

Published in final edited form as:

*Nat Struct Mol Biol.* 2019 June 01; 26(6): 460–470. doi:10.1038/s41594-019-0232-z.

## Structural snapshots of actively transcribing influenza polymerase

Tomas Kouba<sup>1,§</sup>, Petra Drncova<sup>1,§</sup>, Stephen Cusack<sup>1,\*</sup>

<sup>1</sup>European Molecular Biology Laboratory, 71 Avenue des Martyrs, CS 90181, 38042 Grenoble Cedex 9, France

### Abstract

Influenza virus RNA-dependent RNA polymerase uses unique mechanisms to transcribe its single-stranded genomic vRNA into mRNA. The polymerase is initially bound to a promoter comprising the partially base-paired 3' and 5' extremities of the vRNA. A short, capped primer, 'cap-snatched' from a nascent host polymerase II transcript, is directed towards the polymerase active site to initiate RNA synthesis. Here we present structural snapshots, determined by X-ray crystallography and cryo-electron microscopy, of actively initiating influenza polymerase as it transitions towards processive elongation. Unexpected conformational changes unblock the active site cavity to allow establishment of a nine base-pair template-product RNA duplex before the strands separate into distinct exit channels. Concomitantly, as the template translocates, the promoter base-pairs are broken and the template entry region is remodelled. These structures reveal new details of the influenza polymerase active site that will help optimize nucleoside analogs or other compounds that directly inhibit viral RNA synthesis.

### Introduction

The heterotrimeric influenza polymerase, which comprises subunits PA, PB1 and PB2, binds to the conserved 5' and 3' termini (the 'promoter') of each of the eight negative sense single-stranded viral RNA (vRNA) genome segments. It both transcribes and replicates the vRNA in the infected cell nucleus<sup>1–3</sup> and is the major target for novel anti-influenza drugs<sup>2,4,5</sup>. Influenza polymerase employs unique mechanisms to synthesise 5' capped and 3' poly-adenylated viral mRNA. Capped oligomers of 10–15 nucleotides in length, pirated

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Corresponding author: Stephen Cusack, European Molecular Biology Laboratory, Grenoble Cedex 9, France. Tel. (33)476207238, [cusack@embl.fr](mailto:cusack@embl.fr).

§Equal contribution

#### Author contributions:

T.K. performed transcription assays, prepared cryo-EM grids, collected EM data, performed image processing and cryo-EM 3D reconstruction, built initial models and made figures; P.D. expressed, purified and crystallized FluB polymerase; S.C. conceived and supervised the project, collected crystallographic data, did crystallographic analysis, refined atomic models; T.K. and S.C. wrote the manuscript with input from P.D.

#### Competing interests:

T.K., P.D. and S.C. have filed for a patent related to this work.

#### Request for Materials and Correspondence

should be addressed to Stephen Cusack ([cusack@embl.fr](mailto:cusack@embl.fr)).

from nascent Pol II transcripts, are used to prime transcription<sup>6,7</sup>. Polyadenylation is achieved by stuttering of the polymerase on an oligo(U) stretch proximal to the template 5' end<sup>8</sup>. Attempts to structurally characterize the different steps of actively transcribing influenza polymerase<sup>9</sup> have hitherto been unsuccessful. For instance, capped primers with 3' ends (e.g. m<sup>7</sup>GpppNN...AG-3') complementary to the 3' end of the template (3'-UCGUCU...) are most efficient in transcription initiation<sup>10</sup>. However, the structure of the transcription initiation state, in which a short primer-template duplex is expected to be positioned in the active site to allow further nucleotide incorporation, has so far been elusive<sup>9,11</sup>. There are several reasons for the difficulty in trapping this complex, whose formation is the rate-limiting step in RNA synthesis<sup>11</sup> and whose instability can result in abortive product formation<sup>12</sup>. Firstly, the 3' end of the template is flexible as shown by single-molecule FRET<sup>13</sup> and consistent with this, crystal structures either locate it on the surface of the polymerase or threaded through the template entrance channel into the active site cavity<sup>11,14,15</sup>. Secondly, an element of PB1 known as the priming loop, which is thought to stabilize the initiation state for unprimed RNA synthesis (i.e. replication)<sup>16</sup>, normally occupies the active site cavity and would clash with the presumed position of the 3' extremity of the primer<sup>11</sup>. How the priming loop is displaced to allow transcription initiation and processive RNA synthesis to proceed is currently unknown. Thirdly, template translocation requires disruption of the highly stable distal 5'-3' base paired region of the promoter (Fig. 1A, top). This is energetically unfavourable without the formation of compensating product-template base pairs in the active site<sup>11,13</sup>.

Here we employ a modified template that stabilises the initial primer-template interaction and enables visualisation by X-ray crystallography of a transcription initiation-like state. Subsequent addition of only ATP and GTP leads to polymerase stalling after product elongation by five nucleotides with concomitant template translocation, enabling us to visualise by cryo-electron microscopy (cryo-EM) the initiation to elongation transition.

## Results

### Overview of structures obtained

We used an 18+3-mer template in which the native 3' end is artificially extended in length by three nucleotides (3'-*UUAUCGUCU...*, three nucleotide extension in italics) and a 14- or 15-mer capped primer ending in ...AAUAG-3' or ...AAUAGC-3' respectively that could potentially make five or six base pairs with the extended template (Fig. 1B, top). Influenza B (FluB) polymerase was co-crystallized with these RNAs and anisotropic diffraction data were integrated using an ellipsoidal mask<sup>17</sup> to a maximum resolution of 2.9-3.1 Å (Table 1). Structure solution revealed clear electron density for the complete primer-template RNA (Supplementary Fig. 1a), annealed as expected with a four or five base pair duplex between the primer and extended template in the active site cavity, for the 14- or 15-mer capped primer respectively (Fig. 1b). With respect to the native 3' end, these structures mimic an initiation complex since the promoter structure is intact and the template is not translocated<sup>11,14</sup>. Correspondingly, there are either two base-pairs between the 14-mer primer (AG-3') and native template terminal nucleotides (3'-UC), leaving space for an incoming CTP as the next nucleotide, or three base-pairs between the 15-mer primer

(AGC-3') and template (3'-UCG) representing a post-incorporation, pre-translocation state. However, the extra primer-template base-pairs resulting from the artificially extended template imply that the crystal structures also have features of a post-initiation state in which the template has pseudo-translocated by three nucleotides. We also determined a third crystal structure with the 14-mer capped primer and soaked incoming CTP (Table 1). Local changes in the active site resulting from CTP binding and incorporation are described later. In all other respects, notably the partly extruded position of the priming loop (see below), the three crystal structures are the same and we refer to the observed conformation as a mixed-initiation state. Importantly, we were able to show that these complexes are functional for further elongation with only a minor loss of efficiency of the 18+3-mer compared to the native 18-mer template (Fig. 1d, Supplementary Figure 2). In particular, addition of only ATP and GTP extends the 15-mer primer by five nucleotides to form a capped 20-mer product, before lack of CTP causes polymerase stalling (Fig. 1d). This stalled elongation intermediate survives heparin treatment<sup>18,19</sup> and is able to continue the transcription reaction upon addition of CTP showing that it is on pathway (Fig. 1d).

We determined that four hours incubation was optimal for formation of the stalled elongation complex (Supplementary Fig. 2) after which the reaction sample was frozen on cryo-electron microscope (cryo-EM) grids and data permitting 3D single particle reconstruction were collected. Two 3D classes are of particular interest. One structure, at an overall resolution of 3.16 Å (Supplementary Fig. 3–4, Table 2), shows a pre-initiation state with unperturbed promoter base pairing, the template 3' end located on the surface of the protein and the capped RNA primer partially ordered (Fig. 1a). This cryo-EM pre-initiation conformation is similar to previous crystal structures of promoter bound influenza B polymerase, but with the PB1 β-ribbon packing against the promoter rather than directed away from it due to crystal contacts (compare PB1 β-ribbon position in Figs. 1a and 1b). The second cryo-EM structure, at an overall resolution of 3.20 Å (Supplementary Fig. 3–4, Table 2), shows an early elongation state. The template is translocated by five nucleotides, resulting in a remarkable nine base pair template-product duplex in the active site cavity (Fig. 1c, Supplementary Fig. 1c, Supplementary Video 1). Correspondingly the primer is extended by five nucleotides, the cryo-EM map being of sufficient resolution to confirm that not only is the product of the correct length (20-mer) but that the duplex sequence is as expected (Supplementary Fig. 1d). The transition to elongation involves a 4.5° outward rotation of the thumb domain (residues 509-670 of PB1, denoted PB1/509-670) together with the PB2 N1-N2 domains (PB2/54-153) (Supplementary Fig. 5). In addition, template translocation correlates with promoter base pair disruption and reconfiguration of the template entrance and exit channels. These structures (Figs. 1a–c), discussed in more detail below, together with a previous pre-initiation crystal structure<sup>11</sup>, reveal the successive transitions that transcribing influenza polymerase undergoes in progressing from pre-initiation to early elongation.

### **Progressive extrusion of the priming loop allows growth of the product-template duplex**

A priming loop inserted into the active site cavity has been shown to promote initiation of un-primed RNA synthesis for several viral polymerases<sup>20,21</sup>. For influenza, the priming loop plays a mechanistic role in vRNA to cRNA replication<sup>16</sup> and in early stages of

transcription<sup>22,23</sup>. In the cryo-EM structure of the pre-initiation state (Fig. 1a), the complete priming loop (PB1/631-660) is visible in its characteristic  $\beta$ -hairpin conformation (Fig. 2a, Supplementary Fig. 6c), preventing the 3' end of the primer approaching the active site. The tip of the priming loop (648-AHGP) forms a platform above the active site, leaving just enough space to accommodate and align the two nucleoside triphosphates at the +1 (GTP) and -1 (ATP) positions consistent with its importance for *de novo* initiation of vRNA to cRNA replication<sup>16</sup>. In this structure, the expected base pairing between the primer and template cannot occur due to the presence of the priming loop. Instead, the 3' end of the template turns away from the template entrance and binds to the PB1  $\beta$ -ribbon (Fig. 2a), as previously observed<sup>14</sup>. Similarly, the cap proximal part of the primer binds to the PB2 cap binding and midlink domains as described<sup>9</sup> but is not visible beyond the fifth nucleotide (Fig. 2a). In the pre-initiation state, the mobile template 3' end can also flip into the active site adjacent to the tip of the priming loop, which then becomes slightly disordered<sup>11</sup> (Fig. 2b).

In the crystal structures of the mixed initiation complex (Fig. 1b), the primer forms five base pairs (position +1 to -4) with the template and this is only possible due to the partial retraction of the priming loop (Fig. 2c, Supplementary Fig. 6d). The last template base U21(-5 position) turns is however not paired with A10 of the primer; instead, both these bases as well as primer base A8 form a close packed arrangement that butts against residues PB1/652-655 of the repositioned priming loop with conserved PB1/Asp655 forming a salt-bridge with PB2/Arg218 (Supplementary Fig. 6a). The rest of the priming loop (residues 632-637 and 643-656) refolds into and still blocks the template exit channel (see below) with residues 638-642 forming a partially disordered loop projecting into the solvent, near the N-terminus of PB2. The observed position of the priming loop likely does not correspond to that in the true initiation state but to a later post-initiation state, in line with the mixed nature of the structure.

Following addition of ATP and GTP, five nucleotides are incorporated into the product and a nine base pair duplex (positions -1 to -9) is established in the active site cavity (Fig. 1c and 2d). This early elongation complex is in the post-translocation state, with the incoming nucleotide position at +1 being vacant (i.e. similar to the crystal structure with the 14-mer capped RNA primer). Nucleotide addition would thus result in a ten base pair duplex (position +1 to -9) before translocation. To accommodate the extended RNA duplex, the priming loop has to be fully displaced from the active site cavity with 17 residues being extruded into the solvent in a disordered loop that projects towards the PB1-PB2 interface helical bundle (Fig. 2d, Supplementary Fig. 6e). These results show for the first time how the priming loop is successively displaced during early transcription (see also Supplementary Fig. 6b-e).

### **The mechanism of product-template strand separation and template exit channel opening**

The elongation state structure clearly identifies the PB2 helical lid domain (PB2/153-212) as responsible for product-template strand separation in influenza polymerase (Fig. 3a), confirming a previous suggestion<sup>14</sup>. The helical lid lies on top of the active site chamber with helix  $\alpha$ 12 facing the growing product-template duplex. Tyr207 (conserved in all

influenza A and B strains, histidine in influenza C) changes rotamer to stack on the template base of the last base pair, preventing duplex continuation beyond position -9 (Fig. 3a). The stable, nine base pair long product-template duplex in the active site chamber disfavors product dissociation. This likely explains the high processivity of elongating influenza polymerase, compared to the initiation phase where abortive products frequently occur due to the instability of a too short product-template duplex<sup>12</sup>. In the initiation state, the partially extruded priming loop still blocks the template exit channel (Fig. 3b–e). Only in the elongation state, when the priming loop is fully extruded, is there an unobstructed exit channel for the template after strand separation (Fig. 3d,e). There is a sharp kink at the -9 to -10 junction as the template exits and the -10 nucleotide is accommodated in the exit channel without specific interactions (Fig. 3a).

### Conformation and protein-RNA interactions of the template, primer and product

Viral RNA-dependent RNA polymerases have a conserved catalytic core with six characteristic structural motifs (A to F) that are responsible, together with two divalent cations (denoted A and B), for controlling the nucleotide addition cycle<sup>24</sup>. Our structures reveal how these motifs, and other parts of the influenza polymerase core, interact with the primer, template and incoming nucleotide during transcription initiation and elongation (Fig. 4). In the elongation state, the product-template duplex fills the active site cavity and residues lining it make numerous van der Waals and polar interactions to the backbone of both strands (Fig. 4b). The proximal part of the template strand (+1 to -4 position) binds to the fingers (residues 127-136), fingertips (residues 227-229 and 241-249 of motif F) and palm (residues 271-274 and motif B 412-415) domains of PB1. These interactions are conserved in the mixed-initiation complex (Fig. 4a). The distal part of the translocated template (positions -5 to -9), only present in the elongation complex, crosses over to the thumb domain to be contacted by PB1 residues 527-531 as well as conserved PB2 basic residues Arg211, Arg216 and Arg218 which interact with the phosphates at positions -8, -7 and -6 respectively. The template enters the active site by undergoing the characteristic kink at the +1 to +2 junction<sup>25</sup> with the +1 guanine base stacking on PB1/Ile241 and making a polar interaction with PB1/Lys229, both conserved motif F residues (Fig. 4a–c). In the mixed-initiation complex the template then follows a sinuous path back to the double stranded region of the promoter<sup>11,15</sup>, whereas the entire promoter region is remodelled in the elongation complex (see below).

The complete 15-mer capped primer is observed in the corresponding mixed initiation state complex (Supplementary Fig. 1a, 7). Both the cap and the first base of the primer bind in specific pockets formed by the PB2 cap-binding and midlink domain as previously described<sup>9,11</sup>. The following three bases (2-AAU-4) are stacked on each other and sandwiched between PB2/Tyr434 of the cap-binding domain and a salt-bridge between PB2/Arg217 and Glu155 of the PB2 N2 domain (Fig. 4a, Supplementary Fig. 7). Between nucleotides five to ten, the capped primer has an irregular backbone structure that probably will vary according to its exact length and sequence. There is a distinct pocket for the splayed out base of G5, which is stacked between PB2/Arg146 and Arg425 and makes base contacts with PB1/Glu227. Poorly ordered bases C6 and U7 are turned outwards, whereas A8 stacks on A10 at the top of the primer-template duplex region and U9 is bulged out.

Primer bases 11-15 form an A form duplex with the template (+1 to -4). The distal part of the primer backbone interacts with PB1 Tyr24, Arg233, Ser493 and the peptide PB1/506-510, whereas the proximal part interacts with residues PB1/443-445 from motif C and PB1/309-310 from motif A (Fig. 4a).

In the elongation complex, the interactions with the cap proximal region of the primer ( $m^7GpppGAAU$ ) are maintained, with virtually no change in the position of the cap-binding domain, but the next seven nucleotides are disordered, presumably due to bulging out of the product, before connecting with the end of the duplex after strand separation (Fig. 3a). In the nine base pair duplex region, the backbone of nucleotides 16-20 of the product interact as the equivalently placed nucleotides 11-15 in the initial position of primer (Fig. 4a,b), whereas the backbone of primer nucleotides 12-15, due to their translocation, now interact with PB2/35-45 and PB1/124-126 and Lys706 (Fig. 4b).

### The methionine-rich motif B loop reconfigures to stabilise the incoming nucleotide

In the cryo-EM elongation structure the incoming NTP position opposite template base G(+1) is vacant and thus corresponds to the post-translocation state (Fig. 4b). A very similar active site configuration is observed in the crystal structure with the 14-mer primer. Soaking CTP into these crystals reveals the incoming nucleotide (Fig. 4d, Supplementary Fig. 1e), whereas the 15-mer primer crystal structure, in which the primer-template base pair C15-G(+1) is at the active site position, corresponds to a post-incorporation, pre-translocation state (Fig. 4a,c). Indeed, this structure mimics a reaction product complex since there is a pyrophosphate at the position of the triphosphate of the incoming NTP (Fig. 4a,c). In the cryo-EM elongation structure, a presumed  $Mg^{2+}$  ion, coordinated by motif C PB1/Asp445, motif E PB1/Glu490 and the carbonyl oxygen of motif A PB1/Gly304, is found in a position typical for metal A in the inactive 'open' state<sup>25</sup> (Fig. 4b,c). In the 15-mer crystal structure, metal A is not observed, but there is a magnesium at the metal B site, octahedrally coordinated by motif A PB1/Asp305, the carbonyl oxygen of motif A PB1/Gly306, motif C PB1/Asp444, the phosphate of the terminal primer nucleotide C15 and the pyrophosphate (Fig. 4a,c). In the CTP bound, 14-mer primer structure, only metal B is observed. Based on the architectural similarity of viral RNA polymerase active sites, local conformational changes of motifs A and D will be required to reorientate the active site triad of aspartates (PB1/Asp305, Asp444, Asp445) to correctly bind the two metals in a catalytic configuration<sup>25</sup>.

Comparing the post-translocation (NTP site vacant), pre-catalytic (NTP site occupied) and pre-translocation state (NTP incorporated into the product) structures highlights the role of the flexible, methionine-rich motif B loop (PB1/407-GMMMGMF-413, highly conserved in all orthomyxoviruses) in adapting to the presence of a base at the incoming nucleotide position +1 (Fig. 4d). In the pre-translocation and pre-catalytic states (i.e. with CTP), this base stacks on the side-chain of Met410 and also contacts the backbone of Gly411. The carbonyl oxygens of Gly411 and Asp242 hydrogen bond to the 2' hydroxyls of the template nucleotides at respectively the -1 and +1 positions and the side-chain of Asn310 (Motif A) interacts with both hydroxyls of the primer nucleotide at position +1. In the post-translocation state, when the incoming nucleotide site is vacant, motif B loop reconfigures,

with Met410 being buried in a hydrophobic pocket, necessitating adjustments of the side-chains Phe344 and Phe413 as well (Fig. 4d, Supplementary Fig. 8).

### Progression to elongation requires promoter melting and remodelling of the template entrance.

The intact vRNA promoter, as observed in the pre-initiation state, comprises both conserved extremities of the vRNA (Fig. 1a). The 5' nucleotides 1–10 form a compact stem–loop (hook) structure (Supplementary Fig. 1b) that is tightly bound in a pocket formed by the PA and PB1 subunits and is required to activate polymerase functions<sup>15,26</sup>. The 5' nucleotides 11–14 base pair with 3' nucleotides 10–13 forming a stem that project away from the polymerase (Fig. 1a). The intact promoter configuration is stabilized by a network of specific interactions, including stacking and pseudo base pairing of PA/His506 with 5' A10 and 3' G9 respectively, and stacking of PA/Met473 on 3' G9 (Fig. 5a,c). Of particular importance is the short PB1/670–677  $\eta$ -helix, which stabilizes the pronounced turn of the 3' backbone between nucleotides 7–11 (Fig. 5a). Finally, the tip of the PB1  $\beta$ -ribbon loosely engages with end of the promoter duplex from the solvent side in the region of 5' G14–3' C13 (Fig. 5a,c).

A major reorganization of the promoter-binding region occurs during the initiation-elongation transition, a consequence of the promoter melting necessitated by template translocation (Fig. 5b). The entire incoming template strand from nucleotide 7 to the duplex region in the active site is displaced from its initial position, allowing the  $\beta$ -ribbon to rotate by  $\sim 25^\circ$  towards the 5' promoter strand (Fig. 5b–e). The peptide PB1/667–681 also reconfigures and the  $\eta$ -helix is replaced by a short  $\beta$ -strand PB1/675–677 that establishes a three-stranded  $\beta$ -sheet with the  $\beta$ -ribbon, stabilising its new position (Fig. 5b,d). The triple stack of PA/His505 with 5' bases A11 and G12 remains, but with  $\beta$ -ribbon residue PB1/Leu200 now also packing against 5' G12, forcing 5' A13 into a new position (Fig. 5d). An additional consequence of the displacement of template nucleotides 7 to 8 is that peptide PB2/37–44 repositions into the space vacated by these nucleotides (Fig. 5b,e). This movement widens the active site cavity, allowing the product-template helix to grow, with residues from PB2/37–44 interacting with the product strand (Fig. 4b) rather than sterically clashing with it. After the transition to elongation, the incoming template can therefore no longer follow the entry pathway into the active site as observed in the pre-initiation and initiation states. However, the cryo-EM map does not allow a clear definition of the trajectories of the outgoing 5' end beyond nucleotide A13 or the incoming template prior to the +3 position (Fig. 5d), although uninterpretable density exists in the vicinity. It thus remains to be elucidated whether there is some mechanism to prevent interference between the translocating incoming template and the hook proximal 5' strand, which are in close proximity. Interestingly, the remodelled PB1/667–681 region immediately follows the priming loop (PB1/631–660), which itself is being extruded during the initiation to elongation transition (Fig. 5e). This suggests that these two conformational changes might be coupled. However a previous co-crystal structure of FluB polymerase bound only to 5' vRNA nucleotides 1–12 (PDB 5EPI<sup>26</sup>) also shows the triple-stranded configuration of the collapsed  $\beta$ -ribbon and the shift in PB2/37–44 but without exhibiting priming loop extrusion. Thus removal of the 3' template from the promoter is sufficient to induce  $\beta$ -ribbon collapse

whereas priming loop extrusion is rather correlated with product-template helix growth and thumb domain opening as a second step after promoter melting.

## Discussion

Our results, summarized in Figure 6, show that in the early stages of transcription by influenza polymerase the initiation to elongation transition involves progressive priming loop extrusion coupled to widening of the active site cavity. The latter accommodates growth of the product-template duplex to a steady-state of nine base pairs in the post-translocation state. Subsequently, the PB2 helical lid enforces strand separation, directing the template into the newly opened template exit channel, whereas the increasingly bulged capped mRNA transcript eventually forces the release from the cap-binding domain<sup>27</sup>. Concomitantly, template translocation causes promoter disruption, resulting in collapse of the PB1  $\beta$ -ribbon onto the vRNA 5' end and template entry channel remodelling. The steric hindrance of the priming loop to RNA duplex growth is consistent with priming loop deletions increasing transcription and may also explain its role in the process of template backtracking and realignment during early transcription<sup>16,22,23</sup>. The intermediate position of the priming loop that we observe in the mixed-initiation state structure correlates well with a previous proposal, based on the anomalously high  $K_m$  for ATP incorporation at template position 5, that a considerable structural transition had to occur at this stage to permit elongation<sup>12</sup>. Hepatitis C (HepC) virus polymerase has a  $\beta$ -hairpin priming loop similar to that of influenza polymerase and an analogous extrusion mechanism has been implicated in the transition to elongation, although a substantially larger rotation ( $\sim 20^\circ$ ) of the thumb domain is observed<sup>21,28</sup>. However deletion of the priming loop was required to obtain the HepC elongation structure<sup>21,28</sup>. Thus, the influenza polymerase structures reported here, to our knowledge, are the first to show successive configurations of an intact priming loop during the transition to elongation (Fig. 2). Whether the extruded priming loop has another functional role during later steps in transcription remains to be elucidated. The priming loop extrusion is reminiscent of the large displacement of the N-terminal domain of T7 RNA polymerase that permits template-product duplex growth and transcript exit<sup>29</sup>. Similarly, in bacterial and eukaryotic RNA polymerases, the 3.2 region of bacterial  $\sigma$ -factor<sup>30,31</sup> or the topologically corresponding B-reader of transcription factor IIB<sup>32</sup>, respectively, have to be displaced in the initiation to elongation transition.

The nine base pair length of the post-translocation template-product RNA duplex in elongating influenza polymerase is the same as for the RNA-DNA hybrid duplex in elongating bacterial<sup>33</sup> and eukaryotic<sup>34</sup> RNA polymerases, whereas for T7 bacteriophage RNA polymerase, the hybrid is slightly shorter<sup>29,35</sup>. Similarly, in these systems, idiosyncratic structural elements that stack against the last base pair of the duplex play an analogous role to the PB2 helical lid in promoting strand separation. In T7 bacteriophage, bacterial and eukaryotic RNA polymerases, thumb domain helices<sup>29</sup>, the 'lid loop'<sup>33</sup> and 'rudder'<sup>34</sup>, respectively, separate the strands of the product-template duplex directing them into distinct exit channels. However, small viral ssRNA RNA-dependent RNA polymerases, such as those from HepC (PDB 4E7A), Norwalk (PDB 3BSO), entero (PDB 5FNG) or polio (PDB 3OL7) viruses, lack such a feature. For these enzymes, the exact

length of the product-template duplex is not established, nor is the mechanism of strand separation.

This work provides the first characterization of the dynamics of influenza polymerase active site during the nucleotide addition cycle. Of particular note is the flexible methionine-rich motif B loop, which is observed in several conformations during the pre- to post-translocation transition (Fig. 4d, Supplementary Fig. 8a). In one position, PB1/Met410 would clash with the base in the +1 template position. We suggest that the motif B loop pushing on the template strand might be part of the mechanism promoting product and template translocation after nucleotide incorporation. At the other end of the template-product duplex, flipping of Tyr207 could act as a shutter during translocation, allowing individual template bases to sequentially enter the exit tunnel.

Our results contribute the first snapshots towards a structure-based, mechanistic understanding of the complete cap-dependent transcription cycle of influenza polymerase, from initiation to termination. Some features, such as priming loop extrusion and strand separation, are also likely to occur in the same way during unprimed replication. By varying the template and capped primer sequence and/or length in combination with various NTPs or nucleoside analogues, our methodology can be extended to capture multiple RNA synthesis intermediates. Structural analysis of such states, by the complementary techniques of X-ray crystallography and cryo-EM, will not only give further insight into the mechanisms of RNA synthesis by influenza polymerase, but should also be valuable for anti-influenza drug development. In this respect, knowledge of the different configurations of the motif B loop could be of particular help in optimizing specific inhibitors, such as nucleoside analogues, targeting the influenza polymerase RNA synthesis active site<sup>36,37</sup>.

## Methods

### Expression and purification

The influenza B/Memphis/13/03 (FluB) polymerase self-cleaving polyprotein heterotrimer construct was expressed in High Five insect cells as described previously<sup>14</sup>. Frozen cell pellets were re-suspended in lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 10% glycerol, pH 8) containing protease inhibitors (Roche, complete mini, EDTA-free). Following lysis by sonication and centrifugation at 20,000 r.p.m. (JA20/Beckman Coulter) for 45 min at 10°C, the supernatant was precipitated by ammonium sulphate (0.5 g ml<sup>-1</sup>) and centrifuged at 45,000 r.p.m. for 45 min at 10°C (45Ti/Beckman Coulter). The pellet was re-dissolved in lysis buffer, and finally re-centrifuged at the same settings. Cleared supernatant was incubated with nickel resin (His60 NiNTA, Clontech) for one hour at 10°C. Protein was eluted with lysis buffer supplemented with 500 mM imidazole and loaded on a Strep-Tactin matrix (Superflow, IBA). Elution was performed with 2.5 mM d-desthiobiotin in low salt buffer (50 mM Tris pH 8, 250 mM NaCl, 10 % glycerol). Pooled FluB polymerase fractions were filtered with 0.22 µm filter and loaded on a heparin column (HiTrap Heparin HP, GE Healthcare). Elution was performed with a gradient using buffers A and B (2 mM TCEP, 50 mM HEPES, pH 7.5, 150 mM (A) or 1 M NaCl (B), 5% glycerol). Homogeneous monomeric polymerase was pooled and dialysed overnight with 6-8 kDa molecular weight cut-off membrane tubing (Spectra/Por, Spectrum Labs) into 50mM HEPES, 500 mM NaCl,

5% glycerol at pH 7. Finally, the protein was concentrated with Amicon® Ultra-15 (50 KDa cutoff), flash-frozen and stored at -80 °C.

### Cap-dependent polymerase transcription assays

Separated synthetic 18-mer or 18+3-mer 3' and 14-mer 5' ends (IBA) were used as vRNA and synthetic 14- or 15-mer capped RNA (TriLink Biotechnologies) as primer (Table S1, Fig. 1a). For the cap-dependent transcription assay, 0.6  $\mu\text{M}$  FluB polymerase, 0.75  $\mu\text{M}$  vRNA, 0.75  $\mu\text{M}$  capped RNA primer, 50  $\mu\text{M}$  GTP, ATP and 2.5 pM  $\alpha\text{-}^{32}\text{P}\text{-ATP}$  were mixed and incubated in reaction buffer (150 mM NaCl, 50 mM HEPES, pH 7.4, 5 mM  $\text{MgCl}_2$  and 2 mM TCEP) at 30 °C for 0.5, 1, 2, 3, 4, 5 and 6 hours. Samples were separated on a 7 M urea, 20% acrylamide gel in TBE buffer, exposed on a storage phosphor screen and read with a Typhoon scanner. 20-mer, 23-mer and 28-mer RNAs, 2-O'-methylated at first 5' ribose adjacent to the cap and radio-labelled at the 3' end were used as markers (Table S1). For quantitative analysis, transcript bands were analyzed in Quantity One software (BIO-RAD).

The elongation intermediate stability was determined by a transcription assay in the presence of heparin as a competitor similarly to as described<sup>18,19</sup>. The elongation intermediate complexes were reconstituted by mixing 0.2  $\mu\text{M}$  FluB polymerase with 0.25  $\mu\text{M}$  of either synthetic 18-mer or 18+3-mer 3', 14-mer 5' end, synthetic 15-mer capped RNA as a primer, and 100  $\mu\text{M}$  GTP, ATP and 2.5 pM  $\alpha\text{-}^{32}\text{P}\text{-ATP}$  in reaction buffer (150 mM NaCl, 50 mM HEPES, pH 7.4, 5 mM  $\text{MgCl}_2$  and 2 mM TCEP). After four hours pre-incubation at 28 °C, either 100  $\mu\text{M}$  CTP or CTP and heparin at 50  $\mu\text{g}/\text{ml}$  were added to the reaction and incubated for an additional 15 minutes. Positive control samples were incubated with either only GTP and ATP or GTP, ATP and CTP for 4 hours and 15 minutes. Samples were separated on an 8 M urea, 20% acrylamide gel in TBE buffer, exposed on a storage phosphor screen and read with a Typhoon scanner. Competitor test experiments were also conducted, demonstrating that the concentration of heparin (50  $\mu\text{g}/\text{ml}$ ) used was sufficient to completely abolish transcription if present in the reaction before the addition of influenza polymerase (Fig. 1d).

### Crystallization, data collection and structure determination

FluB polymerase at 9  $\text{mg ml}^{-1}$  in dialysis buffer was mixed with 40  $\mu\text{M}$  of the vRNA 5' end 14-mer, 40  $\mu\text{M}$  of the vRNA 3' end 21-mer and 80  $\mu\text{M}$  14-mer or 15-mer capped RNA. RNAs used are listed in Table S1. Hanging drops for crystallization were set up at 4 °C. Marquise-shaped crystals growing up to 200  $\mu\text{m}$  in size appeared in two to three weeks in mother liquor containing 200 mM di-ammonium phosphate and 100 mM sodium acetate between pH 4.0 and 4.4. CTP was soaked for 18 hours at a final concentration of 5 mM in the drop. Crystals were cryoprotected with 30 % glycerol in mother liquor and flash-frozen in liquid nitrogen. Diffraction data were collected on beamlines ID30A-1 (MASSIF, wavelength 0.966 Å) and ID29 (wavelength 0.977 Å) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Anisotropic data were integrated with AUTOPROC/STARANISO using an ellipsoidal mask with resolution cut-off criteria local  $(I/\sigma I) = 1.2$ <sup>17</sup> (Table 1). Crystal structures were solved by molecular replacement with PHASER<sup>38</sup>, using the FluB polymerase structure (PDB 5MSG)<sup>11</sup>, rebuilt with COOT<sup>39</sup>, refined using

REFMAC5<sup>40</sup> with TLS parameters and validated with MolProbity<sup>41</sup> (Table 1). For the capped 15-mer structure, 92.1 (0.4) % of residues were in the Ramachandran favoured (outlier) regions, and the Molprobity score and clash score were 1.72 and 1.47, respectively. For the capped 14-mer structure, 94.4 (0.2) % of residues were in the Ramachandran favoured (outlier) regions, and the Molprobity score and clash score were 1.73 and 1.80, respectively. For the capped 14-mer+CTP structure, 92.2 (0.7) % of residues were in the Ramachandran favoured (outlier) regions and the Molprobity score and clash score were 1.80 and 1.99, respectively.

### Cryo-EM of cap-dependent transcribing complexes

To assemble transcribing complexes, 0.8  $\mu\text{M}$  FluB polymerase, 1  $\mu\text{M}$  5' and 3' vRNA, 1  $\mu\text{M}$  of capped RNA and 50  $\mu\text{M}$  ATP and GTP were incubated for 5 hours at 30 °C in cryo-EM buffer (150 mM NaCl, 50 mM HEPES, pH 7.4, 5 mM MgCl<sub>2</sub> and 2 mM TCEP). The transcription reaction was then transferred to 8 °C and incubated with 50  $\mu\text{M}$  non-hydrolysable cytidine-5'-[( $\alpha,\beta$ )-methylene]triphosphate (CMPCPP) for 10 minutes. Aliquots of 3  $\mu\text{l}$  were applied to glow-discharged Quantifoil R1.2/1.3 Au 300 mesh grids, immediately blotted for 2 s and plunged into liquid ethane using an FEI Vitrobot IV (4 °C, 100% humidity). The grids were loaded into a 300 keV Tecnai Krios (FEI) electron microscope (beamline CM01 at the ESRF) equipped with a K2 Summit direct electron-counting camera positioned after a GIF Quantum energy filter (Gatan). Cryo-EM data were acquired using EPU software (FEI) at a nominal magnification of x165000, with a pixel size of 0.831 Å per pixel. Movies were acquired for 6 seconds at a flux of 8.3 electrons per Å<sup>2</sup> per second, giving a total exposure of ~ 50 electrons per Å<sup>2</sup>. Each movie was fractionated into 40 frames of 150 ms. 4151 movies were acquired at a defocus range from -0.7 to -3.3  $\mu\text{m}$  (Table 2).

### Cryo-EM image processing

All movie frames were aligned using MotionCor2<sup>42</sup> (Supplementary Fig. 3a) and then used for contrast transfer function parameter calculation with Gctf<sup>43</sup>. Initially, particles were selected without a template by Gautomatch (provided by Dr. Kai Zhang, <http://www.mrc-lmb.cam.ac.uk/kzhang>) from a small portion of the data set (~200 movies). The relatively high dose enabled sufficient contrast to allow particle picking even in low defocused movies. This initial small dataset was subjected to reference free 2D-classification using RELION 2.1<sup>44</sup>. Eight representative classes of different views were selected from the two-dimensional averages and used as reference for automatic particle picking for the whole data set by Gautomatch. The resulting ~1,615,000 particles were iteratively subjected to two rounds of 2D-classification (Supplementary Fig. 3b) at a pixel size of 5.16Å and 2.58 Å per pixel, respectively, and particles in classes with poor structural features were removed. The remaining particles (~1,055,000) were then subjected to three-dimensional classifications with image alignment (Supplementary Fig. 4). The first round of 3D-classification was restricted to eight classes and performed using bat FluA polymerase with bound vRNA promoter (PDB entry 4WSB) as a 60 Å low-pass filtered initial model. Classification was done during three rounds of 25 iterations each, using regularization parameter T = 4. During the second and third round, local angular searches were performed at 3.5° and 1.8° to clearly separate structural species. The most abundant 3D-classes were pooled together (~615,000

particles), re-extracted at the pixel size of 0.831 Å per pixel, and a second round of 3D-classification was performed again restricted to eight classes. Four major classes of the full complex of FluB polymerase with all the peripheral domains visible were identified (Supplementary Fig. 4). The final accuracy of the rotational alignment was  $\sim 1^\circ$  and the translational alignment accuracy was  $\sim 0.5$  pixels. In the next step, selected movie frames were aligned using a dose-weighting scheme of MotionCorr<sup>42</sup>. The first frame (of total dose of approximately 1.2 e-/Å<sup>2</sup>) and the last 21 frames (of total dose of  $\sim 26$  e-/Å<sup>2</sup>) were discarded to reduce the total exposure to  $\sim 22.5$  e-/Å<sup>2</sup>, thus limiting the radiation damage in the used frames. Particles from the four 3D-classes of the FluB complex were re-extracted from the newly aligned movies and 3D auto-refined using respective masks in RELION (Supplementary Fig. 4). The results of the 3D auto-refinement were used for further 3D-classification with restricted angular searches into eight classes (Supplementary Fig. 4). The two most well defined 3D-classes, corresponding to the pre-initiation and elongation complexes, were separately 3D auto-refined (Supplementary Fig. 4). The final cryo-EM density maps were generated by the post-processing feature in RELION and sharpened or blurred into MTZ format using CCP-EM<sup>45</sup>. The resolution of the cryo-EM density maps was estimated using the 0.143 gold standard Fourier Shell Correlation (FSC) cut off (Supplementary Fig. 3g). The local resolution was calculated using RELION (Supplementary Fig. 3e,f) and reference-based local amplitude scaling was performed with LocScale<sup>46</sup>. The angular distribution of particles (Supplementary Fig. 3c,d) was calculated using the cryoEF software package<sup>47</sup>.

### Cryo-EM model building and refinement.

The known structure of promoter-bound FluB polymerase (PDB entry 5MSG) was first rigid-body fitted into the cryo-EM density by Molrep<sup>48</sup>. Individual protein sub-domain positions were optimised using the Jigglefit tool<sup>49</sup> in COOT<sup>39</sup>. The vRNA promoter and capped RNA primer atomic models were then manually adapted using COOT. The cryo-EM atomic models were iteratively improved by manual building in COOT and refinement with Phenix real-space-refinement<sup>50</sup>. Validation was performed using the Phenix validation tool and model resolution was estimated at the 0.5 FSC cut off (Supplementary Fig. 3h) (Table 2).

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

We thank EMBL (especially M. Bowler) and ESRF staff for access to X-ray beamlines; the ESRF (especially E. Kandiah), IBS and EMBL staff for access to the ESRF Krios beamline CM01; the EMBL eukaryotic expression and high-throughput crystallisation facilities and the biophysical platform within the EMBL-ESRF-ILL-IBS Partnership for Structural Biology. We thank J. Wandzik (EMBL) for help with transcription experiments and EM-data collection. This work was partly supported by ERC Advanced Grant V-RNA (322586) to SC and ANR grant (ANR-18-CE11-0028) to SC. TK holds a fellowship from the EMBL Interdisciplinary Postdocs (EI3POD) initiative co-funded by Marie Skłodowska-Curie grant agreement No 664726.

## Data Availability

Co-ordinates and structure factors or maps have been deposited in the wwwPDB or EMDB.

FluB polymerase-initiation complex 15-mer primer (X-ray) PDB ID 6QCX

FluB polymerase-initiation complex 14-mer primer (X-ray) PDB ID 6QCW

FluB polymerase-initiation complex 14-mer primer +CTP (X-ray) PDB ID 6QCV

FluB pre-initiation complex with primer (cryo-EM) EMD-4511, PDB ID 6QCS

FluB elongation complex (cryo-EM) EMD-4512, PDB ID 6QCT

## References

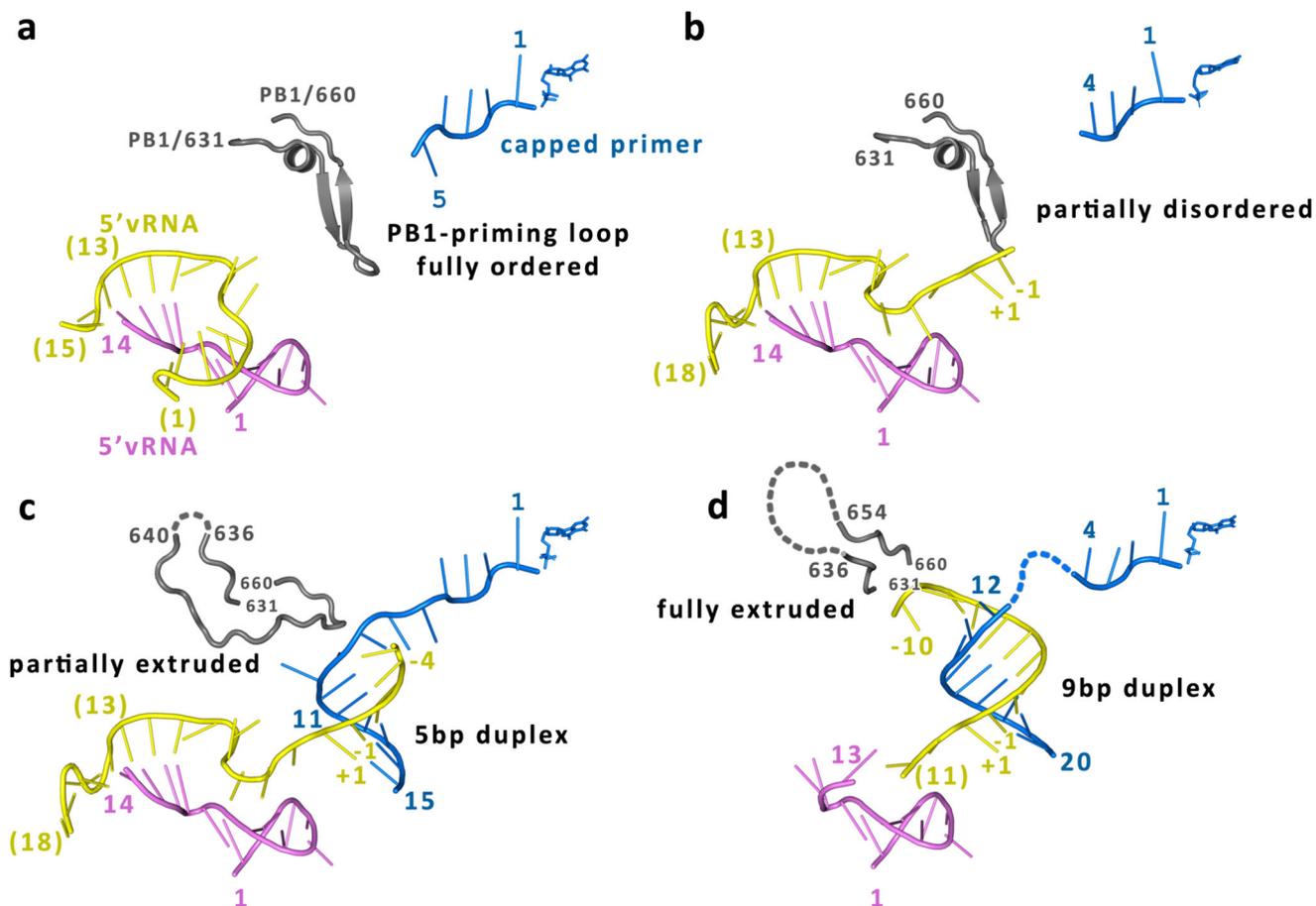
1. Einfeld AJ, Neumann G, Kawaoka Y. At the centre: influenza A virus ribonucleoproteins. *Nature reviews Microbiology*. 2015; 13:28–41. DOI: 10.1038/nrmicro3367 [PubMed: 25417656]
2. Pflug A, Lukarska M, Resa-Infante P, Reich S, Cusack S. Structural insights into RNA synthesis by the influenza virus transcription-replication machine. *Virus Res*. 2017; 234:103–117. DOI: 10.1016/j.virusres.2017.01.013 [PubMed: 28115197]
3. Te Velthuis AJ, Fodor E. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nature reviews Microbiology*. 2016; 14:479–493. DOI: 10.1038/nrmicro.2016.87 [PubMed: 27396566]
4. Zhang J, Hu Y, Musharrafieh RG, Yin H, Wang J. Focusing on the influenza virus polymerase complex: recent progress in drug discovery and assay development. *Current medicinal chemistry*. 2018; doi: 10.2174/0929867325666180706112940
5. Hayden FG, et al. Baloxavir Marboxil for Uncomplicated Influenza in Adults and Adolescents. *The New England journal of medicine*. 2018; 379:913–923. DOI: 10.1056/NEJMoa1716197 [PubMed: 30184455]
6. Plotch SJ, Bouloy M, Ulmanen I, Krug RM. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell*. 1981; 23:847–858. [PubMed: 6261960]
7. Lukarska M, et al. Structural basis of an essential interaction between influenza polymerase and Pol II CTD. *Nature*. 2017; 541:117–121. DOI: 10.1038/nature20594 [PubMed: 28002402]
8. Poon LL, Pritlove DC, Fodor E, Brownlee GG. 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. *Journal of virology*. 1999; 73:3473–3476. [PubMed: 10074205]
9. Pflug A, et al. Capped RNA primer binding to influenza polymerase and implications for the mechanism of cap-binding inhibitors. *Nucleic acids research*. 2018; 46:956–971. DOI: 10.1093/nar/gkx1210 [PubMed: 29202182]
10. Geerts-Dimitriadou C, Zwart MP, Goldbach R, Kormelink R. Base-pairing promotes leader selection to prime in vitro influenza genome transcription. *Virology*. 2011; 409:17–26. DOI: 10.1016/j.virol.2010.09.003 [PubMed: 21051068]
11. Reich S, Guilligay D, Cusack S. An in vitro fluorescence based study of initiation of RNA synthesis by influenza B polymerase. *Nucleic acids research*. 2017; 45:3353–3368. DOI: 10.1093/nar/gkx043 [PubMed: 28126917]
12. Klumpp K, Ford MJ, Ruigrok RW. Variation in ATP requirement during influenza virus transcription. *The Journal of general virology*. 1998; 79(Pt 5):1033–1045. [PubMed: 9603318]
13. Robb NC, et al. Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA. *Nucleic acids research*. 2016; 44:10304–10315. DOI: 10.1093/nar/gkw884 [PubMed: 27694620]

14. Reich S, et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature*. 2014; 516:361–366. DOI: 10.1038/nature14009 [PubMed: 25409151]
15. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature*. 2014; 516:355–360. DOI: 10.1038/nature14008 [PubMed: 25409142]
16. Te Velthuis AJ, Robb NC, Kapanidis AN, Fodor E. The role of the priming loop in influenza A virus RNA synthesis. *Nature microbiology*. 2016; 1:16029. doi: 10.1038/nmicrobiol.2016.29
17. Tickle, IJ, , et al. STARANISO. Global Phasing Ltd. Cambridge; United Kingdom: 2018.
18. Gourse RL. Visualization and quantitative analysis of complex formation between E. coli RNA polymerase and an rRNA promoter in vitro. *Nucleic acids research*. 1988; 16:9789–9809. [PubMed: 3054811]
19. Borukhov S, Sagitov V, Josaitis CA, Gourse RL, Goldfarb A. Two modes of transcription initiation in vitro at the rrnB P1 promoter of Escherichia coli. *The Journal of biological chemistry*. 1993; 268:23477–23482. [PubMed: 8226874]
20. Butcher SJ, Grimes JM, Makeyev EV, Bamford DH, Stuart DI. A mechanism for initiating RNA-dependent RNA polymerization. *Nature*. 2001; 410:235–240. DOI: 10.1038/35065653 [PubMed: 11242087]
21. Appleby TC, et al. Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. *Science*. 2015; 347:771–775. DOI: 10.1126/science.1259210 [PubMed: 25678663]
22. Te Velthuis AJW, Oymans J. Initiation, elongation and realignment during influenza virus mRNA synthesis. *Journal of virology*. 2017; doi: 10.1128/JVI.01775-17
23. Oymans J, Te Velthuis AJW. A Mechanism for Priming and Realignment during Influenza A Virus Replication. *Journal of virology*. 2018; 92 doi: 10.1128/JVI.01773-17
24. te Velthuis AJ. Common and unique features of viral RNA-dependent polymerases. *Cellular and molecular life sciences : CMLS*. 2014; 71:4403–4420. DOI: 10.1007/s00018-014-1695-z [PubMed: 25080879]
25. Shu B, Gong P. Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation. *Proceedings of the National Academy of Sciences of the United States of America*. 2016; 113 E4005–4014 doi: 10.1073/pnas.1602591113 [PubMed: 27339134]
26. Thierry E, et al. Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Molecular cell*. 2016; 61:125–137. DOI: 10.1016/j.molcel.2015.11.016 [PubMed: 26711008]
27. Braam J, Ulmanen I, Krug RM. Molecular model of a eucaryotic transcription complex: functions and movements of influenza P proteins during capped RNA-primed transcription. *Cell*. 1983; 34:609–618. [PubMed: 6616622]
28. Mosley RT, et al. Structure of hepatitis C virus polymerase in complex with primer-template RNA. *Journal of virology*. 2012; 86:6503–6511. DOI: 10.1128/JVI.00386-12 [PubMed: 22496223]
29. Yin YW, Steitz TA. Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science*. 2002; 298:1387–1395. DOI: 10.1126/science.1077464 [PubMed: 12242451]
30. Murakami KS, Masuda S, Darst SA. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science*. 2002; 296:1280–1284. DOI: 10.1126/science.1069594 [PubMed: 12016306]
31. Vassilyev DG, et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*. 2002; 417:712–719. DOI: 10.1038/nature752 [PubMed: 12000971]
32. Kostrewa D, et al. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature*. 2009; 462:323–330. DOI: 10.1038/nature08548 [PubMed: 19820686]
33. Vassilyev DG, Vassilyeva MN, Perederina A, Tahirov TH, Artsimovitch I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*. 2007; 448:157–162. DOI: 10.1038/nature05932 [PubMed: 17581590]
34. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science*. 2001; 292:1876–1882. DOI: 10.1126/science.1059495 [PubMed: 11313499]

35. Liu C, Martin CT. Promoter clearance by T7 RNA polymerase. Initial bubble collapse and transcript dissociation monitored by base analog fluorescence. *The Journal of biological chemistry*. 2002; 277:2725–2731. DOI: 10.1074/jbc.M108856200 [PubMed: 11694519]
36. Jordan PC, Stevens SK, Deval J. Nucleosides for the treatment of respiratory RNA virus infections. *Antiviral chemistry & chemotherapy*. 2018; 26 2040206618764483 doi: 10.1177/2040206618764483 [PubMed: 29562753]
37. Furuta Y, Komeno T, Nakamura T. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*. 2017; 93:449–463. DOI: 10.2183/pjab.93.027
38. McCoy AJ, et al. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40:658–674. DOI: 10.1107/S0021889807021206 [PubMed: 19461840]
39. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography*. 2004; 60:2126–2132. DOI: 10.1107/S0907444904019158 [PubMed: 15572765]
40. Murshudov GN. Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica Section D, Biological crystallography*. 1997; 53:240–255. [PubMed: 15299926]
41. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica Section D, Biological crystallography*. 2010; 66:12–21. DOI: 10.1107/S0907444909042073 [PubMed: 20057044]
42. Zheng SQ, et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods*. 2017; 14:331–332. DOI: 10.1038/nmeth.4193 [PubMed: 28250466]
43. Zhang K. Gctf: Real-time CTF determination and correction. *Journal of structural biology*. 2016; 193:1–12. DOI: 10.1016/j.jsb.2015.11.003 [PubMed: 26592709]
44. Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology*. 2012; 180:519–530. DOI: 10.1016/j.jsb.2012.09.006 [PubMed: 23000701]
45. Burnley T, Palmer CM, Winn M. Recent developments in the CCP-EM software suite. *Acta crystallographica Section D, Structural biology*. 2017; 73:469–477. DOI: 10.1107/S2059798317007859 [PubMed: 28580908]
46. Jakobi AJ, Wilmanns M, Sachse C. Model-based local density sharpening of cryo-EM maps. *eLife*. 2017; 6 doi: 10.7554/eLife.27131
47. Naydenova K, Russo CJ. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat Commun*. 2017; 8:629. doi: 10.1038/s41467-017-00782-3 [PubMed: 28931821]
48. Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta crystallographica Section D, Biological crystallography*. 2010; 66:22–25. DOI: 10.1107/S0907444909042589 [PubMed: 20057045]
49. Brown A, et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta crystallographica Section D, Biological crystallography*. 2015; 71:136–153. DOI: 10.1107/S1399004714021683 [PubMed: 25615868]
50. Afonine PV, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta crystallographica Section D, Structural biology*. 2018; 74:531–544. DOI: 10.1107/S2059798318006551 [PubMed: 29872004]

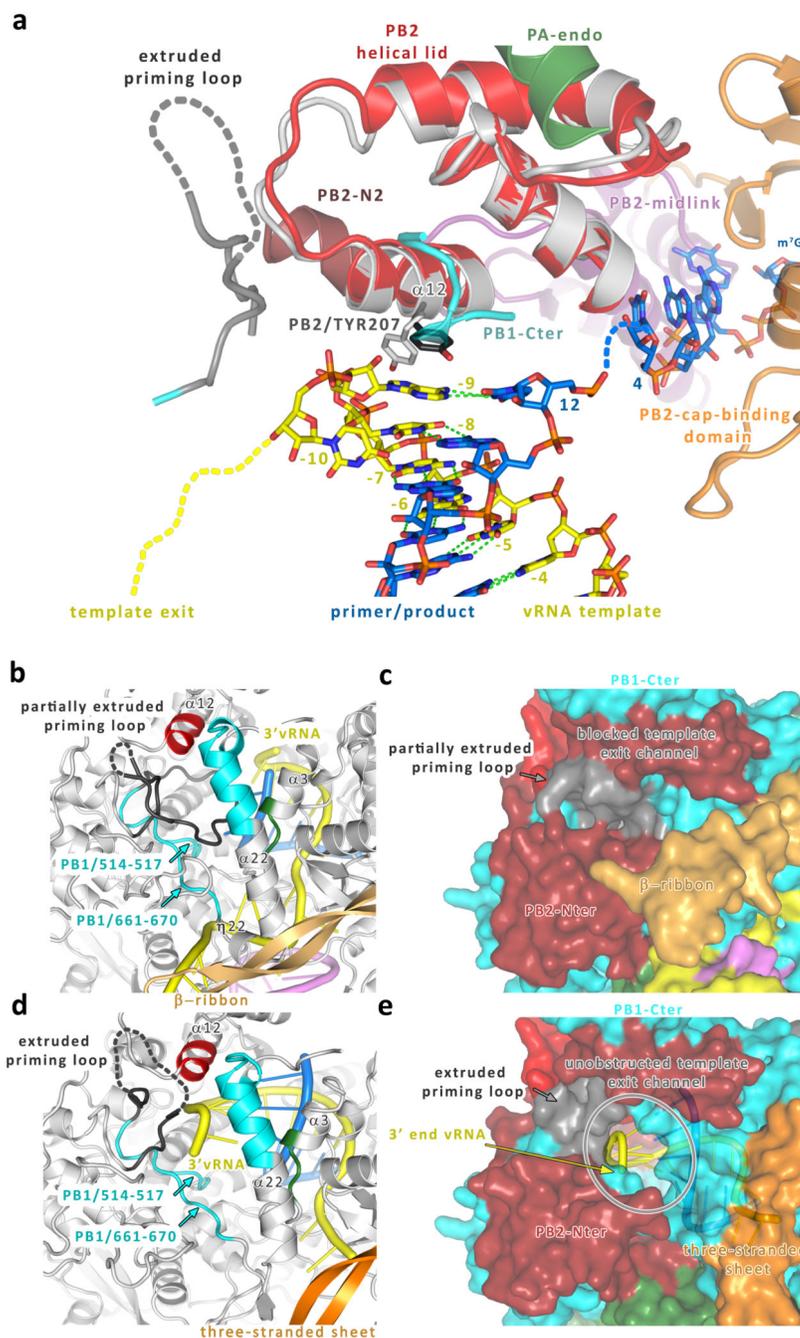


unextended template (bracketed numbers). Nucleotides not visible in the structure are in italics. *Middle*: Structure of the RNA moieties, represented as spheres. *Bottom*: Ribbon diagrams of the pre-initiation (cryo-EM), mixed-initiation (crystal) and elongation (cryo-EM) state complexes coloured according to domain structure<sup>14</sup>. **d**, Elongation intermediate stability and on pathway assay showing that pre-incubated complexes formed with either 18 (left) or 18+3 (right) 3' vRNA templates, together with 15-mer primer and ATP and GTP, are able to continue the transcription reaction upon CTP addition even in the presence of heparin competitor, whereas addition of heparin at the beginning of the reaction completely inhibits RNA synthesis (left-most two lanes in each group). Read-through and endonuclease cleavage products are indicated. M, radiolabelled capped RNA markers, sizes indicated on the left side. Uncropped image is in Supplementary Data Set 1.



**Figure 2. The priming loop extrudes out of the active site in stages during the progression from pre-initiation to elongation state.**

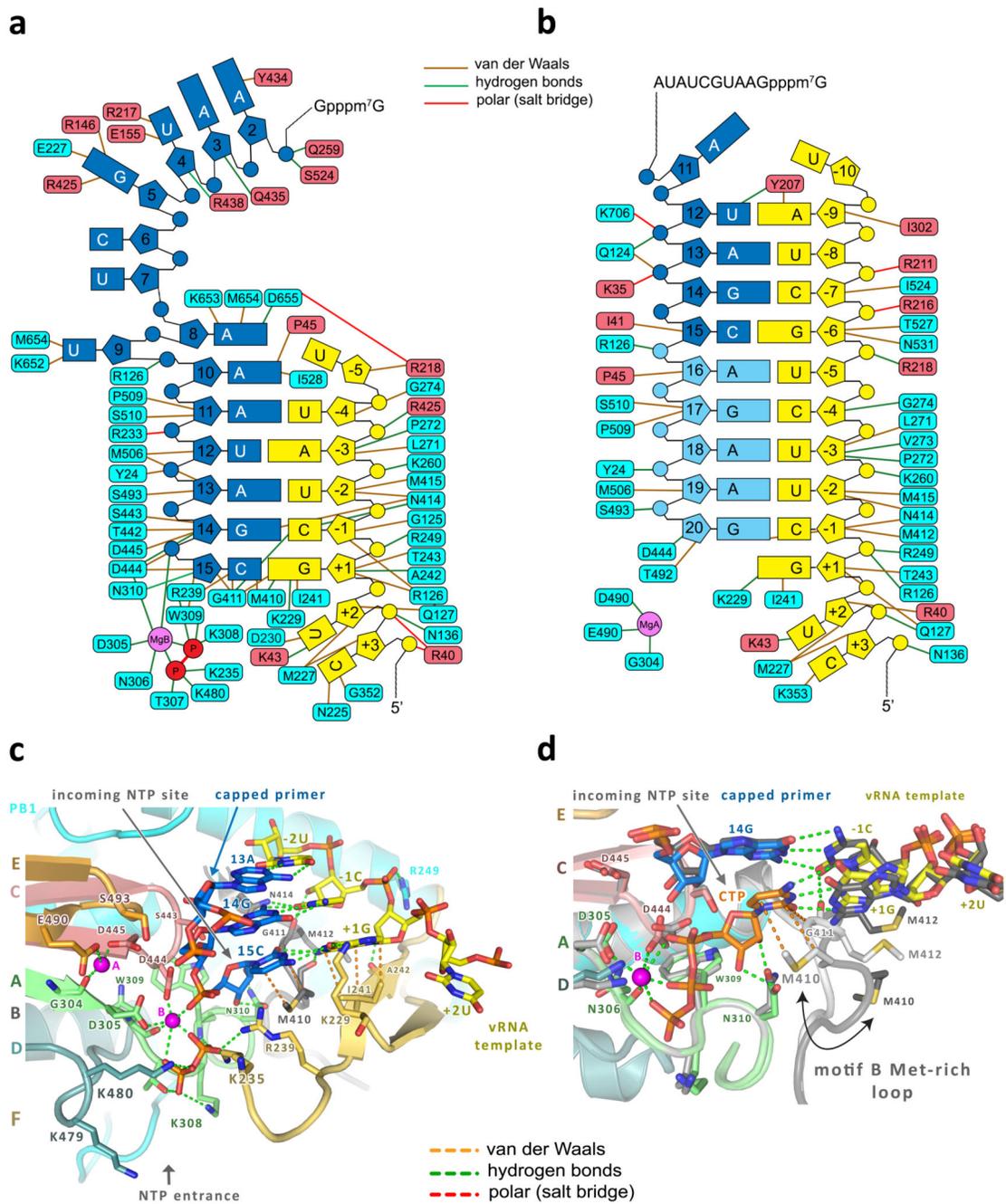
**a**, Pre-initiation state 1 (cryo-EM). The priming loop (grey) is fully ordered in its  $\beta$ -ribbon configuration when neither the template (yellow) nor the primer (blue) are in the active site cavity. Nucleotide numbering as in Fig. 1. **b**, Pre-initiation state 2 (PDB 5MSG<sup>11</sup>). The tip of the priming loop disorders when the template 3' end enters the active site cavity. **c**, Mixed-initiation state (crystal). The partially extruded priming loop accommodates five template-primer base pairs in the active site cavity. **d**, Elongation state (cryo-EM). The priming loop is fully extruded allowing the active site cavity to accommodate an RNA duplex of nine template-product base pairs in the post-translocation state (positions -1 to -9) and ten base pairs in the pre-translocation state (positions +1 to -9).



**Figure 3. Strand separation and template exit channel opening.**

**a.** Superposition of the PB2 helical lid domain in the pre-initiation (light grey) and elongation state (red). The lid is held in place by PB2 N2 domain residues 112-119 (dark red?), PB1 C-terminal residues 656-659 and 699-705 (cyan), the PB1 priming loop (dark grey) and the PA endonuclease residues 89-97 (green). Lid helix  $\alpha 12$  faces the product (blue)-template (yellow) duplex and Tyr207 (black) stacks on the template base of the last base pair, preventing duplex continuation beyond position -9. **b,c.** Cartoon (**b**) and surface (**c**) representations of the initiation state showing that the partially extruded priming loop

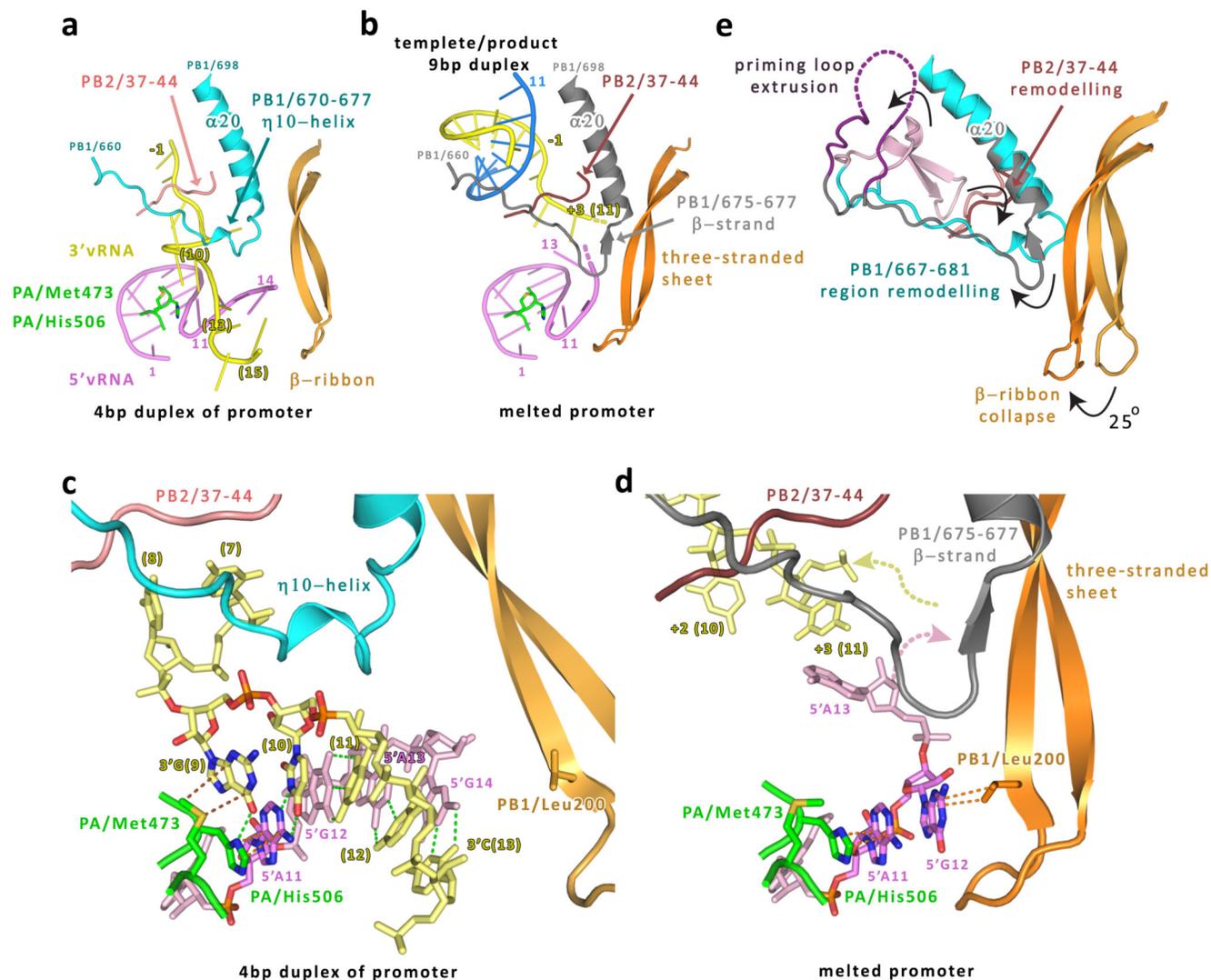
blocks the template exit channel and the five base pair primer-template duplex remains buried in the active site cavity. **d, e.** Cartoon (d) and surface (e) views along the open template exit channel after priming loop extrusion in the elongation state. The template 3' end emerges from the active site cavity upon strand separation. The open channel is lined by parts of the priming loop (dark grey), the following linker PB1/661-670 (cyan), PB1 helix  $\alpha$ 22 (cyan), PB1/514-517 (cyan), the PB2 lid helix  $\alpha$ 22 (red) and PB2/34-37 (green).



**Figure 4. RNA-protein interactions in the active site cavity.**

RNA-protein contacts in the active site cavity of **a**, the pre-translocation mixed-initiation and **b**, the post-translocation elongation states. Protein residues are coloured according to polymerase subunits (PB1:cyan, PB2:ruby). The template (yellow) is numbered with respect to the +1 site, the primer (product) (blue) with respect to the capped 5' end. Interaction type is colour coded as indicated. **c**, View into the PB1 catalytic site showing the conserved RNA-dependent RNA polymerase functional motifs with colour code motif A (green), B (light grey), C (magenta), D (green-cyan), E (orange) and F (orange-yellow). Template (yellow),

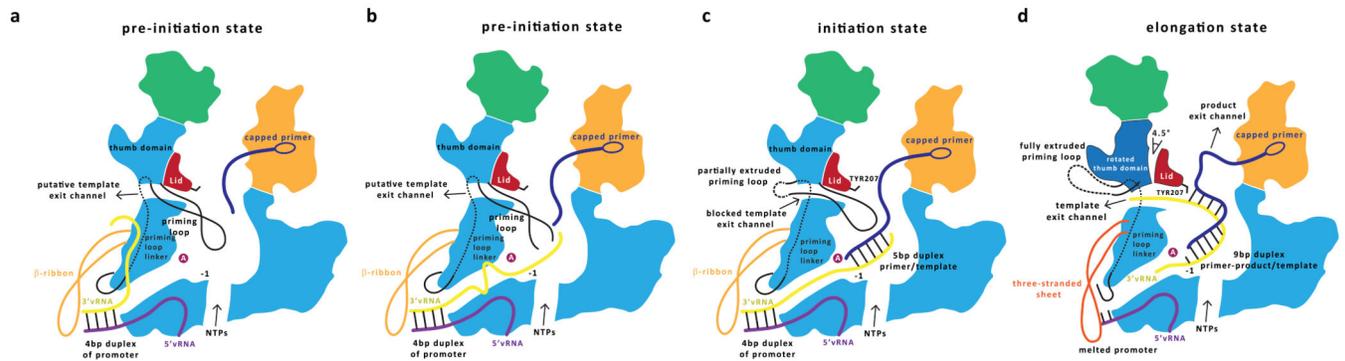
primer (blue), metal B (magenta sphere) and pyrophosphate (orange) are shown as observed in the pre-translocation initiation state. Metal A (magenta sphere), as observed in the post-translocation elongation state, is superimposed. The G+1 template base stacks on Ile241 and makes a polar interaction with Lys229, and its 2' hydroxyl hydrogen bonds to the Ala242 carbonyl oxygen (all motif F residues). The opposite base in the primer (C15) stacks on Met410 (motif B) and its ribose interacts with Asn310 (motif A). **d**, Conformation and interactions of the flexible, methionine-rich motif B loop (PB1/407-413) as observed in occupied (light grey) and vacant (dark grey) active site crystal structures. When incoming CTP (orange) is present, the motif B loop remodels to stabilize the base by stacking with Met410 and the backbone of Gly411. The specific conformation of the motif B loop is further secured by a hydrogen bond interaction of Gly411 carbonyl with the 2' hydroxyl of the template at the -1 position. The +1 NTP position 2' and 3' hydroxyls are sensed by hydrogen bonds to the side-chain of Asn310. When the +1 NTP is absent, Met410 flips into a specific hydrophobic pocket (arrow). See also Supplementary Figure 8.



**Figure 5. Promoter disruption and remodeling of the template entry channel.**

**a**, The promoter 3' (yellow) and 5' (pink) base pairing and sinuous pathway into the active site in the pre-initiation state is stabilized by interactions with PA residues Met473 and His506 (green), the short PB1/670-677  $\eta$ 10-helix (cyan), the PB1  $\beta$ -ribbon (light orange) and PB2/37-44 (salmon). Nucleotide numbering as in Fig. 1. **b**, Upon promoter melting, as observed in the elongation state, the PB1  $\beta$ -ribbon (orange) collapses onto the 5' end vRNA and forms a new three-stranded sheet together with the remodelled  $\eta$ 10-helix residues (grey). PB2/37-44 move to block the position in the initial template channel previously occupied by 3' nucleotides 7-8. **c**, Details of the promoter configuration in the pre-initiation state. The 3' G9 position is stabilized by stacking to PA/Met473 (green) and pseudo-base pairing with PA/His506. 3' A11 packs against PA/Met473. The PB1/670-677  $\eta$ 10-helix stabilizes the tight turn between nucleotides 8 and 9 of the template. **d**, Details of melted promoter state. 5' A11-G12 are sandwiched between PA/His506 and PB1/Leu200 of the collapsed  $\beta$ -ribbon. The paths of the incoming template and outgoing 5' end (arrows) are not defined by the structure. **e**, Superposition of **a** and **b** showing concerted conformational

changes associated with promoter disruption, priming loop extrusion and template-product duplex growth in the active site chamber.



**Figure 6. Schematic of the transitions between the pre-initiation (a,b), initiation (c) and elongation (d) states for transcribing influenza polymerase.**

In the promoter-bound pre-initiation state the template 3' end is either (a) on the polymerase surface or (b) takes a sinuous route into the active site (A in red sphere). Eventually stochastic displacement of the priming loop tip allows primer-templating hybridization and incorporation of the first few nucleotides (c). Template translocation could initially occur by straightening of the template entrance pathway but eventually promoter melting has to occur. The associated collapse of the  $\beta$ -ribbon and closure of the initial template entry tunnel disfavours backtracking, re-formation of the promoter and abortive product formation. NTP incorporation driven translocation provides a counter force to the priming loop, which back extrudes in steps through the active site cavity and out of the template exit channel. Thumb rotation and active cavity opening allow growth of the template-product duplex to the full nine base pairs characteristic of the elongation state before abutting against the PB2 helical lid which forces strand separation (d).

**Table 1**  
**X-ray crystallography data collection and refinement statistics**

	FluB/Memphis polymerase +3' vRNA 1-21 +5' vRNA 1-14 +15-mer capped RNA primer (PDB 6QCX)	FluB/Memphis polymerase +3' vRNA 1-21 +5' vRNA 1-14 +14-mer capped RNA primer (PDB 6QCW)	FluB/Memphis polymerase +3' vRNA 1-21 +5' vRNA 1-14 +14-mer capped RNA primer + CTP (PDB 6QCV)
<b>Data collection <sup>a</sup></b>			
Space group	<i>P</i> 3 <sub>2</sub> 21	<i>P</i> 3 <sub>2</sub> 21	<i>P</i> 3 <sub>2</sub> 21
Cell dimensions??			
<i>a, b, c</i> (Å)	200.4, 200.4, 256.4	200.4, 200.4, 256.5	200.2, 200.2, 256.2
$\alpha, \beta, \gamma$ (°)	90., 90., 90.	90., 90., 90.	90., 90., 90.
Resolution (Å) (anisotropic)	50.0-3.08 (ellipsoidal)	50.0-2.88 (ellipsoidal)	50.0-3.24 (ellipsoidal)
	50.0-3.47 (isotropic)	50.0-3.47 (isotropic)	50.0-3.89 (isotropic)
	(3.27-3.08) <sup>b</sup>	(3.12-2.88) <sup>b</sup>	(3.48-3.24) <sup>b</sup>
<i>R</i> <sub>sym</sub>	0.076 (0.756)	0.177 (1.42)	0.175 (1.19)
<i>R</i> <sub>meas</sub>	0.086 (0.888)	0.185 (1.50)	0.189 (1.28)
<i>I</i> / $\sigma$ ( <i>I</i> )	14.3 (1.7)	8.2 (1.8)	7.6 (1.5)
<i>CC</i> <sub>1/2</sub>	0.999 (0.540)	0.995 (0.677)	0.996 (0.650)
Completeness (%)	75.2 (anisotropic)	60.7 (anisotropic)	67.9 (anisotropic)
Redundancy	4.8 (3.7)	8.2 (10.0)	6.9 (7.3)
<b>Refinement</b>			
Resolution (Å)	3.08-3.47 (anisotropic)	2.88-3.47 (anisotropic)	3.24-3.89 (anisotropic)
No. reflections: work (free)	79129 (4024)	77475 (3942)	60950 (3069)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.202 / 0.226	0.215 / 0.234	0.209 / 0.234
No. atoms	18582	18540	18659
Protein (PA, PB1, PB2)	5760, 5860, 5863	5755, 5843, 5863	5755, 5852, 5863
RNA	1084	1064	1064
Ions	15 (PPi, PO <sub>4</sub> , Mg)	15 (3 x PO <sub>4</sub> )	6 (PO <sub>4</sub> , Mg)
Ligand	-	-	29 (CTP)
<i>B</i> factors (Å <sup>2</sup> ) (all)	123.9	109.9	126.54
Protein (PA, PB2, PB2)	134.5, 111.1, 128.0	121.8, 97.5, 112.5	135.2, 114.2, 131.5
Capped RNA primer	146.6	133.8	160.4
Promoter 3' end	103.8	89.2	106.6
Promoter 5' end	93.0	78.9	94.0
Ions	147.5	130.2	138.6
Ligand	-	-	135.6 (CTP)
R.m.s. deviations			
Bond lengths (Å)	0.002	0.003	0.002
Bond angles (°)	1.141	0.731	1.133

<sup>a</sup>For each structure one crystal was used.

<sup>b</sup>Values in parentheses are for highest-resolution shell.

**Table 2**  
**Cryo-EM data collection, refinement and validation statistics**

	Pre-initiation complex (EMD-4511, PDB 6QCS)	Elongation complex (EMD-4512, PDB 6QCT)
<b>Data collection and processing</b>		
Magnification	165,000	165,000
Voltage (kV)	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	22.5	22.5
Defocus range (µm)	0.7-3.3	0.7-3.3
Pixel size (Å)	0.8311	0.8311
Symmetry imposed	C1	C1
Initial particle images (no.)	1,615,000	1,615,000
Final particle images (no.)	32,640	34,659
Map resolution (Å)	3.16	3.20
FSC threshold	0.143	0.143
Map resolution range (Å)	3.09-5.90	3.05-5.90
<b>Refinement</b>		
Initial model used	PDB 4WSB,5MSG	PDB 4WSB,5MSG
Model resolution (Å)	3.2	3.2
FSC threshold	0.5	0.5
Model resolution range (Å)	3.09-5.90	3.05-5.90
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-63	-62
Model vs map cross correlation	0.86	0.85
<b>Model composition</b>		
Nonhydrogen atoms	18279	18266
Protein residues	2207	2186
RNA residues	33	38
Ligands	3	3
<b><i>B</i> factors (Å<sup>2</sup>)</b>		
Protein	48.36	49.55
RNA	53.04	42.76
Ligand	84.10	84.68
<b>R.m.s. deviations</b>		
Bond lengths (Å)	0.006	0.006
Bond angles (°)	0.750	0.735
<b>Validation</b>		
MolProbity score	2.09	2.14
Clashscore	4.43	4.25
Poor rotamers (%)	5.64	5.87
<b>Ramachandran plot</b>		
Favored (%)	95.68	94.99
Allowed (%)	4.32	5.01

	<b>Pre-initiation complex (EMD-4511, PDB 6QCS)</b>	<b>Elongation complex (EMD-4512, PDB 6QCT)</b>
Disallowed (%)	0	0