



**HAL**  
open science

# The host RNA polymerase II C-terminal domain is the anchor for replication of the influenza virus genome

Tim Krischuns, Benoît Arragain, Catherine Isel, Sylvain Paisant, Matthias Budt, Thorsten Wolff, Stephen Cusack, Nadia Naffakh

## ► To cite this version:

Tim Krischuns, Benoît Arragain, Catherine Isel, Sylvain Paisant, Matthias Budt, et al.. The host RNA polymerase II C-terminal domain is the anchor for replication of the influenza virus genome. Nature Communications, 2024, 15 (1), pp.1064. 10.1038/s41467-024-45205-2 . hal-04563054

**HAL Id: hal-04563054**

**<https://hal.science/hal-04563054>**

Submitted on 29 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# The host RNA polymerase II C-terminal domain is the anchor for replication of the influenza virus genome

Received: 1 August 2023

Accepted: 16 January 2024

Published online: 05 February 2024

Check for updates

Tim Krischuns <sup>1,4</sup>✉, Benoît Arragain <sup>2,4</sup>, Catherine Isel <sup>1</sup>, Sylvain Paisant<sup>1</sup>, Matthias Budt <sup>3</sup>, Thorsten Wolff <sup>3</sup>, Stephen Cusack <sup>2</sup>✉ & Nadia Naffakh <sup>1</sup>✉

The current model is that the influenza virus polymerase (FluPol) binds either to host RNA polymerase II (RNAP II) or to the acidic nuclear phosphoprotein 32 (ANP32), which drives its conformation and activity towards transcription or replication of the viral genome, respectively. Here, we provide evidence that the FluPol-RNAP II binding interface, beyond its well-acknowledged function in cap-snatching during transcription initiation, has also a pivotal role in replication of the viral genome. Using a combination of cell-based and in vitro approaches, we show that the RNAP II C-terminal-domain, jointly with ANP32, enhances FluPol replication activity. We observe successive conformational changes to switch from a transcriptase to a replicase conformation in the presence of the bound RNPAII C-terminal domain and propose a model in which the host RNAP II is the anchor for transcription and replication of the viral genome. Our data open new perspectives on the spatial coupling of viral transcription and replication and the coordinated balance between these two activities.

Influenza viruses are major human pathogens which cause annual epidemics (type A, B and C viruses)<sup>1</sup> and have zoonotic and pandemic potential (type A viruses)<sup>2</sup>. The viral RNA dependent RNA polymerase (FluPol) is a key determinant of viral host-range and pathogenicity, and a prime target for antiviral drugs<sup>3</sup>. It is a ~260 kDa heterotrimer composed of the subunits PB1 (polymerase basic protein 1), PB2 (polymerase basic protein 2) and PA (polymerase acidic protein)<sup>4</sup>. In viral particles each of the eight negative-sense RNA genomic segments (vRNAs) is associated with one copy of the viral polymerase and encapsidated with multiple copies of the nucleoprotein (NP) to form viral ribonucleoproteins (vRNPs)<sup>5</sup>. Upon viral infection, incoming vRNPs are imported into the nucleus in which FluPol performs transcription and replication of the viral genome through distinct primed<sup>6</sup> and unprimed<sup>7</sup> RNA synthesis mechanisms, respectively. Determination of high-resolution FluPol structures by X-ray crystallography

and cryogenic electron microscopy (cryo-EM) have revealed that a remarkable conformational flexibility allows FluPol to perform mRNA transcription<sup>8–12</sup>. Recently, first insights into the molecular process of FluPol genome replication were revealed, however the molecular details remain largely to be determined<sup>13–16</sup>.

Transcription of negative-sense vRNAs into mRNAs occurs via a process referred to as ‘cap-snatching’, whereby nascent 5′-capped host RNA polymerase II (RNAP II) transcripts are bound by the PB2 cap-binding domain<sup>9</sup>, cleaved by the PA endonuclease domain<sup>8</sup>, and used as capped primers to initiate transcription<sup>10</sup>. Polyadenylation is achieved by a non-canonical stuttering mechanism at a 5′-proximal oligo(U) polyadenylation signal present on each vRNA<sup>12,17</sup>. Influenza mRNAs, therefore, harbour 5′-cap and 3′-poly(A) structures characteristic of cellular mRNAs and are competent for translation by the cellular machinery. The accumulation of primary transcripts of

<sup>1</sup>Institut Pasteur, Université Paris Cité, CNRS UMR 3569, RNA Biology of Influenza Virus, Paris, France. <sup>2</sup>European Molecular Biology Laboratory, Grenoble, France. <sup>3</sup>Unit 17 “Influenza and other Respiratory Viruses”, Robert Koch Institut, Berlin, Germany. <sup>4</sup>These authors contributed equally: Tim Krischuns, Benoît Arragain. ✉e-mail: [tim.krischuns@pasteur.fr](mailto:tim.krischuns@pasteur.fr); [cusack@embl.fr](mailto:cusack@embl.fr); [nadia.naffakh@pasteur.fr](mailto:nadia.naffakh@pasteur.fr)

incoming vRNPs leads to de novo synthesis of viral proteins, including FluPol and NP, which are thought to trigger genome replication upon nuclear import<sup>18</sup>.

Replication generates exact, full-length complementary RNAs (cRNAs) which in turn serve as templates for the synthesis of vRNAs. Unlike transcription, replication occurs through a primer-independent, de novo initiation mechanism which occurs either at the first nucleotide of the 3'-vRNA template or at nucleotide 4 and 5 of the 3'-cRNA template followed by realignment prior to elongation<sup>7,19,20</sup>. Nascent cRNAs and vRNAs are encapsidated by NP and FluPol to form cRNPs and vRNPs, respectively<sup>21</sup>. Progeny vRNPs can then perform secondary transcription and replication or, late during the infection cycle, be exported from the nucleus to the cytoplasm to be incorporated into new virions.

FluPol associates dynamically with many cellular proteins, among which the interaction with the host RNAP II transcription machinery<sup>22–24</sup> and the nuclear acidic nuclear phosphoprotein 32 (ANP32)<sup>25</sup> are essential for a productive viral infection. A direct interaction between FluPol and the disordered C-terminal domain (CTD) of the largest RNAP II subunit is essential for 'cap-snatching'<sup>26</sup>. In mammals, the CTD consists of 52 repeats of the consensus sequence Y<sub>1</sub>-S<sub>2</sub>-P<sub>3</sub>-T<sub>4</sub>-S<sub>5</sub>-P<sub>6</sub>-S<sub>7</sub><sup>27</sup>. To perform 'cap-snatching', FluPol binds specifically to the CTD when phosphorylated on serine 5 residues (pS5 CTD), which is a hallmark of RNAP II in a paused elongation state<sup>28</sup>. Structural studies revealed bipartite FluPol CTD-binding sites for influenza type A, B and C viruses<sup>22,23,29</sup>. For type A viruses, both CTD-binding sites, denoted site 1A and site 2A, are on the PA-C-terminal domain (PA-Cter) and consist of highly conserved basic residues that directly interact with pS5 moieties of the CTD<sup>23</sup>.

The host ANP32 proteins consist of a structured N-terminal leucine-rich repeat (LRR) domain and a disordered C-terminal low complexity acidic region (LCAR) and are essential for influenza virus genome replication<sup>30,31</sup>. Differences in avian and mammalian ANP32 proteins represent a major driver of FluPol adaptation upon zoonotic infections of mammalian species with avian influenza viruses<sup>25,31,32</sup>. A type C virus FluPol (Flu<sub>C</sub>Pol) co-structure with ANP32A shows that the ANP32A LRR domain binds to an asymmetrical influenza polymerase dimer, which is presumed to represent the FluPol replication complex<sup>13</sup>. The two FluPol conformations observed in this dimer would represent the RNA-bound replicase (FluPol<sub>(R)</sub>), which synthesises de novo genomic RNA and the encapsidase (FluPol<sub>(E)</sub>), which binds the outgoing replication product and nucleates formation of the progeny RNP<sup>18</sup>. In addition, the ANP32 LCAR domain was shown to interact with NP<sup>15,16</sup>, leading to a model in which the ANP32 LCAR domain recruits NP to nascent exiting RNA, thus, in combination with the FluPol<sub>(E)</sub>, ensuring efficient co-replicative encapsidation of de novo synthesised genomic RNAs.

Structural, biochemical and functional studies to date have led to a model in which the FluPol binds either to the pS5 CTD or to ANP32, driving RNPs towards transcription or replication, respectively<sup>18</sup>. Here, we provide genetic evidence that the FluPol CTD-binding interface is essential not only for transcription but also for replication of the viral genome. We show that the CTD, jointly with ANP32, enhances FluPol replication activity and we demonstrate in structural studies that CTD-binding to the FluPol is consistent with replication activity. We therefore propose a model in which transcription and replication of the influenza virus genome occur in association with the RNAP II CTD thereby allowing an efficient switching between the two activities.

## Results

### The FluPol CTD-binding interface is essential for replication of the viral genome

Previous studies demonstrated that the FluPol PA-Cter domain is crucial for FluPol transcription by mediating the interaction between the FluPol transcriptase (FluPol<sub>(T)</sub>) and the host RNAP II pS5 CTD for

'cap-snatching' through direct protein-protein interaction (Fig. S1A, FluPol<sub>(T)</sub>)<sup>22,23</sup>. Recent cryo-EM structures of Flu<sub>C</sub>Pol suggest that the PA-Cter domain is also involved in the formation of the FluPol replication complex<sup>13</sup>. As there are no structures published for an equivalent Flu<sub>A</sub>Pol complex, we used the Flu<sub>C</sub>Pol structures to generate a model of the Flu<sub>A</sub>Pol replicase-encapsidase complex (Fig. 1A FluPol<sub>(R)</sub> and FluPol<sub>(E)</sub>). According to our model, the key CTD-binding residues are surface exposed and potentially competent for CTD-binding in the FluPol<sub>(R)</sub> (Fig. 1A, left). Strikingly, FluPol<sub>(E)</sub> CTD-binding residues are in close proximity to the ANP32-binding interface (Fig. 1A, right). We therefore hypothesised that the FluPol CTD-binding interface may have a role beyond the described association with the pS5 CTD in the FluPol<sub>(T)</sub> conformation namely in the assembly of the FluPol<sub>(R)</sub>-ANP32-FluPol<sub>(E)</sub> replication complex.

To assess this dual role hypothesis, we examined whether binding of the A/WSN/1933 (WSN) FluPol to human ANP32A (huANP32A) is altered by previously described FluPol CTD-binding mutations (PA K289A, R454A, K635A or R638A)<sup>23</sup>. For this purpose, we made use of cell-based *G. princeps* split-luciferase protein-protein complementation assays as described before in refs. 33,34. As expected from our previous work<sup>23</sup>, steady-state levels of the wild-type (WT) and mutant PA protein are similar or slightly reduced in the case of the PA K635A mutant (Fig. S1B) and each mutation reduces FluPol binding to the CTD (Fig. 1B, grey bars) as well as FluPol activity as measured in a vRNP reconstitution assay (Fig. 1B, hatched bars). Remarkably, they also reduce FluPol binding to huANP32A (Fig. 1B, light blue bars) as well as chicken ANP32A (chANP32A) (Fig. 1B, dark blue bars). We then tested mutations previously shown to impair viral multiplication more strongly, either double mutations (PA K635A + R638A in site 1A, K289 + R454A in site 2A)<sup>23</sup> or mutations in site 2A (PA S420E, Δ550-loop)<sup>22</sup>. Steady-state levels of the mutant PA proteins remain unchanged (Fig. S1C). The observed reductions in CTD-binding signal are in the same 25–50% range as with the first series of single mutants (Fig. S1D, grey bars), suggesting that disruption of a single site, 1A or 2A, largely preserves the CTD-binding signal through the remaining intact binding site. Nevertheless, stronger reductions in huANP32A- and chANP32A-binding and FluPol activity are observed with the second series of mutants compared to the first (Fig. S1D, blue and hatched bars, compared to Fig. 1B), thus reflecting their stronger impact on viral multiplication<sup>22,23</sup>.

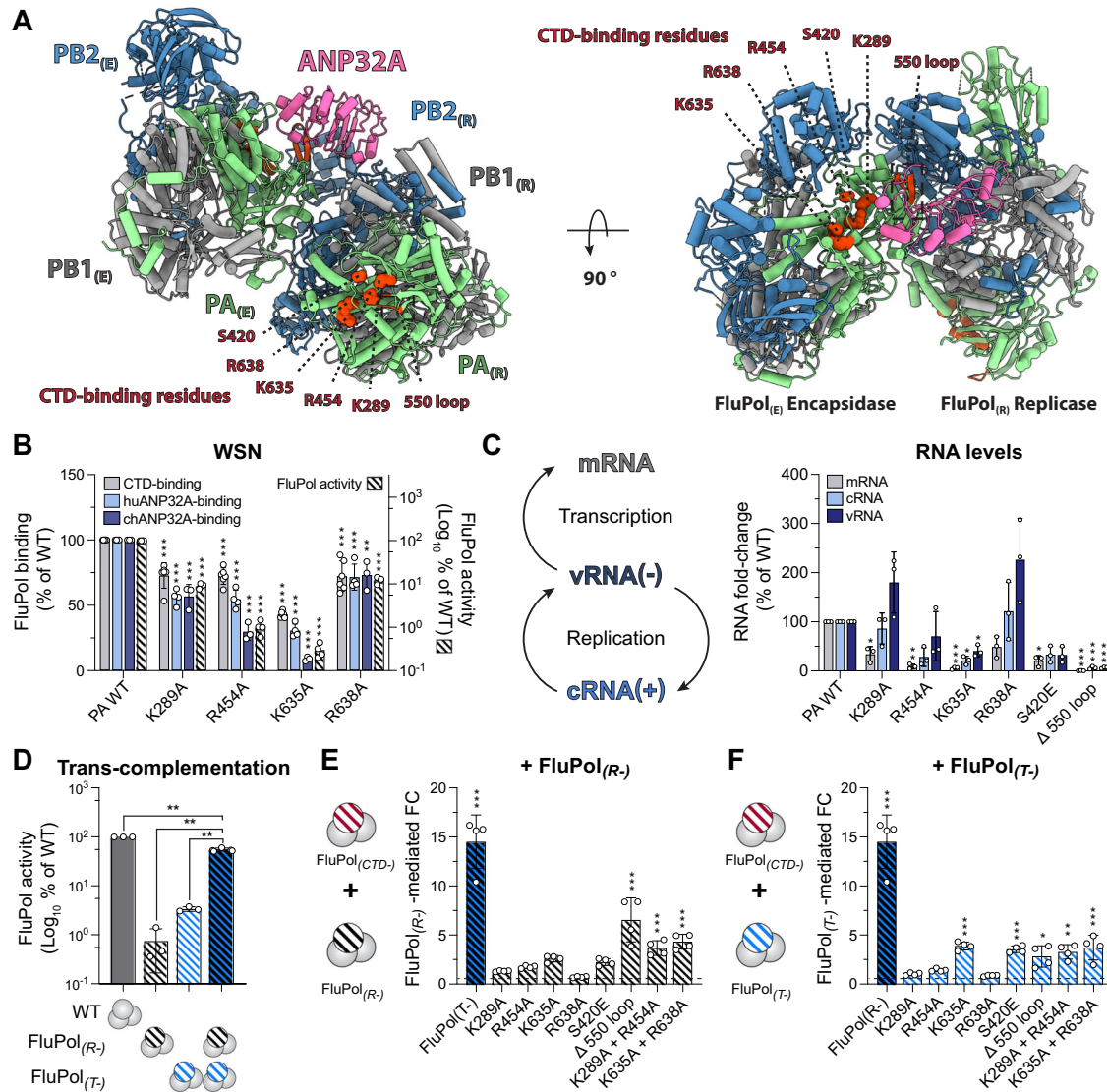
We then quantified the accumulation of viral RNA species (NA mRNA, cRNA and vRNA) in the vRNP reconstitution assay by strand-specific RT-qPCR<sup>35</sup>. All tested mutants show decreased mRNA accumulation levels (Fig. 1C, grey bars). Strikingly, the PA R454A, K635A, S420E and Δ550-loop mutants also show decreased cRNA and vRNA levels, indicating that both transcription and replication are affected (Fig. 1C, light and dark blue bars, respectively). The variations in cRNA and vRNA levels mirror the variations in FluPol chANP32-binding signal shown in Fig. 1B and Fig. S1D. Such a correlation is not observed with the FluPol huANP32-binding signal, most likely because in our *G. princeps* split-luciferase protein-protein complementation assay huANP32A binds FluPol weakly compared to chANP32, in agreement with previous reports<sup>33</sup>. We speculate that in the presence of huANP32 the assay is not discriminant enough to resolve mutant-to-mutant differences. In the presence of the PA K289A and R638A mutants, cRNA and vRNA levels appear unchanged or even increased (Fig. 1C). However, like the R454A and 635A mutants, they show not only a ~80% decrease of mRNA:vRNA ratios (Fig. S1E, grey bars) but also a ~50% decrease of cRNA:vRNA ratios (Fig. S1E, blue bars), indicating an imbalance in the accumulation of replication products compared to the WT FluPol.

To further support the notion that the FluPol CTD-binding interface is essential for replication, we made use of transcription-defective (FluPol<sub>(T-)</sub>), PA D108A<sup>8</sup> and replication-defective (FluPol<sub>(R-)</sub>), PA K664M<sup>36</sup> mutants of the WSN FluPol. When expressed alone, FluPol<sub>(T-)</sub>

and FluPol<sub>(R-)</sub> show low activity, as expected (Fig. 1D). However, when they are both co-expressed, FluPol<sub>(T-)</sub> and FluPol<sub>(R-)</sub> trans-complement each other and the signal exceeds the background defined as the sum of the activity of FluPol<sub>(T-)</sub> and FluPol<sub>(R-)</sub> alone. This trans-complementation system was used to reveal a defect in transcription or in replication (Fig. 1E, F). Upon co-expression with the FluPol<sub>(R-)</sub> mutant, trans-complementation is observed for most of the investigated CTD-binding mutants (Fig. 1E), confirming that they are transcription-defective. Similarly, upon co-expression with the

FluPol<sub>(T-)</sub> mutant, trans-complementation is observed for most of the investigated FluPol CTD-binding mutants (Fig. 1F) demonstrating that a replication defect is introduced when the CTD-binding interface is mutated.

Overall, our findings show that the FluPol CTD-binding interface, beyond its function for FluPol ‘cap-snatching’, is also involved in the FluPol-ANP32 interaction and is essential for genome replication. They are not limited to the WSN strain, as a dual-effect of the mutations PA K289A, R454A, K635A and R638A on CTD-, huANP32A- and



**Fig. 1 | The FluPol CTD-binding interface is essential for replication of the viral genome.** **A** Model of actively replicating Zhejiang-H7N9 Flu<sub>A</sub>Pol replicase-ANP32-encapsidase (FluPol<sub>(R)</sub>-ANP32A-FluPol<sub>(E)</sub>). Ribbon diagram representation with PA (green), PB1 (grey), PB2 (blue) and ANP32A (pink). The model was constructed by superposing Zhejiang-H7N9 elongating Flu<sub>A</sub>Pol domains (PDB: 7QTL)<sup>43</sup> for the FluPol<sub>(R)</sub> and apo-dimer Zhejiang-H7N9 Flu<sub>A</sub>Pol (PDB: 7ZPL)<sup>43</sup> for the FluPol<sub>(E)</sub>, on those of Flu<sub>C</sub>Pol replication complex (PDB: 6XZQ)<sup>13</sup>. ANP32A was left unchanged. Key FluPol CTD-binding residues are highlighted on the FluPol<sub>(R)</sub> (left) and FluPol<sub>(E)</sub> (right). **B** WSN FluPol binding and activity assays in HEK-293T cells. Left Y-axis (linear scale): FluPol (PA WT or indicated PA CTD-binding mutants) binding to the CTD (grey bars), huANP32A (light blue bars) and chANP32A (dark blue bars) was assessed using split-luciferase-based complementation assays. Right Y-axis (logarithmic scale): FluPol activity (hatched bars) was assessed by vRNP reconstitution, using a model vRNA encoding the Firefly luciferase (Fluc-vRNA). The luminescence signals are represented as a percentage of PA WT (mean ± SD, n = 6, 4, 4, 3, \*\*p < 0.002, \*\*\*p < 0.001 (one-way ANOVA; Dunnett’s multiple comparisons test).

**C** WSN vRNPs were reconstituted in HEK-293T cells using the NA vRNA. Steady-state levels of NA mRNA, cRNA and vRNA were quantified by strand-specific RT-qPCR<sup>35</sup> and are represented as a percentage of PA WT (mean ± SD, n = 3, \*p < 0.033, \*\*p < 0.002, \*\*\*p < 0.001, one-way repeated measure ANOVA; Dunnett’s multiple comparisons test). **D–F** FluPol trans-complementation assays upon vRNP reconstitution in HEK-293T cells with a model Fluc-vRNA. **D** Trans-complementation of a transcription-defective (FluPol<sub>(T-)</sub>, PA D108A)<sup>8</sup> and replication-defective (FluPol<sub>(R-)</sub>, PA K664M)<sup>36</sup> FluPol. Luminescence is represented as percentage of PA WT (mean ± SD, n = 3, \*\*p < 0.002, one-way repeated measure ANOVA; Dunnett’s multiple comparisons test). WSN vRNPs were co-expressed with **(E)** a replication-defective (FluPol<sub>(R-)</sub>, PA K664M)<sup>36</sup> or **(F)** a transcription-defective (FluPol<sub>(T-)</sub>, PA D108A)<sup>8</sup> FluPol. Luminescence signals are represented as the fold-change (FC) relative to the background, which is defined as the sum of signals measured when the FluPol<sub>(CTD-)</sub> and FluPol<sub>(R-/T-)</sub> were transfected alone (mean ± SD, n = 4, \*p < 0.033, \*\*p < 0.002, \*\*\*p < 0.001, two-way ANOVA; Sidak’s multiple comparisons test). Source data are provided as a Source data file.

chANP32A-binding is also observed with the FluPol of the human isolate A/Anhui/01/2013 (H7N9) (Anhui-H7N9), a virus of avian origin which harbours the human-adaptation signature PB2 K627 (Fig. S1F, G). Sequence alignments of the PB1, PB2, PA and NP proteins from the influenza A viruses used in this study are shown in Fig. S2.

### Serial passaging of FluPol PA K289A, R454A, K635A, R638A mutant viruses selects for adaptive mutations which restore FluPol binding to the CTD and huANP32A

We previously rescued by reverse genetic procedures recombinant WSN PA mutant IAVs (PA K289A, R454A, K635A, R638A) that were highly attenuated, formed pinhead-sized plaques and had acquired multiple non-synonymous mutations<sup>23</sup>. In order to investigate whether and how the recombinant viruses adapt to defective CTD and ANP32-binding (Fig. 1), we subjected them to eight serial cell culture passages in MDCK-II cells followed by titration and next generation sequencing (NGS) of the viral genome at passage 1 to 4 and 8 (Fig. 2A, p1–4 and p8). The plaque size of the passaged viruses progressively increased (Fig. 2B and S3A), which went along with distinct patterns of adaptive mutations (Source data file), suggesting that cell culture passaging increased the replicative capacity of the PA mutant viruses. True reversions that require at least two nucleotide substitutions did not emerge. The observed second-site mutations are primarily located on the FluPol subunits as well as the NP (Fig. 2C, S3B and S4A) and are mostly surface exposed (Fig. 2D and S4B–D). Several FluPol second-site mutations are located close to the FluPol CTD-binding sites (Fig. 2D, PA D306N, C489R, R490I, R496Q) while some PB2 second-site mutations are located in flexible hinge-subdomains (Fig. 2D, PB2 D253G, S286N, M535I). Overall, the FluPol second-site mutations show no local clustering or common function, indicating that diverse routes of adaptation occurred. NP second-site mutations are also surface exposed, but to our knowledge have not been associated to date with any specific function (Fig. S4D)<sup>37</sup>. The observed reversion pathways most likely represent a fraction of possible reversion pathways, especially as passaging was performed at low MOIs and, therefore, less frequent variants may have been lost.

We first focused on the reversion pathway of PA K289A mutant virus. Recombinant viruses with combinations of FluPol second-site mutations observed along the PA K289A reversion pathway at p1, p4 and p8 recapitulate the progressive increase in plaque size observed during serial passaging (Fig. 2E and S3C), demonstrating that the FluPol second-site mutations are leading to the observed increase in viral replicative capacity. PB2 and PA proteins with second-site mutations which occurred during passaging of PA K289A mutant virus, accumulate at similar levels compared to the WT proteins (Fig. S3D). However, the combinations of FluPol second-site mutations result in sequential increases and decreases of CTD-binding, huANP32A-binding and/or FluPol activity (Fig. 2F). The highest levels of CTD-binding, huANP32A-binding and FluPol activity are observed for the combination of mutations observed at p8 (PA K289A + PB2 D253G + PA C489R), which goes along with the highest replicative capacity observed for the passaged virus at p8 (Fig. 2B, E).

The reversion pathways of the PA R454A, K635A and R638A mutant viruses were analysed in the same way (Fig. S3E). Across the tested combinations of primary FluPol mutations and second-site mutations that occurred during passaging, FluPol activity shows a strong positive correlation with huANP32A-binding (Fig. 2G, left panel) as well as a positive correlation with CTD-binding (Fig. 2G, middle panel) but not with FluPol-binding to other known cellular partners, such as DDX5 (Fig. 2G, right panel) or RED (Fig. S3F)<sup>38</sup>.

Overall, we show that serial cell culture passaging of PA K289A, R454A, K635A and R638A mutant viruses selects for adaptive mutations that rescue viral polymerase activity by restoring FluPol binding to the CTD and/or huANP32A, which agrees with our initial observation

that the investigated FluPol mutants are impaired for CTD-binding as well as ANP32-binding (Fig. 1).

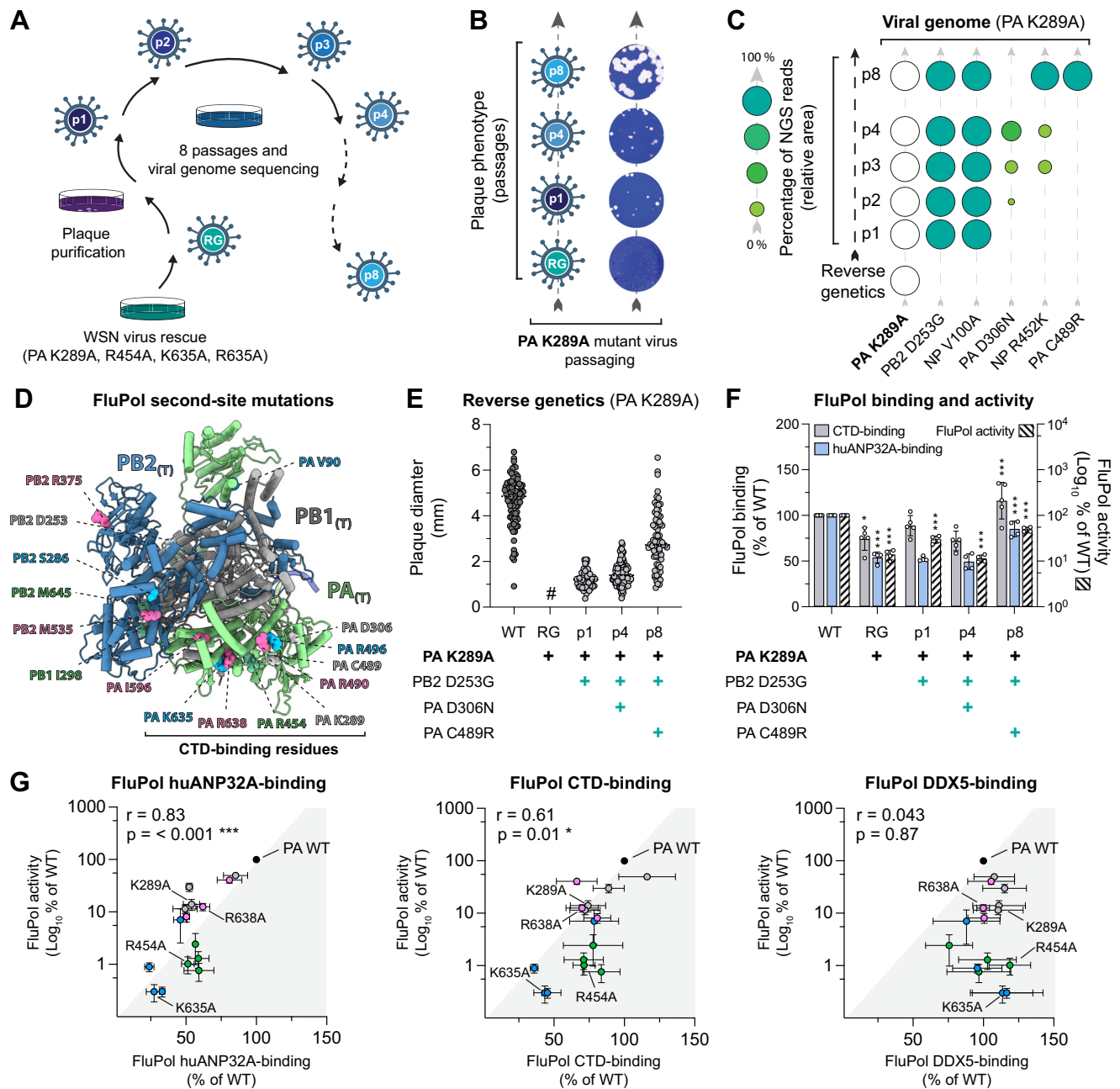
### CTD overexpression leads to an increased binding of FluPol to huANP32

To further document a potential interplay between CTD and ANP32-binding to FluPol, we tested the effect of mCherry-CTD (CTD-WT) overexpression on FluPol binding to huANP32A as well as FluPol oligomerisation in cultured cells. A FluPol binding-deficient CTD in which all serine 5 residues were replaced with alanines was used as a specificity control (CTD-S5A)<sup>22</sup>. mCherry-CTD-WT and -CTD-S5A accumulate to similar levels (Fig. S5A), however mCherry-CTD-WT leads to a significant increase of FluPol huANP32A-binding compared to the FluPol binding-deficient CTD-S5A mutant, suggesting that the CTD facilitates FluPol-binding to huANP32A (Fig. 3A). mCherry-CTD-WT overexpression also leads to significantly increased WSN FluPol oligomerisation compared to the FluPol binding-deficient CTD-S5A mutant (Fig. 3B), as measured in a split-luciferase complementation assay<sup>39</sup>. A similar increase in oligomerisation was also observed with the Anhui-H7N9 and B/Memphis/13/2003 FluPols (Fig. S5B). Various oligomeric FluPol species have been described<sup>13,40,41</sup>. There is evidence that the utilised split-luciferase-based assay, when FluPol is overexpressed in the absence of any other overexpressed viral or cellular protein, detects the symmetrical FluPol dimer species<sup>39,41,42</sup>. Upon CTD overexpression, it is unclear which oligomeric FluPol species is affected. However, the oligomerisation signal is still increased in the presence of the PB2 NEQ71-73AAA mutant, which abolishes FluPol symmetrical dimerisation (Fig. S5C–E)<sup>39,41</sup>, indicating that CTD overexpression affects FluPol oligomers distinct from the symmetrical FluPol dimer. Moreover, the increased signal upon CTD-WT overexpression is less pronounced in HEK-293T cells depleted for huANP32A and huANP32B (ANP32AB KO) (Fig. S5F, G), most clearly so for B/Memphis/13/2003 FluPol, suggesting that the CTD-dependent increase in FluPol oligomerisation is at least partially dependent on huANP32A and/or huANP32B. Altogether, the data show that CTD overexpression leads to increased binding of FluPol to huANP32A as well as FluPol oligomerisation, suggesting that the CTD can bring two or more FluPols in close proximity with each other in conjunction with ANP32A, thereby promoting FluPol replication.

### De novo FluPol RNA synthesis activity is enhanced in vitro in the presence of CTD and ANP32A

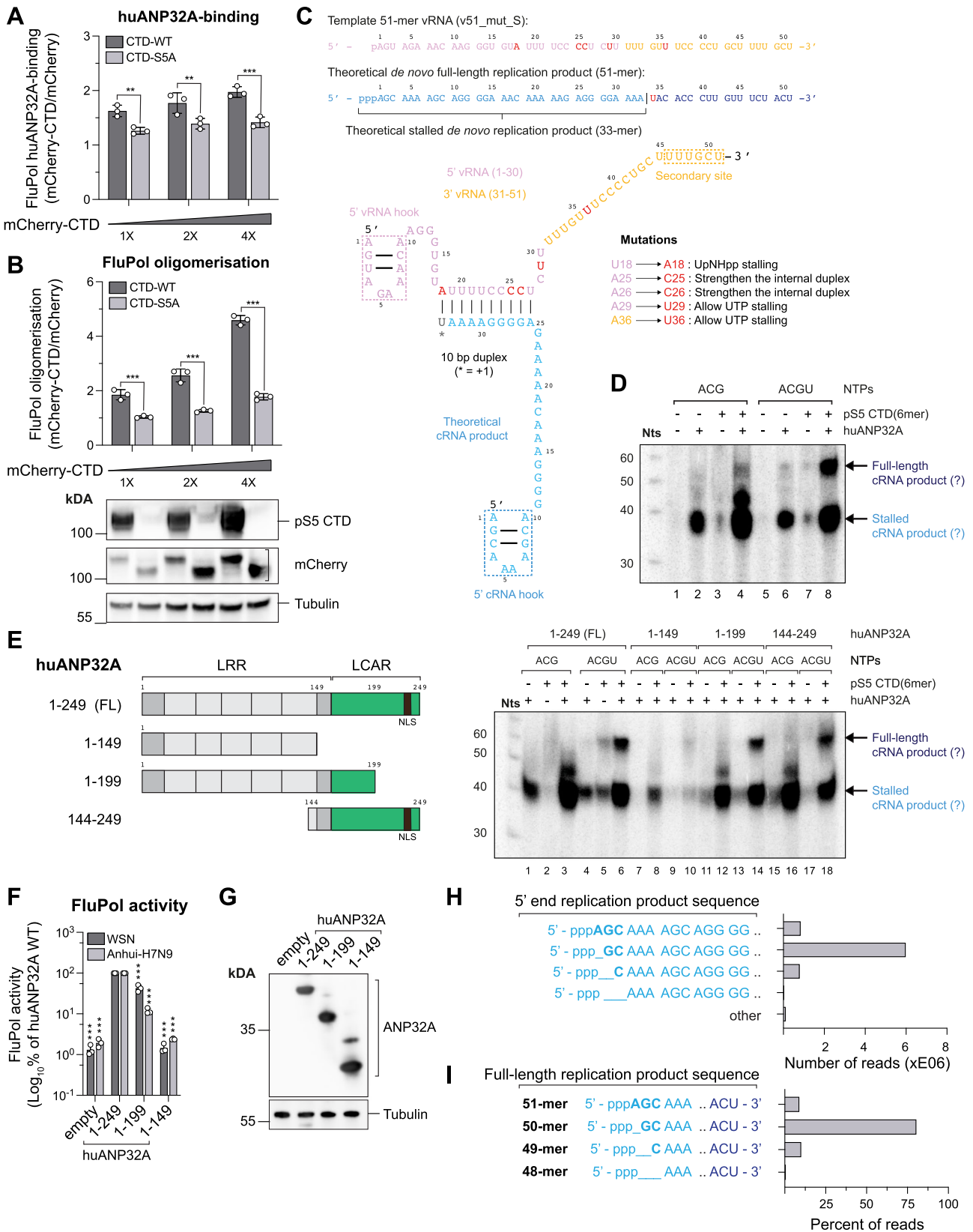
To investigate by biochemical and structural approaches whether in vitro replication is dependent on ANP32A and pS5 CTD, we performed unprimed RNA synthesis assays using recombinant A/Zhejiang/DTID-ZJU01/2013(H7N9) FluPol<sup>43</sup> bound to a 51-mer vRNA loop template (denoted v51\_mut\_S). This polymerase, referred to below as Zhejiang-H7N9, derives from a human isolate of an avian strain that possesses PB2 E627 together with the mammalian adaptation mutation PB2 N701 (Fig. S2)<sup>44</sup>. The v51\_mut\_S template sequence derives from joining the first 30 and last 21 nucleotides (nts) of the respective 5' and 3' vRNA ends of Zhejiang-H7N9 segment 4. It was modified so that a stalled elongation product of 33 nts could be produced by using only AGC nts or replacing UTP in the reaction by non-hydrolysable UpNHpp (Fig. 3C, light blue letters). This 33-mer cRNA product was estimated to be long enough that its 5' hook (1–10) could bind to the capsidase, according to the modelled Zhejiang-H7N9 replicase-encapsidase structure. The addition of all NTPs would in fine produce a full-length 51-mer replication product without poly-adenylation (since the oligo-U sequence is reduced to 4xU due to changing U18 to A)<sup>12</sup>, assuming RNA synthesis proceeded to the 5' end of the template, rather than terminating prior to reading through the 5' hook (Fig. 3C, dark blue letters).

De novo replication assays were performed with WT Zhejiang-H7N9 FluPol, using three (ACG) or four (ACGU) nts, excess apo-polymerase, that could serve as capsidase, recombinant huANP32A and



**Fig. 2 | Serial passaging of mutant WSN viruses with mutations at the FluPol-CTD interface selects for adaptive mutations which restore FluPol binding to the CTD and huANP32A.** **A** Schematic representation of the passaging experiment. **B, C** Passaged PA K289A viruses. **B** Plaque phenotype. Representative images of crystal violet stained plaques assays on reverse genetics (RG), p1, p4 and p8 supernatants are shown. **C** Short-read next generation sequencing of the viral genome. Second site mutations found in  $\geq 10\%$  of reads in at least one passage are shown (green circles). The fraction of reads showing a given mutation is indicated by the circle area and shades of green (schematic on the left:  $<25$ ,  $25-50$ ,  $50-75$  and  $>75\%$  of reads). **D** Ribbon diagram representation of FluPol<sub>T</sub> conformation (A/NT/60/1968, PDB: 6RR7)<sup>41</sup>. Second-site mutations observed during passaging of the PA K289A, R454A, K635A and R638A viruses are indicated in grey, green, blue and pink fonts respectively. **E, F** FluPol second-site mutations observed during passaging of the PA K289A virus. **E** Recombinant viruses with PA K289A and the indicated second-site mutations were produced ( $n = 2$ ). Two independent RG supernatants were titrated, plaque diameters (mm) were measured, each dot represents one

plaque (see Fig. S3C). (#) pinhead-sized plaques. **F** Cell-based FluPol binding and activity assays as in Fig. 1B. Left Y-axis (linear scale): FluPol binding to the CTD (grey bars) and huANP32A (blue bars). Right Y-axis (logarithmic scale): FluPol activity (hatched bars). Luminescence signals are represented as a percentage of WT FluPol. Stars indicate statistical significance when passage N is compared to passage N-1, or RG is compared to WT FluPol (mean  $\pm$  SD,  $n = 5, 4, 4$ , \* $p < 0.033$ , \*\* $p < 0.002$ , \*\*\* $p < 0.001$ , one-way ANOVA; Tukey's multiple comparisons test). **G** For each primary mutation and FluPol genotype observed upon serial passaging, FluPol binding to huANP32A, CTD and DDX5 (x-axis in the left, middle and right panel, respectively) were plotted against the FluPol activity (y-axis). Binding and activity data are represented separately in (F) and Fig. S3E. Combinations of mutations which appeared during passaging of the WSN PA K289A, R454A, K635A and R638A mutant viruses are highlighted in grey, green, blue and pink respectively. (mean  $\pm$  SD,  $n = 4$ ,  $r$ : Pearson correlation coefficient, \* $p < 0.033$ , \*\*\* $p < 0.001$ , two-tailed 95% confidence interval). Source data are provided as a Source data file.



6-repeats pS5 CTD peptide (pS5 CTD(6mer)) (Fig. 3D). In reactions with ACG only (Fig. 3D, lane 1–4), the principal product is around 38–40 nts (with another at ~45 nts), whose significant production is only visible in the presence of huANP32A (Fig. 3D, lane 2) and strongly enhanced when both the CTD and huANP32A are added (Fig. 3D, lane 2 vs 4). When all four nts are added (Fig. 3D, lane 5–8), the presence of CTD

and huANP32A again strongly enhances RNA synthesis with the production of both a similar product as with ACG alone (38–40 nts) and a significant amount of a longer ~55–56 nts product (Fig. 3D, lane 6 vs 8). Although the apparent size of these products does not correspond with the theoretical length of the stalled and full-length products (they are consistently around 5–6 nts higher), nevertheless, we think that

**Fig. 3 | FluPol replication activity is enhanced in the presence of CTD and ANP32A.** WSN FluPol binding to huANP32A (A) and WSN FluPol oligomerisation (B) in the presence of increasing amounts of pS5 CTD, as assessed using split-luciferase-based complementation assays. Plasmids encoding mCherry or mCherry fused to CTD-WT (dark grey bars) or CTD-S5A (serine five residues replaced with alanines, light grey bars) were co-transfected in increasing amounts (mean  $\pm$  SD,  $n = 3$ ,  $^{**}p < 0.002$ ,  $^{***}p < 0.001$ , two-way ANOVA; Sidak's multiple comparisons test). Luminescence signals are represented as fold-changes compared to untagged mCherry co-expression. In (B), cell lysates were analysed by western blot using the indicated antibodies ( $n = 1$ ). C 51-mer vRNA template (v51\_mut\_S) derived from segment 4 of A/Zhejiang/DUID-ZJU01/2013(H7N9)/KJ633805 and used in *in vitro* replication activity assays. The vRNA 5' end (1–30) and 3' end (31–51) are coloured in pink and gold, respectively, with introduced mutations in red. The theoretical *de novo* full-length replication product is coloured in light blue (1–33) and dark blue (35–51), with U34 in red. The expected stalled elongation state is schematically represented. *de novo* replication activity assays of Zhejiang-H7N9 FluPol using v51\_mut\_S, in the presence of either 3 NTPs (AUG) or 4 NTPs (AUGC), with or

without pS5 CTD(6mer) and huANP32A full-length (D,  $n = 3$ ) or deletion mutants (E,  $n = 3$ ). Tentative full-length and stalled replication products are indicated by an arrow. LRR leucine-rich repeat, LCAR low complexity acidic region, NLS nuclear localisation signal. Nts: molecular weight marker. F WSN and Anhui-H7N9 vRNPs were reconstituted in ANP32AB KO cells with a model Fluc-vRNA, and co-expressed with the indicated huANP32A proteins (V5-tagged and fused to SV40-NLS). Luminescence signals are represented as a percentage of huANP32A FL (mean  $\pm$  SD,  $n = 3$ ,  $^{***}p < 0.001$ , two-way ANOVA; Sidak's multiple comparisons test). G Lysates of ANP32AB KO cells transiently expressing the indicated huANP32A proteins were analysed by western blot using an anti-V5 antibody ( $n = 1$ ). H, I RNA-sequencing of Zhejiang-H7N9-4M FluPol *de novo* replication products in the presence of all NTPs, the pS5 CTD(6mer) and huANP32A. H The number of reads is plotted according to the recurrence of the exact 5' cRNA motifs indicated on the left. Reads that do not encompass these motifs are plotted as 'Other'. I The percentages of full-length replication products are plotted according to the 5' terminal nucleotide indicated on the left. Source data are provided as a Source data file.

they correspond to these products for the following reasons. Firstly, for both the AGC+UpNHpp and AGCU reactions, we have obtained high-resolution cryo-EM structures of stalled elongation states, unambiguously showing products of size 33-mer and 35-mer respectively (32-mer/34-mer if RNA synthesis starts at G2 of the template instead of U1, see below) that we presume to correspond to the bands observed. Secondly, for the AGCU reaction, the difference in size of the two bands is 16–18 nts, which corresponds to the expected difference between the stalled and full-length (hook read-through) products. We are unable to explain the origin of the ~45 nts product with AGC only. Despite the remaining uncertainties in the assignment of the products, these results demonstrate the synergistic effect of CTD and huANP32A in significantly enhancing *de novo* cRNA synthesis.

We further investigated which domains of huANP32A are essential for the enhancement of *de novo* RNA synthesis activity in the presence, or not, of pS5 CTD(6mer) peptide. We compared the full-length huANP32A ('FL' lane 1-6) with huANP32A subdomains corresponding to the LRR ('1–149' lane 7-10), the LRR and LCAR N-terminal region ('1–199' lane 11-14), or the LCAR and LRR C-terminal region ('144–249' lane 15–18) (Fig. 3E). While without addition of CTD, only FL huANP32A supported significant replication product synthesis, we show that the 144–199 region of huANP32A is most critical to observe the synergistic CTD- and ANP32-dependent enhancement of replication product synthesis (Fig. 3E, right panel). These *in vitro* findings are consistent with cell-based assays in which FluPol activity was investigated in the presence of the same huANP32A C-terminal deletion mutants (Fig. 3F, G). While the presence of the CTD alone was shown to enhance FluPol transcriptional activity *in vitro*<sup>22,29</sup>, enhancement of FluPol replicational activity requires both the presence of the CTD and the 144–249 domain of ANP32, suggesting that CTD and the ANP32 LCAR domain are jointly priming FluPol for replication.

Instead of using the WT Zhejiang-H7N9 FluPol, which forms a robust symmetrical dimer *in vitro*<sup>22</sup> that could compete with the formation of an active asymmetric FluPol<sub>(R)</sub>-FluPol<sub>(E)</sub> dimer, we subsequently used a Zhejiang-H7N9 FluPol that harbours four mutations (PA E349K, PA R490I, PBI K577G, and PB2 G74R) and is referred to below as the 4 M mutant. PA E349K, PBI K577G, and PB2 G74R were shown to disrupt the symmetrical apo FluPol dimer interface<sup>39</sup>, while PA R490I appeared during passaging of the PA R638A mutant virus and increases FluPol-ANP32A binding in a cell-based assay (Fig. S3E). Indeed, Zhejiang-H7N9 4 M shows enriched monomeric particles while preserving a similar *in vitro* activity when compared to the Zhejiang-H7N9 WT FluPol (Fig. S6A, B). Accordingly, in cell-based assays Anhui-H7N9 4 M shows decreased FluPol oligomerisation, increased binding to huANP32A, while showing activity levels comparable to the WT Anhui-H7N9 FluPol as measured in a vRNP reconstitution assay (Fig. S6C).

To determine whether 5' triphosphorylated full-length replication products were actually synthesised, we performed next-generation sequencing (NGS) of all RNAs present in the reaction mix (Fig. S6B, left panel, lane 8). To simplify the analysis by NGS of this heterogeneous RNA sample, we first degraded non-5' triphosphorylated RNAs with a 5'-phosphate-dependent exonuclease. This removes the v51\_mut\_S template, which carries a 5' monophosphate, and which is perfectly complementary to the expected product and could have formed unwanted double-stranded RNA (Fig. S6D). NGS analysis revealed that the majority of reads (i.e., premature and near full-length products) start with 5'-pppGC... instead of the expected 5'-pppAGC... (Fig. 3H). Despite the presence of many premature termination products (Fig. S6E), around 80% of near full-length products, are 50-mers, lacking the expected 5' pppA, with few 51-mers (5'-pppAGC...) and 49-mers (5'-pppC...) (Fig. 3I). This suggests that, *in vitro*, most RNA synthesis products are initiated *de novo* with a 5'-pppG at the second position of the v51\_mut\_S template (3'-UCG...).

### Structural analysis shows that both replication and transcription are consistent with CTD-binding to the FluPol

Using cryo-EM, we then sought to determine structures of actively replicating Zhejiang-H7N9 4M FluPol in the presence of huANP32A, pS5 CTD(6mer) peptide mimic and excess apo FluPol, by freezing grids with the *de novo* reaction stalled with UpNHpp (Fig. S6B, left panel, lane 4). Several high-resolution structures were obtained from a Krios data collection (Table 1, Fig. 4 and S7). These include FluPol in the recently described intermediate structure, denoted PB2-C(I)/ENDO(T), where the cap-domain is swung out (Fig. 4A, left panel, 3.25 Å resolution)<sup>14,45</sup>. A second structure was obtained at 2.9 Å resolution where FluPol with PB2-C mid-link and cap-binding domains take a replicase-like configuration (PB2-C(R)), with the 627-domain remaining flexible (Fig. 4A, middle panel). However, the PA endonuclease is still in an unrotated transcriptase-like conformation (ENDO(T)), thus precluding the usual interaction with the PB2-NLS domain, as previously observed for FluPol H5N1 replicase conformation (Fig. 4A, right panel)<sup>41</sup>. A third structure was obtained at a higher resolution (2.5 Å), in which PB2-C is not visible (Fig. 4B). These three structures show FluPol in the pre-initiation state mode A, with the 3' end of the v51\_mut\_S template (3'-UCG...) in the active site and the priming loop fully extended and well ordered (Fig. 4C, D and S8A). The template is positioned with C2 and G3 respectively at the -1 and +1 positions, whilst only the phosphate of U1 is visible (Fig. 4C, D). In addition, a stalled elongation state was obtained at 2.9 Å resolution in which FluPol encloses a template-product duplex, with incoming UpNHpp at the +1 active site position, and the 3' vRNA end bound to the secondary site (Fig. 4E and S8B). Finally, from a similar *de novo* reaction mix but in the



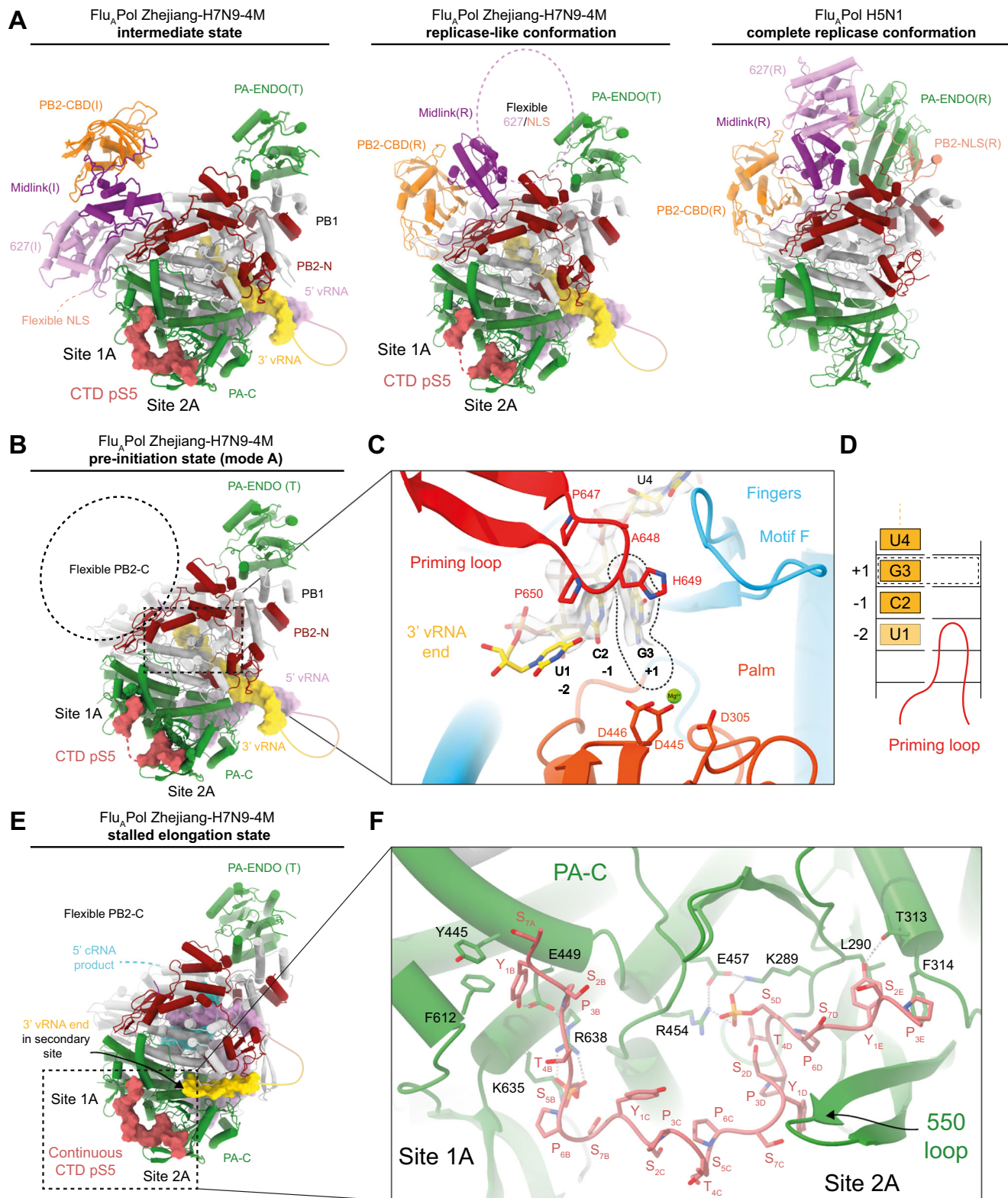
**Table 1 | Cryo-EM structures data collection, refinement and validation statistics**

Sample	A/H7N9 polymerase with Pol II pS5 CTD mimic bound in site 1A/2A			
Structure	Pre-initiation state in replicase (R)-like conformation PB2-C(R), ENDO(T)	Pre-initiation state core ENDO(T)	Pre-initiation state in intermediate (I)-conformation PB2-C(I), ENDO(T)	Elongation state stalled with UpNHpp ENDO(T)
PDB ID	PDB ID 8PMO	PDB ID 8PNP	PDB ID 8R3L	PDB ID 8PNQ
EMDB ID	EMD-17755	EMD-17782	EMD-18872	EMD-17783
<b>Data collection and processing</b>				
Microscope	ThermoFisher Krios TEM			
Voltage (kV)	300			
Camera	Gatan K3 direct electron detector mounted on a Gatan Bioquantum LS/967 energy filter			
Magnification	105000			
Nominal defocus range ( $\mu\text{m}$ )	-0.8/-2.0			
Electron exposure ( $e^-/\text{\AA}^2$ )	40			
Number of frames collected (no.)	40			
Number of frames processed (no.)	40			
Pixel size ( $\text{\AA}$ )	0.84			
Initial micrographs (no.)	6000			
Final micrographs (no.)	5883			
<b>Refinement</b>				
Particles per class (no.)	33395	260565	17121	14878
Map resolution ( $\text{\AA}$ ), 0.143 FSC	2.9	2.5	3.25	2.9
Model resolution ( $\text{\AA}$ ), 0.5 FSC	3.0	2.5	3.3	2.9
Map sharpening B factor ( $\text{\AA}^2$ )	-20	-30	-20	-20
Map versus model cross-correlation (CCmask)	0.8428	0.8901	0.8179	0.8826
<b>Model composition</b>				
Non-hydrogen atoms	16673	14464	17878	14917
Protein residues	1997	1716	2149	1727
Nucleotide residues	31	29	31	45
Water	-	43	-	-
Ligands	1 Mg	-	-	2 Mg, 1 UpNHpp
<b>B factors (<math>\text{\AA}^2</math>)</b>				
Protein	74.00	67.24	89.19	70.02
Nucleotide	72.82	60.96	90.96	83.39
Ligand	48.38	-	-	68.99
Water	-	33.97	-	-
<b>RMS deviations</b>				
Bond lengths ( $\text{\AA}$ )	0.003	0.003	0.002	0.005
Bond angles ( $^\circ$ )	0.475	0.499	0.478	0.465
<b>Validation</b>				
MolProbity score	1.53	1.35	1.53	1.40
All-atom clash score	4.79	4.32	6.41	3.86
Poor rotamers (%)	2.03	1.24	0.21	2.01
<b>Ramachandran plot</b>				
Favoured (%)	97.77	97.70	97.00	98.24
Allowed (%)	2.23	2.24	3.00	1.76
Outliers (%)	0.00	0.06	-	0.00

presence of all NTPs (Fig. S6B, left panel, lane 8), we obtained a structure at 3.43  $\text{\AA}$  resolution of Zhejiang-H7N9 4 M FluPol in a self-stalled pre-termination state (Table 2 and Fig. S9). In this post-transfer structure, with the PPi reaction product visible, template U17 base pairs with A35 of the product cRNA at the +1 active site position, with a maximally taut connection of the template from the hook to the active site, as previously described<sup>12</sup>. This structure shows that even when all NTPs are present, under these conditions self-stalling prior to the hook can occur, as seen in the in vitro activity assays (Fig. 3D, E and S6A, B).

In all structures, the pS5 CTD is observed bound in sites 1A and 2A. Continuous density between the two sites, with 25 connecting residues coming from five CTD repeats, is observed for the pre-initiation state and the elongation structures (Fig. 4F), whereas for the replicase-like structure, the connectivity is lost between both sites. Despite extensive cryo-EM data analysis, huANP32A was never visualised nor the putative replicase-encapsidase dimer.

Altogether, these results show that in vitro, de novo RNA synthesis assays result in a heterogeneous mix of FluPol conformations



intermediate between the full replicase and transcriptase configurations, reflecting the flexibility of the peripheral PB2-C and PA-N domains (Fig. S7). Importantly, a replicase-like initiation complex (PB2-C(R)/ENDO(T)) is observed for the first time, characterised by the radically different position of the cap-binding domain compared to the transcriptase conformation (Fig. 4A, middle panel)<sup>16</sup>, which was never observed during extensive studies of cap-dependent transcription<sup>12</sup>. However, the elongation state stalled with UpNHpp, in which the

translocated 3' end of the template has reached the secondary binding site (Fig. 4E), is similar to the previously described pre-termination transcription state since PB2-C is not visible<sup>12</sup>. Finally, all structures have the CTD peptide mimic bound in both sites 1A and 2A, suggesting that both replication and transcription are consistent with CTD-binding (Fig. 4A, B and E). The observation of the intermediate state (denoted PB2-C(I)/ENDO(T)) in the context of de novo RNA synthesis lends support to the suggestion that it is an important intermediate in

**Fig. 4 | FluPol replication is consistent with CTD-binding to the FluPol.** **A** Left: Cartoon representation of Flu<sub>A</sub>Pol Zhejiang-H7N9-4M in the intermediate conformation (obtained in this study) as previously described<sup>44,45</sup>, bound to pS5 CTD. Middle: Flu<sub>A</sub>Pol Zhejiang-H7N9-4M in a replicase-like conformation bound to pS5 CTD. Right: Flu<sub>A</sub>Pol replicase conformation from A/duck/Fujian/01/2002(H5N1) (PDB: 6QPF)<sup>41</sup>. FluPols are aligned on the PB1 subunit. Flu<sub>A</sub>Pol Zhejiang-H7N9-4M PA-ENDO remains in a transcriptase conformation (PA-ENDO(T)). In Flu<sub>A</sub>Pol Fujian-H5N1 structure, it rotates and interacts with the PB2-NLS domain (PA-ENDO(R)). PA is coloured in green, PB1 in light grey, PB2-N in dark red, PB2-CBD in orange, PB2 mid-link in purple, PB2 627 in plum, PB2 NLS in salmon. The pS5 CTD is coloured in red, displayed as surface, with discontinuity between sites 1A/2A shown as a dotted line. PB2-627/NLS domains flexibility is highlighted as a dotted circle. RNAs are displayed as surfaces. The 5' vRNA end is coloured in pink, the 3' vRNA end in yellow. Flexible nts are represented as solid line. **B** Cartoon representation of Flu<sub>A</sub>Pol Zhejiang-H7N9-4M structure in the pre-initiation state mode A. Colour code

is identical to (A). PB2 C-terminal domains (PB2-C) are flexible, highlighted as a dotted circle. **C** Close-up view on the 3' vRNA end in Flu<sub>A</sub>Pol Zhejiang-H7N9-4M active site. The Coulomb potential map of the template is shown. 3'-U1 remains unseen in the map. The 3'-G3 is in the +1 active site, highlighted by a dotted line. The priming loop is coloured in red, the palm domain in orange with the catalytic aspartic acids displayed, coordinating a Mg<sup>2+</sup> ion, in green. The motif F is coloured in blue. **D** Schematic representation of the 3' vRNA end terminal nucleotides active site position, as seen in (C). **E** Cartoon representation of Flu<sub>A</sub>Pol Zhejiang-H7N9-4M in stalled elongation state. The colour code is identical to A. PB2-C domains are flexible. The cRNA de novo replication product is displayed as surface when visible or as a dotted line when flexible. The 3' vRNA end is bound to the secondary site. **F** Close-up view on the pS5 CTD interaction with Flu<sub>A</sub>Pol Zhejiang-H7N9-4M PA-C domain. FluPol residues interacting with the pS5 CTD are shown. Each CTD pS5 repeat is indicated.

the transition from transcriptase to replicase conformation<sup>45</sup>. Indeed, our structural results suggest that this transition might occur in the following sequence of steps: PB2-C(T)/ENDO(T), PB2-C(I)/ENDO(T), PB2-C(R)/ENDO(T), PB2-C(R)/ENDO(R), with the first and last being the full transcriptase and replicase, respectively. We suggest that the final step, ENDO(T) to ENDO(R) is more likely to happen after PB2-C(R) is established, which liberates the PB2-NLS domain so that it can interact with ENDO(R). Furthermore, we suggest that the full replicase conformation (including the otherwise flexible 627-domain) is only fully stabilised in the context of the replicase-encapsidase complex.

The pre-initiation state structures show that at least in vitro, and using highly purified recombinant FluPol, the preferred position of the 3' end of the vRNA template is with C2, rather than U1, at the -1 position. This leads to formation of pppGpC at the beginning of the product, consistent with the NGS results, or alternatively, would allow efficient priming by pppApG, if this dinucleotide was already available from another source. Internal initiation at position 2 of vRNA has been previously described<sup>20,47</sup>, but other authors report pppApG formation at position 1<sup>7</sup>. We note that initiation at position 2 of vRNA implies that the 5' end of the cRNA product lacks A1 and thus probably forms a less stable hook structure since the A1:A10 non-canonical base pair cannot form. This could explain why we do not observe in cryo-EM the cRNA hook bound in a putative encapsidase, even though there is excess apo-polymerase. This could further explain why the putative asymmetric FluPol<sub>(R)</sub>-FluPol<sub>(E)</sub> dimer is also not observed, since it seems likely that in vitro, for the Flu<sub>A</sub>Pol, the FluPol<sub>(R)</sub>-FluPol<sub>(E)</sub>-ANP32 complex is not so stable, unlike the case of Flu<sub>C</sub>Pol, where the FluPol<sub>(R)</sub>-FluPol<sub>(E)</sub> complex is stable even in the absence of ANP32<sup>13</sup>. These observations suggest that other cellular or viral factors and the RNP context might be required to recapitulate true terminal initiation of vRNA replication and further work is required to clarify this issue<sup>36</sup>.

### Restoration of CTD-binding interface rescues FluPol replication activity and enhances FluPol binding to huANP32A

Second-site mutations PA C489R and PB2 D253G were selected during passaging of the recombinant WSN virus with a PA K289A mutation (Fig. 2C) and conferred elevated FluPol-binding to the CTD as well as to huANP32A (Fig. 2F, grey and blue bars). PA C489 is in close proximity to the phosphoserine binding site in CTD-binding site 2A (Fig. 5A). The mutation PA C489R could therefore plausibly compensate for the loss of the positive charge of PA K289A and rescue the interaction with the CTD. Indeed, structural analysis by cryo-EM of Zhejiang-H7N9 WT FluPol in the symmetric dimeric form (Table 2 and Fig. S10) and bearing the double mutation PA K289A + C489R confirms that C489R points towards the phosphoserine binding site, although at a slightly greater distance than K289 (Fig. 5B, C). Despite this, the structure did not reveal CTD-binding to the Zhejiang-H7N9 PA K289A + C489R FluPol in this conformation. Therefore, we sought to analyse the

functional impact of the single PA C489R reversion on the viral phenotype, and in cellular assays, on FluPol activity and FluPol-binding to the CTD and huANP32A (Fig. 5D–I).

A recombinant WSN virus with the PA K289A + C489R mutations shows larger plaques compared to the PA K289A mutant virus, demonstrating that the PA C489R second-site mutation provides a significant growth advantage to PA K289A mutant virus (Fig. 5D and S11A). Indeed, WSN FluPol activity as measured in a vRNP reconstitution assay increases significantly when the PA C489R mutation is combined with PA K289A (Fig. 5E). This increase correlates with a restored mRNA:vRNA ratio (Fig. 5F and S11B) as well as an increased FluPol-binding to the CTD (Fig. 5G). Importantly, steady-state levels of the WT and mutant PA proteins are similar (Fig. S11C). These observations suggest that charge restoration in CTD-binding site 2A of the WSN FluPol in the transcriptase conformation rescues the interaction with the host RNAP II for 'cap-snatching' in cell-based assays and during live-virus infection. In cell-based assays, the Anhui-H7N9 FluPols with PA K289A and PA K289A + C489R show similar trends to those observed with the WSN FluPol counterparts, although with much less pronounced phenotypes (Fig. S11D). Therefore, we speculate that the fact that CTD-binding is not restored in vitro for Zhejiang-H7N9 PA K289A + C489R could be due to strain-specificities (Fig. S2) and/or to the presence of cellular factors that favour binding of the CTD to the PA K289A + C489R and are missing in the in vitro assay.

Beyond the rescue of FluPol transcription, cRNA accumulation levels increase significantly when PA C489R is combined with PA K289A (Fig. S11B), which goes together with a rescue of the imbalanced cRNA:vRNA ratio associated with the PA K289A mutant (Fig. 5H), and with increased huANP32A-binding (Fig. 5I). Taken together our observations suggest that the second-site mutation PA C489R has another functional impact, namely restoring FluPol replication by enhancing huANP32A-binding.

Similar results are obtained for the PA C453R revertant which has been shown to restore FluPol-binding to the CTD in site 1A by compensating the loss of positive charge of the PA R638A mutation (Fig. S12A)<sup>23,48</sup>. PA C453R rescues the attenuated viral plaque phenotype (Fig. S12B, C) as well as the reduced FluPol activity (Fig. S12D) associated with PA R638A, while it does not affect PA steady-state levels (Fig. S12E). Strikingly, PA C453R enhances huANP32A-binding when tested in combination with PA R638A (Fig. S12F), mimicking at CTD-binding site 1A the effects of PA C489R at site 2A.

The mechanism of rescue of ANP32A-binding by the PA C489R and PA C453R mutations remains unclear and could possibly involve the FluPol<sub>(R)</sub> and/or FluPol<sub>(E)</sub> moieties. The mutations could enhance binding of the pS5 CTD in site 1A or 2A in the FluPol<sub>(R)</sub>, leading to an enhanced priming of the replicase activity in conjunction with ANP32. They could as well increase binding of ANP32 through direct interactions at the FluPol<sub>(E)</sub>-ANP32 interface.

**Table 2 | Cryo-EM structures data collection, refinement and validation statistics (continued)**

Name of structure	A/H7N9 K289A/C489R revertant Apo-dimer with promoter	A/H7N9 self-stalled termination product complex with bound CTD
PDB ID	PDB ID 8POH	PDB ID 8R3K
EMDB ID	EMD-17792	EMD-18871
<b>Data collection and processing</b>		
Microscope	ThermoFisher Glacios	ThermoFisher Glacios
Voltage (kV)	200	200
Camera	Falcon IV/Selectris X	Falcon IV/Selectris X
Magnification	130k	130k
Nominal defocus range ( $\mu\text{m}$ )	-0.8/-2.0	-0.8/-2.0
Electron exposure ( $\text{e}^-/\text{\AA}^2$ )	40	40
Number of fractions processed (no.)	24	24
Pixel size ( $\text{\AA}$ )	0.907	0.878
Micrographs (no.)	5092	3509
<b>Refinement</b>		
Particles per class (no.)	278761	29072
Map resolution ( $\text{\AA}$ ), 0.143 FSC	3.26	3.43
Model resolution ( $\text{\AA}$ ), 0.5 FSC	3.38	3.40
Map sharpening B factor ( $\text{\AA}^2$ )	-140	-55
Map versus model cross-correlation (CCmask)	0.7456	0.8472
<b>Model composition</b>		
Non-hydrogen atoms	20729	14287
Protein residues	2479	1718
Nucleotide residues	38	41
Water	-	-
Ligands	1 Mg	2 Mg, PPI
<b>B factors (<math>\text{\AA}^2</math>)</b>		
Protein	105.84	55.88
Nucleotide	123.86	61.19
Ligand	33.48	44.02
Water	-	-
<b>R.M.S. deviations</b>		
Bond lengths ( $\text{\AA}$ )	0.003	0.003
Bond angles ( $^\circ$ )	0.562	0.507
<b>Validation</b>		
MolProbity score	1.84	1.44
All-atom clash score	6.79	6.44
Poor rotamers (%)	2.37	0.07
<b>Ramachandran plot</b>		
Favoured (%)	96.91	97.59
Allowed (%)	3.09	2.41
Outliers (%)	0.0	0.0

## Discussion

### A model for RNAP II CTD-anchored transcription and replication of the influenza virus genome

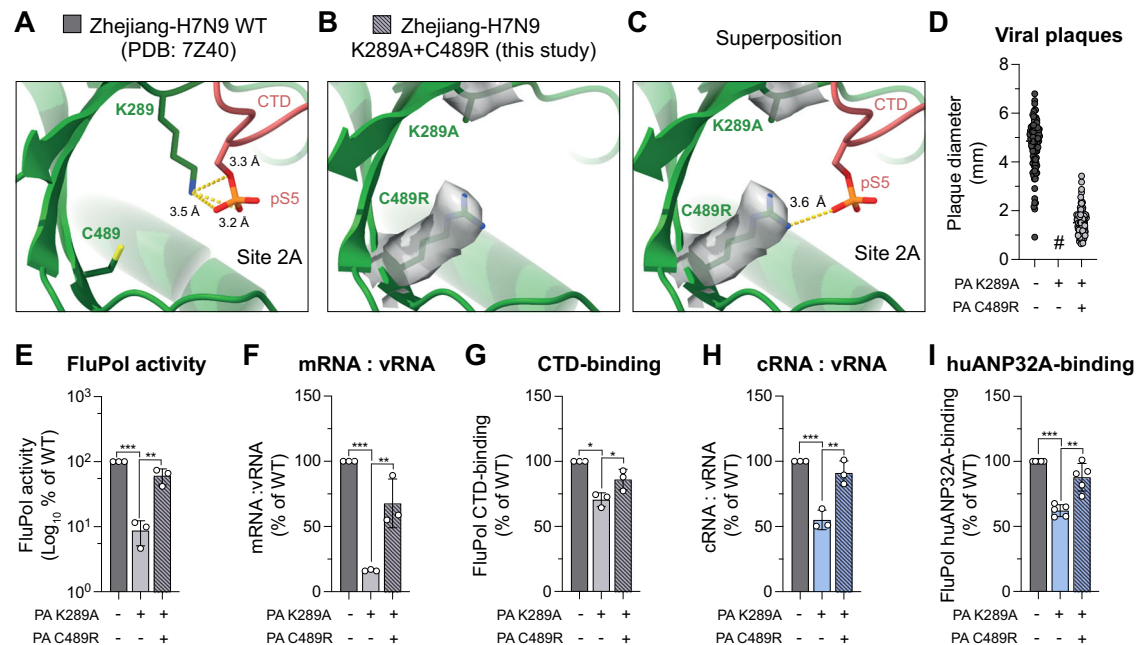
There is long-standing and extensive evidence that binding of FluPol to the pS5 CTD repeats of host RNAP II is essential for transcription of viral mRNAs<sup>26,49,50</sup>. Here, we show that CTD-binding not only stabilises

the transcriptase conformation of the influenza virus polymerase (FluPol<sub>(T)</sub>)<sup>22,29</sup>, but also enhances replication of the viral genome, in conjunction with the host protein ANP32A, which was recently described to bridge two FluPol moieties in an asymmetric replicase-encapsidase dimer (FluPol<sub>(R)</sub>-FluPol<sub>(E)</sub>)<sup>13</sup>. Our findings open new perspectives on the spatial coupling of viral transcription and replication and the coordinated balance between these two activities.

Early in infection, incoming vRNPs serve as a template for both primary mRNA transcription and the first round of cRNA synthesis<sup>51</sup>. Based on our findings, we propose a model in which incoming vRNPs switch from primary transcription to cRNA synthesis while remaining bound to the RNAP II CTD (Fig. 6A). It has been long understood that primary transcription requires FluPol<sub>(T)</sub> association with the RNAP II CTD to enable 'cap-snatching' for the priming of viral mRNA synthesis in the FluPol<sub>(T)</sub> conformation (Fig. 6B, C). Our structural and biochemical data reveal that a vRNA- and CTD-bound FluPol can adopt a FluPol<sub>(R)</sub> conformation and initiate vRNA to cRNA synthesis de novo, and that CTD-binding to the FluPol enhances cRNA synthesis. Importantly, during an ongoing infection the switch from a transcribing FluPol<sub>(T)</sub> to a replicating FluPol<sub>(R)</sub> can occur only after de novo synthesised FluPol and NP are imported into the nucleus (Fig. 6D) and an asymmetric FluPol<sub>(R)</sub>-ANP32A-FluPol<sub>(E)</sub> complex is assembled, ensuring encapsidation of the nascent viral RNA by the FluPol<sub>(E)</sub> moiety in conjunction with the NP. Our cell-based and in vitro assays show that the RNAP II CTD and ANP32 jointly enhance cRNA synthesis. The data support a model in which the replicating FluPol<sub>(R)</sub> remains associated to pS5 CTD repeats while being assembled into a FluPol<sub>(R)</sub>-ANP32A-FluPol<sub>(E)</sub> complex (Fig. 6E, F). The CTD is a low complexity disordered region and there is evidence that it drives phase separation and the formation of biomolecular condensates that concentrate RNAP II and transcription regulatory factors<sup>52</sup>. In infected cells, the RNAP II condensates may in addition concentrate the FluPol and its viral and cellular partners, allowing CTD-bound FluPols to switch efficiently between transcription or replication depending on the relative availability of, e.g. 5' capped RNAs, NP and/or apo FluPol (Fig. 6G), and perform both activities in a coordinated manner.

Although a structure for the asymmetric FluPol<sub>(R)</sub>-ANP32-FluPol<sub>(E)</sub> replicase-encapsidase dimer has so far been obtained only for vRNA-bound FluPol<sub>(R)</sub>, it is generally thought that the same type of asymmetric dimer is formed by a cRNA-bound FluPol<sub>(R)</sub><sup>18</sup>. Therefore, the model proposed for incoming vRNPs (Fig. 6) may possibly be extended to neo-synthesised progeny vRNPs or cRNPs, consistent with vRNA synthesis being detected mostly in the chromatin-associated fractions<sup>33,54</sup>. In our model, RNPs could, therefore, remain anchored to the same CTD repeats and undergo several rounds of transcription (vRNPs) or genome replication (vRNPs or cRNPs). Progeny RNPs might in turn associate to the same or a close CTD, thereby generating a viral factory for RNA synthesis. Further studies will be needed to investigate the precise relationships between replicating vRNP, cRNPs and host RNAP II, given that cRNPs do not perform 'cap-snatching' and require transactivation by an apo FluPol, and could therefore be subjected to different regulatory networks<sup>21,41</sup>.

Our model that the RNAP II CTD provides a recruitment platform for the different types of monomeric and multimeric polymerases raises the question of how the CTD is coordinating transcription versus replication in the context of chromatin in infected cells. Influenza infection was shown to strongly dysregulate RNAP II activities, as evidenced by a decreased RNAP II occupancy in gene bodies downstream of the transcription start site<sup>49,55</sup>, a widespread RNAP II termination defect and subsequent transcriptional read-through<sup>55-57</sup>, as well as a proteolytic degradation of the RNAP II at late stages of infection<sup>58,59</sup>. FluPol-binding to the CTD is dependent on S5 phosphorylation<sup>22,23,26</sup>, and our data show that pS5 CTD-binding is compatible with both transcription and replication activity of the FluPol. Consistently, a large proportion of vRNPs were found to colocalize with pS5 RNAP II at



**Fig. 5 | Restoration of CTD-binding interface rescues FluPol replication activity and enhances FluPol binding to huANP32A.** **A** Cartoon representation of the pS5 CTD-bound to FluPol Zhejiang-H7N9 in site 2A (PDB: 7Z4O)<sup>22</sup>. PA subunit is coloured in green, pS5 CTD in red. PA K289-C489 residues are displayed. Putative hydrogen bonds are drawn as yellow dashed lines. Distances are indicated. **B** Cartoon representation of Flu<sub>A</sub>Pol Zhejiang-H7N9 PA K289A + C489R (obtained in this study). The Coulomb potential map of PA/K289A-C489R is shown. **C** Superposition of the pS5 CTD, extracted from the structure shown in (A), with the Flu<sub>A</sub>Pol Zhejiang-H7N9 PA K289A + C489R structure shown in (B). The putative hydrogen bond between PA C489R and pS5 is shown. Distance is indicated. **D–I** Phenotypes associated with the WSN FluPol PA K289A primary mutation and PA C489R second-site mutation. **D** Plaque phenotype of recombinant WSN mutant viruses produced by reverse genetics ( $n = 2$ ), analyzed as in Fig. 2E (see Fig. S11A) (#) pinhead-sized plaques. **E** WSN FluPol activity was measured by vRNP reconstitution

in HEK-293T cells, using a model Fluc-vRNA. Luminescence signals are represented as a percentage of PA WT. **F** WSN vRNPs were reconstituted in HEK-293T cells using the NA vRNA segment. The steady-state levels of NA mRNA and vRNA were quantified by strand-specific RT-qPCR<sup>35</sup>, normalised to GAPDH by the  $2^{-\Delta\Delta CT}$  method<sup>21</sup> and are presented as ratios of mRNA to vRNA levels relative to PA WT (RNA levels are shown in Fig. S11B). **G** WSN FluPol binding to the CTD was assessed using a split-luciferase-based complementation assay. Luminescence signals are represented as a percentage of PA WT. **H** Accumulation levels of cRNA and vRNA in a vRNP reconstitution assay were determined by strand-specific RT-qPCR<sup>35</sup> as in (F). Ratios of cRNA to vRNA levels relative to PA WT are shown (RNA levels are shown in Fig. S11B). **I** huANP32A-binding to WSN FluPol was determined as in (G). (mean  $\pm$  SD,  $n = 3, 3, 3, 3, 5$ , \* $p < 0.033$ , \*\* $p < 0.002$ , \*\*\* $p < 0.001$ , one-way ANOVA; Dunnett's multiple comparisons test). Source data are provided as a Source data file.

6 hours post-infection, a stage when viral replication is becoming predominant over transcription<sup>55</sup>. S5 phosphorylation is considered as the hallmark modification of RNAP II initiation<sup>28</sup>. Because there is a short time window for the influenza polymerase to gain access to capped nascent RNAP II transcripts before the high affinity nuclear cap-binding complex binds the capped RNA, 'cap-snatching' is thought to occur in association with promoter-proximal paused RNAP II<sup>22–24</sup>. However, recent studies show that pS5 CTD can also occur during RNAP II elongation<sup>60</sup>, and it remains unknown to what extent the CTD phosphorylation dynamics is altered upon influenza infection. Therefore, in principle, pS5 CTD-bound FluPol could perform viral replication in association with initiating and/or elongating RNAP II. Investigating whether or not transcribing and replicating FluPol complexes are associated to the same phosphoforms of the CTD, in terms of complex phosphorylation patterns of residues S<sub>2</sub>, S<sub>5</sub>, S<sub>7</sub> but also Y<sub>1</sub> and P<sub>4</sub><sup>61</sup>, and/or localised to the same RNAP II condensates, could help address this question. Moreover, further investigation are required to gain a better understanding of how FluPol transcription and replication complexes interact with the full-length CTD and other RNAP II domains, and to what extent the CTD and/or host factors recruited by the CTD are controlling the balance between transcription and replication. Whether the viral-induced degradation of RNAP II observed at late time-points during influenza infection<sup>58,59</sup> could favour the release of neo-synthesised vRNPs from the chromatin and their nuclear export also warrants investigation. Fully addressing these questions will require to overcome technical challenges such as structural analyses on the FluPol in the context of vRNPs and cRNPs, or high-resolution

imaging in live-infected cells to visualise the subnuclear localisation of FluPol transcription and replication.

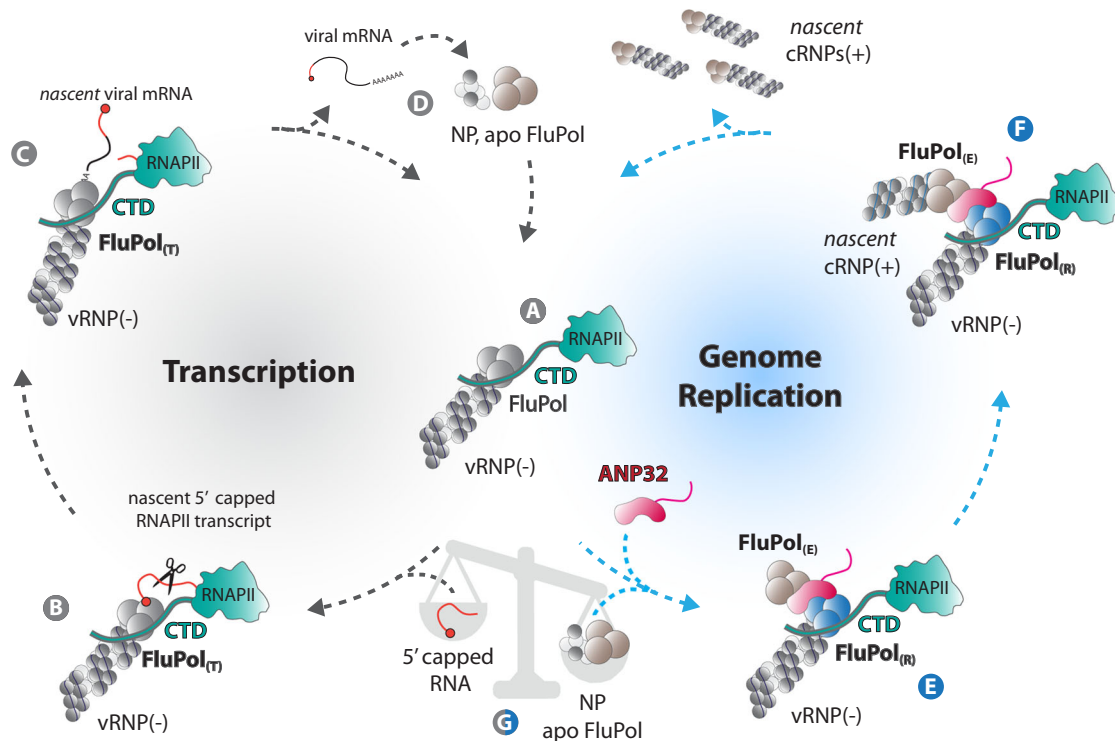
## Methods

### Cells

HEK-293T cells (ATCC CRL-3216 and ATCC CRL-11268) were grown in complete Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% foetal bovine serum (FBS) and 1% penicillin-streptomycin (Gibco). MDCK cells (provided by the National Influenza Center, Paris, France) were grown in Modified Eagle's medium (MEM, Gibco) supplemented with 5% FBS and 1% penicillin-streptomycin.

### Plasmids

A/WSN/33 (WSN) ORFs were encoded in pcDNA3.1-PB2, -PB1, -PA and pCI-NP plasmids<sup>22,23</sup>. The A/Anhui/1/2013 (Anhui-H7N9) pCAGGS-PB2, -PB1, -PA and NP plasmids were kindly provided by R. Fouchier (Erasmus Medical Center, Netherlands)<sup>62</sup>. The B/Memphis/13/2003 ORFs were encoded in pcDNA3.1-PB2 and PA plasmids<sup>9</sup>. The WSN reverse genetics plasmids were kindly provided by G. Brownlee (Sir William Dunn School of Pathology, Oxford, UK)<sup>63</sup>. For vRNP reconstitution assays, a pPoll-Firefly plasmid encoding the Firefly luciferase sequence in negative polarity flanked by the 5' and 3' non-coding regions of the IAV NS segment was used and the pTK-Renilla plasmid (Promega) was used as an internal control. For RNA quantifications in vRNP reconstitution assays by strand-specific qRT-PCR a pPR7-WSN-NA reverse genetic plasmid was used<sup>64</sup>. For the WSN and Anhui-H7N9 pCI-PB1-G1 and pCI-PB1-G2 plasmids used for *G. princeps* split-luciferase-based



**Fig. 6 | A model for RNAP II CTD-anchored transcription and replication of the influenza virus genome.** A model for RNAP II CTD-anchored transcription and replication of the influenza virus genome. Upon influenza virus infection, incoming vRNPs are imported into the nucleus and bind to the host RNAP II CTD through bipartite interaction sites on the FluPol (A). This intimate association allows FluPol, in the transcriptase conformation (FluPol<sub>T</sub>), to cleave short capped oligomers derived from nascent RNAP II transcripts in a process referred to as ‘cap-snatching’ to initiate primary transcription of viral mRNAs (B). Polyadenylation is achieved by a non-canonical mechanism involving stuttering of the viral polymerase at a polyadenylation signal (C). The 5' and 3' vRNA extremities always remain bound to the polymerase which allows efficient recycling from the termination to the initiation state (C–A). Upon translation of viral mRNAs, de novo synthesised FluPols in an apo state (not viral RNA-bound) and NPs are imported into the nucleus (D). The apo

FluPol, in conjunction with the host factor ANP32, associates with the parental CTD-associated FluPol<sub>T</sub> and triggers its conformational transition into a replicating FluPol<sub>R</sub>, to form an asymmetric FluPol<sub>R</sub>-FluPol<sub>E</sub> dimer (E) where FluPol<sub>E</sub> is encapsidating the newly synthesised cRNA in conjunction with NP (F). The FluPol<sub>R</sub>-ANP32-FluPol<sub>E</sub> replication complex remains associated to the RNAP II through direct binding of the CTD to FluPol<sub>R</sub>. Anchoring of the parental vRNP to the CTD allows it to engage into successive cycles of either viral genome replication or mRNA transcription, depending on the availability of NP, apo FluPol and/or nascent capped oligomers derived from actively transcribing RNAP II (G). Such switching between both activities allows efficient adaptation to waving levels of de novo synthesised vRNP components in the nucleus of an infected cell and is key to ensure a correct balance between genome replication and mRNA transcription.

protein-protein complementation assays, the PB1 ORF was tagged C-terminally with the *G. princeps* fragments<sup>65</sup>. The Anhui-H7N9 PA open-reading frame (ORF) was subcloned into the pCI vector and the PA-X ORF was deleted when indicated by introducing silent mutations at the site of ribosomal frameshifting<sup>66</sup>. For CTD-binding assays, a C-terminal stretch of RPB1 (108 amino acids) in conjunction with 14 CTD repeats (termed CTD) were tagged C-terminally with the *G. princeps* fragments<sup>34</sup>. For ANP32A-binding assays, ANP32A was tagged C-terminally with the *G. princeps* fragments<sup>67</sup>. The pCI-mCherry-CTD-WT and mCherry-CTD-SSA plasmids were generated by replacing the G2 sequence in pCI-G2-CTD constructs<sup>22</sup>. ANP32A encoding the amino acids 1-249 (full-length), 1-199 and 1-149 were subcloned into the pCI vector with a C-terminal V5-tag and SV40 nuclear localisation signals (NLS). DDX5 and RED ORFs were encoded in the pCI vector and tagged N-terminally with the *G. princeps* fragments<sup>38,68</sup>. All mutations were introduced by an adapted QuikChange site-directed mutagenesis (Agilent Technologies) protocol<sup>69</sup>. The ORFs were verified by Sanger sequencing. Plasmid sequences are provided as a source data file and are available with annotations at [<https://doi.org/10.5281/zenodo.10462746>].

**Generation of HEK-293T ANP32A and ANP32B knockout cells**  
CRISPR-Cas9-mediated HEK-293T (ATCC CRL-11268) ANP32A and ANP32B knockout cells (HEK-293T ANP32AB KO) and a control cell line

(HEK-293T CTRL) were generated by lentiviral transduction. Guide RNAs (gRNAs) targeting ANP32A (5'-ACCGTCAGGTGAAAGAAGCTTGTCC-3' and 5'-AAACGGACAAGTTCTTTCACCTGA-3'), ANP32B (5'-ACCGGAGCTGAGGAACCGGACCC-3' and 5'-AAACGGGTCCGGTTCCTCAGCTCC-3') and a non-targeting control (5'-ACCGGTATTACTGATATGGTGGG-3' and 5'-AAACCCCAATATCAGTAATAC-3') were annealed and cloned into the BsmBI site of pSicoR-CRISPR-PuroR (a kind gift from Robert Jan Lebbink<sup>70</sup>). Replication-incompetent lentiviral particles were generated by transient transfection of HEK-293T cells with the pSicoR-CRISPR-PuroR plasmid harbouring gRNAs targeting either ANP32A, ANP32B or a non-targeting control gRNA as well as with the packaging plasmids pMD2.G and psPAX2 (gifts from Didier Trono, Addgene plasmids #12259 and #12260, respectively). Supernatants were harvested 72 hpt, centrifuged to remove cellular debris and passed through 0.45 μm sterile filters. Fresh HEK-293T cells were plated in 6 well plates pre-coated with poly-D-lysine and were infected with 3 ml of lentiviral supernatant by constant centrifugation at 700 g for 40 min in the presence of 10 μg/ml Polybrene (Sigma-Aldrich). At 72 hpi, the medium was exchanged to selection medium containing 1 μg/ml puromycin (PAA Laboratories) for one week. Surviving cells were then cloned by limiting dilution in 96-well plates. Individual cell clones were grown and ANP32A and ANP32B protein levels in cellular lysates of individual clones was assessed by immunoblotting. To generate HEK-293T cells with a double KO for ANP32A and ANP32B, two rounds of targeting were performed.

### Protein complementation assays

HEK-293T cells were seeded in 96-well white plates (Greiner Bio-One) the day before transfection using polyethylenimine (PEI-max, #24765-1 Polysciences Inc). Cells were lysed 20–24 hpt in Renilla lysis buffer (Promega) for 45 min at room temperature (RT) under steady shaking. *G. princeps* luciferase enzymatic activity due to luciferase reconstitution was measured on a Centro XS LB960 microplate luminometer (Berthold Technologies, MikroWin Version 4.41) using a reading time of 10 s after injection of 50  $\mu$ l Renilla luciferase reagent (Promega). Mean relative light units (RLUs) of technical triplicates of *G. princeps* split-luciferase-based protein-protein complementation assays are represented. In the graphs each dot represents an independently performed biological replicate while at least three experiments were performed in each case. For FluPol-binding assays cells were always co-transfected with plasmids encoding the viral polymerase subunits (PB2, PB1, PA) tagged to one part of the *G. princeps* luciferase and the target protein tagged to the other part. For FluPol oligomerisation PB2, PA, PB1-G1 and PB1-G2 were co-transfected and when indicated an expression plasmid for mCherry, mCherry-CTD-WT or mCherry-CTD-SSA was added while the total amount of transfected plasmid was adjusted using an empty control plasmid. Plasmid combinations, orientations of tags as well as plasmid quantities used for transfections in a given experiment are available as a source data file.

### vRNP reconstitution assays

HEK-293T cells were seeded in 96-well white plates (Greiner Bio-One) the day before transfection. Cells were co-transfected with plasmids encoding the vRNP protein components (PB2, PB1, PA, NP), a pPoll-Firefly plasmid encoding a negative-sense viral-like RNA expressing the Firefly luciferase and the pTK-Renilla plasmid (Promega) as an internal control. For FluPol trans-complementation assays, equal amounts of expression plasmids expressing the trans-complementing PA mutants were co-transfected. The fold-change shown for FluPol trans-complementation assays represents the increase above the integrated background of the respective FluPol mutants alone. For FluPol activity rescue experiments in HEK-293T ANP32AB KO cells, expression plasmids encoding FL or truncated versions of ANP32A were co-transfected. Firefly luciferase activity due to viral polymerase activity and Renilla luciferase activity due to RNAP II activity were measured using the Dual-Glo Luciferase Assay system (Promega) according to the manufacturer's instructions. The graphs represent viral polymerase activity normalised to RNAP II activity. Luciferase activities were measured in technical duplicates at 20–24 hpt or at 48 hpt when indicated. In the graphs each dot represents an independently performed biological replicate while at least three experiments were performed in each case. For the quantification of mRNA, cRNA and vRNA levels, HEK-293T cells were seeded in 12-well plates and transfected with plasmids encoding the vRNP protein components (PB2, PB1, PA, NP) and a low amount of WSN-NA RNA expressing plasmid (5 ng/well). Total RNA was isolated at 24 hpt with RNeasy Mini columns according to the manufacturer's instructions (RNeasy Kits, Qiagen) and strand-specific RT-qPCRs were performed<sup>35</sup>. Briefly, total RNA was reverse transcribed using primers specific for NA mRNA, cRNA, vRNA and, when indicated, the cellular glyceraldehyde 3-phosphate dehydrogenase (GADPH) with SuperScript™ III Reverse Transcriptase (Invitrogen) and quantified using SYBR-Green (Roche) with the Light-Cycler 480 system® (Roche, Light Cycler Software 1.5.0.39). RNA levels were normalised to GAPDH when indicated and analysed using the  $2^{-\Delta\Delta CT}$  as described before<sup>71</sup>. Plasmid combinations and plasmid quantities used for transfections in a given experiment are available as a source data file.

### Antibodies and immunoblots

Total cell lysates were prepared in RIPA cell lysis buffer<sup>72</sup>. Proteins were separated by SDS-PAGE using NuPAGE™ 4–12% Bis-Tris gels

(Invitrogen) and transferred to nitrocellulose membranes which were incubated with primary antibodies directed against pS5 CTD (clone 3E8, Active Motif, 1:1000), mCherry (26765-1-AP, Proteintech, 1:000), V5 (SV5-Pk1, Thermo Fisher, 1:5000), Tubulin (B-5-1-2, Sigma-Aldrich, 1:10000), PA<sup>73</sup> (1:2500), PB2 (GTX125925, GeneTex, 1:5000), ANP32A (AV40203, Sigma-Aldrich, 1:2500), ANP32B (EPR14588, AbCam, 1:2500) and subsequently with HRP-tagged secondary antibodies (Jackson Immunoresearch). Membranes were revealed with the ECL2 substrate according to the manufacturer's instructions (Pierce) and chemiluminescence signals were acquired using the ChemiDoc imaging system (Bio-Rad) and analysed with ImageLab (Bio-Rad). Uncropped gels are provided as a source data file.

### Production and characterisation of recombinant viruses

The recombinant PA mutant influenza viruses (PA K289A, R454A, K635A, R638A) used for serial cell culture passaging were generated by reverse genetics followed by plaque purification under agarose overlay and one round of viral amplification (P1)<sup>23</sup>. For the serial passaging, three dilutions of the supernatant collected at passage N (1:10, 1:100 and 1:1000) were used to infect MDCK cells at passage N+1. The supernatant corresponding to one of these dilutions (the lowest dilution that resulted in the infectious titre greater than or equal to titre at passage N) was selected to infect MDCK cells at passage N+2. These conditions resulted in MOIs in the 0.001 or 0.0001 range, depending on the passages. Upon infection, MDCK cells were incubated for 3 days at 37 °C in DMEM containing TPCK-Trypsin (Sigma) at a final concentration of 1  $\mu$ g/mL. Viral supernatants were titrated on MDCK cells in a plaque assay<sup>74</sup>. Viral RNA was extracted from 140  $\mu$ l of viral stocks using the QIAamp Viral RNA Mini kit (Qiagen) according to the manufacturer's instructions. For next-generation sequencing of the full viral genome reverse transcription and amplification of the eight genomic segments were performed using the RT-qPCR protocol adapted by the National Influenza Center (Institut Pasteur)<sup>75</sup>. Next-generation sequencing was performed by the 2P2M facility at Institut Pasteur using the Nextera XT DNA Library Preparation kit (Illumina), the NextSeq 500 sequencing systems (Illumina) and the CLC Genomics Workbench 9 software (Qiagen) for analysis. Recombinant viruses with a selected subset of observed second-site mutations were generated by reverse genetics with an adapted protocol<sup>76</sup>. In brief, HEK-293T cells were seeded in 6-well plates the day before transfection with the WSN reverse genetics plasmids<sup>63</sup> in Opti-MEM (Gibco) using FuGene®6 (Promega) according to the manufacturer's instructions. The next day, MDCK were added to the wells in 500  $\mu$ l Opti-MEM containing 0.5  $\mu$ g/mL TPCK trypsin (Sigma). The following days, 500  $\mu$ l Opti-MEM containing 1 and 2  $\mu$ g/mL TPCK trypsin, respectively, were added. One day later, the supernatants were harvested, centrifuged and stored at –80 °C. The reverse genetics supernatants were titrated on MDCK cells and plaque diameters were measured upon staining with crystal violet using Fiji (ImageJ2 Version 2.3.0/1.3q)<sup>77</sup>.

### Influenza virus polymerase Zhejiang-H7N9 (WT, 4 M, PA K289A + C489R)

The pFastBac Dual vector encoding for the A/Zhejiang/DTID-ZJU01/2013 (H7N9) polymerase heterotrimer subunits, PA (Uniprot: M9TI86), PB1 (Uniprot: M9TLW3), and PB2 (Uniprot: X5F427) was used as a starting point (Zhejiang-H7N9 WT)<sup>43</sup>. The mutations PA E349K, PA R490I, PB1 K577G and PB2 G74R were introduced by a combination of PCRs and Gibson assembly (Zhejiang-H7N9 4 M). A similar approach has been applied to clone the FluPol Zhejiang-H7N9 PA K289A + C489R mutant. Plasmid sequences were confirmed by sanger sequencing for each polymerase subunit. All primer and plasmid sequences are available upon request.

FluPols Zhejiang-H7N9 (WT/4M/PA K289A + C489R) were produced using the baculovirus expression system in *Trichoplusia ni* High five cells. For large-scale expression, cells at 0.8-1E06 cells/mL

concentration were infected by adding 1% of virus. Expression was stopped 72 to 96 h after the day of proliferation arrest and cells were harvested by centrifugation (1,000 g, 20 min at 4 °C).

Cells were disrupted by sonication for 4 min (5 s ON, 20 s OFF, 40% amplitude) on ice in lysis buffer (50 mM HEPES pH 8, 500 mM NaCl, 0.5 mM TCEP, 10% glycerol) with cOmplete EDTA-free Protease Inhibitor Cocktail (Roche). After lysate centrifugation at 48,384 g during 45 min at 4 °C, ammonium sulphate was added to the supernatant at 0.5 g/mL final concentration. The recombinant protein was then collected by centrifugation (45 min, 4 °C at 70,000 g), re-suspended in the lysis buffer, and the procedure was repeated another time. FluPol Zhejiang-H7N9 was then purified using His60 NiNTA Superflow resin (Takara Bio) from the soluble fraction. Bound proteins were subjected to two sequential washes using (i) the lysis buffer supplemented by 1 M NaCl and (ii) the lysis buffer supplemented by 50 mM imidazole. Remaining bound proteins were eluted using the lysis buffer supplemented by 500 mM imidazole. Fractions with FluPol Zhejiang-H7N9 were pooled and directly subjected to a strep-tactin affinity purification (IBA, Superflow). Bound proteins were eluted using the lysis buffer supplemented by 2.5 mM d-desthiobiotin and protein-containing fractions were pooled and diluted with an equal volume of buffer (50 mM HEPES pH 8, 0.5 mM TCEP, 10% glycerol) before loading on a third affinity column HiTrap Heparin HP 5 mL (Cytiva). A continuous gradient of lysis buffer supplemented with 1 M NaCl was applied over 15 CV, and FluPol Zhejiang-H7N9 was eluted as single species at -800 mM NaCl. Pure and nucleic acid free FluPol Zhejiang-H7N9 was dialysed overnight in a final buffer (50 mM HEPES pH 8, 500 mM NaCl, 2 mM TCEP, 5% glycerol), concentrated with Amicon Ultra-15 (50 kDa cutoff), flash-frozen and stored at -80 °C for further use.

#### Influenza virus nucleoprotein Anhui-H7N9 monomeric mutant (R416A)

The influenza virus NP Anhui-H7N9 gene was amplified from a pCAGGS plasmid (gift from Ron Fouchier)<sup>62</sup> and introduced by Gibson assembly in a pLIB vector with an N-terminal double strep-tag followed by a human rhinovirus (HRV) 3C protease cleavage site. The NP R416A mutation was introduced by a combination of PCRs and Gibson assembly. Introduction of the desired mutation was confirmed by sanger sequencing. Monomeric influenza virus Anhui-H7N9 NP R416A was produced using the baculovirus expression system in Trichoplusia ni High five cells. For large-scale expression, cells at 0.8-1E06 cells/mL concentration were infected by adding 1% of virus. Expression was stopped 72-96 h after the day of proliferation arrest and cells were harvested by centrifugation (1,000 g, 20 min at 4 °C).

Cells were disrupted by sonication for 4 min (5 s ON, 15 s OFF, 40% amplitude) on ice in lysis buffer (50 mM HEPES pH 8, 500 mM NaCl, 2 mM TCEP, 5% glycerol) with cOmplete EDTA-free Protease Inhibitor Cocktail (Roche). After lysate centrifugation at 48,384 g for 45 min at 4 °C, the soluble fraction was loaded on a StrepTrap HP 5 mL (Cytiva). Bound proteins were eluted using the lysis buffer supplemented by 2.5 mM d-desthiobiotin. Protein-containing fractions were pooled and diluted with an equal volume of buffer (50 mM HEPES pH 8, 2 mM TCEP, 5% glycerol) before loading on an affinity column HiTrap Heparin HP 5 mL (Cytiva). A continuous gradient of lysis buffer supplemented with 1 M NaCl was applied over 15 CV, and Anhui-H7N9 NP R416A was eluted as single species at -500 mM NaCl without nucleic acids ( $A_{260/280}$ : 0.6). Pure and nucleic acid free Anhui-H7N9 NP R416A was dialysed overnight in a final buffer (50 mM HEPES pH 8, 300 mM NaCl, 1 mM TCEP, 5% glycerol) together with N-terminal his-tagged HRV 3C protease (ratio 1:5 w/w). Tag-cleaved Anhui-H7N9 NP R416A was subjected to a last Ni-sepharose affinity to remove the HRV 3C protease, further concentrated with Amicon Ultra-15 (10 kDa cutoff), flash-frozen and stored at -80 °C for later use.

#### Acidic nuclear phosphoprotein 32A

Human and chicken ANP32A genes (GeneScript) were introduced in a pETM11 vector with an N-terminal 6xHis-tag followed by a Tobacco Etch Virus (TEV) protease cleavage site. ANP32A constructs were expressed in BL21(DE3) E.coli cells. Expression was induced when absorbance reached 0.6, with 1 mM IPTG, incubated for 4 h at 37 °C. Cells were harvested by centrifugation (1,000 g, 20 min at 4 °C).

Cells were disrupted by sonication for 5 min (5 s ON, 15 s OFF, 50% amplitude) on ice in lysis buffer (50 mM HEPES pH 8, 150 mM NaCl, 5 mM beta-mercaptoethanol (BME)) with cOmplete EDTA-free Protease Inhibitor Cocktail (Roche). After lysate centrifugation at 48,384 g for 45 min at 4 °C, the soluble fraction was loaded on a HisTrap HP 5 mL column (Cytiva). Bound proteins were subjected to a wash step using the lysis buffer supplemented by 50 mM imidazole. Remaining bound proteins were eluted using the lysis buffer supplemented by 500 mM imidazole. Fractions containing ANP32A were dialysed overnight in the lysis buffer (50 mM HEPES pH 8, 150 mM NaCl, 5 mM BME) together with N-terminal his-tagged TEV protease (ratio 1:5 w/w). Tag-cleaved ANP32A protein was subjected to a Ni-sepharose affinity column to remove the TEV protease, further concentrated with Amicon Ultra-15 (3 kDa cutoff) and subjected to a Size-Exclusion Chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva) in a final buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 2 mM TCEP. Fractions containing exclusively ANP32A were concentrated with Amicon Ultra-15 (3 kDa cutoff), flash-frozen and stored at -80 °C for later use.

Human ANP32A truncation constructs (1-199 and 144-249) were generated, expressed and protein purified as previously described ref. 78. The human ANP32A 1-149 construct was a gift from Cynthia Wolberger (Addgene plasmid #67241<sup>79</sup>) and was expressed and protein purified as previously described<sup>78</sup>.

#### De novo FluPol replication activity

Synthetic vRNA loop ('v51\_mut\_S') (5'-pAGU AGA AAC AAG GGU GUA UUU UCC CCU CUU UUU GUU UCC CCU GCU UUU GCU -3') (IDT) was used for all in vitro replication activity assays. For all de novo replication activity assays, 2.4 μM FluPol Zhejiang-H7N9 (WT or 4 M) were mixed with (i) 0.8 μM v51\_mut\_S and/or (ii) 8 μM ANP32A and/or (iii) 16 μM pS5 CTD(6mer) (respective molar ratio: 3 FluPols: 1 template: 10 ANP32: 20 pS5 CTD). Reactions were launched at 30 °C for 4 h by adding ATP/GTP/CTP/UTP (AGCU) or only ATP/GTP/CTP (AGC), α-32P ATP (PerkinElmer) and MgCl<sub>2</sub> in a final assay buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 2 mM TCEP, 100 μM/NTP, 1 mM MgCl<sub>2</sub>, 0.05 μCi/μl α-32P ATP. Reactions were stopped by adding 2X RNA loading dye, heating 5 min at 95 °C and immediately loaded on a 20% TBE-7M urea-polyacrylamide gel. Each gel was exposed on a storage phosphor screen and read with an Amersham Typhoon scanner (Cytiva). For each gel the decade markers system (Ambion) was used.

#### Next generation sequencing of in vitro FluPol replication products

To confirm the identity of the replication products (full-length cRNA product and stalled cRNA product) and address the discrepancy between expected size and urea-PAGE migration, sequencing of the total reaction product was performed from a de novo FluPol replication assay. 2.4 μM FluPol Zhejiang-H7N9 4 M were mixed with (i) 0.8 μM v51\_mut\_S, (ii) 8 μM ANP32A and (iii) 16 μM pS5 CTD(6mer) (respective molar ratio: 3 FluPols: 1 template: 10 ANP32A: 20 pS5 CTD). Reaction was launched at 30 °C for 4 h by adding ATP/GTP/CTP/UTP and MgCl<sub>2</sub> in a final assay buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 2 mM TCEP, 100 μM/NTP, 1 mM MgCl<sub>2</sub>. After completion of the reaction, proteins were removed using the Monarch RNA Cleanup Kit (NEB) and 5'-mono-phosphorylated v51\_mut\_S templates were specifically digested with a terminator 5'-phosphate-dependent exonuclease (Biosearch technologies) and subjected again to the



Monarch RNA Cleanup Kit (NEB). Total remaining RNAs were used for next-generation sequencing.

Sample concentration and fragment size distribution were checked with the RNA Pico assay (Bioanalyzer, Agilent). 6 ng of sample was treated with 2 units T4 Polynucleotide Kinase (NEB) and incubated at 37 °C for 30 min, followed by heat inactivation at 65 °C for 20 min. The treated samples were then taken into library preparation following the Takara SMARTer small RNA sequencing kit according to the manufacturer's instructions with 13 cycles of PCR. The final library fragment size distribution was checked with the High Sensitivity Bioanalyzer assay. 8 pM of library was sequenced on a MiSeq to generate 76 bp single-end reads.

Resulting total reads (13,073,371) were trimmed using Cutadapt v2.3<sup>80</sup> with settings recommended by the library preparation kit (-u 3 -a 'AAAAAAAAA') omitting the filter for only reads longer than 20 bases (10,991,864). Trimmed reads were subsequently used as an input for pattern lookup. All reads exactly starting by these motifs have been kept for further analysis: AGCAAAGCA/GCAAAGCA/CAAAGCA/AAAAGCA (7,776,959 reads). Each read starting exactly by these motifs 5'-AGCAAAGCAGGGG/5'-GCAAAGCAGGGG/5'-CAAAGCAGGGG/5'-AAAAGCAGGGG were counted and plotted (Fig. 3H). Each read matching the exact sequence of a full-length replication product starting by 5'-AGC AAA...ACU-3' (51-mer), 5'-GC AAA... ACU-3' (50-mer), 5'-C AAA... ACU-3' (49-mer) or 5'-\_\_ AAA... ACU-3' (48-mer) were counted and plotted as percentage (Fig. 3I).

## Electron microscopy

**Sample 1.** De novo pre-initiation and stalled elongation states were trapped by mixing 2.4 μM FluPol Zhejiang-H7N9 4 M with (i) 0.8 μM v51\_mut\_S, (ii) 8 μM chANP32A, (iii) 16 μM pS5 CTD(6mer) and (iv) 4 μM Anhui-H7N9 NP R416A in a final buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 2 mM TCEP, 1 mM MgCl<sub>2</sub>, ATP/GTP/CTP at 100 μM/NTP, and 100 μM of non-hydrolysable UpNHpp (Jena Bioscience).

De novo reaction mix was incubated for 4 h at 30 °C. The sample was centrifuged for 5 min at 11,000 g and kept at 4 °C before proceeding to grids freezing.

For grids preparation, 1.5 μl of sample was applied on each sides of plasma cleaned (Fischione 1070 Plasma Cleaner: 1 min 10 s, 90% oxygen, 10% argon) grids (UltraAufoil 1.2/1.3, Au 300). Excess solution was blotted for 3–5 s, blot force 0, 100% humidity, at 4 °C, with a Vitrobot Mark IV (ThermoFisher) before plunge-freezing in liquid ethane.

Automated data collection for sample 1 was performed on a TEM Titan Krios G3 (Thermo Fisher) operated at 300 kV equipped with a K3 (Gatan) direct electron detector camera and a BioQuantum energy filter, using EPU. Coma and astigmatism correction were performed on a carbon grid. Micrographs were recorded in counting mode at a ×105,000 magnification giving a pixel size of 0.84 Å with defocus ranging from -0.8 to -2.0 μm. Gain-normalised movies of 40 frames were collected with a total exposure of -40 e<sup>-</sup>/Å<sup>2</sup>.

**Sample 2.** The self-stalled elongation state was trapped using similar method as 'sample 1', but with UTP instead of the non-hydrolysable UpNHpp. The subsequent grid preparation and freezing steps are similar.

Automated data collection for sample 2 was performed on a TEM Glacios (Thermo Fisher) operated at 200 kV equipped with a F4i (Thermo Fisher) direct electron detector camera and a Selectris X energy filter, using EPU. Coma and astigmatism correction were performed on a carbon grid. Micrographs were recorded in counting mode at a ×130,000 magnification giving a pixel size of 0.878 Å with defocus ranging from -0.8 to -2.0 μm. EER movies were collected with a total exposure of -40 e<sup>-</sup>/Å<sup>2</sup>.

**Sample 3.** For Zhejiang-H7N9 PA K289A + C489R structures, 0.8 μM FluPol Zhejiang-H7N9 PA K289A + C489R was mixed with (i) 0.8 μM

v51\_mut\_S and 16 μM pS5 CTD(6mer) in a final buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 2 mM TCEP. After 1 h incubation on ice, the sample was centrifuged for 5 min, 11,000 g and kept at 4 °C before proceeding to grids freezing.

For grid preparation, 1.5 μl of sample was applied on each sides of plasma cleaned (Fischione 1070 Plasma Cleaner: 1 min 10 s, 90% oxygen, 10% argon) grids (UltraAufoil 1.2/1.3, Au 300). Excess solution was blotted for 3–5 s, blot force 0, 100% humidity, at 4 °C, with a Vitrobot Mark IV (ThermoFisher) before plunge-freezing in liquid ethane.

Automated data collection was performed on a TEM Glacios (ThermoFisher) operated at 200 kV equipped with a F4i (ThermoFisher) direct electron detector camera and a Selectris X energy filter, using EPU. Coma and astigmatism correction were performed on a carbon grid. Micrographs were recorded in counting mode at a ×130,000 magnification giving a pixel size of 0.907 Å with defocus ranging from -0.8 to -2.0 μm. EER movies were collected with a total exposure of -40 e<sup>-</sup>/Å<sup>2</sup>.

## Image processing

For the TEM Titan Krios dataset (sample 1), movie drift correction was performed using Relion's Motioncor implementation, with 7 × 5 patch, using all movie frames<sup>81</sup>. All additional initial image processing steps were performed in cryoSPARC v3.3<sup>82</sup>. CTF parameters were determined using 'Patch CTF estimation', realigned micrographs were then manually inspected and low-quality images were manually discarded. To obtain an initial 3D reconstruction of FluPol Zhejiang-H7N9 4 M, particles were automatically picked on few hundreds micrographs using a circular blob with a diameter ranging from 100 to 140 Å. Particles were extracted using a box size of 360 × 360 pixels<sup>2</sup>, 2D classified and subjected to an 'ab-initio reconstruction' job. The best initial model was further used to prepare 2D templates. Template picking was then performed using a particle diameter of 120 Å and particles extracted from dose-weighted micrographs. Successive 2D classifications were used to eliminate particles displaying poor structural features. All remaining particles were then transferred to Relion 4.0. Particles were divided in subset of 300k to 500k particles and subjected to multiple 3D classification with coarse image-alignment sampling using a circular mask of 180 Å. For each similar FluPol conformation, particles were grouped and subjected to 3D masked refinement followed by multiple 3D classification without alignment or using local angular searches. Once particles properly classified, bayesian polishing was performed and re-extracted 'shiny' particles were subjected to a last 3D masked refinement. Post-processing was performed in Relion 4.0 using an automatically or manually determined B-factor. For each final map, reported global resolution is based on the FSC 0.143 cut-off criteria. Local resolution variations were estimated in Relion 4.0. Detailed image processing information is shown in Fig. S7.

For the TEM Glacios datasets (sample 2 and 3), EER fractionation was set to 24, giving -1e<sup>-</sup>/Å<sup>2</sup> per resulting fraction. Movie drift correction was performed using Relion's Motioncor implementation, with 5 × 5 patch<sup>81</sup>. All additional initial image processing steps were performed in cryoSPARC v4.0 or v4.3<sup>82</sup>. CTF parameters were determined using 'Patch CTF estimation', realigned micrographs were then manually inspected and low-quality images were manually discarded. To obtain an initial 3D reconstruction, particles were automatically picked on few hundreds micrographs using a circular blob with a diameter ranging from 100 to 140 Å. Particles were extracted using a box size of 380 × 380 pixels<sup>2</sup> (sample 2) or 340 × 340 pixels<sup>2</sup> (sample 3), 2D classified and subjected to an 'ab-initio reconstruction' job. The best initial model was further used to prepare 2D templates. Template picking was then performed using a particle diameter of 120 Å and particles extracted from dose-weighted micrographs. Successive 2D classifications were used to eliminate particles displaying poor structural features.

For sample 2, all remaining particles were subjected to a 'non-uniform refinement' job followed by a '3D classification' job. Particles

in a self-stalled elongation state were then subjected to a last non-uniform refinement. Post-processing was performed in Relion 4.0 using a manually determined B-factor. Reported global resolution is based on the FSC 0.143 cut-off criteria. Local resolution variations were estimated in Relion 4.0. Detailed image processing information is shown in Fig. S9.

For sample 3, all remaining particles were subjected to an ‘Heterogeneous refinement’ job, with three symmetrical dimers and three monomers as initial 3D models. Particles from well-resolved dimeric 3D classes were subjected to a final non-uniform refinement. Post-processing was performed in Relion 4.0 using an automatically determined B-factor. Reported global resolution is based on the FSC 0.143 cut-off criteria. Local resolution variations were estimated in Relion 4.0. Detailed image processing information is shown in Fig. S10.

### Model building and refinement

Using the FluPol Zhejiang-H7N9 elongation complex as starting point (PDB: 7QTL)<sup>43</sup>, atomic models were constructed by iterative rounds of manual model building with COOT (version 0.9.8) and real-space refinement using Phenix (version 1.20.1)<sup>83</sup>. Validation was performed using Phenix (version 1.20.1). Model resolution according to the cryo-EM map was estimated at the 0.5 FSC cutoff. Figures were generated using ChimeraX (version 1.6.1)<sup>84</sup>.

### Statistics

Statistics were performed with the GraphPad Prism software (Version 9.5.1 (528)), as indicated in the figure legends.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The coordinates and EM maps generated in this study have been deposited in the Protein Data Bank and the Electron Microscopy Data Bank: Influenza A/Zhejiang-H7N9 polymerase (4 M mutant) in replicase-like conformation in pre-initiation state with RNAP II pS5 CTD peptide mimic bound in site 1A/2A, PDB 8PMO and EMD-17755; Influenza A/Zhejiang-H7N9 polymerase (4 M mutant) in pre-initiation state with continuous RNAP II pS5 CTD peptide mimic bound in site 1A/2A PDB 8PNP and EMD-17782; Influenza A/Zhejiang-H7N9 polymerase (4 M mutant) in elongation state with continuous RNAP II pS5 CTD peptide mimic bound in site 1A/2A PDB 8PNQ and EMD-17783; Influenza A/Zhejiang-H7N9 polymerase (PA K289A + C489R) symmetric dimer bound to the promoter PDB 8POH and EMD-17792. Influenza A/H7N9 polymerase in pre-initiation state, intermediate conformation (I) with PB2-C(I), ENDO(T), and Pol II pS5 CTD peptide mimic bound in site 1A/2A PDB 8R3L and EMD-18872; Influenza A/H7N9 polymerase in self-stalled pre-termination state, with Pol II pS5 CTD peptide mimic bound in site 1A/2A PDB 8R3K and EMD-18871. The NGS raw reads generated in this study have been deposited in the European nucleotide archive under accession code ERP149587 [<https://www.ebi.ac.uk/ena/browser/view/PRJEB64419>]. The sequence files of plasmids used in this study have been deposited in the Zenodo repository [<https://doi.org/10.5281/zenodo.10462746>]. The raw data generated in this study are provided in the Source Data File. Source data are provided with this paper.

### References

- Krammer, F. et al. Influenza. *Nat. Rev. Dis. Prim.* **4**, 3 (2018).
- Harrington, W. N., Kackos, C. M. & Webby, R. J. The evolution and future of influenza pandemic preparedness. *Exp. Mol. Med.* **53**, 737–749 (2021).
- Mifsud, E. J., Hayden, F. G. & Hurt, A. C. Antivirals targeting the polymerase complex of influenza viruses. *Antivir. Res.* **169**, 104545 (2019).
- Wandzik, J. M., Kouba, T. & Cusack, S. Structure and function of influenza polymerase. *Cold Spring Harb. Perspect. Med.* **11**, a038372 (2021).
- Eisfeld, A. J., Neumann, G. & Kawaoka, Y. At the centre: influenza A virus ribonucleoproteins. *Nat. Rev. Microbiol.* **13**, 28–41 (2015).
- Plotch, S. J., Bouloy, M. & Krug, R. M. Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro. *Proc. Natl Acad. Sci. USA* **76**, 1618–1622 (1979).
- Deng, T., Vreede, F. T. & Brownlee, G. G. Different de novo initiation strategies are used by influenza virus RNA polymerase on its cRNA and viral RNA promoters during viral RNA replication. *J. Virol.* **80**, 2337–2348 (2006).
- Dias, A. et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914–918 (2009).
- Guilligay, D. et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* **15**, 500–506 (2008).
- Kouba, T., Drncová, P. & Cusack, S. Structural snapshots of actively transcribing influenza polymerase. *Nat. Struct. Mol. Biol.* **26**, 460–470 (2019).
- Pflug, A. et al. Capped RNA primer binding to influenza polymerase and implications for the mechanism of cap-binding inhibitors. *Nucleic Acids Res.* **46**, 956–971 (2018).
- Wandzik, J. M. et al. A structure-based model for the complete transcription cycle of influenza polymerase. *Cell* **181**, 877–893.e21 (2020).
- Carrique, L. et al. Host ANP32A mediates the assembly of the influenza virus replicase. *Nature* **587**, 638–643 (2020).
- Keown, J. R. et al. Mapping inhibitory sites on the RNA polymerase of the 1918 pandemic influenza virus using nanobodies. *Nat. Commun.* **13**, 251 (2022).
- Wang, F. et al. The C-terminal LCAR of host ANP32 proteins interacts with the influenza A virus nucleoprotein to promote the replication of the viral RNA genome. *Nucleic Acids Res.* **50**, 5713–5725 (2022).
- Camacho-Zarco, A. R. et al. Multivalent dynamic colocalization of avian influenza polymerase and nucleoprotein by intrinsically disordered ANP32A reveals the molecular basis of human adaptation. *J. Am. Chem. Soc.* **145**, 20985–21001 (2023).
- Poon, L. L., Pritlove, D. C., Fodor, E. & Brownlee, G. G. Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J. Virol.* **73**, 3473–3476 (1999).
- Zhu, Z., Fodor, E. & Keown, J. R. A structural understanding of influenza virus genome replication. *Trends Microbiol.* **31**, 308–319 (2023).
- Oymans, J. & Te Velthuis, A. J. W. A mechanism for priming and realignment during influenza A virus replication. *J. Virol.* **92**, e01773–17 (2018).
- Zhang, S., Wang, J., Wang, Q. & Toyoda, T. Internal initiation of influenza virus replication of viral RNA and complementary RNA in vitro. *J. Biol. Chem.* **285**, 41194–41201 (2010).
- York, A., Hengrung, N., Vreede, F. T., Huiskonen, J. T. & Fodor, E. Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc. Natl Acad. Sci. USA* **110**, E4238–E4245 (2013).
- Krischuns, T. et al. Type B and type A influenza polymerases have evolved distinct binding interfaces to recruit the RNA polymerase II CTD. *PLoS Pathog.* **18**, e1010328 (2022).
- Lukarska, M. et al. Structural basis of an essential interaction between influenza polymerase and Pol II CTD. *Nature* **541**, 117–121 (2017).
- Rialdi, A. et al. The RNA exosome syncs IAV-RNAPII transcription to promote viral ribogenesis and infectivity. *Cell* **169**, 679–692.e14 (2017).

25. Long, J. S. et al. Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* **529**, 101–104 (2016).
26. Engelhardt, O. G., Smith, M. & Fodor, E. Association of the influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. *J. Virol.* **79**, 5812–5818 (2005).
27. Harlen, K. M. & Churchman, L. S. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **18**, 263–273 (2017).
28. Core, L. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* **33**, 960–982 (2019).
29. Serna Martin, I. et al. A mechanism for the activation of the influenza virus transcriptase. *Mol. Cell* **70**, 1101–1110.e4 (2018).
30. Nilsson-Payant, B. E., tenOever, B. R. & Te Velthuis, A. J. W. The host factor ANP32A is required for influenza A virus vRNA and cRNA synthesis. *J. Virol.* **96**, e0209221 (2022).
31. Swann, O. C., Rasmussen, A. B., Peacock, T. P., Sheppard, C. M. & Barclay, W. S. Avian influenza A Virus polymerase can utilize human ANP32 proteins to support cRNA but not vRNA synthesis. *mBio.* e0339922 (2023).
32. Baker, S. F., Ledwith, M. P. & Mehle, A. Differential splicing of ANP32A in birds alters its ability to stimulate RNA synthesis by restricted influenza polymerase. *Cell Rep.* **24**, 2581–2588.e4 (2018).
33. Mistry, B. et al. Elucidating the interactions between influenza virus polymerase and host factor ANP32A. *J. Virol.* **94**, e01353–19 (2020).
34. Morel, J. et al. The influenza virus RNA-Polymerase and the host RNA-polymerase II: RPB4 is targeted by a PB2 domain that is involved in viral transcription. *Viruses* **14**, 518 (2022).
35. Kawakami, E. et al. Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. *J. Virol. Methods* **173**, 1–6 (2011).
36. Cusack, S., Naffakh, N., Krischuns, T. & Drncova, P. High-resolution structure of a replication-initiation like configuration of influenza polymerase active site visualises the essential role of a conserved dibasic motif in the PA subunit. 2023.11.20.567839 Preprint at <https://doi.org/10.1101/2023.11.20.567839v1> (2023).
37. Tang, Y.-S., Xu, S., Chen, Y.-W., Wang, J.-H. & Shaw, P.-C. Crystal structures of influenza nucleoprotein complexed with nucleic acid provide insights into the mechanism of RNA interaction. *Nucleic Acids Res.* **49**, 4144–4154 (2021).
38. Fournier, G. et al. Recruitment of RED-SMU1 complex by influenza A Virus RNA polymerase to control Viral mRNA splicing. *PLoS Pathog.* **10**, e1004164 (2014).
39. Chen, K.-Y., Santos Afonso, E. D., Enouf, V., Isel, C. & Naffakh, N. Influenza virus polymerase subunits co-evolve to ensure proper levels of dimerization of the heterotrimer. *PLoS Pathog.* **15**, e1008034 (2019).
40. Chang, S. et al. Cryo-EM structure of influenza virus RNA polymerase complex at 4.3 Å resolution. *Mol. Cell* **57**, 925–935 (2015).
41. Fan, H. et al. Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature* **573**, 287–290 (2019).
42. Sheppard, C. M. et al. An Influenza A virus can evolve to use human ANP32E through altering polymerase dimerization. *Nat. Commun.* **14**, 6135 (2023).
43. Kouba, T. et al. Direct observation of backtracking by influenza A and B polymerases upon consecutive incorporation of the nucleoside analog T1106. *Cell Rep.* **42**, 111901 (2023).
44. Chen, Y. et al. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. *Lancet* **381**, 1916–1925 (2013).
45. Li, H. et al. An intermediate state allows influenza polymerase to switch smoothly between transcription and replication cycles. *Nat. Struct. Mol. Biol.* **30**, 1183–1192 (2023).
46. Thierry, E. et al. Influenza polymerase can adopt an alternative configuration involving a radical repacking of PB2 domains. *Mol. Cell* **61**, 125–137 (2016).
47. Vreede, F. T. & Brownlee, G. G. Influenza virion-derived viral ribonucleoproteins synthesize both mRNA and cRNA in vitro. *J. Virol.* **81**, 2196–2204 (2007).
48. Fodor, E., Mingay, L. J., Crow, M., Deng, T. & Brownlee, G. G. A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase promotes the generation of defective interfering RNAs. *J. Virol.* **77**, 5017–5020 (2003).
49. Chan, A. Y., Vreede, F. T., Smith, M., Engelhardt, O. G. & Fodor, E. Influenza virus inhibits RNA polymerase II elongation. *Virology* **351**, 210–217 (2006).
50. Newcomb, L. L. et al. Interaction of the influenza A virus nucleocapsid protein with the viral RNA polymerase potentiates unprimed viral RNA replication. *J. Virol.* **83**, 29–36 (2009).
51. Vreede, F. T., Jung, T. E. & Brownlee, G. G. Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *J. Virol.* **78**, 9568–9572 (2004).
52. Cramer, P. Organization and regulation of gene transcription. *Nature* **573**, 45–54 (2019).
53. Takizawa, N., Watanabe, K., Nouno, K., Kobayashi, N. & Nagata, K. Association of functional influenza viral proteins and RNAs with nuclear chromatin and sub-chromatin structure. *Microbes Infect.* **8**, 823–833 (2006).
54. López-Turiso, J. A., Martínez, C., Tanaka, T. & Ortín, J. The synthesis of influenza virus negative-strand RNA takes place in insoluble complexes present in the nuclear matrix fraction. *Virus Res* **16**, 325–337 (1990).
55. Bauer, D. L. V. et al. Influenza virus mounts a two-pronged attack on host RNA polymerase II transcription. *Cell Rep.* **23**, 2119–2129.e3 (2018).
56. Heinz, S. et al. Transcription elongation can affect genome 3D structure. *Cell* **174**, 1522–1536.e22 (2018).
57. Zhao, N. et al. Influenza virus infection causes global RNAPII termination defects. *Nat. Struct. Mol. Biol.* **25**, 885–893 (2018).
58. Rodríguez, A., Pérez-González, A. & Nieto, A. Influenza virus infection causes specific degradation of the largest subunit of cellular RNA polymerase II. *J. Virol.* **81**, 5315–5324 (2007).
59. Vreede, F. T., Chan, A. Y., Sharps, J. & Fodor, E. Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells. *Virology* **396**, 125–134 (2010).
60. Prudêncio, P., Savaisaar, R., Rebelo, K., Martinho, R. G. & Carmo-Fonseca, M. Transcription and splicing dynamics during early *Drosophila* development. *RNA* **28**, 139–161 (2022).
61. Zaborowska, J., Egloff, S. & Murphy, S. The pol II CTD: new twists in the tail. *Nat. Struct. Mol. Biol.* **23**, 771–777 (2016).
62. Mänz, B. et al. Multiple natural substitutions in avian influenza A virus PB2 facilitate efficient replication in human cells. *J. Virol.* **90**, 5928–5938 (2016).
63. Fodor, E. et al. Rescue of influenza A virus from recombinant DNA. *J. Virol.* **73**, 9679–9682 (1999).
64. Vieira Machado, A., Naffakh, N., Gerbaud, S., van der Werf, S. & Escriou, N. Recombinant influenza A viruses harboring optimized dicistronic NA segment with an extended native 5' terminal sequence: induction of heterospecific B and T cell responses in mice. *Virology* **345**, 73–87 (2006).
65. Cassonnet, P. et al. Benchmarking a luciferase complementation assay for detecting protein complexes. *Nat. Methods* **8**, 990–992 (2011).
66. Jagger, B. W. et al. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199–204 (2012).
67. Long, J. S. et al. Species specific differences in use of ANP32 proteins by influenza A virus. *eLife* **8**, e45066 (2019).
68. Munier, S., Rolland, T., Diot, C., Jacob, Y. & Naffakh, N. Exploration of binary virus-host interactions using an infectious protein complementation assay. *Mol. Cell Proteom.* **12**, 2845–2855 (2013).

69. Zheng, L., Baumann, U. & Reymond, J.-L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, e115 (2004).
70. van Diemen, F. R. et al. CRISPR/Cas9-mediated genome editing of herpesviruses limits productive and latent infections. *PLoS Pathog.* **12**, e1005701 (2016).
71. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).
72. Krischuns, T. et al. Phosphorylation of TRIM28 enhances the expression of IFN- $\beta$  and proinflammatory cytokines during HPAIV Infection of human lung epithelial cells. *Front. Immunol.* **9**, 2229 (2018).
73. Da Costa, B. et al. Temperature-sensitive mutants in the influenza A virus RNA polymerase: alterations in the PA linker reduce nuclear targeting of the PB1-PA dimer and result in viral attenuation. *J. Virol.* **89**, 6376–6390 (2015).
74. Matrosovich, M., Matrosovich, T., Garten, W. & Klenk, H.-D. New low-viscosity overlay medium for viral plaque assays. *Virol. J.* **3**, 63 (2006).
75. Watson, S. J. et al. Viral population analysis and minority-variant detection using short read next-generation sequencing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120205 (2013).
76. Sjaastad, L. E. et al. Distinct antiviral signatures revealed by the magnitude and round of influenza virus replication in vivo. *Proc. Natl Acad. Sci. USA* **115**, 9610–9615 (2018).
77. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
78. Camacho-Zarco, A. R. et al. Molecular basis of host-adaptation interactions between influenza virus polymerase PB2 subunit and ANP32A. *Nat. Commun.* **11**, 3656 (2020).
79. Huyton, T. & Wolberger, C. The crystal structure of the tumor suppressor protein pp32 (Anp32a): structural insights into Anp32 family of proteins. *Protein Sci.* **16**, 1308–1315 (2007).
80. Kechin, A., Boyarskikh, U., Kel, A. & Filipenko, M. cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput Biol.* **24**, 1138–1143 (2017).
81. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
82. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
83. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D. Struct. Biol.* **74**, 531–544 (2018).
84. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

## Acknowledgements

We thank Ron Fouchier (Erasmus MC Department of Viroscience, Netherlands), George Brownlee (Oxford University), Bernard Delmas (INRAE, France), Robert-Jan Lebbink (Utrecht University, Netherlands) and Sandie Munier (Institut Pasteur, France) for sharing plasmids and antibodies. We thank Martin Pelosse for support in using the Eukaryotic Expression Facility at EMBL Grenoble; Aldo R. Camacho-Zarco and Martin Blackledge for providing huANP32A truncated constructs proteins; Daphne Walter and Vladimir Benes from the EMBL GeneCore facility for the next generation sequencing data; Wojtek Galej, Sarah Schneider and Romain Linares for access to the Glacios at EMBL Grenoble; Guy Schoehn and Eleftherios Zarkadas for access to the Glacios at IBS Grenoble; Aymeric Peuch for support with using the joint EMBL-IBS computer clusters. We acknowledge the European Synchrotron

Radiation Facility and PSB for access to the Titan Krios CMO1, and Michael Hons for assistance with data collection. We thank the Pasteur International Bioresources Network (PIBNet) for genome sequencing of recombinant viruses. This work was funded by the ANR grant FluTranscript (ANR-18-CE11-0028, S.C. and N.N.), the ANR LabEx IBEID (ANR-10-LABX-62-IBEID, N.N. and T.K.), and the German Research Foundation (DFG) grant SFB-TR84 (project B2, T.W.). This work used the platforms at the Grenoble Instruct-ERIC Center (ISBG; UMS 3518 CNRS CEA-UGA-EMBL) with support from the French Infrastructure for Integrated Structural Biology (FRISBI; ANR-10-INSB-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003) within the Grenoble Partnership for Structural Biology. The IBS Electron Microscope facility is supported by the Auvergne Rhône-Alpes Region, the Fonds Feder, the Fondation pour la Recherche Médicale and GIS-IBISA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

TK, SC and NN conceived the project. TK and BA performed most of the experiments. CI performed serial passaging of mutant viruses. SP provided research assistance. MB and TW generated and provided essential reagents. TK, BA, CI, SC and NN oversaw and interpreted data. TK, BA, SC and NN wrote the manuscript with input from all authors. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-45205-2>.

**Correspondence** and requests for materials should be addressed to Tim Krischuns, Stephen Cusack or Nadia Naffakh.

**Peer review information** *Nature Communications* thanks Alex Chin, Haitian Fan and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024