THE
EMBO
JOURNAL

Manuscript EMBO-2012-83802

Structure of a bacterial type IV secretion core complex at subnanometer resolution

Angel Rivera-Calzada, Rémi Fronzes, Christos G. Savva, Vidya Chandran, Pei W. Lian, Toon Laeremans, Els Pardon, Jan Steyaert, Han Remaut, Gabriel Waksman and Elena V. Orlova

Corresponding author: Elena Orlova, Birkbeck College

Review timeline:

Submission date:	01 November 2012
Editorial Decision:	04 December 2012
Revision received:	31 January 2013
Accepted:	19 February 2013

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.)

Editor: David del Alamo

1st Editorial Decision 04 December 2012

Thank you for your patience while your manuscript has been reviewed. We have just now received the full set of reports from the referees, which I copy below. As referee reports are quite explicit I will not repeat their arguments here, but as you will see, they agree that your manuscript is highly interesting and their comments are rather positive.

Given the referees' recommendations, I would like to invite you to submit a revised version of the manuscript. Please be aware that your revised manuscript must address the referees' concerns and their suggestions should be taken on board. It is 'The EMBO Journal' policy to allow a single round of revision only and, therefore, acceptance or rejection of your study will depend on the completeness of your responses included in the next, final version of the manuscript.

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 initiative, please visit our website: http://www.nature.com/emboj/about/process.html

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed. Should you

foresee a problem in meeting this three-month deadline, please let us know in advance and we may be able to grant an extension.

Thank you again for the opportunity to consider your work for publication. I look forward to your revision.

Please, do no hesitate to contact me in case you have any further question, need further input or any problem arises during the revision process.

REFEREE REPORTS:

Referee #1:

The manuscript by Rivera-Calzada et al reports a reanalysis of the bacterial type IV secretion core complex by cryo-EM to improve the resolution of the reconstruction and a novel, higher resolution, analysis of a proteolytically generated smaller complex.
The new data reveal many interesting and novel features in the assembly and as such the manuscript appears to meet the criteria for broad interest required for publication in EMBO J. The major caveat this referee feels is with the attempt to model the VirB9 and Vir B10 which seem unnecessary and introduces many potential errors to the paper. The issue arises because neither domain shares any sequence identity with any proteins of known structure and the authors take a fairly naive approach to modelling them. They make no critical assessments of how reliable the models that have been obtained are (this referee suspects they are very unreliable).
For VirB9 all that is shown is that the domain they identify using Ab labelling and by the differences between the xtal structure and the elastase treated map is of the correct volume for this amount of protein. To demonstrate that they don't really have to introduce the potential for error that comes with de novo modelling of the protein structure. Even more so with their modelling of the N-terminal helices of B10 which must correspond to the difference between the elastase treated and full length maps. Both of these attempts to model the structure within the density don't really add anything - the difference between the two maps clearly suggests which density is likely to be the domains and the modelling simply adds potential for error. This referee would suggest removing these speculative models and publishing the simpler interpretation of the data.

A few minor points could be addressed to further improve the manuscript:

Minor Points
1. This referee is not an expert in EM methodology but the work appears sound and appropriately reported. It would be good for the statement about hand determination to be incorporated in the main manuscript and to be backed up by the correlations in the two hands that allowed determination of hand. In general the correlations between the structures fitted and the maps also appear to be rather low c.f. other papers. Some comment from the authors on why this is so would also be useful.
2. Fig 1. - if the masses of the proteolytic fragments have solely been determined by their position in the SEC trace this should be stated as this is a less than optimal way to determine mass. Can the authors be confident that only the N-termini have been truncated?
3. why were the nanobodies added at a sub-stoichiometric ratio to the complex? If added at a higher ratio then reconstruction of a nanobody-core-complex assembly would be possible and would reveal more information about the location of the epitope.

Typographical changes
page 3, line 8 swap form and likely to read - ....likely form an inner....
page 7, line 12 up from bottom replace is with are .... there ARE AN extra....

Referee #2:

This manuscript is based on the structures of two multiprotein assemblies from the bacterial type IV secretion system (T4SS)- a full-length pKM101 core complex and, at higher resolution, a truncated complex derived from it by limited proteolysis. There is no doubt that other, recently reported,

studies of the T4SS by cryoEM and crystallography have moved our understanding of the process to a new level, and have been ground-breaking. This manuscript seeks to build on that excellent work, by extending and improving the resolution of the derived cryoEM structures. The EM data are of good quality and the work has been carried out to a high technical standard. The authors seek to improve the resolution of their data to the point where atomic models for individual components can be modeled into the electron density, and they describe some progress in doing this, particularly for the I layer, where such information has been lacking to date. The lack of experimentally-derived 3D structures for the TraO and TraF N-termini is a serious limitation here, however, and the authors have to fall back on molecular modelling (presumably because those components were refractory to crystallization or NMR structure determination).

The main drawback of the manuscript is that it lacks a principal central conclusion. In extending the resolution of the EM reconstructions, the findings are incremental on previously published work, but provide relatively little new and important information. The abstract refers to 'new structural insights'- and there are certainly some- but their implications for our understanding of the mechanism of T4S are harder to identify. The most interesting conclusion is the central role which VirB10 plays as a 'signal transmitter' between the ATPases and other components in the assembly. This topic is addressed in the Discussion but is incomplete without additional experimental evidence to elaborate and substantiate this claim. The manuscript also contained a rather high proportion of typographical and grammatical errors, which limited its fluency; I have suggested corrections for some below.

Major
1. Title: 'Subnanometer resolution structure..' would be more accurate
2. P6 Modelling of TraO and TraF N-termini: it is difficult for the reader to gauge how reliable these models are. How do the authors know that they are sufficiently reliable to be used for rigid body fitting? Some additional experimental data would help- eg CD spectra- if the recombinant domains are tractable to study.
3. There is no figure of the final predicted TraO/VirB9NT structure from I-TASSER which was subsequently used for fitting; it would be useful to include this in the Supplementary information, at least.
4. The three predicted helices from TraF/VirB10CT only make up a small proportion of the total mass (see Supplementary Figure 2b). How can the authors be sure that these three helices are indeed responsible for the density features they have attributed to them?
5. P9 Nanobody labelling: is the resolution of this experiment sufficient to distinguish, as the authors conclude at the end of the second paragraph, that TraO/VirB9NT is located 'on the outer part of the IL'? How would the results differ if it were located on the inner part instead? The difficulty here is that the labelling complex- a nanobody bound to an anti-His antibody- is much larger than the labelled domain itself.
6. The labelling of the bands in the gel in Figure 1b) needs to be clearer- it's not clear which band is TraN. Are the other bands contaminants?

Minor
1. P3 line 8 'likely form an inner membrane subcomplex.'
2. P3 line 9 'The extracellular pilus is composed of major and minor pilins, termed VirB2 and VirB5, respectively.'
3. P3 line 14 'Because of its central position, the complex participates actively in T4S substrate transfer through the bacterial envelope.'
4. P4, five lines from bottom '...very close to the measured molecular weight indicated above.'
5. P7 line 22 'Although there are an extra 9 amino acids present at the N-terminus of TraF/VirB10CT in..'
6. P8 line 2 'These might be attributed to the acetyl moiety of VirB7'- I think the authors mean the lipid moiety of VirB7 (which could contribute significant density)?
7. P8 fourth line from bottom 'However, the remaining density is not sufficiently big...' This sentence is rather clumsily phrased- I think the authors mean that the volume determined from the density at a particular sigma value is insufficient to accommodate the linker polypeptide between the helices (which, I agree, seems reasonable).
8. P9 line 8 Why are the nanobodies named '#CA4271' on this line, but a different naming convention seems to be used further down 'NBCA4271' (line 12)? There is further potential for confusion in the next paragraph, where the particular nanobody is simply referred to as 'NB'. A

consistent nomenclature is needed.
9. P12 line 13 '....the solution was gel filtered as described previously..'
10. P14 line 2 '...according to the 0.5 criterion level...'
11. P14 line 3 'To compare two structures, levels of low frequencies and their resolution in the structures were adjusted to 12.4 ≈.' This sentence is rather clumsily phrased- I think the authors mean that they effectively filtered the resolution of the CCelastase structure to 12.4A (i.e. that of FLCC).
12. P14 line 8 'The sequences of TraO/VirB9NT....'
13. Figure 1c) what column was used- it is not clear from the M&M. It is good practice to insert the elution positions of the mass standards used in calibration.

Referee #3:

In this manuscript, Waksman and Orlova and colleagues have presented refined structures of the core complex composed of the TraN, O, and F proteins encoded by the E. coli pKM101 transfer system. This work follows previous reports of cryoEM structure of the core complex and an X-ray structure of the outer half of the core complex. The pKM101 transfer system is a member of the family of type IV secretion systems (T4SSs) functioning in many different bacterial species. This group has pioneered studies aimed at solving the structure of these complex membrane machines and here they have added important new elements to this structural detail. Specifically, they solved a portion of the core complex at 8.5 angstrom resolution and the entire complex at 12 angstroms. By fitting the X-ray structure of the outer portion of the complex - termed the O layer - into the refined CryoEM structure, structural detail of most densities in the latter can be identified. More interestingly, a comparison of the refined full-length complex with a complex missing the N-terminal region of the TraF subunit generated important new information about the portion of the complex located within and near the bacterial inner membrane. Besides predictions from the structure comparisons, the authors provide additional evidence with nanobody binding that the N-terminus of the TraO subunit is positioned along the periphery of the TraF subunit. The resulting structure is completely novel among characterized transporter complexes, and generates testable hypotheses for future structure-function assignments. Overall, I have only a few minor comments for the authors to consider, in the interest of improving an already outstanding manuscript:

1. Regarding the composition of the cap, the two alpha helices of TraF seem to project inwardly leaving substantial density in the CryoEM structure unidentified. The authors speculate this is composed of the fatty acid moieties of the TraN lipoprotein. With this arrangement, however, the alpha helices should embed into but not across the outer membrane, in contrast to previous predictions by this group. Is it still thought that the TraN helices form the outer membrane pore? Also, could the unidentified density be due to a projection from the TraO subunit? Previous work by this group suggested that a C-terminal domain of TraO subunit might extend across the outer membrane.

2. Fig. 2. Suggest changing the color and thickness of the circles/ovals highlighting the regions of interest in panels C and F to thin/black. Especially in panel F, the current format blocks one from seeing the continuity of this domain along the inner wall of the chamber.

3. Fig. 2C. The middle platform is a very interesting new feature of this structure and it's unfortunate that the authors couldn't identify its content. It's of course appealing to think this could be involved in regulating the channel gating, but it also seems important to rule out the possibility that its an artifact found only in the elastase complex. Are there any other data that can bear on the biological significance of this structure?

Fig. 3e. The color coding is different in the legend vs the figure. Also, in the legend, the alpha-helical domain is referred to as an antennae projection - was this nomenclature used previously in the paper?

Fig. 4. The top of the core complex is closed with density - is this real? If so, it would suggest that the cap is gated with a domain that is currently unidentified. If this exists in the CryoEM structure, why has it not been discussed previously?

Answers to Referee #1:

R1Q1 ... The new data reveal many interesting and novel features in the assembly and as such the manuscript appears to meet the criteria for broad interest required for publication in EMBO J. The major caveat this referee feels is with the attempt to model the VirB9 and VirB10 which seem unnecessary and introduces many potential errors to the paper. The issue arises because neither domain shares any sequence identity with any proteins of known structure and the authors take a fairly naive approach to modelling them.

R1A1. We thank the first referee for his/her positive comments and recognition of the significance of our results. He/She considers them of broad interest and recommends publication in the EMBO Journal. We find the comments both important and interesting. However, we respectfully disagree that modelling of the secondary elements of VirB9 and VirB10 is "unnecessary and introduces many potential errors".

Firstly, it is important to realise that, in the field of electron microscopy, it is standard practice that all structures (particularly novel ones) be analysed and validated using three types of additional experiments: i- labelling of defined parts of the structure with antibodies or by other means, ii- imaging mutants of the structure where some defined parts have been removed in order to locate these parts, and iii- fitting of atomic models into the structure. That is exactly what was done in our study. We have obtained two structures: the full length core complex (FLCC) and a truncated version (the $CC_{elastase}$ complex) where the N-terminal domain of TraF/VirB10 was selectively removed); we have used labelling of the FLCC with NBs to locate the N-terminal domain of TraO/VirB9; the sub-nanometre details were confirmed by fitting of atomic models for the N-terminal domains of both proteins. Therefore, our results and analysis abide by the most exacting modern standards currently applied for a state-of-the-art investigation of a cryo-EM structure.

Secondly, modern secondary structure and three-dimensional structure prediction programmes have become very reliable. Thus, modelling and fitting of predictive atomic models (the only option in the absence of experimentally-derived models) provides not only an important validation tool for the interpretation of cryo-EM maps, but also, conversely, the quality of the fit provides an important evaluation of the predicted models. On both issues, we were able to generate reliable models for most of the IL parts, both the modelling and fitting scores being very high, indicating high confidence in the prediction and the fitting.

To convince the reviewers of the reliability of our modelling and fitting results, we have now expanded the old Supplementary Fig. 2 (now Supplementary Fig. 3). We now provide scores and cross-correlation coefficients in both old Supplementary Figs. 2 and 3 (now Supplementary Figs. 3 and 4). In the discussion, we have added the following: "*In this report, we have re-analysed the FLCC at higher resolution and compared this more detailed structure to that of another complex where a part of the IL, TraF/VirB10$_{NT}$, was selectively removed. Using immuno-labelling selectively targeting the remaining component of the IL, TraO/VirB9$_{NT}$, we unambiguously located and identified this part of the IL. Finally, our analysis was confirmed by fitting of predicted atomic structures into cryo-EM density maps. Given the high reliability of modern three-dimensional structure prediction programmes, fitting of predictive atomic models (the only option in the absence of experimentally-derived models) provides an important validation tool for the interpretation of cryo-EM maps. Conversely, the quality of the fit provides an important evaluation of the predicted models. On both issues, we were able to generate reliable models for most of the IL parts*."

R1Q2. For VirB9 all that is shown is that the domain they identify using Ab labelling and by the differences between the xtal structure and the elastase treated map is of the correct volume for this amount of protein. To demonstrate that they don't really have to introduce the potential for error that comes with de novo modelling of the protein structure.

R1A2. As pointed out in our response to this Referee's first comment, modelling of protein structures and their fitting into electron density maps are essential validation tools. TraO/VirB9 has an approximate length of 300 amino acids, and the structure of the C-terminal fragment containing the last 134 residues has been solved by X-ray crystallography. The very first 23 amino acids of the TraO/VirB9 N-terminal domain (1- 159 aa) correspond to a predicted signal sequence that targets TraO/VirB9 to the periplasm (Petersen TN, *et al.*, *Nature Methods*, 8:785-786, 2011), and therefore the signal peptide is not present in the CC, as this part of the protein is removed during the translocation of VirB9 across the inner membrane. We therefore focused our *in silico* analysis on the

sequence comprising residues 24 to 159. The analysis has revealed two well defined regions: one from 24 to 135 aa, and the second from 136 to 159 aa. The second region does not show any clear secondary structural elements but the first region has an unambiguous beta-sheet organisation that was identified by several secondary structure prediction algorithms. All tertiary structure prediction programmes we used consistently predicted a beta-sandwich structure that could easily be fitted into the experimentally-derived electron density maps. The fitting of the pseudo atomic model had to satisfy two important additional constraints: the location of its C-terminus had to be in close proximity to the O-layer, and no steric clashes should arise from 14-fold symmetrisation. Our model was next refined using flexible fitting into the experimentally-derived EM density maps. The final model fitted the cryo-EM map with an excellent cross-correlation coefficient of 0.82. Thus, our model of TraO/VirB9$_{NT}$ satisfies the most stringent requirements and should be presented. Further details, notably the cross-correlation coefficients have been added.

R1Q3. Their modelling of the N-terminal helices of B10 must correspond to the difference between the elastase treated and full length maps. These attempts to model the structure within the density don't really add anything - the difference between the two maps clearly suggests which density is likely to be the domains and the modelling simply adds potential for error. This referee would suggest removing these speculative models and publishing the simpler interpretation of the data.

R1A3. We do not think that the proposed modelling of the N-terminal helices of VirB10 should be removed: modelling of secondary structures has become a standard requirement in the EM field to validate EM structures and to verify consistency between EM maps and hypothetical atomic models of the proteins, especially in the absence of pre-existing crystal or NMR structures.

The FLCC EM structure at higher resolution reveals the presence of inner columns that form the inner part of the I-layer. The difference map demonstrates that these inner structures are composed of TraF/VirB10$_{NT}$. Typically columns of densities in EM maps correspond to alpha-helices. To examine the possibility that the columns of electron density we observe inside the core complex might be contributed by alpha-helices in TraF/VirB10 (1-170 aa), we performed an *in silico* analysis of the sequence using the same secondary structure prediction programmes we used for TraO/VirB9. The major part of the TraF/VirB10$_{NT}$ was shown to be unstructured and the three dimensional structure modelling of the complete N-terminal domain was not possible. However, all algorithms reliably predicted the presence of three helical peptides that were compatible with the diameter and the length of the observed columns in the cryo-EM map. Thus, the modelling provides a useful validation of our EM structure. However, we agree that there is little structural information regarding the unstructured sequences between these alpha-helical elements and this has been acknowledged in our manuscript.

Minor Points

R1Q4. It would be good for the statement about hand determination to be incorporated in the main manuscript and to be backed up by the correlations in the two hands that allowed determination of hand. In general the correlations between the structures fitted and the maps also appear to be rather low c.f. other papers. Some comment from the authors on why this is so would also be useful.

R1A4. The requested information has been added. We have incorporated details of our analysis in a new supplementary figure (Supplementary Fig. 2) and a statement has been added in the supplementary info text.

R1Q5. Fig 1. - if the masses of the proteolytic fragments have solely been determined by their position in the SEC trace this should be stated as this is a less than optimal way to determine mass. Can the authors be confident that only the N-termini have been truncated?

R1A5. We agree that SEC does not provide an accurate mass but we observed that the SEC trace was shifted right compared to the FLCC and left compared to the OL complex, indicating that the MW of the CC$_{elastase}$ complex was intermediate between the two previously-characterised complexes. A more accurate MW was derived when the elastase-digested complex was analysed by SDS-PAGE. We observed that the two main bands corresponding to TraN/VirB7 and TraO/VirB9 remain unaffected and therefore these two proteins are not digested by elastase. However TraF/VirB10 is affected. A Western-blot against the Strep-tag present at the C-terminus of TraF/VirB10 clearly confirms that this tag is present in the band corresponding to the elastase-digested TraF/VirB10 (data not shown) and therefore there are no additional cuts at the C-terminus of TraF/VirB10 in the elastase-digested CC. Thus, N-terminal sequencing of this band yields its exact composition, allowing us to derive a more accurate MW for the CC$_{elastase}$ complex. We have

now added on page 4: “the integrity of the C-terminus of TraF/VirB10$_{CT}$ was confirmed by Western blotting analysis of the Strep-tag present at the C-terminus of the TraF/VirB10 construct”.

R1Q5.Why were the nanobodies added at a sub-stoichiometric ratio to the complex? If added at a higher ratio then reconstruction of a nanobody-core-complex assembly would be possible and would reveal more information about the location of the epitope.

R1A5. We respectfully disagree. Increasing the concentration of NBs to a higher ratio of NB versus CC leads to aggregation. Indeed, there are 14 sites for NB binding per core complex. When those binding sites are saturated with NBs, addition of the anti-His tag antibodies (directed against the His-tag present on the NBs) leads to the anti-His antibodies cross-reacting with two or more adjacent NBs and thus triggers aggregation of the CC:NB:anti-His complex, thereby preventing structure analysis by EM. Formation of such aggregates at high NB:CC molar ratio was easily monitored using gel filtration. For ratios of NB versus CC above 7:1, we saw a clear increase in the void peak, indicating aggregation. Only by using a 7:1 or lower ratios, were we able to obtain non aggregative preparations of the CC:NB:anti-His complex and these were used for EM visualisation.

R1Q5. Typographical changes: page 3, line 8 swap form and likely to read - ....likely form an inner; page 7, line 12 up from bottom replace is with are .... there ARE AN extra

R1A5. We have corrected the typographical errors.

<u>Answers to Referee #2:</u>

We thank the second Referee for his/her appreciation of our EM structural analysis. The Referee agrees that modelling and fitting of atomic models into the EM maps is useful and provides the first structural characterisation of the inner layer.

R2Q1 The main drawback of the manuscript is that it lacks a principal central conclusion.

R2A1. We respectfully but strongly disagree with this statement. For the first time, we were able to locate and trace the entire sequence of VirB10 in the EM structure of the core complex. Our new structures clearly demonstrate how TraF/VirB10 traverses the whole length of the FLCC complex. They provide for the first time the structural basis of VirB10 function as an energy transducing protein powering the process of substrate translocation. By having mapped out the entire VirB10 structure in the context of a fully-assembled core complex, we now know which structural path the energy transduction process must follow to effect conformational changes during substrate translocation. This is definitely a breakthrough in the field. The structures have also revealed several new features such as the presence of the middle platform that can be a point of TraO/VirB9 and/or TraF/VirB10 interaction with the substrate.

However, we realise that these important conclusions might have not been sufficiently emphasized in the first version of our manuscript. In this revised version, we have rewritten the first few paragraphs of the discussion and we believe that the new version now provides clearer conclusions as to the importance of our new structures.

R2Q2 In extending the resolution of the EM reconstructions, the findings are incremental on previously published work, but provide relatively little new and important information.

R2A2. Here again we respectfully but strongly disagree with the Referee. As the Referee has acknowledged, little was known on the organisation of the inner layer of the core complex. Without dramatic improvements in the resolution of the structures it would have been impossible to identify the positions and structural features of TraO/VirB9 and TraF/VirB10 within the IL of the CC. As a matter of fact, the structure of the IL presented here is just as important as the structure of the OL we published in 2009 in a full article in Nature. The structures presented here provide an unprecedented level of details on the IL: before the present results, we had no idea how the IL was organised and where its components were located and how they were organised. The structures presented here provide fundamental insights into IL organisation and structure and is therefore NOT an incremental finding but indeed a breakthrough.

R2Q3. The abstract refers to 'new structural insights'- and there are certainly some- but their implications for our understanding of the mechanism of T4S are harder to identify. The most interesting conclusion is the central role which VirB10 plays as a 'signal transmitter' between the

ATPases and other components in the assembly. This topic is addressed in the Discussion but is incomplete without additional experimental evidence to elaborate and substantiate this claim. The manuscript also contained a rather high proportion of typographical and grammatical errors, which limited its fluency.

R2A3. Here again, we disagree. Before this work, the mechanism by which VirB10 could act as a transmitter/transducer of ATP-driven conformational changes was very limited because we had no knowledge of the structure of VirB10 in the IL. The work presented here provides the first experimental evidence that VirB10 lines the interior of the core complex throughout the entire length of the cell envelope. This is an essential piece of information and a crucial step towards the elucidation of the mechanism by which VirB10 senses conformational changes in the IM ATPses.

R2Q4. Title: 'Subnanometer resolution structure…' would be more accurate

R2A4. We have changed the title to : "Structure of a bacterial type IV secretion core complex at subnanometer resolution"

R2Q5. Modelling of TraO and TraF N-termini: it is difficult for the reader to gauge how reliable these models are. How do the authors know that they are sufficiently reliable to be used for rigid body fitting? Some additional experimental data would help- eg CD spectra- if the recombinant domains are tractable to study.

R2A5. The first Referee raised similar concerns (see comments R1A1). We now provide additional information and details on the procedure we followed to model and fit the various structural models. The discussion has been amended accordingly.

R2Q6 There is no figure of the final predicted TraO/VirB9NT structure from I-TASSER which was subsequently used for fitting; it would be useful to include this in the Supplementary information, at least.

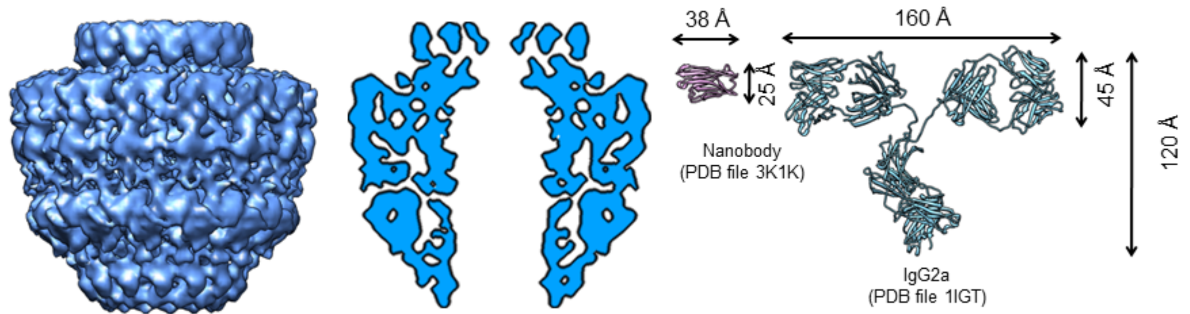
R2A6. We thank the Referee for this advice and the requested figure is now added as Supplementary Fig. 4.

R2Q7. The three predicted helices from TraF/VirB10$_{NT}$ only make up a small proportion of the total mass (see Supplementary Figure 2b). How can the authors be sure that these three helices are indeed responsible for the density features they have attributed to them?

R2A7. The cryo-EM map of the FLCC has 14 clearly defined columns that form the inner part of the I-layer and which were not revealed previously. To understand which part of TraF/VirB10$_{NT}$ this column-like densities might correspond to, we have carried out an *in silico* analysis of secondary structures in TraF/VirB10$_{NT}$. Three helical regions were reliably identified using three programs to predict secondary elements in TraF/VirB10$_{NT}$. The modelling/fitting of the TraF/VirB10$_{NT}$ demonstrated the consistency between the length and diameter of the predicted helices and that of the observed columns of densities. See our more detailed answer to comment R1A3 of the first Referee and modified Supplementary Fig. 3b.

R2Q8. Nanobody labelling: is the resolution of this experiment sufficient to distinguish, as the authors conclude at the end of the second paragraph, that TraO/VirB9NT is located 'on the outer part of the IL'? How would the results differ if it were located on the inner part instead? The difficulty here is that the labelling complex- a nanobody bound to an anti-His antibody- is much larger than the labelled domain itself.

R2A8. The labelling experiments were performed with the FLCC, which exhibits a 55 Å diameter opening at the cytoplasmic base of the core complex. This opening is big enough to allow the access of nanobodies inside the IL (the nanobody dimensions are 38 Å x 25 Å). However the complex NB:Anti-His is much larger: as demonstrated in the figure below, the size of the combined NB:Anti-His complex is 198 Å x 120 Å, which is too large to access the inner chamber of the IL. In the figure below (which we have not incorporated in the revised version of our manuscript), we show the three components FLCC, NB and IgG antibody at the same scale to clarify this point.



R2Q9. The labelling of the bands in the gel in Figure 1b) needs to be clearer- it's not clear which band is TraN. Are the other bands contaminants?

R2A9. We have modified Fig. 1b in order to make the labelling of the bands clearer. The two fainter minor bands are not contaminants but correspond to TraO, a fact that we have established by performing N-terminal sequencing. We now present a new Fig. 1b with clearer labels for the two main bands and for the two minority by-products of elastase digestion.

<u>Minor issues</u>.

We are very grateful to the Referee for proofing of our writing.

1. P3 line 8 'likely form an inner membrane subcomplex.'

-> corrected

2. P3 line 9 'The extracellular pilus is composed of major and minor pilins, termed VirB2 and VirB5, respectively.'

-> corrected

3. P3 line 14 'Because of its central position, the complex participates actively in T4S substrate transfer through the bacterial envelope.' -> corrected

4. P4, five lines from bottom '...very close to the measured molecular weight indicated above.'

-> corrected

5. P7 line 22 'Although there are an extra 9 amino acids present at the N-terminus of TraF/VirB10CT in..'

-> corrected

6. P8 line 2 'These might be attributed to the acetyl moiety of VirB7'- I think the authors mean the lipid moiety of VirB7 (which could contribute significant density)?

The additional density could correspond to both the lipid moieties of the core complex and the detergent molecules associated with them. This has been clarified in the manuscript.

7. P8 fourth line from bottom 'However, the remaining density is not sufficiently big...' This sentence is rather clumsily phrased- I think the authors mean that the volume determined from the density at a particular sigma value is insufficient to accommodate the linker polypeptide between the helices (which, I agree, seems reasonable).

-> corrected

8. P9 line 8 Why are the nanobodies named '#CA4271' on this line, but a different naming convention seems to be used further down 'NBCA4271' (line 12)? There is further potential for confusion in the next paragraph, where the particular nanobody is simply referred to as 'NB'. A consistent nomenclature is needed.

-> corrected

9. P12 line 13 '....the solution was gel filtered as described previously..'

-> corrected

10. P14 line 2 '...according to the 0.5 criterion level...'

-> corrected

11. P14 line 3 'To compare two structures, levels of low frequencies and their resolution in the structures were adjusted to 12.4 Å.' This sentence is rather clumsily phrased- I think the authors mean that they effectively filtered the resolution of the CCelastase structure to 12.4A (i.e. that of FLCC).

–> corrected

12. P14 line 8 'The sequences of TraO/VirB9NT....'

-> corrected

13. Figure 1c) what column was used- it is not clear from the M&M. It is good practice to insert the elution positions of the mass standards used in calibration.

-> corrected

Answers to Referee #3:

This group has pioneered studies aimed at solving the structure of these complex membrane machines and here they have added important new elements to this structural detail. Specifically, they solved a portion of the core complex at 8.5 angstrom resolution and the entire complex at 12 angstroms. By fitting the X-ray structure of the outer portion of the complex - termed the O layer - into the refined CryoEM structure, structural detail of most densities in the latter can be identified. More interestingly, a comparison of the refined full-length complex with a complex missing the N-terminal region of the TraF subunit generated important new information about the portion of the complex located within and near the bacterial inner membrane. Besides predictions from the structure comparisons, the authors provide additional evidence with nanobody binding that the N-terminus of the TraO subunit is positioned along the periphery of the TraF subunit. The resulting structure is completely novel among characterized transporter complexes, and generates testable hypotheses for future structure-function assignments.

We thank the Referee for his/her high assessment of our study.

I have only a few minor comments for the authors to consider, in the interest of improving an already outstanding manuscript:

R3Q1. Regarding the composition of the cap, the two alpha helices of TraF seem to project inwardly leaving substantial density in the CryoEM structure unidentified. The authors speculate this is composed of the fatty acid moieties of the TraN lipoprotein. With this arrangement, however, the alpha helices should embed into but not across the outer membrane, in contrast to previous predictions by this group. Is it still thought that the TraN helices form the outer membrane pore? Also, could the unidentified density be due to a projection from the TraO subunit? Previous work by this group suggested that a C-terminal domain of TraO subunit might extend across the outer membrane.

R3A1. In the FLCC, the two alpha helices of TraF do not project inwardly. There is additional density and this density is interpreted as being part of the TraN lipoprotein. So the Referee's interpretation is not quite correct. In Chandran et al (Nature 2009) we unambiguously demonstrated that the two alpha helices project through the membrane, since a FLAG tag inserted between the two helices is exposed at the surface of bacteria. This interpretation remains unchanged in light of the new structures we show here.

R3Q2. Fig. 2. Suggest changing the colour and thickness of the circles/ovals highlighting the regions of interest in panels C and F to thin/black. Especially in panel F, the current format blocks one from seeing the continuity of this domain along the inner wall of the chamber.

R3A2. We have modified Fig. 2 according to the Referee's request.

© EMBO

R3Q3. Fig. 2C. The middle platform is a very interesting new feature of this structure and it's unfortunate that the authors couldn't identify its content. It's of course appealing to think this could be involved in regulating the channel gating, but it also seems important to rule out the possibility that it's an artefact found only in the elastase complex. Are there any other data that can bear on the biological significance of this structure?

R3A3. The middle platform is a new feature particularly visible in the $CC_{elastase}$ complex and thus might have become prominent in this complex as a result of elastase action. We speculate that it may be related to the middle part of the TraO/VirB9 (136 to 159 aa) a structure which we were not able to predict. This part of VirB9 might adopt a different conformations with and without TraF/$VirB10_{NT.}$ We are sure that this is not an artefact since the densities are well defined and it may be that this part in the FLCC structure is moved slightly down to form an interface with the central region of TraF/VirB10. However this hypothesis will need to be confirmed in future investigations.

R3Q4. Fig. 3e. The colour coding is different in the legend vs the figure. Also, in the legend, the alpha-helical domain is referred to as an antennae projection - was this nomenclature used previously in the paper?

R3A4. We have modified Fig. 3e according to the Referee's request.

R3Q5. Fig. 4. The top of the core complex is closed with density - is this real? If so, it would suggest that the cap is gated with a domain that is currently unidentified. If this exists in the CryoEM structure, why has it not been discussed previously?

R3A5. The top of the CC does contain an opening to the extracellular media of an approximate diameter of 20 Å with a 10 Å constriction underneath, as shown in Figs. 2e and 2f, and therefore it is not closed. However, in Fig. 4a, in order to better show the proposed atomic model fitting inside the electron density, we decided to show a central section with an increased slab thickness compared to the one shown in Fig. 2f: because of the increased slab thickness and positioning, the rear wall of the 10 Å constriction is now visible, giving the misleading appearance that the complex is closed. However, had we opted for a slab thickness similar to the one used in Figs. 2e and 2f, the atomic models would not have been clearly visible. To clarify the figure, the rear wall of the constriction is now indicated and in a lighter grey colour.

Acceptance 19 February 2013

Thank you very much for your patience while your study has been evaluated. We have now heard back from two of the original referees, whom we asked to review your manuscript and I am pleased to inform you that both reviewers agree on the suitability of your paper for publication in The EMBO Journal. While I copy below their reports for your information, I consider that no further action is required and your manuscript has therefore been accepted for publication.

Please note the suggestion of referee #2 regarding the link between the EM and the PDB databanks. We require for coordinates of structures to be published in the corresponding public databases, which you have done in this case. Further linking of data among databases is not required for publication, but I personally think it is an excellent idea.

Thank you for your contribution to The EMBO Journal and congratulations on a successful publication.

REFEREE REPORTS:

Referee 1

The manuscript has been improved in response to the initial points raised. This referee is still not entirely convinced by the modelling but the manuscript better discusses the advantages and disadvantages.

Referee 2

The manuscript describes the structure of a type IV secretion complex obtained by electron microscopy. The importance of the type IV secretion system is well established, and recent structural work has been spectacularly successful in contributing to our current knowledge of the way in which DNA and protein is transported. This work is essentially an extension of the EM structure of the complex published by Fronzes et al (Science, 2009). The key difference is an improvement in resolution, from 15 to 12.4A for the whole complex, and to 8.5A for a sub-structure obtained by protease digestion. Overall, the work has been executed to a high standard: the data are of good quality and the methods employed are state-of-the-art for this type of investigation.

The manuscript is much improved after revision. My main criticism was that the paper lacked an obvious central conclusion. The authors' responses have helped to clarify this point: the major conclusion is the identification of the location of the VirB10 component, demonstration that it spans the periplasm and therefore is capable of forming a link between the cytoplasmic ATPases and secretion events within the complex chamber. This is complemented by verification of the location of the N-terminus of the VirB9 component to the outer sheath of the CC. I still think that these main conclusions need to be incorporated into the abstract in some way: even in its current form, it still gives little information about the biological significance of the work.

The work is not without its drawbacks, however. The structures of the N-terminal segments of VirB9 and VirB10 have been generated by modelling processes. In spite of what the authors say, confidence in the accuracy of these structures is inevitably diminished, compared with experimental structure determination of these components by X-ray crystallography or NMR. Probably the least convincing part of the docked model for the complex is the identification of the three predicted helices from the VirB10 N-terminus, as shown in Fig 4. If I read the manuscript correctly, these are docked into an electron density map which is at 12.4A resolution i.e. it is part of the structure which is removed by elastase digestion (Fig 1). This would generally be regarded as low resolution for the detection of alpha helices, and is not helped by the fact that a large part of this section of VirB10 is predicted to be disordered (Supp Fig 3b). Apart from the antibody labelling experiment, it is also worth pointing out that the electron microscopy and modelling work is not supported by much biochemical data or, indeed, any in vivo data. This rather weakens the authors' evidence base for the central message of the manuscript.

Minor points

1. Are the authors proposing to deposit the coordinates of their docked model in the PDB? They will probably be aware that it is now possible to link from a map deposition in the EM data bank to an atomic model in the PDB.
2. p21, legend to Fig 4. The text refers to 'tentative docking..... with high confidence...'. If the docking is indeed 'tentative', I would recommend that it is omitted from the manuscript. As explained above, this is probably the weakest part of the docking process.