THE EMBO JOURNAL

Cryo-EM structure of native human uromodulin, a zona pellucida module polymer

Alena Stsiapanava, Chenrui Xu, Martina Brunati, Sara Zamora-Caballero, Céline Schaeffer, Marcel Bokhove, Ling Han, Hans Hebert, Marta Carroni, Shigeki Yasumasu, Luca Rampoldi, Bin Wu, and Luca Jovine
DOI: 10.15252/embj.2020106807

Corresponding author(s): Luca Jovine (luca.jovine@ki.se) , Bin Wu (wubin@ntu.edu.sg)

Review Timeline:

Submission Date:	16th Sep 20
Editorial Decision:	25th Sep 20
Revision Received:	7th Oct 20
Accepted:	8th Oct 20

Editor: Hartmut Vodermaier

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. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Please note that this manuscript was previously reviewed at another journal. Since the original reviews are not subject to *The EMBO Journal's* transparent review process policy, these initial reports and author response to them cannot be published here.

1st Editorial Decision 25th Sep 2020

Thank you again for submitting your manuscript together with previous reports and responses from another journal for our editorial consideration. As discussed earlier, we have now consulted with a trusted cryo-EM reviewer of our own journal regarding the methodological issues raised by the previous referees. Our arbitrator had access to both the latest (revised) manuscript and to the original comments and your response to them, as well as to the raw maps and model coordinates that you provided. As you will see from our expert's comments copied below, the arbitrating referee considers the density maps well sufficient for building, refining and interpreting the presented atomic models. At the same time, our referee is not convinced that the density figures as currently presented do full justice to this, and recommends (as detailed in the comments below) to revise the figures particularly in this regard. Furthermore, the arbitrator shares one specific interpretational concern of original referee 3 regarding Figure 6, and recommends that this should be presented much more cautiously (possibly also shortening this part of the currently very extensive discussion section).

In light of these overall supportive comments, we shall be happy to publish this work in The EMBO Journal, once the specific issues emphasized by our arbitrating referee have been incorporated into a final version of the manuscript. Addressing also the following important editorial points at this stage should greatly facilitate expedited consideration and processing after resubmission:

Referee #1:

The general concerns about referees 1 and 3 were about the interpretability or "buildability" of the map, in particular the so-called "density modified map". Therefore, having access to the maps and built model is indeed very helpful. Having inspected the densities, it is clear that one can properly refine an atomic model within the map. One additional consideration: as pointed out by referee 3 point 1 that the authors only docked an X-ray model. This is partially true as they needed to carry out modifications and change the connectivities in the assembled filament. In support of the author's interpretation, the fact that a large part of the model is from the X-ray structure gives better confidence in the presented model even if certain parts of the cryo-EM map are not as well resolved. This still does not diminish the presented work as it was now successfully determined in its assembled filamentous context. The authors used the relevant parts of the existing X-ray structure to restrain their building and refinement, which is good practice. The validation statistics are all within reason for that resolution. Interestingly, as you probably know in late August, another group also published the structure of the uromodulin filament (Stanisich et al., 2020). The structures agree very well (RMSD <1 Å at this resolution fine) and this fact in itself give strong confidence that Stsiapanava also present a reliable model from the density.

My recommendation is that authors improve the figures, with a particular emphasis on convincing the reader that this map has a lot of details that are interpretable on a side-chain level, as it is commonly done for cryo-EM papers, e.g. separation of beta-strands, large bulky side chain density with an atomic model and superimposed density. This is very much missing and at the moment the reader is left wondering whether the quality of the map justifies the conclusions.

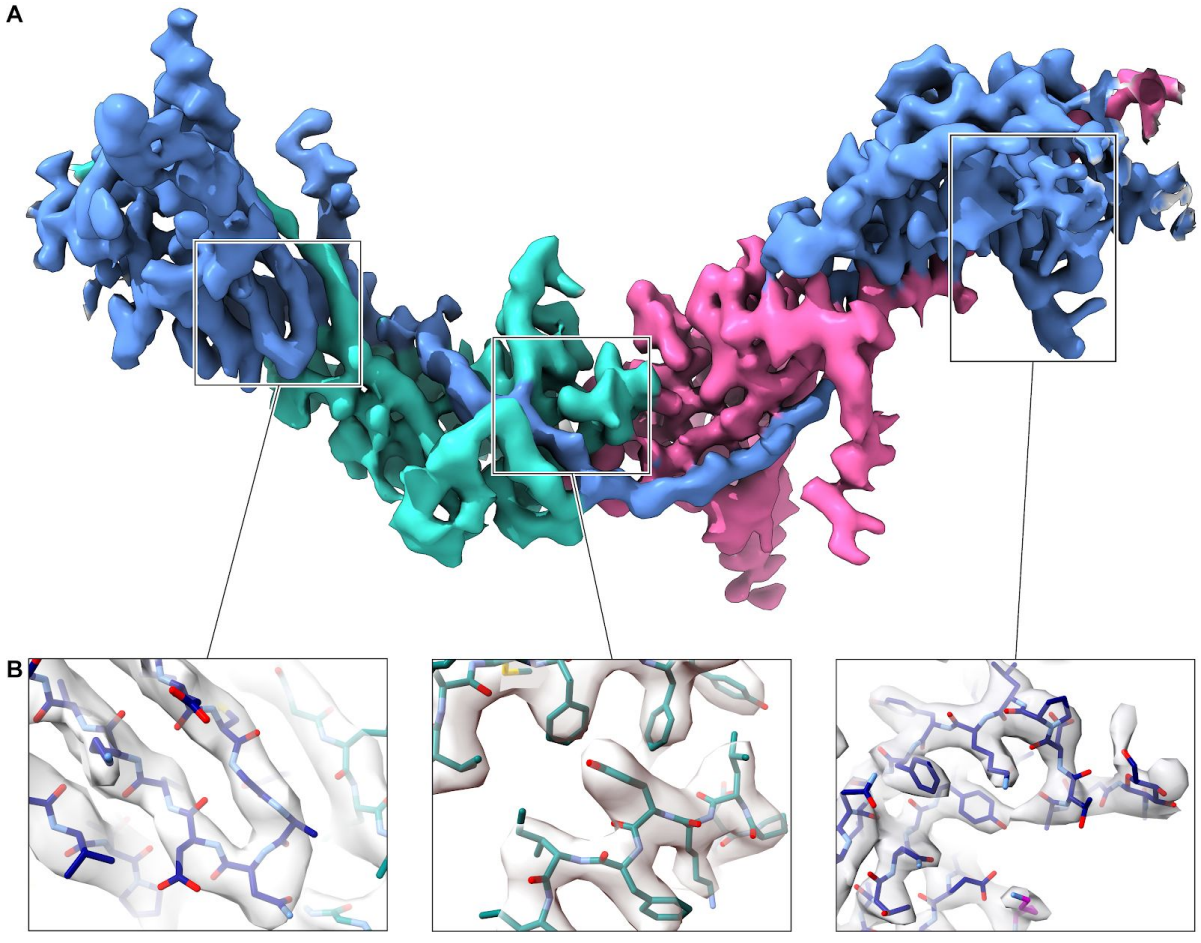
I share one additional concern about the model interpretation based on negative stain images in Figure 6 (referee 3, point 10). When left as is, it is indeed far-fetched and cannot be easily supported unless they resolve the structure. The authors should exercise more caution on this and revert to a more schematic interpretation as in Figure 7.

Response to the Referee's Comments

We thank the Referee for carefully analyzing our maps, with the conclusion that the associated atomic model has been reliably built. Here are answers to the two specific comments made by the Referee:

My recommendation is that authors improve the figures, with a particular emphasis on convincing the reader that this map has a lot of details that are interpretable on a side-chain level, as it is commonly done for cryo-EM papers, e.g. separation of beta-strands, large bulky side chain density with an atomic model and superimposed density. This is very much missing and at the moment the reader is left wondering whether the quality of the map justifies the conclusions.

We see the point of the Referee and are grateful for this suggestion. However, because the density regions shown in Fig 1 and associated Movies EV1-4 were specifically chosen to match certain areas of the structure described in the text, we thought that it would not be optimal to simply replace them. Therefore, we decided to comply with the Referee's request by generating an additional Expanded View Figure (Fig EV2) connected with Fig 1. The new figure specifically highlights the quality of the 3.4 Å map, both globally (top panel) and in detail (bottom panels):



The bottom panels provide examples of the features mentioned by the Reviewer by showing separation of β-strands (bottom left panel), as well as details of different map regions where large bulky side chains are found (bottom center and right panels).

I share one additional concern about the model interpretation based on negative stain images in Figure 6 (referee 3, point 10). When left as is, it is indeed far-fetched and cannot be easily supported unless they resolve the structure. The authors should exercise more caution on this and revert to a more schematic interpretation as in Figure 7.

We understand the concern of the Referee but, if possible, we would prefer not to replace the current UMOD sheet model representation in Fig 6 (assembled by combining multiple copies of the experimental UMOD filament map) with a more schematic depiction. This is for two different reasons. First, we think that it is important to show that the experimental map of the individual filament is in fact compatible with the proposed sheet architecture. Second, replacing the current map-based representation in Fig 6B with a schematic one (such as the one that we included in the Synopsis figure) would generate an additional issue. This is because the model depicted in Fig 6B provides an explanation for the fuzzy density that we experimentally observed in the black areas between the sheet filaments shown in the bottom panel of Figure 6A, by suggesting that this corresponds to the long glycan chains attached to N396 and N513 (arrows in Fig 6C). It is hard to envisage how to convincingly make this point, without showing a model that is based on an actual experimental map with density for such sugars. At the same time, we think that it would be even more confusing to opt for a hybrid solution where we replaced Fig 6B with a scheme but left Fig 6C (which makes an important point in favour of the model) as it is.

Since the worry here is that readers may mistakenly interpret what is shown in Figs 6B and C as an actual experimental map of the sheet (as opposed to a model of the latter, made up by assembling copies of the experimental map of a single filament), we have addressed the concern of the Referee by making the nature of the model and its limitations clearer in the legend of Fig 6. Specifically, we have modified the title of the legend so that it starts with "Modelling" ("Modelling of UMOD sheet architecture and ZP module filament binding sites for UPEC and sperm.") and edited the legend of Fig 6B so that it begins as follows:

UMOD sheet model, generated by antiparallel juxtaposition of multiple copies of the $UMOD_{fl}$ composite map, according to the 2D information shown in panel A. Due to lack of information in the direction perpendicular to the plane of the sheets, the exact position of adjacent filaments relative to this plane is unknown; for simplicity, the depicted model has been generated by assuming that the cores of the filaments making up the sheet lie on the same plane.

In our view, keeping the representation as it is but pointing out the above in the figure legend is a good compromise that makes the most of the experimental information that we have on the system, while at the same time making it clear to the reader that this information is incomplete in one direction and, thus, should not be confused with a real 3D model (although it

is most likely quite accurate because, in order to make a sheet, filaments must still be close enough in the perpendicular direction so that they can physically interact with each other).

1st Revision - Editorial Decision 8th Oct 2020

Thank you for submitting your final revised manuscript for our consideration. I have now assessed your responses and modifications, and I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Luca Jovine, Bin Wu
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2020-106807R

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- ➔ the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ➔ figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- ➔ graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- ➔ if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be iustified
- ➔ Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- ➔ a specification of the experimental system investigated (eg cell line, species name).
- ➔ the assay(s) and method(s) used to carry out the reported observations and measurements
- ➔ an explicit mention of the biological and chemical entity(ies) that are being measured.
- ➔ an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- ➔ the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- ➔ a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- ➔ a statement of how many times the experiment shown was independently replicated in the laboratory.
- ➔ definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	No statistical methods were used to predetermine sample size. Representative micrographs for Figs 1A-E, 5I, 6A, EV1B,D were chosen by eye.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Two independent SEC-MALS experiments were carried out for each protein. Each experiment included independent expression, purification and measurement of ZI-1,2 or ZI-3, rather than just a repeated analysis of the same material. All measurements of each protein agreed but, for clarity, the results of a single experiment are presented in Appendix Fig S3.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	NA

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Mouse purified anti-HA.11 Epitope Tag (Cat# 901502, Biolegend, lot number B242907). Rabbit affinity isolated anti-Flag antibody (F7425, Sigma, lot number 086M4803V). Rat monoclonal antibody Anti-HA High Affinity (clone 3F10) (Cat# 11867423001, Roche, lot number10145700). Mouse monoclonal antibody anti-KDEL (10C3) (ADI-SPA-827, Enzo, lot number 09011541).
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	HEK293T: laboratory of Prof. A. Radu Aricescu (University of Oxford, UK; now at the MRC Laboratory of Molecular Biology, Cambridge, UK) (PMID 3031469). HEK293S GnTi-: ATCC cat. no. CRL-3022 (PMID 12370423). MDCK: ATCC cat. no.CCL-34 (PMID 5918973). Cell lines were authenticated and checked for mycoplasma contamination by the sources described above; additionally, we confirmed that HEK293T, HEK293S and MDCK were mycoplasma-free by using a PCR Mycoplasma Test Kit II (Applichem cat. no.A8994).

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	Fish unfertilized egg coats fragments were provided by Prof. Shigeki Yasumasu (Sophia University, Tokyo), who prepared them from oocytes of Japanese rice fish/medaka (Oryzias latipes) according to a published protocol (Yasumasu et al. 2010).
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	Animal procedures were approved by the ethics committee of Sophia University (approval number 2016-006).
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	Part of this study depended on the analysis of egg coat material obtained from unfertilized oocytes of Japanese rice fish. Because these animals were only used to obtain eggs, apart from the details described above the mentioned guidelines do not apply to this work.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	The human subject is one of the authors of the manuscript (L.J.). No ethical approval was deemed necessary by the human subject's department, as he used his own urine.
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA (please see point 11)
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The "Data availability" section of the manuscript contains accession codes for the cryo-EM maps and atomic model coordinates generated during this study, which were deposited in the Electron Microscopy Data Bank (EMDB)and the Protein Data Bank (PDB), respectively.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	As detailed in "Data availability", the Python scripts used for UMOD filament picking have been deposited in the public repository GitHub.

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA