THE EMBO JOURNAL

Manuscript EMBO-2015-92651

Unexpected Features and Mechanism of Heterodimer Formation of a Herpesvirus Nuclear Egress Complex

Ming F Lye, Mayuri Sharma, Kamel El Omari, David J Filman, Jonathan P Schuermann, James M Hogle and Donald M Coen

Corresponding author: Donald M. Coen and James M. Hogle, Harvard Medical School

Review timeline:

Submission date:	23 July 2015
Editorial Decision:	17 August 2015
Revision received:	11 September 2015
Accepted:	30 September 2015

Editor: Karin Dumstrei

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 17 August 2015

Thank you for submitting your manuscript to The EMBO Journal. Your study has now been seen reviewed by three good experts in the field who also reviewed the back-to-back submission from Bigalke and Heldwein.

As you can see from the referee comments below, the referees appreciate the analysis and support publication here. They raise a number of constructive comments that should not involve too much additional work to sort out. Given the referee comments, I would t like to invite you to submit a suitable revised manuscript. You can use the link below to upload the revised version. Let me know if we need to discuss any of the points in more details.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: http://emboj.embopress.org/about#Transparent_Process

A similar decision was reached for the Bigalke and Heldwein submission. Could you please coordinate with Katya Heldwein and colleagues so that we get the revised versions back around the same time - Thanks!

REFEREE REPORTS

Referee #1:

Lye et al describe crystal structures and functional analyses of HCMV UL53 constructs, both free and as a non-native HCMV-MCMV chimera in complex with UL50, thereby reconstituting the nuclear egress complex (NEC). Together with the structure of MCMV M50 from a subset of these same authors, and the companion paper from Bigalke and Heldwein, we now have multiple views of herpes viral NEC subunits and complexes. The overall similarities of these structures reassures one that the structural biology being reported here is fundamentally correct, and that the functional, but non-native H/MCMVChm53 UL53 construct is not giving misleading results. This paper also uniquely presents a detailed analysis of the Bergerat fold, as well as structure-based mutational analyses of the UL50 zinc finger and UL50-UL53 heterodimer interface mutations on UL50-UL53 colocalization (Figure 4) and viral replication.

This is generally high quality work on an important viral system. It is also complementary to the companion manuscript in some ways, but doesn't go as far because the Bigalke and Heldwien manuscript describes the structures of two different NEC complexes and also shows how the HSV-1 NEC complex can form hexagonal lattices.

Comments
1) The manuscript could be shortened considerably to make it more readable. The authors and editor should consider reducing sections that describe comparisons with other homologous structures (p8-12 and p16-17), structural details that make for difficult text (17-20), rather speculative sections about possible differing functions of conserved segments (p28) and drug target sites (p29) and alignment details (e.g., I would move Figure 2 into supplemental material - surely Figures 3 and 5, which are more visual, are sufficient).

2) Figure 4: The co-localization results seem clear in the images shown, but the authors should also quantify the degree of co-localization in the different cases. This would demonstrate that the images are representative of co-localization in multiple cells, and also solidify partial phenotypes like the one seen for the S(118)I(220)-AA mutation. It would also be helpful if the figure contained better structural illustrations of the residues being mutated (which is done well for the zinc finger element, but not for the different interface mutations). There should also be some designation of which protein is being mutated in the different panels within Figure 4C itself (though this is clear in the caption).

3) The authors should clarify precisely what they mean by the statement "This is consistent with virus-specific differences in activities of NECs and their subunits in membrane remodeling and vesiculation in cells and in vitro" (p. 26-27). Do the authors mean to suggest that not all NECs will form hexagonal coats like those seen for the HSV-1 NEC? If so, I think they need to explain better their rationale for thinking this (because this seems unlikely to me). If not, then they need to clarify exactly what they are trying to say.

Referee #2:

This manuscript provides important structures of the proteins forming the nuclear egress complex (NEC) of cytomegalovirus (CMV), a beta-herpesvirus that is a very important human pathogen. The CMV NEC is composed of proteins UL50 and UL53. Following a previous report on the structure of the nucleoplasmic domain of UL50 of the murine CMV (MCMV) by NMR, the authors now report several crystal structures of fragments of UL53 from the human virus (HCMV), as well as a UL50/UL53 heterodimer (no the intact protein but a fragment of each that was shown to make a stable complex). To obtain the crystals, the authors were force to use a chimeric construct, in which parts of the MCMV sequence of UL53 was used. But the authors go on to show that this same constructs, when introduced via a bacmid to make green fluorescent virus, does gives a viable virus, indicating that the observed chimeric NEC complex is functional and that the observed interactions are relevant. Comparison of the NMR structure of isolated MCMV UL50 with that of the HCMV UL50 in complex with UL53 shows that there is a conformational change in the C-terminal helix of UL50, illustrating an interesting feature of the formation of the complex. Furthermore, the authors

identify a clear structural relation of both NEC components with proteins belonging to a large family of ATP binding proteins (the GHKL family), which has important evolutionary implications. Although it is not clear why these proteins would be related, this observation is intriguing and may in the long term become more obvious, as the authors discuss. The authors also identify regions in the C-terminal end of UL53 that could be responsible for protein-protein interactions during formation of the NEC coat, which are very interesting when comparing to the oligomeric structures of alpha-herpesvirus counterpart described in the accompanying manuscript. The manuscript is well written and the structures are of good quality, even if for one of the crystal forms the Rfree is a little high and the corresponding R factor too low in comparison. Also, the authors provide mutational studies addressing the role of the most important residues identified in the structures, validating the hypotheses generated. Together with the accompanying manuscript, these results provide meaningful and significant advances on this quite poorly understood process, which is how these large capsids exit the nucleus to reach the cytoplasm without using the nuclear pores. The interaction sites identified also constitute relevant targets for antivirals. In this context, I strongly support publication in the EMBO Journal, if the authors can address the comments listed below.

1. In the introduction, please provide the percentage amino acid sequence identity among homologs (UL53, M53 and UL31; UL50, M50 and UL34).

2. In Figures 1, 3, 5 , EV1: label N- and C-terminal ends of the constructs

3. It would be clearer if the diagram presented in Figs. 1D and S1B was made to scale, to give a better idea of the regions missing from the protein (right now, it looks like the 80 residues or so missing at either end of UL53 were much less compared to the rest of the protein).

4. The space groups of the various crystals used for data collection and structure determination should be quoted in the main text. Also, Table I indicates that a complete data set to 2.5 Å resolution was collected for the UL53(70-292) construct. Yet this structure was refined only against a different data set that goes only to 3Å resolution. The rationale for this should be better explained.

5. The description of the structures is a little confusing. Do the "two 'faces' of beta-strands" mean the two beta sheets of the taco? Or do the authors mean the outer face of each beta-sheet? This paragraph in page 9 should be re-written to make it clearer. Also at the beginning of the Discussion, where the authors write: "Additionally, the organization of strands into an A and a B face are conserved in both NEC subunits", they should at least specify that they mean beta strands (somehow they do not talk about beta sheets...). The specification of "beta strand" here is particularly important since the authors then explain that: "monomers of the complex form lateral strands of NEC" (page 24).

6. The text is also confusing when it describes the four conserved regions (CR1 through CR4) in UL53 homologous across the Herpesviridae family. It reads as if there were four clearly conserved regions within a mostly unconserved protein. But looking at Fig. 2A, the conserved regions are separated by one or two amino acids (and maybe a loop of about 10-15 aa between CR3 and CR4). In this context, it looks more like a single central core region that is conserved with a more variable N-terminal extension, instead of four clearly separate conserved regions. I think the wording should be changed to reflect this. Also, although the authors say that "each CR forms a discrete part of the three-dimensional structure", these do not seem to make clear separate domains (for example, the yellow helices in Fig. 3D seem to make a nice helix bundle with the green helices, for instance, whereas the N-terminal helix within CR1 is clearly a separate domain from the rest; the Zn binding domain would also be split between separate regions). Maybe the authors can say that, for historical reasons, it is convenient to separate the polypeptide chain in these four regions, instead of presenting them as four separate and clearly independent conserved regions.

7. The whole paragraph explaining the folding across the CRs is too detailed and hard to read (this is also true for the paragraph describing the UL50 fold further downstream). It should be rewritten in a simpler way, referring to the topology diagrams of Fig. 3. Instead of coloring by CR, the authors could ramp-color the polypeptide chain from N- to C-terminal ends, which would make it easy to follow the chain along the ribbon. Also, given that the authors identify the "Bergerat fold" within both UL53 and UL50, it would be good to have the two proteins with only the region corresponding

to this fold highlighted in a bright color, and the rest in grey, for example, as done for M50 and UL50 in Fig. 3G and H, but with the Bergerat fold in the same orientation for clarity.

8. The ITC studies of Chm53 with UL50(1-169) would be more informative if they were run in parallel with UL53(61-292), and with longer N-terminal extensions of UL53. Also, Fig. EV2 shows that additional green cells appear after day 5, indicating that the virus propagates, but there are fewer cells than in WT and than in the rescuing mutants E75R and M82AR. It would be important to have more information about the behavior of the Chm53-like mutant virus, to make sure that the observed complex is really functional.

9. Please consider using the same nomenclature for corresponding secondary structure elements in M50 and UL50, to make it easier to compare. Additional beta strands or alpha (or 3/10) helices in UL can be primed or have a "x" sub-index, to highlight that they are insertions. There is no reason to explain the Bergerat fold again for UL50, when it has been done for M50.

10. Can the actual residues of UL50 in the alpha1 helix be modeled onto the Chm53 moiety of the NEC? Are the substituted residues exposed to solvent, or are they part of the contact (this is addressed only partially in the text).

Referee #3:

Overview:

This is a very well written manuscript by Lye et al that reports the structure of UL50 and UL53, which form the nuclear egress complex (NEC) of human cytomegalovirus (HCMV) a beta-herpesvirus. The NEC is required for viral egress, and so determination and analysis of the structures of these two proteins and the complex that they form represent important milestones in understanding of herpesvirus morphogenesis. The authors present the structure of UL53 by itself, the core of which resembles certain nucleotide binding proteins and contains a Zn finger. They then used a well considered, function-guided biochemical approach to redesign UL50 to facilitate assembly of a NEC complex that was suitable for structure determination. The structure of the complex reveals the mechanism of heterodimerization. NEC contact interfaces in the crystal structure correlate very well with previous mutagenesis studies. The authors also analyse mutations designed based on their structures, focusing on Zn coordination and heterodimer contact sites. This study makes an important contribution through analysis of the NEC complex; the presented analysis of the UL53 structure and the Bergerat fold is of more specialized interest.

Major comments:

While it is worth noting the Bergerat fold and its similarity to nucleotide binding proteins, the significance of this finding to the mechanism of virus assembly and budding is not really established in the manuscript.

There is no direct experimental evidence that the metal bound to UL53 is indeed Zn.

Crystallography: Table 1 indicates relatively high bond angle rmsds (> 2 Å), and poor Ramachandran distributions for two of the structures. These indicate overweighting of the x-ray terms during refinement; I suggest a few more rounds of rebuilding and refinement with optimized weights to tighten up model geometry.

Referee #1:

Lye et al describe crystal structures and functional analyses of HCMV UL53 constructs, both free and as a non-native HCMV-MCMV chimera in complex with UL50, thereby reconstituting the nuclear egress complex (NEC). Together with the structure of MCMV M50 from a subset of these same authors, and the companion paper from Bigalke and Heldwein, we now have multiple views of herpes viral NEC subunits and complexes. The overall similarities of these structures reassures one that the structural biology being reported here is fundamentally correct, and that the functional, but non-native H/MCMVChm53 UL53 construct is not giving misleading results. This paper also uniquely presents a detailed analysis of the Bergerat fold, as well as structure-based mutational analyses of the UL50 zinc finger and UL50-UL53 heterodimer interface mutations on UL50-UL53 colocalization (Figure 4) and viral replication.

This is generally high quality work on an important viral system. It is also complementary to the companion manuscript in some ways, but doesn't go as far because the Bigalke and Heldwien manuscript describes the structures of two different NEC complexes and also shows how the HSV-1 NEC complex can form hexagonal lattices.

We thank the referee for the helpful comments and suggestions. Below are the point-by-point responses:

Comments
1) The manuscript could be shortened considerably to make it more readable. The authors and editor should consider reducing sections that describe comparisons with other homologous structures (p8-12 and p16-17), structural details that make for difficult text (17-20), rather speculative sections about possible differing functions of conserved segments (p28) and drug target sites (p29) and alignment details (e.g., I would move Figure 2 into supplemental material - surely Figures 3 and 5, which are more visual, are sufficient).

The text on p.8-12 and 16-20 have been revised to be more concise. Figure 2 is now Appendix S2 in the Appendix. In the Discussion, we have deleted the sentence regarding conserved region (CR) 3 (p.28, second paragraph), but retained the section on CR1 (p.27, last paragraph) as our results directly bear on its functions, and added a sentence to the paragraph on CR2-4 indicating how the structure may be relevant to interactions of CR4 with other proteins (p.28, second paragraph). We have retained the paragraph on how the mechanism of heterodimerization is germane to drug design, as we feel this is an appropriate level of speculation for a Discussion (p.28 – 29, under the heading "Implications of the mechanism of heterodimer formation for drug discovery").

2) Figure 4: The co-localization results seem clear in the images shown, but the authors should also quantify the degree of co-localization in the different cases. This would demonstrate that the images are representative of co-localization in multiple cells, and also solidify partial phenotypes like the one seen for the S(118)I(220)-AA mutation. It would also be helpful if the figure contained better structural illustrations of the residues being mutated (which is done well for the zinc finger element, but not for the different interface mutations). There should also be some designation of which protein is being mutated in the different panels within Figure 4C itself (though this is clear in the caption).

We have calculated Manders coefficients from multiple cells for each sample to quantify the degree of co-localization between WT and mutant versions of FLAG-53 and HA-50 in cells shown in Figure 3C (Note: As Figure 2 was moved to the Appendix, all the proceeding figures were thus moved up and renumbered. The Figure 4 that the reviewer was referring to is now Figure 3). These values are shown in Figure EV2.

The residues at the NEC interaction interface that were mutated in the co-localization studies are shown in Figure 5 as they were discussed later in the main text. We now cite these parts of Figure 5 when discussing the co-localization results – Figure 5E for the UL50 2A and 4A mutants (p.21, second paragraph); Figure 5D for the E75A, M82A, and L79A mutants (p.21, third paragraph); Figure 5F for the $S_{118}I_{120}$-AA mutant (p.20, first paragraph under "Effects of substitutions affecting the interaction interface").

In Figure 3C, the protein being mutated in each panel (from top to bottom) is indicated to the left and is now color-coded as designated in the top panel (UL53 is shown in red while UL50 is in green).

3) The authors should clarify precisely what they mean by the statement "This is consistent with virus-specific differences in activities of NECs and their subunits in membrane remodeling and vesiculation in cells and in vitro" (p. 26-27). Do the authors mean to suggest that not all NECs will form hexagonal coats like those seen for the HSV-1 NEC? If so, I think they need to explain better their rationale for thinking this (because this seems unlikely to me). If not, then they need to clarify exactly what they are trying to say.

We agree with the referee that we cannot exclude either possibility – that the NECs across different sub-families adopt different mechanisms of oligomerization at the membrane, or, more likely, that they all form hexagonal coats. However, if they all form hexagonal coats, our analysis raises the possibility that the contacts involved may differ among viruses, consistent with published differences in membrane remodeling and vesiculation. We have thus rewritten this part of the text to reflect this in the second paragraph on p.26.

We again thank the referee for the close reading of the manuscript and thoughtful comments.

Referee #2:

This manuscript provides important structures of the proteins forming the nuclear egress complex (NEC) of cytomegalovirus (CMV), a beta-herpesvirus that is a very important human pathogen. The CMV NEC is composed of proteins UL50 and UL53. Following a previous report on the structure of the nucleoplasmic domain of UL50 of the murine CMV (MCMV) by NMR, the authors now report several crystal structures of fragments of UL53 from the human virus (HCMV), as well as a UL50/UL53 heterodimer (no the intact protein but a fragment of each that was shown to make a stable complex). To obtain the crystals, the authors were force to use a chimeric construct, in which parts of the MCMV sequence of UL53 was used. But the authors go on to show that this same constructs, when introduced via a bacmid to make green fluorescent virus, does gives a viable virus, indicating that the observed chimeric NEC complex is functional and that the observed interactions are relevant. Comparison of the NMR structure of isolated MCMV UL50 with that of the HCMV UL50 in complex with UL53 shows that there is a conformational change in the C-terminal helix of UL50, illustrating an interesting feature of the formation of the complex. Furthermore, the authors identify a clear structural relation of both NEC components with proteins belonging to a large family of ATP binding proteins (the GHKL family), which has important evolutionary implications. Although it is not clear why these proteins would be related, this observation is intriguing and may in the long term become more obvious, as the authors discuss. The authors also identify regions in the C-terminal end of UL53 that could be responsible for protein-protein interactions during formation of the NEC coat, which are very interesting when comparing to the oligomeric structures

of alpha-herpesvirus counterpart described in the accompanying manuscript. The manuscript is well written and the structures are of good quality, even if for one of the crystal forms the Rfree is a little high and the corresponding R factor too low in comparison. Also, the authors provide mutational studies addressing the role of the most important residues identified in the structures, validating the hypotheses generated. Together with the accompanying manuscript, these results provide meaningful and significant advances on this quite poorly understood process, which is how these large capsids exit the nucleus to reach the cytoplasm without using the nuclear pores. The interaction sites identified also constitute relevant targets for antivirals. In this context, I strongly support publication in the EMBO Journal, if the authors can address the comments listed below.

We thank the referee for the helpful comments and suggestions. Below are the point-by-point responses:

1. In the introduction, please provide the percentage amino acid sequence identity among homologs (UL53, M53 and UL31; UL50, M50 and UL34).

We thank the referee for this suggestion, but feel that this information would be better provided in the legend for the UL50 and UL53 structure-based multiple sequence alignment in Appendix S2A and S2B. (Note: Appendix S2A and S2B were previously Figure 2A and 2B. We moved Figure 2 into the Appendix following the advice of another reviewer. All the figures after Figure 2 previously were thus moved up and renumbered. (E.g., what was Figure 3 before is now Figure 2).

2. In Figures 1, 3, 5, EV1: label N- and C-terminal ends of the constructs

The N- and C-terminal ends in the above figures (now 1, 2, 4 and EV1 (as well as Figure 5)) have now been labeled.

3. It would be clearer if the diagram presented in Figs. 1D and S1B was made to scale, to give a better idea of the regions missing from the protein (right now, it looks like the 80 residues or so missing at either end of UL53 were much less compared to the rest of the protein).

The diagrams in Figures 1D and S1B have been corrected and are now shown to scale. Thank you to the referee for bringing this to our attention!

4. The space groups of the various crystals used for data collection and structure determination should be quoted in the main text. Also, Table I indicates that a complete data set to 2.5 Å resolution was collected for the UL53(70-292) construct. Yet this structure was refined only against a different data set that goes only to 3Å resolution. The rationale for this should be better explained.

The space groups for the crystals are listed in the main text of the Results section (p.7, in the first paragraph under the heading "Structure determination of UL53" for UL53$^{72-292}$ and UL53$^{84-292}$, and on p.15, first paragraph under the heading "The NEC crystal structure" for the NEC in the revised text).

We have now incorporated the rationale for structure solution of UL53$^{72-292}$ at 3Å on p.7, first paragraph under "Structure determination of UL53". Briefly, the electron density was more readily interpretable using data from a single crystal at 3Å resolution.

5. The description of the structures is a little confusing. Do the "two 'faces' of beta-strands" mean the two beta sheets of the taco? Or do the authors mean the outer face of each beta-sheet? This paragraph in page 9 should be re-written to make it clearer. Also at the beginning of the Discussion, where the authors write: "Additionally, the organization of strands into an A and a B face are

conserved in both NEC subunits", they should at least specify that they mean beta strands (somehow they do not talk about beta sheets...). The specification of "beta strand" here is particularly important since the authors then explain that: "monomers of the complex form lateral strands of NEC" (page 24).

Yes, the "two 'faces' of b-strands" in the above sentence do mean the two beta sheets of the taco. This sentence on p.9 in the first paragraph under the heading "UL53 has a unique topology that incorporates the Bergerat fold" has been rewritten as "two sheets of b-strands, each forming a 'face'". The designation of an A and B face was described previously for M50 (Leigh et al, 2015) and used here to describe a similar arrangement of the b-strands in UL53. The sentence described above by the referee has now been modified (shown underlined) to "Additionally, the organization of two sheets of b-strands into an A and a B face are conserved in both NEC subunits." (p.22, first paragraph of the Discussion section). This is for consistency in the usage of b-strand to describe a single b-strand, and sheets to describe the multi-b-stranded A and B faces.

6. *The text is also confusing when it describes the four conserved regions (CR1 through CR4) in UL53 homologous across the Herpesviridae family. It reads as if there were four clearly conserved regions within a mostly unconserved protein. But looking at Fig. 2A, the conserved regions are separated by one or two amino acids (and maybe a loop of about 10-15 aa between CR3 and CR4). In this context, it looks more like a single central core region that is conserved with a more variable N-terminal extension, instead of four clearly separate conserved regions. I think the wording should be changed to reflect this. Also, although the authors say that "each CR forms a discrete part of the three-dimensional structure", these do not seem to make clear separate domains (for example, the yellow helices in Fig. 3D seem to make a nice helix bundle with the green helices, for instance, whereas the N-terminal helix within CR1 is clearly a separate domain from the rest; the Zn binding domain would also be split between separate regions). Maybe the authors can say that, for historical reasons, it is convenient to separate the polypeptide chain in these four regions, instead of presenting them as four separate and clearly independent conserved regions.*

We agree with the view of the referee that the sentence "each CR forms a discrete part of the three-dimensional structure" may be misconstrued as each CR forming a domain. We have thus removed it from the main text as well as rewritten this second paragraph under the heading "UL53 has a unique overall topology that incorporates the Bergerat fold" on p.9 describing the CRs to better reflect the results of our structural and proteolysis studies (Figure S1A), which support a stable, proteolysis-resistant core domain comprising the main body (residues 84 to 292) of UL53 separate from the N-terminal heterodimerization domain. The color-coded mapping of the CRs (defined previously by Lotzerich et al, 2006) onto the UL53 structure are not meant to indicate domains, but to facilitate a structural interpretation of previous and future mutational and functional studies that base their construct design on these CR designations.

7. The whole paragraph explaining the folding across the CRs is too detailed and hard to read (this is also true for the paragraph describing the UL50 fold further downstream). It should be rewritten in a simpler way, referring to the topology diagrams of Fig. 3. Instead of coloring by CR, the authors could ramp-color the polypeptide chain from N- to C-terminal ends, which would make it easy to follow the chain along the ribbon. Also, given that the authors identify the "Bergerat fold" within both UL53 and UL50, it would be good to have the two proteins with only the region corresponding to this fold highlighted in a bright color, and the rest in grey, for example, as done for M50 and UL50 in Fig. 3G and H, but with the Bergerat fold in the same orientation for clarity.

This description of the CRs on p.9 has been rewritten and now refers mostly to the topology diagram (Figure 2A) as suggested by the referee for consistency, as well as the UL53 structure on which the CRs are mapped (Figure 1E). Additionally, the UL53 structure in Figure 1E is now labeled with the secondary structure elements to aid in tracing the continuity of the structure. We have maintained the block coloring (by CRs) of UL53 as it depicts the UL53

structure in context of the regions examined in previous biochemical and functional studies as mentioned in our response to comment number 6.

The UL53, M50, and UL50 structures in Figure 2D (left), 2G, and 2H respectively are all shown in the same orientation with respect to their Bergerat fold (circled). For ease of comparison, we have now added a figure of UL53 (Figure 2D (bottom)) in a similar color scheme as that for M50 and UL50 (with the Bergerat fold highlighted and the rest of the structure shown in gray) as suggested by the referee.

8. The ITC studies of Chm53 with UL50(1-169) would be more informative if they were run in parallel with UL53(61-292), and with longer N-terminal extensions of UL53. Also, Fig. EV2 shows that additional green cells appear after day 5, indicating that the virus propagates, but there are fewer cells than in WT and than in the rescuing mutants E75R and M82AR. It would be important to have more information about the behavior of the Chm53-like mutant virus, to make sure that the observed complex is really functional.

We agree with the referee and have included ITC data for UL50$^{1-169}$ with Chm53, UL53$^{50-292}$ (reported previously by Sam et al, 2009 and repeated in this study), and UL53$^{1-292}$ to evaluate the effects of residue substitution (to generate Chm53), as well as N-terminal truncation mutations on UL53 on NEC formation. The binding affinities of Chm53 and UL53$^{50-292}$ for UL50$^{1-169}$ are comparable (1.0 ± 0.08 mM and 0.42 ± 0.05 mM respectively). Interestingly, the presence of the full length N-terminal region (UL53$^{1-292}$ construct) resulted in a decreased binding affinity (K_d of 3.8 ± 0.8 mM) suggesting a regulatory effect of the N-terminal portion on UL53 activity. This information has now been included in the first paragraph under "Construction and expression of the HCMV NEC for crystallization" on p.15.

As UL53$^{61-292}$ has the same starting boundaries as Chm53, it would be ideal to compare the binding affinities between these two constructs with UL50. However, we were unable to perform the ITC study as UL53$^{61-292}$ precipitated after buffer-exchange into the TCEP-containing ITC buffer. We mention this observation in the first paragraph under "Construction and expression of the HCMV NEC for crystallization" on p.15.

For the Chm53 virus, we have shown that it spread through the culture and, when harvested, was infectious based on plaque assays. We have also confirmed that it has the correct sequence. This information has now been incorporated as the second paragraph under "Construction and expression of the HCMV NEC for crystallization" on p.15.

9. Please consider using the same nomenclature for corresponding secondary structure elements in M50 and UL50, to make it easier to compare. Additional beta strands or alpha (or 3/10) helices in UL can be primed or have a "x" sub-index, to highlight that they are insertions. There is no reason to explain the Bergerat fold again for UL50, when it has been done for M50.

To facilitate the structural comparison between M50 and UL50 in this study, a-helix or b-strand insertions in the structures are now indicated by a prime (') as shown in Figures 2, 4, 5, and Appendix Figure 2B, their corresponding figure legends, and in the main text. The description of the Bergerat fold for UL50 has been removed to avoid redundancy.

10. Can the actual residues of UL50 in the alpha1 helix be modeled onto the Chm53 moiety of the NEC? Are the substituted residues exposed to solvent, or are they part of the contact (this is addressed only partially in the text).

The reviewer raises an excellent point. We have now included a figure (Appendix Figure S3) with the five UL53 residues (H62, D63, I67, R69, E70) that were substituted overlaid with their respective M50 homologs (S62, E63, V67, Q69, R70) used in the Chm53 construct. Of these five residues, only V67 is involved in heterodimer contacts (forming a hydrophobic interaction

surface with other residues on the UL53 N-terminal helices that face a complementary hydrophobic surface on the UL50 a4 helix (moveable jaw) as discussed in the second paragraph on p.19). This information has been inserted into the legend for Figure S3.

We again thank the referee for the close reading of the manuscript and thoughtful comments.

Referee #3:

Overview:

This is a very well written manuscript by Lye et al that reports the structure of UL50 and UL53, which form the nuclear egress complex (NEC) of human cytomegalovirus (HCMV) a beta-herpesvirus. The NEC is required for viral egress, and so determination and analysis of the structures of these two proteins and the complex that they form represent important milestones in understanding of herpesvirus morphogenesis. The authors present the structure of UL53 by itself, the core of which resembles certain nucleotide binding proteins and contains a Zn finger. They then used a well considered, function-guided biochemical approach to redesign UL50 to facilitate assembly of a NEC complex that was suitable for structure determination. The structure of the complex reveals the mechanism of heterodimerization. NEC contact interfaces in the crystal structure correlate very well with previous mutagenesis studies. The authors also analyse mutations designed based on their structures, focusing on Zn coordination and heterodimer contact sites. This study makes an important contribution through analysis of the NEC complex; the presented analysis of the UL53 structure and the Bergerat fold is of more specialized interest.

We thank the referee for the helpful comments and suggestions. Below are the point-by-point responses:

Major comments:

1. While it is worth noting the Bergerat fold and its similarity to nucleotide binding proteins, the significance of this finding to the mechanism of virus assembly and budding is not really established in the manuscript.

We have shown that both helices (a1 and a2) of the Bergerat fold in UL50 are directly involved in forming the NEC (discussed under "The NEC interaction interface," in the first paragraph on p.18, and shown in Figures 2C, 2H, and 5A), which is critical for viral viability and nuclear egress.

2. There is no direct experimental evidence that the metal bound to UL53 is indeed Zn.

We thank the referee for suggesting such an experiment, and have now included an elemental analysis of a frozen sample of UL53$^{72-292}$ using energy-dispersive X-ray spectroscopy (EDS) in which the only significant peaks present from the scan corresponds to that of zinc (Appendix Figure S5). We have added Jonathan Schuermann, who performed this experiment, as an author.

3. Crystallography: Table 1 indicates relatively high bond angle rmsds (> 2 Å), and poor Ramachandran distributions for two of the structures. These indicate overweighting of the x-ray terms during refinement; I suggest a few more rounds of rebuilding and refinement with optimized weights to tighten up model geometry.

We have improved the refinement statistics of the NEC and UL53$^{71-292}$ crystal structures, as reported in Table 1. The changes in the final structures compared with the previous models are negligible and do not affect the structural interpretations.

Aside from the changes detailed above, we have also included new data showing that rescued derivatives of the zinc finger mutants are viable (p.14 and Figure EV2), which indicates that the substitutions, rather than some other mutations in those viruses, are responsible for ablating virus replication. We have also modified the Abstract to mention this.

We again thank the referee for the close reading of the manuscript and thoughtful comments.

2nd Editorial Decision 30 September 2015

Thank you for submitting your revised manuscript to The EMBO Journal. Your study has now been seen by referee #3. As you can see the referee appreciates the introduced changes and supports publication here. I am therefore very pleased to accept the manuscript.

REFEREE REPORT

Referee #3:

The authors have addressed my previous comments. Together with the accompanying paper, this is an important advance.

