THE EMBO JOURNAL

Manuscript EMBO-2011-77165

Crystal structure of the Haemophilus influenzae Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation

Guoyu Meng, Nicole Spahich, Roma Kenjale, Gabriel Waksman and Joseph W. St. Geme, III

Corresponding author: Guoyu Meng, Rui-Jin Hospital affiliated to Shanghai JiaoTong University School of Medicine

Review timeline:

Submission date:	31 January 2011
Revision received:	28 February 2011
Revision received:	24 June 2011
Editorial Decision:	08 July 2011
Revision received:	09 July 2011
Accepted:	15 July 2011

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 28 February 2011

Dear Dr. Meng,

Thank you very much for submitting your research manuscript for consideration to The EMBO Journal editorial office.
You will find reports from three expert scientists attached to this E-Mail. As you will see, these scientists find the subject matter n principle very interesting and the structural work of high standards. Still, all three referees express concerns related to both experimentation as well as presentation of your structural results on the Hap adhesin potentially revealing a mechanism for oligomerization and thus biofilm formation. The major issues of ref#1 and #3 are (1) insufficient experimental support for the model (please refer to particularly point 13 of the comments) and (2) the proposal seems to enable oligomerization only in one dimension, an observation very much unlikely based on the biology of biofilm formation. All referees thus demand significant further modeling as well as experimental efforts to proof the significance of the deduced structural predictions.

As these are from the perspective and scope of a rather more general journal relatively severe concerns, I do urge you to invest the necessary time and experimental efforts to address them satisfactorily.
Conditioned on such modifications/expansions, we would be delighted to re-assess a modified version for potential publication in The EMBO Journal.

I also have to formerly remind you that it is EMBO_J policy to allow a single round of major

revisions and that the decision on acceptance of the work entirely depends on the content and strength of the final submission.

Yours sincerely,

Editor
The EMBO Journal

REFEREE REPORTS:

Referee #1:

The manuscript of Meng et al., Crystal structure of Haemophilus.... reports the crystal structure of Hap adhesin, and based on the crystal packing where two Hap molecules interact in trans position they propose a mechanism for intercellular Hap-mediated interactions leading to biofilm formation. While the manuscript deals with an important natural phenomenon some of the presented predictions would benefit of better experimental evidence (see point 13 below).

1. My impression is that there are too many figures in the manuscript. All of them do not serve the purpose, and some are difficult to interpret.
2. Figure 1. Authors include Figure 1 in the Introduction: Figure 1A has upper and lower panel, and figures 1B and 1C show results. In addition, figure 1A legend contains information that should be included in the introduction. I suggest that panels B and C are moved elsewhere and only Figure 1A lower and upper panel remain, perhaps as new Figure 1A and 1B. Relocate SEM pictures to results with proper explanation in the text, and move electron density picture as trivial to SI. In addition, for Figure 1C the sigma value for electron density was not given and the reference on p6 ".. The structure was solved... as search templates (Figure 1C)" is not meaningful.
3. Figure 2. Figure 2B could be moved to supplementary information.
4. p7 last sentence of middle paragraph, the authors could try to explain more clearly the interactions between autochaperone and beta-helix domains by beta-augmentation.
5. Figure 3: Serine protease domain. A: information about SD1-SD4 could be implemented in Figure 2 using arrows to point the SDs. The orientation of SP domain in respect to beta-spine is illustrated already in Figure 2, electrostatics are in Figure 3D, Figure 3B is impossible to read - rmsd of the molecules (X ≈ with N Calpha) would suffice, so Figures 3A and 3B could be removed. Finally, the authors should provide more detailed information how SPLI was simulated to the structure. This information has not been presented in any detail.
6. Supplementary figure 2. Panels C-F. No explanation is provided which residues belong to SP/other part of the molecule, and, in addition, the numbers are difficult to read and sticks are impossible to see. How about a table of the interactions? Panels A-B are ok, but could they be rather be part of SFig 1?
7. Supplementary figure 3: this information should be included in figure 3D.
8. Serine protease domain interactions to the rest of the molecule are described in four different figures in stereo, but in none of them the interaction partner is specified. As the SDs are mentioned to form a platform for SP domain, it would be informative to know the nature of the interactions, distances, networking etc. Now the information is in the figures and hard to see (labels and sticks are far too small). Also, although structure comparison is made to other SP domains, nothing is mentioned about conservation of the residues forming the binding site.
9. Ten sentences are started with "As shown in the Figure..." Repetition.
10. How about conservation of the residues forming the hydrophobic core of beta-helix? Is every layer/turn similar or different, and if so, how different?
11. p.11 end of first paragraph. Aggregation/dimerization contacts: "secondary interaction site", just to confirm, with the information provided, does it illustrate a crystal contact?
12. The SEM picture from the intro could be implemented in the beginning of the "Functional characterization of self-associating activity" with a proper text referring to the result.
13. Is there any evidence for self-association in trans in addition to crystal packing? The manuscript would benefit if the multimerization model could be verified. The only experimental

evidence is presented in Fig.5 where the effect of a set of deletions is analyzed by tube setting assay. To provide more direct evidence there are several alternative strategies. This might be possible for example by using a FRET-like assay or by coating latex-particles with purified Hap (similar to Yersinia Invasin by Isberg). If latex aggregation can be achieved by wt Hap the assay can be easily used to test modified Haps, i.e., by introducing substitutions to identified interacting residues. Since Ala-substitutions did not work, would substitutions with bulkier amino acids introduce steric hindrance? Also changing the charges could be a feasible approach. Finally, introducing a bigger SD2 (from Hbp) to Hap should prevent aggregation.

14. p 13 line 6. Should this refer to Suppl Fig 2?
15. Move Figure 7 to Supplemental information, and even there one could wonder if less images could be used to make a point?
16. The model presented in figures 4 and 8 gives an impression that the multimerization grows only in one direction while one would expect that it would also grow to both sides as a 2D-lattice. Can this be modeled based on the available data?
17. Suppl Figure 4 legend. Recent structure Nishimura 2010 is mentioned there, but the authors could move it and the the text into discussion.
18. p. 14-15. Discussion on evolutionary relationships with a sequence alignment with zero discussion about it was used as the only supporting argument to conclude finding a missing link between type 1 and 2 monomeric autotransporters. To make such a conclusion more data should be provided.
19. Figure 7C and D are unnecessary, as 7B already states the similarity, and the oligomeric structure is already seen in Figure 4 and 6. The message is clear from the text.
20. Figure 3 legend, end. Is simulated SLPI shown in green or grey? Legend tells grey, in figure it is green.
21. p4 E coli first time should be written Escherichia coli
22. p4,p6 SP is abbreviated for signal peptide on p4, line 7 from bottom, but not used a little later, line 3 from bottom. The same abbreviation is used for chromatography but not explained on p6.
23. p21 ClustalW instead of ClustW
24. p22 OD(600) or OD600
25. p 22 line 4. The final model... Should "residues 920 and" be deleted? Table 1 lists 783 water molecules.

Referee #2:

The authors present a structural study of the H. influenzae Hap self-associating autotransporter. Based upon the packing observed in the crystals of this protein, the authors suggest a model for the mechanism of biological self-association of this protein that is compatible with the role of this protein in biofilm formation.

In general, this is a well-done piece of work with conclusions and a resultant model that are consistent with the biological data on hand. I do have some minor comments/criticisms that I think should be addressed.

1. Figure 2A. Is the length of the overhang illustrated in the head to tail packing (trans configuration) consistent with the length of the C-terminal cell anchor and the extracellular portion of the beta-domain that is absent in the structure? This is key to the validity of the observed packing being biologically relevant. I didn't see it mentioned in the text and it would be helpful if these predicted and measured distances were provided in Figure 2 A.

2. The paper discusses several oligomeric model structures in addition to the modeled complex with SLPI, however there is no mention as to how the models were generated. Were the molecules manually docked? Was any sort of energy minimization carried out on the docked complexes? In addition, the authors utilize a comparison of the Hap structure with a homology model of AIDA-I in their discussion. Again, no mention of how this homology model was generated or statistical analysis of its quality is presented. I think these oversights should be corrected.

3. While it may be common knowledge in the field, it is never expressly stated that the cleavage of Hap from the cell surface is an intermolecular event. Coupled to this thought, how does the proteolytic cleavage from the loss of inhibition by SLPI and cleavage from cell anchoring fit into the model proposed? I.e. within the aggregate that is proposed, how do the authors envision the loss of adhesion? Based upon the model, do the authors propose that release from cell attachment via proteolysis would occur within the aggregate or would the molecules have to disassociate prior to cleavage by a soluble population of Haps?

4. Other than the lack of electron density for the region from 977-1036, what is the evidence that this region is cleaved during purification/secretion rather than it simply being disordered? This is not made any clearer by the conflicting notation in the methods (p. 21) that the region from 977-1036 was absent due to disorder and not cleavage.

5. I believe on page 9 it would be more correct to state that Haps preferentially, not selectively, cleaves after a leucine residue as, in the next sentence, the authors describe cleavage of substrates with a phenylalanine residue at P1. This cleavage would be impossible if the enzyme had selectivity for only a P1 leucine.

6. Hydrophilic edges are seen in structures of other autotransporters. This should be noted. This observation would further support the notion that H-bonding of the domains is not likely important in the self-association pathway. Similarly, it should also be stated that the hydrophobic core of the beta helix is not unique to Hap as it is a hallmark of the structural motif and seen in other autotransporters.

7. Is Haps self-association pH dependant as has been shown for Ag43? If so how would that dependence be explained by the proposed model, which relies upon van der Waal complimentarity driving self-association?

8. Based upon the crystallographic nature of the work, the word choice (more ordered) to describe the nature of the beta-helix in Haps when compared to HBP and IgA protease is likely to be confusing. To most crystallographers casually reading the paper, more ordered implies that the temperature-factors are lower in this domain than the other structures. Contrary to this impression, I think the authors mean that the beta-helix adopts a more regular structure with flatter sides etc., when compared to the same domain of the other two proteases.

9. I agree with the authors that the observation that the deletion mutants are located on the outer membrane of the cells expressing them does imply that the pathway of secretion is unaffected but this result in itself does not support the conclusion that the overall conformation of the proteins is unaffected.

10. With respect to the bacterial aggregation assays it should be made clear that the deletion mutants are done in the S243A background. Further, the lack of error bars associated with Figure 5C makes the author's argument for differences between the mutants much less convincing.

11. Finally, with respect to the model, from a thermodynamic point of view I am confused by the fact that if the thermodynamic force for self-association is so strong, how is it that the protein remains soluble in vitro and does not form insoluble aggregates? It appears that only when Hap is attached to the cell surface is the free energy such that the proposed self-association is favorable. If this is correct, do the authors envision that the C-terminal anchoring tail is involved in a nucleating function that when cleaved is absent and allows for disassociation of the biofilm, or in the case of the soluble protein, prevents its aggregation?

12. p.13, The IgA protease structure that is referenced is from H. influenzae, not Neisseria.

13. Crystallographic Refinement.
What is the justification for using TLS refinement? Did its inclusion improve the R/Rfree? If so what were the R/Rfree values before and after TLS refinement? While noted in the methods section, the space group and unit cell parameters, molecules per ASU, and Ramachandran statistics are missing from Table 1 and should be present there along with the other data and refinement statistics.

14. Figures
In general many figures need to be re-thought as they will be impossible to interpret at a reduced size in the journal. For example, in Figure 4a, the intermolecular distance labels are too small. Further, the distance between Q940 and N577 should be labeled. In figure 7A, the sequence alignment is not legible even at its current size.

In figure 1A, the coloring of the cartoon model of Hap adhesion, aggregation and invasion is confusing. Based upon the domain coloring above this cartoon in the same panel, one would think that the signal peptide domain corresponds to the red oval in the lower panel and Haps corresponds to the blue rectangle. In the figure legend for panel C, the sigma level for the maps should be provided as well as the fact that the figure is a stereographic view.

In supplementary figures 1 and 2, the rendering of IgAP protease in panel B is incorrect. The beta-strands are rendered as coils rather than ribbons, in contrast to the rendering of Hap and HBP in the superposition in panel A.

Referee #3:

This paper describes the crystal structure of the passenger domain of Hap, a self-associating autotransporter protein. These proteins are of great clinical interest, and to date few autotransporter structures have been solved. The subject matter is therefore suitable for the journal, and the crystallography has been carried out competently. My only concern is that the principal conclusion of the paper, that the crystal packing reflects the in vivo interactions by which the protein self-associates, is not adequately supported. The figures strongly suggest to me that the packing interface only explains polymerization in one dimension, not two, and the functional data do not strongly support the conclusion reached on the basis of the molecular model. Given the importance of the model, the first of this group of autotransporters to be crystallized, there is a case for publication but I think revision is advisable taking into account the comments below.

1) The cleavage of Hap could be introduced more fully in the introduction. Does the protein self-cleave though its serine protease domain, and if so where does it cut?

2) SLPI is not adequately introduced. If this protein has a known crystal structure it should be cited in the introduction.

3) Page 8. What is a plucked plane?

4) Figure 4 showing the primary interaction site is poor. A stereo figure is needed and the residues shown as ball-and-stick models should be labeled.

5) On page 11 it is suggested that the dimer interaction buries 1173 sq Angstroms of surface area, but the burial of 3 interfaces buries almost exactly 6 times as much. Given that the N terminal region (residues 725-977) is known not to promote polymerization, it is suspicious that such a large buried SA is reported for the tetramer. I worry that a mistake has been made in using the CONTACT program and surfaces have been counted twice. What area is attributed to the non-functional contact through residues 725-977, and how does this surface differ from the "primary interaction site"?

6) The main concern I have is that the F1-F2 edge to F2 face packing can only explain growth of a polymer in one dimension. I found Figure 6 very unclear (it duplicates Figure 4 in any case). The authors have not shown that molecules lying in one plane can interact with other molecules lying in a separate plane facing them to give a two dimensional lattice. The packing is reported to be "unprecedented" but also that the packing is the same as observed in an Hbp mutant (PDB 3AK5). Downloading this model only increased my concern that the packing interaction described is one-dimensional. Perhaps I have misunderstood, but it seems to me that another interaction is required to create a two-dimensional sheet.

7) In figure 5 it is suggested that deletions around the primary interaction site give a significant loss of self-association (page 12). The settling assay however shows substantially greater self-association

than the control. Is it possible the deletion mutants are prone to unfold and interact through non-specific apolar interactions?

8) If Hap self-association does occur through a largely apolar interaction as suggested (page 17) then the polymerization will be temperature dependent, and much weaker at 4 degrees C than 37 degrees. Analytical ultracentrifugation or laser light scattering should give an indication of association at chosen temperatures. It is remarkable that the protein can be concentrated stably, but this does allow precise techniques to be used. Not all the self-association experiments with mutants are reported, but it would be interesting to see how regions known not to affect self-association map onto the molecular surface.

9) Figure 3. Is color saturation for the electrostatic potential really at +/- 63 kT?! That seems very high, what software was used?

1st Revision - Authors' Response 24 June 2011

REVIEWER 1

1. My impression is that there are too many figures in the manuscript. All of them do not serve the purpose, and some are difficult to interpret.
Response: In the revised manuscript, the figures have been streamlined and reorganized to make the presentation clearer to readers.

2. Figure 1. Authors include Figure 1 in the Introduction: Figure 1A has upper and lower panel, and figures 1B and 1C show results. In addition, figure 1A legend contains information that should be included in the introduction. I suggest that panels B and C are moved elsewhere and only Figure 1A lower and upper panel remain, perhaps as new Figure 1A and 1B. Relocate SEM pictures to results with proper explanation in the text, and move electron density picture as trivial to SI. In addition, for Figure 1C the sigma value for electron density was not given and the reference on p6 ". The structure was solved... as search templates (Figure 1C)" is not meaningful.
Response: As suggested by the referee, we have changed Figure 1. Figure 1B has been moved to become the new Figure 6A, and Figure 1C has been moved to become the new Supplementary Figure 1A. In addition, the sigma value for electron density is now given in the figure legend of Supplementary Figure 1. The sentence in question on page 6 has been changed to read: "The structure of Hap_S was solved by molecular replacement, using the published E. coli Hbp (Otto et al, 2005) and H. influenzae IgA1 protease coordinates (Johnson et al, 2009) as initial search templates (Supplementary Figure 1A)."

3. Figure 2. Figure 2B could be moved to supplementary information.
Response: Figure 2B has been moved to become the new Supplementary Figure 1B, and the residues in Sub-Domains in this figure have been highlighted with green boxes.

4. p7 last sentence of middle paragraph, the authors could try to explain more clearly the interactions between autochaperone and beta-helix domains by beta-augmentation.
Response: Text describing the interaction between the autochaperone domain and the beta-helix stem has been added on page 7: "The β67-β69 pair contributes three hydrogen bonds and the β68-β71 pair contributes four hydrogen bonds to this junction. As a result of these interactions, the autochaperone domain appears to be a short extension of the β-helix by adding two extra strands on each face."

5. Figure 3: Serine protease domain. A: information about SD1-SD4 could be implemented in Figure 2 using arrows to point to the SDs. The orientation of SP domain in respect to beta-spine is illustrated already in Figure 2, electrostatics are in Figure 3D, Figure 3B is impossible to read - rmsd of the molecules (X Å with N Calpha) would suffice, so Figures 3A and 3B could be removed. Finally, the authors should provide more detailed information how SPLI was simulated to the structure. This information has not been presented in any detail.
Response: The revised Figure 2 now has labels for SD1-SD4, and the old Figure 3A has been changed to become the new Supplementary Figure 1C. These two figures should provide a clear

illustration of the overall structure of Hap_S and the intra-molecular relationship between the SDs and the rest of the Hap_S structure. Figure 3B has been removed, and information about the RMSD values of the molecules being compared is included on Page 9. As requested, we have added a new section in the "Materials and Methods" section on page 24 to describe the docking procedure: "The docking is carried out based on the following two observations: 1) SLPI is known to be the common inhibitor of Hap_S and human neutrophil elastase proteolytic activity; 2) the Hap serine protease domain shares the same fold and a nearly identical active site with human neutrophil elastase. The initial model of SLPI:Hap_S was obtained via structural superimposition of the serine protease domains of Hap_S and human neutrophil elastase. The resulting complex was then subjected to energy minimization using SYBYL (Tripos, St. Louis, MO) to remove structural clashes derived from the manual docking. The (SLPI:Hap_S)n multimer was generated using the same methodology described above."

6. Supplementary Figure 2. Panels C-F. No explanation is provided which residues belong to SP/other part of the molecule, and, in addition, the numbers are difficult to read and sticks are impossible to see. How about a table of the interactions? Panels A-B are ok, but could they be rather be part of SFig 1?
Response: In the revised manuscript, we have converted Supplementary Figure 2 panels C-F to the new Supplementary Figure 3 and have provided an improved presentation. The font and stick sizes have been made larger to improve visibility. Residues in the serine protease domain are labeled in black, and residues in other parts of the Hap_S structure are labeled in red. This color-coding together with the explanations in the figure legend makes it easier for the reader to understand the intra-molecular interaction surrounding the proteolytic domain. Supplementary Figure 2 panels A-B have been combined with the old Supplementary Figure 1 as the new Supplementary Figure 2 in the revised manuscript.

7. Supplementary Figure 3: this information should be included in Figure 3D.
Response: In response to the referee's suggestion, Supplementary Figure 3 has been changed to the new Figure 3A.

8. Serine protease domain interactions to the rest of the molecule are described in four different figures in stereo, but in none of them the interaction partner is specified. As the SDs are mentioned to form a platform for SP domain, it would be informative to know the nature of the interactions, distances, networking etc. Now the information is in the figures and hard to see (labels and sticks are far too small). Also, although structure comparison is made to other SP domains, nothing is mentioned about conservation of the residues forming the binding site.
Response: See our reply to point 6 above. In the new figure, the interaction partners are specified both in the figure legend and in the figure. In addition, a two-color labeling scheme has been used for interactions from different regions. As suggested by the referee, the intra-molecular interactions observed in Hap_S are compared with the intra-molecular contacts in Hbp and IgA1 structures. As shown in the new Supplementary Figure 2BC, this mode of intra-molecular contact is conserved in the Hap, Hbp, and IgA1 autotransporters. This information has been added to the revised manuscript in page 8: "Interestingly, this mode of interaction appears to be conserved in Hap, Hbp, and IgA1 protease autotransporters (Supplementary Figure 2CD)." The residues that are strictly conserved among Hap, Hbp, and IgA1 are underlined in the new Supplementary Figure 3.

9. Ten sentences are started with "As shown in the Figure..." Repetition.
Response: As suggested, this repetition has been reduced.

10. How about conservation of the residues forming the hydrophobic core of beta-helix? Is every layer/turn similar or different, and if so, how different?
Response: We appreciate the referee's suggestion. A novel feature of the Hap_S structure has been identified. In the revised manuscript, a new figure (Supplementary Figure 4) has been added to show the conserved pattern in the inner-core of the b-helix, which is also part of the SAAT domain. The following text describing and discussing this finding has been added on page 11: "The inner core of the SAAT domain is packed with relatively conserved hydrophobic/aromatic residues along the axis of the β-helix, including the sequence motif (I/L)XLXXXXX(A/F)X(V/L), in which X represents a random amino acid. As shown in Supplementary Figure 4, Hap_S appears to adopt a strictly conserved inner core to assemble/extend the β-helix region of the SAAT domain. The orientations of the side-chains of the Ile, Leu, Ala, and Val residues in the β-helix turns of residues

732-742 and residues 772-782 are nearly identical. The same applies to the Val, Leu, Phe, and Leu residues in the β-helix turns of residues 752-762 and 791-801. Interestingly, this repetitive structural feature is not seen in the inner core of the Hbp and IgA1 autotransporters, which lack SAAT activity, provoking the thought that this feature may play an important structural role shaping the prism-like morphology that appears to be critical for the SAAT activity."

11. p.11 end of first paragraph. Aggregation/dimerization contacts: "secondary interaction site", just to confirm, with the information provided, does it illustrate a crystal contact?
Response: The available crystal packing is thoroughly reviewed in the Discussion, including reference to new biochemical data. The structural data show that Hap possesses a self-complementary surface that enables inter-molecular oligomerization in a trans-configuration (Figure 4-5). This observation is in good agreement with the available biological studies, leading to the proposed inter-molecular oligomerization model. In this model, hydrogen bonding is not essential for overall binding. Furthermore, published data showed that Q940-N557 is not required for Hap-Hap interaction, leading to the hypothesis that the Q940-N557 hydrogen bond might contribute to overall binding but is not essential, as elimination of this bond via mutagenesis had little effect on Hap-Hap interaction. Hence, the region surrounding the Gln-Asp H-bonding is considered a secondary interaction site.

12. The SEM picture from the intro could be implemented in the beginning of the "Functional characterization of self-associating activity" with a proper text referring to the result.
Response: The SEM image has been moved to the new Figure 6.

13. Is there any evidence for self-association in trans in addition to crystal packing? The manuscript would benefit if the multimerization model could be verified. The only experimental evidence is presented in Fig.5 where the effect of a set of deletions is analyzed by tube setting assay. To provide more direct evidence there are several alternative strategies. This might be possible for example by using a FRET-like assay or by coating latex-particles with purified Hap (similar to Yersinia Invasin by Isberg). If latex aggregation can be achieved by wt Hap the assay can be easily used to test modified Haps, i.e., by introducing substitutions to identified interacting residues. Since Ala-substitutions did not work, would substitutions with bulkier amino acids introduce steric hindrance? Also changing the charges could be a feasible approach. Finally, introducing a bigger SD2 (from Hbp) to Hap should prevent aggregation.
Response: As suggested by the referee, we have performed additional testing of Hap adhesive activities by coating latex beads with wild-type and mutant proteins and then examining the beads for aggregation (new Figure 7). The results of these experiments support the studies examining whole bacteria by SEM and in tube settling assays (Figure 6). In particular, beads coated with wild type Hap_S formed aggregates when viewed by light microscopy (Figure 7C). In contrast, beads coated with Hap_SD751-827 (lacking four rungs of helical turns in the SAAT domain) failed to aggregate. With this information in mind, we have now demonstrated Hap_S self-associating properties by three different approaches, namely SEM of whole bacteria, tube settling assays using whole bacteria, and latex beads coated with purified protein. All of the functional data presented in this report and in previous publications supports the conclusion that the Hap_S-Hap_S interaction revealed by crystal packing is biologically relevant and represents a genuine snapshot of Hap_S-mediated interbacterial interaction.

As with other self-associating autotransporters (Sherlock, et al, 2004; Cote and Mourez, 2011; Klemm et al, 2006; Sheets and St Geme, 2010), in our previous studies the tube settling assay proved to be a good model to study Hap_S-mediated cell-cell adhesion (Hendrixson et al, 1998; Fink et al, 2003). Bacteria lacking Hap are not able to form bacterial aggregates and have a much slower rate of settling. As suggested by the referee, we have mutated bulky amino acids (such as Trp) and charged amino acids (such as Arg) at the 1st interaction site. These mutations were intended to introduce steric or electrostatic hindrance but had little impact on Hap_S-Hap_S interaction. In fact, this observation is not surprising given that the minimum inter-molecular distance at the 1st interaction site is 4-6 Å (Figure 4). To test the hypothesis that Hap_S has a certain degree of tolerance of amino acid variation at sites of self-association, we have computationally mutated Asp/Asn ladders to Trp-ladders or Arg-ladders (in both molecules of the trans Hap-Hap dimer). As depicted in the figure below, the Trp-ladders and Arg-ladders do not show obvious structural clashes. Interestingly, there is no apparent sequence identity between the Hap_S SAAT domain and the SAAT-domains in other SAAT proteins, including Ag43, TibA, and AIDA-I, further

suggesting that SAAT-type interaction is dictated by overall architecture and self-complementary surfaces.

In additional work, we attempted to insert extra sequences in the 1st interface. Unfortunately, random insertion introduced presumed abnormalities in folding and the mutant proteins suffered premature degradation, preventing further functional studies.

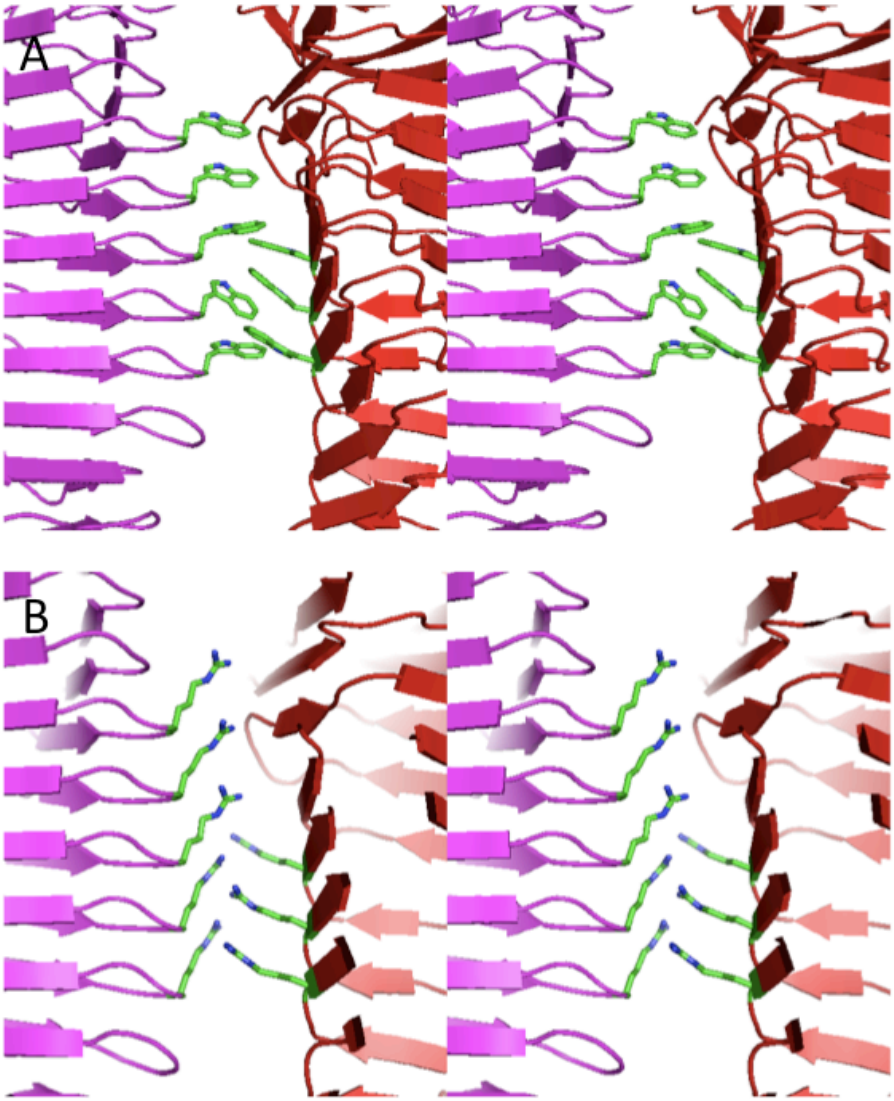

Figure 1. Simulated Trp ladder and Arg ladder in the 1st interaction site.

14. p 13 line 6. Should this refer to Suppl Fig 2?

Response: We appreciate the referee's careful review and have made the suggested change.

15. Move Figure 7 to Supplemental information, and even there one could wonder if less images could be used to make a point?

Response: As suggested, Figure 7 has been changed to the new Supplementary Figure 6.

16. The model presented in figures 4 and 8 gives an impression that the multimerization grows only in one direction while one would expect that it would also grow to both sides as a 2D-lattice. Can this be modeled based on the available data?

Response: Hap is an autotransporter protein and contains an outer membrane anchoring domain. It has been shown this domain might "move" within the outer membrane, triggered by the interactions involving the surface-associated N-terminal passenger domain, Hap_S. To clarify this point, we have added the following text on pages 18-19: "Furthermore, it has been shown that the membrane-anchoring domain in autotransporter proteins is mobile in lipid bilayers (Jose & Meyer, 2007). Hence, it is reasonable to envisage that the mobile Hap_S domain forms multiple arrays of oligomers at the cross-section of cell junctions upon bacterial aggregation. These "lines" of Hap_S multimer can act like "stitches" to seal cells together, leading to biofilm formation."

17. Suppl Figure 4 legend. Recent structure Nishimura 2010 is mentioned there, but the authors could move it and the the text into discussion.

Response: Text describing the recent publication by Nishimura, 2010 has been included in the revised manuscript on page 15: "Indeed the structure of an Hbp mutant illustrates this structural clash (Nishimura et al., 2010). As shown in Supplementary Figure 5, deletion of sub-domain 2 in Hbp enables assembly into a SAAT-like packed structure."

18. p. 14-15. Discussion on evolutionary relationships with a sequence alignment with zero discussion about it was used as the only supporting argument to conclude finding a missing link between type 1 and 2 monomeric autotransporters. To make such a conclusion more data should be provided.

Response: In addition to the sequence alignment, we have included several lines of evidences to support our claim. Based on sequence alignments, Hap_S is recognized as a type 1 monomeric autotransporter. Indeed, the Hap_S structure shows great similarity to Hbp and IgA1 protease (which are also type 1 autotransporters). However, only Hap_S has self-adhesive activity (a unique activity otherwise found exclusively in type 2 autotransporters). The structure of Hap_S highlights unique alterations that enable self-associating activity despite similarity to type 1 autotransporters, namely, an SD1 molecular velcro, a smaller SD2, a relatively regular SAAT domain, and an absent C-terminal helix at the Hap_S-Hap_S interface). The differences between Hap and other type 1 autotransporters is further supported by the crystal packing observed in the Hbp mutant structure. Based on these observations, we believe that it is reasonable to highlight this evolutionary relationship.

19. Figure 7C and D are unnecessary, as 7B already states the similarity, and the oligomeric structure is already seen in Figure 4 and 6. The message is clear from the text.

Response: As suggested by the referee, Figure 7C and Figure 7D have been removed.

20. Figure 3 legend, end. Is simulated SLPI shown in green or grey? Legend tells grey, in figure it is green.

Response: We appreciate the referee's attention to detail, and we have corrected the typographical error in the figure legend.

21. p4 E coli first time should be written Escherichia coli

Response: As suggested, we have changed the first reference to E. coli to Escherichia coli.

22. p4, p6 SP is abbreviated for signal peptide on p4, line 7 from bottom, but not used a little later, line 3 from bottom. The same abbreviation is used for chromatography but not explained on p6.

Response: In the revised manuscript we have used SP to refer only to signal peptide and have spelled out "sepharose cation-exchange" chromatography where appropriate.

23. p21 ClustalW instead of ClustW

Response: We have changed ClustW to ClustalW.

24. p22 OD(600) or OD600

Response: As suggested, in the revised manuscript we have changed OD600 to OD(600).

25. p 22 line 4. The final model... Should "residues 920 and" be deleted? Table 1 lists 783 water molecules.

Response: As suggested, we have changed this sentence to read: "The final model of Hap_S contains 920 residues and 783 water molecules."

REVIEWER 2

1. Figure 2A. Is the length of the overhang illustrated in the head to tail packing (trans configuration) consistent with the length of the C-terminal cell anchor and the extracellular portion of the beta-domain that is absent in the structure? This is key to the validity of the observed packing being biologically relevant. I didn't see it mentioned in the text and it would be helpful if these predicted and measured distances were provided in Figure 2 A.

Response: The predicted distance of the C-terminal cell anchor and the extracellular portion of the b-domain is consistent with the measured distance of the overhang illustrated in the head to tail packing (Supplementary Figure 7), providing further validation of the observed packing. In the revised manuscript, we have added the following text on page 16: "Second, the measured distance of the overhang in the trans configuration is consistent with the predicted length of the linker loop attached to the cell membrane (Supplementary Figure 7). Assuming that the average distance of Cα-Cα in a stretched conformation is about 2.3 Å, we can estimate that the theoretical length of

residues 977-1036, the bridge between the C-terminal end of Hap_S and the cell anchor (Hap_b) is ~135 Å, much longer than the overhang in the trans configuration, giving ample space for Hap-Hap interactions."

2. The paper discusses several oligomeric model structures in addition to the modeled complex with SLPI, however there is no mention as to how the models were generated. Were the molecules manually docked? Was any sort of energy minimization carried out on the docked complexes? In addition, the authors utilize a comparison of the Hap structure with a homology model of AIDA-I in their discussion. Again, no mention of how this homology model was generated or statistical analysis of its quality is presented. I think these oversights should be corrected.
Response: In the revised manuscript, we have expanded the "Materials and Methods" section to explain, 1) how SLPI was docked into the Hap structure, and 2) how the homology model of AIDA-I was generated.

3. While it may be common knowledge in the field, it is never expressly stated that the cleavage of Hap from the cell surface is an intermolecular event. Coupled to this thought, how does the proteolytic cleavage from the loss of inhibition by SLPI and cleavage from cell anchoring fit into the model proposed? I.e. within the aggregate that is proposed, how do the authors envision the loss of adhesion? Based upon the model, do the authors propose that release from cell attachment via proteolysis would occur within the aggregate or would the molecules have to disassociate prior to cleavage by a soluble population of Haps?
Response: In the revised manuscript, the intermolecular cleavage event is now expressly stated on page 5: "The serine protease domain mediates intermolecular autoproteolysis at L1036-N1037, L1046-T1047, F1077-A1078, and F1067-S1068 (termed the primary, secondary, tertiary, and quaternary cleavage sites, respectively), resulting in release of the Hap_S passenger domain from the bacterial surface and modulating bacterial adherence and aggregation (Fink & St Geme, 2003; Hendrixson & St Geme, 1998)." Intermolecular cleavage of Hap is believed to be an important event that is influenced by the concentration of surface associated Hap_S and that controls Hap-mediated bacterial aggregation. This information is consistent with the proposed polymerisation/depolymerisation model. In this model, the Hap_S monomer with better solubility and smaller size for penetration might function than the Hap_S multimer in autoproteolysis and invasion. This has led to hypothesis in page 19: "This information implies that formation of a Hap_S-Hap_S dimer and a mega-Dalton multimer might be a thermodynamic process constantly associated with a polymerization/depolymerization mechanism, reminiscent of growth of an actin filament. ... In the context of H. influenzae pathogenicity, cleaved and released Hap_S is a monomer and is highly soluble, potentially an advantage in migrating through the extracellular matrix to cleave host proteins."

4. Other than the lack of electron density for the region from 977-1036, what is the evidence that this region is cleaved during purification/secretion rather than it simply being disordered? This is not

made any clearer by the conflicting notation in the methods (p. 21) that the region from 977-1036 was absent due to disorder and not cleavage.

Response: In the revised manuscript, we have rephrased this section to improve clarity. The reason that the region corresponding to residues 977-1036 is not available for model building is stated on page 7: "Residues 1-25 represent the signal peptide and hence are not present in the structure of Hap_S*. Residues 265-272, 851-873, and 977-1036 yielded no electron density, either because they are disordered (265-272, 851-873) or were cleaved during the process of Hap secretion or purification (977-1036), making them unavailable for model building."*

5. I believe on page 9 it would be more correct to state that Haps preferentially, not selectively, cleaves after a leucine residue as, in the next sentence, the authors describe cleavage of substrates with a phenylalanine residue at P1. This cleavage would be impossible if the enzyme had selectivity for only a P1 leucine.

Response: As suggested, in the revised manuscript we have replaced the word "selectively" with the word "preferentially."

6. Hydrophilic edges are seen in structures of other autotransporters. This should be noted. This observation would further support the notion that H-bonding of the domains is not likely important in the self-association pathway. Similarly, it should also be stated that the hydrophobic core of the beta helix is not unique to Hap as it is a hallmark of the structural motif and seen in other autotransporters.

Response: We have added text on page 11 stating that hydrophilic edges are a common feature in the structure of autotransporters. As suggested by the referee, we have added this point to the discussion to argue that inter-molecular H-bonding is likely not important in the self-association pathway. As for the inner core, Hap uses a repetitive sequence to assemble the SAAT domain (see the reply to referee 1, point 10), giving rise to a regular prism-like morphology that mediates formation of Hap_S*-*Hap_S *dimers and enables higher order oligomerization through the F1-F2 edge and F2 face.*

7. Is Hap_S self-association pH dependent as has been shown for Ag43? If so how would that dependence be explained by the proposed model, which relies upon van der Waal complementarity driving self-association?

Response: We have no information about whether Hap-Hap self-association is pH dependent. The low sequence identity (<20%) between Hap and Ag43 limits interpretation on the pH-dependent self-association activity observed in Ag43. For example, besides the self-complementary surface observed in the Hap structure, Ag43 may also have other factors, such as specific electrostatic patches, to regulate the self-association pathway. Furthermore, it may be reasonable to envisage that the Hap C-terminal linker (i.e. residues 977-1036) adopts a different conformation that could prevent SAAT activity under certain pH conditions.

8. Based upon the crystallographic nature of the work, the word choice (more ordered) to describe the nature of the beta-helix in Hap_S when compared to Hbp and IgA1 protease is likely to be confusing. To most crystallographers casually reading the paper, more ordered implies that the temperature-factors are lower in this domain than the other structures. Contrary to this impression, I think the authors mean that the beta-helix adopts a more regular structure with flatter sides etc., when compared to the same domain of the other two proteases.

Response: We appreciate the referee's insightful point and have changed "more ordered" to "more regular."

9. I agree with the authors that the observation that the deletion mutants are located on the outer membrane of the cells expressing them does imply that the pathway of secretion is unaffected but this result in itself does not support the conclusion that the overall conformation of the proteins is unaffected.

Response: The deletions are guided by the crystal structure. Based on the structure, deletion of one or two rungs of the beta-helix does not change the overall conformation (as the inner core of the beta-helix is repetitive). In marked contrast to other mutants (such as insertion of sequence into the beta-helix), the fact that the deletion mutants are stable in the outer membrane supports the claim on page 13 that "the deletions have little effect on the overall conformation or secretion of the protein."

10. With respect to the bacterial aggregation assays it should be made clear that the deletion mutants are done in the S243A background. Further, the lack of error bars associated with Figure 5C makes the author's argument for differences between the mutants much less convincing.

Response: In the revised manuscript, we have clarified that the deletion mutants were generated in the S243 background. In addition, we have added error bars in the revised figure.

11. Finally, with respect to the model, from a thermodynamic point of view I am confused by the fact that if the thermodynamic force for self-association is so strong, how is it that the protein remains soluble in vitro and does not form insoluble aggregates? It appears that only when Hap is attached to the cell surface is the free energy such that the proposed self-association is favorable. If this is correct, do the authors envision that the C-terminal anchoring tail is involved in a nucleating function that when cleaved is absent and allows for disassociation of the biofilm, or in the case of the soluble protein, prevents its aggregation?

Response: In our previous studies, we found that Hap_S in solution can form multimers at high protein concentration (Hendrixson and St Geme, 1998). We think that the local protein concentration in the membrane could trigger Hap multimerization. We have no information at present about whether the C-terminal linker loop (containing about 50 residues) plays a role in intercellular oligomerization. One could envisage that the attachment of Hap to the outer membrane might help to promote a productive-conformation (making sure this domain/loop would not be in the way of oligomerization) that allows the C-terminal SAAT domains to interact with each other.

However, without biological evidence, this point remains too speculative and hence is not included in the discussion.

12. p.13, The IgA protease structure that is referenced is from H. influenzae, not Neisseria.
Response: We appreciate the referee's careful review of the manuscript and have corrected this error.

13. Crystallographic Refinement.
What is the justification for using TLS refinement? Did its inclusion improve the R/Rfree? If so what were the R/Rfree values before and after TLS refinement? While noted in the methods section, the space group and unit cell parameters, molecules per ASU, and Ramachandran statistics are missing from Table 1 and should be present there along with the other data and refinement statistics.
Response: Indeed, using TLS has improved the Rfree factor by 0.3%. The R/Rfree values by refmac 5 (CCP4) before and after TLS refinement are as follows: 18.6, 23.3 (before TLS refinement) and 18.4, 23.0 (after TLS refinement). As suggested by the referee, the space group and unit cell parameters, molecules per ASU, and Ramachandran statistics have been added to the revised Table 1.

14. Figures
In general many figures need to be re-thought as they will be impossible to interpret at a reduced size in the journal. For example, in Figure 4A, the intermolecular distance labels are too small. Further, the distance between Q940 and N577 should be labeled. In Figure 7A, the sequence alignment is not legible even at its current size.
Response: As suggested by the referee, we have modified the figures to improve their readability and interpretation.

In Figure 1A, the coloring of the cartoon model of Hap adhesion, aggregation and invasion is confusing. Based upon the domain coloring above this cartoon in the same panel, one would think that the signal peptide domain corresponds to the red oval in the lower panel and Haps corresponds to the blue rectangle. In the figure legend for panel C, the sigma level for the maps should be provided as well as the fact that the figure is a stereographic view.
Response: The color-coding in Figure 1 has been changed to avoid confusion. In the revised manuscript, the sigma level of the maps has been included in the figure legend.

In supplementary figures 1 and 2, the rendering of IgAP protease in panel B is incorrect. The beta-strands are rendered as coils rather than ribbons, in contrast to the rendering of Hap and HBP in the superposition in panel A.
Response: We appreciate the referee's comment, and we have corrected the rendering of IgA1 protease in the revised figures.

REVIEWER 3

1) The cleavage of Hap could be introduced more fully in the introduction. Does the protein self-cleave though its serine protease domain, and if so where does it cut?
Response: As suggested, in the revised manuscript we have added information about the intermolecular cleavage of Hap on page 5: "The serine protease domain mediates intermolecular autoproteolysis at L1036-N1037, L1046-T1047, F1077-A1078, and F1067-S1068 (termed the primary, secondary, tertiary, and quaternary cleavage sites, respectively), resulting in release of the Hap_S passenger domain from the bacterial surface and modulating bacterial adherence and aggregation (Fink & St Geme, 2003; Hendrixson & St Geme, 1998)."

2) SLPI is not adequately introduced. If this protein has a known crystal structure it should be cited in the introduction.
Response: In the revised manuscript we have added citation of the SLPI structure in the Introduction on page 5.

3) Page 8. What is a plucked plane?
Response: In the revised manuscript we have replaced the term "a plucked plane" with "a bended plane."

4) Figure 4 showing the primary interaction site is poor. A stereo figure is needed and the residues shown as ball-and-stick models should be labeled.
Response: In the revised manuscript, a new stereo figure showing the primary interaction has been added and the residues in the primary interaction site are now labeled.

5) On page 11 it is suggested that the dimer interaction buries 1173 sq Angstroms of surface area, but the burial of 3 interfaces buries almost exactly 6 times as much. Given that the N terminal region (residues 725-977) is known not to promote polymerization, it is suspicious that such a large buried SA is reported for the tetramer. I worry that a mistake has been made in using the CONTACT program and surfaces have been counted twice. What area is attributed to the non-functional contact through residues 725-977, and how does this surface differ from the "primary interaction site"?
Response: Taking the trans-dimer as an example, the total surface area of a single Hap molecule is 36,242 sq Angstroms. The self-associating surface of 1173 sq Angstroms is obtained by 36,242 X 2 - (total surface area of a trans dimer, 71,311 sq Angstroms). The SA surface of a tetramer, 7054 sq Angstroms, is obtained by 36,242 X 4 - (total surface area of a trans dimer, 137,914 sq Angstroms). These calculations have been rechecked, and the values are correct. Regarding residues 725-977, this region is the SAAT domain, and the F2 face and the F1-F2 edge in this region contribute to the binding, i.e. polymerization. The non-functional contact through residues 725-977 lies in the F1 and F3 faces that are on the same side of the serine protease domain. Based on the surface alone, it is

not clear why Hap uses the F2 face and the F1-F2 edge for polymerization. However, based on the overall structure, polymerization via the F2 face and the F1-F2 edge, but not via other region, can bring other parts of the structure to interact to contribute to the overall SA surfaces and binding. This point has been clarified on pages 12-13.

6) The main concern I have is that the F1-F2 edge to F2 face packing can only explain growth of a polymer in one dimension. I found Figure 6 very unclear (it duplicates Figure 4 in any case). The authors have not shown that molecules lying in one plane can interact with other molecules lying in a separate plane facing them to give a two dimensional lattice. The packing is reported to be "unprecedented" but also that the packing is the same as observed in an Hbp mutant (PDB 3AK5). Downloading this model only increased my concern that the packing interaction described is one-dimensional. Perhaps I have misunderstood, but it seems to me that another interaction is required to create a two-dimensional sheet.

Response: See reply to referee 1, point 16. The Hap_S packing presented in this manuscript is unknown to the field of autotransporter biology, including in the publication describing the Hbp mutant. This point is not surprising, as there is no suggestion/hypothesis thus far linking Hbp to SAAT-type autotransporters. In the Hbp mutant, Hap-Hap like interaction is hidden among many random interactions derived from crystal packing. As shown in the Supplementary figure 5, the gap between Hbp molecules in a SAAT-like interaction is much larger than that in Hap-Hap association. In comparison, our structure clearly shows a dominant crystal packing via its C-terminal SAAT region. The trans configuration is perfectly agreeable to its role as a cell linker. Actually, the Hap structure was determined before the recent publication of the Hbp mutant. Only after seeing the Hap-Hap like self-association, we were able to identify a similar, but much looser crystal packing in the Hbp mutant. In addition, until our report is published, there is no atomic structure providing insight into how bacteria can come together to form microcolonies. Based on this information, we consider our finding to be novel and "unprecedented".

7) In Figure 5 it is suggested that deletions around the primary interaction site give a significant loss of self-association (page 12). The settling assay however shows substantially greater self-association than the control. Is it possible the deletion mutants are prone to unfold and interact through non-specific apolar interactions?

Response: If this were the case, the protein should be less stable and sensitive to proteolysis. The fact that mutant 751-770 is stable and mimics the SAAT activity of wild type Hap_S further supports the notion that the overall architecture and self-associating surface is important for SAAT activity.

8) If Hap self-association does occur through a largely apolar interaction as suggested (page 17) then the polymerization will be temperature dependent, and much weaker at 4 degrees C than 37 degrees. Analytical ultracentrifugation or laser light scattering should give an indication of association at chosen temperatures. It is remarkable that the protein can be concentrated stably, but this does allow precise techniques to be used. Not all the self-association experiments with mutants are

reported, but it would be interesting to see how regions known not to affect self-association map onto the molecular surface.

Response: As suggested, we have performed DLS on purified Hap_S at a concentration of 3 mg/ml. The results are shown below. At this protein concentration, Hap_S appeared to form larger aggregate as the temperature increased. Furthermore, when a stringent filter of 0.02 μm was used, the light scattering stayed at a water-like baseline suggesting, that Hap_S tends to form large soluble aggregates at this concentration. Considered together with the characterization by the DLS assay (Supplementary Table 1), gel filtration, and native gel electrophoresis, this information has led to the hypothesized polymerization/depolymerization model. The significance of polymerization/depolyermization is discussed in the context of H. influenzae pathogenesis.

The mutants that do not affect SAAT activity are discussed on page 12 in the revised manuscript.

Supplementary Table 1. Dynamic light scattering assay of purified Hap_S

Temperature (°C)	Rd* (nM)	Polydispersity (%)	Baseline	SOS
4	986.1	33.2	0.097	1.029
10	1447.7	33.6	1.000	0.981
16	1509.7	42.2	0.999	1.708
22	1717.3	48.8	0.999	3.253
37	1902.6	56.1	1.003	8.722

*Hydrodynamic radius (Rd) was calculated from 200 independent measurements.

9) Figure 3. Is color saturation for the electrostatic potential really at +/- 63 kT?! That seems very high, what software was used?

Response: We appreciate the referee's careful review of the manuscript. As suggested by the reviewer, APBS plug-in in Pymol is used to estimate the electrostatic potential scale in the surface diagrams. The information of the variance in the electrostatic surface potential is updated on Page 2, in the SI section of the revised manuscript: "The surface is colored according to the electrostatic surface potential (negative charges $-4K_BT$ in red and positive charges $+4K_BT$ in blue, with linear interpolation in between). ... The surface diagrams were prepared using programs APBS (Baker et al, 2001) and Pymol."

We thank the reviewers for their insightful and helpful comments, which have enabled us to prepare a significantly stronger manuscript. Given the enthusiastic response to the first version of this manuscript, we hope the revised version is now acceptable for publication in EMBO J.

Best regards,

Guoyu Meng

2nd Editorial Decision 08 July 2011

Dear Dr. Meng,

I just received comments on your revised manuscript from one of the original referees. This scientist noticed a problem in establishing the electrostatic potential scale. I kindly ask you to resolve this issue and provide us with the ultimate version of your study to enable final acceptance.

Yours sincerely,

Senior Editor
The EMBO Journal

REFEREE REPORT:

Referee #2:

With one exception the corrections made to the manuscript have addressed my previous concerns. My remaining concern relates to the fact that the authors have not adequately addressed reviewer #3's prior concern with the electrostatic surface rendering. The rendering given in the manuscript using the default settings in Pymol as stated in the response letter is completely qualitative in its surface potential scale and as stated previously by reviewer 3. The range the authors give would correspond to approximately -2V to +2V which is not at all realistic. If the authors are going to report an electrostatic potential scale for the figure, they must install the APBS plugin into Pymol in order for them to obtain a realistic representation of the variance in the electrostatic surface potential.

2nd Revision - Authors' Response 09 July 2011

Response: We appreciate the referee's careful review of the manuscript. As suggested by the reviewer, APBS plug-in in Pymol is used to estimate the electrostatic potential scale in the surface diagrams. The information of the variance in the electrostatic surface potential is updated on Page 2, in the SI section of the revised manuscript: "*The surface is colored according to the electrostatic surface potential (negative charges -4KBT in red and positive charges +4KBT in blue, with linear interpolation in between).
…**The surface diagrams were prepared using programs APBS (Baker et al, 2001) and Pymol.**"

