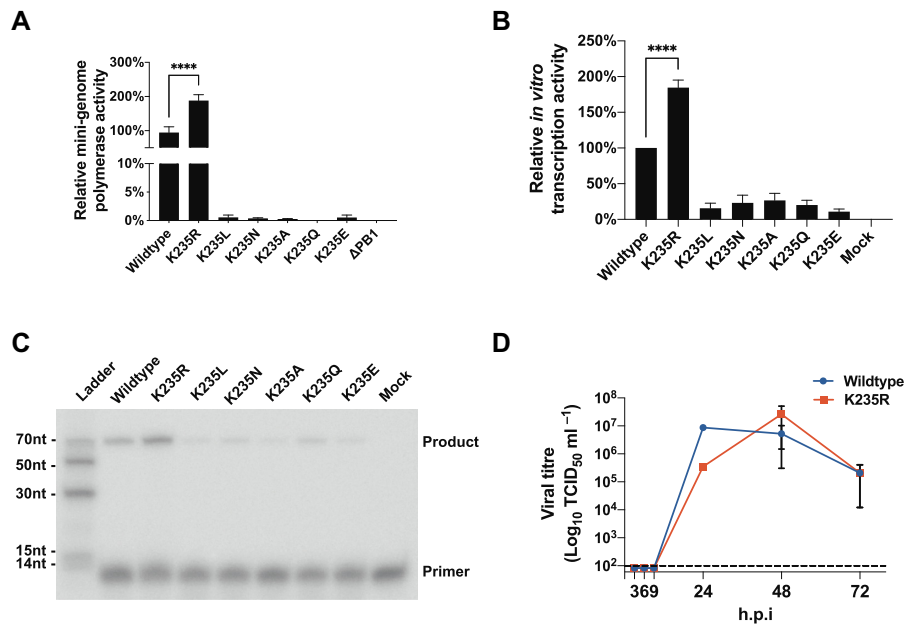
Figure 3. Lys 235 Arg mutation increases the IAV polymerase activity. (**A**) Cell-based dual-luciferase minigenome assay of WT and various K235 mutants. HEK-293T cells were co-transfected with PA, PB2, NP, NS-firefly luciferase, Gaussian luciferase and PB1 plasmids coding one of the WT or K235 mutant PB1 gene. The relative polymerase activity was represented as firefly luciferase intensity/Gaussian luciferase intensity and normalized to WT (100%). Error bars represent standard deviation ($n = 3$) followed by one-way ANOVA with Dunnett's multiple comparison test, ****$P < 0.0001$. (**B**) *in vitro* transcription assay of purified WT and K235 mutant polymerases. The [α-32P] GTP labeled 11 nt capped RNA was extended on a 30 nt RNA. The band intensity from K235 mutants was quantified and normalized to WT. Error bars represent standard deviation ($n = 3$) followed by one-way ANOVA with Dunnett's multiple comparison test, ****$P < 0.0001$. (**C**) The denaturing urea polyacrylamide gel image of the *in vitro* transcription assay of purified WT and K235 mutant polymerases. The reactions were incubated for 1 h at 30°C. Unincorporated primer and full product were indicated in the graph. (**D**) WT and K235R live virus rescue and growth kinetics measurement. WT and K235R recombinant viruses were rescued by co-transfecting cells with the eight plasmids coding each gene of A/WSN/1933 strain and inoculated to MDCK cells at MOI = 0.01. The culture media were collected 3, 6, 9, 24, 48 and 72 h after infection. The TCID$_{50}$ titer concentration was measured and plotted against hours post infection (h.p.i) for WT (blue) and K235R (red). The K235R mutant exhibited similar growth kinetics to WT. The dashed line represents the detection limit of the TCID$_{50}$ assay.

We systematically performed single-turnover NTP incorporation kinetics assays to quantify the incorporation rate constant (k_{pol}) and the apparent substrate dissociation constant ($K_{d,app}$) for site-specific incoming cognate and noncognate NTP. k_{pol} is contributed by the catalysis and associated isomerization of the RdRp active site, while $K_{d,app}$ is a result of NTP loading and binding to the active site (see Supplementary Sections S1.5–1.8 for details). Furthermore, we calculated the substrate specificity ($k_{pol}/K_{d,app}$) to measure the polymerase transcription efficiency and discrimination factor (substrate specificity of cognate over noncognate NTP) to profile the fidelity in both wild-type and K235R mutant RdRps. Although similar kinetics characterization studies have been performed on eukaryotic RNA polymerases (27), to our knowledge, this is the first time that a quantitative measurement of fidelity was carried out for the IAV heterotrimer polymerase complex. Here, we provided quantitative fidelity measurements by comparing the substrate specificity among different types of base pairing. In particular, we explored two different types of incoming NTP:template base pairing (Supplementary Figure S10B), including CTP:G(template) for correct base pairing and UTP:G(template) for wobble base pairing. We chose the wobble base pair because the transition mutations from C to U and G to A are the most frequent types of mutations according to reported IAV mutation bias in the influenza genome sequencing analysis (61,62). Intriguingly, our re-

sults show that in the determination of CTP:G(template), the CTP incorporation rate in K235R RdRp ($k_{pol} = 620.0$ min^{-1}) significantly increased by 2.8-fold compared to the wild-type RdRp ($k_{pol} = 219.3$ min^{-1}) (Figure 4A; Supplementary Figure S6 and Supplementary Table S1). However, the $K_{d,app}$ for CTP incorporation in K235R is 186.0 ± 77.4 μM, comparable with the 126.4 ± 56.0 μM in wild-type. Therefore, the specificity of CTP:G(template) of K235R and wild-type are 3.3 and 1.7 min$^{-1}\mu$M^{-1}, respectively (Figure 4C). These results suggest that the K235R mutant exerts increased transcription efficiency for the cognate CTP by 1.9-fold.

We next determined the specificity for UTP incorporation in both the wild-type and K235R mutant RdRps. The incorporation rate for K235R mutant ($k_{pol} = 7.8$ min^{-1}) is slightly higher than that for wild-type ($k_{pol} = 6.8$ min^{-1}). However, $K_{d,app}$ for wobble base pair in K235R (448.8 ± 88.2 μM) is obviously higher than that in wild-type RdRp (163.2 ± 38.7 μM) (Figure 4B and Supplementary Figure S6). Therefore, the transcription efficiency for the noncognate UTP incorporation in K235R mutant ($k_{pol}/K_{d,app} = 0.02$ min$^{-1}\mu$M^{-1}) is lower than that in the wildtype ($k_{pol}/K_{d,app} = 0.04$ min$^{-1}\mu$M^{-1}) (Figure 4C and Supplementary Table S1). Based on the specificity for UTP, we computed the discrimination factor for both wild-type and K235R, which serves as a quantitative measure-
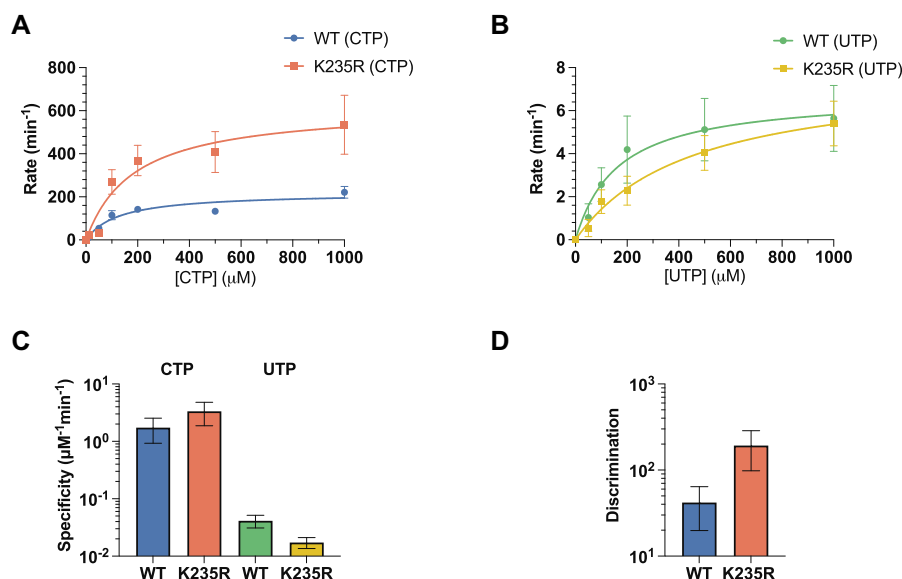
Figure 4. The single-turnover NTP incorporation assay of wild-type and K235R mutant RdRps on CTP:G(template) and UTP:G(template) pairings. (**A**) The observed incorporation rates obtained from the fast phase against various CTP, and (**B**) UTP concentrations were fitted into hyperbolic equations to derive the k_{pol} and $K_{d,app}$ values for wild-type (WT) and K235R mutant RdRps. (**C**) The specificity for WT and K235R RdRps in CTP and UTP incorporation was determined by $k_{cat}/K_m = k_{pol}/K_{d,app}$. The *y*-axis is on a log-10 scale. (**D**) The discrimination of CTP:G(template) over UTP:G(template) for WT and K235R RdRps. Discrimination was determined by specificity$_{CTP:G}$/specificity$_{UTP:G}$. Error bars represent deviations in nonlinear regression analysis.

ment for the transcription fidelity by computing the ratio of specificities for cognate CTP:G(template) and wobble UTP:G(template). As the K235R RdRp has higher efficiency to incorporate the cognate CTP but a lower efficiency to incorporate the noncognate UTP than the wild-type, the transcription fidelity for K235R RdRp is around 4.6-fold higher than that for wild-type RdRp (Figure 4D).

To further explore the molecular mechanisms by which K235R enhanced the incorporation rate (k_{pol}) in comparison with wild-type (Figure 5A), we performed MD simulations to survey the active site of wild-type and K235R mutant RdRps (five 100-ns simulations for each system, see Supplementary Sections 1.9–1.10 for details). In the wild-type RdRp with the cognate CTP, the minimum distance between the β-phosphate oxygen atoms of CTP and the nitrogen atom of the amine group of K235 is centered around 8 Å (Figure 5C), too far to allow K235 to be involved in phosphoryl transfer reaction for NTP incorporation. However, the arginine substitution of K235 significantly decreases its distance to the β-phosphate group (∼5 Å) (Figure 5C) and facilitates its participation in the NTP incorporation. Furthermore, the K235R mutation well maintains other critical distances for catalysis (Supplementary Figure S12, see Supplementary Section S1.11 for details), and simultaneously maintained or imposed extra interactions with the P$_\alpha$ & P$_\gamma$ groups of CTP (Supplementary Figure S13, see Supplementary Sections 1.10 and 1.11 for details). Overall, the K235R mutant not only preserves distances that are critical to catalysis but also forms extra interactions between K235R and CTP to further stabilize the catalytically active conformation. This provides the mechanistic interpretation for the higher k_{pol} in K235R (Figure 5A). When UTP is present at the active site to form the wobble pair with the template

G, the K235R mutant also shows a closer distance to the β-phosphate group of UTP than wild-type (Figure 5D). However, this distance fluctuates more obviously and has a fractional distribution in the region of 6–7 Å, consistent with the experimental observation that the K235R mutant has a less evident increase in k_{pol} for UTP:G(template) wobble pair than the CTP:G(template) canonical pair (Figure 5A).

Overall, the distance analysis from MD simulations explains the closer distance between K235R and substrate NTP can lead to a faster incorporation rate (k_{pol}). It is noted that the sufficient space in the active site of IAV RdRp is important to ensure the K235R mutant with a longer side chain while maintaining the positive charge could work together with R239 to approach the triphosphate moiety of CTP without imposing any steric hindrance, thereby exerting a higher propensity for withdrawing electron density from the phosphate and facilitating the phosphodiester bond formation. This also explains the discrepancy in the transcription activity of K235R mutant IAV RdRp and other RdRp with the corresponding mutation. For example, a previous study identified a mutant HIV-1 with a K65R substitution on the p66 subunit of reverse transcriptase (RT). This mutant showed increased fidelity and resistance to nucleoside reverse transcriptase inhibitors (NRTIs) but resulted in a significant reduction in retroviral transcription activity (63). The K65 in HIV-1 RT is structurally equivalent to K235 in IAV RdRp (Supplementary Figures S1A, S14A and S14C), where the disturbed enzyme activity of K65R was probably caused by a bulkier side chain of the substituted arginine in the highly compacted active site (Supplementary Figure S14A and S14B, see Supplementary Section S1.12 for details). This is in sharp contrast to IAV RdRp, wherein the active site has enough space to accom-
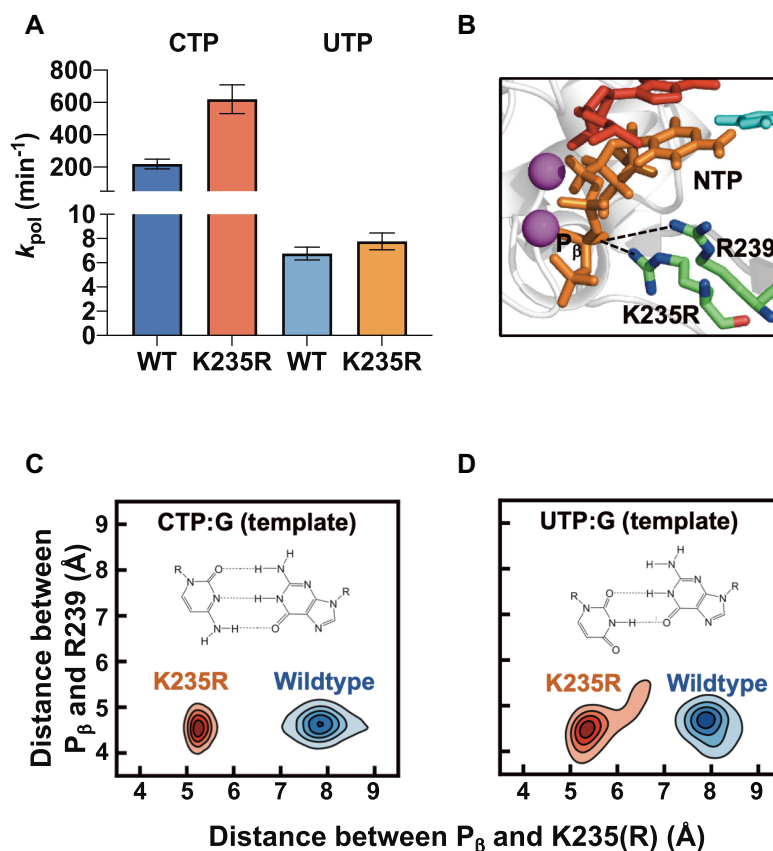
Figure 5. K235R mutation accelerates the incorporation of the cognate CTP. (**A**) Polymerization rates (k_{pol}) for cognate CTP and noncognate CTP in wild-type and K235R mutant RdRps. Error bars represent deviations in nonlinear regression analysis. (**B**) Diagram showing the R239 and K235R mutant residues interact (shown in dashed lines) with the β-phosphate group of the NTP. (**C**) The contour plot of the distance between the P_β atom of NTP and K235/K235R versus that between the β-phosphate (P_β) atom of NTP and R239 when CTP is in the active site. (**D**) Similar to (C) but for noncognate UTP in the active site. In (C and D), the hydrogen bonding pattern for the CTP:G pair and UTP:G wobble pair were depicted as the inset. See Supplementary Section S1.10 for details about the distance calculations.

modate K235R and R239 without steric hindrance (Supplementary Figure S14D).

We noticed that the specificity and fidelity are also contributed by $K_{d,app}$ in addition to k_{pol}. The experimentally measured $K_{d,app}$ is mainly contributed by two terms: NTP loading and NTP binding to the active site. For NTP loading, it is challenging to directly quantify its contribution to $K_{d,app}$. For NTP binding, we have computed the free energy of NTP binding to the active site in the revision of this manuscript and further examined the impact of K235R on the cognate versus noncognate NTP binding (Supplementary Figure S15, see Supplementary Section S1.13 for details). Our results clearly indicate the less favourable binding free energy of UTP compared to CTP (the difference is ~3.4 kJ/mol, Supplementary Figure S15B). We note that this difference is still at the level of thermal fluctuation (~1.3 $k_B T$), nevertheless it is converged in our alchemical free energy calculations (Supplementary Figure S15C). Further structural analysis indicated the conformational flexibility is most pronounced in the K235R mutant RdRp with noncognate UTP (Supplementary Figures S16 and S17, see Supplementary Section S1.14 for details), possibly due to larger fluctuations of the phosphate tail induced by the relatively unstable wobble base pair. This observed conforma-

tional flexibility together with the results of the free energy calculations explains how the K235R mutation deters the UTP binding (i.e. with a larger $K_{d,app}$, Supplementary Table S1).

We further assessed if K235R could exhibit better discrimination for NTP analogs. The reaction was initiated by pairing the guanosine template by either 3′-deoxycytidine-5′-triphosphate (3′dCTP) or 2′-fluoro-2′-deoxycytidine-5′-triphosphate (2′F-dCTP). As expected, K235R mutant RdRp incorporates both the 3′dCTP and 2′F-dCTP at much lower efficiency compared to wild-type RdRp (Supplementary Figure S18). These results further support our discovery that the K235R mutant has higher transcription fidelity than the wild-type to distinguish the cognate NTP from the mismatched NTP and the NTP analogs.

Altogether, the single-turnover NTP incorporation kinetics assays allow us to quantify the transcription fidelity of K235R mutant against the wild-type RdRp from the derived incorporation constant (k_{pol}) and apparent substrate dissociation constant ($K_{d,app}$) for the cognate CTP and the noncognate UTP. The longer side chain imposed by the K235R mutant renders the residue closer to the triphosphate moiety of CTP and thereby considerably facilitates the incorporation of the cognate CTP (faster k_{pol}), leading
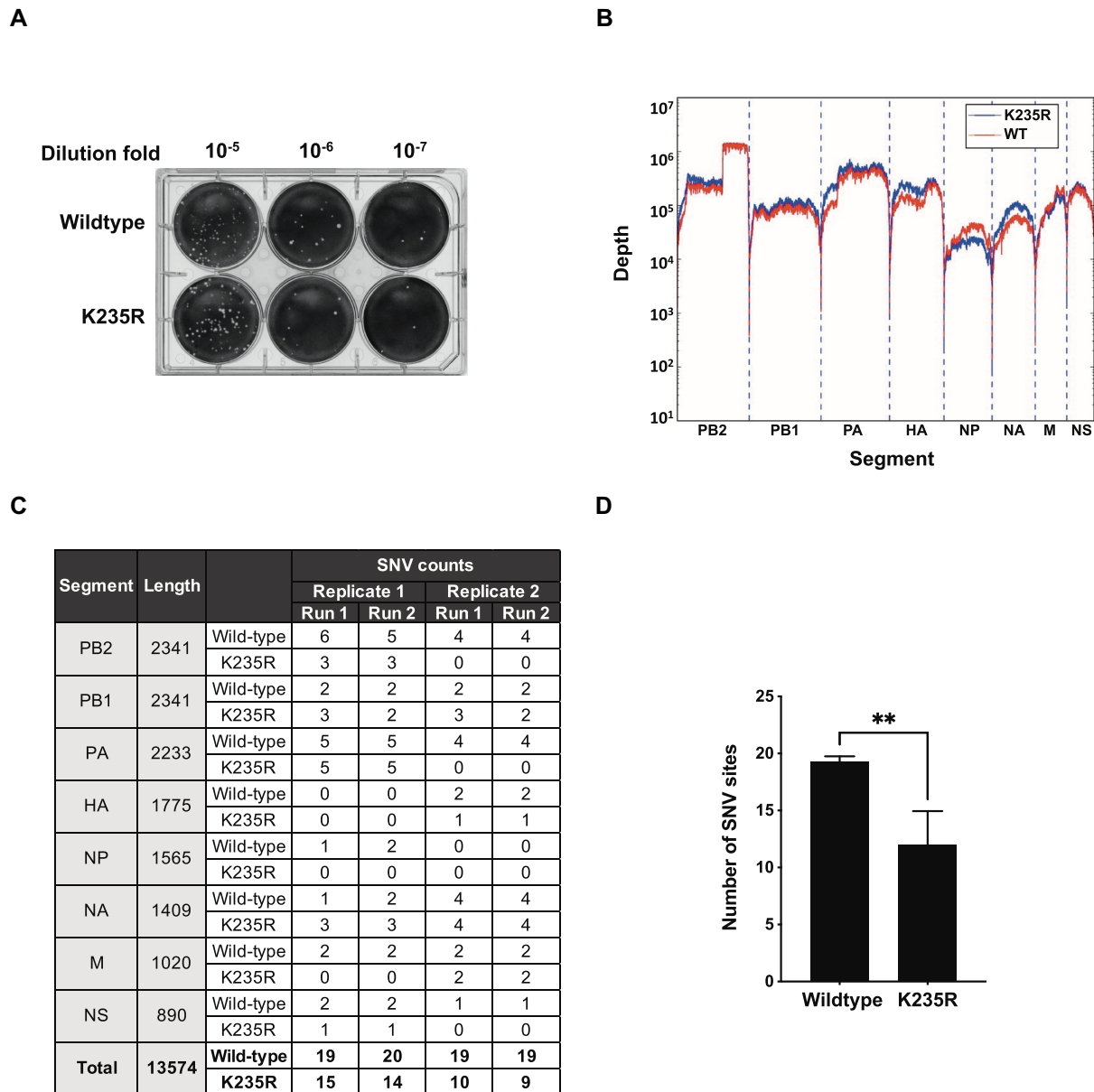
A

B

C

Segment	Length		SNV counts			
			Replicate 1		Replicate 2	
			Run 1	Run 2	Run 1	Run 2
PB2	2341	Wild-type	6	5	4	4
		K235R	3	3	0	0
PB1	2341	Wild-type	2	2	2	2
		K235R	3	2	3	2
PA	2233	Wild-type	5	5	4	4
		K235R	5	5	0	0
HA	1775	Wild-type	0	0	2	2
		K235R	0	0	1	1
NP	1565	Wild-type	1	2	0	0
		K235R	0	0	0	0
NA	1409	Wild-type	1	2	4	4
		K235R	3	3	4	4
M	1020	Wild-type	2	2	2	2
		K235R	0	0	2	2
NS	890	Wild-type	2	2	1	1
		K235R	1	1	0	0
Total	13574	Wild-type	19	20	19	19
		K235R	15	14	10	9

D

Figure 6. NGS sequencing on wild-type and K235R mutant virus genome. (**A**) Plaque assay of P3 wild-type and K235R recombinant IAV. Full confluent MDCK cells were inoculated by diluted P3 wild-type and K235R recombinant viruses with agar overlay for 3 days. The plaques were visualized by crystal violet stain. (**B**) The genome-wide depth distribution of sequencing reads from P3 wild-type and K235R viral genome. The average sequencing depths of both wild-type and K235R are $\sim 10^5$ per replicate. (**C**) The Supplementary Table Summary of SNV counts screened from each gene segment of the extracted P3 wild-type and K235R mutant IAV genome RNA. (**D**) The number of SNV was screened and reported in the bar chart. Error bars are ± SD, $n = 4$, paired two-tailed t-tests, **$P < 0.01$.

to higher transcription efficiency than the wild-type. Meanwhile, the mutant shows lower transcription efficiency to incorporate the noncognate UTP mainly due to the higher $K_{d,app}$ for the K235R mutant. To our knowledge, this is the first time that the incorporation rate of cognate and noncognate NTP for IAV RdRp is systematically assessed. More importantly, our findings suggest that the higher fidelity of K235R mutant is originated from higher transcription efficiency for incorporating cognate NTP over noncognate NTP. This discovery is distinct from previous studies where a more accurate NTP incorporation is always accompanied by a slower transcription activity (64–67). However,

it should be noted that IAV RdRp is reported to introduce mismatches in the *in vitro* transcription with depleted NTP pool (9,68). Hence, in our single-turnover NTP incorporation assay, other than the primary extension products initiated from the third base (G), the mismatched products extended from the second base (C) may also contribute to the single-cycle extension signal for CTP or UTP incorporation, which could probably interfere with the measurement.

Our discovery of the K235R mutant IAV may have significant implications to vaccine development by overcoming two main challenges. First, most high-fidelity mutant RNA viruses were identified through blind passages of the virus

in the presence of antiviral drugs. Second, to harness fidelity variants as stable attenuated virus strains for the vaccine design, the fidelity variant mutant viruses should possess high fidelity and high activity. However, thus far, the accurate transcription of identified high-fidelity viruses always exhibits reduced polymerase activity in various degrees. This is exemplified not only in the IAV (9) but also in other viruses, such as PV (67), HIV-1 (63), Foot and Mouth Disease virus (66), West Nile virus (64) and Coxsackievirus (65). To address these challenges, we here adopted a mechanistic-insight driven approach to discover a novel variant in IAV RdRp. To achieve this, we employed the single-turnover NTP incorporation assay, which can selectively measure the specificity of cognate versus noncognate NTP tightly to quantify fidelity, and coupled with all-atom MD simulations to explore the molecular explanation on the mechanism of altered incorporation rate. We showed here, as the first exception, that a single Lysine-to-Arginine mutation on the IAV RdRp can enhance both the incorporation rate and fidelity. In the future, we anticipate that this mutation holds great promise for the development of vaccines owing to its high fidelity for fast attenuation in infection while maintaining a high profile in viral replication for mass production.

Mutational frequencies carrying the PB1-K235R mutation measured by NGS

To confirm the high-fidelity phenotype of live viruses with the PB1-K235R mutation, we rescued the wild-type and PB1-K235R mutant A/WSN/1933(H1N1) recombinant viruses from pHW2000 plasmids. The viral RNA was extracted from two biological replicates of intact virus during their third passage (P3) with both titers reached 10^7 plaque forming unit (PFU)/ml (Figure 6A). The entire genome was reverse-transcribed and amplified according to a previously published protocol (43) and sequenced on the MGI DNBSEQ-G400 platform. After removing the adapter, the sequencing reads were mapped to the reference IAV genome. The average sequencing depths of both wild-type and K235R are $\sim 10^5$ per replicate (Figure 6B). With the high coverage of both systems, we were confident to further compute SNVs by following a previously published method (69). As shown in Figure 6C and D, we found the number of SNV sites in wild-type live virus is 19.25 ± 0.50 (mean $\pm$ SD), whereas the number of SNV sites in K235R is 12 ± 2.94 (mean $\pm$ SD). The NGS results indicated that the K235R mutant live virus had a significantly lower mutation rate than that of the wild-type live virus, which are also consistent with our findings based on biochemical assays and confirm that K235R is indeed a high-fidelity mutant at the live virus level.

In this study, we show that the K235R mutant IAV RdRp exhibits increased fidelity and activity compared to the wild-type K235 IAV RdRp. We expect that the K235R substitution will not become prevalent because the virus's high fidelity RdRp may prevent it from initiating the required mutational changes, rendering it incapable of adapting to its environment. As a result, the IAV genome will not be dominated by the high-fidelity mutant K235R. On the contrary, the wild-type K235 is preserved because it exhibits high polymerase activity and optimum fidelity without impair-

ing viral fitness. Indeed, the high-fidelity phenotype K235R is only present in a small number of swine and avian strains (e.g. isolation of A/swine/Guangdong/wxl/2004(H9N2) (70)) but not in human isolations, implying that K235R renders the virus incapable of effective cross-species transmission.

CONCLUSION

In this study, we utilized a multiple-level approach by integrating enzymatic, virological assays, MD simulations to elucidate that the highly conserved R239 and K235 are essential for transcription activity and fidelity in IAV. More intriguingly, we found that K235 is critical to maintaining transcription activity in IAV due to its positively charged side chain. When altering the side chain of K235 while preserving its positive charge, we showed that K235R exhibited a much-enhanced NTP incorporation rate for cognate NTP and a higher transcription fidelity by distinguishing cognate NTP against the mismatched NTP and NTP analogs. These remarkable properties obtained from a single lysine-to-arginine mutation have not been characterized before, which renders this mutant a promising target for the structural-base design of attenuated vaccine development and a model to study the transcription rate and fidelity control of viral polymerases.

DATA AVAILABILITY

The raw NGS sequencing data for both wildtype and K235R mutant IAV live viruses are available in the NCBI sequence read archive (SRA) with accession number: PRJNA737597, Link: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA737597.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Velthuis,te, Long,A.J.W. and Barclay,W.S. (2018) Assays to measure the activity of influenza virus polymerase. In: *Influenza Virus, Methods and Protocols*. Springer, NY, Vol. **1836**, pp. 343–374.
2. Wu,J., Liu,W. and Gong,P. (2015) A structural overview of RNA-dependent RNA polymerases from the flaviviridae family. *IJMS*, **16**, 12943–12957.
3. Bruenn,J.A. (2003) A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res.*, **31**, 1821–1829.
4. Castro,C., Smidansky,E., Maksimchuk,K.R., Arnold,J.J., Korneeva,V.S., Götte,M., Konigsberg,W. and Cameron,C.E. (2007) 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.
5. Castro,C., Smidansky,E.D., Arnold,J.J., Maksimchuk,K.R., Moustafa,I., Uchida,A., Götte,M., Konigsberg,W. and Cameron,C.E. (2009) Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat. Struct. Mol. Biol.*, **16**, 212–218.
6. 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. USA*, **111**, E3335–E3342.
7. Gong,P. and Peersen,O.B. (2010) Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. USA*, **107**, 22505–22510.
8. Shu,B. and Gong,P. (2016) Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation. *Proc. Natl. Acad. Sci. USA*, **113**, E4005–E4014.
9. Wandzik,J.M., Kouba,T., Karuppasamy,M., Pflug,A., Drncová,P., Provaznik,J., Azevedo,N. and Cusack,S. (2020) A structure-based model for the complete transcription cycle of influenza polymerase. *Cell*, **181**, 877–893.
10. Garforth,S., Lwatula,C. and Prasad,V. (2014) The lysine 65 residue in HIV-1 reverse transcriptase function and in nucleoside analog drug resistance. *Viruses*, **6**, 4080–4094.
11. Cheney,I.W., Naim,S., Lai,V.C.H., Dempsey,S., Bellows,D., Walker,M.P., Shim,J.H., Horscroft,N., Hong,Z. and Zhong,W. (2002) Mutations in NS5B polymerase of hepatitis C Virus: Impacts on in vitro enzymatic activity and viral RNA replication in the subgenomic replicon cell culture. *Virology*, **297**, 298–306.
12. Abdelnabi,R., de Morais,A.T.S., Leyssen,P., Imbert,I., Beaucourt,S., Blanc,H., Froeyen,M., Vignuzzi,M., Canard,B., Neyts,J. *et al.* (2017) Understanding the mechanism of the broad-spectrum antiviral activity of favipiravir (T-705): Key role of the F1 motif of the viral polymerase. *J. Virol.*, **91**, e00487-17.
13. Yang,X., Liu,X., Musser,D.M., Moustafa,I.M., Arnold,J.J., Cameron,C.E. and Boehr,D.D. (2017) Triphosphate reorientation of the incoming nucleotide as a fidelity checkpoint in viral RNA-dependent RNA polymerases. *J. Biol. Chem.*, **292**, 3810–3826.
14. Ranjith-Kumar,C.T., Sarisky,R.T., Gutshall,L., Thomson,M. and Kao,C.C. (2004) De novo initiation pocket mutations have multiple effects on hepatitis C virus RNA-dependent RNA polymerase activities. *J. Virol.*, **78**, 12207–12217.
15. te Velthuis,A.J.W. (2014) Common and unique features of viral RNA-dependent polymerases. *Cell. Mol. Life Sci.*, **71**, 4403–4420.
16. Arnold,J.J. and Cameron,C.E. (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg2+. *Biochemistry*, **43**, 5126–5137.
17. Yu,J., Da,L.-T. and Huang,X. (2014) Constructing kinetic models to elucidate structural dynamics of a complete RNA polymerase II elongation cycle. *Phys. Biol.*, **12**, 016004.
18. Wang,M., Li,R., Shu,B., Jing,X., Ye,H.-Q. and Gong,P. (2020) Stringent control of the RNA-dependent RNA polymerase translocation revealed by multiple intermediate structures. *Nat. Commun.*, **11**, 2605.
19. Cheung,P.P.-H., Rogozin,I.B., Choy,K.-T., Ng,H.Y., Peiris,J.S.M. and Yen,H.-L. (2015) Comparative mutational analyses of influenza A viruses. *RNA*, **21**, 36–47.
20. Steel,J. and Lowen,A.C. (2014) Influenza A virus reassortment. *Curr. Top Microbiol. Immunol.*, **385**, 377–401.
21. Naito,T., Mori,K., Ushirogawa,H., Takizawa,N., Nobusawa,E., Odagiri,T., Tashiro,M., Ohniwa,R.L., Nagata,K. and Saito,M. (2017) Generation of a genetically stable high-fidelity influenza vaccine strain. *J. Virol.*, **91**, 1776.
22. Vignuzzi,M., Wendt,E. and Andino,R. (2008) Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.*, **14**, 154–161.
23. Cheung,P.P.H., Watson,S.J., Choy,K.-T., Sia,S.F., Wong,D.D.Y., Poon,L.L.M., Kellam,P., Guan,Y., Malik Peiris,J.S. and Yen,H.-L. (2014) Generation and characterization of influenza A viruses with altered polymerase fidelity. *Nat. Commun.*, **5**, 4794.
24. Goldhill,D.H., Velthuis,te, Fletcher,A.J.W., Langat,R.A., Zambon,P., Lackenby,M. and Barclay,W.S. (2018) The mechanism of resistance to favipiravir in influenza. *Proc. Natl. Acad. Sci. USA*, **115**, 11613–11618.
25. Hoffmann,E., Neumann,G., Kawaoka,Y., Hobom,G. and Webster,R.G. (2000) A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc Natl Acad Sci USA*, **97**, 6108–6113.
26. te Velthuis,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.
27. Kellinger,M.W., Song,C.-X., Chong,J., Lu,X.-Y., He,C. and Wang,D. (2012) 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.*, **19**, 831–833.
28. Zamyatkin,D.F., Parra,F., Alonso,J.M.M., Harki,D.A., Peterson,B.R., Grochulski,P. and Ng,K.K.-S. (2008) Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.*, **283**, 7705–7712.
29. Butcher,S.J., Grimes,J.M., Makeyev,E.V., Bamford,D.H. and Stuart,D.I. (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, **410**, 235–240.
30. Huang,H., Chopra,R., Verdine,G.L. and Harrison,S.C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, **282**, 1669–1675.
31. Takeshita,D. and Tomita,K. (2012) Molecular basis for RNA polymerization by Qβ replicase. *Nat. Struct. Mol. Biol.*, **19**, 229–237.
32. Dong,R., Pan,S., Peng,Z., Zhang,Y. and Yang,J. (2018) mTM-align: a server for fast protein structure database search and multiple protein structure alignment. *Nucleic Acids Res.*, **46**, W380–W386.
33. Dong,R., Peng,Z., Zhang,Y. and Yang,J. (2018) mTM-align: an algorithm for fast and accurate multiple protein structure alignment. *Bioinformatics*, **34**, 1719–1725.
34. Schrödinger PyMOL: The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
35. Brister,J.R., Ako-Adjei,D., Bao,Y. and Blinkova,O. (2015) NCBI viral genomes resource. *Nucleic Acids Res.*, **43**, D571–D577.
36. Bao,Y., Bolotov,P., Dernovoy,D., Kiryutin,B., Zaslavsky,L., Tatusova,T., Ostell,J. and Lipman,D. (2008) The influenza virus resource at the National Center for Biotechnology Information. *J. Virol.*, **82**, 596–601.
37. Edgar,R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.*, **5**, 113–119.
38. Trikha,J., Filman,D.J. and Hogle,J.M. (1999) Crystal structure of a 14 bp RNA duplex with non-symmetrical tandem GxU wobble base pairs. *Nucleic Acids Res.*, **27**, 1728–1739.
39. 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.
40. 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.
41. 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.
42. 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.

43. 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. *J. Virol.*, **83**, 10309–10313.

44. 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.

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

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

47. 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.

48. Surana,P., Satchidanandam,V. and Nair,D.T. (2013) RNA-dependent RNA polymerase of Japanese encephalitis virus binds the initiator nucleotide GTP to form a mechanistically important pre-initiation state. *Nucleic Acids Res.*, **42**, 2758–2773.

49. Butcher,S.J., Grimes,J.M., Makeyev,E.V., Bamford,D.H. and Stuart,D.I. (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, **410**, 235–240.

50. Reich,S., Guilligay,D. and Cusack,S. (2017) An in vitrofluorescence based study of initiation of RNA synthesis by influenza B polymerase. *Nucleic Acids Res.*, **45**, 3353–3368.

51. Hengrung,N., El Omari,K., Serna Martin,I., Vreede,F.T., Cusack,S., Rambo,R.P., Vonrhein,C., Bricogne,G., Stuart,D.I., Grimes,J.M. *et al.* (2015) Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature*, **527**, 114–117.

52. Peng,Q., Liu,Y., Peng,R., Wang,M., Yang,W., Song,H., Chen,Y., Liu,S., Han,M., Zhang,X. *et al.* (2019) Structural insight into RNA synthesis by influenza D polymerase. *Nat. Microbiol.*, **4**, 1750–1759.

53. Jia,H. and Gong,P. (2019) A structure-function diversity survey of the RNA-Dependent RNA polymerases from the positive-strand RNA viruses. *Front. Microbiol.*, **10**, 1945.

54. Mildvan,A.S. (1997) Mechanisms of signaling and related enzymes. *Proteins*, **29**, 401–416.

55. Jung,T.E. and Brownlee,G.G. (2006) A new promoter-binding site in the PB1 subunit of the influenza A virus polymerase. *J. Gen. Virol.*, **87**, 679–688.

56. Leahy,M.B., Dessens,J.T., Weber,F., Kochs,G. and Nuttall,P.A. (1997) The fourth genus in the Orthomyxoviridae: sequence analyses of two Thogoto virus polymerase proteins and comparison with influenza viruses. *Virus Res.*, **50**, 215–224.

57. Černý,J., Černá Bolfíková,B., Valdés,J.J. and Růžek,D. (2014) Evolution of tertiary structure of viral RNA dependent polymerases. *PLoS ONE*, **9**, e96070.

58. Selisko,B., Papageorgiou,N., Ferron,F. and Canard,B. (2018) Structural and functional basis of the fidelity of nucleotide selection by flavivirus RNA-Dependent RNA polymerases. *Viruses*, **10**, 59–24.

59. Arias,A., Arnold,J.J., Sierra,M., Smidansky,E.D., Domingo,E. and Cameron,C.E. (2008) Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. *J. Virol.*, **82**, 12346–12355.

60. Gohara,D.W., Arnold,J.J. and Cameron,C.E. (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): kinetic, thermodynamic, and structural analysis of ribonucleotide selection. *Biochemistry*, **43**, 5149–5158.

61. Goldhill,D.H., Langat,P., Xie,H., Galiano,M., Miah,S., Kellam,P., Zambon,M., Lackenby,A. and Barclay,W.S. (2019) Determining the mutation bias of favipiravir in influenza virus using next-generation sequencing. *J. Virol.*, **93**, 448.

62. Pauly,M.D., Procario,M. and Lauring,A.S. (2017) The mutation rates and mutational bias of influenza A virus. bioRxiv doi: https://doi.org/10.1101/110197, 21 February 2017, preprint: not peer reviewed.

63. Garforth,S.J., Domaoal,R.A., Lwatula,C., Landau,M.J., Meyer,A.J., Anderson,K.S. and Prasad,V.R. (2010) K65R and K65A substitutions in HIV-1 reverse transcriptase enhance polymerase fidelity by decreasing both dNTP misinsertion and mispaired primer extension efficiencies. *J. Mol. Biol.*, **401**, 33–44.

64. Van Slyke,G.A., Arnold,J.J., Lugo,A.J., Griesemer,S.B., Moustafa,I.M., Kramer,L.D., Cameron,C.E. and Ciota,A.T. (2015) Sequence-Specific fidelity alterations associated with west nile virus attenuation in mosquitoes. *PLoS Pathog.*, **11**, e1005009.

65. McDonald,S., Block,A., Beaucourt,S., Moratorio,G., Vignuzzi,M. and Peersen,O.B. (2016) Design of a genetically stable high fidelity coxsackievirus B3 polymerase that attenuates virus growth in vivo. *J. Biol. Chem.*, **291**, 13999–14011.

66. Rai,D.K., Segundo,Diaz-San, Campagnola,F., Keith,G., Schafer,A., Kloc,E.A., A.,de Los Santos, Peersen,T. and Rieder,E. (2017) Attenuation of Foot-and-mouth disease virus by engineered viral polymerase fidelity. *J. Virol.*, **91**, 1.

67. Vignuzzi,M., Stone,J.K. and Andino,R. (2005) Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Res.*, **107**, 173–181.

68. Oymans,J. and te Velthuis,A.J.W. (2018) A mechanism for priming and realignment during influenza A virus replication. *J. Virol.*, **92**, e01773-17.

69. Borges,V., Pinheiro,M., Pechirra,P., Guiomar,R. and Gomes,J.P. (2018) INSaFLU: an automated open web-based bioinformatics suite 'from-reads' for influenza whole-genome-sequencing-based surveillance. *Genome Med.*, **10**, 46.

70. Zhang,Y., Aevermann,B.D., Anderson,T.K., Burke,D.F., Dauphin,G., Gu,Z., He,S., Kumar,S., Larsen,C.N., Lee,A.J. *et al.* (2017) Influenza Research Database: An integrated bioinformatics resource for influenza virus research. *Nucleic Acids Res.*, **45**, D466–D474.