THE EMBO JOURNAL

Manuscript EMBO-2009-72476

How baculovirus polyhedra fit square pegs into round holes to robustly package viruses

Xiaoyun Ji, Geoff Sutton, Gwyndaf Evans, Danny Axford, Robin Owen, David I. Stuart

Corresponding author: David Stuart, University of Oxford

Review timeline:

Submission date:	07 September 2009
Editorial Decision:	05 October 2009
Revision received:	23 October 2009
Accepted:	30 October 2009

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 05 October 2009

Thank you very much for submitting your research manuscript for consideration to The EMBO Journal editorial office. Three referees have assessed the quality and insight that your work obviously provided as all of them are rather positive in their conclusion. Despite a few necessary modifications that relate to toning down the phylogenetic relationship (ref#1), clarification to some figures (ref#2) and trying to focus the text, particularly discuss implications for crystal growth and the importance of the identical crystal lattice (ref#3) I am happy to offer submission of a revised version of your work in the near future. I also have to remind you that it is EMBO_J policy to allow a single round of revisions only, which means that the final decision on publication depends on the content of the final version of your manuscript.

Thank you for the opportunity to consider your work for publication. I look forward to reading the revised manuscript.

Yours sincerely,

Editor
The EMBO Journal

REFEREE REPORTS

Referee #1 (Remarks to the Author):

This manuscript describes the 3D structure of the baculovirus polyhedrin, a protein tailored to make

a crystalline array within the nucleus of the infected cells. These crystals are stabilized by a network of disulfide bonds. Infectious baculovirus particles are enclosed and protected within these crystals until reaching the midgut of a new insect larva. The crystals, or "polyhedra" are extremely resistant, and can only be dissolved at a pH above 10, which is found the insect midgut. The polyhedrin molecule is folded similarly to the capsid protein of picorna-like insect viruses, and the authors propose that they are likely to have diverged from some common ancestor gene. In contrast, the baculovirus polyhedrin shows only overall similarities to its counterpart of the cytoplasmic polyhedrosis virus (CPV), which packs in an identical 3D lattice. The reported similarities in shape are not maintained to the level of the topological arrangement of the strands, and the question of divergent of convergent evolution of the two proteins remains open.
This is a beautiful structure, with the crystallographic statistics attesting of its high quality. The comparison to other proteins of known structure, the picornavirus coat proteins and CPV polyhedrin, provide new and important information about the possible evolution pathways of these viral proteins. In my view, this manuscript is totally appropriate for publication in the EMBO Journal, provided that the authors address the following issues.

1. In page 8 the authors explain that "Since residues 32-48 could not be located in the electron density map there is an ambiguity in connecting the N-terminal residues 1-31 to the rest of the subunit. Because the unresolved residues lie close to the crystallographic origin, where numerous symmetry axes converge, 12 different links are feasible". I think that it is important to expand this a little bit more, since the authors appear to have picked, arbitrarily, only one of the many alternative connections. Since it is the specific lattice that has biological meaning, different connections at this point may lead to different patterns of lattice packing, via possible strand swapping, etc. It may be that the packing analysis provided in further sections is not complete if the particular connectivity that they have chosen is not the right one. For instance, residues 1-31 could be forming an important interaction further away, to a subunit not being contacted by the one to which it belongs in the presently assumed connection. This should be considered in their discussion further down in the paper

2. At the end of page 10, the authors explain that 86 Calpha carbons of CPV and baculovirus polyhedrin superpose with rms deviation of 3.5Å. What score is obtained when comparing with the DALI server? I ask this because the authors provide the DALI scores for the comparison with the picornavirus core proteins, but not to this one. Also, in this page (line 8), I think that "VP2 of cricket paralysis virus" is a typo, should be VP3

3. From page 11 on, the authors use the term "monomer" a number of times, which implies that there is a monomeric species of polyhedrin, alhtough the authors actually show that it is trimeric. In this context, it is probably less confusing to speak of the protomer, or trimer subunit, instead of a monomer in the trimer (since, if the protein is a trimer, it is not a monomer, and vice versa)

4. The phylogenetic tree provided in Figure 3 is very nice and illustrative, but may be a bit misleading because it is not complete. It would be important to see where other coat proteins folded as jelly-roll fall in it (from T=3 viruses, for instance) and also proteins from cellular origin, like the TNF. Since it is not clear if the cellular jelly-rolls and those in viral coats have emerged independently or have a common origin, this could perhaps help better delimiting the "twilight zone" and allow for a more in-depth discussion of the evolution of these viral proteins.

5. The discussion is pretty much centered on the evolution of these proteins, which is interesting but also a little risky. It is clear that the authors detect homology between the baculovirus polyhedrin and the picornavirus coat proteins. But can they say that that an insect picornavirus coat protein is a plausible origin for the baculovirus polyhedrin, without considering the opposite possibility? Or that the gene for these proteins could have been picked from a common cellular pool, at different times during evolution? Furthermore, since evolution is not necessarily parcimonious, more complex alternative can be equally plausible. At the end of page 18, the authors also discuss about the possible convergent or divergent evolution of the CPV and the baculovirus coat protein, and they lean heavily toward the second option. But what would be the driving force that would make a protein that is already capable to make such a lattice (I23 and 109Å parameter) to diverge so much, changing its topological arrangement of beta-strands, etc., when it already has the required properties? Is it not easier to assume that it is the lattice has some special features - even considering

the case that entomopox viruses have a much larger protein that makes a different lattice - and that two proteins from different origins may have converged to making such a particular lattice? Again, by adding to Figure 3 the branches corresponding to proteins of cellular origin that are similar to the jelly rolls but for which it is not certain that they may be homologs could give some idea to orient this discussion, which appears a little bit biased to me in the current version.

Referee #2 (Remarks to the Author):

This paper describes the crystal structures of native polyhedra from Autographa californica multiple nucleopolyhedrovirus (3.0 Å resolution) as well as polyhedra formed by a mutant of the polyhedrin protein (1.84 Å resolution). The crystallography appears to be well done, but is it really true that a change in cell dimensions from 102.6 to 101.6 will change the number of unique reflections to 3.0 Å by 10 % ? The major result is that the polyhedrin protein is similar to viral jellyroll coat proteins but with extensions that allows the protein to form a very densely packed cubic crystal. The apparently very similar polyhedra from cytoplasmic polyhedrosis virus are formed by a protein with a somewhat different topology. Still the viruses may use a similar mechanism for release of the virus particles based on deprotonation of tyrosine residues at high pH. The polyhedrin protein has flexible regions and these are suggested to be important for the interactions with the non-symmetrical virus particles.

The paper is clear and easy to follow. The discussion about the evolution of the polyhedrin protein from an insect picornalike virus is based on a not very striking structural similarity and, as is often the case with viruses, no traces of sequence similarity. The suggested relation is a clear possibility, but it is difficult to see how this can be proved or used for a better understanding of the mechanisms used by the polyhedrin.

Some further remarks:

1. The title is catchy but in my opinion does not describe the main contents of the paper very well.

2. Page 4, line 15. Dodecameric?

3. It is a little confusing that the authors point out that the polyhedrin protein is most similar to VP3 proteins, but still use VP2 of cricket paralysis for the comparison. (It is surprising that does DALI pick up a similarity with VP2 of Cricket paralysis virus, rather than VP3.) Does Figure 3 indicate that polyhedrin is part of a VP2 branch but more similar to picorna VP3 because these proteins have changed less than VP2s?

4. If one would like to see to what extent the elements of the structures are in the same position both in space and in the sequence, the stereo drawings in Fig. 2 are difficult to follow. Some numbers or letters would maybe help.

5. Page 11, line 7. The reference here to Fig 1F seems to be wrong.

6. Page 16, line 6. What is "specialist high pH"?

Referee #3 (Remarks to the Author):

This work is technically impressive, since the in vivo grown crystals of baculovirus (AcMNPV) polyhedrin are even smaller than those used to determine the CPV polyhedrin structure. They authors solved the structure using SeMet labelling in insect culture, SAD phasing on a mutant followed by molecular replacement to the native with smaller crystals, and many data sets were merged in each case. The results also reflect the excellence of design and experiment at the beam line. The illustrations, especially Fig. 1, are instructive and beautiful.
The findings are novel: AcMNPV is structurally related to the picornavirus VP3 proteins and the cricket paralysis virus coat protein, but dissimilar to the CPV polyhedrin, and yet AcMNPV is the closest structural relative of CPV polyhedrin. The levels of packing interactions of AcMNPV are explained very clearly. The overabundance of Tyr near subunit interface is suggested as a cause of

requirement of high pH for dissociation.
The text is however not well focused to make the author-intended conclusions stand out. From the abstract it is not entirely clear what has been done. There are two points that have been touched upon but can be discussed more clearly. One is why the polyhedrin crystal can grow around the occlusion, which has biotechnological applications as well as interest for crystal growth. The second is that AcMNPV and CPV polyhedrin have identical space group as well as unit cell dimensions. Some explanation of the latter will satisfy the curiosity raised by the abstract. I recommend publication after revision of the text.
Specific points:
p.12: "to form cavities" Please clarify the relation to the subunit structure.
p.15: "though NLS interactions (Fig. 1F)." Information is absent in Fig. 1F.

1st Revision - Authors' Response 23 October 2009

We are grateful for the thoughtful remarks from all three referees. We will deal with individual comments below but would like to point out that the most significant changes we have made are in the abstract, which has been modified to more clearly reflect the findings of this study, and the discussion which has been restructured and somewhat expanded in places to enable the principle threads to be more clearly presented. We believe that the changes we have made address all the substantive comments.

Specific answers to referee's remarks:

Referee #1

Q1. " *In page 8 the authors explain that "Since residues 32-48 could not be located in the electron density map there is an ambiguity in connecting the N-terminal residues 1-31 to the rest of the subunit. Because the unresolved residues lie close to the crystallographic origin, where numerous symmetry axes converge, 12 different links are feasible". I think that it is important to expand this a little bit more, since the authors appear to have picked, arbitrarily, only one of the many alternative connections. Since it is the specific lattice that has biological meaning, different connections at this point may lead to different patterns of lattice packing, via possible strand swapping, etc. It may be that the packing analysis provided in further sections is not complete if the particular connectivity that they have chosen is not the right one. For instance, residues 1-31 could be forming an important interaction further away, to a subunit not being contacted by the one to which it belongs in the presently assumed connection. This should be considered in their discussion further down in the paper"*

We agree with the referee and have taken the opportunity to expand on this subject at the appropriate place in the manuscript (p11). In addition we have added Supplementary Figure 4 for further clarification.

Q2. "*At the end of page 10, the authors explain that 86 Calpha carbons of CPV and baculovirus polyhedrin superpose with rms deviation of 3.5Å. What score is obtained when comparing with the DALI server? I ask this because the authors provide the DALI scores for the comparison with the picornavirus core proteins, but not to this one. Also, in this page (line 8), I think that "VP2 of cricket paralysis virus" is a typo, should be VP3*."

DALI: We have amended Supplementary Table I to include the CPV1 polyhedrin and altered the text to reference this (p10).
VP2/3: infact VP2 and VP3 are both almost equally similar to the polyhedrin and we have amended the text to remove the focus on VP3.

Q3. "*From page 11 on, the authors use the term "monomer" a number of times, which implies that there is a monomeric species of polyhedrin, alhtough the authors actually show that it is trimeric. In*

this context, it is probably less confusing to speak of the protomer, or trimer subunit, instead of a monomer in the trimer (since, if the protein is a trimer, it is not a monomer, and vice versa)"

We agree and have replaced the term "monomer" with "subunit".

Q4.*" The phylogenetic tree provided in Figure 3 is very nice and illustrative, but may be a bit misleading because it is not complete. It would be important to see where other coat proteins folded as jelly-roll fall in it (from T=3 viruses, for instance) and also proteins from cellular origin, like the TNF. Since it is not clear if the cellular jelly-rolls and those in viral coats have emerged independently or have a common origin, this could perhaps help better delimiting the "twilight zone" and allow for a more in-depth discussion of the evolution of these viral proteins."*

This is a very sensible suggestion. Figure 3 has been modified, we have now included the coat proteins from a T=3 virus as well as TNF. The results are as expected and the main text has been modified to reference this – finally the results also inform the revised discussion.

Q5. "*The discussion is pretty much centered on the evolution of these proteins, which is interesting but also a little risky. It is clear that the authors detect homology between the baculovirus polyhedrin and the picornavirus coat proteins. But can they say that that an insect picornavirus coat protein is a plausible origin for the baculovirus polyhedrin, without considering the opposite possibility? Or that the gene for these proteins could have been picked from a common cellular pool, at different times during evolution? Furthermore, since evolution is not necessarily parcimonious, more complex alternative can be equally plausible. At the end of page 18, the authors also discuss about the possible convergent or divergent evolution of the CPV and the baculovirus coat protein, and they lean heavily toward the second option. But what would be the driving force that would make a protein that is already capable to make such a lattice (I23 and 109Å parameter) to diverge so much, changing its topological arrangement of beta-strands, etc., when it already has the required properties? Is it not easier to assume that it is the lattice has some special features - even considering the case that entomopox viruses have a much larger protein that makes a different lattice - and that two proteins from different origins may have converged to making such a particular lattice? Again, by adding to Figure 3 the branches corresponding to proteins of cellular origin that are similar to the jelly rolls but for which it is not certain that they may be homologs could give some idea to orient this discussion, which appears a little bit biased to me in the current version."*

We thank the referee for these suggestions. Figure 3 has been modified as suggested and the discussion overhauled to try to lay out the arguments more clearly (eg to discuss what the special lattice features might be). We hope that the discussion is now not unacceptably biased.

Referee #2

The referee notes that the change in the number of unique reflections exceeds what might be expected for a 1% change in cell dimensions. We were also surprised by this (3% would seem a more plausible change than 10%) nevertheless it would appear that this must be a sampling effect, we attach at the end of this document the output from Scalepack, to reassure the referee.

Q1. *"The title is catchy but in my opinion does not describe the main contents of the paper very well."*

We agree that much of the discussion in the paper concerns other issues, however the net result is to produce a protein lattice that has the remarkable ability to package massive pleiomorphic particles and we would not wish to lose sight of that in the title.

Q2. "*Page 4, line 15. Dodecameric?"*

Quite right, this has now been corrected.

Q3. "*It is a little confusing that the authors point out that the polyhedrin protein is most similar to VP3 proteins, but still use VP2 of cricket paralysis for the comparison. (It is surprising that does DALI pick up a similarity with VP2 of Cricket paralysis virus, rather than VP3.) Does Figure 3 indicate that polyhedrin is part of a VP2 branch but more similar to picorna VP3 because these proteins have changed less than VP2s?"*

We have mostly addressed these points in our response to Referee 1, points 2 and 4. In our experience SHP can provide a more reliable analysis than DALI when there is significant structural variation. It is probably unwise to place too much weight on fine details of branch points, and we have slightly toned down the text.

Q4. *"If one would like to see to what extent the elements of the structures are in the same position both in space and in the sequence, the stereo drawings in Fig. 2 are difficult to follow. Some numbers or letters would maybe help."*

We recognize that the stereo drawings in Figure 2 were difficult to see. We have now amended them so that the areas of interest now standout from the rest of the molecules – the other portions of the molecule are shown as ribbons (not cartoons) and semi-transparent. This makes the figure far easier to 'read'. Due to the complexity of the figure we find that labels or numbers tend to confuse the eye.

Q5."*Page 11, line 7. The reference here to Fig 1F seems to be wrong."*

The reference to Figure 1F is correct. It is referring to the annotation lines above the sequence which show the various interactions taking place in the structure.

Q6. *"Page 16, line 6. What is "specialist high pH"?"*

Apologies, this was obscure. We have amended the text to make it clearer – it is not just pH, there is a dependency on buffer type – ie specialised to carbonate buffered insect gut.

Referee #3

General points: *"The text is however not well focused to make the author-intended conclusions stand out. From the abstract it is not entirely clear what has been done. There are two points that have been touched upon but can be discussed more clearly. One is why the polyhedrin crystal can grow around the occlusion, which has biotechnological applications as well as interest for crystal growth. The second is that AcMNPV and CPV polyhedrin have identical space group as well as unit cell dimensions. Some explanation of the latter will satisfy the curiosity raised by the abstract. I recommend publication after revision of the text. "*

The abstract has been re-structured along the lines suggested.
The discussion has also been substantially modified to answer these helpful suggestions.

p.12: "*"to form cavities" Please clarify the relation to the subunit structure."*

We have added two supplementary figures (numbers 4 and 6) to clarify.

p.15: ""*though NLS interactions (Fig. 1F)." Information is absent in Fig. 1F."*

Figure 1F has been amended to show the NLS.

Scalepack output

G25D Se-Met

Shell Lower limit	Upper limit	I/Sigma in resolution shells: No. of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	6.46	5	8	14	19	30	60	111	289	400
6.46	5.13	6	15	25	39	51	95	158	218	376
5.13	4.48	2	10	16	20	31	61	105	267	372
4.48	4.07	6	10	18	29	43	87	137	234	371
4.07	3.78	10	20	30	42	63	106	181	192	373
3.78	3.56	10	26	39	60	82	133	234	131	365
3.56	3.38	7	12	25	41	73	156	234	136	370
3.38	3.23	8	23	54	74	112	201	291	68	359
3.23	3.11	10	32	65	96	145	228	314	55	369
3.11	3.00	20	42	81	117	168	249	335	31	366
All hkl		84	198	367	537	798	1376	2100	1621	3721

Shell Lower limit	Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	6.46	1.2	2.0	3.5	4.7	7.5	15.0	27.7	72.1	99.8
6.46	5.13	1.6	4.0	6.6	10.4	13.6	25.3	42.0	58.0	100.0
5.13	4.48	0.5	2.7	4.3	5.4	8.3	16.4	28.2	71.8	100.0
4.48	4.07	1.6	2.7	4.9	7.8	11.6	23.5	36.9	63.1	100.0
4.07	3.78	2.7	5.4	8.0	11.3	16.9	28.4	48.5	51.5	100.0
3.78	3.56	2.7	7.1	10.7	16.4	22.5	36.4	64.1	35.9	100.0
3.56	3.38	1.9	3.2	6.8	11.1	19.7	42.2	63.2	36.8	100.0
3.38	3.23	2.2	6.4	15.0	20.6	31.2	56.0	81.1	18.9	100.0
3.23	3.11	2.7	8.7	17.6	26.0	39.3	61.8	85.1	14.9	100.0
3.11	3.00	5.5	11.5	22.1	32.0	45.9	68.0	91.5	8.5	100.0
All hkl		2.3	5.3	9.9	14.4	21.4	37.0	56.4	43.6	100.0

Wild-type Native

Shell Lower limit	Upper limit	I/Sigma in resolution shells: No. of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	6.46	18	41	66	85	118	192	296	53	349
6.46	5.13	25	57	102	133	183	282	336	6	342
5.13	4.48	23	44	67	91	133	219	306	20	326
4.48	4.07	24	54	91	114	162	231	314	26	340
4.07	3.78	26	60	96	134	187	279	325	7	332
3.78	3.56	29	70	120	163	226	295	323	2	325
3.56	3.38	31	75	136	185	255	322	344	3	347
3.38	3.23	37	98	157	208	251	306	322	0	322
3.23	3.11	46	115	183	236	290	311	320	0	320
3.11	3.00	74	143	213	259	298	320	326	0	326
All hkl		333	757	1231	1608	2103	2757	3212	117	3329

Shell Lower limit	Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	6.46	5.0	11.5	18.4	23.7	33.0	53.6	82.7	14.8	97.5
6.46	5.13	7.3	16.7	29.8	38.9	53.5	82.5	98.2	1.8	100.0
5.13	4.48	7.0	13.5	20.5	27.8	40.7	67.0	93.6	6.1	99.7
4.48	4.07	7.1	15.9	26.8	33.5	47.6	67.9	92.4	7.6	100.0
4.07	3.78	7.8	18.1	28.9	40.4	56.3	84.0	97.9	2.1	100.0
3.78	3.56	8.9	21.5	36.9	50.2	69.5	90.8	99.4	0.6	100.0
3.56	3.38	8.9	21.6	39.2	53.3	73.5	92.8	99.1	0.9	100.0
3.38	3.23	11.5	30.3	48.6	64.4	77.7	94.7	99.7	0.0	99.7
3.23	3.11	14.3	35.8	57.0	73.5	90.3	96.9	99.7	0.0	99.7
3.11	3.00	22.7	43.9	65.3	79.4	91.4	98.2	100.0	0.0	100.0
All hkl		10.0	22.7	36.8	48.1	62.9	82.5	96.1	3.5	99.6