nature portfolio

Peer Review File

Transcription regulation of African swine fever virus: dual role of M1249L

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The manuscript describes a very extensive work on ASFV vRNAP and its interaction with transcription factors in particular the late protein M1249L. It comprises an analysis of the ASFV virus transcriptome, and the time course of the M1249L interaction with vRNAP during different stages of infection and its role as transcriptional activator. Cryo-EM structures show the more or less tight association of M1249L with vRNAP

A serendipitous co-purification of endogenous M1249L with ASFV RNAP from infected cells showed its role of a transcriptional activator/regulator and led to an important and exciting finding.
This double role as a ruler protein and a role as transcription factor or transcriptional regulator should be highlighted in the discussion. M1249L is apparently a rare example of a protein with two states, an extended alpha-helical conformation as observed in capsid assembly where it serves as a staple or ruler protein and where only domain 1 and 6 are folded. In contrast up to 6 folded domains can be visible when it is associated with vRNAP. Due to this two states, structure prediction using Alphafold is likely to be unreliable and clean and well-documented experimental structures are more than required.

In the amount of experimental materiel and the large number of models presented in the manuscript, it is very hard to find the relevant information to judge the quality of the model of M1249L. The relevant information is scattered over different validation reports:
For example, when I use the straight-forward approach to look at the “Extended Data Table 3:” ASFV RNA polymerase-M1249L complex complete (8YQY), I get to a complex of a moderate resolution of 3.68 A, which shows a poor fit of the M1249L model to density for almost all of the chain. Probably more precise structural information is hidden in other maps, but in which ones ?
If I take entry 8YQT with a nominal resolution of 2.56 A, M1249L is only well defined from res. 672 to res. 866. and another subunit, D339L, is also very poorly defined in this structure.

The structure of M1249L should be presented domain by domain together with the local resolution of the best defined model and the fit to the density in order to enable the reader to judge the reliability of the model. The information of the Extended data figure 6 needs to be completed. The pdb code of the entry, which defines this domain structure of each of the 6 domains, should be given. In the Supplementary material, a file containing the relevant parts of the validation report should be compiled, in order to be able to judge the quality of the M1249L model. Currently, this information is diluted in 29 structures and their validation reports.

Also the question arises, whether the structure of domain 1 and 6 is compatible with the virus core structure by Wang et al., 2019.
The structure of the other subunits probably confirms the previously published structure of recombinantly expressed core of the vRNAP without M1249L (Pilotto, 2024). This agreement should be discussed.

Further comments are made referring to specific lines of the manuscript.

Line 92. Here 3 structures are introduced: core, core complexed with endogenous DNA and core complexed with M1249L. How do they related to the 27 states and the 22 complexes presented elsewhere in article and Supporting Material?

Line 128-132: The identification of the mass increment as M1249L should be described. Complementary gel electrophoresis and MS or ES-MS analysis is necessary is necessary for the identification. VRPB2 and M1249L seem to co-migrate on PAGE, how has the latter been identified?

line 150-154: References to supporting material and pdb entries should be clearer.

Line 157-159. The methodology is not explained.

Line 187: The structure seems to be unique in the terms of the double functions, but apparently a number of the domains resemble different transcription factors. So the structure is not really “unique” in these terms.

Line 239: High similarity with TFII2 from … Part of the sentence missing?

Line 242: Domain II and domain III of what ?

Line 243: Looking at the figure 3l, it is difficult to find any similarities. In the discussion of the acidic hairpin, more evidence for a role should be given as two subsequent acidic residues in a protein are not rare and there could be a pure coincidence!

Line 253: The first ten lines of the discussion are a repetition of the introduction section. They can be omitted in order to gain some space, for example to discuss more the relation of M1249L role as ruler protein and a transcription factor and to discuss the quality of the structures and their dynamics more in detail.

Line 266: The relation with swine RNAP has to be supported further. It is more likely that ASFV RNAP is just similar to eukaryotic RNAP, with VACV being evolutionarily further distant. To prove an relationship with the host RNAP, more phylogenetic analysis is required to show that the relationship is indeed closer than for other eukaryotic RNAPs.

Line 269: The double role of M1249L is the central result of the work and should be discussed more in detail.

line 497: The methods does not agree with Extended Data Figure 2, for example “Following homogeneous and local refinement in cryoSPARC, two maps at resolutions of 2.64 Å (state 3) and 2.97 Å (state 4) were obtained. “ Subsequently, state 6 is not present in Extended data figure 2 and it is not indicated how the map named “local” has been obtained and what makes it different from state 3.

lines 516: A huge amount of data is mentioned here but where are the links to The Supporting Material

section explaining the processing steps used and which are exactly the results obtained from these data?

Line 526: “classification generated a total of 1,000 classes from datasets of various periods “ The number of classes seems to be excessive.

Line 531 Table 3 seems to be missing, or is it Extended Data table 3 ?

Line 574: “HEK293T cells were transfected with plasmids of p72-Fluc RNAP reporter system in…” This should be more precise. I suppose the 8 vRNAP are encoded by individual transfected plasmids together with the p72-Fluc plasmid.

Figure 1b: Please indicate how M1249L has been identified.

Figure 2: pdb identifiers should be included

Figure 4i: The abbreviations should be explained in the legend

Globally, I am concerned about the size of a lot of the panels, which are too small to be well readable.

I have also some doubts how far claimed resolutions are always real, but to judge this, I would have to sift through 29 validation reports.

The body of the article should refer to the pdb entries. Apparently, 29 pdb structures and the corresponding maps are associated with the article. For all figures and supplementary figures, the corresponding pdb entries should be indicated.

The structure of the core RNAP should be compared with the structure of Pilotto et al. and the agreement should be discussed.

Extended Data Figure 6: panel C:
The secondary structure elements of each domain should be boxed to improve readability. Linkers should also be shown together with their secondary structure elements.

Extended Data Figure 7: The right panel is unintelligible.

Extended Data Figure 8: The M1249L C-tail is difficult to follow. The yellow hydrophobic region add more to confusion than they are clarifying.

Extended Data Table 1: I do not understand the relation between data sets on core and complex 1-21 with the different states presented in the text. pdb codes should be used as unique identifiers within the article.

The presence or absence of domains and complex partners is a relative term for cryo-EM structure as they are accompanied by a progressive loss of resolution, a reduced level of the density and finally a

clear absence. Distinguishing between an absence and an disorder of a domain is not an easy task. The amount of material presented should be revisited for its relevance to the message of the article and should be reorganized to improve the ease of reading. With the provided material, it is impossible to judge the quality of the subunit and domain structures and the main messages are getting swamped in a multitude of presented conformations blurring considerably the image.

Reviewer #2 (Remarks to the Author):

This manuscript describes the structure of the African swine fever virus (ASFV) polymerase complexed with the viral protein M1249L, which is also a capsid component. The authors demonstrate co-purification of M1249L with the polymerase and use cryo-EM analysis to reveal its interaction with the enzyme. Luciferase assays suggest that M1249L enhances polymerase activity.

Several key concerns warrant further investigation. The authors describe multiple structures and propose two (maybe more) functional states for M1249L. However, interpreting five distinct structures without clear functional evidence requires additional support.

The manuscript readability and figure clarity has to be be improved. The figures are too complex and difficult to follow and will benefit from simplification for better interpretation.

For M1249L, the manuscript lacks details on the D1-6 domain structures. Can AlphaFold predict these structures considering M1249L's role as a capsid skeleton protein with a fiber-like, 100 nm long configuration? How do these domains compare to their counterparts in human, yeast, and vaccinia viruses?

The authors mention solving the ASFV-vRNAP complex structure within 48 hours, followed by structures from material purified at different time points - and then try to interpret function from these. Also it's unclear which PDB code corresponds to which state. While 23 distinct reconstructions suggest various M1249L-vRNAP binding states, this might simply reflect flexible, heterogeneous binding.

Analysis of one PDB structure (PDBID:8YQZ, DNA complex) reveals potential issues with map interpretation. While the overall model-to-map fit appears reasonable, specific regions lack clear density support and require remodeling. Potential oversharpening of the map is a concern (I havent looked at the other structures but imagine would have similar concerns).

The modeling of specific chain regions (Chain F:82-339, Chain A:439-508, Chain A:63-96, 131-155) lacks sufficient density support. Repositioning residues in Chain A (439-508) might allow chain break completion. The weak RNA density suggests oversharpening or limitations in current modeling. If these regions were modeled based on modified maps (e.g., deepEMhancer), presenting those maps would strengthen the interpretation.

The title of Figure 3 ("Functional dissection of M1249L") could be misleading without stronger functional

evidence.

In summary, the manuscript raises intriguing questions about M1249L's interaction with the ASFV polymerase. However, interpretations of the EM data and proposed functional states require further clarification and supporting evidence.

Point-to-point response to reviewers' comments

Reviewer #1 (Remarks to the Author):

The manuscript describes a very extensive work on ASFV vRNAP and its interaction with transcription factors in particular the late protein M1249L. It comprises an analysis of the ASFV virus transcriptome, and the time course of the M1249L interaction with vRNAP during different stages of infection and its role as transcriptional activator. Cryo-EM structures show the more or less tight association of M1249L with vRNAP

Our response: This is a conclusion of our manuscript. No response is needed.

A serendipitous co-purification of endogenous M1249L with ASFV RNAP from infected cells showed its role of a transcriptional activator/regulator and led to an important and exciting finding. This double role as a ruler protein and a role as transcription factor or transcriptional regulator should be highlighted in the discussion. M1249L is apparently a rare example of a protein with two states, an extended alpha-helical conformation as observed in capsid assembly where it serves as a staple or ruler protein and where only domain 1 and 6 are folded. In contrast up to 6 folded domains can be visible when it is associated with vRNAP. Due to this two states, structure prediction using Alphafold is likely to be unreliable and clean and well-documented experimental structures are more than required.

Our response: We thank the reviewer's suggestion as to highlight the double role of M1249L both in capsid assembly and RNA transcription regulation. We have made revision to the discussion section, adding more in-depth analysis and discussion about the double role of M1249L during ASFV infection.

As to the reviewer's concern about the reliability of M1249L structure, the reconstructed structure of the corresponding domains (D1-D6) are modeled to fit the final refined maps of high resolution (within 3.0 Angstrom) instead of just as a rigid body fitting of AlphaFold prediction model. Not only do the main chains tracing conforms to the map, but also the side chains fit well with the local maps as well. And it is noteworthy that the D1 domain has no resemblance with the AlphaFold model and is *ab initio* modeled manually. In conclusion, we do not soly rely on AlphaFold prediction and the structure of M1249L is reliably built, refined and validated based on the actual density maps. In order to show the structural reliability, we have added local high-resolution map demonstrations and model-map fitness to the Extended Data Figure 3.

In the amount of experimental materiel and the large number of models presented in the manuscript, it is very hard to find the relevant information to judge the quality of the model of M1249L. The relevant information is scattered over different validation reports:

For example, when I use the straight-forward approach to look at the "Extended Data Table 3:" ASFV RNA polymerase-M1249L complex complete (8YQY), I get to a complex of a moderate resolution of 3.68 A, which shows a poor fit of the M1249L model to density for almost all of the chain. Probably more precise structural information is hidden in other maps, but in which ones?

If I take entry 8YQT with a nominal resolution of 2.56 A, M1249L is only well defined from res. 672 to res. 866. and another subunit, D339L, is also very poorly defined in this structure.

Our response: We agree that due to our extensive classification of various conformations, the

number of final reported M1249L structures complexed with ASFV RNA polymerase, as well as the maps, is relatively large. However, these structures/maps reflect the intrinsic flexibility of M1249L and should play important roles during ASFV RNA transcription, during which several domains needs to move away for DNA entry and RNA exit, as described in our manuscript.
Additionally, our reported structures involving M1249L is reliable. The structures of the individual domains of M1249L were modeled based on high-quality domain maps. To support this, we have included the domain maps in Extended Data Figure 3. Furthermore, we have provided the structure validation data for each domain in Extended Data Table 1, which includes information on chemical bond/length/chimral information and model-map correlation values.

The structure of M1249L should be presented domain by domain together with the local resolution of the best defined model and the fit to the density in order to enable the reader to judge the reliability of the model. The information of the Extended data figure 6 needs to be completed. The pdb code of the entry, which defines this domain structure of each of the 6 domains, should be given. In the Supplementary material, a file containing the relevant parts of the validation report should be compiled, in order to be able to judge the quality of the M1249L model. Currently, this information is diluted in 29 structures and their validation reports.
Our response: We thank the reviewer for the suggestion. Each domain of M1249L has been presented together with its respective local resolution map and fitness to the density, as included in Extended Data Figure 3. As for Extended Data figure 7 (former Extended Data figure 6), pdb entry codes as well as the EMD numbers have been included for clarification. As mentioned above, we have provided the structure validation data for each domain in Extended Data Table 1, which includes information on chemical bond/length/chiral information and model-map correlation values.

Also the question arises, whether the structure of domain 1 and 6 is compatible with the virus core structure by Wang et al., 2019.
Our response: The structure of domain 1 and 6 is bears no similarity with the virus core structure (Wang et al., 2019). Due to the low resolution of the ASFV core structure (7-8 angstrom), it is insufficient for model building of M1249L in the core by then. Therefore we have included the maps of M1249L in ASFV core (Wang et al., 2019) to compare with the M1249L in vRNAP in this manuscript (see Extended Data Figure 10). Also, the lobe 1 of M1249L in ASFV core is likely to correspond to the NTD in vRNAP complex, which is missing in map due to its structural flexibility.

The structure of the other subunits probably confirms the previously published structure of recombinantly expressed core of the vRNAP without M1249L (Pilotto, 2024). This agreement should be discussed.
Our response: We thank the reviewer for the suggestion. The other subunits of vRNAP has been compared between our modeled structure and those reported previously (Pilotto, 2024), with a highly similar overall tructure (r.m.s.d. of Cα atoms ~0.85 angstrom). The structural comparison has been added to the Extended data Figure 5 and the manuscript.

Further comments are made referring to specific lines of the manuscript.

Line 92. Here 3 structures are introduced: core, core complexed with endogenous DNA and core complexed with M1249L. How do they related to the 27 states and the 22 complexes presented elsewhere in article and Supporting Material?

Our response: We thank the reviewer for pointing out the expression that might cause misunderstanding. The vRNAP core corresponds to PDB-8YQV and EMD-39507; the vRNAP-DNA complex corresponds to PDB-8YQZ and EMD-39511; the rest 5 structures and 26 maps correspond to vRNAP-M1249L complex at different conformations. We have rephrased this sentence in the modified manuscript.

Line 128-132: The identification of the mass increment as M1249L should be described. Complementary gel electrophoresis and MS or ES-MS analysis is necessary is necessary for the identification. VRPB2 and M1249L seem to co-migrate on PAGE, how has the latter been identified?

Our response: We thank the reviewer for the suggestion. The mass increment as M1249L in the Mass photometry measurement result has been added to the manuscript. And the manuscript already includes gel electrophoresis and MS Fingerprinting results for the identification of components in purified ASFV vRNAP. Due to the close similarity of MW of VRPB2 and M1249L, they do appear as one band on SDS-PAGE (Figure 1b), and are co-identified by MS Fingerprinting, the result of which could be retrieved on line (http://www.matrixscience.com/cgi/master_results.pl?file=..%2Fdata%2F20231109%2FFTocfnueh.dat).

line 150-154: References to supporting material and pdb entries should be clearer.

Our response: We have added the pdb entry numbers for clarity.

Line 157-159. The methodology is not explained.

Our response: Due to the space limit in Results section, we did not explain in detail for the methodology. Please see the method section (under the title of “Image processing”) for the methodology details.

Line 187: The structure seems to be unique in the terms of the double functions, but apparently a number of the domains resemble different transcription factors. So the structure is not really “unique” in these terms.

Our response: The reason we initially described the structures of the M1249L domains as “unique” is that no homologous structures could be identified through structural similarity searches with Dali or VAST. The M1249L domains exhibit generally overall similarities to other transcription factor structures (in ways such as similar binding positions, functional motifs, and Zinc finger compositions). Additionally, there are some similarities between ASFV’s M1249L and the RAP94 protein from vaccinia virus. Therefore, we have revised the description in the manuscript, replacing “unique” with “special.”

Line 239: High similarity with TFII2 from … Part of the sentence missing?

Our response: We have rephrased the sentence to “By structure comparison, we found that D5

exhibits high degree of similarity with TFIIS of RNAPs from human (PDB: 8A40) and yeast (PDB: 1Y1V).” for clarity.

Line 242: Domain II and domain III of what?
Our response: Here the Domain II and domain III refer to the second and third domain regions of TFIIS. We have rephrased this sentence in the manuscript to avoid misunderstanding.

Line 243: Looking at the figure 3l, it is difficult to find any similarities. In the discussion of the acidic hairpin, more evidence for a role should be given as two subsequent acidic residues in a protein are not rare and there could be a pure coincidence!
Our response: Regarding the structural comparison between M1249L-D5 and TFIIS, while their structures appear different at the first sight, there are three evident similarities: (1) both bind at the NTP channel position of RNA polymerase, (2) both contain similar functional motifs of Zinc finger structures, and (3) both include two consecutive D-E acidic amino acid motif. Explaining these similarities as pure coincidences would imply a very low probability of all these conditions occurring simultaneously. To avoid over-interpretation, we have tone down our phrases in the manuscript to present this more delicately.

Line 253: The first ten lines of the discussion are a repetition of the introduction section. They can be omitted in order to gain some space, for example to discuss more the relation of M1249L role as ruler protein and a transcription factor and to discuss the quality of the structures and their dynamics more in detail.
Our response: We thank the reviewer for the constructive suggestion. The first part of discussion has been reduced, with the repetition of the introduction removed. Further discussion about the multiple roles of M1249L during viral infection has been added to the discussion section.

Line 266: The relation with swine RNAP has to be supported further. It is more likely that ASFV RNAP is just similar to eukaryotic RNAP, with VACV being evolutionarily further distant. To prove an relationship with the host RNAP, more phylogenetic analysis is required to show that the relationship is indeed closer than for other eukaryotic RNAPs.
Our response: We thank the reviewer for the suggestion. The ASFV vRNAP is indeed more closely related to host RNAP than its viral homolog of VACV. As suggested by the reviewer, we have done phylogenetic analysis (Extended Figure 6.b) based on RPB1&RPB2 subunits sequences, which proved our proposed hypothesis.

Line 269: The double role of M1249L is the central result of the work and should be discussed more in detail.
Our response: We thank the reviewer for the suggestion. As mentioned above, we have added discussion of the multiple roles of M1249L in discussion.

line 497: The methods does not agree with Extended Data Figure 2, for example “Following homogeneous and local refinement in cryoSPARC, two maps at resolutions of 2.64 Å (state 3) and 2.97 Å (state 4) were obtained. “ Subsequently, state 6 is not present in Extended data figure 2 and it is not indicated how the map named “local” has been obtained and what makes it different from

state 3.
Our response: We have corrected and rephrased the sentences here. The “State 6” here should be “State 5” and has been corrected. “Local” refers to the local refinement map of “State 3”

lines 516: A huge amount of data is mentioned here but where are the links to The Supporting Material section explaining the processing steps used and which are exactly the results obtained from these data?
Our response: link to Extended data figure 4 and Extended data table 1have been added. The exact results obtained from these data have been described in Result section, under the title of “Structure determination of ASFV core vRNAP and vRNAP-M1249L complex”.

Line 526: “classification generated a total of 1,000 classes from datasets of various periods “ The number of classes seems to be excessive.
Our response: We agree that the phrasing of “1000 classes” is excessive, and was simply calculated mathematically by 250 classes times 4 time periods (12/16/24/48h protein samples). The actual final classification is manually checked, and reduced to 25. We have made corrections to the sentences.

Line 531 Table 3 seems to be missing, or is it Extended Data table 3 ?
Our response: The original “Table 3” is actually “Extended Data table 1”. We have corrected it.

Line 574: “HEK293T cells were transfected with plasmids of p72-Fluc RNAP reporter system in…” This should be more precise. I suppose the 8 vRNAP are encoded by individual transfected plasmids together with the p72-Fluc plasmid.
Our response: We thank the reviewer for this suggestion. Yes, the eight vRNAP subunits and M1249L are each encoded by an individual plasmid, together with the p72-Fluc plasmid. We have added the information to this Method section.

Figure 1b: Please indicate how M1249L has been identified.
Our response: M1249L was identified by MS Fingerprinting analysis with the SDS-PAGE stained band. We have added this information in the figure legend.

Figure 2: pdb identifiers should be included
Our response: The pdb and EMD identifiers have been included in the updated Figure 2.

Figure 4i: The abbreviations should be explained in the legend
Our response: We thank the review for this suggestion. Explanations for the abbreviations have been included in the legend now.

Globally, I am concerned about the size of a lot of the panels, which are too small to be well readable.
Our response: Due to the complex structure of the ASFV vRNAP-M1249L, we included a large number of panels to better present the various domains of M1249L and their structural comparisons and interaction analysis, elucidating its potential role in viral RNA transcription. Our

original images were sufficiently large and of high resolution; however, converting them into a PDF format during initial submission caused a decrease in resolution. Following the reviewer's suggestion, we have enlarged the panels in the new figures.

I have also some doubts how far claimed resolutions are always real, but to judge this, I would have to sift through 29 validation reports.
Our response: For the reviewer's reference, we have uploaded the validation report，as well as the maps and pdb files online for downloading if needed.（https://www.dropbox.com/scl/fo/h713bg9l8p82euvz2jxgo/AAeXUj-2TB8yLXcn_Ltei-o?rlkey=x3m7vk6qwbz3ej66b6z4pvojw&st=u68b7vhi&dl=0）.

The body of the article should refer to the pdb entries. Apparently, 29 pdb structures and the corresponding maps are associated with the article. For all figures and supplementary figures, the corresponding pdb entries should be indicated.
Our response: There might be some misunderstanding from the reviewer. Most of the maps have low resolution and are insufficient for building structures, but this does not affect our main discussion and analysis. In total, we constructed 7 PDBs and 28 maps. We have included their enty numbers in the main text, figures, and legends of the manuscript for clarity.

The structure of the core RNAP should be compared with the structure of Pilotto et al. and the agreement should be discussed.
Our response: We thank the reviewer for this suggestion. Comparison with the structure from Pilotto et al has been included (see Extended Data Figure 5a) and discussed in our revised manuscript.

Extended Data Figure 6: panel C:
The secondary structure elements of each domain should be boxed to improve readability. Linkers should also be shown together with their secondary structure elements.
Our response: We have redrawn the secondary structure elements of M1249L, with domains and linkers shown more clearly in the revised Extended Data Figure 7 (former Extended Data Figure 6).

Extended Data Figure 7: The right panel is unintelligible.
Our response: We have adjusted the coloring and presentation of the right panel of Extended Data Figure 8 (former Extended Data Figure 7) to make it more intelligible.

Extended Data Figure 8: The M1249L C-tail is difficult to follow. The yellow hydrophobic region add more to confusion than they are clarifying.
Our response: We have revised the Extended Data Figure 9(former Extended Data Figure 8), with the M1249L C-tail structure shown in bold and yellow hydrophobic region more transparent, so as to reduce the vagueness.

Extended Data Table 1: I do not understand the relation between data sets on core and complex 1-21 with the different states presented in the text. pdb codes should be used as unique identifiers

within the article.
Our response: As suggested by the reviewer, we have included the pdb codes in the revised Extended Data Table 1.

The presence or absence of domains and complex partners is a relative term for cryo-EM structure as they are accompanied by a progressive loss of resolution, a reduced level of the density and finally a clear absence. Distinguishing between an absence and an disorder of a domain is not an easy task.
Our response: Yes, we agree with the reviewer that the presence/absence of each domain of M1249L could be due to those mentioned reasons and is hard to distinguish which reason is dominant. That's why we have classified the particles into 250 classes initially, so as to better characterize the various conformations of M1249L.

The amount of material presented should be revisited for its relevance to the message of the article and should be reorganized to improve the ease of reading. With the provided material, it is impossible to judge the quality of the subunit and domain structures and the main messages are getting swamped in a multitude of presented conformations blurring considerably the image.
Our response: We have taken the suggestions from the reviewer and made revision to our manuscript including: modification to the wording, adding Extended data figures 3 for further explanation with clear details, more specific references to PDB/EMD codes in main text and figures.

Reviewer #2 (Remarks to the Author):

This manuscript describes the structure of the African swine fever virus (ASFV) polymerase complexed with the viral protein M1249L, which is also a capsid component. The authors demonstrate co-purification of M1249L with the polymerase and use cryo-EM analysis to reveal its interaction with the enzyme. Luciferase assays suggest that M1249L enhances polymerase activity.
Our response: This is a conclusion of our manuscript. No response is needed.

Several key concerns warrant further investigation. The authors describe multiple structures and propose two (maybe more) functional states for M1249L. However, interpreting five distinct structures without clear functional evidence requires additional support.
Our response: In addition to extensive and in-depth structural analysis, our manuscript also includes ASFV transcriptome analysis, cell-based functional assays and a comparison of activity differences across various M1249L truncations, which should elucidate the roles of different domains of M1249L (Figure 4i). Our five distinct structures primarily clarify functions based on structural analysis, containing a significant amount of information. Due to space limitations, fully elaborating on all of this within a single article would extend beyond the scope of this manuscript and we think it better to further explore these aspects in future work/publications.

The manuscript readability and figure clarity has to be be improved. The figures are too complex

and difficult to follow and will benefit from simplification for better interpretation.
Our response: We have modified the manuscript and figure panels to strengthen its demonstration, and PDB/map entry indications have been added as well.

For M1249L, the manuscript lacks details on the D1-6 domain structures. Can AlphaFold predict these structures considering M1249L's role as a capsid skeleton protein with a fiber-like, 100 nm long configuration? How do these domains compare to their counterparts in human, yeast, and vaccinia viruses?
Our response: We thank the reviewer for this constructive suggestion. As suggested by the reviewer, we have included details of structural description of D1-6 domains of M1249L in the legend of Extended date Figure 7. On the other hand, AlphaFold cannot predict the conformation of M1249L within the capsid, possibly due to strong interactions with proteins such as p72 affecting the folding of M1249L. The M1249L protein is unique in ASFV, with no homolog counterpart found in other species e.g. human or yeast, except that its binding mode with vRNAP bears certain resemblance with Rap94 protein found in vaccinia viruses.

The authors mention solving the ASFV-vRNAP complex structure within 48 hours, followed by structures from material purified at different time points - and then try to interpret function from these. Also it's unclear which PDB code corresponds to which state. While 23 distinct reconstructions suggest various M1249L-vRNAP binding states, this might simply reflect flexible, heterogeneous binding.
Our response: We thank the reviewer for pointing this out. As suggested, we have included the PDB and EMD codes to the revised figures, legends and main text for clarity. We believe that our constructed structures and maps are not only due to the different stages of M1249L binding to vRNAP in ASFV but also potentially because of flexible/heterogeneous binding. Additionally, M1249L is composed of multiple domains connected by flexible linkers, naturally characterized with flexibility. Regardless, the multiple structures we have resolved can provide valuable insights into the regulatory mechanisms of M1249L during the viral transcription process. Still, our discussion primarily revolves around high-resolution maps and the structures of complexes with specific interactions.

Analysis of one PDB structure (PDBID:8YQZ, DNA complex) reveals potential issues with map interpretation. While the overall model-to-map fit appears reasonable, specific regions lack clear density support and require remodeling. Potential oversharpening of the map is a concern (I havent looked at the other structures but imagine would have similar concerns).
Our response: The fact that specific regions lack clear density support is due to the local domain movement/flexibility which leads to weaker density at these regions. In order to build a solid model, those parts with weaker local density map are constructed with guidance from other higher resolution maps calculated by domain fitting followed by real-space refinement. As for the reviewer’s concern of map oversharpening, during the structure determination of 8YQZ (DNA complex), in order to reveal more structural details in regions of the map with higher resolution, the entire map was sharpened using the RELION software without manually adjusting the B-factor. These specific regions become clearer after applying a low-pass filter.

The modeling of specific chain regions (Chain F:82-339, Chain A:439-508, Chain A:63-96, 131-155) lacks sufficient density support. Repositioning residues in Chain A (439-508) might allow chain break completion. The weak RNA density suggests oversharpening or limitations in current modeling. If these regions were modeled based on modified maps (e.g., deepEMhancer), presenting those maps would strengthen the interpretation.

Our response: We think the reviewer is referring to (Chain F:82-339, Chain B:439-508, Chain B:63-96, 131-155). As mentioned above, that specific regions lack clear density support is due to the local domain movement/flexibility which leads to weaker density. To ensure the accuracy of our constructed structure, we first built and refined the structure of each subunit in the high-resolution map, then fitted it into other slightly weaker maps for structural adjustments. In addition, for regions with weak density, we used a low-pass filter to enhance the signal-to-noise ratio (SNR) and aid with the structural refinement. On the other hand, we do not recommend usage of deepEMhancer to modify the maps, because this is likely to generate artifacts to the map.

The title of Figure 3 ("Functional dissection of M1249L") could be misleading without stronger functional evidence.

Our response: We thank the reviewer for this suggestion, and have modified the title of Figure 3 to "Structural analysis of M1249L as multiple temporary transcription factors".

In summary, the manuscript raises intriguing questions about M1249L's interaction with the ASFV polymerase. However, interpretations of the EM data and proposed functional states require further clarification and supporting evidence.

Our response: As suggested by the reviewer, we have revised the main text to include in-depth analysis and discussion on different conformations of M1249L and its roles in RNA transcription and other viral functions. In terms of functional studies, we combined transcriptomic data analysis, cell experiments (see Fig. 1) using different M1249L constructs besides the abundant structural data. We believe we have sufficient evidence to demonstrate the functions of different domains of M1249L and their roles in RNAP regulation. In addition, we have also toned down the conclusions, focusing primarily on structural comparison and analysis. Given the complexity and the multitude of structures involved, more detailed complex structure studies and functional characterize would be more suitable to be carried out in following works.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

My comments and requests have been taken into account and I thank the authors for improving and clarifying the manuscript considerably.
There are two remaining minor points:
- Looking at Extended Data Figure 5a, here seems to be a substantial difference in vRPB7 compared to the Pilotto structure. This should be stated.
-Concerning the discussion of structural similarities there seems to be a contradiction between a result of a DALI or VAST search (line 230), which did not show any homology and a homology between TFIIB and domain 5 discussed below (line 292).
"Our response: We have rephrased the sentence to S tructure comparison reveals that D5 exhibits certain similarity to TFIIS of RNAPs from human PDB : 8 A 40) and yeast PDB : 1 Y 1 V for clarity (lines 292 293)" of the rebuttal.
Is this explained by a use of DALI and VAST on the full-length structure whereas the TFII homology was inferred up by a domain-wise search or the intermediate structure of the VACV homologue ?
- The phylogenetic analysis (Extended data Figure 6b) is still incomplete.
A larger variety of eukaryotic RNAP sequences should have been included such H. sapiens, M. musculus, D. melanogaster or N. tabacum as a plant exemple. The presented tree only shows an early divergence of vaccinia and ASFV RNAP from the eukaryotes with evolved eukaryotes on one branch and the primitive S. cerivisae on the other branch, but the analysis of more sequences are required to validate a proximity with S scrofa compared to other evolved eukaryotes.

Reviewer #2 (Remarks to the Author):

My comments have been partially addressed. I still think there is too much emphasis that the structures are all functionally relevant - to me there is little evidence for this. But in the main my concerns have been addressed.

Point-to-point response to reviewers' comments

Reviewer #1 (Remarks to the Author):

My comments and requests have been taken into account and I thank the authors for improving and clarifying the manuscript considerably.

Our response: We sincerely appreciate the valuable and constructive feedback provided by the reviewer regarding our former manuscript, as well as the reviewer's acknowledgment of our revised manuscript.

There are two remaining minor points:

- Looking at Extended Data Figure 5a, here seems to be a substantial difference in vRPB7 compared to the Pilotto structure. This should be stated.

Our response: We thank the reviewer for this suggestion. The structures of vRPB7 subunits between ours and Pilotto are almost identical except for only a couple of loops with low density maps. Therefore, we think it appropriate to incorporate the similarity of this subunit into the whole structure comparison (line 175-178 in our revised manuscript).

-Concerning the discussion of structural similarities there seems to be a contradiction between a result of a DALI or VAST search (line 230), which did not show any homology and a homology between TFIIB and domain 5 discussed below (line 292)."Our response: We have rephrased the sentence to Structure comparison reveals that D5 exhibits certain similarity to TFIIS of RNAPs from human PDB : 8 A 40) and yeast PDB : 1 Y 1 V for clarity (lines 292 293)" of the rebuttal. Is this explained by a use of DALI and VAST on the full-length structure whereas the TFII homology was inferred up by a domain-wise search or the intermediate structure of the VACV homologue ?

Our response: We thank the reviewer for pointing this out. The structural similarity search with online tools DALI or VAST is based on automatic full-length structure comparison/search, therefore it is difficult to feedback structures with low similarities. This difficulty is especially enhanced by the fact that domain 5 shares only partial and local similarity with TFIIS, which was discovered by manual inspection and analysis. For clarity, we have modified the sentence to "Manual structural comparison and sequence analysis reveals that D5 exhibits certain similarity to TFIIS of RNAPs from human(PDB: 8A40) 35 and yeast(PDB: 1Y1V)." (line 259-260 in revised manuscript)

- The phylogenetic analysis (Extended data Figure 6b) is still incomplete.

A larger variety of eukaryotic RNAP sequences should have been included such H. sapiens, M. musculus, D. melanogaster or N. tabacum as a plant exemple. The presented tree only shows an early divergence of vaccinia and ASFV RNAP from the eukaryotes with evolved eukaryotes on one branch and the primitive S. cerivisae on the other branch, but the analysis of more sequences are required to validate a proximity with S scrofa compared to other evolved eukaryotes.

Our response: We thank the reviewer for the constructive suggestion. We have included the RNA polymerase sequences from H. sapiens, M. musculus and D. melanogaster to the former phylogenetic analysis result (in revised Supplementary Figure 6b). The result validated our statement that "ASFV vRNAP adopts a structure more closely resembling eukaryotic RNAPs than

its NCLDV homolog of VACV, as confirmed by phylogenetic analysis" (line 180-182 in revised manuscript).

Reviewer #2 (Remarks to the Author):

My comments have been partially addressed. I still think there is too much emphasis that the structures are all functionally relevant - to me there is little evidence for this. But in the main my concerns have been addressed.

Our response: We thank the reviewer for acknowledgment of our revised manuscript. In the course of our cryo-EM structural investigation of the ASFV vRNAP-M1249L complex, we have uncovered a great collection of conformational diversity within the M1249L subunit, a multidomain protein linked by flexible regions. We deduce that these varied conformations not only reflect the inherent flexibility and conformational plasticity of the complex (as the reviewer notes, not all conformations may correspond to specific functions), but are also valuable for understanding the precise mechanisms by which they regulate the transcription process with ASFV vRNAP. We firmly believe that presenting this collection of conformations to our readers will facilitate subsequent research into the specific roles and regulatory mechanisms of these conformations during transcription.