THE EMBO JOURNAL

Manuscript EMBO-2016-96174

Sen1 has unique structural features grafted on the architecture of the Upf1-like helicase family

Bronislava Leonait, Zhong Han, Mr. Jerome Basquin, Mr. Fabien Bonneau, Domenico Libri, Odil Porrua and Elena Conti

Corresponding authors: Odil Porrua, Institut Jacques Monod; Elena Conti, Max Planck Institute of Biochemistry

Review timeline:

Submission date:	25 November 2016
Editorial Decision:	13 January 2017
Revision received:	28 February 2017
Editorial Decision:	28 February 2017
Revision received:	06 March 2017
Accepted:	09 March 2017

Editor: Anne Nielsen

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 13 January 2017

Thank you for submitting your manuscript for consideration by the EMBO Journal and my apologies for the extended duration of the review period, brought on by the recent holidays. Your study has now been seen by three referees and their comments are shown below.

As you will see from the reports, all three referees highlight the technical quality of your work and express interest in the findings reported in your manuscript. They consequently support publication here, pending minor revision that largely relates to text changes.

Given the referees' positive recommendations, I would like to invite you to submit a revised version of the manuscript, addressing the comments of all three reviewers. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance of your manuscript will therefore depend on the completeness of your responses in this revised version.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: http://emboj.embopress.org/about#Transparent_Process

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as

soon as possible upon publication of any related work, to discuss how to proceed. Should you foresee a problem in meeting this three-month deadline, please let us know in advance and we may be able to grant an extension.

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

REFEREE REPORTS

Referee #1:

This is a nice, simple paper that presents a structure for the Sen1 helicase C-terminal region, which is sufficient for helicase and termination activiteis. The structure shows the expected two RecA domains and some accessory domains. Comparison to the related Upf1 protein shows that the orientation of the accessory domains are different and may be fixed by a "brace". Based on some RNAse footprinting biochemical experiments, it is proposed that the "prong" is involved in termination and in melting RNA-DNA hybrids. Finally, some disease-causing point mutations found in the human senataxin protein are made in Sen1 and shown to affect folding or activity. Overall, there are important insights into Sen1 provided by the paper that will be helpful to everybody working on Sen1 and the entire class of RNA helicases.

A few minor points:

1. It's reasonable to speculate about a role for the prong in unwinding and termination. However, I wonder if the loss of activity in the LP deletion could be because of reduced RNA affinity (as expected from the smaller footprint) rather than a direct effect on the enzymatic steps. Although the L1549D point mutant argues against this, it could be worth doing the RNA binding titration (as in Fig 6) for these mutants to rule it out.

2. I'm confused about one part of the model in the Discussion. It's clear why the unwinding module would be on opposite ends of the RNA channel depending on the helicase polarity, but I don't understand the claim that this position could itself be "responsible" for the polarity of the unwinding. Since the movement along RNA is presumably due to the movement of the two RecA modules, how would the prong or CH domain change the direction? If the authors stand by this proposal, maybe an additional diagram would be useful for illustrating how they think this works.

Referee #2:

The manuscript by Leonaite et al. analyses the structure of the Sen1 helicase from S. cerevisiae, which is compared with Upf1 and IGHMBP2, other members of the Superfamily 1B. Based on the structure, the authors perform biochemical and mutational studies, to correlate the structure of Sen1 and its functional properties. In addition, disease-causing mutations in senataxin are interpreted with the structural information provided by the yeast ortholog.

Overall, the work is well performed and provides novel information about the structural bases of Sen1 function in transcription termination. The comparison between the three helicases, their structure and functional properties, provides an interesting perspective to correlate common and specific functions of each protein with their structure.

Minor points

- Despite the interest of the manuscript, the main message the authors want to provide (how the structure of Sen1 illuminates on how/why Sen1 can perform a distinct function in transcription termination) is lost throughout the text, and the discussion could be improved so that this message (why/how) is the main take home message.

- Related with this, some parts of the Results section provide too many structural details, and this is

sometimes difficult to follow. Maybe some simplification, without affecting the core of the description of results, could help readers with interest in helicases but who are not expert structural biologists.

Referee #3:

Conti and co-workers describe the crystal structure of a fragment (~ 44%) of the yeast RNA helicase Sen1. The crystallized fragment contains the helicase core, the beta-barrel domain, and the small stalk and prong domains, all of which are typical for SF1 RNA helicases. The Sen1 fragment also reveals a distinct "brace" segment, which appears to be conserved among eukaryotic Sen1 orthologs. The authors further show that the crystallized Sen1 fragment catalyzes unwinding and transcription termination in vitro and they probe the functional significance of several identified structural domains and residues. Finally, they test the functional importance of residues in human SETX that are mutated in diseases and that can now be mapped directly on the Sen1 structure.

The paper is very well crafted. It is clearly written, structure and biochemical experiments are well integrated, and the presented data are of high quality. One might question the degree of overall conceptual advance of the work, which seems to be limited to the discovery of the brace segment. Most of the SETX mutations could - most likely- be mapped on a homology model of SETX made from Upf1 or IGHMBP2 structures, but having a Sen1 structure is somewhat better. Of course, judgments of conceptual advance are somewhat arbitrary.

With respect to the substance of the paper, this reviewer found nothing of significance warranting augmentation or alteration before publication. The ms. is a solid contribution to our understanding of SF1 RNA helicases in general, and of Sen1 and orthologs in particular.

1st Revision - authors' response 28 February 2017

We are very grateful to the Referees for their positive comments and their constructive points. Below is the point-by-point response with the changes we have implemented to address the specific points raised by Reviewers 1 and 2.

Referee #1:

1. It's reasonable to speculate about a role for the prong in unwinding and termination. However, I wonder if the loss of activity in the LP deletion could be because of reduced RNA affinity (as expected from the smaller footprint) rather than a direct effect on the enzymatic steps. Although the L1549D point mutant argues against this, it could be worth doing the RNA binding titration (as in Fig 6) for these mutants to rule it out.

We have tested the RNA-binding properties of the L1549D mutant and of the entire LP deletion. Removal of the LP region ($Sen1_{Hel}$D1461-1554, which includes L1549) showed similar RNA-binding properties as compared to the wild-type (**new Figure EV3**). $Sen1_{Hel}$L1549D showed only a mild decrease (~2-fold) (**new Figure EV3**). These data are consistent with the similar levels of unwinding activity between the mutants and wild type, pointing to a specific effect rather than unspecific loss of RNA binding.

2. I'm confused about one part of the model in the Discussion. It's clear why the unwinding module would be on opposite ends of the RNA channel depending on the helicase polarity, but I don't understand the claim that this position could itself be "responsible" for the polarity of the unwinding. Since the movement along RNA is presumably due to the movement of the two RecA modules, how would the prong or CH domain change the direction? If the authors stand by this proposal, maybe an

additional diagram would be useful for illustrating how they think this works.

We agree with the Reviewer, and changed the statement to: "*Melting elements on opposite sides of the helicase core together with the movements of the two RecA domains in response to ATP hydrolysis could thus underpin the opposite unwinding polarities of Upf1-like and of Ski2-like RNA helicases.*"

Referee #2:

- Despite the interest of the manuscript, the main message the authors want to provide (how the structure of Sen1 illuminates on how/why Sen1 can perform a distinct function in transcription termination) is lost throughout the text, and the discussion could be improved so that this message (why/how) is the main take home message.

We now elaborate more in the final discussion paragraph the key determinants for the specific function of the Sen1 helicase domain in transcription termination, namely the *particular conformation that the "brace" imposes to the accessory domains as well as the distinctive characteristics of the "prong"*. In support of this model, we have included in the revised version of the manuscript *in vivo* data showing that *in the context of the full-length protein the LP deletion lead to lethality and provoked major transcription termination defects in vivo* (**new Figure EV4**).

- Related with this, some parts of the Results section provide too many structural details, and this is sometimes difficult to follow. Maybe some simplification, without affecting the core of the description of results, could help readers with interest in helicases but who are not expert structural biologists.

We have streamlined the description of the structure and eliminated many of the structural details from the Results section.

2nd Editorial Decision 28 February 2017

Thank you for submitting a revised version of your manuscript. I have now gone through the point-by-point response you provided to address the referee concerns and I happy to inform you that your study is now in principle ready to be accepted for publication here. However, before we can go on to officially accept your manuscript there are a few editorial issues concerning text and figures that I need you to address in a final revision:

-> Please include a brief conflict of interest statement in the manuscript.

-> Could you please provide a brief title for Table 1?

-> We noticed that you refer to Appendix Figure S1 in the manuscript but that there is no appendix uploaded. Is this supposed to be Figure EV1 instead?

-> Since there is only one file in the Appendix, I would encourage you to label this table as Table EV1 (and to update the callouts in the text accordingly)

-> Please rename the 'Experimental Procedures' in the main manuscript to 'Materials and methods'

-> We noticed that the image resolution for the blot in fig 1C is fairly low and if you have a higher resolution scan of the same gel I would encourage you to use that instead.

-> Please make sure that all figures displaying statistics analysis have information on the number of replicas and the nature of the error bars indicated in the figure legend (currently missing in fig 1B, 5D and 6D). In addition, we noticed that the data in fig 6E is displayed with error bars although the data derives from two independent experiments according to the legend. We generally require that n{less than or equal to}3 for statistical analysis and I would therefore encourage you to display the two data series here instead.

-> We generally encourage the publication of source data, particularly for electrophoretic gels and blots, with the aim of making primary data more accessible and transparent to the reader. We would need 1 file per figure (which can be a composite of source data from several panels) in jpg, gif or PDF format, uploaded as "Source data files". The gels should be labelled with the appropriate figure/panel number, and should have molecular weight markers; further annotation would clearly be useful but is not essential. These files will be published online with the article as a supplementary "Source Data". Please let me know if you have any questions about this policy.

2nd Revision - authors' response 06 March 2017

Thanks again for considering our manuscript for publication in EMBO J.
We have just submitted an amended version of our manuscript EMBOJ-2016-96174R containing the following modifications:

1. Manuscript main text:

- Materials and Methods instead of Experimental procedures.

-Statement of no conflict of interest.

-Modified figure legends to include detailed statistical information on experiments in all panels.

-Brief mention to supplementary methods in the Appendix.

2. Modified table 1 including a title.

3. Modified figures 1, 5 and 6 in which the experiments performed in two replicates display the values of one of the replicates (both values are provided in an excel file as source data).

In addition, we have uploaded the following files:

1. Appendix containing Figure S1, Table S1 and supplementary methods.

2. Source data files for figures 1, 4, 5 and 6.

3rd Editorial Decision 09 March 2017

Thank you for submitting the final revision of your manuscript, I am pleased to inform you that it has now been accepted for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name:	Elena Conti and Odil Porrua
Journal Submitted to:	EMBO J
Manuscript Number:	EMBOJ-2016-96174

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

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

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

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	NA
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	NA
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	NA
Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

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

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	PDB Coordinates Accession Number 5MZN
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	We refer not only to the papers but also to the PDB codes in the Figure legends where we made figures with these coordinates
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

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