EMBO reports

# Structural study of UFL1-UFC1 binding uncovers importance of UFL1 N-terminal helix for ufmylation

Sayanika Banerjee, Julia Varga, Manoj Kumar, Guy Zoltsman, Shahar Rotem-Bamberger, Einav Cohen-Kfir, Michail Isupov, Rina Rosenzweig, Ora Schueler-Furman, and Reuven Wiener
**DOI: 10.15252/embr.202356920**

*Corresponding author(s): Ora Schueler-Furman (ora.furman-schueler@mail.huji.ac.il) , Reuven Wiener (reuvenw@ekmd.huji.ac.il)*

## Review Timeline:


| | |
|---|---|
| Submission Date: | 31st Jan 23 |
| Editorial Decision: | 15th Mar 23 |
| Revision Received: | 14th Jun 23 |
| Editorial Decision: | 24th Aug 23 |
| Revision Received: | 18th Sep 23 |
| Editorial Decision: | 19th Oct 23 |
| Revision Received: | 23rd Oct 23 |
| Accepted: | 25th Oct 23 |

*Editor: Martina Rembold*


## 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.)

Dear Ora,

Thank you for the submission of your manuscript to our journal. I apologize for the delay in handling your manuscript, but we have meanwhile received the full set of referee reports that is copied below.

As you will see, while the referees agree that the study is potentially interesting, they also all point out that it requires significant revision before it can be considered for publication here. The major concerns regard the lack of validation of the structural models and the minimal functional data provided. From the referee comments it is clear that a major revision would be required before the study can be considered for publication in our journal. It will be essential to address all concerns from referee 1 and referee 3, to provide further functional and mechanistic insight and to verify the proposed interactions using mutagenesis. Also UFM1 should be modelled into the system, as proposed by referee 1 and referee 3. It will however not be required to expand the work to animal models (referee 2, point 3).

Given the potential interest of your findings, I would like to give you the opportunity to address the concerns and would be willing to consider a revised manuscript with the understanding that the referee concerns must be fully addressed and their suggestions (as detailed above and in their reports) taken on board.

Should you decide to embark on such a revision, acceptance of the manuscript will depend on a positive outcome of a second round of review and I should also remind you that it is EMBO reports policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript.

We realize that it is difficult to revise to a specific deadline. In the interest of protecting the conceptual advance provided by the work, we recommend a revision within 3 months (June 15th). Please discuss the revision progress ahead of this time with me if you require more time to complete the revisions. I can extend the revision duration to 4-5 months.

I am also happy to discuss the revision further via e-mail or a video call, if you wish.

You can either publish the study as a short report or as a full article. For short reports, the revised manuscript should not exceed 27,000 characters (including spaces but excluding materials & methods and references) and 5 main plus 5 expanded view figures. The results and discussion sections must further be combined, which will help to shorten the manuscript text by eliminating some redundancy that is inevitable when discussing the same experiments twice. For a normal article there are no length limitations, but it should have more than 5 main figures and the results and discussion sections must be separate. In both cases, the entire materials and methods must be included in the main manuscript file.

*****IMPORTANT NOTE:
We perform an initial quality control of all revised manuscripts before re-review. Your manuscript will FAIL this control and the handling will be DELAYED if the following APPLIES:

1) A data availability section providing access to data deposited in public databases is missing. If you have not deposited any data, please add a sentence to the data availability section that explains that.

2) Your manuscript contains statistics and error bars based on n=2. Please use scatter blots in these cases. No statistics should be calculated if n=2.

When submitting your revised manuscript, please carefully review the instructions that follow below. Failure to include requested items will delay the evaluation of your revision.*****

When submitting your revised manuscript, we will require:

1) a .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure).
Please download our Figure Preparation Guidelines (figure preparation pdf) from our Author Guidelines pages https://www.embopress.org/page/journal/14693178/authorguide for more info on how to prepare your figures.

3) a .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point responses to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

4) a complete author checklist, which you can download from our author guidelines (). Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

5) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript (). Please find instructions on how to link your ORCID ID to your account in our manuscript tracking system in our Author guidelines
()

6) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online. A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc... in the text and their respective legends should be included in the main text after the legends of regular figures.

- For the figures that you do NOT wish to display as Expanded View figures, they should be bundled together with their legends in a single PDF file called *Appendix*, which should start with a short Table of Content. Appendix figures should be referred to in the main text as: "Appendix Figure S1, Appendix Figure S2" etc. See detailed instructions regarding expanded view here:

- Additional Tables/Datasets should be labeled and referred to as Table EV1, Dataset EV1, etc. Legends have to be provided in a separate tab in case of .xls files. Alternatively, the legend can be supplied as a separate text file (README) and zipped together with the Table/Dataset file.

7) Before submitting your revision, primary datasets (and computer code, where appropriate) produced in this study need to be deposited in an appropriate public database (see < https://www.embopress.org/page/journal/14693178/authorguide#dataavailability>).
Specifically, we would kindly ask you to provide public access to the structural data.

Please remember to provide a reviewer password if the datasets are not yet public.

The accession numbers and database should be listed in a formal "Data Availability " section (placed after Materials & Method) that follows the model below (see also < https://www.embopress.org/page/journal/14693178/authorguide#dataavailability>). Please note that the Data Availability Section is restricted to new primary data that are part of this study.

# Data availability

The datasets (and computer code) produced in this study are available in the following databases:

- RNA-Seq data: Gene Expression Omnibus GSE46843 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46843)
- [data type]: [name of the resource] [accession number/identifier/doi] ([URL or identifiers.org/DATABASE:ACCESSION])

*** Note - All links should resolve to a page where the data can be accessed. ***

8) At EMBO Press we ask authors to provide source data for the main figures. Our source data coordinator will contact you to discuss which figure panels we would need source data for and will also provide you with helpful tips on how to upload and organize the files.

Additional information on source data and instruction on how to label the files are available .

9) The journal requires a statement specifying whether or not authors have competing interests (defined as all potential or actual interests that could be perceived to influence the presentation or interpretation of an article). In case of competing interests, this must be specified in your disclosure statement. Further information: https://www.embopress.org/competing-interests

10) Figure legends and data quantification:
The following points must be specified in each figure legend:

- the name of the statistical test used to generate error bars and P values,
- the number (n) of independent experiments (please specify technical or biological replicates) underlying each data point,
- the nature of the bars and error bars (s.d., s.e.m.)

- If the data are obtained from n {less than or equal to} 5, show the individual data points in addition to the SD or SEM.
- If the data are obtained from n {less than or equal to} 2, use scatter blots showing the individual data points.

Discussion of statistical methodology can be reported in the materials and methods section, but figure legends should contain a basic description of n, P and the test applied.

See also the guidelines for figure legend preparation:
https://www.embopress.org/page/journal/14693178/authorguide#figureformat

- Please also include scale bars in all microscopy images.

11) Our journal encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at .

12) As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

I look forward to seeing a revised form of your manuscript when it is ready and please let me know if you have questions or comments regarding the revision.

Yours sincerely,

Martina Rembold, PhD
Senior Editor
EMBO reports

**********

Referee #1:

This manuscript reports studies aimed at defining interactions among the three enzymes responsible for the post-translational modification known as ufmylation, in which a ubiquitin-like modifier protein is attached to other proteins. Ufmylation is involved in DNA damage response, protein translation, and possibly other cellular processes, but to date little is known about this system. The authors used a combination of structural prediction by the deep learning program AlphaFold2, experimental structural techniques, and biochemistry to focus on the ligase ("E3") component, UFL1, and the conjugating enzyme ("E2") UFC1. The main conclusions of the study are that 1) UFL1 bears no resemblance to any of the known structural classes of Ub E3 ligases, 2) an N-terminal helix of UFL1 binds to the E2, and 3) the E1 and E3 compete for binding to the E2. While the data presented support these conclusions, a lack of validation of the structural models and minimal functional data left this reader feeling unsatisfied. For example, there was no attempt to validate the interactions shown in Fig. 2 through mutagenesis. Also, the two functional assays presented (Fig. 1f,g) are neither explained nor labeled sufficiently to impart any useful information. In the Discussion, the authors use a lot of words in the first paragraph to describe the relationship among the E1, E2, and E3 and seem to suggest that this is somehow novel. Nowhere are we told that the ufmylation process appears to work in much the same way as the ubiquitylation system or the SUMOylation system. For the non-expert, this would seem an important point to make. The manuscript would also benefit from a more thorough description of ufmylation in the Introduction, including the fact that the C-terminus of UFM1 is processed before it can be attached, presumably to a lysine residue (?). Where are the proteins expressed? Overall, the manuscript feels a bit thin and adds to our understanding in a fairly incremental way. Bringing UFM1 itself into the modeling and structural studies could add more insight. For example, can a stable UFC1-UFM1 conjugate be generated? How does this change in the presence of UFL1? Additional detailed comments for the authors are offered below.

Specific Comments.

1. As it is still early days in the use and presentation of AlphaFold generated models, it is important to adopt careful language

when referring to its output and its utility. Early in the manuscript you say that the structural models from AlphaFold can be used to guide design of protein constructs (by removing regions predicted to be disordered). But your results showing the importance of the N-terminal helix are contrary to that notion.

2. It would be interesting to know whether there were any hints in the sequence of DDRGK1 that would have suggested the presence of the winged-helix domain? Are there similar complementary structures from other systems? On what would Alphafold model this complex? Does Alphafold model the fusion protein the same as the two proteins entered separately in AlphaFold_multimer?

3. It's not clear on what the conclusion that the proteins "do not elute as soluble aggregates" in the gel filtration experiment is based. This conclusion requires better documentation. Furthermore, it's surprising that 20 residues changes elution as much as it does. Could this indicate further structural differences?

4. Given that the crystal structure of the WH domain was solved using the AlphaFold model as the starting structure, it doesn't make a lot of sense to do a detailed analysis of the difference between the two. Furthermore, the resolution of 3.1 Å doesn't warrant detailed analysis.

It is stated that "In the crystal structure the last 18 amino acids are flexible and are not detected in the electron density map." That logic is backwards. In the crystal, the last 18 residues are not visible and may therefore be flexible. Or there may be high static heterogeneity.

5. The functional assays as presented, with minimal labeling or description are uninformative. What is the strong band at ~30 kDa? Why is a His-tag used in one assay and a flag-tag used in the other? What is being ufmylated in these assays? Are UFM1 chains expected? Where would they run? In Fig 1g, why do no complex and complex look the same? What are the bands in the no ATP?

6. On page 10, the titration would suggest that UFC1-UFL1 binding is stronger than UFC1-UBA5.
What sequence of UFL1 was added for competition? If the whole protein, what happens if a peptide equivalent to the ~50 residues of UBA5 is added? Again, how might the presence of UFM1 alter these interactions?

Fig4 legend does not indicate titrated ratios, requiring reader to look it up: UBA5 @ 150µM, UFC1 @300µM, DDRGK1-UFL1 @700µM.

Could titrations be done of 15N UFC1 and equivalent peptides of UBA5 vs UFL1 to get relative Kds?

Referee #2:

(1) The manuscript uncovered an interesting, helix-mediated regulatory mechanism in ufmylation: UFL1 N-terminal helix competes with the E1 (UBA5) C-terminal helix in binding to UFC1 and drives ufmylation, and it is an interesting finding using DeepMind/AlphaFold2 predictions and in vitro analysis.
(2) Why did you omit the regulatory role of E3 co-factor LZAP in this mechanism? How do you exclude the effects of LZAP in your system prediction and in vitro analysis? Could you supplement these evidences in your research?
(3)Could you supplement the morphological evidences of cell lines or animal models to support your conclusion?

Referee #3:

In this manuscript, Banerjee et al. focus on the understanding mechanisms of the enzymatic machinery responsible for UFMylation. The major focus is on how the E2 conjugating enzyme UFC1 works with the E3 UFL1 and its cofactor DDRGK1 to promote UFMylation. The authors first used AlphaFold2 modeling to design a UFL1/DDRGK1 fusion construct which was then used to determine the first crystal structure of this complex. The fusion construct was also used as a tool for in vitro UFMylation assays in which it was discovered that the N-terminal helix of UFL1 is important for promoting UFMylation. The authors next demonstrate a direct interaction between the UFL1 N terminal helix and UFC1 and again used AlphaFold2 to gain insights into the molecular details of this interaction. The binding surface for the UFL1 helix on UFC1 was further mapped using NMR and suggest that the UFC1 surface that engages the UFL1 helix overlaps with the same UFC1 surface that engages a helix in the structurally similar interaction between UFC1 and the E1 Uba5. Overall, the importance of UFMylation to biology is becoming increasingly apparent and this is a nice multidisciplinary study that represents a significant conceptual advance in our understanding of how the UFM1 E2 and E3 function together to promote catalysis. With that said, there are a few important issues that should be addressed to make this manuscript an even stronger candidate for publication in EMBO reports:

1) The authors nicely demonstrate that the N-terminal helix of UFL1 mediates an interaction with UFC1 by ITC and NMR and

that the presence of the helix promotes UFMylation in vitro, in the context of the UFL1/DDRGK1 fusion. Can the authors determine whether this interaction simply brings the E2 and E3 in proximity or whether this interaction stimulates the catalytic activity of the E2 using something akin to the UFC1-UFM1 lysine discharge assays? This would provide deeper mechanistic insights into the role of this interaction in UFMylation. It would be interesting to know whether the UFL1 helix alone can promote UFMylation as shown in Fig. 1f.

2) If technically feasible, it would be informative to assess whether a UFC1-UFM1 thioester mimetic has an altered affinity for UFL1/DDRGK1 compared to UFC1 alone. This would provide additional insights into the mechanistic importance of competition between Uba5/UFL1 for UFC1 and whether the UFM1 might contact the E3.

3) It would be nice to expand a bit in the Discussion on how this emerging noncanonical E3 mechanism fits into the broader context of known E3 mechanisms in other Ubl pathways. Might UFL1 hold UFC1-UFM1 thioester intermediate in a 'closed' conformation to promote catalysis akin to other Ubl pathways (could AlphaFold2 provide insights)? As it pertains to the topic of UFL1 and Uba5 competition for UFC1 noted in the first and second paragraphs of the discussion, this is where the UFM1 that is linked via thioester bond to UFC1 may come into play. Note the thioester switch model for E1-E2 transthioesterification in the ubiquitin system (Yuan et al., Nature Comms 2022) and in the Nedd8 system (Huang et al., Nature 2007) and how such a phenomenon may be relevant to the UFM1 pathway. Finally, if the relative concentrations of Uba5, UFC1, and UFL1 in the cell are known it would be nice to comment on how this relates to the observed affinities for these interactions in vitro and the mechanistic implications of E1/E3 competition for E2.

Minor points:
• Overall, the figures could benefit from more detailed labeling, especially Fig. 1f and g.
• Advise against using yellow color for UFL1 as it is very hard to see in print.

Referee #1:

*This manuscript reports studies aimed at defining interactions among the three enzymes responsible for the post-translational modification known as ufmylation, in which a ubiquitin-like modifier protein is attached to other proteins. Ufmylation is involved in DNA damage response, protein translation, and possibly other cellular processes, but to date little is known about this system. The authors used a combination of structural prediction by the deep learning program AlphaFold2, experimental structural techniques, and biochemistry to focus on the ligase ("E3") component, UFL1, and the conjugating enzyme ("E2") UFC1. The main conclusions of the study are that 1) UFL1 bears no resemblance to any of the known structural classes of Ub E3 ligases, 2) an N-terminal helix of UFL1 binds to the E2, and 3) the E1 and E3 compete for binding to the E2. While the data presented support these conclusions, a lack of validation of the structural models and minimal functional data left this reader feeling unsatisfied. For example, there was no attempt to validate the interactions shown in Fig. 2 through mutagenesis.*

We thank the reviewer for his/her supportive and constructive comments. Addressing these has significantly improved the manuscript.

We acknowledge the reviewer's concern regarding the lack of model validation. To address this issue, we performed experiments by generating model-based mutations in both UFL1 and UFC1. Subsequently, we evaluated the impact of these mutations on binding affinity and ufmylation activity (Fig. 2G-I). As anticipated, the mutations localized on the binding surface significantly disrupted both binding (Fig. 2G) and activity (Fig. 2I).

*Also, the two functional assays presented (Fig. 1f,g) are neither explained nor labeled sufficiently to impart any useful information.*

We improved the functional assay in the following ways: 1) for both the fusion construct as well as the complex we detected activity not only by western blot analysis using anti -Myc (current Fig. 1F, H; UFL1 has a Myc tag), but also by using anti-FLAG (current Fig. 1E, G UFM1 has a FLAG tag); 2) In all blots, we have labeled the corresponding bands and their respective identities (current Fig. 1E-H).

*In the Discussion, the authors use a lot of words in the first paragraph to describe the relationship among the E1, E2, and E3 and seem to suggest that this is somehow novel. Nowhere are we told that the ufmylation process appears to work in much the same way as the ubiquitylation system or the SUMOylation system. For the non-expert, this would seem an important point to make.*

In the revised manuscript, we have merged the results and discussion sections, leading to the removal of the mentioned paragraph. However, in the introduction, we emphasize the similarity of UFM1 conjugation to the established processes observed in ubiquitin and other ubiquitin-like proteins.

*The manuscript would also benefit from a more thorough description of ufmylation in the Introduction, including the fact that the C-terminus of UFM1 is processed before it can be attached, presumably to a lysine residue (?). Where are the proteins expressed?*

We have added in the introduction a description about UFM1 maturation, and where UFM1 is expressed.

*Overall, the manuscript feels a bit thin and adds to our understanding in a fairly incremental way. Bringing UFM1 itself into the modeling and structural studies could add more insight.*

We have followed the reviewer's suggestion and incorporated charged UFC1 into our study. We demonstrated that in the presence of DDRGK1, there is preferential binding to the charged UFC1 compared to the corresponding uncharged form. This observation was confirmed through both pulldown experiments and isothermal titration calorimetry (ITC) (Fig. 5C). Specifically, our ITC results revealed a 10-fold lower dissociation constant (Kd) for the fusion protein in the presence of charged UFC1. Additionally, we investigated the impact of DDRGK1-UFL1 on the stability of the thioester and observed its destabilizing effect (Fig. 5A).

*For example, can a stable UFC1-UFM1 conjugate be generated? How does this change in the presence of UFL1?*

We generated a stable UFC1-UFM1 conjugate using an isopeptide bond. Specifically, we introduced a mutation at position C116, replacing it with lysine. Next, we utilized UBA5 to create a linkage between UFC1 and UFM1 through an isopeptide bond. This stable conjugate was then employed in binding experiments with DDRGK1-UFL1 (Fig. 5C). Furthermore, for discharge experiments, we generated UFC1 charged with UFM1 by utilizing a thioester bond. We then showed that the DDRGK1-UFL1 stimulates UFC1 discharge (Fig. 5A).

Additional detailed comments for the authors are offered below.

Specific Comments.

*1. As it is still early days in the use and presentation of AlphaFold generated models, it is important to adopt careful language when referring to its output and its utility. Early in the manuscript you say that the structural models from AlphaFold can be used to guide design of protein constructs (by removing regions predicted to be disordered). But your results showing the importance of the N-terminal helix are contrary to that notion.*

To generate a construct that will form a globular structure, it is advised to remove regions that are modeled with low confidence (low pLDDT). This is what guided us in the design of the fusion construct.

However, to identify possible interactions that are not mediated by two defined domains, one would specifically look for isolated secondary structure elements predicted with low confidence. Such elements have the potential to be stabilized upon binding to their partner. As such, the UFL1 N-terminal helix of low pLDDT was a strong hint for interaction with a partner, in this case UFC1.

Thus, both high pLDDT, as well as low pLDDT (when combined with a defined secondary structure) provide important information to guide experiments.
Overall, this study has clearly shown how AlphaFold can boost research. Thanks to the AlphaFold models, we were for example able to determine the regions to use for the design of fusion constructs, and also to use the models as starting points for the determination of interface hotspots that are crucial for binding, using computational alanine scanning. All these models were experimentally validated.

*2. It would be interesting to know whether there were any hints in the sequence of DDRGK1 that would have suggested the presence of the winged-helix domain?*

DDRGK1 contains a full winged helix (WH) domain preceding the half WH that is complemented by UFL1. This full domain has been solved for mouse DDRGK1 (PDB ID: 1wi9, sequence identity 76%). The region that contains half of the WH domain that is complemented by UFL1 has not been solved.

*Are there similar complementary structures from other systems? On what would Alphafold model this complex?*

We performed a structure-based search in the PDB using DALI (*Holm 2022 NAR 50:W210)*, with a single and double repeat structure as query (including the complemented one) to see how many of the solved structures have winged-helix domain repeats and could be used as an implicit template (for the actual prediction, we did not use any templates from PDB70). Out of the resulting 2,020 structures with WH domains, only 6 contained at least two WH domain repeats (with sequence identities ranging between 9% and 16%), but none of them were complemented as in the UFL1-DDRGK1 system. There are only 3 hits with an alignment shorter than 55 residues with the query (the length that would cover less than one domain), but

none of them are interacting with another chain. Taken together these results, and our literature review, we have not yet found any other systems that could contain a complemented WH domain.

We have searched for templates that could have been used to model the interaction if AlphaFold2 would not have been available. Using SwissModel's sensitive HHBlits-based template search, we indeed found that the interacting region would have been correctly identified as harboring half, unfinished WH domains both for UFL1 and DDRGK1. Since there are a large number (962) of PDB structures with at least one WH domain and a range of sequence identity to the repeat (2-76%), it is safe to assume that some of these structures were part of the training dataset of AlphaFold2. This could imply that AlphaFold2 learned that these sequences belong to the WH domain fold, and implicitly performed a quick and efficient homology modeling, even without actual templates being provided.

*Does Alphafold model the fusion protein the same as the two proteins entered separately in AlphaFold_multimer?*

Our model of the structure of the fusion protein shows high similarity to the model generated for the complex composed of individual proteins (backbone RMSD= 0.18 Å for this region).

*3. It's not clear on what the conclusion that the proteins "do not elute as soluble aggregates" in the gel filtration experiment is based. This conclusion requires better documentation. Furthermore, it's surprising that 20 residues changes elution as much as it does. Could this indicate further structural differences?*

We concur with the reviewer's suggestion that additional validation is required to ascertain that the fusion proteins do not form aggregates in solution. In response, we

conducted a Size-Exclusion Chromatography coupled with Multi-Angle Light Scattering (SEC-MALS) experiment to measure the molecular weight of each fusion protein in solution (Fig. 1C). The obtained molecular weights align well with the expected values based on their respective amino acid sequences, indicating that they do not aggregate in solution. Having determined the molecular weights, we attribute the significant disparity in the elution profiles of these proteins to the profound influence of the N-terminus on protein migration, rather than protein oligomerization.

*4. Given that the crystal structure of the WH domain was solved using the AlphaFold model as the starting structure, it doesn't make a lot of sense to do a detailed analysis of the difference between the two. Furthermore, the resolution of 3.1 Å doesn't warrant detailed analysis.*

We followed the reviewer's suggestion and removed the detailed comparison of the crystal structure and the Alphafold model.

*It is stated that "In the crystal structure the last 18 amino acids are flexible and are not detected in the electron density map." That logic is backwards. In the crystal, the last 18 residues are not visible and may therefore be flexible. Or there may be high static heterogeneity.*

We have corrected the sentence "In the crystal structure the last 18 amino acids (UFL1 amino acids 183-200) are not detected, suggesting a flexible region".

*5. The functional assays as presented, with minimal labeling or description are uninformative. What is the strong band at ~30 kDa? Why is a His-tag used in one assay and a flag-tag used in the other? What is being ufmylated in these assays? Are UFM1 chains expected? Where would they run? In Fig 1g, why do no complex and complex look the same? What are the bands in the no ATP?*

We would like to thank the reviewer for bringing up these important points. We made significant modifications to our activity assays. Specifically, we have implemented

the use of anti-Myc antibody for the detection of UFL1 (Fig. 1H) or DDRGK1-UFL1 (Fig. 1F), and anti-FLAG antibody for the detection of UFM1 (Fig. 1E, G). We no longer rely on the His-tag for detection purposes. Additionally, we have ensured clear labeling of all bands, providing explicit correspondence between each band and its respective protein.

*6. On page 10, the titration would suggest that UFC1-UFL1 binding is stronger than UFC1-UBA5.What sequence of UFL1 was added for competition? If the whole protein, what happens if a peptide equivalent to the ~50 residues of UBA5 is added?* Again, how might the presence of UFM1 alter these interactions?

Based on our previous publications, we have conducted several measurements of UBA5 fragments binding to UFC1. The wild-type UBA5 (1-404), UBA5 (333-404), and UBA5 Δ349-389 fragments exhibit dissociation constants (Kd) of 0.9 μM, 2.5 μM, and 5.5 μM, respectively (Soudah *et al.* JMB 2019; Kumar *et al.* Nat Comm 2021). These values are within the range of DDRGK1-UFL1 (1-200) binding to UFC1, as demonstrated in this study.

Furthermore, we conducted simulations to explore the competitive binding of UBA5's C-terminal and UFL1's N-terminal helices for UFC1. The simulations revealed that the UFL1 peptide effectively outcompeted UBA5, which aligns with our experimental findings. Specifically, we directly measured the Kd of UFC1 with the UFL1 N-terminal helix to be 1.4 μM. This value is stronger than the previously published Kd for the interaction between UFC1 and the C-terminus of UBA5 (333-404) (2.5 μM).

However, the main parameter that increases the binding affinity of UFL1 is the charging of UFC1 with UFM1, upon which affinity increases by 10-fold, resulting in a Kd value of 0.23 μM. Overall, our data suggests that the affinity of UFL1 for UFC1 is

similar to its affinity for UBA5. However, once UFC1 is charged, UFL1 can outcompete UBA5 and facilitate the transfer of UFM1 to the target protein.

*Fig4 legend does not indicate titrated ratios, requiring reader to look it up: UBA5 @ 150μM, UFC1 @300μM, DDRGK1-UFL1 @700μM.*

We have added the protein concentrations in the figure legend (now Fig. 3).

*Could titrations be done of 15N UFC1 and equivalent peptides of UBA5 vs UFL1 to get relative Kds?*

While it is theoretically possible to employ NMR titrations for determining Kds, the binding process between UFC1 and UFL1 occurs on the intermediate-slow exchange NMR time scale. Consequently, ITC provides a more accurate measurement of affinity, and thus, we utilized ITC to obtain the following Kds: UFC1 with DDRGK1-UFL1 (Kd=2.3μM) and UFC1 with a peptide corresponding to the UFL1-N-terminus (Kd=1.4μM). In our previous reports, we utilized ITC to determine the Kds of UBA5 with UFC1 (Kd=0.94μM; Soudah *et al.* JMB 2019) and UFC1 with the C-terminus of UBA5 (Kd=2.5μM; Kumar *et al*. Nat Comm 2021). As evident, all Kd values fall within a similar range. However, when we assessed the Kd of DDRGK1-UFL1 in the presence of charged UFC1, we observed a decrease to 0.23 μM, indicating a preferential binding of DDRGK1-UFL1 to charged UFC1.

Referee #2:

*(1) The manuscript uncovered an interesting, helix-mediated regulatory mechanism in ufmylation: UFL1 N-terminal helix competes with the E1 (UBA5) C-terminal helix in binding to UFC1 and drives ufmylation, and it is an interesting finding using DeepMind/AlphaFold2 predictions and in vitro analysis.*

We would like to thank the reviewer for this comment.

*(2) Why did you omit the regulatory role of E3 co-factor LZAP in this mechanism? How do you exclude the effects of LZAP in your system prediction and in vitro analysis? Could you supplement these evidences in your research?*

The primary discovery presented in this manuscript revolves around the significance of the UFL1 N-terminus in its binding to UFC1. This observation was initially made using our fusion protein containing DDRGK1 and UFL1. To further validate the relevance of this finding in the native ternary complex comprising UFL1, DDRGK1, and LZAP, we went on to purify this complex and demonstrate that the N-terminus of UFL1 is indeed essential for its activity. Hence, we firmly believe that the findings derived from our initial investigation using the fusion protein extend to the native complex involving LZAP.

*(3)Could you supplement the morphological evidences of cell lines or animal models to support your conclusion?*

These are indeed intriguing points that we plan to investigate in future studies. However, at present, they are beyond the scope of this manuscript.

Referee #3:

*In this manuscript, Banerjee et al. focus on the understanding mechanisms of the enzymatic machinery responsible for UFMylation. The major focus is on how the E2 conjugating enzyme UFC1 works with the E3 UFL1 and its cofactor DDRGK1 to promote UFMylation. The authors first used AlphaFold2 modeling to design a UFL1/DDRGK1 fusion construct which was then used to determine the first crystal structure of this complex. The fusion construct was also used as a tool for in vitro UFMylation assays in which it was discovered that the N-terminal helix of UFL1 is important for promoting UFMylation. The authors next demonstrate a direct interaction between the UFL1 N terminal helix and UFC1 and again used AlphaFold2 to gain insights into the molecular details of this interaction. The binding surface for the UFL1 helix on UFC1 was further mapped using NMR and suggest that the UFC1 surface that engages the UFL1 helix overlaps with the same UFC1 surface that engages a helix in the structurally similar interaction between UFC1 and the E1 Uba5. Overall, the importance of UFMylation to biology is becoming increasingly apparent and this is a nice multidisciplinary study that represents a significant conceptual advance in our*

*understanding of how the UFM1 E2 and E3 function together to promote catalysis. With that said, there are a few important issues that should be addressed to make this manuscript an even stronger candidate for publication in EMBO reports:*

We would like to thank the reviewer for these supportive comments.

*1) The authors nicely demonstrate that the N-terminal helix of UFL1 mediates an interaction with UFC1 by ITC and NMR and that the presence of the helix promotes UFMylation in vitro, in the context of the UFL1/DDRGK1 fusion. Can the authors determine whether this interaction simply brings the E2 and E3 in proximity or whether this interaction stimulates the catalytic activity of the E2 using something akin to the UFC1-UFM1 lysine discharge assays? This would provide deeper mechanistic insights into the role of this interaction in UFMylation.*

We thank the reviewer for bringing up this important point. Following his/her suggestion, we conducted experiments to investigate the impact of DDRGK1-UFL1 on the stability of charged UFC1. As illustrated in the newly added Fig. 5, our findings demonstrate that DDRGK1-UFL1 promotes the discharge of charged UFC1. This observation is consistent with previous studies involving ubiquitin and E3 and E2 enzymes.

*It would be interesting to know whether the UFL1 helix alone can promote UFMylation as shown in Fig. 1f.*

We fully agree that testing the effect of the helix alone on ufmylation is very interesting. However, since in our assay we use UFL1 not only as an E3 but also as substrate, the lack of lysine residue at the N-terminus prevents it from serving as a substrate (We did test the effect of a peptide corresponding to the N-terminus on discharge of UFC1 but did not find a significant effect).

*2) If technically feasible, it would be informative to assess whether a UFC1-UFM1 thioester mimetic has an altered affinity for UFL1/DDRGK1 compared to UFC1 alone. This would*

*provide additional insights into the mechanistic importance of competition between Uba5/UFL1 for UFC1 and whether the UFM1 might contact the E3.*

Following the reviewer's suggestion, we have successfully generated a stable UFC1-UFM1 conjugate using an isopeptide bond. Subsequently, we utilized this conjugate to investigate its binding of DDRGK1-UFL1. Initially, through pull-down experiments, we confirmed that the fusion protein exhibits preferential binding to charged UFC1. Furthermore, utilizing ITC, we discovered that the affinity of this interaction is approximately 10-fold higher than the affinity of DDRGK1-UFL1 for uncharged UFC1. These results are now summarized in Fig. 5.

It is worth noting that this preferential binding was observed specifically in the fusion protein containing DDRGK1 spanning residues 87 - 314 (termed DDRGK1ext-UFL1 in our manuscript), but not in the fusion protein starting at residue 207. This observation aligns with a recent publication suggesting the presence of a UFM1 binding site in DDRGK1 spanning residues 116-214 (Ishimura *et al.,* Nat Comm 2022).

*3) It would be nice to expand a bit in the Discussion on how this emerging noncanonical E3 mechanism fits into the broader context of known E3 mechanisms in other Ubl pathways. Might UFL1 hold UFC1-UFM1 thioester intermediate in a 'closed' conformation to promote catalysis akin to other Ubl pathways (could AlphaFold2 provide insights)?*

In the revised version, we have combined the Result and Discussion sections. However, on *page 9* of the revised manuscript, we mention the possibility that the binding of DDRGK1-UFL1 to UFC1 may contribute to stabilizing a closed conformation of the charged UFC1. We note that despite our efforts, we were not able to generate a model with AlphaFold2 that could provide us with insights into the

potential occurrence of a closed conformation upon binding of DDRGK1-UFL1 to UFC1.

*As it pertains to the topic of UFL1 and Uba5 competition for UFC1 noted in the first and second paragraphs of the discussion, this is where the UFM1 that is linked via thioester bond to UFC1 may come into play. Note the thioester switch model for E1-E2 transthioesterification in the ubiquitin system (Yuan et al., Nature Comms 2022) and in the Nedd8 system (Huang et al., Nature 2007) and how such a phenomenon may be relevant to the UFM1 pathway.*

To date, our understanding of the structural mechanism underlying the transthiolation process in UFM1 is relatively limited compared to well-studied systems such as Ub or other UBLs. However, within our text, we raise the possibility that, similar to Ub and other UBLs, conformational changes occur during the transthiolation reaction. These conformational changes can subsequently impact the affinity between UBA5 and UFC1, thus influencing the overall dynamics of the UFM1 conjugation pathway and the competition with the E3.

*Finally, if the relative concentrations of Uba5, UFC1, and UFL1 in the cell are known it would be nice to comment on how this relates to the observed affinities for these interactions in vitro and the mechanistic implications of E1/E3 competition for E2.*

This is an intriguing idea that we were also interested in exploring. However, due to the anchoring of UFL1 to the ER membrane through DDRGK1, interpreting the effect of enzyme concentrations on ufmylation becomes challenging, especially when compared to freely present UFC1 and UBA5. Nonetheless, on page 9 of the revised manuscript, we highlight the significance of enzyme localization as another mechanism that can regulate the competition between UBA5 and UFL1 for UFC1 binding.

Minor points:

*• Overall, the figures could benefit from more detailed labeling, especially Fig. 1f and g.*

We have made significant updates to the figures in our manuscript, particularly the activity blots. We have taken great care to clearly label each band, ensuring that they are properly defined and easily identifiable.

*• Advise against using yellow color for UFL1 as it is very hard to see in print.*

We have modified the colors of the structural figures.

Dear Ora, dear Reuven,

Thank you for the submission of your revised manuscript to EMBO reports. As you know, we have received the two enclosed reports on it. Referee 3 supports publication of the revised manuscript but referee 1 raises concerns regarding the activity and discharge assays provided.

Thank you also for your patience, while we have discussed the remaining concerns from referee 1 as well as your response to these with the entire editorial team, referee 3 and an independent editorial advisor who is familiar both with the field and with our journal and its scope.

Referee 1 raised concerns regarding the discharge data presented in Figure 5A. The discharge appears to be slow, as evidenced by the lack of appearance of either free UFM1 or free UFC1 in the timepoints analysed. Referee 3 agreed with the statement that the discharge is slow but also appreciated that you did not overstate the results on page 8 and in the discussion on page 9in his/her view. Referee 1 suggested in his/her report to repeat the discharge assays with the extended version of DDRGK1 (DDRGK1[ext]) to test whether UFL1 might function similar to HECT E3 ligases. In your response you point to the recently published study by Joshua Peter, Yogesh Kulathu and colleagues (https://doi.org/10.15252/embj.2022111015) and note that this study already ruled out a HECT-like mechanism and provided evidence that UFL1 acts similar to RING E3 ligases. Referee 3 considered these arguments valid and further noted: "[...] while it does not reveal the precise molecular mechanism for UFL1 E3 ligase activity, this study has provided significant insights into the molecular basis for E2/E3 recognition. I think working out the precise E3 ligase mechanism is going to be an exciting focus of future investigation, as this E3 appears to function though a unique mechanism compared to canonical RING E3 ligases."

The second major concern from referee 1 relates to the activity assays presented in Figure 1 E-H. While in panel E a reasonable amount of product is detected, the assays using Myc-tagged fusion construct shown in panel F and G detect a very small amount of product representing only a small fraction of the substrate. The referee is concerned that either the anti-Myc antibody recognizes the UFM1-modified fusion protein with low affinity or that the absence of the UFM1-binding region of DDRGK1 accounts for the low activity. Referee 1 suggests repeating the assay with DDRGK1[ext]. Referee 3 agreed with this suggestion.

As noted above, I have also discussed these remaining concerns and your response to these with an independent expert editorial advisor.

The advisor commented:
"I can see the point that the main message is the role of the N-terminal helix of UFL1 in the reaction, and for that the discharge experiments do not directly help. The quality of the ITC data is very high and with the X-ray and NMR together makes a nice story, but the activity assays are indeed rather limited, relying on single point assays. In fact, I agree with the reviewer that the importance of this helix for activity is really only shown under suboptimal conditions, in the absence of the UFM1 interaction region of DDRGK1, except for the full length triple complex pull downs of fig 1 G-H. These were done by parallel expression and it is clear that UFL1deltaN protein levels were substantially lower than those for wt, both from the ponceau, and the blot, and given the thin line for the product with the WT, this could just be enough to prevent seeing anything for the mutant. All in all, this means that the paper would be strengthened by activity assays with the fusion construct that contains the UFM1 interaction domain (DDRGK1ext-UFL1) in the presence and absence of the N-terminal helix, to confirm its importance under more active conditions. If they do that, adding some time and/or concentration series rather than these single point assays could further strengthen the data.

Regarding the discharge data in Figure 5, the advisor suggested to "try it with the longer DDRGK1 construct, with both reducing and non-reducing gel: if there is significant discharge to the E2 you could even use the production of non-reducible E2-UFM1 as read-out".

**********

Taking all this information into account, the referee reports, the feedback from you and referee 3 as well as the opinion of our advisor and that there is in essence an agreement that further assays are required to support the activity assays in Figure 1, we would like to give you the chance to further revise your study by performing the experiments suggested by referee 1 and the advisor, i.e., repeat the assay with the DDRGK1ext-UFL1.

Further experiments to bolster the discharge experiments in Figure 5 are recommended but not mandatory. Please discuss these data also in the context of the data provided by Peter et al. Please also address all other technical concerns and provide a point-by-point response.

The revised manuscript will be seen again by referee 3 and the advisor.

**********

In addition to these experiments, there are also a few things from the editorial side that require your attention:

- In order to provide enough space for the description of all results and for the discussion of these, we will publish your study as a full Article. Please provide separate Results and Discussion sections.

- Please provide up to five keywords.

- Please provide the EV figures as separate high-resolution files, i.e., one file per figure and remove the legends from the EV figure files.

- Please add callouts to Fig. EV2B.

- Figures and their panels should be called out in a numerical and alphabetical order. In this context we note that Fig EV1D-H are called out after EV2C. If possible, please rearrange these panels.

- The manuscript sections are in the wrong order. Please order them like this:
Title page - Abstract - Introduction - Results - Discussion - Materials and Methods - Acknowledgements - Disclosure and competing interests statement - References - Figure legends - Tables and their legends (not EV tables) - Expanded View Figure legends

- The EV tables should be uploaded individually. Please remove the EV table legends from the manuscript file.

- I attach to this email a related manuscript file with comments by our data editors. Please address all comments and upload a revised file with tracked changes with your final manuscript submission.

- In the Author Checklist you state that the Core Facility was acknowledged in Table EV1, which however seems not to include such a statement/reference.

- The data shown in Figure EV1A and EV2A look very similar to each other (A: UFL1 and B: DDRGK1). In case the same data are shown in these two panels, please clearly state so in the legend of Fig. EV2.

- Finally, EMBO Reports papers are accompanied online by A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is 550x300-600 pixels large (width x height) in PNG for JPG format. You can either show a model or key data in the synopsis image. Please note that the size is rather small and that text needs to be readable at the final size. Please send us this information along with the revised manuscript.

We look forward to seeing a revised version of your manuscript as soon as possible.

With kind regards,

Martina

Martina Rembold, PhD
Senior Editor
EMBO reports

**********

Referee #1:

The authors have responded to the suggestions and comments made by the original review and the revised manuscript is improved over the original. The logic of the studies and model building is clearly presented and the results from those are now validated. As requested, the authors have included some in vitro biochemical assays to demonstrate the activity of their constructs and to test certain of their hypotheses. But here, there are issues, some of which are technical and others of which are more fundamental (see the Point-by-Point section below). The central point stems from the low in vitro activity observed, both in terms of products formation (Figure 1) and E2 discharge (Figure 5). The authors either fail to recognize that their in vitro system and conditions are yielding unusually low activities or have chosen not to comment on this. The "good news" is that the low activity under their current assay design likely hints at a novel mechanism of action; the "bad news" is that without further

characterization, the work in falls short of being much more than a validation of an AlphaFold prediction. Answers to the questions raised below will shed light on potentially unique features of the UFM1-modifying system as well as what features are in common with the better understood ubiquitin (and other UBL)-modifying systems. A study that fully addresses the interesting possibilities would represent an important contribution to current understanding; the present manuscript falls short of this.

Point-by-point critique:

Figure 1.
Panel E. From the Methods section, it appears that reaction mixtures were quenched in beta-ME-containing SDS page buffer but then run on "non-reducing PAGE." What is the logic behind that? The bME in the quench buffer will destroy any thiol-linked UFM1 conjugates, leaving only isopeptide-linked (presumably) UFM1 products. Is that the intention? If so, why run non-reducing gels?

Use of Western blots in Panels E - H help simplify the presentation but make it difficult to assess how much product is being produced. For example, there appears to be a reasonable amount of "products" formed in Panel E (but how much of each species is actually modified? One would need to see Coomassie stained gels for that.). But the amount of product detected in the anti-Myc blots for DDRGK1-UFL1-UFM1 is a very small fraction of the "substrate" in panels F and G. Is this an issue with the anti-Myc antibody not recognizing UFM1-modified fusion protein? How strong is the band in a blot for the FLAG-UFM1? Or might the low reactivity be due to using a construct that does not contains the UFM1-binding region of DDRGK1, introduced later in the manuscript? Assays using the extended version of DDRGK1 will be very insightful.

Figure 2 (and other figures with ITC data).
There is no mention of the number of replicates performed for the ITC and no statistical/error analysis is presented throughout the manuscript. Some of the differences reported are rather small, so it's important to report the uncertainties associated with the values being compared.

Figure 5.

There is remarkably little discharge in the assays presented, as evidenced by the lack of appearance of either free UFM1 or free UFC1 in the timepoints. This is surprising given that the assays were performed at pH 8, where discharge from hydrolysis alone is expected to be rapid. In Panel A, an enormous amount of the E2~UFM1 conjugate was used in order to visualize the appearance of weak bands of free UFM1 and UFC1 (Lanes 5-7). This either means that 1) the conjugate is very stable even in the presence of Lys or 2) the E2 readily self-ufymlates to form an isopeptide-linked UFM1-UFC1 species that migrates in the same position as the thio-ester-linked conjugate in the gels. Running each time point on both non-reducing and reducing gels may help clarify this point. It's important to parse this issue out as it has different consequences in terms of mechanism. Once again, part of the issue (if indeed the discharge is very slow) may be the omission of the additional extension on DDRGK1 shown to be important in Figure 5 Panel B & C. This is a critical point as it has direct bearing on the mechanism.

If the discharge assays still show weak discharge in the context of the extension, this raises other mechanistic alternatives. Can the conjugate discharge completely to Cys (see Wenzel et al., 2011, referenced in the manuscript)? To other nucleophiles? Parallel assays to monitor the rate of hydrolysis should be performed. If discharge occurs only to Cys, this would suggest a HECT-like mechanism, raising the question of whether UFL1 has an active-site Cys. Thus, the issues raised here are more than just technical--they are fundamental to understanding this interesting system. A study that fully addresses these interesting possibilities would represent an important contribution to current understanding.

RESULTS & DISCUSSION

The authors have chosen to combine Results and Discussion into one section. This is a matter mostly of taste and either format can be effective. However in its current form, there is virtually no "discussion" offered other than the final paragraph of the section and it left this reader feeling quite unsatisfied.

Referee #3:

In this revised manuscript, Banerjee et al. have adequately addressed all of the reviewer concerns. The considerable efforts to validate the structural modeling and extend the resulting mechanistic insights are appreciated. This reviewer fully supports publication of the revised manuscript in EMBO Reports.

Dear Editor,

We appreciate and thank you for the in depth assessment of our revised manuscript. We have carefully addressed the Reviewers concerns and performed the suggested experiments to improve our manuscript.
We have added activity assays using extended DDRGK1-UFL1 fusion fragments that contain the UFM1 binding site. We constructed one fusion protein with and one without the N-terminal helix of UFL1. We have also tested whether the extended fragment improves the stimulation of discharge using these constructs.
Our results show clearly that (1) the N-terminal helix of UFL1 is crucial for ufmylation activity, even in the presence of the UFM1 binding site in DDRGK1; and (2) the UFM1 binding site does NOT improve stimulation of UFC1~UFM1 discharge.
Below we provide a detailed response to all the concerns raised by the reviewers (in red).

We hope that the extended and improved manuscript can now be accepted for publication.

Thank you!

Ora Schueler-Furman and Reuven Wiener

---

Referee #1:

The authors have responded to the suggestions and comments made by the original review and the revised manuscript is improved over the original. The logic of the studies and model building is clearly presented and the results from those are now validated. As requested, the authors have included some in vitro biochemical assays to demonstrate the activity of their constructs and to test certain of their hypotheses.

We would like to thank the reviewer for his/her supportive feedback.

But here, there are issues, some of which are technical and others of which are more fundamental (see the Point-by-Point section below). The central point stems from the low in vitro activity observed, both in terms of products formation (Figure 1) and E2 discharge (Figure 5). The authors either fail to recognize that their in vitro system and conditions are yielding unusually low activities or have chosen not to comment on this. The "good news" is that the low activity under their current assay design likely hints at a novel mechanism of action; the "bad news" is that without further characterization, the work in falls short of being much more than a validation of an AlphaFold prediction. Answers to the questions raised below will shed light on potentially unique features of the UFM1-modifying system as well as what features are in common with the better understood ubiquitin (and other UBL)-modifying systems. A study that fully addresses the interesting possibilities would represent an important contribution to current understanding; the present manuscript falls short of this.

We understand the reviewer's concerns and have followed his/her suggestions, as detailed in our point-by-point responses to the critiques.

Point-by-point critique:

Figure 1.
Panel E. From the Methods section, it appears that reaction mixtures were quenched in beta-ME-containing SDS page buffer but then run on "non-reducing PAGE." What is the logic behind that? The bME in the quench buffer will destroy any thiol-linked UFM1 conjugates, leaving only isopeptide-linked (presumably) UFM1 products. Is that the intention? If so, why run non-reducing gels?

We would like to thank the reviewer for highlighting this point. As the reviewer correctly comprehended, the experiment was conducted in the presence of BME, leading to the cleavage of all thioester bonds. The reference to non-reducing PAGE in our initial report was indeed an error. This textual mistake has been rectified in the revised version.
We would like to clarify that we employed reducing gels supplemented with BME for this experiment.

Use of Western blots in Panels E - H help simplify the presentation but make it difficult to assess how much product is being produced. For example, there appears to be a reasonable amount of "products" formed in Panel E (but how much of each species is actually modified? One would need to see Coomassie stained gels for that.). But the amount of product detected in the anti-Myc blots for DDRGK1-UFL1-UFM1 is a very small fraction of the "substrate" in panels F and G. Is this an issue with the anti-Myc antibody not recognizing UFM1-modified fusion protein? How strong is the band in a blot for the FLAG-UFM1?

We concur with the reviewer's observation that the yield of the ufmylated product is relatively low. However, it is evident that this ufmylated product is entirely absent when UFL1 lacks the N-terminus (Figure 1E).

Or might the low reactivity be due to using a construct that does not contains the UFM1-binding region of DDRGK1, introduced later in the manuscript? Assays using the extended version of DDRGK1 will be very insightful.

We agree with the reviewer's suggestion that conducting ufmylation tests in the presence of the extended fragment with and without the N-terminus would further emphasize the significance of the N-terminus, even when DDRGK1 contains the UFM1-binding site. We have therefore repeated the experiment with an extended version of DDRGK1. The results in Figure 5D demonstrate clearly that in the absence of the UFL1 N-terminus, the extended fragment remains non-functional.

There is no mention of the number of replicates performed for the ITC, and no statistical/error analysis is presented throughout the manuscript. Some of the differences reported are rather small, so it's important to report the uncertainties associated with the values being compared.

Details on the number of replicates carried out for the ITC experiments are now provided in the Methods Section. Table EV3 includes details about the fits for one representative ITC experiment. All replicates are provided in a source file.
For the discharge experiment involving the small fusion protein, we have included p-values to underscore the significance of our findings.

There is remarkably little discharge in the assays presented, as evidenced by the lack of appearance of either free UFM1 or free UFC1 in the timepoints. This is surprising given that the assays were performed at pH 8, where discharge from hydrolysis alone is expected to be rapid. In Panel A, an enormous amount of the E2~UFM1 conjugate was used in order to visualize the appearance of weak bands of free UFM1 and UFC1 (Lanes 5-7). This either means that 1) the conjugate is very stable even in the presence of Lys or 2) the E2 readily self-ufymlates to form an isopeptide-linked UFM1-UFC1 species that migrates in the same position as the thio-ester-linked conjugate in the gels. Running each time point on both non-reducing and reducing gels may help clarify this point. It's important to parse this issue out as it has different consequences in terms of mechanism. Once again, part of the issue (if indeed the discharge is very slow) may be the omission of the additional extension on DDRGK1 shown to be important in Figure 5 Panel B & C. This is a critical point as it has direct bearing on the mechanism.

We followed the reviewer's suggestion and conducted a discharge experiment with the extended fragment. Additionally, we employed BME at the final time point to assess the level of ufmylated UFC1. Our data indicate that the presence of the UFM1 binding site on DDRGK1 does not enhance the stimulation of the discharge of charged UFC1 by free Lys. In addition, we do not detect a large increase in ufmylated UFC1 in the presence of the fusion proteins (see Figure EV4 D).

If the discharge assays still show weak discharge in the context of the extension, this raises other mechanistic alternatives. Can the conjugate discharge completely to Cys (see Wenzel et al., 2011, referenced in the manuscript)? To other nucleophiles? Parallel assays to monitor the rate of hydrolysis should be performed. If discharge occurs only to Cys, this would suggest a HECT-like mechanism, raising the question of whether UFL1 has an active-site Cys. Thus, the issues raised here are more than just technical--they are fundamental to understanding this interesting system. A study that fully addresses these interesting possibilities would represent an important contribution to current understanding.

We wholeheartedly agree with the reviewer's observation that the weak discharge suggests a potential mechanism for UFL1 as an HECT or RBR-like E3 ligase. However, this hypothesis has been rigorously examined by the Kulathu group through systematic mutagenesis of UFL1 cysteines. Their work definitively demonstrated that UFL1 lacks an active site cysteine, a fundamental requirement for an HECT-type E3 ligase mechanism.

RESULTS & DISCUSSION

The authors have chosen to combine Results and Discussion into one section. This is a matter mostly of taste and either format can be effective. However in its current form, there is virtually no "discussion" offered other than the final paragraph of the section and it left this reader feeling quite unsatisfied.

We have now separated the Results from the Discussion section.

Referee #3:

In this revised manuscript, Banerjee et al. have adequately addressed all of the reviewer concerns. The considerable efforts to validate the structural modeling and extend the resulting mechanistic insights are appreciated. This reviewer fully supports publication of the revised manuscript in EMBO Reports.

We would like to thank the reviewer for supporting our manuscript for publication in EMBO Reports.

Dear Ora, dear Reuven,

Thank you for the submission of your revised manuscript to EMBO reports. As you know, we have received the two enclosed reports on it. Referee 3 supports publication of the revised manuscript but referee 1 raises concerns regarding the activity and discharge assays provided.

Thank you also for your patience, while we have discussed the remaining concerns from referee 1 as well as your response to these with the entire editorial team, referee 3 and an independent editorial advisor who is familiar both with the field and with our journal and its scope.

Referee 1 raised concerns regarding the discharge data presented in Figure 5A. The discharge appears to be slow, as evidenced by the lack of appearance of either free UFM1 or free UFC1 in the timepoints analysed. Referee 3 agreed with the statement that the discharge is slow but also appreciated that you did not overstate the results on page 8 and in the discussion on page 9in his/her view. Referee 1 suggested in his/her report to repeat the discharge assays with the extended version of DDRGK1 (DDRGK1[ext]) to test whether UFL1 might function similar to HECT E3 ligases. In your response you point to the recently published study by Joshua Peter, Yogesh Kulathu and colleagues (https://doi.org/10.15252/embj.2022111015) and note that this study already ruled out a HECT-like mechanism and provided evidence that UFL1 acts similar to RING E3 ligases. Referee 3 considered these arguments valid

and further noted: "[...] while it does not reveal the precise molecular mechanism for UFL1 E3 ligase activity, this study has provided significant insights into the molecular basis for E2/E3 recognition. I think working out the precise E3 ligase mechanism is going to be an exciting focus of future investigation, as this E3 appears to function though a unique mechanism compared to canonical RING E3 ligases."

The second major concern from referee 1 relates to the activity assays presented in Figure 1 E-H. While in panel E a reasonable amount of product is detected, the assays using Myc-tagged fusion construct shown in panel F and G detect a very small amount of product representing only a small fraction of the substrate. The referee is concerned that either the anti-Myc antibody recognizes the UFM1-modified fusion protein with low affinity or that the absence of the UFM1-binding region of DDRGK1 accounts for the low activity. Referee 1 suggests repeating the assay with DDRGK1[ext]. Referee 3 agreed with this suggestion.

As noted above, I have also discussed these remaining concerns and your response to these with an independent expert editorial advisor.

The advisor commented:
"I can see the point that the main message is the role of the N-terminal helix of UFL1 in the reaction, and for that the discharge experiments do not directly help. The quality of the ITC data is very high and with the X-ray and NMR together makes a nice story, but the activity assays are indeed rather limited, relying on single point assays. In fact, I agree with the reviewer that the importance of this helix for activity is really only shown under suboptimal conditions, in the absence of the UFM1 interaction region of DDRGK1, except for the full length triple complex pull downs of fig 1 G-H. These were done by parallel expression and it is clear that UFL1deltaN protein levels were substantially lower than those for wt, both from the ponceau, and the blot, and given the thin line for the product with the WT, this could just be enough to prevent seeing anything for the mutant. All in all, this means that the paper would be strengthened by activity assays with the fusion construct that contains the UFM1 interaction domain (DDRGK1ext-UFL1) in the presence and absence of the N-terminal helix, to confirm its importance under more active conditions.

We completely share the advisor's concerns regarding the activity assay and the necessity of the N-terminus when the fusion protein already contains the UFM1 binding site of DDRGK1. Therefore, we conducted ufmylation tests in the presence of the extended fusion protein both with and without the N-terminus. As anticipated, in the absence of the UFL1 N-terminus, we did not observe any ufmylation of the extended fragment (Figure 5D).

If they do that, adding some time and/or concentration series rather than these single point assays could further strengthen the data.

We have added a series of time points in Figure EV4C, showing the increment in ufmylation product over time.

Regarding the discharge data in Figure 5, the advisor suggested to "try it with the longer DDRGK1 construct, with both reducing and non-reducing gel: if there is significant discharge to the E2 you could even use the production of non-reducible

E2-UFM1 as read-out".

We have performed discharge experiments in the presence of the extended fusion (see Figure EV4E). Our data show that there is no enhancement in discharge in the presence of the extended fragment.

**********

Taking all this information into account, the referee reports, the feedback from you and referee 3 as well as the opinion of our advisor and that there is in essence an agreement that further assays are required to support the activity assays in Figure 1, we would like to give you the chance to further revise your study by performing the experiments suggested by referee 1 and the advisor, i.e., repeat the assay with the DDRGK1ext-UFL1.

Further experiments to bolster the discharge experiments in Figure 5 are recommended but not mandatory. Please discuss these data also in the context of the data provided by Peter et al. Please also address all other technical concerns and provide a point-by-point response.

The revised manuscript will be seen again by referee 3 and the advisor.

In addition to these experiments, there are also a few things from the editorial side that require your attention:
We have addressed all the issues below.

- In order to provide enough space for the description of all results and for the discussion of these, we will publish your study as a full Article. Please provide separate Results and Discussion sections.

Done

- Please provide up to five keywords.

DDRGK1, UFC1 , AlphaFold2 , ufmylation UFL1

- Please provide the EV figures as separate high-resolution files, i.e., one file per figure and remove the legends from the EV figure files.

Done.

- Please add callouts to Fig. EV2B.

Could you please provide more details? It was not clear what needs to be added here.

- Figures and their panels should be called out in a numerical and alphabetical order. In this

context we note that Fig EV1D-H are called out after EV2C. If possible, please rearrange these panels.

Fig EV1D-H are mentioned in Fig 1, while FigEV2C is mentioned in the text but belongs to Fig2 that deals with the interaction of fusion with UFC1. We therefore think that the order is right.

- The manuscript sections are in the wrong order. Please order them like this:
Title page - Abstract - Introduction - Results - Discussion - Materials and Methods - Acknowledgements - Disclosure and competing interests statement - References - Figure legends - Tables and their legends (not EV tables) - Expanded View Figure legends

Done

- The EV tables should be uploaded individually. Please remove the EV table legends from the manuscript file.

Done

- I attach to this email a related manuscript file with comments by our data editors. Please address all comments and upload a revised file with tracked changes with your final manuscript submission.

We have addressed the comments, namely, we have corrected the links to structure and structural model depositories, we have added a definition of the dashed line in Figure 3A, and a legend to the lower panel in Figure 5A.

- In the Author Checklist you state that the Core Facility was acknowledged in Table EV1, which however seems not to include such a statement/reference.

We have added an acknowledgement to the ESRF beamline.

- The data shown in Figure EV1A and EV2A look very similar to each other (A: UFL1 and B: DDRGK1). In case the same data are shown in these two panels, please clearly state so in the legend of Fig. EV2.
These two Figures differ in that the first shows a prediction of the complex between UFL1 and DDRGK1 (dimer), while the second shows a prediction of the complex between UFL1, DDRGK1 and UFC1 (trimer). While the first part of these plots is similar, UFC1 is present only in the second.

- Finally, EMBO Reports papers are accompanied online by A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is 550x300-600 pixels large (width x height) in PNG for JPG format. You can either show a model or key data in the synopsis image. Please note that the size is rather small and that text needs to be readable at the final size. Please send us this information along with the revised manuscript.

We have included both a summary and a graphical abstract.

We look forward to seeing a revised version of your manuscript as soon as possible.

With kind regards,

Martina

Manuscript number: EMBOR-2023-56920V3
Title: Structural study of UFL1-UFC1 binding uncovers importance of UFL1 N-terminal helix for ufmylation
Author(s): Sayanika Banerjee, Julia Varga, Manoj Kumar, Guy Zoltsman, Shahar Rotem-Bamberger, Einav Cohen-Kfir, Isupov Isupov, Rina Rosenzweig, Ora Schueler-Furman, and Reuven Wiener

Dear Ora, dear Reuven

Thank you for your patience while we have reviewed your revised manuscript. As you will see from the reports below, former referee 3 and the advisor (listed as referee 4) are now all positive about its publication in EMBO reports. I am therefore writing with an 'accept in principle' decision, which means that I will be happy to accept your manuscript for publication once a few minor issues/corrections have been addressed, as follows.

- Please address the remaining concerns from the advisor/referee 4.

- Please add callouts to Fig. EV2B.
Explanation: In the text and legends you refer to EV2A, EV2, EV2C, and EV2D. But EV2B is never mentioned. Is there any instance where you describe data that is in EV2B? At this point, please add a reference to this panel. All figure panels need to be called out at some point in the text.

- Figure EV1 panels A and B appear to have rather low resolution. The text in panels C - G appears pixelated and the text size differs from that in panel H

- Figure EV2 panels C and D have different text types and the text appears in part pixelated.

- I also attach your manuscript with a few minor corrections and comments. I have also added the keywords to the title page and I moved the figure legends after the references. Please address all comments in the file and then upload it as final manuscript file.

- Please provide a short summary (1-2 sentences) and up to 4 bullet points, describing the key findings of your study.

Once you have made these minor revisions, please use the following link to submit your corrected manuscript:

Link Not Available

If all remaining corrections have been attended to, you will then receive an official decision letter from the journal accepting your manuscript for publication in the next available issue of EMBO reports. This letter will also include details of the further steps you need to take for the prompt inclusion of your manuscript in our next available issue. In order to publish your article in our December issue, please submit by October 27th.

Thank you for your contribution to EMBO reports.

Kind regards,

Martina

Martina Rembold, PhD
Senior Editor
EMBO reports

**********

Referee #3:

In this revised manuscript, Banerjee et al. have adequately addressed the remaining reviewer comments and have further strengthened the major conclusions of their paper. Along with previously published results from the Kulathu group, the results of experiments conducted during this revision suggest that UFL1 harbors a novel mechanism of action that will be an interesting focus of future investigation. Lastly, separating the Discussion from the Results provides the reader with a clearer understanding of how the results of this study fit into the broader context of the ubiquitin/ubiquitin-like protein field. This reviewer fully supports publication of the revised manuscript in Embo Reports.

Referee #4:

The authors have addressed the main points adequately

The PDB file uploaded does not make clear what is DDRGK and what is ligase, and it is not clear in the validation report either so it would be check if the official PDB file is more explicit, or to communicate with PDB and change the documentation

The authors addressed all editorial comments.

Ora Schueler-Furman
Hebrew University
Microbiology and Molecular Genetics
Ein-Karem
Jerusalem, _ 91120
Israel

Dear Dr. Schueler-Furman,

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. As you are aware, this File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: emboreports@embo.org]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

Thank you again for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Kind regards,

Martina Rembold, PhD
Senior Editor
EMBO reports

**********

THINGS TO DO NOW:

Please note that you will be contacted by Wiley Author Services to complete licensing and payment information. The required 'Page Charges Authorization Form' is available here: https://www.embopress.org/pb-assets/embo-site/er_apc.pdf - please download and complete the form and return to embopressproduction@wiley.com

EMBO Press participates in many Publish and Read agreements that allow authors to publish Open Access with reduced/no publication charges. Check your eligibility: https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/affiliation-policies-payments/index.html

You will receive proofs by e-mail approximately 2-3 weeks after all relevant files have been sent to our Production Office; you should return your corrections within 2 days of receiving the proofs.

Please inform us if there is likely to be any difficulty in reaching you at the above address at that time. Failure to meet our deadlines may result in a delay of publication, or publication without your corrections.

All further communications concerning your paper should quote reference number EMBOR-2023-56920V4 and be addressed to emboreports@wiley.com.

Should you be planning a Press Release on your article, please get in contact with emboreports@wiley.com as early as possible, in order to coordinate publication and release dates.

# EMBO Press Author Checklist

| | |
|---|---|
| Corresponding Author Name: Ora Schueler-Furman and Reuven Wiener | **USEFUL LINKS FOR COMPLETING THIS FORM** |
| Journal Submitted to: EMBO Reports | The EMBO Journal - Author Guidelines |
| Manuscript Number: 2023-56920V1 | EMBO Reports - Author Guidelines |
| | Molecular Systems Biology - Author Guidelines |
| | EMBO Molecular Medicine - Author Guidelines |

**Reporting Checklist for Life Science Articles (updated January**

This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your

**Please note that a copy of this checklist will be published alongside your article.**

**Abridged guidelines for 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.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical
- if n<5, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data

**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 χ2 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.

**Please complete ALL of the questions below.**
**Select "Not Applicable" only when the requested information is not relevant for your study.**

## Materials

| Newly Created Materials | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| New materials and reagents need to be available; do any restrictions apply? | Yes | Materials and methods |

| Antibodies | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| For **antibodies** provide the following information:<br>- Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number<br>- Non-commercial: RRID or citation | Yes | Materials and methods |

| DNA and RNA sequences | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Short novel DNA or RNA including primers, probes**: provide the sequences. | Yes | Table EV4 |

| Cell materials | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Cell lines:** Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID. | Yes | Materials and methods |
| **Primary cultures:** Provide species, strain, sex of origin, genetic modification status. | Not Applicable | |
| Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination. | Not Applicable | |

| Experimental animals | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Laboratory animals or Model organisms:** Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID. | Not Applicable | |
| **Animal observed in or captured from the field:** Provide species, sex, and age where possible**.** | Not Applicable | |
| Please detail **housing and husbandry conditions**. | Not Applicable | |

| Plants and microbes | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Plants:** provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens). | Not Applicable | |
| **Microbes:** provide species and strain, unique accession number if available, and source. | Yes | Materials and methods |

| Human research participants | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants. | Not Applicable | |

| Core facilities | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| If your work benefited from core facilities, was their service mentioned in the acknowledgments section? | Yes | Table EV1 |

## Design

| Study protocol | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|

| | | |
|---|---|---|
| If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI. | Not Applicable | |
| Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable. | Not Applicable | |

| **Laboratory protocol** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Provide DOI OR other citation details if **external detailed step-by-step protocols** are available. | **Yes** | Materials and methods |

| **Experimental study design and statistics** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Include a statement about **sample size** estimate even if no statistical methods were used. | **Yes** | figure legend |
| Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described? | Not Applicable | |
| Include a statement about **blinding** even if no blinding was done. | Not Applicable | |
| Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established?<br><br>If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification. | Not Applicable | |
| For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared? | **Yes** | figure legend |

| **Sample definition and in-laboratory replication** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| In the figure legends: state number of times the experiment was **replicated** in laboratory. | **Yes** | figure legend |
| In the figure legends: define whether data describe **technical or biological replicates**. | **Yes** | figure legend |

## Ethics

| **Ethics** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. | Not Applicable | |
| Studies involving **human participants**: 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. | Not Applicable | |
| Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained. | Not Applicable | |
| Studies involving experimental **animals**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations. | Not Applicable | |
| Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why. | Not Applicable | |

| **Dual Use Research of Concern (DURC)** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm | Not Applicable | |
| If you used a select agent, is the security level of the lab appropriate and reported in the manuscript? | Not Applicable | |
| If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript? | Not Applicable | |

## Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

| **Adherence to community standards** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided. | Not Applicable | |
| 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. | Not Applicable | |
| 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. | Not Applicable | |

## Data Availability

| **Data availability** | **Information included in the manuscript?** | **In which section is the information available?** (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section? | **Yes** | Data Avaliability section |
| Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement? | Not Applicable | |
| Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided? | **Yes** | Data Avaliability section |
| If publicly available data were reused, provide the respective **data citations in the reference list**. | **Yes** | |