

Cryo-EM structures of Nipah virus polymerases and high-throughput RdRp assay development enable anti-NiV drug discovery

Corresponding Author: Dr Bo Liang

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The manuscript by Chen et al., presents the cryo-EM structure of the full-length apo-Nipah virus (NiV, Malaysia strain) polymerase complex comprising the viral L protein and a tetramer of the viral P protein, at an overall 2.92 Å resolution. As with several other NNS virus L-P complexes, only the RdRp and PRNTase (CAP) domains are observed, the CD, MTase and CTD domains are missing, presumably due to flexibility. In addition, the authors have determined an L-P complex structure from the NiV Bangladesh strain at 3.43 Å resolution, using a truncated L protein lacking the CD, MTase and CTD domains, as well as a partial dimer of the full-length Malaysia strain complex. Even though the Malaysia and Bangladesh proteins are 98.2 % identical, the Bangladesh strain is more pathogenic. The authors first show that their recombinant complexes are active in template-dependent RNA synthesis and develop a high-throughput assay that could be used to screen for small molecule inhibitors. They then describe the various structures, focussing on the differences observed between the strains, the contacts made by the P protein to L, and on the identification of a putative dimer, in which it is proposed (based on low resolution density) that the CD domain from a second L protein is docked onto the RdRp of the first L, as previously observed for human parainfluenza virus type 3 (hPIV3).

Given the high mortality of this emerging virus, which causes local serious local epidemics, these first structures of the polymerase complex are an important step forward to deeper understanding of the molecular basis for replication and transcription of the viral genome and in providing a target for structure-assisted drug design. On the other hand, the structure of the apo- L-P complex is very similar to other paramyxovirus L-P complexes. Furthermore, the manuscript is rather descriptive providing few new insights into the mechanism of action of this kind of polymerase complex.

Major issues.

(1) The section on RNA synthesis assays (lines 95-118) skirts over a number of issues that should be commented on and explained better, especially for the uninitiated. In Figure 1b, why is there such a striking difference between the truncated NiV-B protein, which synthesises a single product, whereas the full-length protein produces a multitude of products? Is it the same for the NiV-M strain? If the authors think this is due to the heterogeneity of the full-length protein and dimerization, then this should be discussed. Also, the full-length M and B strains are different In Fig 1b and Fig S1a and b. Why does the NiV-B strain produce more long products? What is back-priming, which products are 'back-primed' products and what is the relevance of the diagrams at the bottom of Figure 1c? Given the diversity of products, what justifies saying that Le and TrC10-25 are the 'preferred templates'? Why are many of the products stalled at 14 nts? What is meant by a 4-fold and 6-fold signal window and how are these numbers derived? What is the tool compound? The radioactive gel in Fig. 1h shows still considerable activity at 100 micromolar, whereas the others assays are essentially down to zero. The radioactive gel should be removed or done at more concentrations so that it can be quantified. Concerning the values of the apparent Km and NTP, are these different for NiV-M and NiV-B, as this could give insight into the difference in pathogenicity (see also point (2)).

(2) Structural differences between fl NiV-M L-P and tr NiV-B L-P. Could some of the observed structural differences be due to the fact that NiV-M L is full-length (which is more heterogeneous) and NiV-B L is truncated? It would be better to make the comparison with the same construct for each strain. It is speculated that some of the strain differences could affect RdRp activity. It would be better to have some test of this, at least determination of the apparent Km and NTP for both strains (with comparable constructs).

(3) For the L-P dimer, the quality of the AF prediction of the NiV CD domain needs to be given (this should be standard for any AF model that is used!). A clear diagram showing the fit of this domain into the extra density needs to be given, to convince the reader that it is true, independently of the similarity to the hPIV3 structure.

(4) For the working model, is this transcription or replication? How would they differ ?

Reviewer #2

(Remarks to the Author)

The manuscript by Chen et al., presents a set of new in-vitro assays to assess Nipah polymerase activity. They developed two assays that have the benefit of not relying on radioactive NTPs and are scalable in 384 wells plate to facilitate future high throughput compounds screening. They also describe cryo-EM structures of the apo-Nipah virus polymerase complex for both the Malaysia and Bangladesh strains. The structures are composed of the viral L protein (full-length or truncated) and a tetramer of the viral P protein, at reported resolutions ranging from 2.9Å to 3.5Å. As it is often the case for most of L-P structures from other related viruses, the CD, MTase and CTD domains are missing, presumably due to their intrinsic flexibility. The authors also observed extra density in one of the structures which they attributed to the connecting domain (CD) of a second L protein and elaborated a speculative model about the function of L-P dimers.

Nipah virus is a major pathogen, highly lethal, with potential risk of creating outbreak. Therefore, a better understating of the mechanism behind key viral factors is an asset to prepare against these potential outbreaks. Nevertheless, regarding this manuscript, also the developed in-vitro assays seem robust and will allow the initial high-throughput screening of compounds targeting the active site, the mechanistic insight gained from this study is minimal. The structural work only resolved the RdRP and PRNTase domains bound to P which turned out to be very similar to that of related viruses. There is no template/product RNA used to capture intermediate reactions and there is an absence of functional data to validate the interfaces observed in the structures (e.g. between L and CD domain or L and P). There is no real connection between the structural study and the developed in-vitro assays beyond confirming the activity of the recombinantly expressed complex. Finally, most of the discussion and mechanistic hypothesis are driven by existing literature and not from data generated for this manuscript.

Beyond the insight of this paper, I have concerns about the quality of some of the cryo-EM data and the conclusion derived from them, especially regarding the "L:P dimer".

Major points.

(1) The cryo-EM data for the Malaysian strain are very good while the cryo-EM map for the Bangladesh strain is a bit distorted but still good enough for model building. However, the refinement and model statistics presented in supplementary table 1 show that the quality of the model building is poor. The % of rotamers outliers should be close to 0 while it is at 3.83% and 6.1% for the Malaysia and Bangladesh strains, respectively. The clash scores are also high respectively to the reported resolution and further work is absolutely needed to improve them. This raise concerns about the confidence in the analysis done on the local variation of residues not conserved between the Malaysia and Bangladesh strains described in lines 183-193, where the authors comment about the movement of side chains by couple of Angstrom.

(2) Overall, the "L:P dimer" observation is extrapolated and not supported enough by the data as explained in the following points.

(3) There is no data that supports confidently that the observed extra density belongs to the CD. The overall quality of the map is good however the signal for the observed extra density is really poor and doesn't allow an unambiguous fit. The attribution of the density and the identification of an "L:P dimer" relies exclusively on the PIV-3 structures recently published (PDBID: 8KDB) rather than the actual data. In the PIV-3 structure, the authors captured a fully resolved L (including the CD) bound to an extra CD which could only come from another L. This is not the case here and the extra density could also correspond to the CD from the same L.

(4) Line 261-266, the authors justified that the CD density was unlikely to come from the same L by saying it was too far to connect with the PRNTase domain. However, this statement is not correct. As the authors didn't provide any model for the CD, I followed their method and used the PIV3 structure onto which I superimposed the CD domain predicted using AF3. At the suggested threshold and only on the unsharpened map, less than half of the domain was inside the density (and there is no density for this domain in the sharpened map). Nevertheless, by measuring the distance between residues P1446 (last residue modelled in the linker) and S1486 (which is at the end of the N-terminal helix of the CD), the distance is about 68 Ang. The linker between the domains is then composed of 39 residues which makes it possible to connect the two even if they belong to the same polymerase. Of note, the absence of pdb model deposited or provided by authors for the "L:P dimer map" suggests that the authors are also not confident in the interpretation of this extra density.

(5) Regarding the use of 2D classes to support the "L:P dimer" hypothesis (line 268-274), the image on the left fig. 5d shows an RdRp domain + some extra density while the image on the right fig. 5d shows an RdRp domain + some extra density and a blurry density for an adjacent RdRp domain. Based on the cartoon diagram next to the images and the yellow arrows, this clearly demonstrates that a monomer (image of the left panel) is having the same extra density than the image on the right, further invalidating the "L:P dimer" hypothesis. In addition, as mentioned by the authors lines 250-252, the sample on the grid was highly aggregated and it is therefore very likely that the adjacent particles correspond to particles that were stochastically close on the micrographs as it is often the case. This becomes even more obvious when compared to 2D class averages of real polymerase dimers from other NSV.

(6) Better structural data and/or some mutagenesis of the interface between the RdRp and CD domain are absolutely

required to first validate that the density is the CD and to assess whether it has any functional relevance. L-P activity in a cell-based minireplicon assay is likely to be the most relevant assay to assess the functionality.

Minor points

(1) The literature seems to mainly use PRNTase domain instead of Cap domain. I think keeping the nomenclature consistent will improve the clarity.

(2) The cartwheel model is not the only existing one and there is no evidence in the literature that it is the most likely model. The authors should reference all possible models.

(3) Although the Nipah L-P complex hasn't been published in peer reviewed journal, there is already two pre-prints available since couple of months that might be worth discussing or at least referenced (<https://doi.org/10.1101/2024.05.29.596445>, <https://www.researchsquare.com/article/rs-4663080/v1>).

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The manuscript has been considerably improved based on the response to the referees very pertinent criticisms. Some new data has been included (and some previous poor data excluded), explanations are clearer, the analysis more rigorous and the conclusion more restrained.

However there are still a few issues that need to be addressed in a final version.

(1) In Figure 1a and b, comparative gels are shown for NiV-M and NiV-B. These should be displayed on the same vertical scale, if possible, to facilitate the comparison.

(2) In Figure 5 the fit of the CD to the extra density would be made more convincing if the same two projections were used as in Figure S7b to exclude the MT and CTD domains.

(3) Add NiV to the summary Table in Figure S10h.

(4) In Supp. Table 1, the statistics for the fl M L-P-CD (PDB:9MZH) structure are particularly poor (e.g. Ramachandran favoured 87.73 %). For a 3.67 Å structure like this, one would normally start from a higher resolution similar structure, remove parts for which there is not decent density and just do rigid body refinement. If more general refinement is done one risks to degrade the structure, as seems to be happening here. Unless there are major differences in the structure, which should be highlighted. In addition all the structures have rather high clash scores, particularly the truncated ones.

Reviewer #2

(Remarks to the Author)

Thanks to the authors for going through my comments and addressing them. I can see that most of the issues highlighted have been taken out of the manuscript or addressed. It is still a difficult manuscript to review in line with the current literature. Indeed, there are already 7 publications presenting the same cryo-EM structure of Nipah-M in complex with P, most of them with a much higher resolution and map quality in addition to a fully resolved polymerase elongation complex with the RNA template and product. Altogether, these publications already allowed a deeper understanding of the virus polymerase than what is described in this manuscript.

As I said in my previous review, there is no real connection between the structural study and the developed in-vitro assays, and the novel insight gain from the structural work is very minimal. The only possible novelty is about the position of the CD domain in one of the structures. But the insight is very limited by the poor structural data where only a single alpha helix could be built properly in density and the lack of functional data supporting whether such interaction has any functional relevance.

Beyond the lack of novel insight in this manuscript, I still have two major concerns about the cryo-EM data:

- The map of FI NiV-M-L-P-CD and therefore the model derived from this map is expanded compared to all the other PDB models from this manuscript but also already existing structures. It is easy to visualize by aligning all the models to one side of the protein structure (like Aas 200-250) and looking at the difference in size. This is something not so uncommon in the cryo-EM field and is the result of using the wrong pixel size for the cryo-EM processing. The authors must revise their processing and model building using the correct pixel for this data set. I also noticed that the magnification in the Supp. Table 1 for this data set is incorrect, 110,000x doesn't exist on 300 kV microscope.

- The data for the Bangladesh strain are not good enough for publication and that is reflected by the high clashscore of 16.8. The density is lacking continuity even at secondary structure level. However, judging from the data processing flowchart, this is something that the authors could improve by doing a focused refinement on the RdRp-PRNTase to minimize the impact of the flexible P.

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

Thanks to the authors for looking into my comments. Overall, the model building has been greatly improved. There are still some issues with the supplementary table one and the data collection parameters.

For the tr NiV-M L-P dataset, a pixel size of 0.927 seems to correspond to the calibration of a Selectris-Falcon IV detector rather than a Gatan K3 camera and energy filter like mentioned in the methods. This needs to be corrected.

Regarding the FI NiV-M L-P-CD, the authors greatly improved the map, especially around the connecting domain. However, I'm sorry for being insistent but the provided pixel size value of 0.874 Å/pixel is still suspicious. The average calibrated pixel size value for a K3 detector on Titan Krios across facilities around the world seems to be around 0.83 Å/pixel with of course some microscope dependent deviation. A quick online search illustrates this:

0.83 Å/pixel - The University of Tokyo, Japan
0.845 Å/pixel - University of California, Berkeley
0.825 Å/pixel - New York Structural Biology Center
0.834 Å/pixel - MRC Laboratory of Molecular Biology
0.831 Å/pixel - eBIC - Diamond Light Source
0.834 Å/pixel - Janelia Research Campus
0.835 Å/pixel - Stanford-SLAC Cryo-EM Center
0.827 Å/pixel - Max Planck Institute of Biochemistry
0.84 Å/pixel - Yale University West Campus Cryo-EM Facility
0.825 Å/pixel - NCCAT at the New York Structural Biology Center

The pixel size for your expended structure is 0.874 Å/pixel, which correspond to a deviation of about 5% relative to the mean 0.83 Å/pixel. If I now look at your structures and superimpose the tr NiV-M L-P to the FI NiV-M L-P-CD on residues 200-250 as mentioned before, I can measure the distance between the C α of two residues from each side of the active site. I randomly took Phe218 and Leu873. In the tr NiV-M L-P the distance is 56.7 Å while in the FI NiV-M L-P-CD it is 59.65 Å, which correspond to a deviation of 4.95%, comparable to the deviation between the provided pixel size and the mean pixel size for this camera/microscope/magnification configuration.

From a biological point of view, there are plenty of example of viral polymerase structures solved in various conditions and such distortion of the active site doesn't seem to have been observe. To not report something nonphysiological, I would strongly advise the authors to contact the facility where the dataset has been collected and ask them to collect a high-resolution dataset of protein standard such as ApoF. Using existing structure in the pdb, it will quickly allow you to accurately determine the exact pixel size for this magnification, which will be beneficial for you and them.

Using existing Nipah polymerase structures solved in the same conformation than the FI NiV-M L-P-CD (e.g. pdb 9BDQ or 9FUX), one can try estimating what the pixel size should be by looking at the percentage of included atoms inside the map for every possible voxel. By doing that, the estimated pixel size is 0.839 Å/pixel using 9BDQ and 0.842 Å/pixel using 9FUX, in agreement with expected pixel size on a K3-EF detector fitted on a titan Krios.

Interestingly, once using the newly estimated pixel size. The fitting of the connecting domain predicted with AF3 is significantly better than before. The quality of the map is sufficient to built and refine most of the connecting domain as opposed to a single helix. Looking at the processing flowchart, there is still room for improvement to get better maps. First, using the proper calibrated pixel size will allow the authors to have a better per particle defocus estimation. They can also perform a reference-based motion correction and a local refinement on the core + CD after subtracting the density for the P proteins which should improve the density for the both the core and the CD.

Version 3:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

Thank you to the authors for taking my feedback into consideration and double checking the pixel size of their dataset with the facility. Having an estimated spherical aberration close to 2.7 mm after adjusting the pixel size of 0.84 Ang/px is an extra confirmation that this is the correct approximation. The map and model are better and allow a proper interpretation of the interface with the CD domain. I don't have any further concerns.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewer 1:

The manuscript by Chen et al., presents the cryo-EM structure of the full-length apo-Nipah virus (NiV, Malaysia strain) polymerase complex comprising the viral L protein and a tetramer of the viral P protein, at an overall 2.92 Å resolution. As with several other NNS virus L-P complexes, only the RdRp and PRNTase (CAP) domains are observed, the CD, MTase and CTD domains are missing, presumably due to flexibility. In addition, the authors have determined an L-P complex structure from the NiV Bangladesh strain at 3.43 Å resolution, using a truncated L protein lacking the CD, MTase and CTD domains, as well as a partial dimer of the full-length Malaysia strain complex. Even though the Malaysia and Bangladesh proteins are 98.2 % identical, the Bangladesh strain is more pathogenic. The authors first show that their recombinant complexes are active in template-dependent RNA synthesis and develop a high-throughput assay that could be used to screen for small molecule inhibitors. They then describe the various structures, focussing on the differences observed between the strains, the contacts made by the P protein to L, and on the identification of a putative dimer, in which it is proposed (based on low resolution density) that the CD domain from a second L protein is docked onto the RdRp of the first L, as previously observed for human parainfluenza virus type 3 (hPIV3).

Given the high mortality of this emerging virus, which causes local serious local epidemics, these first structures of the polymerase complex are an important step forward to deeper understanding of the molecular basis for replication and transcription of the viral genome and in providing a target for structure-assisted drug design. On the other hand, the structure of the apo- L-P complex is very similar to other paramyxovirus L-P complexes. Furthermore, the manuscript is rather descriptive providing few new insights into the mechanism of action of this kind of polymerase complex.

Response:

We thank the reviewer for the comments on our work. We have revised the manuscript accordingly and listed our response below:

Major issues.

(1) The section on RNA synthesis assays (lines 95-118) skirts over a number of issues that should be commented on and explained better, especially for the uninitiated.

In Figure 1b, why is there such a striking difference between the truncated NiV-B protein, which synthesises a single product, whereas the full-length protein produces a multitude of products? Is it the same for the NiV-M strain? If the authors think this is due to the heterogeneity of the full-length protein and dimerization, then this should be discussed.

Thank you for the comments. In Fig. 1b, the product difference between the truncated and full-length NiV-B polymerase is due to their different RdRp activities. The truncated (L RdRp + PRNTase domains & full-length P) generates fewer bands that are all longer than the template because it has back priming activity on the Le25 and TrC25 template. The template size is 25 nt and the products are 26-28 nt. Back priming activity means that the RNA template forms a hairpin

through base pairing at its 3' end and RNA synthesis is initiated from the RNA double strand that is formed in the process. This is now better explained in the text and we have updated Fig. 1b. In contrast, the full-length protein has both back priming activity and de-novo RdRp activity as evidenced by products that are longer and shorter than 25 nt. In the literature, this phenomenon has been reported for RSV and EBOV (PMID: 25010481, 33028717, 34914795, 37699521). To better understand if this difference in activity was due to the difference in strains being used in this assay or was a result of the L protein truncation, we also generated the truncated protein of NiV-M (L RdRp + PRNTase domains & full-length P). A comparison of both full length and truncated proteins in this assay showed consistent results: the lack of de novo RdRP activity in the truncated protein. Based on the predicted secondary RNA structures of the Le25 and TrC25 RNA, we only supplied ^{32} P-ATP in the RdRp assay and further corroborated the back priming activity of all four L-P constructs on the TrC25 RNA. In contrast, the Le25 sequence is predicted to form a hairpin with a mismatch at the 3' end and as a result RNA template extension is compromised.

This is now updated in Fig. 1b & c and reflected in the track change version of our manuscript.

We think the RdRP activity difference is not due to the heterogeneity of the full-length protein or dimerization, as we used the L-P protein samples that were eluted in well-separated peaks after 10 mL in Superose 6 column for all the assays, if not specified.

Also, the full-length M and B strains are different In Fig 1b and Fig S1a and b. Why does the NiV-B strain produce more long products?

We appreciate the reviewer's comment. We think this difference is due to different protein conditions including the freeze-thaw cycles and concentrations used for the Fig 1b and Fig S1a and b. We repeated the experiments for the Fig S1a and b by using same protein concentration (~200 nM L-P), our results show similar products using full-length proteins of the 2 different strains (revised Fig. S1d).

What is back-priming, which products are 'back-primed' products and what is the relevance of the diagrams at the bottom of Figure 1c?

Thank you for the comments. As mentioned above, back priming is a mechanism used by some RdRps to initiate RNA synthesis due to the base pairing formation at the 3' end of the RNA template. Therefore, the back-primed products are longer than the RNA template itself, for example, in revised Fig. 1b, the products longer than 25 nt (~26-28 nt) are the back-primed ones. The bottom diagrams of Fig 1c was to show the predicted 3' base pairing in the RNA templates we used for the assays and indicate the products sizes. Now we moved it to the revised Fig. 1c to better explain the back priming activity we found.

For further validation, based on the RNA secondary structure predication we supplied α - ^{32} P-ATP with Le25 or TrC25 templates (revised Fig. 1c), our results clearly show that both fl and tr

polymerase of two strains have strong back priming activity on TrC25, but it's very weak on the Le25, which is likely due to the unpaired "U" at the 3' end of Le25.

Given the diversity of products, what justifies saying that Le and TrC10-25 are the 'preferred templates'?

Thank you for the comments. We tried different lengths of Le/TrC RNA, and found Le and TrC 10-25 were the templates that worked to produce multiple product bands. We justified the preferred templates by comparing the product band intensity. We realized this may not be proper, so we have removed the sentence including the claim about the "preferred templates".

Why are many of the products stalled at 14 nts?

Thank you for the comments. We have tried both α - ^{32}P -GTP and α - ^{32}P -ATP and found this phenomenon. We are not sure why this is happening. In literature, similar results were reported for the RSV polymerase (PMID: 25010481 and 33028717). We propose that the dominant bands ($\leq 14\text{nt}$) are likely abortive initiation products that are generated by NiV polymerase that could not move away from the RNA promoter (PMID: 25010481).

What is meant by a 4-fold and 6-fold signal window and how are these numbers derived? What is the tool compound? The radioactive gel in Fig. 1h shows still considerable activity at 100 micromolar, whereas the others assays are essentially down to zero. The radioactive gel should be removed or done at more concentrations so that it can be quantified.

Thank you for the comments. A 4-fold and 6-fold signal window refers to the ratio of the fluorescent or luminescent signal at the concentration of enzyme tested and to the fluorescent or luminescent signal in the absence of protein. The tool compound is a proprietary compound from the Novartis library. We have removed the radioactive gel result in Fig. 1h.

Concerning the values of the apparent K_m and NTP, are these different for NiV-M and NiV-B, as this could give insight into the difference in pathogenicity (see also point (2)).

Thank you for the comments. The K_m values for NTP and RNA are comparable between the truncated NiV-M and NiV-B L-P. The NiV-B apparent K_m for RNA and NTP were determined as $\sim 5.6 \mu\text{M}$ and $\sim 22.7 \mu\text{M}$, respectively. The NiV-M apparent K_m for RNA and NTP were approximately $3.3 \mu\text{M}$ and $21 \mu\text{M}$, respectively. We don't have any evidence that there is a difference in activity between the polymerase of both strains and therefore cannot speculate on the pathogenicity of these two strains.

(2) Structural differences between fl NiV-M L-P and tr NiV-B L-P.

Could some of the observed structural differences be due to the fact that NiV-M L is full-length (which is more heterogeneous) and NiV-B L is truncated? It would be better to make the comparison with the same construct for each strain. It is speculated that some of the strain

differences could affect RdRp activity. It would be better to have some test of this, at least determination of the apparent K_m and NTP for both strains (with comparable constructs).

Thank you for the comments. We agree with the suggestion that it would be better to make the comparison with the same construct for each strain. We have generated the truncated NiV-M protein and determined its cryo-EM structure. Based on our new and improved structural data, we have updated the structural comparison in revised Fig. 3. We also determined the apparent K_m for NTP and RNA for both strains using the truncated proteins.

This is now updated in Fig. 1, 3 and S1 and reflected in the track change version of our manuscript.

(3) For the L-P dimer, the quality of the AF prediction of the NiV CD domain needs to be given (this should be standard for any AF model that is used!). A clear diagram showing the fit of this domain into the extra density needs to be given, to convince the reader that it is true, independently of the similarity to the hPIV3 structure.

Thank you for the comments. We have provided the quality of the AF models in the Method part using the mean predicted local distance difference test (pLDDT) and the uncertainty. The mean pLDDT of the CD domain model is 89.31, and the uncertainty is 4.98%.

We have optimized our data processing and determined the structure of fl NiV-M L-P-CD, which enables us to show the fit of a 22-residue fragment of the CD domain to the extra density. We also made AF models for the MT domain (mean pLDDT = 81.62; uncertainty = 6.7%) and CTD domain (mean pLDDT = 88.83; uncertainty = 1.35%) and found no reasonable fit between these two domains' models and the extra density.

This is now updated in Fig. 5 and S7 and reflected in the track change version of our manuscript.

(4) For the working model, is this transcription or replication? How would they differ ?

Thank you for the comments. This working model was for replication and we highlighted that the P-NTD and the N⁰ work for the encapsulation of the nascent RNA. Transcription means the RdRp generates individual transcripts for genes which are 5' capped and methylated and polyadenylated at the 3'. Replication is for the whole genome of the virus. Our working model in Fig. 6 was largely based on our hypothesis of the L-P dimer. Since we were not able to sufficiently corroborate the formation of L-P dimers in follow up experiments, we would like to withdraw Fig. 6c.

Reviewer 2

The manuscript by Chen et al., presents a set of new in-vitro assays to assess Nipah polymerase activity. They developed two assays that have the benefit of not relying on radioactive NTPs and are scalable in 384 wells plate to facilitate future high throughput compounds screening. They also describe cryo-EM structures of the apo-Nipah virus polymerase complex for both the

Malaysia and Bangladesh strains. The structures are composed of the viral L protein (full-length or truncated) and a tetramer of the viral P protein, at reported resolutions ranging from 2.9Å to 3.5Å. As it is often the case for most of L-P structures from other related viruses, the CD, MTase and CTD domains are missing, presumably due to their intrinsic flexibility. The authors also observed extra density in one of the structures which they attributed to the connecting domain (CD) of a second L protein and elaborated a speculative model about the function of L-P dimers.

Nipah virus is a major pathogen, highly lethal, with potential risk of creating outbreak. Therefore, a better understating of the mechanism behind key viral factors is an asset to prepare against these potential outbreaks. Nevertheless, regarding this manuscript, also the developed in-vitro assays seem robust and will allow the initial high-throughput screening of compounds targeting the active site, the mechanistic insight gained from this study is minimal. The structural work only resolved the RdRP and PRNTase domains bound to P which turned out to be very similar to that of related viruses. There is no template/product RNA used to capture intermediate reactions and there is an absence of functional data to validate the interfaces observed in the structures (e.g. between L and CD domain or L and P). There is no real connection between the structural study and the developed in-vitro assays beyond confirming the activity of the recombinantly expressed complex. Finally, most of the discussion and mechanistic hypothesis are driven by existing literature and not from data generated for this manuscript.

Beyond the insight of this paper, I have concerns about the quality of some of the cryo-EM data and the conclusion derived from them, especially regarding the “L:P dimer”.

Response:

We want to thank you for your comments on our work. We appreciate your efforts and constructive suggestions in this review.

Major points.

(1) The cryo-EM data for the Malaysian strain are very good while the cryo-EM map for the Bangladesh strain is a bit distorted but still good enough for model building. However, the refinement and model statistics presented in supplementary table 1 show that the quality of the model building is poor.

The % of rotamers outliers should be close to 0 while it is at 3.83% and 6.1% for the Malaysia and Bangladesh strains, respectively. The clash scores are also high respectively to the reported resolution and further work is absolutely needed to improve them. This raise concerns about the confidence in the analysis done on the local variation of residues not conserved between the Malaysia and Bangladesh strains described in lines 183-193, where the authors comment about the movement of side chains by couple of Angstrom.

Thank you for your comments. We have improved our structures on the refinement and model statistics, now the % of rotamers outliers for the Malaysia full-length and Bangladesh truncated L-P are 0.65% and 0%, respectively. We could also slightly decrease the clash scores for these

two structures. Thank you for pointing out the potential issues with our analysis on the local residues' structural variation between the different NiV strains, we found that may not be accurate based on our updated structures and have removed the content on side chains' differences in lines 183-193 from the old manuscript.

In this revision, we determined two more structures: the truncated NiV-M L-P and the full-length NiV-M L-P structure containing part of the CD domain. We have tried our best to improve the models and updated the cryo-EM structure statistics in Supplementary-Table.1.

(2) Overall, the “L:P dimer” observation is extrapolated and not supported enough by the data as explained in the following points.

Thank you for the comments. We have updated the manuscript to reflect the response to the comments.

(3) There is no data that supports confidently that the observed extra density belongs to the CD. The overall quality of the map is good however the signal for the observed extra density is really poor and doesn't allow an unambiguous fit. The attribution of the density and the identification of an “L:P dimer” relies exclusively on the PIV-3 structures recently published (PDBID: 8KDB) rather than the actual data. In the PIV-3 structure, the authors captured a fully resolved L (including the CD) bound to an extra CD, which could only come from another L. This is not the case here, and the extra density could also correspond to the CD from the same L.

Thank you for the critical comments. We optimized the image processing strategy by adding a 3D classification step in cryoSPARC, we could select 75,538 particles for obtaining a 3.67 Å map and build in the extra density an 18-residues long helix plus flanking residues that belong to the CD domain (S1624 - S1645). We show that the AF CD model fits the extra density but the MT/CTD domain model cannot. We agree that we do not have direct evidence to support the identification of the “L:P dimer”. In the revised version we don't claim the dimer. We focus on the discovery of the binding and partial interface between the RdRp-PRNTase and CD domain of NiV L protein.

This is now updated in Fig. 5 & S7 and reflected in the track change version of our manuscript.

(4) Line 261-266, the authors justified that the CD density was unlikely to come from the same L by saying it was too far to connect with the PRNTase domain. However, this statement is not correct. As the authors didn't provide any model for the CD, I followed their method and used the PIV3 structure onto which I superimposed the CD domain predicted using AF3. At the suggested threshold and only on the unsharpened map, less than half of the domain was inside the density (and there is no density for this domain in the sharpened map). Nevertheless, by measuring the distance between residues P1446 (last residue modelled in the linker) and S1486 (which is at the end of the N-terminal helix of the CD), the distance is about 68 Ang. The linker between the domains is then composed of 39 residues which makes it possible to connect the two even if they belong to the same polymerase. Of note, the absence of pdb model deposited or provided by

authors for the “L:P dimer map” suggests that the authors are also not confident in the interpretation of this extra density.

Thank you for your efforts in reviewing our original manuscript. As answered in last point of response, we have improved the cryo-EM density containing the extra density for CD to a 3.67 Å resolution. Using the improved density map, first, the AF CD domain model could fit the extra density at ChimeraX map level=0.15; second, we built an atomic model containing RdRp-PRNTase of L, P protein and the residues S1624–S1645 of the CD domain of L. We named this structure as fl NiV-M L-P-CD and deposited the PDB model (PDB ID: 9MZH).

For the distance measurement between the C terminus of the PRNTase domain in our structure and the N terminus of the CD domain model, we apologize that our previous selection of the end residues may not be appropriate, making the claim not accurate. We agree with the reviewer's modeling that it is possible that the extra CD domain comes from the same polymerase. In this revision, we used P1446 (last residue modeled in the linker) and L1481 (the N terminal residue in the newly reported crystal structure of the CD domain, PDB ID: 9FTF), the distance is 70.8 Å, which is possible for a 34-residue link to connect the two domains in a single polymerase.

This is now updated in Fig. S7f and reflected in the track change version of our manuscript.

(5) Regarding the use of 2D classes to support the “L:P dimer” hypothesis (line 268-274), the image on the left fig. 5d shows an RdRp domain + some extra density while the image on the right fig. 5d shows an RdRp domain + some extra density and a blurry density for an adjacent RdRp domain. Based on the cartoon diagram next to the images and the yellow arrows, this clearly demonstrates that a monomer (image of the left panel) is having the same extra density than the image on the right, further invalidating the “L:P dimer” hypothesis.

In addition, as mentioned by the authors lines 250-252, the sample on the grid was highly aggregated and it is therefore very likely that the adjacent particles correspond to particles that were stochastically close on the micrographs as it is often the case. This becomes even more obvious when compared to 2D class averages of real polymerase dimers from other NSV.

Thank you for the comments. We realized that our original Fig. 5d may not be able to support our L-P dimer hypothesis, and it is possible that the adjacent particles in our 2D class averages are from the particles that are randomly close with each other. We have removed the Fig. 5a & d. We also removed the Fig. S7 that involved using 2D classes of two or more RdRp-PRNTase ring shaped particles.

In this revision, we use three 2D class averages to show it matches the three corresponding projections of our fl NiV-M L-P-CD structure superposed with the AF model of CD, thus supporting our model building of the fragment of CD domain in our fl NiV-M L-P-CD structure. We don't claim that it comes from the same polymerase or a second one.

This is now updated in Fig. S7c and reflected in the track change version of our manuscript.

(6) Better structural data and/or some mutagenesis of the interface between the RdRp and CD domain are absolutely required to first validate that the density is the CD and to assess whether it has any functional relevance. L-P activity in a cell-based minireplicon assay is likely to be the most relevant assay to assess the functionality.

Thank you for the comments. As mentioned above, we have improved our structural data of the partial interface between the RdRp-PRNTase and the CD domains. We could build a long α -helix of CD domain into the density and found hydrophobic interactions in the partial interface, as shown in the revised Fig. 5 and S7. We cannot rule out the possibility that this conformation of the NiV polymerase containing the RdRp-PRNTase – CD interface is an artifact, so we don't intend to imply any biological function relevance.

Minor points

(1) The literature seems to mainly use PRNTase domain instead of Cap domain. I think keeping the nomenclature consistent will improve the clarity.

Thank you for the comments. We have changed “Cap domain” to “PRNTase domain” in the text.

(2) The cartwheel model is not the only existing one and there is no evidence in the literature that it is the most likely model. The authors should reference all possible models.

Thank you for the comments. Our working model in the original Fig. 6 was based on our hypothesis on an L-P dimer formation. Since now we found our data is not sufficient to support the identification of L-P dimers, we would like to withdraw the proposed working model.

(3) Although the Nipah L-P complex hasn't been published in peer reviewed journal, there is already two pre-prints available since couple of months that might be worth discussing or at least referenced (<https://doi.org/10.1101/2024.05.29.596445>, <https://www.researchsquare.com/article/rs-4663080/v1>).

Thank you for the comments. We have referenced the published papers (including these two preprints) on the cryo-EM structures of the NiV L-P complex in Discussion:

Additional groups have made efforts to understand the structure of the NiV polymerase complex. Structures with resolutions ranging from 1.85 - 3.19 Å have been made available. The published structures of the complex and the identified interactions between the L protein and P protein align with and enhance our findings. However, the L protein back-priming capabilities and structural analysis between two NiV strains presented here provides unique structural and functional insights.

Reviewer #1 (Remarks to the Author):

The manuscript has been considerably improved based on the response to the referees very pertinent criticisms. Some new data has been included (and some previous poor data excluded), explanations are clearer, the analysis more rigorous and the conclusion more restrained. However there are still a few issues that need to be addressed in a final version.

Response:

We thank the reviewer for the comments. We have addressed all the issues and listed our responses point-by-point below.

Major issues.

(1) In Figure 1a and b, comparative gels are shown for NiV-M and NiV-B. These should be displayed on the same vertical scale, if possible, to facilitate the comparison.

Thank you for the comments. For the Figure 1a, we have rerun the SDS-PAGE using the same protein marker and updated the image for the FL and TR L-P of NiV-B. Our primary purpose with Figure 1b is to show the de novo and back priming RdRp activities in FL L-P and only the back priming products with the TR L-P from NiV-M and NiV-B. We propose to use the same vertical scale in the Fig 1c and S1d for the comparison of these RdRp activities of the L-P complexes from the two NiV strains.

(2) In Figure 5 the fit of the CD to the extra density would be made more convincing if the same two projections were used as in Figure S7b to exclude the MT and CTD domains.

Thank you for this key comment. We have optimized Figure 5b to better show the fit between the CD domain model and the extra density, by using the same two projections that were used in Figure S7b.

(3) Add NiV to the summary Table in Figure S10h.

Thank you for the comments. We have used both of our NiV-M and NiV-B L-P structures in the summary and added "NiV" to the summary table.

(4) In Supp. Table 1, the statistics for the fl M L-P-CD (PDB:9MZH) structure are particularly poor (e.g. Ramachandran favoured 87.73 %). For a 3.67 Å structure like this, one would normally start from a higher resolution similar structure, remove parts for which there is not decent density and just do rigid body refinement. If more general refinement is done one risks to degrade the structure, as seems to be happening here. Unless there are major differences in the structure, which should be highlighted. In addition all the structures have rather high clash scores, particularly the truncated ones.

We appreciate the reviewer's comment. In the beginning, we tried the suggested method for the modeling of FL M L-P-CD by using the FL NiV-M L-P structure in this study, however, the model

does not fit into the map of FL M L-P-CD. We shifted the coordinates of many residues a bit in the starting model and ran the refinement to get the final model in the first revision.

In this revision, we have improved the model of FL M L-P-CD to have Ramachandran favoured 92.37 %, allowed 7.55% and clash score at 6.91. We found there is a major difference that the RdRp-PRNTase of FL M L-P-CD expands a bit from the FL M L-P structure. We have highlighted this finding in the main text and Figure 5d.

We have refined the structures of the truncated NiV M & B L-P complexes, now the clash scores are 8.46 and 7.52, respectively.

Reviewer #2 (Remarks to the Author):

Thanks to the authors for going through my comments and addressing them. I can see that most of the issues highlighted have been taken out of the manuscript or addressed. It is still a difficult manuscript to review in line with the current literature. Indeed, there are already 7 publications presenting the same cryo-EM structure of Nipah-M in complex with P, most of them with a much higher resolution and map quality in addition to a fully resolved polymerase elongation complex with the RNA template and product. Altogether, these publication already allowed a deeper understanding of the virus polymerase than what it is described in this manuscript.

As I said in my previous review, there is no real connection between the structural study and the developed in-vitro assays, and the novel insight gain from the structural work is very minimal. The only possible novelty is about the position of the CD domain in one of the structures. But the insight is very limited by the poor structural data where only a single alpha helix could be built properly in density and the lack of functional data supporting whether such interaction has any functional relevance.

Response:

We thank the reviewer for the comments. We have cited all 7 publications on the NiV L-P structures and its elongation complex. We acknowledge the reviewer's concern regarding the lack of connection between the structural study and the developed in-vitro assays, however, we are confident that the findings reported here, especially the structure of FL NiV L-P-CD and the in-vitro biochemical assays, can lay the foundation for future studies on NiV polymerase, as well as henipavirus drug discovery.

We present here for the first time two biochemical plate-based assays that will be instrumental in the high-throughput screening of novel chemical matter. There are no known direct antiviral drugs for henipaviruses, making the drug discovery even more urgent.

Another unique finding of ours is the back priming RdRp activity for both the full-length and truncated versions of NiV-M and NiV-B L-P, which was not reported yet and therefore showcases what specific RNA sequence is required for this activity. In this revision, we further conduct AF modeling and validate that the TrC19 RNA containing the predicted hairpin structure at the 3' end

could be stabilized and positioned by the RdRp-PRNTase of NiV L for the 3' extension, as shown in the Fig S8.

Additionally, these in vitro assays could elucidate how the potential interface between the CD and RdRp-PRNTase domains regulates NiV RdRp activity in real-time. In the future, we could evaluate the impact of mutating or disrupting the RdRp-PRNTase-CD interface on the RdRp and back priming activity as well, using our developed in vitro biochemical assays.

Beyond the lack of novel insight in this manuscript, I still have two majors concerns about the cryo-EM data:

- The map of FI NiV-M-L-P-CD and therefore the model derived from this map is expended compared to all the other pdb models from this manuscript but also already existing structures. It is easy to visualize by aligning all the models to one side of the protein structure (like Aas 200-250) and looking at the difference in size. This is something not so uncommon in the cryo-EM field and is the result of using the wrong pixel size for the cryo-EM processing. The authors must revise their processing and model building using the correct pixel for this data set. I also noticed that the magnification in the Supp. Table 1 for this data set is incorrect, 110,000x doesn't exist on 300 kV microscope.

Thank you for pointing out the issues. We realized that we overlooked the magnification in the Supp. Table 1 and have revised it to 105,000. We have corrected the pixel size used in image processing in the 1st revision (1.05 Å) to the correct value 0.874 Å and re-processed the dataset.

In fact, we found the structure of the FL NiV-M L-P-CD is a little bigger than the FL NiV-M L-P in this study, by aligning the residue 200-250 of the L protein as suggested. We double checked on the two different pixel sizes used for dataset collection of the two structures at two different locations: 0.809 Å for FL NiV-M L-P, and 0.874 Å for FL NiV-M L-P-CD. We clarified that these two structures were obtained from two different batches of purification with different pHs. We added in the text that this expansion is likely due to the higher pH (8.0) used for the FL NiV-M L-P-CD than the pH (7.6) used for the FL NiV-M L-P.

We updated the main text, Figure 5 and Supp Figure 2 to reflect the revision and finding.

- The data for the Bangladesh strain are not good enough for publication and that is reflected by the high clashscore of 16.8. The density is lacking continuity even at secondary structure level. However, judging from the data processing flowchart, this is something that the authors could improve by doing a focused refinement on the RdRp-PRNTase to minimize the impact of the flexible P.

Thank you for the comments. We have refined the structure of TR NiV-B L-P. The clash score is now 7.52. In the data processing, we intentionally focused on the P so that we could obtain a better map on the L-P interface. The regions that tend to be disordered and/or discontinuous are largely confined to the peripheral regions that lie outside the primary regions of interest.

We attempted a more focused refinement on the RdRp-PRNTase. However, this did not lead to a map with measurable improvement and the disordered and discontinuous regions still remained. We decided to focus on structure refinement and adhere to stricter guidelines for model building. We are confident that the model represents the map data as best possible.

Reviewer #2 (Remarks to the Author):

Thanks to the authors for looking into my comments. Overall, the model building has been greatly improved. There are still some issues with the supplementary table one and the data collection parameters.

Response:

Thank you for the reviewer for the positive comment on our manuscript revision. We sincerely thank the reviewer for the continuously constructive comments for correcting and improving our manuscript. We have addressed all the issues and listed our responses point-by-point below.

For the tr NiV-M L-P dataset, a pixel size of 0.927 seems to correspond to the calibration of a Selectris-Falcon IV detector rather than a Gatan K3 camera and energy filter like mentioned in the methods. This needs to be corrected.

Thank you for the comments. We have checked with the cryo-EM facility for this dataset collection and corrected the camera to "TFS Selectris-Falcon 4i camera".

Regarding the FI NiV-M L-P-CD, the authors greatly improved the map, especially around the connecting domain. However, I'm sorry for being insistent but the provided pixel size value of 0.874 Å/pixel is still suspicious. The average calibrated pixel size value for a K3 detector on Titan Krios across facilities around the world seems to be around 0.83 Å/pixel with of course some microscope dependent deviation. A quick online search illustrates this:

0.83 Å/pixel - The University of Tokyo, Japan
0.845 Å/pixel - University of California, Berkeley
0.825 Å/pixel - New York Structural Biology Center
0.834 Å/pixel - MRC Laboratory of Molecular Biology
0.831 Å/pixel - eBIC - Diamond Light Source
0.834 Å/pixel - Janelia Research Campus
0.835 Å/pixel - Stanford-SLAC Cryo-EM Center
0.827 Å/pixel - Max Planck Institute of Biochemistry
0.84 Å/pixel - Yale University West Campus Cryo-EM Facility
0.825 Å/pixel - NCCAT at the New York Structural Biology Center

The pixel size for your expended structure is 0.874 Å/pixel, which correspond to a deviation of about 5% relative to the mean 0.83 Å/pixel. If I now look at your structures and superimpose the tr NiV-M L-P to the FI NiV-M L-P-CD on residues 200-250 as mentioned before, I can measure the distance between the C α of two residues from each side of the active site. I randomly took Phe218 and Leu873. In the tr NiV-M L-P the distance is 56.7 Å while in the FI NiV-M L-P-CD it is 59.65 Å, which correspond to a deviation of 4.95%, comparable to the deviation between the provided pixel size and the mean pixel size for this camera/microscope/magnification configuration.

From a biological point of view, there are plenty of example of viral polymerase structures solved in various conditions and such distortion of the active site doesn't seem to have been observe. To not report something nonphysiological, I would strongly advise the authors to contact the facility where the dataset has been collected and ask them to collect a high-resolution dataset of protein standard such as ApoF. Using existing structure in the pdb, it will quickly allow you to accurately determine the exact pixel size for this magnification, which will be beneficial for you and them.

Using existing Nipah polymerase structures solved in the same conformation than the FI NiV-M L-P-CD (e.g. pdb 9BDQ or 9FUX), one can try estimating what the pixel size should be by looking at the percentage of included atoms inside the map for every possible voxel. By doing that, the estimated pixel size is 0.839 Å/pixel using 9BDQ and 0.842 Å/pixel using 9FUX, in agreement with expected pixel size on a K3-EF detector fitted on a titan Krios.

We appreciate the reviewer's critical comments on our structure determination and cryo-EM dataset collection parameters.

The reviewer's comment suggests that there may be some error in our reported pixel size (0.874 Å) for the FL NiV-M L-P-CD dataset, which was collected at the cryo-EM facility at Purdue University under the grant of Midwest Consortium (NIH U24 GM116789, acknowledged in the manuscript), we have contacted the cryo-EM facility on this matter. They used the gold diffraction data to calibrate the beam and pixel size (0.874 Å), which is a standard way. They believe it is correct but acknowledge that a 5% error is possible. We searched the literature with the grant number (GM116789) and found that the same magnification was reported:

PMID	Voltage	camera	Microscope	Magnification	Pixel size (Å)	total dose (e/Å ²)
37393310	300 kV	Gatan K3	Titan Krios	105,000	0.872	60.5
36192627	300 kV	Gatan K2 / K3	Titan Krios	105,000 / -	0.66 / 0.56	40 / 40
31227607	300 kV	K2 summit Gatan	Titan Krios	105,000	0.692	40.8
32832689	300 kV	K2 detector	Titan Krios	105,000	0.69	60

The paper (PMID: 37393310) used almost the same pixel size as ours. However, our dataset was collected at the end of the electron microscope's usage life, and the microscope has been retired. Therefore, a dataset of protein standards can not be collected.

Nevertheless, as the reviewer and the manager at the Purdue cryo-EM facility advised, we have estimated the pixel size to be 0.84 Å by inspecting the fitting between the reported Nipah polymerase structures (PDB IDs: 9BDQ and 9FUX) that have same conformations into our map and improved our structure by re-processing the dataset with a pixel size of 0.84 Å/px. The re-extracted particles greatly improved our map's resolution from 3.64 Å to 2.99 Å, suggesting the current pixel size (0.84 Å) is accurate. We also validated the applied pixel size (0.84 Å/px) by using the "fit spherical aberration" function in global CTF refinement job of cryoSPARC and obtaining a spherical aberration value at 2.7153.

We have revised our text, figures (Fig. 5 and S2), and table to reflect the correction and improvement of the structure of FL NiV-M L-P-CD.

Interestingly, once using the newly estimated pixel size. The fitting of the connecting domain predicted with AF3 is significantly better than before. The quality of the map is sufficient to built and refine most of the connecting domain as opposed to a single helix. Looking at the processing flowchart, there is still room for improvement to get better maps. First, using the proper calibrated pixel size will allow the authors to have a better per particle defocus estimation. They can also perform a reference-based motion correction and a local refinement on the core + CD after subtracting the density for the P proteins which should improve the density for the both the core and the CD.

Thank you for the comments and helpful suggestions for us to improve our dataset processing and model building on the FL NiV-M L-P-CD. As mentioned above, we have used an accurate pixel size to re-process the dataset. For this point of comment, we also tried the suggested image processing methods, including the reference-based motion correction and local refinement using a mask on the RdRp + CD in cryoSPARC. As expected, our improved FL NiV-M L-P-CD structure at 2.99 Å resolution has a lower Molprobit score (1.67 vs 1.85 in last revision) and a higher Ramachandran plot favored (96.93% vs 92.37% in last revision), demonstrating a better model building and map quality as well. With this, we could describe the interface between the RdRp core and the CD domain, which could be shown to be conserved in the Paramyxoviridae.