Structural basis of poxvirus transcription:
Vaccinia RNA polymerase complexes

Clemens Grimm[1]*, Hauke S. Hillen[2]*, Kristina Bedenk[1], Julia Bartuli[1], Simon Neyer[2], Qian Zhang[4], Alexander Hüttenhofer[5], Matthias Erlacher[5], Christian Dienemann[2], Andreas Schlosser[6], Henning Urlaub[3], Bettina Böttcher[1,6], Aladar A. Szalay[1,4], Patrick Cramer[2]§ and Utz Fischer[1,4,7]§

[1]Department of Biochemistry and Cancer Therapy Research Center (CTRC), Theodor Boveri-Institute, University of Würzburg, Am Hubland, 97074 Würzburg, Germany

[2]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

[3]Bioanalytical Mass Spectrometry Group, Max-Planck-Institute of Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany and Bioanalytics, Institute for Clinical Chemistry, University Medical Center Göttingen, Robert Koch Straße 40, 37075 Göttingen

[4]Genelux Corporation, 3030 Bunker Hill Street, San Diego, CA 92109, USA

[5]Medical University of Innsbruck, Biocenter, Innrain 80/82, 6020 Innsbruck, Austria

[6]Rudolf-Virchow-Center, University of Würzburg, Josef-Schneider-Str. 4, 97080 Würzburg, Germany

[7]Helmholtz Institute for RNA-based Infection Research (HIRI), 97080 Würzburg, Germany

*These authors contributed equally
§ Lead contact.
Correspondence:

utz.fischer@biozentrum.uni-wuerzburg.de; pcramer@mpibpc.mpg.de

SUMMARY

Poxviruses encode a multi-subunit DNA-dependent RNA polymerase (vRNAP) that carries out viral gene expression in the host cytoplasm. We report cryo-EM structures of core and complete vRNAP enzymes from *Vaccinia* virus at 2.8 Å resolution. The vRNAP core enzyme resembles eukaryotic RNA polymerase II (Pol II), but also reveals many virus-specific features, including the transcription factor Rap94. The complete enzyme additionally contains the transcription factor VETF, the mRNA processing factors VTF/CE and NPH-I, the viral core protein E11, and host tRNA^{Gln}. This complex can carry out the entire early transcription cycle. The structures show that Rap94 partially resembles the Pol II initiation factor TFIIB, that the vRNAP subunit Rpo30 resembles the Pol II elongation factor TFIIS, and that NPH-I resembles chromatin remodelling enzymes. Together with the accompanying paper (Hillen et al., this issue of Cell), these results provide the basis for unravelling the mechanisms of poxvirus transcription and RNA processing.

INTRODUCTION

The eukaryotic nucleus contains the machineries for DNA replication and gene transcription. Many viruses rely for their replication and transcription on factors of the host cell and therefore require at least a transient nuclear phase to ensure viral propagation. A remarkable exception amongst eukaryotic DNA viruses are the members of the *Poxviridae* family, whose replication and transcription are confined to the cytoplasm (Moss, 2013). These processes require virus-encoded factors for the production of mature mRNAs from the viral genome. Such cytosolic gene expression events were extensively studied for *Vaccinia* virus, a non-pathogenic prototype of the *Poxviridae* family. These studies uncovered a virus-encoded multisubunit RNA polymerase (vRNAP) and an array of associated factors that ensure the expression of the viral genome (Broyles, 2003; Kates and McAuslan, 1967; Munyon et al., 1967).

Upon infection *Vaccinia* virus enters the cell via micropinocytosis and becomes uncoated (Chi and Liu, 2012; Moss, 2012). Whereas the viral genome is silent in these initial events, all subsequent steps of the replication cycle are dependent on viral transcription and translation processes. Poxviruses coordinate the different processes

of DNA replication and virion formation through timing of expression of individual genes grouped into early, intermediate and late classes (Baldick and Moss, 1993). Accordingly, early genes encode factors involved in events that shortly follow infection, such as viral DNA replication and intermediate gene expression, whereas later processes of the infection cycle, such as virion assembly, require the expression of intermediate and late class gene products.

vRNAP consists of eight subunits encoded by early viral genes and termed according to their apparent molecular masses Rpo147, Rpo132, Rpo35, Rpo30, Rpo22, Rpo19, Rpo18 and Rpo7 (Liu et al., 2010). These subunits show varying degrees of homology to subunits of Pol II, suggesting an evolutionary relationship with the host transcription apparatus (**Table S1**) (Zimmermann et al., 2018). At the level of amino acid residues, the two largest subunits (i.e. Rpo147 and Rpo132) have been reported to be approximately 20% identical to RPB1 and RPB2 of Pol II, respectively (Ahn et al., 1990; Ahn et al., 1992; Amegadzie et al., 1992; Amegadzie et al., 1991; Broyles and Moss, 1986; Knutson and Broyles, 2008). To date, there is no structural information on vRNAPs and their complexes.

vRNAP has the catalytic potential to synthesize RNA in a DNA-dependent manner. However, *in vivo* it requires additional factors in order to become specifically directed to viral early, intermediate and late class genes. Early transcription has been studied most extensively and shown to require the heterodimeric *Vaccinia* early transcription factor (VETF), which interacts with early promoters upstream and downstream of the initiation site (Barnes et al., 2015). Together with Rap94, VETF mediates the recruitment of vRNAP to its promotors and its transition into active elongation (Liu et al., 2010). Rap94 has also been proposed to connect vRNAP with VETF and NPH-I to facilitate termination (Zimmermann et al., 2018). Other virus-encoded proteins are used to add a 5'-terminal m^7G-cap and a 3'-terminal poly(A)-tail to viral RNAs. They include the heterodimeric *Vaccinia* termination factor/capping enzyme (VTF/CE), consisting of subunits D1 and D12, and the termination factor NPH-I, which acts together with a poly(A) polymerase to form polyadenylated 3'-ends. Whether these factors are part of defined functional vRNAP complexes is unknown.

Here we describe the isolation of two distinct vRNAP complexes from human cells infected by *Vaccinia* virus: The ~500 kDa vRNAP core enzyme and the ~900 kDa complete enzyme with six additional viral proteins plus $tRNA^{Gln}$ from the host. We determine the structures of these two complexes by cryo-electron microscopy (cryo-

EM). Whereas the core complex represents the active core RNA polymerase, the complete enzyme apparently represents the packaged machinery containing the factors for early gene transcription. The structures reveal similarities and differences between the viral cytoplasmic transcription apparatus and the nuclear RNA polymerase machinery. Our results form the basis for unravelling the molecular mechanisms of poxvirus gene transcription and RNA processing, and enabled structure determination of functional vRNAP complexes reported in the accompanying paper (Hillen et al., this issue of Cell).

RESULTS

Purification of *Vaccinia* vRNAP complexes

We developed a purification strategy for the isolation of vRNAP complexes based on the recombinant *Vaccinia* virus strain GLV-1h439. This virus is derived from the Vaccinia Lister strain GLV-1h68 and expresses a C-terminally HA/FLAG tagged vRNAP subunit Rpo132 (**Figure S1A**). GLV-1h439 multiplied at similar rates as the untagged parental GLV-1h68 strain upon infection of HeLa cells, suggesting that the tag on Rpo132 does not interfere with the transcriptional activity and replication of the virus (**Figure S1B**).

For affinity-purification of vRNAP, HeLaS3 cells were infected with GLV-1h439. Extract from infected cells was then subjected to purification on an anti-FLAG column and tagged Rpo132, along with its interacting partners, was eluted with FLAG peptide (**Figure S1C**). The eluate was separated by gel electrophoresis (**Figure 1A**) and analysed by mass spectrometry (**Table S2**). All known subunits of the vRNAP core enzyme (**Table S1**) as well as the transcription factor Rap94, the capping enzyme VTF/CE (D1/D12), the termination factor NPH-I, and the early transcription factor subunits VETF-l and VETF-s (A7/D11) were enriched in the GLV-1h439 elution. None of these factors was enriched in a control purification performed with extracts from cells infected with the untagged virus (**Figure 1A**). This purification also identified the viral core protein E11L and the host tRNA^{Gln} (**Table S3**) as new factors associated with the *Vaccinia* virus transcription apparatus.

Vaccinia RNAP complexes are functional

When the eluate was analyzed by sucrose gradient centrifugation and mass spectrometry, two major complexes became apparent. The lighter complex contained

all subunits of the vRNAP core enzyme including sub-stoichiometric amounts of Rap94 (**Figure 1B**). Biochemical characterization showed that this complex represents the catalytically active RNA polymerase core enzyme, as it is capable of elongating an RNA primer *in vitro* (**Figure 1C**). However, no transcriptional activity was detected on an artificial gene under the control of a fully double-stranded viral promoter (**Figure 1D and Figure S1D**), confirming that the core enzyme requires additional factors for initiation.

The second, heavier complex contained all subunits of the core enzyme and additionally VTF/CE, NPH-I, VETF-l, VETF-s, E11L and tRNA^{Gln} (**Figure 1B**). This complex was capable of early promoter-dependent transcription initiation, elongation and termination at a viral termination signal *in vitro* (**Figures 1C and 1D**). Taken together, the first complex represents the catalytically active core vRNAP enzyme, whereas the second complex represents a complete enzyme that comprises core vRNAP and viral transcription and RNA processing factors, and is competent of carrying out all steps of the early *Vaccinia* transcription cycle.

Structure of *Vaccinia* core vRNAP

We analyzed the core vRNAP by single-particle cryo-EM and obtained a reconstruction at 2.8 Å resolution (**Figure S2A-G and Table S4**). The high resolution allowed for placement and adjustment of homology models or *de novo* modelling of all eight subunits. The reconstruction showed additional densities, which were found to stem from Rap94 according to chemical cross-linking (**Figure S2H and Table S6**). Focused classification and refinement yielded improved maps that allowed us to model two domains of Rap94 on opposite sides of the polymerase. The resulting structure of the vRNAP core enzyme has good stereochemical quality (**Table S4)** and contains all eight core vRNAP subunits, four structural zinc ions, the catalytic magnesium ion A, and two domains of Rap94.

The structure shows that core vRNAP resembles multisubunit RNA polymerases in eukaryotic cells, and in particular Pol II (**Figure 2 and Movie S1**). Based on structural and sequence homology, we annotated domains in all subunits in accordance to their counterparts in *S.cerevisiae* Pol II, which serves as a paradigm for eukaryotic multisubunit RNA polymerases (**Figure 2, Figures S3-S5**) (Bond and Schüttelkopf, 2009). The two large subunits, Rpo147 and Rpo132, form two sides of a central cleft that holds the active center, giving vRNAP the typical bilobal appearance

of multisubunit RNA polymerases found in all three domains of life (**Figure 2B**) (Armache et al., 2005). Subunits Rpo35 and Rpo7 form a subassembly on the back of the polymerase body that contacts both large subunits (**Figure 2C**).

The entry path for the DNA duplex to the cleft is lined by two 'jaws' formed by Rpo147 and subunit Rpo22 (**Figure 2C**). Rpo22 assembles with subunits Rpo19 and Rpo18 on the periphery of the polymerase (**Figure 2C**). Rpo18 protrudes slightly from the polymerase body, forming a stalk. At its base, Rpo18 is anchored to the polymerase body and Rpo19, which in turn bridges to Rpo22. Rpo30 is only partially visible in the structure and binds with its N-terminal domain on the outside of the enzyme, near the 'funnel' domain of Rpo147 (**Figure 2B**). The *Vaccinia*-specific transcription factor Rap94 is likewise only partially visible in the core vRNAP structure, with two of its domains (Domain 2 and C-terminal domain) binding to the periphery of the polymerase on opposite sides of the cleft (**Figure 2B**).

vRNAP contains a conserved core

Seven of the eight core vRNAP subunits show structural homology to subunits found in Pol II, albeit their degree of similarity differs (**Figure 3A and Table S1**). Therefore, we carried out a structure-based comparison between vRNAP and *S. cerevisiae* Pol II that provides insight into the functional roles for the individual subunits of vRNAP (**Figure 3 and Figures S3-S5**) (Armache et al., 2003). The two large subunits Rpo147 and Rpo132, which form the body of the polymerase, are highly similar to their Pol II counterparts Rpb1 and Rpb2, respectively (**Figure 3B** and **Figures S3 and S4**). In particular, the active center and nucleic acid-binding regions are structurally conserved. The active site is formed by an invariant DxDxD motif in Rpo147 that binds the catalytic metal ion A (**Figure 2, Movie S1 and Figure S3**) and is flanked by the bridge helix of Rpo147 that traverses the cleft (**Figure 2B**). However, both Rpo147 and Rpo132 lack several regions and are smaller compared to the yeast counterparts (**Figure 3B and Figures S3 and S4**).

In all known multisubunit RNA polymerases, the two large subunits are anchored to a dimeric platform at the back of the enzyme, formed by Rpb3 and Rpb11 in the case of Pol II (Cramer et al., 2001; 2000; Engel et al., 2013; Fernández-Tornero et al., 2013; Hirata et al., 2008; Hoffmann et al., 2015; Zhang et al., 1999). The vRNAP subunit Rpo35 combines features of both Rpb3 and Rpb11 in one polypeptide (**Figure 2C and 3A and Figure S5**). It contains a Rpb3-like N-terminal part and a C-terminal

part that is similar to Rpb11. However, it lacks the zinc binding motif and the regions responsible for interactions with Rpb12 and Rpb10 in Pol II (**Figure S5A**), consistent with the absence of a Rpb12-like subunit in vRNAP. The corresponding location of Rpb12 on vRNAP is instead occupied by a helical insertion in Rpo35. Rpo7 interacts with Rpo35 and closely resembles the Pol II subunit Rpb10, both in structure and location in the enzyme complex (**Figure 2C and S5A**). The C-terminal tail of Rpo7, however, extends further, forming additional interactions with Rpo35 and Rpo132. The Rpo35/Rpo7 subassembly therefore represents the viral equivalent to the Rpb3/10/11/12 subassembly in Pol II and the α_2 homodimer in bacterial RNA polymerases (Zhang et al., 1999).

Rpo22 structurally resembles Rpb5 and is located at a similar position (**Figure 2B and 3A**), as predicted previously (Knutson and Broyles, 2008). Rpo19 is a structural and functional homolog of the Pol II subunit Rpb6. As for the latter, the N-terminal tail of Rpo19 is mobile and hence invisible in our structure (**Figure 2A**). The regions flanking the conserved assembly domain of Rpo19 (α1a and α3) are unique to the viral enzyme. Also, helix α1a forms a contact to Rpo22 that is not observed between the corresponding Pol II subunits Rpb5 and Rpb6 (**Figure 2C and Figure S5B**). The foot domain of Rpo147 lacks some regions found in its Pol II counterpart, and this space is partially occupied by the Rpo19 α1a helical insertion (**Figure S3**). In summary, this detailed comparison of vRNAP to Pol II shows that the enzyme core is largely conserved between vRNAP and other multisubunit polymerases.

Vaccinia-specific polymerase periphery

The structure-based comparison also demonstrates that the enzyme surface deviates substantially from that of other multisubunit RNA polymerases (**Figure 3B**). In particular, vRNAP does not contain counterparts to the Pol II surface subunits Rpb4, Rpb8, Rpb9 and Rpb12 (**Figure 3A**). Moreover, differences in related subunits of vRNAP and Pol II also map to the surface of the enzymes (**Figure 3B**). For example, the clamp core domain in the largest subunit is smaller in vRNAP, but larger and involved in transcription factor interactions in Pol II (Bernecky et al., 2017; Martinez-Rucobo et al., 2011; Plaschka et al., 2016). Likewise, the jaw and foot domains in the largest subunit Rpo147 are also smaller. Rpo147 also does not possess the long and repetitive C-terminal domain (CTD) found in its Pol II counterpart Rpb1. Instead, it contains a short C-terminal tail ('C-tail') (res. 1259-1286) (**Figure S3**), which is mobile

in the core vRNAP structure and hence not visible. The second largest subunit Rpo132 lacks several small regions and contains a few insertions compared to its Pol II counterpart Rpb2. It has an extended carboxy-terminal tail ('C-tail') that emerges from the clamp and wraps arounds the polymerase, traversing across subunit Rpo19 and towards the foot domain of Rpo147 (**Figure 2C and Figure S4**).

The jaws of vRNAP, formed by Rpo147 and Rpo22, also show unique features. Whereas the C-terminal assembly domain of Rpo22 is highly conserved, its jaw domain adopts a unique fold (**Figure 2C and 3B**) and lacks the 'TPSA'-motif found in its Pol II counterpart Rpb5 that interacts with downstream DNA (**Figure S5B**) (Bernecky et al., 2016). The opposite side of the jaw, formed by Rpo147, is smaller and adopts a different orientation than in Pol II. Near this domain, the unique viral subunit Rpo30 binds at the rim of the cleft (**Figure 2B and Figure 3B**). Rpo30 does not have a counterpart in Pol II, but its N-terminal domain (NTD) is located in a similar position on the polymerase as the dissociable Pol II elongation factor TFIIS (**Figure 3A**), which Rpo30 has been suggested to functionally resemble based on sequence analysis (Ahn et al., 1990; Hagler and Shuman, 1993).

A prominent unique feature of vRNAP is its one-subunit stalk, formed by Rpo18, which is homologous to the Pol II subunit Rpb7. Eukaryotic nuclear RNA polymerases I, II and III and archeal RNA polymerase all contain a heterodimeric stalk (Armache et al., 2005; Engel et al., 2013; Fernández-Tornero et al., 2013; Hirata et al., 2008; Hoffmann et al., 2015). In Pol II, the stalk is comprised of subunits Rpb4 and Rpb7 (Armache et al., 2003) and is involved in multiple protein interactions with transcription factors during different stages of the transcription cycle (Bernecky et al., 2017; Plaschka et al., 2016; Vos et al., 2018). The overall fold of Rpo18 is virtually identical to Rpb7, except for a smaller C-terminal region (**Figure 2C and S5B**). Rpo18 uses its tip domain to bind the polymerase core with conserved structural elements (**Figure S5B**). The Rpo18 tip domain may restrict movement of the clamp, as proposed for Rpb7 (Armache et al., 2003). In comparison to the Rpb4-Rpb7 stalk, the C-terminal domain of Rpo18 appears tilted towards the polymerase as it protrudes from the enzyme surface (**Figure 3A**). In summary, these comparisons suggest that the surface of vRNAP has evolved specialized features, likely to facilitate interactions with virus-specific transcription factors.

Transcription factor Rap94 spans the vRNAP cleft

The core vRNAP structure contains the poxvirus-specific transcription factor Rap94 bound to the enzyme periphery (**Figure 2B**). Rap94 may be involved in the recognition of early viral promoters (Ahn et al., 1994) and transcription termination (Christen et al., 2008). However, no structural information is available for Rap94 and sequence-based homology searches do not detect substantial homology to any known proteins. The two Rap94 domains resolved in the core vRNAP structure occupy distant locations on the polymerase surface on opposite sides of the cleft. One of these Rap94 domains, which we refer to as Domain 2 (D2), comprises residues 107-292 and binds to the top of the vRNAP clamp, interacting with both Rpo147 and Rpo132 (**Figure 2B and Movie S1**). It is located close to Rpo18 and may stabilize the stalk in the observed orientation. It consists of a β sheet flanked by helical regions on either side and shows no structural similarity to factors known to interact with the clamp of Pol II. The carboxy-terminal domain (CTD) of Rap94 comprises residues 637-795 and is located at the lobe of Rpo132 (**Figure 2B**). The CTD contacts the protrusion domain with a β sheet (res. 661-686). The fold of the Rap94 CTD does not resemble known Pol II transcription factors. The two Rap94 domains are connected via extended linkers that wrap around the polymerase like a belt (**Figure 2A and 2B**). These linkers traverse the binding sites of Pol II subunits that are absent in vRNAP, including the C-ribbon domain of Rpb9 and the Zn-binding motif of Rpb12. The central region of Rap94 (res. 317-587) is not visible in the core vRNAP structure.

Structure of complete *Vaccinia* vRNAP

We next determined the structure of the complete vRNAP that contains additional transcription and RNA processing factors. We collected a cryo-EM dataset from the pooled fractions 15-17 of the gradient shown in **Figure 1B,** which yielded a reconstruction at 2.8 Å resolution (**Figure S6**). The core vRNAP model could be unambiguously docked into the reconstruction with minor adjustments. We also placed a newly determined crystal structure of the E11 core protein (**Figure S7C, Table S5**) into the density. Then we docked the crystal structure of VTF/CE (Kyrieleis et al., 2014). The location of the bound tRNA^{Gln} could also be identified. The remaining density regions were traced *de novo* and included NPH-I, the Rap94 N-terminal domain (NTD) and central region, the Rpo30 C-terminal region, a compact domain of VETF-I comprising residues 365-436 (VETF-I$^{365-436}$, **Figure 4 and Movie S2**), and several linker regions. The refined atomic model displays excellent stereochemistry

(**Table S4**). The complete vRNAP structure comprises 15 polypeptides and tRNA^{Gln}. It adopts an oval-shaped, bilobal structure with overall dimensions of 220 Å x 150 Å x 130 Å (**Figure 4B and Movie S2**). Whereas one lobe is formed by the core vRNAP enzyme, the other lobe contains the additional factors E11, VTF/CE, NPH-I, VETF and Rap94 regions that are not resolved in the core vRNAP structure.

Rap94 forms a bridge between vRNAP core and additional factors

The complete vRNAP structure shows well-defined density for all parts of Rap94, which interacts with bound factors. In addition to the two domains observed in the core vRNAP structure, the NTD (res. 1-94) and the central region (res. 325-580) of Rap94 are well defined. The Rap94 domains are distributed over the entire complex and are connected by extended linker regions (**Figure 4A and 5A)**. Linker 1 (L1; res. 94-107) connects the NTD to Domain 2. Linker 2 (L2; res. 292-325) emerges from Domain 2 next to the Rpo18 stalk and extends towards Rpo19, passing the C-terminal tail of Rpo147 (**Figure 5B**). It then continues along the polymerase dock domain to the back of vRNAP. On the other side of the cleft, Linker 3 (L3; res. 581-637) extends near the wall and protrusion domains of Rpo132, where it traverses the binding site of Rpb12 in Pol II. L3 then extends through a groove formed by the wall and external domains of Rpo132, and to the funnel helices of Rpo147 and the Rap94 CTD (**Figure 5C**).

The N-terminal region of Rap94 interacts with the C-terminal region of NPH-I. Together they fold into a domain-like module that contacts VTF/CE and that we termed the 'CE connector' (CEC). The CEC forms a wedge between the TPase/GTase and the MTase domains of VTF/CE, keeping the two domains apart by 10 Å compared to the VTF/CE crystal structure (**Figure 5D**). A further contact between Rap94 and NPH-I is buttressed by the dimeric E11 core protein (**Figures 5E**). Domain 2 of Rap94 adapts tRNA^{Gln} to the core vRNAP (**Figure 5F**). In contrast to the situation in the core vRNAP structure, the C-tail of Rpo147 is ordered in the complete vRNAP and adopts an extended structure that tethers VTF/CE (**Figure 5B**). Thus, Rap94 is highly modular and serves as a scaffold to assemble the complete vRNAP complex.

The Rap94 central region resembles Pol II initiation factor TFIIB

The central region of Rap94 in the complete vRNAP (res. 325-580) is reminiscent of a large portion of the Pol II initiation factor TFIIB (**Figure 5G**) and we therefore term it 'B-homology region'. It comprises a B-ribbon element (res. 325-371), a B-reader

hairpin (res. 372-385), a B-linker (res. 386-396), and a B-cyclin domain (res. 397-580). In particular, the zinc ribbon fold and the zinc binding site in the B-ribbon are well conserved between Rap94 and TFIIB. However, the N-terminal part of the B-ribbon is formed by two unique helices in Rap94 that participate in Zn coordination via H328 instead of a cysteine. The B-linker and B-reader appear smaller compared to their TFIIB counterparts, but occupy comparable locations between the dock and clamp domains of the polymerase (Sainsbury et al., 2013). The B-cyclin domain of Rap94 corresponds to the N-terminal cyclin domain of TFIIB with respect to its fold and location. Thus, the B-homology region in Rap94 occupies a similar location as TFIIB in Pol II transcription initiation complexes (Plaschka et al., 2016; Sainsbury et al., 2013), suggesting that Rap94 may function like TFIIB during transcription initiation.

Subunit Rpo30 distantly resembles the Pol II elongation factor TFIIS

The structures show that the core vRNAP subunit Rpo30 shares similarities with eukaryotic TFIIS, as suggested based on sequence analysis (Ahn et al., 1990; Hagler and Shuman, 1993). The Rpo30 N-terminal domain (res. 23-130) binds to the rim of the polymerase funnel (**Figure 6A**), at the location occupied by TFIIS domain II on Pol II (Kettenberger et al., 2003; 2004). Despite their similar location, these domains differ in sequence and structure. In particular, the Rpo30 N-terminal domain contains an insertion (res. 50-100) that wraps around the base of the jaw domain and meanders into the cleft towards the trigger loop, a mobile element of the active center (**Figure 6A, inset**). The N-terminal domain of Rpo30 is connected to a linker region that extends to the Rpo147 funnel helices, forming a short single-turn helical segment (**Figure 6A**).

The C-terminal domain of Rpo30 (res. 131-259) shows sequence similarity to domain III of TFIIS, a zinc ribbon that inserts into the polymerase pore to reach the active site of the enzyme (**Figure S7A**) (Kettenberger et al., 2003). This domain is mobile in both of our structures, but it can likely insert into the polymerase pore and reach the vRNAP active site, as observed for domain III of TFIIS (**Figure 6A**) (Kettenberger et al., 2003; 2004). This domain can trigger nucleolytic RNA cleavage at the Pol II active site, and *Vaccinia* vRNAP has been shown to harbour nucleolytic activity and this has been suggested to be conferred by Rpo30 (Hagler and Shuman, 1993). Thus, Rpo30 contains an N-terminal domain that binds to the polymerase in a

manner reminiscent of domain II of TFIIS, and a mobile C-terminal domain that likely uses a TFIIS-like mechanism to trigger RNA cleavage at the vRNAP active site.

Rpo30 places its phosphorylated C-tail in the active center

Rpo30 additionally contains a C-terminal tail (C-tail; res. 200-259) that is not resolved in the core vRNAP structure but is clearly visible in the complete vRNAP structure (**Figure S7A**). This tail inserts into the pore of the polymerase, running past the active site and into the region that is predicted to interact with the DNA-RNA hybrid at the floor of the active center cleft (**Figure 6B**). The interactions that hold the C-tail in place are centered around three phosphorylated SP sequence motifs for which we find clear density peaks that allowed us to obtain an atomic model for this Rpo30 region. Although the function of the Rpo30 C-tail remains unknown, structural superposition with a Pol II elongation complex (Gnatt et al., 2001) show that it would interfere with binding of the DNA-RNA hybrid and thus impair formation of a transcribing complex. In the accompanying paper (Hillen et al., this issue of Cell) we show that the DNA-RNA hybrid indeed binds at the expected position and would clash with the Rpo30 C-tail. This indicates that the Rpo30 C-tail must be displaced for transcription.

Termination factor NPH-I resembles chromatin remodelers

The complete vRNAP structure also contains the *Vaccinia* termination factor NPH-I, consisting of N- and C-terminal domains (N-lobe and C-lobe, respectively). NPH-I is located with its N-lobe near the RNA exit pore of vRNAP (**Figure 4B and 7A**). A structural homology search shows a striking similarity to the chromatin remodeler INO80 of the SNF2 family (Eustermann et al., 2018) (**Figure 7B**), confirming previous predictions (Henikoff, 1993). SNF2 family proteins are ATP-driven motors with two lobes that are connected by one (INO80, **Figure S7B, middle panel**) or two (SNF2, **Figure S7B, right panel**) extended 'brace' helices, and two protrusions that facilitate DNA interactions. The lobes of NPH-I are connected by a single brace helix, and the C-lobe contains the 'protrusion II' found in members of the SNF2 family (**Figure S7B, left panel**). An additional common feature is the surface at the inside of the 'brace' formed by the two helicase domains, which is lined by stretches of conserved amino acid motifs denoted as motif I-VI (**Figure 7B, left panel**). The motif II (Walker B) sequence qualifies NPH-I as a DExH helicase and is strictly conserved over all members of the poxviridae family (Deng and Shuman, 1998). NPH-I additionally

contains a unique C-terminal region (res. 561-639) that contacts the NTD of Rap94 as part of the CEC through multiple interactions, including an inter-protein β–sheet. NPH-I may therefore have evolved from a common ancestor of the SNF2 family and has adapted to its virus-specific function by the acquisition of its C-terminal domain.

Host tRNA^{Gln} is an integral component of the complete vRNAP

A peculiar feature of the complete vRNAP complex is the presence of the host tRNA^{Gln}. RNA sequencing identified the isoacceptor tRNAs GlnTTG and GlnCTG as the predominant species (**Table S3**). We therefore modelled the tRNA as tRNA-GlnTTG (chr17.trna16-GlnTTG, termed tRNA^{Gln}). The binding site of this tRNA molecule is located on the periphery and the acceptor arm points away from the center of the complex (**Figure 4B**). Only weak density could be detected for the acceptor arm of the tRNA, as it is not supported by any protein contacts and hence partially mobile. tRNA^{Gln} contacts Domain 2 of Rap94, which forms a broad interface with the anticodon- and D-arm (**Figure 5F**). This interaction displays no prominent contacts to particular bases in this area and hence does not confer binding specificity. However, the anticodon loop of tRNA^{Gln} is oriented in a manner that it can be specifically read out by the NPH-I N-lobe (**Figure 7C)** and VETF-I$^{365-436}$ (**Figure 7D**), which may confer specificity for tRNA^{Gln}. Due to the many observed interactions of tRNA^{Gln}, it is likely important for the stability of the complete vRNAP complex.

The initiation factor VETF is anchored to complete vRNAP

The *Vaccinia* initiation factor VETF is known to bind promoter DNA up- and downstream of the transcription start site (TSS) during initiation of early transcription (Broyles et al., 1991). In the complete vRNAP structure, we observe a central domain of the large VETF subunit (VETF-I$^{365-436}$). This domain has a novel fold that is stabilized by three disulfide bonds and provides the connection between tRNA^{Gln}, the TPase module of VTF/CE and the Rpo18 stalk of the vRNAP core enzyme (**Figure 7A and 7D**). Although only this domain of the 710 amino acid VETF–I peptide chain is visible in our density, it is likely that the entire heterodimeric protein is anchored to the complex by this means, as we detect VETF-I and VETF-s in stoichiometric amounts in the sucrose gradient peak fraction (**Figure 1B**). Consistent with this, VETF has been described as a stable heterodimer of VETF-I and VETF-s (Broyles and Moss, 1988). It is likely that during promoter recognition there are major rearrangements in the

complete vRNAP that lead to a positioning of mobile VETF regions onto the promoter DNA.

DISCUSSION

Here we present a purification procedure for endogenous *Vaccinia* vRNAP complexes from infected cells and report the first structures of core and complete vRNAP complexes. A comparison to cellular enzymes, in particular eukaryotic Pol II, confirms the common evolutionary origin of multisubunit RNA polymerases and suggests functions of various vRNAP subunits during transcription. Whereas the two large subunits and the active center cleft are generally conserved, peripheral domains, subunits and factors display virus-specific features.

In particular, the viral factor Rap94 associates with vRNAP and contains a central region that resembles the Pol II initiation factor TFIIB and is thus likely involved in transcription initiation. Further, the subunit Rpo30 distantly resembles the Pol II elongation factor TFIIS and likely confers RNA cleavage activity to vRNAP. Such nucleolytic activity appears conserved among multisubunit RNA polymerases and allows for rescue of the transcription machinery in case of backtracking or misincorporation (Fish and Kane, 2002). Whereas the protein that facilitates transcript cleavage is stably associated with Pol I and Pol III (Engel et al., 2013; Fernández-Tornero et al., 2013; Hoffmann et al., 2015; Neyer et al., 2016), Pol II requires the auxiliary factor TFIIS (Kettenberger et al., 2003). A similar function is fulfilled by the transcript cleavage factors GreA and GreB in bacterial transcription (Borukhov et al., 1993; Opalka et al., 2003; Polyakov et al., 1998; Stebbins et al., 1995). Rpo30 also contains a C-terminal tail that is specific to poxviridae and not found in other large DNA viruses (Mirzakhanyan and Gershon, 2017). Phosphorylation of this tail region occurs in packaged virions (Ngo et al., 2016) and it can occupy the vRNAP active site, raising the possibility that this is a regulatory modification. A comparable observation has been made in the apo form of Pol I, in which a peptide region of the largest subunit occupies the active center cleft (Engel et al., 2013; Fernández-Tornero et al., 2013).

A striking feature of vRNAP is the C-terminal tail located on the largest subunit Rpo147. Whereas this tail is flexible in the core vRNAP complex, it binds to the capping enzyme in the complete vRNAP structure. Although structurally not related, the vRNAP C-tail may thus resemble the Pol II CTD with respect to its function in capping enzyme recruitment, although the Pol II CTD more generally acts as an integration hub for

transcription-coupled processes (Harlen and Churchman, 2017; Jasnovidova and Stefl, 2013). The CTD recruits various factors during different phases of transcription in a phosphorylation-dependent manner (Buratowski, 2009; Hsin and Manley, 2012) and is also involved in recruitment of the capping enzyme (Cho et al., 1997; Fabrega et al., 2003; McCracken et al., 1997; Noe Gonzalez et al., 2018). In the accompanying paper, we show that the Rpo147 C-tail acts as a tether and alters structure upon rearrangements in the complete vRNAP complex that accompany the formation of an active co-transcriptional capping complex (Hillen et al., this issue of Cell).

The additional factors observed in the complete vRNAP structure are unique to the viral machinery. Rap94 acts as an integral building block of the complete vRNAP, as it bridges the interaction between the polymerase and the associated factors. Consistent with this, a loss of this factor leads to generation of virions that lack vRNAP (Zhang et al., 1994). Rap94 binds NPH-I and locks VTF/CE away from the vRNAP core. The structural similarity and location of the Rap94 central region to TFIIB hint at a functional role during transcription initiation. Consistent with this, Rap94 domain 2 occupies a position that resembles the location of the initiation factor TFIIE in the Pol II pre-initiation complex (PIC) (Plaschka et al., 2016), and the Rap94 CTD is found at a location that is congruent with that of TFIIF in Pol II initiation complexes (He et al., 2016; Plaschka et al., 2016). Based on its biochemical composition and activity it is likely that the complete vRNAP complex represents a unit that is packaged into viral progenies and used for early viral transcription upon virus entry into a host cell.

Our structures also rationalize known functional data. Antibodies directed against an epitope within the CEC of Rap94 inhibit the formation of the PIC *in vitro* (Mohamed et al., 2002), underlining the importance of Rap94 for transcription initiation. Likewise, mutations and deletions within the NPH-I portion of the CEC inhibit termination without affecting its ATPase activity (Mohamed and Niles, 2000; Piacente et al., 2003). For early transcription termination, a sequence motif in the transcribed mRNA triggers the ATPase activity of the ssDNA helicase NPH-I (Broyles, 2003). It has previously been demonstrated that both Rap94 and VTF/CE are involved in the recognition of the termination motif, which may pause the elongating polymerase (Christen et al., 2008; Luo et al., 1995; Tate and Gollnick, 2015). NPH-I may then cause transcript extrusion from the active site by a 5' to 3' translocase activity on the non-template strand (Hindman and Gollnick, 2016; Tate and Gollnick, 2011). Provided the observed location of the CEC near the putative RNA exit tunnel is relevant for a

termination intermediate, CEC may be involved in the recognition of the termination signal. Finally, our finding that NPH-I structurally resembles chromatin remodeling ATPases supports the forward translocation model of *Vaccinia* transcription termination.

We also identified the homodimeric viral core protein E11 as a stoichiometric component of the complete vRNAP. The structure suggests that E11 is a major contributor to the stability of the complete vRNAP. E11 is a late viral product and two temperature-sensitive mutants have been previously identified to map to its gene (Kato et al., 2008; Wang and Shuman, 1996). One of these, G66R, does not affect the virus morphogenesis but rather leads to the formation of non-infectious viral particles under non-permissive conditions (Wang and Shuman, 1996). According to the crystal structure of E11, this G66R mutant maps to a tight beta-hairpin and is likely to be a structural mutant. Of note, temperature sensitive mutations in VETF-s and Rap94 have been reported to lead to a defect in protein packaging into mature virions (Kane and Shuman, 1992; Li et al., 1994). These findings are consistent with the idea that the complete vRNAP is the unit that is incorporated into viral progenies and initiates early transcription immediately after virus internalization during the infection cycle.

The incorporation of an uncharged host tRNA^{Gln} molecule into a transcription complex is so far unprecedented. The tRNA^{Gln} forms an integral part of the complete vRNAP particle, and a presumed loss of tRNA^{Gln} is therefore likely to destabilize the complete vRNAP complex. These observations suggest that it might be part of a regulatory mechanism to synchronize the *Vaccinia* replication cycle to the metabolic status of the host cell. It is interesting in this regard to note that viral replication critically depends on the amino acid glutamine as the primary energy source (Fontaine et al., 2014). It is hence a possibility that the complete vRNAP forms in the late phase of viral infection when glutamine becomes limiting and uncharged tRNA^{Gln} accumulates.

Vaccinia virus transcription serves as a paradigm for the molecular biology of nucleo-cytoplasmic large DNA viruses, which include poxviruses and the African Swine Fever Virus. Unlike most other viruses, which rely on the host transcription machinery, they utilize a virus-encoded multisubunit RNA polymerase, which contains a conserved core in different virus taxa (Koonin and Yutin, 2001; Mirzakhanyan and Gershon, 2017). The vRNAP structures presented here provide the first structural insight into the transcription machinery of poxviridiae. This provides a framework for future studies aimed at a mechanistic characterization of the viral transcription cycle.

In particular, snapshots of vRNAP initiation, elongation and termination will shed light on the transitions that occur during these processes and decipher the mechanisms by which the virus-specific factors mediate transcription. As a first step in this direction, we provide structures of transcribing and co-transcriptional capping complexes of *Vaccinia* vRNAP in the accompanying paper (Hillen et al., submitted in parallel).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

C.G. processed the complete vRNAP cryo-EM data, built and refined the model, analyzed the model and prepared figures. H.S.H. processed the core vRNAP cryo-EM data, built and refined the model, analyzed the model and prepared figures. K.B. established the purification procedure of the vRNAP complexes. J.B. performed functional vRNAP assays, purified the vRNAP complexes and prepared figures. S.N. prepared cryo-EM samples of core vRNAP and acquired cryo-EM data. A.A.S. and Q.Z. generated the recombinant *Vaccinia* virus strain and provided virus for vRNAP purification. C.D. assisted with cryo-EM data collection and processing. A.S. and H.U. performed MS analysis and XLMS, respectively. A.H. and M.E. performed RNAseq analysis of complete vRNAP. B.B. performed initial cryo-EM analysis of the complete

vRNAP complex and assisted with cryo-EM data collection of the complete vRNAP complex. P.C., A.A.S. and U.F. designed and supervised the project and acquired funding. C.G., H.S.H., P.C. and U.F. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no conflicts of interest. Q.Z. is a full-time employee, A.S. and U.F. were former salaried employees of Genelux Corporation. Q.Z., A.S. and U.F. are company shareholders of Genelux Corporation.

FIGURE LEGENDS

Figure 1: Purification and characterization of *Vaccinia* virus RNA polymerase complexes. See also Figure S1.

(**A**) Purification of Rpo132 and its associated proteins from GLV-1h439 infected cells using anti-FLAG affinity chromatography. Mock purification was performed from cells infected with untagged GLV-1h68. Specific proteins from the GLV-1h493 elution were resolved on SDS-gels and identified by mass spectrometry.

(**B**) Anti-FLAG eluate from cell extracts infected with GLV-1h439 was separated on a 10-30% sucrose gradient and proteins visualized by silver staining on SDS-PAGE.

(**C**) RNA extension assay with a nucleic acid scaffold mimicking an elongation complex transcription bubble.

(**D**) Transcription assay with a linearized pSB24 template containing a *Vaccinia* virus early promoter and early gene termination signal.

Figure 2: Structure of core *Vaccinia* RNAP. See also Figures S2-S5 and Movie S1.

(**A**) Schematic depiction of vRNAP subunits. Functional domains are annotated based on structure-based sequence alignment with *S.cerevisiae* RNA polymerase II (Armache et al., 2005; Cramer et al., 2001). Regions not visible in the core vRNAP structure are shown transparently.

(**B**) Structure of the core *Vaccinia* RNA Polymerase enzyme. The protein is shown in cartoon depiction, with helices depicted as cylinders. Subunits are colored as in (A). The active site metal A and bound structural Zink ions are shown as spheres. **See also Figure S2**.

(**C**) Cartoon depiction of *Vaccinia* RNAP subunits with structural details shown. Rpo147 and Rpo132 domains are colored as indicated in (A). The location of the subunits in the enzyme is indicated schematically.

Figure 3: Comparison of *Vaccinia* RNA polymerase to *S.cerevisiae* Pol II. See also Figures S3-S5.

(**A**) Comparison of subunit composition between core vRNAP and S.cerevisiae Pol II (PDB ID: 1WCM) (Armache et al., 2005). The enzymes are depicated in schematic surface representation. Homologous subunits are indicated in the table and colored accordingly.

(**B**) Detailed comparison of core vRNAP (left) and *S.cerevisiae* Pol II (right) (PDB ID: 1WCM) (Armache et al., 2005). The largely conserved core is depicted as schematic surface in grey and the differing regions are depicted as cartoon. Regions specific to vRNAP are shown in green, regions specific to Pol II in red. Regions located at the back of the enzyme are labelled transparently. **See also Figures S3-S5**.

Figure 4: Structure of the complete vRNAP complex. See also Figure S6.

(**A**) Schematic depiction of the additional Vaccinia transcription factors VTF/CE, VETF-I, E11 and NPH-I contained in the complete vRNAP complex with domains indicated. Rpo30 and Rap94 are also present in the core vRNAP complex.

(**B**) Overview of the complete vRNAP model, color coding as in A. vRNAP is shown in grey. The orientation of the view in the left panel is related to the view in in the left panel of Figure 2B by an approximately 30° rotation counter-clockwise around the viewing axis followed by an approximately 30° rotation counter-clockwise vertical rotation. The protein is shown in cartoon depiction, with helices depicted as cylinders. **See also Figure S6.**

Figure 5: Rap94 and its role in the complete vRNAP complex.

(**A**) Location of Rap94 in the complete vRNAP structure. The whole model is shown as transparent grey solvent accessible surface with Rap94 shown as solid cartoon. The active site metal A is shown as sphere.

(**B**) Details of the Rpo147 C-tail and the Rap94 linker 2 (L2). These two elements are shown in worm mode, the rest of the model as solvent accessible surface. The Rpo147 C-tail was visible as a diffuse corridor in the cryo-EM density and was manually modelled as Cα trace for this figure. The quality of the density for this element did not allow assignment of side chains, therefore this stretch is omitted in the deposited model.

(**C**) The extended Rap94 linker 3 (L3, shown as worm) connects the B-cyclin domain to the CTD and binds into a cleft on the cRNAP core. The model except for Rap94-L3 and the Rpo147 C-tail is shown as solvent accessible surface.

(**D**) Close-up view of the CEC and its interactions with VTF/CE and the NPH-I helicase module. Proteins are shown as cartoon with coloring as in Figure 4.

(**E**) Details of the E11-Rap94 interactions.

(**F**) Details of the Rap94 domain 2 interactions.

(**G**) Comparison of the Rap94 B-homology region (top) to the corresponding elements of yeast TFIIB (PDB ID 4BBR) (Sainsbury et al., 2013) (bottom).

Figure 6: Structure and interactions of subunit Rpo30. See also Figure S7A.

(**A**) Comparison of *Vaccinia* Rpo30 and *S.cerevisiae* TFIIS. The proteins are depicted schematically with domains indicated. The position of Rpo30 on the core vRNAP complex is shown on the left, with the rest of the enzyme shown as transparent surface representation with coloring as in Figure 2A. The position of TFIIS in the Pol II reactivation intermediate complex (PDB ID: 3PO3) (Cheung and Cramer, 2011) is shown on the right, with the rest of the enzyme shown as transparent surface representation.

(**B**) Cross section through the solvent-accessible surface of the complete vRNAP complex model in the area of the active center cleft. The phosphorylated C-terminal tail of Rpo30 is shown in orange as sticks with the phosphate-moieties shown as purple spheres. The Rap94 B-reader is shown as green worm.

Figure 7: Interactions of NPH-I and VETF in the complete vRNAP complex. See also Figure S7B.

(**A**) Location of VETF, NPH-I, E11 and tRNA^{Gln} in complete vRNAP. The whole model is shown as transparent grey solvent accessible surface with the factors shown as solid cartoon models. Color coding as in Figure 4.

(**B**) Details of the NPH-I fold and location of its helicase motifs (left). Comparison to INO80 (right) (PDB 6FHS) (Eustermann et al., 2018). Corresponding regions are coloured identically.

(**C**) Details if the NPH-I interactions with the tRNA anticodon loop.

(**D**) Details of the VETF-I fold and its tRNA interactions. Disulfide bridges are shown as sticks.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1: Purification and activity of vRNAP complexes. Related to Figure 1.

(**A**) Schematic representation of modified *Vaccinia* virus genes. A DNA fragment encoding a HA-FLAG-tag was fused in GLV-1h439 to the 3'-end of A24R, allowing the expression of C-terminally tagged Rpo132.

(**B**) Replication of GLV-1h439 in comparison to its parental virus GLV-1h68. Virus titer was determined for the indicated time points from infected cells and cell culture supernatant, respectively.

(**C**) Schematic representation of the purification strategy.

(**D**) Scheme of the pSB24 template (top) and nucleic-acid scaffold with RNA in red, template DNA in blue, and non-template chain in light pink (bottom) as used for the transcription assays in Figures 1C and 1D.

Figure S2: Structure determination of core vRNAP. Related to Figure 2.

(**A**) Exemplary cryo-EM micrograph of the core vRNAP dataset.

(**B**) The 32 best aligning class averages from unsupervised 2D classification.

(**C**) Cryo-EM processing workflow for structure determination.

(**D**) Focussed classification and refinement workflow for improved local maps.

(**E**) Fourier Shell Correlation (FSC)-plots for cryo-EM reconstructions used.

(**F**) Angular distribution plot for the global reconstruction of core vRNAP.

(**G**) Local resolution estimation for the global reconstruction of core vRNAP as implemented in Relion.

(**H**) Bis(sulfosuccinimidyl)suberate (BS3) crosslinks identified by mass spectrometry used for positioning of Rap94 domains. (Left) Overview of the core vRNAP structure with regions where strong crosslinks occurred indicated. (Indent 1-3) Proteins are shown in cartoon representation with coloring as in Figure 2. Crosslinked lysine residues are shown as sticks. Selected strong crosslinks are shown as lines.

Figure S3: Structure-based sequence alignment of Rpo147 and *S. cerevisae* Rpb1. Related to Figure 3.

(**A**) Schematic depiction of *Vaccinia* Rpo147 and the homologous *S.cerevisiae* Pol II subunit Rpb1 with domains indicated. Insertions and deletions are indicated by

connecting lines, with differing regions shown with dashed lines. Regions with differing fold are indicated by crossed connecting lines.

(**B**) Structure-based sequence alignment with secondary structure elements depicted and colored according to domains as in Figure 2A and 2C. Sheet regions are shown as arrows, helical region as cylinders. Invariant residues are colored in dark blue and conserved residues in light blue. Regions differing in fold are colored in green (vRNAP-specific) and red (Pol II-specific). The alignment was generated with MSAProbs (Liu et al., 2010) within the MPI Bioinformatics Toolkit (Zimmermann et al., 2018), visualized using Aline (Bond and Schüttelkopf, 2009) and manually edited by comparison to the *S.cerevisiae* Pol II structure (PDB 1WCM) (Armache et al., 2005). In Rpo147, helices $\alpha 8$ and $\alpha 9$ in the polymerase clamp core domain are shortened. Helices $\alpha 27$, $\alpha 28$, $\alpha 32$ and $\alpha 34$, which are located in the foot domain of Rpb1, are absent. The jaw domain is substantially reduced, lacking Rpb1 regions 1158-1188 and 1245-1253.

Figure S4: Structure-based sequence alignment of Rpo132 and *S. cerevisae* Rpb2. Related to Figure 3.

(**A**) Schematic depiction of *Vaccinia* Rpo132 and the homologous *S.cerevisiae* Pol II subunit Rpb2 with domains indicated. Insertions and deletions are indicated by connecting lines, with differing regions shown with dashed lines. Regions with differing fold are indicated by crossed connecting lines.

(**B**) Structure-based sequence alignment with secondary structure elements depicted and colored according to domains as in Figure 2A and 2C. Sheet regions are shown as arrows, helical region as cylinders. Invariant residues are colored in dark blue and conserved residues in light blue. Regions differing in fold are colored in green (vRNAP-specific) and red (Pol II-specific). The alignment was generated with MSAProbs (Liu et al., 2010) within the MPI Bioinformatics Toolkit (Zimmermann et al., 2018), visualized using Aline (Bond and Schüttelkopf, 2009) and manually edited by comparison to the *S.cerevisiae* Pol II structure (PDB 1WCM) (Armache et al., 2005). Helices $\alpha 7$ and $\alpha 8$ in the lobe domain are extended in the Rpo132. In the protrusion domain, the region between $\alpha 11$ and $\alpha 12$ differs between the yeast and viral proteins. The most prominent differences are located in the external domains, in particular in the regions between $\beta 16$ and $\beta 17$, $\alpha 16$ and $\alpha 17$, and between $\alpha 19$ and $\beta 24$. The region following $\beta 28$ (res. 784-797), which contacts upstream DNA in the yeast Pol II (Barnes et al., 2015), is reduced and adopts a different conformation in the viral enzyme.

Figure S5: Structure-based sequence alignment of Rpo35, Rpo22, Rpo19, Rpo18 and Rpo7 with corresponding *S. cerevisae* Pol II subunits. Related to Figure 3.

Structure-based sequence alignments with secondary structure elements depicted and colored according to domains as in Figure 3. Sheet regions are shown as arrows, helical region as cylinders. Invariant residues are colored in dark blue and conserved residues in light blue. Regions differing in fold are colored in green (vRNAP-specific) and red (Pol II-specific). The alignment was generated with MSAProbs (Liu et al., 2010) within the MPI Bioinformatics Toolkit (Zimmermann et al., 2018), visualized using Aline (Bond and Schüttelkopf, 2009) and manually edited by comparison to the *S.cerevisiae* Pol II structure (PDB 1WCM) (Armache et al., 2005).

(**A**) Schematic depiction of *Vaccinia* Rpo35 and Rpo7 and the homologous *S.cerevisiae* Pol II subunits Rpb3, Rpb11 and Rpb10 with domains indicated and structure-based sequence alignment between the proteins. Insertions and deletions are indicated by connecting lines, with differing regions shown with dashed lines. Regions with differing fold are indicated by crossed connecting lines. The region resembling the non-conserved domain of Rpb3 responsible for interactions with Rpb10 and Rpb12 is reduced in Rpo35, with the Zn-binding motif lacking altogether.

(**B**) Schematic depiction of *Vaccinia* Rpo22, Rpo19 and Rpo18 and the homologous *S.cerevisiae* Pol II subunits Rpb5, Rpb6 and Rpb7 with domains indicated and structure-based sequence alignments. Depiction as in (A). Like Rpb7, Rpo18 binds to the polymerase core via its K1 helical turn and its tip loop in the amino terminal tip domain. These elements form a wedge between the N-terminal region of Rpo147, the switch 5 region, the Rpo132 anchor, and helix α1 of Rpo19, all of which are conserved between *Vaccinia* and Pol II. The Rpo18 tip domain may therefore restrict movement of the clamp, as proposed for Rpb7 in Pol II (Armache et al., 2003). The C-terminal domain of Rpo19 forms a β-barrel-like structure but appears tilted towards the polymerase body compared to Rpb4/7.

Figure S6: Structure determination of complete vRNAP. Related to Figure 4.

(**A**) Exemplary cryo-EM micrograph of the complete vRNAP complex dataset.

(**B**) Selected Class averages from unsupervised 2D classification in Relion.

(**C**) Cryo-EM processing workflow for structure determination.

(**D**) Local resolution estimates mapped to the cryo EM density isosurface representation.

(**E**) Angular particle orientation map.

(**F**) Fourier Shell Correlation (FSC)-plot.

Figure S7: Sequence alignment of Rpo30 and *S. cerevisiae* TFIIS and structural details of NPH-I and E11. Related to Figure 4, 5 and 6.

(**A**) Structure-based sequence alignment of Rpo30 and *S. cerevisiae* TFIIS with secondary structure elements depicted and colored according to domains as in Figure 6. Sheet regions are shown as arrows, helical region as cylinders. Invariant residues are colored in dark blue and conserved residues in light blue. Regions differing in fold are colored in green (vRNAP-specific) and red (Pol II-specific). The alignment was generated with MSAProbs (Liu et al., 2010) within the MPI Bioinformatics Toolkit (Zimmermann et al., 2018), visualized using Aline (Bond and Schüttelkopf, 2009) and manually edited by comparison to the *S. cerevisiae* Pol II structure (PDB 1WCM) (Armache et al., 2005). The Zink-binding regions are highlighted in pink and the conserved acidic residues of TFIIS that enter the Pol II active site (DEP motif) are highlighted in green.

(**B**) Fold and topology of the E11 crystal structure. Topology (left). Fold and secondary structure elements in cartoon style (right). The two protomers of the homodimer are in orange and yellow, respectively.

(**C**) Comparison of the ATPase domains of NPH-I to those of the chromatin remodelers INO80 (PDB 6FHS) (Eustermann et al., 2018) and SNF2 (from PDB ID 5X0X) (Liu et al., 2017). The characteristic structural elements are color-coded and labelled.

STAR METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Utz Fischer (utz.fischer@biozentrum.uni-wuerzburg.de). All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Male African green monkey kidney fibroblasts (CV-1) were purchased from the American Type Culture Collection (ATCC No. CCL-70). Human HeLa S3 cells (female) were cultured in a 37°C incubator equilibrated with 5% CO2 and 95% humidified atmosphere. Both cell lines were cultured in DMEM (Gibco) supplemented with 10 % FCS and 1% Penicillin/Streptomycin.

METHOD DETAILS

Generation of recombinant Vaccinia virus GLV-1h439

GLV-1h439 was derived from GLV-1h68 with a HA-tag and FLAG-tag inserted at the end of A24R gene (encoding vRNAP subunit Rpo132). For insertion of the HA/FLAG-doubletag, an A24R transfer vector was constructed. DNA fragments (termed A and B), flanking about 500 bps of each side of the insertion site of the A24R gene were first amplified via PCR with primers A24R-5/A23R-tag3 (product A) and A25Ltag-5/A25L-3 (product B) as listed in the **Key Resource Table**. A second round of PCR linked A and B fragments into product C with primers A24R-5 and A25L-3. The PCR product C was cloned into the pCR-Blunt II-TOPO vector using Zero Blunt TOPO PCR cloning Kit (Invitrogen). The resulting construct pCRII-A24Rtag4 was sequence confirmed. A p7.5E-gpt cDNA fragment (*E.coli* xanthine-guanine phosphoribosyltransferase gene under the control of vaccinia 7.5 early promoter), released by Xba I and Pst I restriction digest from the TK transfer vector, was then subcloned into pCRII-A24Rtag4. The gpt selection-expression cassette was located outside the *Vaccinia* virus DNA that directs homologous recombination into the virus genome, allowing for transient dominant selection of vaccinia recombinants (Falkner and Moss, 1990). The final construct A24Rtag-gpt2 was sequence confirmed, and used to make recombinant virus GLV-1h439, with GLV-1h68 as the parental virus.

Viral replication analysis

Replication of recombinant GLV-1h439 and GLV-1h68 was performed using a standard plaque assay(Cotter et al., 2017). HeLa S3 cells were grown in 24-well plates and infected with virus at a multiplicity of infection (MOI) of 1. After incubation for 1 h at 37 °C, medium was replaced by fresh growth medium and samples were collected 2, 24, 48 and 72 h post viral infection (hpi). After three freeze-thaw cycles, lysates were titrated by plaque assay on CV-1 cells. The assay was performed in triplicate and all samples were measured in duplicates.

vRNAP purification

For purification of vRNAP from infected cells, Hela S3 cells were grown in 15-cm plates up to 80-90% of confluence. The cells were infected with purified GLV-1h439 with a MOI of 1.2. After 24 h the cells were pelleted and resuspended in lysis buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1,5 mM MgCl2, 0.5 % [v/v] NP-40, 1 mM DTT, and complete EDTA-free protease inhibitor cocktail [Sigma-Aldrich]). For vRNAP purification, the extract was incubated for 3 h at 4°C with 200 μl anti-FLAG Agarose (Sigma). Beads were washed four times with buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 1.5 mM MgCl2, 0.1 % [v/v] NP-40 and 1 mM DTT and equilibrated with elution buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 1.5 mM MgCl2 and 1 mM DTT). The bead-bound proteins were eluted with 3 x FLAG peptide, resolved in 12 % Bis-Tris gels and visualized by silver staining. For purification of native vRNAP, the eluate of the anti-FLAG column was concentrated to 1 mg/ml and layered on top of a 10% - 30% sucrose gradient and centrifuged for 16 h and 35.000 rpm at 4 °C in a Beckman 60Ti swing-out rotor. Gradient fractions were fractionated manually, separated by SDS-PAGE and proteins visualized by silver staining.

Initiation assay

Plasmid pSB24, containing a G-less cassette downstream of a synthetic vaccinia virus early promoter was generously provided by Dr. Steven Broyles (Purdue University). Construction of the pSB24 vector with Vaccinia virus early termination signal was described in (Luo et al., 1991). Briefly, by standard genetic manipulation, the sequence from BamHI site to HindIII site of the pSB24 was replaced with the duplex oligonucleotides. The insert sequences include three tandem copies of Vaccinia early termination signal. A typical *in vitro* transcription had a volume of a 50μl and contained

40 mM Tris-HCl, pH 7.9, 1 mM DTT, 2mM spermidine, 6 mM MgCl2, 1 mM ATP, 1 mM CTP, 1 mM GTP, 0.1mM UTP, 20 µCi $\alpha[^{32}P]$-UTP [6000 Ci/mmol], 80 µM SAM, 400 ng of NdeI-linearised pSB24 template as well as purified core or complete vRNAP (Luo et al., 1991). The reaction was incubated at 30 °C for the indicated time points before RNA was extracted and precipitated with isopropanol. Transcripts were analysed by denaturing gel electrophoresis and visualised by autoradiography.

Mass spectrometry analysis

For protein identification by in-gel digestion each gel lane was cut into 15 slices. The gel bands were destained with 30 % acetonitrile in 0.1 M NH_4HCO_3 (pH 8.0), shrunk with 100 % acetonitrile, and dried in a vacuum concentrator (Concentrator 5301, Eppendorf, Germany). Digests were performed with 0.1 µg trypsin per gel band overnight at 37 °C in 0.1 M NH_4HCO_3 (pH 8.0). After removing the supernatant, peptides were extracted from the gel slices with 5 % formic acid, and extracted peptides were pooled with the supernatant. Nano LC-MS/MS analyses were performed on an Orbitrap Fusion (Thermo Scientific) equipped with a PicoView Ion Source (New Objective) and coupled to an EASY-nLC 1000 (Thermo Scientific). Peptides were loaded on capillary columns (PicoFrit, 30 cm x 150 µm ID, New Objective) self-packed with ReproSil-Pur 120 C18-AQ, 1.9 µm (Dr. Maisch) and separated with a 30-minute linear gradient from 3% to 30% acetonitrile and 0.1% formic acid and a flow rate of 500 nl/min. Both MS and MS/MS scans were acquired in the Orbitrap analyzer with a resolution of 60,000 for MS scans and 15,000 for MS/MS scans. HCD fragmentation with 35 % normalized collision energy was applied. A Top Speed data-dependent MS/MS method with a fixed cycle time of 3 seconds was used. Dynamic exclusion was applied with a repeat count of 1 and an exclusion duration of 30 seconds; singly charged precursors were excluded from selection. Minimum signal threshold for precursor selection was set to 50,000. Predictive AGC was used with AGC a target value of 2e5 for MS scans and 5e4 for MS/MS scans. EASY-IC was used for internal calibration. Data analysis was performed against UniProt Vaccinia Virus database with PEAKS 8.5 software (Bioinformatics Solution Inc.) with the following parameters: parent mass tolerance: 8 ppm, fragment mass tolerance: 0.02 Da, enzyme: trypsin, variable modifications: oxidation (M), pyro-glutamate (N-term. Q), phosphorylation (STY), carbamidomethylation (C). Results were filtered to 1% PSM-FDR by target-decoy approach.

Cross-linking mass spectrometry (XLMS)

Protein cross-linking of purified complexes and subsequent mass spectrometry was performed as described previously (Vos et al., 2018). Briefly, samples were crosslinked with BS3 (Thermo Fisher Scientific) and incubated for 30 min at 30 °C. The reaction was quenched by adding 100 mM Tris-HCl pH 7.5 and 20 mM ammonium bicarbonate (final concentrations) and incubation for 15 min at 30 °C. Proteins were precipitated with 300 mM sodium acetate pH 5.2 and four volumes of acetone overnight at −20 °C. The protein was pelleted by centrifugation, briefly dried, and resuspended in 4 M urea and 50 mM ammonium bicarbonate. Crosslinked proteins were reduced with DTT and alkylated (Vos et al., 2016). After dilution to 1 M urea with 50 mM ammonium bicarbonate (pH 8.0), the crosslinked protein complex was digested with trypsin in a 1:50 enzyme-to-protein ratio at 37 °C overnight. Peptides were acidified with trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v), desalted on MicroSpin columns (Harvard Apparatus) following manufacturer's instructions and vacuum-dried. Dried peptides were dissolved in 50 µl 30% acetonitrile/0.1% TFA and peptide size exclusion (pSEC, Superdex Peptide 3.2/300 column on an ÄKTAmicro system, GE Healthcare) was performed to enrich for crosslinked peptides at a flow rate of 50 µl min−1. Fractions of 50 µl were collected. Fractions containing the crosslinked peptides (1–1.7 ml) were vacuum-dried and dissolved in 2% acetonitrile/0.05% TFA (v/v) for analysis by LC–MS/MS.

Crosslinked peptides were analyzed as technical duplicates on an Orbitrap Fusion or Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific), coupled to a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific) equipped with an in-house-packed C18 column (ReproSil-Pur 120 C18-AQ, 1.9 µm pore size, 75 µm inner diameter, 30 cm length, Dr. Maisch GmbH). Samples were separated applying the following 58 min gradient: mobile phase A consisted of 0.1% formic acid (v/v), mobile phase B of 80% acetonitrile/0.08% formic acid (v/v). The gradient started with 5% B, increasing to 8% B on Fusion and 15% on Fusion Lumos, within 3 min, followed by 8–42% B and 15–46% B within 43 min accordingly, then keeping B constant at 90% for 6 min. After each gradient the column was again equilibrated to 5% B for 6 min. The flow rate was set to 300 nl min−1. MS1 spectra were acquired with a resolution of 120,000 in the Orbitrap covering a mass range of 380–1580 m/z. Injection time was set to 60 ms and automatic gain control target to 5×105. Dynamic exclusion covered

10 s. Only precursors with a charge state of 3–8 were included. MS2 spectra were recorded with a resolution of 30,000 in the Orbitrap, injection time was set to 128 ms, automatic gain control target to 5 × 104 and the isolation window to 1.6 *m*/*z*. Fragmentation was enforced by higher-energy collisional dissociation at 30%.

Raw files were converted to mgf format using ProteomeDiscoverer 1.4 (Thermo Scientific, signal-to-noise ratio 1.5, 1,000–10,000 Da precursor mass). For identification of crosslinked peptides, files were analyzed by pLink (v. 1.23), pFind group (Yang et al., 2012) using BS3 as crosslinker and trypsin as digestion enzyme with maximal two missed cleavage sites. Carbamidomethylation of cysteines was set as a fixed modification, oxidation of methionines as a variable modification. Searches were conducted in combinatorial mode with a precursor mass tolerance of 5 Da and a fragment ion mass tolerance of 20 p.p.m. The used database contained all proteins within the complex. The false discovery rate was set to 0.01. Results were filtered by applying a precursor mass accuracy of ±10 p.p.m. Spectra of both technical duplicates were combined and evaluated manually.

RNAseq analysis

Libraries were generated from the isolated RNA fraction following the Ion Torrent™ Ion Total RNA-seq kit v2 (Thermo Fisher; Art. No. 4475936) protocol with the following modifications. Before the libraries were generated 40 ng of the gel-purified RNA was digested with 10U of RNAse T1 (Thermos Fisher; Art. No. EN0541) for 1 minute at room temperature. After PCI extraction and ethanol precipitation, the RNA was pre-treated with 5 U Antarctic phosphatase (New England Biolabs; Art. No. M0289) for 30 minutes at 37°C. After heat inactivation at 65°C, the RNA was phosphorylated by 20 U T4 polynucleotide kinase (New England Biolabs; Art. No. M0201) for 60 minutes at 37°C. Adapter ligation was carried out for 16 hours at 16°C followed by an incubation of 10 minutes at 50°C. Reverse transcription (RT) was performed employing SuperScript™ III with incubations at 42°C, 50°C and 55°C for 45, 15 and 10 minutes, respectively. The RT reactions were purified and the cDNA was amplified by Platinum PCR SuperMix High Fidelity. Resulting libraries were sequenced using a Ion Proton (Ion Torrent TM) with high-Q.

Structure determination of core vRNAP

Following sucrose gradient purification, fraction 11 (**Figure 1B**) was diluted 1:50 and concentrated in a Vivaspin concentrator to a concentration of roughly 50 µg/ml to remove the sucrose. For cryo-EM analysis the sample was centrifuged for 2h at 21,000g and diluted 1:1 in a buffer containing 20mM HEPES, pH 7.5, 200mM $(NH_4)_2SO_4$, 1mM $MgCl_2$ and 5mM 2-mercaptoethanol. 4µl of sample were applied to glow discharged UltrAu 2/2 (Quantifoil) grids at 4 °C and 95% humidity in a Vitrobot (FEI Company), blotted for 8.5s at blot force 14 and plunge-frozen in liquid ethane. Cryo-EM data was collected on a Titan Krios G2 electron microscope (FEI Company) operated at 300 kV with a K2 direct electron detection device operated in counting mode (Gatan) and an energy filter (Gatan) set to a slit width of 15 eV. Movie stacks of 39 frames were acquired with a total dose of 55 $e^-/Å^2$ in counting mode at a nominal magnification of 165,000x, corresponding to a calibrated pixel size of 0.81 Å/pixel. Dose weighting and motion correction was performed using MotionCor2 (Zheng et al., 2017). Per-micrograph contrast-transfer function (CTF) estimation was done using Gctf (Zhang, 2016), as implemented in Relion (Scheres, 2012; Zivanov 2018). A subset of 4,065 particles was manually picked from the micrographs and used for reference-free 2D classification in Relion and the resulting class averages were used to generate reference projections. These were then used as templates for automated particle picking using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/).

A total of 479,618 particles were extracted with a box size of 300 pixels in Relion and subjected to reference-free 2D classification followed by initial global 3D refinement using the *B.taurus* Pol II elongation complex structure as reference (EMD 3218) (Bernecky et al., 2016), which yielded a reconstruction at 3.1 Å overall resolution (**Figure S2**). Further 3D classification revealed two distinct states of vRNAP corresponding to 'open' and 'closed' state clefts, similar to the motion observed previously for Pol II (Cramer et al., 2000; 2001). As the two reconstructions did not show any further differences and the closed state class contained more particles, this class was used for further refinement. Per-particle CTF and motion correction was performed on this particle subset using Warp (Tegunov and Cramer, 2018) and CTF and beam tilt refinement was additionally performed using Relion. The resulting final reconstruction from 3D refinement in Relion achieved an overall resolution of 2.8 Å after post-processing with a sharpening B-factor of -79 $Å^2$. This cryo-EM density was of excellent quality, with clear sidechain densities for the majority of the complex and occasional density for bound ions. However, we refrained from modelling ions or

waters, with the exception of the catalytic metal ion A as its location and identity can be inferred from previous crystallographic studies as well as the structural zinc ions which are each complexed by four cysteine or histidine residues. In addition to the well-resolved core, the cryo-EM map showed fragmented densities on either side of the vRNAP cleft which were not of sufficient quality for model building. To improve these regions, soft masks encompassing them were cut out from the global reconstruction that was previously low-pass filtered to 10 Å. Focussed 3D classification using these masks and the particle subset used in the global refinement was then used to identify particle subpopulations with strong occupancy in the desired region. These particle subpopulations were then subjected to focussed 3D refinement, which was initially run without a reference mask until the refinement converged to local searches, from where on the respective mask was provided for alignment of particles within the masked region. Post-processing of these maps was performed in Relion using the same soft masks also used in focussed classification and refinement. This approach yielded improved densities for the previously poorly resolved regions.

The initial model of core vRNAP was constructed by docking homology models of Rpo147 and Rpo132 generated by Swissmodel (Biasini et al., 2014) into the cryoEM density, followed by manual rebuilding of all residues in Coot (Emsley et al., 2010). Subunits Rpo35, Rpo22, Rpo19, Rpo18 and Rpo7 were built *de novo* in Coot. The density for the most distal strands of Rpo18 was weak and improved only moderately upon focussed classification and refinement, thus indicating potential mobility. Subunit Rpo30 was built *de novo* in the improved map obtained by focussed refinement for its binding region. Cross-linking coupled to mass-spectrometry indicated that the initially fragmented densities remaining on either side of the cleft represent Rap94 (**Figure S2H**), and these regions could be built *de novo* after focussed classification and refinement in the respective maps. The Rap94 linker regions L2 and L4 could be partially built *de novo* in the global reconstruction. After fitting of all models, very weak density remained at the back of vRNAP, which corresponds to the B-homology domain of Rap94. Extensive focussed classification and refinements efforts on this region yielded improved maps around the B-ribbon and B-cyclin domains, but these were not of sufficient quality for reliable model building and we thus omitted these parts from the core vRNAP model. In total, the structure contains models for Rpo147 (UniProt B9U1I2; res. 2-207; 217-1268), Rpo132 (UniProt B9U1Q1; res. 8-122; 126-418; 422-448; 458-789; 797-825; 841-1162), Rpo35 (UniProt B9U1R2; res. 3-305), Rpo22

(UniProt B9U1I0; res. 1-184); Rpo19 (UniProt B9U1M4; res. 61-164), Rpo18 (UniProt B9U1K4; res. 2-108; 136-159), Rpo7 (UniProt B9U1G3; res. 2-62), Rpo30 (UniProt B9U1D1; res. 23-62; 67-151) and Rap94 (UniProt B9U1I7; res. 106-134; 160-316; 588-619; 627-650; 655-795). The structure was refined using phenix.real_space_refine (Adams et al., 2010) against a composite map generated from the global refinement map and the focussed refinement map using phenix.combine_focused_maps by weighting the individual parts according to their cross-correlation with the model. To validate this approach, the model was similarly refined against the locally sharpened density obtained during the Relion local resolution estimation, which yielded comparable final results. The final structure displays excellent stereochemistry, as verified by Molprobity (Chen et al., 2010) (**Table S4**).

Figures were created with PyMol (Schrodinger,, LLC, 2015) and UCSF Chimera (Pettersen et al., 2004). Angular distribution plots were created using a tool distributed with Warp (Tegunov and Cramer, 2018). Sequence identity scores in Table S1 were calculated using Ident and Sim (https://www.bioinformatics.org/sms2/ident_sim.html) (Stothard, 2000) with the structure-based alignments as input.

Structure determination of complete vRNAP

Sample were prepared as for the core vRNAP. For cryo-EM data collection, R 1.2/1.3 holey carbon grids (Quantifoil) were glow discharged for 90 s (Plasma Cleaner model PDC-002. Harrick Plasma Ithaca, NY/USA) at medium power and 3.5 µl of C2 sample was applied inside a Vitrobot Mark IV (FEI) at 4°C and 100% relative humidity. The grids were blotted for 3 s and with blot force 5 and plunged into liquid ethane. The Cryo-EM datasets were collected with a Thermo Fisher Titan Krios G3 and a Falcon III camera (Thermo-Fischer). Data was acquired with EPU at 300 keV and a primary magnification of 75,000 (calibrated pixel size 1.0635 Å) in movie-mode with 25 fractions per movie and integrating the electron-signal. The total exposure was 50 e/Å² over an exposure time of 4.5 s with 2 exposures per hole.

Dose-weighted, motion-corrected sums of the micrograph movies were calculated with Motioncorr2 (Zheng et al., 2017). The contrast-transfer function of each micrograph was fitted with CTFFind4 (Rohou and Grigorieff, 2015). An initial set of 1,500 particles was selected manually and subjected to a 2D-Classification in Relion3-beta (Zivanov et al., 2018). 12 reasonable class averages were selected as templates for subsequent automated particle picking within Relion and 256,452 particles were

picked from 2,224 micrographs. The dataset was then cleaned up by four cycles of 2D classification and particle sorting followed by manual selection of classes based on the appearance of their class averages resulting in a final dataset of 190,000 good particles. A subset of 20,000 particles was used to generate an initial model. An initial 3D classification with Relion yielded two major classes which differed obviously in the density for VTF/CE, and were subjected to 3D refinement. The class of the large particle yielded a 3.3 Å reconstruction. A second round of automated particle picking was performed with projections from the reconstruction of the large particle as picking templates and yielded a dataset of 858,702 particles. This dataset was then cleaned up by four cycles of 2D classification and particle sorting followed by manual particle selection resulting in a final dataset of 618,338 good particles. A 3D classification of this dataset yielded only highly similar classes and the reconstruction using the full, unclassified dataset yielded the highest resolution of 2.98 Å. With further per-particle CTF refinement including a per-dataset beam-tilt refinement and per-particle motion-correction ('polishing') within Relion3 we obtained a reconstruction with 2.75 Å resolution.

For model building and refinement, the complete vRNAP density was unambiguously docked with the previously built core vRNAP model, the crystallographic models of VTF/CE (PDB ID 4CKB) (Kyrieleis et al., 2014), the E11 homodimer and bacterial tRNA[Gln] extracted from PDB entry 1GSG (Rould et al., 1989). The residual density for VETF-I[365-436], NPH-I, and Rap94 was assigned and traced manually within Coot (Emsley et al., 2010) with the guidance of secondary structure predictions from PsiPred (Jones, 1999) and XLMS data. The final model was refined with Phenix.real_space_refine including an ADP refinement step. During refinement secondary structure, mild Ramachandran and reference model restraints from the VTF/CE and E11 crystallographic models were imposed. After a further cycle of manual inspection and automated refinement, water molecules were placed with Coot and a final round of refinement with Phenix.real_space_refine was applied.

X-Ray structure determination of E11

Bacterially overexpressed, hexa-histidine tagged E11 protein was bound to Ni-NTA-Agarose, eluted with 200 mM Imidazole and dialysed against TBS. The tag was cleaved with tobacco etch virus protease and a final gel filtration chromatography was performed. Crystals were obtained with the hanging drop vapour diffusion method with

reservoir solution containing 20% PEG 4000. For crystallographic phase determination the crystals were derivatized with sodium ethylmercurithiosalicylate and a SAD experiment was performed at beamline MX1/P13 of the PETRA III storage ring of the Deutsches Elektronen-Synchrotron (DESY). Phasing and initial model building were performed with Phenix.autosol. The model was then refined against a native dataset collected at the same beamline with Phenix.refine and completed manually within Coot. After three more cycles of manual corrections and automated refinement including water placement and TLS refinement, the R-factors converged.

QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

DATA AND CODE AVAILABILITY

Coordinate files for the core and complete vRNAP complex structures were deposited with the Protein Data Bank and are available under accession codes 6RIC and 6RFL, respectively. The cryo-EM density maps for the core and complete vRNAP complexes were deposited with the Electron Microscopy Data Base (EMDB) and are available under accession codes EMD-4888 and EMD-4868, respectively. Coordinate and structure files for the E11 crystal structure were deposited with the Protein Data Bank and are available under accession code 6RFG.

SUPPLEMENTAL VIDEOS, DATA AND TABLES

Movie S1: Structure of core *Vaccinia* RNA polymerase. Related to Figure 2 and Figure 3

Animation of the core vRNAP structure. Proteins are depicted as surface and as cartoon representation, with coloring as in Figure 2. The sub-assemblies of vRNAP are highlighted in identical views as in Figure 2C.

Movie S2: Structural analysis of the complete vRNAP complex. Related to Figures 4-7.

Animation of the core vRNAP structure. Proteins and tRNA are depicted as in Figures 4-7.

Table S1: Subunit composition of vRNAP. Related to Figure 1.

Table S2: Mass Spectrometric identification of vRNAP subunits. Related to Figure 1.

Table S3: Identification of the tRNA species bound to complete vRNAP by RNAseq analysis. Related to Figure 1.

Table S4: Cryo-EM data collection, refinement and validation statistics. Related to Figure 2 and Figure 4.

Table S5: X-Ray data collection, refinement and validation statistics. Related to Figure 4.

Table S6: XL-MS of Rap94 to vRNAP. Related to Figure S2.

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Ahn, B.Y., Gershon, P.D., and Moss, B. (1994). RNA polymerase-associated protein Rap94 confers promoter specificity for initiating transcription of vaccinia virus early stage genes. J. Biol. Chem. *269*, 7552–7557.

Ahn, B.Y., Gershon, P.D., Jones, E.V., and Moss, B. (1990). Identification of rpo30, a vaccinia virus RNA polymerase gene with structural similarity to a eucaryotic transcription elongation factor. Mol. Cell. Biol. *10*, 5433–5441.

Armache, K.-J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. Pnas *100*, 6964–6968.

Armache, K.-J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J. Biol. Chem. *280*, 7131–7134.

Baldick, C.J., and Moss, B. (1993). Characterization and temporal regulation of mRNAs encoded by vaccinia virus intermediate-stage genes. Journal of Virology *67*, 3515–3527.

Barnes, C.O., Calero, M., Malik, I., Graham, B.W., Spåhr, H., Lin, G., Cohen, A.E., Brown, I.S., Zhang, Q., Pullara, F., et al. (2015). Crystal Structure of a Transcribing RNA Polymerase II Complex Reveals a Complete Transcription Bubble. Molecular Cell *59*, 258–269.

Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J.M., and Cramer, P. (2016). Structure of transcribing mammalian RNA polymerase II. Nature *529*, 551–554.

Bernecky, C., Plitzko, J.M., and Cramer, P. (2017). Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp. Nature Publishing Group *24*, 809–815.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research *42*, W252–W258.

Bond, C.S., and Schüttelkopf, A.W. (2009). ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr. D Biol. Crystallogr. *65*, 510–512.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from E. coli. Cell *72*, 459–466.

Broyles, S.S., and Moss, B. (1988). DNA-dependent ATPase activity associated with vaccinia virus early transcription factor. J. Biol. Chem. *263*, 10761–10765.

Broyles, S.S., Li, J., and Moss, B. (1991). Promoter DNA contacts made by the vaccinia virus early transcription factor. J. Biol. Chem. *266*, 15539–15544.

Broyles, S.S. (2003). Vaccinia virus transcription. J. Gen. Virol. *84*, 2293–2303.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Molecular Cell *36*, 541–546.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cheung, A.C.M., and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature *471*, 249–253.

Chi, P.-I., and Liu, H.-J. (2012). Molecular Signaling and Cellular Pathways for Virus Entry. International Scholarly Research Notices *2013*, 1–8.

Cho, E.J., Takagi, T., Moore, C.R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev. *11*, 3319–3326.

Christen, L.A., Piacente, S., Mohamed, M.R., and Niles, E.G. (2008). Vaccinia virus early gene transcription termination factors VTF and Rap94 interact with the U9 termination motif in the nascent RNA in a transcription ternary complex. Virology *376*, 225–235.

Cotter, C.A., Earl, P.L., Wyatt, L.S., and Moss, B. (2017). Preparation of Cell Cultures and Vaccinia Virus Stocks. Curr Protoc Protein Sci *89*, 5.12.1–5.12.18.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science *292*, 1863–1876.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science *288*, 640–649.

Deng, L., and Shuman, S. (1998). Vaccinia NPH-I, a DExH-box ATPase, is the energy coupling factor for mRNA transcription termination. Genes Dev. *12*, 538–546.

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

Engel, C., Sainsbury, S., Cheung, A.C., Kostrewa, D., and Cramer, P. (2013). RNA polymerase I structure and transcription regulation. Nature *502*, 650–655.

Eustermann, S., Schall, K., Kostrewa, D., Lakomek, K., Strauss, M., Moldt, M., and Hopfner, K.-P. (2018). Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. Nature *556*, 386–390.

Fabrega, C., Shen, V., Shuman, S., and Lima, C.D. (2003). Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. Molecular Cell *11*, 1549–1561.

Falkner, F.G., and Moss, B. (1990). Transient dominant selection of recombinant vaccinia viruses. Journal of Virology *64*, 3108–3111.

Fernández-Tornero, C., Moreno-Morcillo, M., Rashid, U.J., Taylor, N.M.I., Ruiz, F.M., Gruene, T., Legrand, P., Steuerwald, U., and Müller, C.W. (2013). Crystal structure of the 14-subunit RNA polymerase I. Nature *502*, 644–649.

Fish, R.N., and Kane, C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. Biochim. Biophys. Acta *1577*, 287–307.

Fontaine, K.A., Camarda, R., and Lagunoff, M. (2014). Vaccinia Virus Requires Glutamine but Not Glucose for Efficient Replication. Journal of Virology *88*, 4366–4374.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876–1882.

Hagler, J., and Shuman, S. (1993). Nascent RNA cleavage by purified ternary complexes of vaccinia RNA polymerase. J. Biol. Chem. *268*, 2166–2173.

Harlen, K.M., and Churchman, L.S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat. Rev. Mol. Cell Biol. *18*, 263–273.

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. Nature *533*, 359–365.

Henikoff, S. (1993). Transcriptional activator components and poxvirus DNA-dependent ATPases comprise a single family. Trends in Biochemical Sciences *18*, 291–292.

Hillen, H.S., Bartuli, J., Grimm, C., Dienemann, C., Bedenk, K. Urlaub, H., Szalay, A., Fischer U. and Cramer P. (2019). Structural basis of Poxvirus transcription: transcribing and capping complexes. This issue of Cell.

Hindman, R., and Gollnick, P. (2016). Nucleoside Triphosphate Phosphohydrolase I (NPH I) Functions as a 5" to 3" Translocase in Transcription Termination of Vaccinia Early Genes. J. Biol. Chem. *291*, 14826–14838.

Hirata, A., Klein, B.J., and Murakami, K.S. (2008). The X-ray crystal structure of RNA polymerase from Archaea. Nature *451*, 851–854.

Hoffmann, N.A., Jakobi, A.J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W.J.H., Sachse, C., and Müller, C.W. (2015). Molecular structures of unbound and transcribing RNA polymerase III. Nature *528*, 231–236.

Hsin, J.-P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. *26*, 2119–2137.

Jasnovidova, O., and Stefl, R. (2013). The CTD code of RNA polymerase II: a structural view. Wiley Interdiscip Rev RNA *4*, 1–16.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195–202.

Kane, E.M., and Shuman, S. (1992). Temperature-sensitive mutations in the vaccinia virus H4 gene encoding a component of the virion RNA polymerase. Journal of Virology *66*, 5752–5762.

Kates, J.R., and McAuslan, B.R. (1967). Poxvirus DNA-dependent RNA polymerase. Pnas *58*, 134–141.

Kato, S.E.M., Moussatche, N., D'Costa, S.M., Bainbridge, T.W., Prins, C., Strahl, A.L., Shatzer, A.N., Brinker, A.J., Kay, N.E., and Condit, R.C. (2008). Marker rescue mapping of the combined Condit/Dales collection of temperature sensitive vaccinia virus mutants. Virology *375*, 213–222.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347–357.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Molecular Cell *16*, 955–965.

Knutson, B.A., and Broyles, S.S. (2008). Expansion of poxvirus RNA polymerase subunits sharing homology with corresponding subunits of RNA polymerase II. Virus Genes *36*, 307–311.

Koonin, E.V., and Yutin, N. (2001). Nucleo-cytoplasmic Large DNA Viruses (NCLDV) of Eukaryotes (Chichester, UK: American Cancer Society).

Kyrieleis, O.J.P., Chang, J., la Peña, de, M., Shuman, S., and Cusack, S. (2014). Crystal structure of vaccinia virus mRNA capping enzyme provides insights into the mechanism and evolution of the capping apparatus. Structure *22*, 452–465.

Li, J., Pennington, M.J., and Broyles, S.S. (1994). Temperature-sensitive mutations in the gene encoding the small subunit of the vaccinia virus early transcription factor impair promoter binding, transcription activation, and packaging of multiple virion components. Journal of Virology *68*, 2605–2614.

Liu, X., Li, M., Xia, X., Li, X., and Chen, Z. (2017). Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. Nature *544*, 440–445.

Liu, Y., Schmidt, B., and Maskell, D.L. (2010). MSAProbs: multiple sequence alignment based on pair hidden Markov models and partition function posterior probabilities. Bioinformatics *26*, 1958–1964.

Luo, Y., Hagler, J., and Shuman, S. (1991). Discrete functional stages of vaccinia virus early transcription during a single round of RNA synthesis in vitro. J. Biol. Chem. *266*, 13303–13310.

Luo, Y., Mao, X., Deng, L., Cong, P., and Shuman, S. (1995). The D1 and D12 subunits are both essential for the transcription termination factor activity of vaccinia virus capping enzyme. Journal of Virology *69*, 3852–3856.

Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C.M., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. Embo J. *30*, 1302–1310.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev. *11*, 3306–3318.

Mirzakhanyan, Y., and Gershon, P.D. (2017). Multisubunit DNA-Dependent RNA Polymerases from Vaccinia Virus and Other Nucleocytoplasmic Large-DNA Viruses: Impressions from the Age of Structure. Microbiol. Mol. Biol. Rev. *81*, e00010–e00017.

Mohamed, M.R., and Niles, E.G. (2000). Interaction between nucleoside triphosphate phosphohydrolase I and the H4L subunit of the viral RNA polymerase is required for vaccinia virus early gene transcript release. J. Biol. Chem. *275*, 25798–25804.

Mohamed, M.R., Christen, L.A., and Niles, E.G. (2002). Antibodies directed against an epitope in the N-terminal region of the H4L subunit of the vaccinia virus RNA polymerase inhibit both transcription initiation and transcription termination, in vitro. Virology *299*, 142–153.

Moss, B. (2012). Poxvirus cell entry: how many proteins does it take? Viruses *4*, 688–707.

Moss, B. (2013). Poxvirus DNA replication. Cold Spring Harb Perspect Biol *5*, a010199–a010199.

Munyon, W., Paoletti, E., and Grace, J.T. (1967). RNA polymerase activity in purified infectious vaccinia virus. Pnas *58*, 2280–2287.

Neyer, S., Kunz, M., Geiss, C., Hantsche, M., Hodirnau, V.-V., Seybert, A., Engel, C., Scheffer, M.P., Cramer, P., and Frangakis, A.S. (2016). Structure of RNA polymerase I transcribing ribosomal DNA genes. Nature *540*, 607–610.

Ngo, T., Mirzakhanyan, Y., Moussatche, N., and Gershon, P.D. (2016). Protein Primary Structure of the Vaccinia Virion at Increased Resolution. Journal of Virology *90*, 9905–9919.

Noe Gonzalez, M., Sato, S., Tomomori-Sato, C., Conaway, J.W., and Conaway, R.C. (2018). CTD-dependent and -independent mechanisms govern co-transcriptional capping of Pol II transcripts. Nat Comms *9*, 3392.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., and Darst, S.A. (2003). Structure and Function of the Transcription Elongation Factor GreB Bound to Bacterial RNA Polymerase. Cell *114*, 335–345.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605–1612.

Piacente, S.C., Christen, L.A., Mohamed, M.R., and Niles, E.G. (2003). Effect of selected mutations in the C-terminal region of the vaccinia virus nucleoside triphosphate phosphohydrolase I on binding to the H4L subunit of the viral RNA polymerase and early gene transcription termination in vitro. Virology *310*, 109–117.

Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. Nature *533*, 353–358.

Polyakov, A., Richter, C., Malhotra, A., Koulich, D., Borukhov, S., and Darst, S.A. (1998). Visualization of the binding site for the transcript cleavage factor GreB on Escherichia coli RNA polymerase. J. Mol. Biol. *281*, 465–473.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Rould, M.A., Perona, J.J., Soll, D., and Steitz, T.A. (1989). Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. Science *246*, 1135-1142.

Sainsbury, S., Niesser, J., and Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. Nature *493*, 437–440.

Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530.

Schrodinger,, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.

Stebbins, C.E., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb, A., and Darst, S.A. (1995). Crystal structure of the GreA transcript cleavage factor from Escherichia coli. Nature *373*, 636–640.

Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. BioTechniques *28*, 1102–1104.

Tate, J., and Gollnick, P. (2011). Role of forward translocation in nucleoside triphosphate phosphohydrolase I (NPH I)-mediated transcription termination of vaccinia virus early genes. J. Biol. Chem. *286*, 44764–44775.

Tate, J., and Gollnick, P. (2015). The role of vaccinia termination factor and cis-acting elements in vaccinia virus early gene transcription termination. Virology *485*, 179–188.

Tegunov, D., and Cramer, P. (2018). Real-time cryo-EM data pre-processing with Warp. bioRxiv 338558.

Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature *560*, 607–612.

Vos, S.M., Pöllmann, D., Caizzi, L., Hofmann, K.B., Rombaut, P., Zimniak, T., Herzog, F., and Cramer, P. (2016). Architecture and RNA binding of the human negative elongation factor. Elife *5*, e14981.

Wang, S.P., and Shuman, S. (1996). A temperature-sensitive mutation of the vaccinia virus E11 gene encoding a 15-kDa virion component. Virology *216*, 252–257.

Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.-X., Chen, H.-F., et al. (2012). Identification of cross-linked peptides from complex samples. Nat. Methods *9*, 904–906.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell *98*, 811–824.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhang, Y., Ahn, B.Y., and Moss, B. (1994). Targeting of a multicomponent transcription apparatus into assembling vaccinia virus particles requires RAP94, an RNA polymerase-associated protein. Journal of Virology *68*, 1360–1370.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., and Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J. Mol. Biol. *430*, 2237–2243.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*, 163.
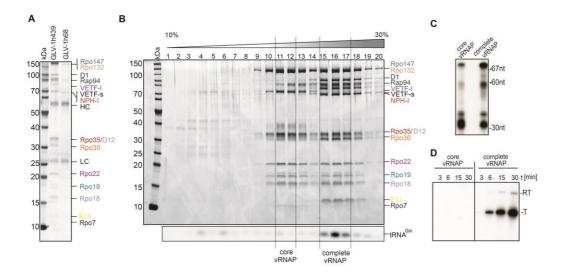
Fig. 1: Purification and characterization of *Vaccinia* virus RNA polymerase complexes
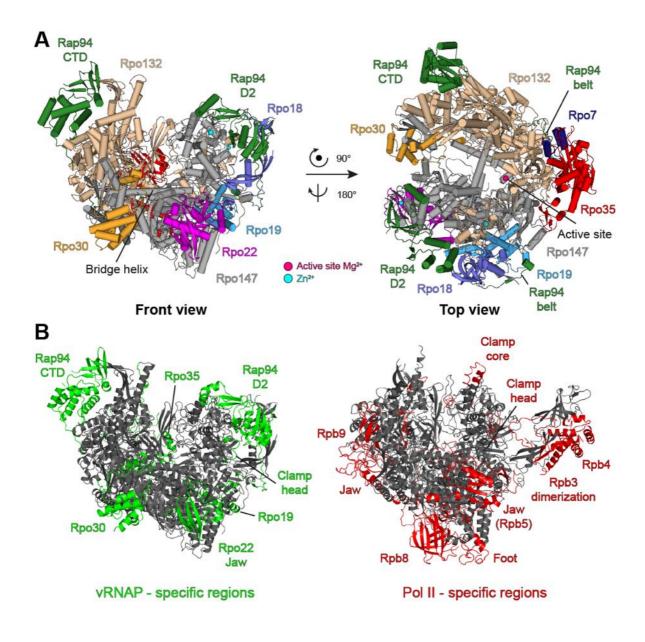
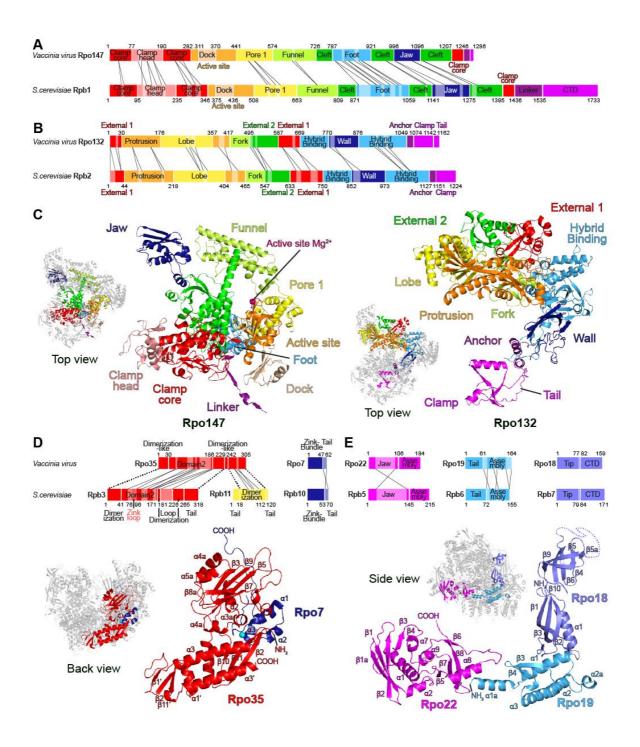**Figure 2: Structure of core Vaccinia RNAP and comparison to RNA Pol II**

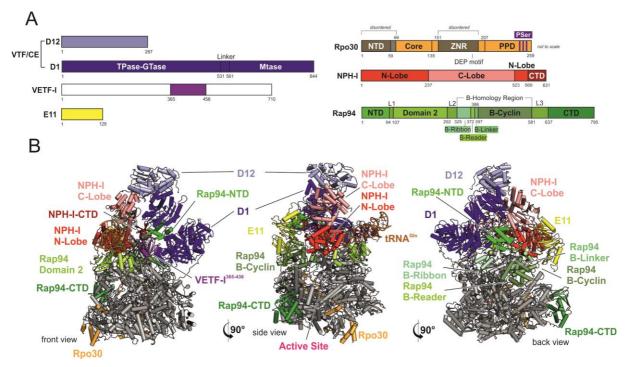**Figure 3: Vaccinia RNA Polymerase subunit composition and comparison to Pol II.**

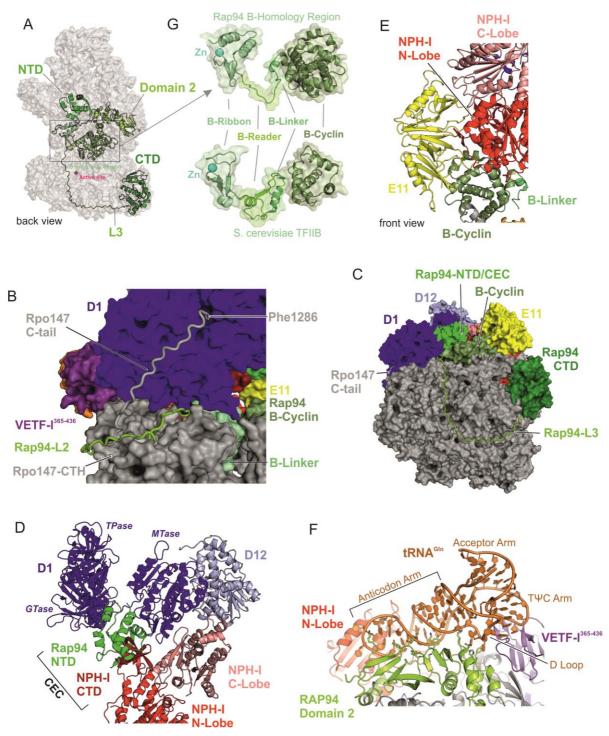**Figure 4: Structure of the complete vRNAP complex.**
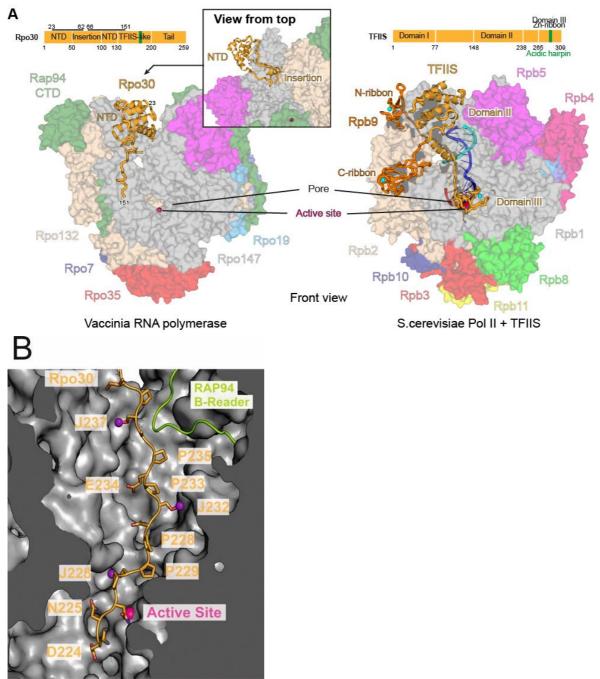
Figure 5: Rap94 and its role in complete vRNAP.
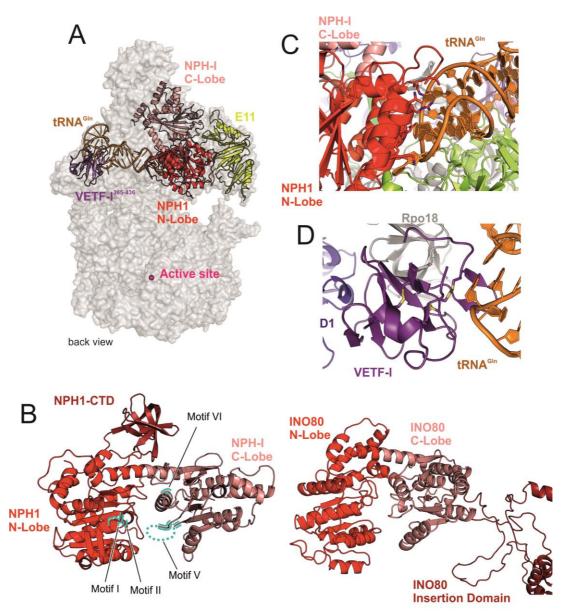
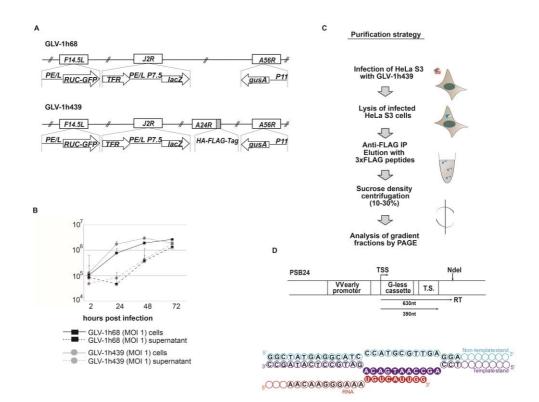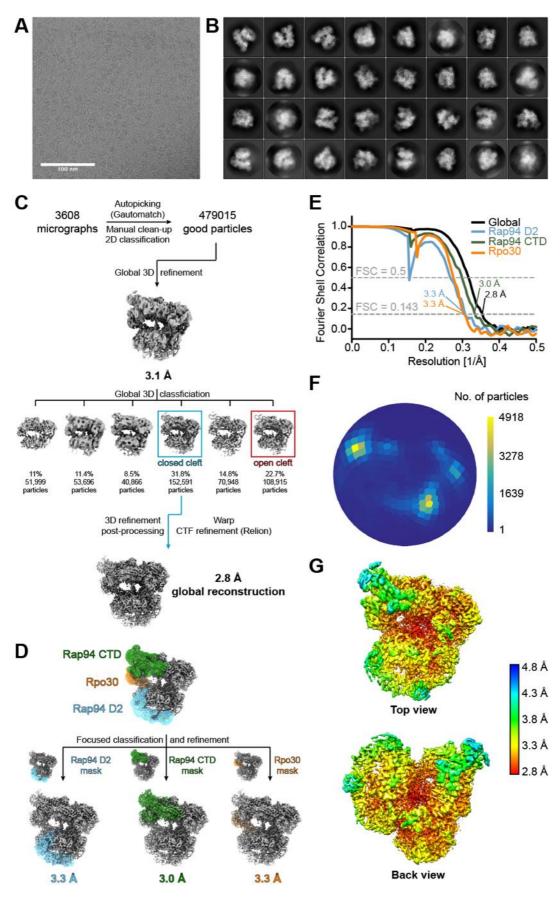**Fig. 6: Structure and interactions of subunit Rpo30.**

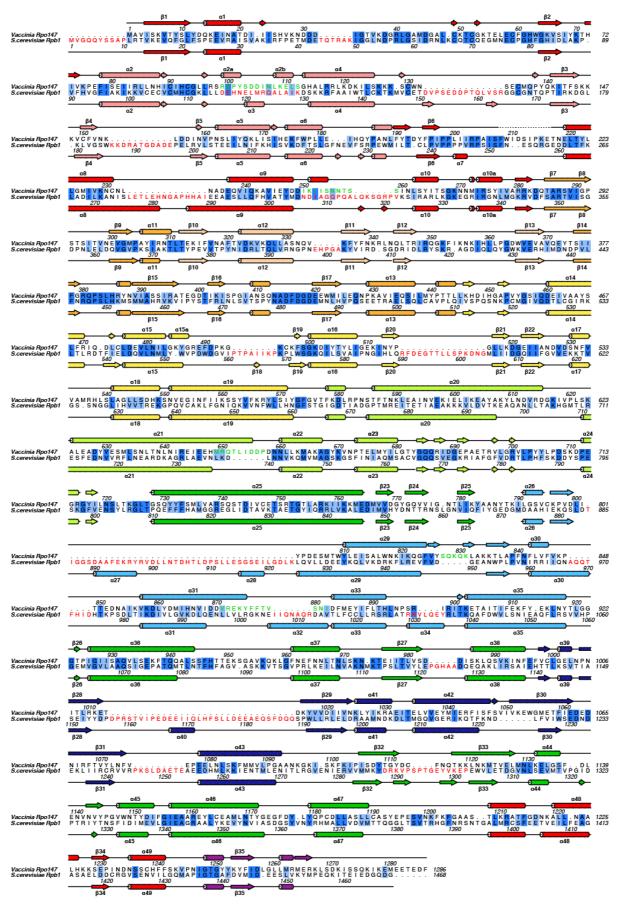**Figure 7: Interactions of NPH-I and VETF in the complete vRNAP complex.**
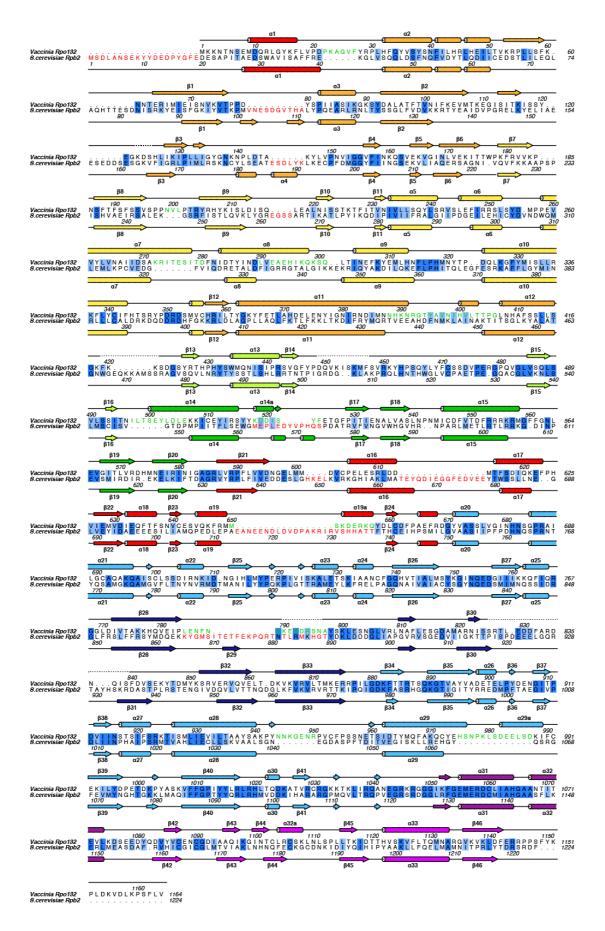
SUPPLEMENTAL FIGURES

Supplemental Figure 1: Purification and activity of vRNAP complexes. Related to Figure 1.
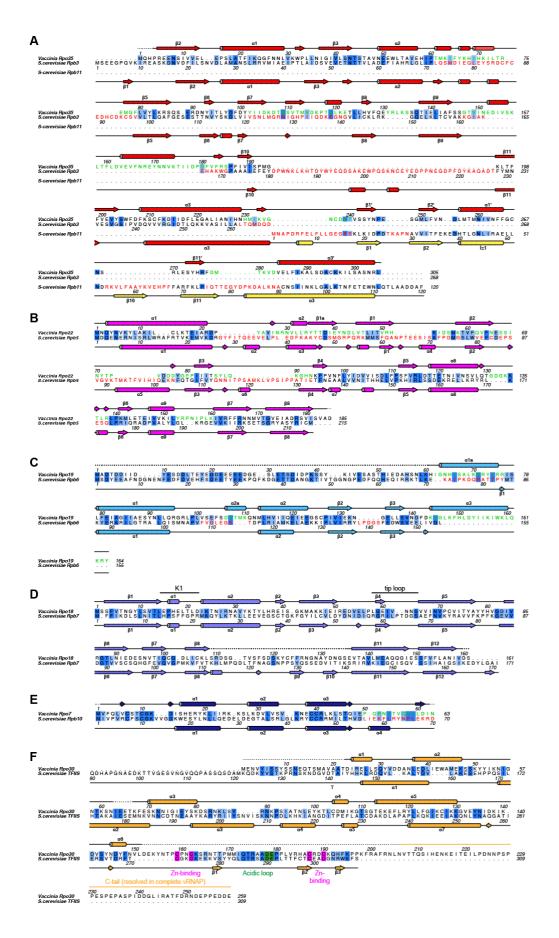
Supplemental Figure 2: Structure determination of core vRNAP. Related to Figure 2.
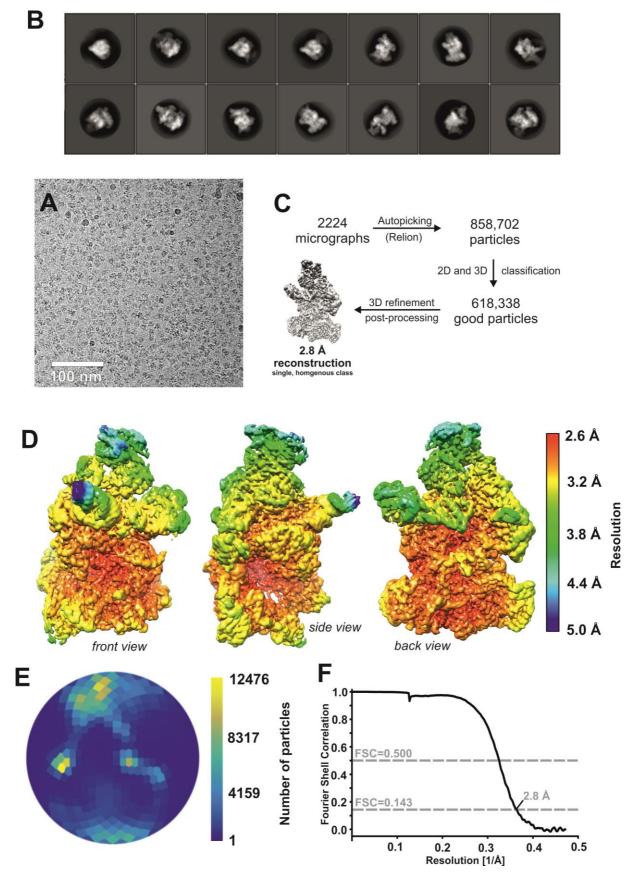
Supplemental Figure 3: Structure-based sequence alignment of Rpo147 and Rpb1. Related to Figure 2B and 3C.
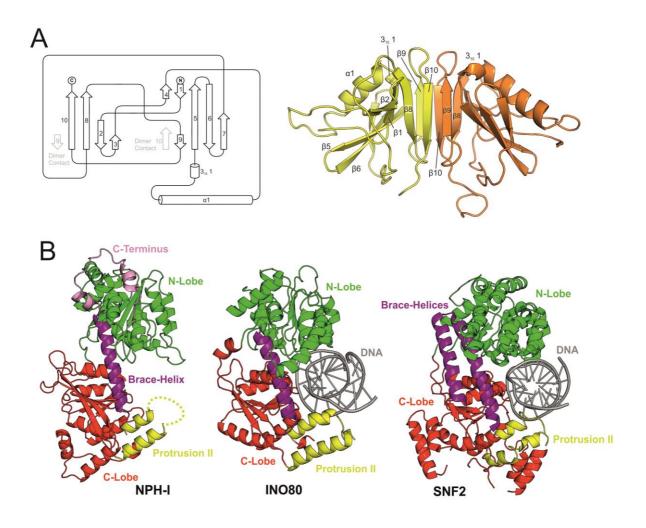
Supplemental Figure 4: Structure-based sequence alignment of Rpo132 and Rpb2. Related to Figure 2B and 3C.

Supplemental Figure 5: Structure-based sequence alignment of Rpo35, Rpo19, Rpo18 and Rpo7. Related to Figure 2B and 3D-F.

Supplemental Figure 6: Structure determination of complete vRNAP. Related to Figure 4.

Supplemental Figure 7: NPH-I and E11. Related to Figure 4 and Figure 5.

SUPPLEMENTARY TABLES

Supplementary Table 1: Subunit composition of vRNAP

vRNAP	RNA Polymerase II homologue	Molecular Weight (kDa, vRNAP)
Rpo147	RPB1	146.9
Rpo132	RPB2	133.4
Rpo35	RPB3 / RPB11	35.4
Rpo30	TFIIS	29.8
Rpo22	RPB5	21.3
Rpo19	RPB6	19.0
Rpo18	RPB7	17.9
Rpo7	RPB10	7.3
Rap94	-	93.6

Supplementary Table 2: Mass Spectrometric identification of vRNAP subunits

Accession number	P-Value	Coverage (%)	#Peptides	Avg. Mass	Description
P20504|RP147_VACCC	796.1	86	467	146831	DNA-directed RNA polymerase 147 kDa polypeptide OX=10249 GN=RPO147 PE=3 SV=1
P20979|MCEL_VACCC	772.6	91	415	96733	mRNA-capping enzyme catalytic subunit OX=10249 GN=D1R PE=3 SV=1
P68694|RP132_VACCC	772.4	94	426	133364	DNA-directed RNA polymerase 133 kDa polypeptide OX=10249 GN=RPO132 PE=3 SV=1
P20637|NTP1_VACCC	617.6	87	241	72366	Nucleoside triphosphatase I OX=10249 GN=NPH1 PE=3 SV=1
P68439|RAP94_VACCC	598.8	81	216	93559	RNA polymerase-associated transcription-specificity factor RAP94 OX=10249 GN=RAP94 PE=2 SV=1
P20635|ETF2_VACCC	585.3	82	200	82329	Early transcription factor 82 kDa subunit OX=10249 GN=VETFL PE=3 SV=1
P21082|RP30_VACCC	560.3	98	146	29796	DNA-directed RNA polymerase 30 kDa polypeptide OX=10249 GN=RPO30 PE=3 SV=1
P20634|ETF1_VACCC	543.8	87	187	73801	Early transcription factor 70 kDa subunit OX=10249 GN=VETFS PE=3 SV=1
P21082-2|RP30_VACCC_Iso	535.9	98	133	27975	Isoform Short of DNA-directed RNA polymerase 30 kDa polypeptide OX=10249 GN=RPO30
P20980|MCES_VACCC	471.7	95	111	33353	mRNA-capping enzyme regulatory subunit OX=10249 GN=D12L PE=3 SV=1
P21087|RP35_VACCC	449.5	79	87	35394	DNA-directed RNA polymerase 35 kDa subunit OX=10249 GN=RPO35 PE=3 SV=1
P68608|RP22_VACCC	446.0	100	105	21342	DNA-directed RNA polymerase 22 kDa subunit OX=10249 GN=RPO22 PE=3 SV=1
P68610|RP19_VACCC	423.8	99	85	18996	DNA-directed RNA polymerase 19 kDa subunit OX=10249 GN=RPO19 PE=2 SV=1
P20988|A11_VACCC	412.8	77	85	36135	Protein A11 OX=10249 GN=A11R PE=2 SV=1
P20642|P4A_VACCC	403.0	44	40	102284	Major core protein 4a precursor OX=10249 GN=A10L PE=3 SV=1
P20643|P4B_VACCC	399.3	56	40	72625	Major core protein 4b OX=10249 GN=A3L PE=3 SV=1
P21010|D5_VACCC	380.0	37	36	90369	Primase D5 OX=10249 GN=D5R PE=3 SV=1
P20498|I1_VACCC	379.2	71	48	35841	Telomere-binding protein I1 OX=10249 GN=I1L PE=2 SV=1
P20998|VITF3_VACCC	356.4	64	41	44638	Intermediate transcription factor 3 large subunit OX=10249 GN=VITF3L PE=2 SV=1
P21034|RP18_VACCC	355.6	99	61	17901	DNA-directed RNA polymerase 18 kDa subunit OX=10249 GN=RPO18 PE=2 SV=1
P68441|D13_VACCC	355.6	43	28	61891	Scaffold protein D13 OX=10249 GN=D13L PE=3 SV=1
P20638|F13_VACCC	342.0	43	26	41824	Envelope protein F13 OX=10249 GN=F13L PE=3 SV=1
P20982|VLTF2_VACCC	340.3	88	41	16950	Viral late gene transcription factor 2 OX=10249 GN=VLTF2 PE=2 SV=1
P21043|C10_VACCC	322.5	40	13	38502	Protein C10 OX=10249 GN=C10L PE=3 SV=1
P20538|H5_VACCC	311.0	63	22	22270	Late transcription elongation factor H5 OX=10249 GN=H5R PE=2 SV=1
P20985|A6_VACCC	310.8	56	32	43128	Protein A6 OX=10249 GN=A6L PE=3 SV=1
P68449|E11_VACCC	304.0	78	32	14899	Protein E11 OX=10249 GN=E11L PE=2 SV=1
P20499|I3_VACCC	299.2	50	22	29981	Protein I3 OX=10249 GN=I3L PE=2 SV=1
P21090|C23L_VACCC	297.3	57	14	26385	Inactive chemokine-binding protein OX=10249 GN=C23L PE=3 SV=1
P20509|DPOL_VACCC	294.0	24	23	116903	DNA polymerase OX=10249 GN=POL PE=1 SV=1
P21047|E6_VACCC	290.1	38	25	66670	Protein E6 OX=10249 GN=E6R PE=2 SV=1
P21080|E2_VACCC	281.2	30	23	85917	Protein E2 OX=10249 GN=E2L PE=3 SV=1
P68318|VLTF3_VACCC	273.6	57	26	26289	Viral late gene transcription factor 3 OX=10249 GN=VLTF3 PE=1 SV=1

P20534|A18_VACCC	272.6	31	24	56737	Transcript termination protein A18 OX=10249 GN=A18R PE=3 SV=1
P21030|G9_VACCC	270.4	45	15	38786	Myristoylated protein G9 OX=10249 GN=G9R PE=1 SV=3
P20640|M1_VACCC	267.6	40	22	54183	Ankyrin repeat protein M1 OX=10249 GN=M1L PE=3 SV=1
P20986|VTF3S_VACCC	266.9	53	24	33576	Intermediate transcription factor 3 small subunit OX=10249 GN=VITF3S PE=2 SV=1
P21081|E3_VACCC	265.8	67	23	21504	Protein E3 OX=10249 GN=E3L PE=2 SV=1
P21022|G1_VACCC	262.6	29	19	67911	Metalloendopeptidase G1 OX=10249 GN=G1L PE=3 SV=1
P20993|A16_VACCC	259.8	32	15	43562	Virion membrane protein A16 OX=10249 GN=A16L PE=1 SV=3
P21079|PAP1_VACCC	259.5	38	24	55580	Poly(A) polymerase catalytic subunit OX=10249 GN=PAPL PE=3 SV=1
P20497|H3_VACCC	257.7	48	20	37458	Envelope protein H3 OX=10249 GN=H3L PE=2 SV=1
P21081-2|E3_VACCC_Iso	254.3	59	18	17515	Isoform Short of Protein E3 OX=10249 GN=E3L
P21008|D2_VACCC	253.9	93	18	16904	Core protein D2 OX=10249 GN=D2L PE=2 SV=1
P20535|A27_VACCC	253.3	65	24	12616	14 kDa fusion protein OX=10249 GN=A27L PE=3 SV=1
P20492|DNLI_VACCC	252.5	27	17	63390	DNA ligase OX=10249 GN=LIG PE=1 SV=1
P20983|A4_VACCC	252.1	40	13	30846	39kDa core protein OX=10249 GN=A4L PE=2 SV=1
P68454|F17_VACCC	251.4	78	16	11336	Core phosphoprotein F17 OX=10249 GN=F17R PE=3 SV=1
P20502|NPH2_VACCC	251.1	30	22	77659	RNA helicase NPH-II OX=10249 GN=NPH2 PE=2 SV=1
P21028|G7_VACCC	249.3	39	20	41921	Assembly protein G7 OX=10249 GN=G7L PE=2 SV=1
P68445|B7_VACCC	241.7	48	15	21312	Protein B7 OX=10249 GN=B7R PE=3 SV=1
P20989|A12_VACCC	235.9	60	16	20519	25 kDa core protein A12L OX=10249 GN=A12L PE=2 SV=1
P21053|F12_VACCC	232.7	17	10	73184	Protein F12 OX=10249 GN=F12L PE=3 SV=1
P20540|L1_VACCC	232.0	37	14	27307	Protein L1 OX=10249 GN=L1R PE=1 SV=3
P20508|CAHH_VACCC	229.7	41	13	35327	Cell surface-binding protein OX=10249 GN=D8L PE=2 SV=1
P20495|DUSP_VACCC	227.4	55	13	19672	Dual specificity protein phosphatase H1 OX=10249 GN=H1L PE=2 SV=1
P68634|DUT_VACCC	225.3	78	11	16264	Deoxyuridine 5'-triphosphate nucleotidohydrolase OX=10249 GN=DUT PE=3 SV=1
P21031|L3_VACCC	220.8	24	11	40646	Protein L3 OX=10249 GN=L3L PE=3 SV=1
P21070|A52_VACCC	220.0	39	14	22710	Protein A52 OX=10249 GN=A52R PE=3 SV=1
P21062|SEMA_VACCC	217.9	22	11	45741	Semaphorin-like protein A39 OX=10249 GN=A39R PE=1 SV=1
P21093|O1_VACCC	216.0	19	16	77578	Protein O1 OX=10249 GN=O1L PE=3 SV=1
P21097|3BHS_VACCC	214.9	36	10	39366	3 beta-hydroxysteroid dehydrogenase/Delta 5-->4-isomerase OX=10249 GN=A44L PE=3 SV=1
P68695|PROF_VACCC	214.5	62	12	15052	Profilin OX=10249 GN=A42R PE=3 SV=1
P20531|SPI1_VACCC	206.9	24	10	40387	Serine proteinase inhibitor 1 OX=10249 GN=SPI-1 PE=3 SV=1
P68315|RP07_VACCC	205.8	76	18	7288	DNA-directed RNA polymerase 7 kDa subunit OX=10249 GN=RPO7 PE=2 SV=1
P21077|SANS_VACCC	205.4	23	11	40952	Surface antigen S OX=10249 GN=B19R PE=3 SV=1
P21052|F11_VACCC	202.6	20	6	39647	Protein F11 OX=10249 GN=F11L PE=3 SV=1
P20981|VP8_VACCC	202.0	38	13	28461	Core protein VP8 OX=10249 GN=L4R PE=2 SV=1
P68463|I6_VACCC	197.1	26	10	43454	Telomere-binding protein I6 OX=10249 GN=I6L PE=3 SV=1
P68461|GLRX2_VACCC	195.0	49	11	13987	Glutaredoxin-2 OX=10249 GN=G4L PE=3 SV=1
P21009|D3_VACCC	193.2	24	12	27974	Core protein D3 OX=10249 GN=D3R PE=2 SV=1
P20496|H2_VACCC	193.2	32	13	21515	Protein H2 OX=10249 GN=H2R PE=3 SV=1
P21049|E8_VACCC	191.9	28	12	31936	Protein E8 OX=10249 GN=E8R PE=2 SV=1
P21066|A46_VACCC	191.6	46	11	24684	Protein A46 OX=10249 GN=A46R PE=3 SV=1
P20537|K4_VACCC	190.2	17	7	48877	Phospholipase-D-like protein K4 OX=10249 GN=K4L PE=3 SV=1

P21033|MCE_VACCC	188.9	28	11	38918	Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase OX=10249 GN=PAPS PE=3 SV=1
P20501|I7_VACCC	188.2	20	11	49040	Core protease I7 OX=10249 GN=I7L PE=2 SV=1
P68612|VLTF1_VACCC	185.4	21	8	29911	Late transcription factor 1 OX=10249 GN=VLTF1 PE=2 SV=1
P21057|A34_VACCC	184.5	42	11	19529	Protein A34 OX=10249 GN=A34R PE=3 SV=1
P21069|A51_VACCC	182.4	21	6	37724	Protein A51 OX=10249 GN=A51R PE=3 SV=1
P21017|F8_VACCC	180.5	85	9	7819	Protein F8 OX=10249 GN=F8L PE=3 SV=1
P21037|VC02_VACCC	179.7	13	6	59246	Kelch repeat protein C2 OX=10249 GN=C2L PE=3 SV=1
P68616|A33_VACCC	178.6	29	5	20506	Protein A33 OX=10249 GN=A33R PE=1 SV=1
P21054|N1_VACCC	175.7	47	7	13962	Protein N1 OX=10249 GN=N1L PE=1 SV=1
P68443|B6_VACCC	175.3	40	10	20142	Protein B6 OX=10249 GN=B6R PE=3 SV=1
P21042|C9L_VACCC	174.8	10	7	74664	Ankyrin repeat protein C9L OX=10249 GN=C9L PE=4 SV=1
P68618|A36_VACCC	173.4	37	8	25133	Protein A36 OX=10249 GN=A36R PE=3 SV=1
P20997|RUVV_VACCC	172.9	47	10	20732	Resolvase A22 OX=10249 GN=A22R PE=3 SV=1
P20995|A20_VACCC	170.6	16	10	49187	DNA polymerase processivity factor component A20 OX=10249 GN=A20R PE=1 SV=1
P21060|A37_VACCC	167.9	19	7	29896	Protein A37 OX=10249 GN=A37R PE=3 SV=1
P21039|C5_VACCC	167.0	25	8	24490	Protein C5 OX=10249 GN=C5L PE=3 SV=1
P68599|C7_VACCC	166.1	39	6	17999	Interferon antagonist C7 OX=10249 GN=C7L PE=3 SV=1
P20505|B1_VACCC	164.3	33	9	34302	B1 kinase OX=10249 GN=VPK1 PE=2 SV=1
P21040|C6_VACCC	161.9	38	8	17367	Protein C6 OX=10249 GN=C6L PE=3 SV=1
P21115|B5_VACCC	160.0	9	4	35109	Plaque-size/host range protein OX=10249 GN=PS/HR PE=3 SV=1
P21114|A26_VACCC	160.0	11	5	37320	Putative A-type inclusion protein OX=10249 GN=A26L PE=3 SV=1
P21073|KBTB1_VACCC	158.6	13	8	64719	Kelch repeat and BTB domain-containing protein 1 OX=10249 GN=KBTB1 PE=3 SV=1
P21074|A57_VACCC	157.8	44	7	17448	Guanylate kinase homolog OX=10249 GN=A57R PE=3 SV=3
P68693|KTHY_VACCC	157.7	20	5	23219	Thymidylate kinase OX=10249 GN=TMK PE=1 SV=1
P68697|TOP1_VACCC	157.5	21	6	36666	DNA topoisomerase 1B OX=10249 GN=TOP1 PE=2 SV=1
P21026|G5_VACCC	157.4	15	8	49873	Putative nuclease G5 OX=10249 GN=G5R PE=2 SV=1
P68592|A17_VACCC	153.9	35	6	22999	Virion membrane protein A17 precursor OX=10249 GN=A17L PE=3 SV=1
P21027|G6_VACCC	152.8	33	6	18949	Uncharacterized protein G6 OX=10249 GN=G6R PE=3 SV=1
P20493|RIR2_VACCC	152.1	10	5	36999	Ribonucleoside-diphosphate reductase small chain OX=10249 GN=F4L PE=2 SV=1
P21067|A47_VACCC	151.8	20	6	28335	Protein A47 OX=10249 GN=A47L PE=3 SV=1
P20999|B2_VACCC	150.3	23	6	24628	Protein B2 OX=10249 GN=B2R PE=3 SV=1
P21095|VPK2_VACCC	150.2	20	11	52159	Serine/threonine-protein kinase 2 OX=10249 GN=VPK2 PE=3 SV=1
P21083|J5_VACCC	148.1	31	7	15158	Protein J5 OX=10249 GN=J5L PE=3 SV=1
P20503|RIR1_VACCC	146.6	7	5	87754	Ribonucleoside-diphosphate reductase large subunit OX=10249 GN=I4L PE=2 SV=1
P68690|GLRX1_VACCC	144.7	47	7	12355	Glutaredoxin-1 OX=10249 GN=O2L PE=3 SV=1
P68450|F1_VACCC	144.4	23	7	26367	Protein F1 OX=10249 GN=F1L PE=3 SV=1
P21013|F3_VACCC	143.9	10	5	55717	Kelch repeat protein F3 OX=10249 GN=F3L PE=3 SV=1
P20536|UNG_VACCC	137.7	31	7	25080	Uracil-DNA glycosylase OX=10249 GN=UNG PE=1 SV=1
P21001|B4_VACCC	133.3	17	11	65272	Ankyrin repeat protein B4 OX=10249 GN=B4R PE=3 SV=1
P21014|F5_VACCC	133.2	8	3	36479	36 kDa major membrane protein F5 OX=10249 GN=F5L PE=3 SV=1
P68459|G3_VACCC	132.0	50	8	12803	Protein G3 OX=10249 GN=G3L PE=3 SV=1
P21116|IL1BP_VACCC	128.9	8	4	36670	Interleukin-1-binding protein OX=10249 GN=B16R PE=3 SV=2

P21032|J1_VACCC	123.8	11	3	17876	Protein J1 OX=10249 GN=J1R PE=2 SV=1
P68632|A28_VACCC	123.5	23	7	16329	Envelope protein A28 homolog OX=10249 GN=A28L PE=3 SV=1
P0CK20|A25_VACCC	121.4	36	5	8890	Protein A2.5 OX=10249 GN=A2.5L PE=2 SV=1
P21021|F16_VACCC	121.3	11	3	26605	Protein F16 OX=10249 GN=F16L PE=3 SV=1
P68639|VCP_VACCC	118.1	7	2	28629	Complement control protein C3 OX=10249 GN=C3L PE=1 SV=1
P21132|SODL_VACCC	115.3	26	5	13795	Cu-Zn superoxide dismutase-like protein OX=10249 GN=A45R PE=3 SV=1
P21065|A43_VACCC	114.0	15	3	22635	Protein A43 OX=10249 GN=A43R PE=3 SV=1
P68467|K7_VACCC	113.6	23	4	17468	Protein K7 OX=10249 GN=K7R PE=1 SV=1
P68614|A32_VACCC	111.6	15	5	30904	Protein A32 OX=10249 GN=A32L PE=3 SV=2
P21098|KRB2_VACCC	109.4	13	5	33366	Probable serine/threonine-protein kinase B12 OX=10249 GN=B12R PE=3 SV=1
P21011|D9_VACCC	108.7	24	7	25024	mRNA-decapping protein D9 OX=10249 GN=D9R PE=2 SV=1
P68457|G2_VACCC	107.5	22	6	25744	Late transcription elongation factor G2 OX=10249 GN=G2R PE=3 SV=1
P21016|F7_VACCC	107.4	41	5	10938	Protein F7 OX=10249 GN=F7L PE=3 SV=1
P20539|H7_VACCC	105.4	18	4	16912	Late protein H7 OX=10249 GN=H7R PE=2 SV=1
P21058|A35_VACCC	105.1	11	3	20005	Protein A35 OX=10249 GN=A35R PE=2 SV=1
P21012|D10_VACCC	101.8	8	3	28934	mRNA-decapping protein D10 OX=10249 GN=D10R PE=2 SV=1

Supplemental Table 3: Identification of the bound tRNA species bound to complete vRNAP by RNAseq analysis.

	number of reads	**%**
$tRNA^{Gln}(TTG)$	633503	98.2
$tRNA^{Gln}(CTG)$	7057	1.1
$tRNA^{Arg}(ACG)$	3337	0.5
$tRNA^{Arg}(TCG)$	1745	0.2

Supplementary Table 4: Data collection and model refinement statistics.

Cryo-EM data collection, refinement and validation statistics

	Core vRNAP (EMD-xxxx) (PDB xxxx)	Complete vRNAP (EMD-4868) (PDB 6RFL)
Data collection and processing		
Magnification	165,000	75,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	55	50
Defocus range (µm)	-0.35 to -3.25	-1.4 to -2.7
Pixel size (Å)	0.81	1.06
Symmetry imposed	C1	C1
Initial particle images (no.)	479,618	858,702
Final particle images (no.)	152,591	618,338
Map resolution (Å)	2.8	2.75
FSC threshold	0.143	0.143
Map resolution range (Å)	2.8 – 4.8	2.6 – 5.0
Refinement		
Model resolution (Å)	2.9	2.9
FSC threshold	0.5	0.5
Map sharpening B factor (Å^2)	-89	-39
Model composition		
Non-hydrogen atoms	29671	51788
Protein/RNA residues	3663 / 0	6194 / 63
Ligands / Water	4 x Zn, 1 x Mg / -	4 x Zn, 1 x Mg / 61
B factors (Å^2)		
Protein/RNA	53.5 / -	55.3 / 75.5
Ligand / Water	89.9 / -	49.0 / 30.2
R.m.s. deviations		
Bond lengths (Å)	0.004	0.005
Bond angles (°)	0.743	0.756
Validation		
MolProbity score	1.50	2.15
Clashscore	5.48	4.30
Poor rotamers (%)	0.12	4.82
Ramachandran plot		
Favored (%)	96.71	93.7
Allowed (%)	3.29	6.2
Disallowed (%)	0.00	0.05

X-Ray data collection, refinement and validation statistics

	E11-native (PDB 6RFG)	E11-Hg
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions a,b,c (Å)		59.8, 75.5, 138.8
Resolution (Å)	40.0 - 1.90 (2.00 - 1.90)*	40.0 – 2.29 (2.43 – 2.29)* [40.0-6.86]**
Rsym (%)	16.4 (163.5)	6.4 (24.7)
$I/\sigma(I)$	11.1 (1.0)	17.6 (5.2)
Completeness (%)	96.6 (87.4)	94.4 (69.3)
Redundancy	6.4 (6.1)	4.8 (3.9)
CC1/2 (%)	99.7 (34.9)	99.8 (92.6)
SigAno		0.99 (0.80) [2.24]
Anomalous correlation (%)		19.0 (1.0) [75.0]
Refinement (Phenix.refine)		
Resolution (Å)	40.0 - 1.90 (2.00 - 1.90)	
No. of reflections	48722 (4328)	
Rwork / Rfree (%)	20.7 / 22.9 (33.6 / 38.1)	
Model composition		
Non-hydrogen atoms	4584	
Protein residues	520	
Water molecules	333	
RMS deviations		
Bond lengths (Å)	0.003	
Bond angles (°)	0.78	
Validation		
All-atom clashscore	2.47	
Rotamer outliers	1.64 %	
C-beta deviations	0	
Ramachandran plot		
Outliers	0.2	
Allowed	2.9 %	
Favored	96.9 %	

* Number in () parentheses for highest resolution shell
** Number in [] parentheses for lowest resolution shell

Supplementary Table 5: Primers used in construction of the A24R transfer vector

1st PCR	Template: GLV-1h68 viral DNA	
Product A	A24R-5	5'-CCTAGTAGTAACGAAACATCCATCGATAC-3'
	A24Rtag-3	5'-CCTTGTAATCAGCGTAATCTGGAACATCGTATGGGTACA TGGTCACCAGAAAAGACGGCTTGAG-3'
Product B	A25Ltag-5	5'-GTTCCAGATTACGCTGATTACAAGGATGACGACGATAAG GCAGCATAATATTCTAGTTTGGTAGTAGATACATATCAATATCATCA-3'
	A25L-3	5'-TCGATCGCCTTCAGAAAGACTG-3'
2nd PCR	Template: PCR products A and B, gel purified	
Product C	A24R-5	5'-CCTAGTAGTAACGAAACATCCATCGATAC-3'
	A25L-3	5'-TCGATCGCCTTCAGAAAGACTG-3'