

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection ImageStudio Lite version 5.2.5

Data analysis

Softwares used
 R (v4.3.2 and v4.4.0)
 GraphPad Prism (v9)
 Image Studio (v5.2.5, LI-COR)
 PEAR (v0.9.6)
 Fiji (v2.14.0)
 ImageJ (v2.14.0)
 pyCRAC (v1.4.5 and v1.5.2)
 Flexbar (v3.5.0)
 STAR (v2.7.10a and v2.7.11b)
 igvtools (v2.12.3)
 PEAKachu (v 0.2.0)
 DESeq2 package (v1.42.1 and v1.48.2)
 featureCounts (v2.0.0)
 Bowtie2 (v2.3.5.1 and v2.5.3)
 FastQC (v0.11.4)
 Perl (v5.32.1)
 SAMtools (v1.19 and v1.9)
 BAMtools (v2.5.2)

Spectronaut (v18.3, Biognosys)
 OpenMS pipeline NuXL (v1)
 OpenMS TOPPViewer (v2.6.0)
 AlphaFold (v3)
 bedSort (v469)
 bedGraphToBigWig (v469)
 BEDTools (v2.27.1)
 rMATS-turbo (v4.3.0)
 maser (v1.26.0)
 pheatmap (v1.0.13)
 ggplot2 (v4.0.1 and v3.5.2)
 UpSetR (v1.4.0)
 ComplexUpset (v1.3.6)
 ggsashimi (v1.1.5)
 TopSpin (v3.5 pl7)
 TopSpin (v4.3.0)
 biomaRt (v2.64.0)
 GenomicFeatures and GenomicRanges (v1.60.0)
 AnnotationDbi (v1.70.0)
 Biobase (v2.68.0)
 GenomInfoDb (v1.44.3)
 IRanges (v2.42.0)
 tibble (v3.3.0)
 gridExtra (v2.3)
 ggpubr (v0.6.2)
 RColorBrewer (v1.1-3)
 dplyr (v1.1.4)
 tidyr (v1.3.1)
 MikroWin 300 SL (v5.63)
 ImageQuant (v8.1)
 Typhoon FLA9500 Control Software (v1.0)
 FluorEssence Software (v3.9)
 Leica Application Suite X (v4.6.1.27508)
 MAFFT (v7)

Code availability

The code developed for optimized intron retention analysis is available via GitHub [<https://doi.org/10.5281/zenodo.18197506>].

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The deep sequencing data generated during this study have been deposited in Gene Expression Omnibus (GEO) database [<http://www.ncbi.nlm.nih.gov/geo/>] under the following accession codes:

GSE270752 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE270752>]: CRAC datasets for THUMPD2-His6-2xFLAG and the His6-2xFLAG tag.

GSE270753 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE270753>]: 3' RACE datasets for THUMPD2-His6-2xFLAG and the His6-2xFLAG tag.

GSE303357 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE303357>]: RNA-seq datasets for WT and THUMPD2 KO treated with siNT or siLARP7_1.

The proteomics data generated during this study have been deposited at the ProteomeXchange Consortium via the PRIDE partner repository [<https://www.ebi.ac.uk/pride/>] under the following accession code and submission reference:

PXD053510: IP-MS datasets for THUMPD2-His6-2xFLAG and the His6-2xFLAG tag.

PXD053619: Protein-RNA cross-linking MS dataset for THUMPD2-His6-2xFLAG.

Source data are provided with this paper.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Not applicable.

Reporting on race, ethnicity, or

Not applicable.

other socially relevant groupings

Population characteristics

Not applicable.

Recruitment

Not applicable.

Ethics oversight

Not applicable.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences☐ Behavioural & social sciences☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Statistical determination of sample size was not performed. In the case of non-quantitative experiments, sample size was determined to demonstrate clear reproducibility of the results (up to three independent experiments). Where the outcomes of experiments were quantified, at least three independent replicates were performed and statistical analyses of the data were done using appropriate statistical tests (detailed in the Methods section and Figure Legends).

Data exclusions

Data derived from experiments in which technical or human errors arose were excluded. Exclusion criteria were not pre-established. Exclusions were based on observations during the experimental procedures and the quality of the generated data.

Replication

Reproducibility was determined by replications and statistical analyses where appropriate. The number of replicates of each experiment presented in each figure are given in the figure legend. Once robust protocols were established, data generated in experiments without technical errors were highly reproducible.

Randomization

Randomization was not performed during biochemical and cell biological experiments. Covariates were controlled by growing cells under standardized conditions and utilizing standardized extraction techniques for purifying components e.g. proteins, RNAs, etc.

Blinding

Blinded experiments were not performed in this study. Experiments mostly compared specific treatments of otherwise comparable samples. To conduct the experiments, it was necessary for the investigators to be aware of the treatment/condition applied. Appropriate cellular and biochemical controls were included in each experiment/replicate.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants |

Methods

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used

See Supplementary Data 11.

Validation

Antibodies were validated by western blotting. Signals corresponding to the appropriate sizes of the target proteins were detected (see Figs. 1f, 3i, l, 4g, 6i-l, S1a). The antibodies against THUMPD2 and LARP7 were validated in this study by comparing extracts from RNAi-treated cells or cells genomically deleted of target genes (see Figs. 3l and 4g). The THUMPD2 antibody has previously been used in PMID: 37283053. The LARP7 antibody was KO/KD validated by the manufacturer and has been used in 13 peer-reviewed publications listed on the

manufacturer's website (<https://www.ptglab.com/products/LARP7-Antibody-17067-1-AP.htm?srsId=AfmBOoq006pWdFRuy9F1MDxThGA4aA6fnuvv2PTokbTf1atk0p-LdCvS#publications>). To validate the anti-FLAG antibody, extracts from HEK293 cells expressing or not expressing tagged versions of specific proteins were analyzed (see Figs1f, 3i, l). This antibody has been used in 8220 peer-reviewed manuscripts cited on the manufacturer's website (https://www.sigmaaldrich.com/DE/en/search/f3165?focus=papers&page=1&perpage=30&sort=relevance&term=F3165&type=citation_search).

All other antibodies used for western blotting in this study have previously been used in peer-reviewed publications listed on the manufacturer's websites:

Anti-Tubulin (Rabbit): (https://www.ptglab.com/products/TUBA1B-Antibody-11224-1-AP.htm?srsId=AfmBOOrEpYG1OLgagVljOpjKjK2MUPQDInjrGL3TU5fse_-WBGiuokVZ#publications)

Anti-Fibrillarin: (<https://www.scbt.com/p/fibrillarin-antibody-g-8?srsId=AfmBOoqj7xkN1UKN1q0u4mfl-5gQ9O-MvEojelJVPtmLaM8TaNK5roi>)

Anti-TRMT112: <https://www.scbt.com/p/trmt112-antibody-f-7?srsId=AfmBOopzHULwliOxWOPTKDIwSpK7-LTYNtV50UBRUstmtJLe51w4hzWR>

Anti-EFTUD2: (<https://www.fortislife.com/products/primary-antibodies/rabbit-anti-eftud2-snrp116-antibody/BETHYL-A300-957>)

Anti-MEPCE: (https://www.ptglab.com/products/MEPCE-Antibody-14917-1-AP.htm?srsId=AfmBOopgzsicLAjEdzJdWPdw4JRKfd53RUV_fd2Srmcd6d8WjplEMr8)

Anti-PRPF4: (https://www.ptglab.com/products/PRPF4-Antibody-10728-1-AP.htm?srsId=AfmBOooMkQLZxoZHDlpHU_hZXko8c6SBRpAowmbCcwh6M1I5sbyy4BpB)

Anti-PRPF3: (https://www.ptglab.com/products/PRPF3-Antibody-10106-1-AP.htm?srsId=AfmBOoqF-gEJCeFJngTnrMPaauNNMI1kzc-EBZIG_3But-xk2itfkptz)

Anti-Tubulin (Mouse): (https://www.sigmaaldrich.com/DE/en/product/sigma/t6199?srsId=AfmBOoqkL4XQ_StCSTXBafQE3IRGOJinvJyJl3WdJDXABoeZkEIJB_6)

Antibodies used for immunofluorescence have been previously verified:

Anti-UTP14A: PMID: 38065954

Anti-SART3: PMID: 38985674

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	Cell lines used in this study are based on the HEK293 Flp-In T-Rex cell line (ThermoFisher Scientific; cat. no R78007) or HCT116 (ECACC; cat no 91091005).
Authentication	CRISPR-Cas-mediated genome disruption of THUMP2 was confirmed by sequencing of genomic DNA, RNA-seq and western blotting (PMID: 37283053). Appropriate integration of cassettes for inducible expression of GFP-THUMP2, THUMP2-His6-2xFLAG or their derivatives were verified by western blotting (see Figs. 1f, 2a,b,e,f, 3g,h,i,l and S1a), and cells were maintained under appropriate antibiotic-mediated selection.
Mycoplasma contamination	Cell lines were routinely checked for mycoplasma contamination and confirmed to be negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Plants

Seed stocks	Not applicable.
Novel plant genotypes	Not applicable.
Authentication	Not applicable.

2'-O-methylation-dependent installation of N2-methylguanosine in the U6 internal stem loop facilitates efficient spliceosome assembly

Corresponding Author: Dr Katherine Bohnsack

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Kleiber et al.

2'-O-methylation-dependent installation of N2-methylguanine in the U6 internal loop facilitates efficient spliceosome assembly

In this manuscript, the authors have biochemically and functionally characterized the enzyme complex that m2G-modifies the U6 snRNA at position 72. Although this link has been made before, the authors show that this is the only modification that THUMP2 generates. THUMP2 is enriched in Cajal bodies, it associates with the adaptor protein TRMT112 and binds to U6 at later stages of U6 RNP assembly. The authors further use crosslinking experiments to map the interaction of THUMP2 to U6 and TRMT112. U6 is contacted at multiple regions in the N and the C terminal part of the protein, particularly the THUMP and the MTase domain. TRMT112 was mapped to the C terminus of the protein. They further generated a DNAzyme that allows for the site-specific identification of m2G72. THUMP2 knock out cells clearly show reduced modification at this position and complementation assays revealed that SAM binding and TRMT112 binding is required for methylation in cells. However, in vitro, the reaction was not efficient and therefore the authors investigated the contribution of associated 2'-O-methylation events. Indeed, pre-2'-O-methylated substrates were efficiently modified by the reconstituted enzyme complex in vitro. Using LARP7 knock down, which globally reduced U6 2'-O-methylation, the authors further show that prior U6 2'-O-methylation is also required in cells. Finally, the authors find that alternative splicing events are affected by THUMP2 knock out, LARP7 knock down and a stronger effect when both are inactivated simultaneously. This appears to be due to less efficient di- and tri-snRNP formation.

This is a very solid and clear manuscript. All experiments are competently carried out, presented and interpreted. Although the link between THUMP2 and U6 m2G72 has already been made, the manuscript goes far beyond and presents a thorough characterization of the entire pathway including experiments in cells, generation of new tools such as DNAzymes as well as biochemical in vitro reconstitution of the enzymatic activities. I have only a few comments that could be considered.

1. Fig. 1A: about 60% of the reads are crosslinked to lncRNAs or mRNAs (which is also seen in the negative control). It is clear that no modification sites besides the one in the ISL of U6 were found. However, THUMP2 may bind to other transcripts without methylating them. Have some of the mRNAs or lncRNAs been validated to clearly show that this is indeed background noise of the CRAC protocol?

2. Figure 4: 2'-O-methylation of the upper stem should lead to a stronger pairing and thus a more stable stem. To further test this, the first A-U pair (I guess 64) could be swapped to a G-C pair, which might overrule the strict requirement of a 2'-O-methylation at this position.

It is also interesting that the system seems to be not optimal since pairing of position 65 generates a much better substrate in vitro. What would be the evolutionary selection against such an optimal substrate? This could be further discussed.

3. The gradient shown in Fig. 5 are not fully clear. First, U4 is hardly visible and maybe a different probe could be used. U

snRNAs are also well stained by silver-staining protocols and since these are the major RNAs in nuclear lysates, this might be enough to visualize them.

The authors could perform the gradients under THUMP2 and LARP7 depletion conditions. This may enhance the effect and it may become clearer in these assays.

Reviewer #2

(Remarks to the Author)

In the manuscript by Bohnsack et al, the authors describe their systematic and structured approach to deciphering the installation of m2G in U6 snRNA and its biological consequences. I am impressed by the variety and at the same time high quality of the methods used by the whole team to answer important questions in the field of spliceosome assembly. The authors have used various and orthogonal methods and convincingly present data for the importance of m2G and its Nm interdependence in U6 snRNA. I see no need for additional experiments. From my perspective everything is covered. Yet, the Materials and Methods sections is not complete and requires careful reading and addition of information to ensure reproducibility.

- "Quantification of U6 snRNA modifications by LC-HRMS" is lacking detail. E.g. the digestion enzymes are not mentioned, columns supplier is not mentioned, which timsTOF was used, flow rate of chromatography, the description of the gradient is unclear. The IP-MS data analysis section is more clear, but still misses some details, e.g. column information is missing, the injection volume, was a trap used?

Further, I have found the manuscript difficult to read and digest. Therefore, I suggest substantial editing of the main text and the discussion. I have made a collection of instances in which I struggled in hope that the manuscript can be advanced and more easy to parse.

- Some sentences require English editing. E.g. line 157 ff and line 471 ff.

- line 252: The current title is confusing and I suggest changing it to "Predicted RNA contacts localize to the THUMP and MTase domains" or something similar to emphasise the modelling character of the section

- line 376: The sentence "Bovine proteins were used as HsTHUMP2 expressed poorly" is unclear. If I understand correctly, the authors expressed the bovine orthologs of THUMP2 (and TRMT112?) because recombinant expression of the human protein was inefficient. This should be clarified explicitly to avoid misinterpretation.

- line 385 ff: The sentence spanning lines 384–389 is overly long and difficult to follow. I recommend breaking it into shorter sentences. The same is true for lines 252 ff, 276 ff, 333 ff, 422 ff and 511 ff.

- The discussion is more like a recap or summary. Interpretation is limited and missing or weakly developed. Method limitations (e.g. structural modeling and DNAzyme sensitivity and its indirect character) are not discussed.

- N2,2-7-trimethylated cap (m3G): this nucleotide is commonly abbreviated as m227G according to modomics

I am confident that the work presented in this manuscript is a great fit for the broad readership of Nature Communications and I recommend publication after text revision.

Reviewer #3

(Remarks to the Author)

The manuscript by Kleiber et al. titled "2'-O-methylation-dependent installation of N2-methylguanosine in the U6 internal stem loop facilitates efficient spliceosome assembly" investigates how the m72G modification of the U6 snRNA is established by the THUMP2 enzyme and its reliance on the prior installation of 2'O methyl modifications on U6 snRNA by LARP7.

The manuscript has three major claims: 1- THUMP2 associates with U6 snRNA during late stages of its biogenesis; 2- m2G72 methylation requires 2'O methylation of U6 snRNA; and 3- m2G72 and 2'O methylations of U6 affect splicing interdependently.

Overall, the manuscript is well written with only minor writing errors and a few missing citations. The figures are well prepared and easy to understand. The materials and methods sections are detailed enough to reproduce the experiment. The manuscript makes a strong case for most of the claims and provides robust experimental data that clearly demonstrate the function of 2'O methylations in THUMP2 activity in vitro. It shows that the m72G modification is likely a late-stage modification in U6 biogenesis. The authors demonstrate that the absence of m72G likely affects spliceosome assembly and that the absence of THUMP2 and LARP7 impacts pre-mRNA splicing. The claims regarding the effects on pre-mRNA splicing could be supported more by additional analysis, as currently the effect of 2'O methylations is only tested in vitro. Nevertheless, overall, I believe this manuscript presents important data that will help the field better understand THUMP2 function and overall U6 snRNA biology.

Major comments

1- Line 149 – specify which human cells are over-expressing and note that they are over-expressing.

2- Line 149, I couldn't find any evidence or data showing this immunoprecipitations, its efficiency and its specificity in main or supplemental figures. It is important to demonstrate that protein and RNA immunoprecipitation are working effectively and specifically through western blots. Are the authors immunoprecipitating only a small percentage of the tagged protein or most of it? Similarly, how can they show that UV crosslinking is working effectively?

3- Similarly, Table S1 shows the overall reads mapped to different classes of RNAs or specific snRNAs, but it lacks information on the total reads (unmapped) from each experiment, the percentage mapped, read lengths, and so on.

4- Line 274, the main claim of the manual modelling is that it performs better than the AlphaFold-generated model. Therefore,

can the authors be more specific than “close proximity to THUMP2”? And following from this, why isn't G72 within the active site?

5- Figure 3i, similar to an earlier point about IP efficiency, makes it difficult to assess how effective their IP is since the input bands at 0.3% are stronger than the IP at ten times higher loading. The RNA binding data also suggest that THUMP2 binds only a very small fraction of U6 at any given time. Is that correct? Does the D329A mutant retain U6 longer?

6- Figure 4 provides a strong argument for the roles of 2'O methylation in G72 methylation by THUMP2. Additionally, in this figure, the authors demonstrate the use of a novel dnzyme-based method for detecting 2'O methylations. This is really good.

7- In Figure 5, authors make a good start to splicing analysis, but overall this section of the manuscript is weak. It was already known that the absence of THUMP2 and LARP7 affects splicing to some extent, and authors' data confirm this, showing there may be a stronger effect when both are missing. However, this does not tell us much about the specific role these modifications play in pre-mRNA splicing or why some genes are affected while others remain unaffected.

8- Figure 5i-k provides valuable insights into the mechanistic understanding of how these modifications influence spliceosome assembly. However, given its current format and the information provided, it is challenging to determine whether the conclusions are fully supported by the data. For example, in 5i, not only does U6 accumulate in fractions 10-17, but there is also a significant difference in U6 levels. Nevertheless, the authors show in figure s4 that the levels are the same. Additionally, there is little observable change in U4 or U5 levels or fractionation profiles. Furthermore, compared to Mabin et al., who performed similar fractionation, why is free U6 so low (PMID: 34234030)? Lastly, if less U6 is available for tri-snRNP assembly, why is the effect on splicing not more pronounced, impacting all transcripts and introns?

9- Can the authors show whether U6atac is affected in Figure 5i? The expectation is that U6atac should not be affected.

10- I don't like asking for additional complex experiments, but authors could consider testing whether spliceosome assembly, specifically of U6, is similarly affected in LARP7 KD cells compared to THUMP2. Currently, the 2'O methylation requirement for m7g is only evident from in vitro experiments. This could provide in vivo mechanistic insight into whether 2'O methylation can mimic THUMP2 absence. Alternatively, the wording around these results should more clearly indicate that the dependence is only evident in vitro.

11- I would like to see full-size gels of all Western and Northern blots as a supplementary file to ensure that the data is not selectively presented.

12- I appreciate the authors for providing a detailed materials and methods section.

Minor Comments

1- All prime symbols are incorrect. An apostrophe is not a prime symbol, and please ignore my use of one in my review report for ease of writing.

2- Line 118 – Citation for the METTL16 being the U6 snRNA m6A methyltransferase should include PMID: 28525753 and PMID: 29262316, and U6 snRNA A43 being involved in splicing should also cite PMID: 36409063 and PMID: 38808663.

3- Line 147 – Is a punctuation mark missing before “so”?

4- In Figure 2e, the GFP panel shows that the cell on the middle left has an incorrectly placed arrow.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have adequately and competently responded to the few minor comments that I had raised on their previous version of the manuscript. Thus, I am satisfied with the response to my comments.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I am satisfied with the provided revision and I recommend the manuscript for publication.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

Authors have made substantial revisions, including additional experiments that provide further support for their conclusions and additional information about the role of U6 2'O methylations.

They have addressed all the points I raised, and I appreciate the detailed, point-by-point response to our comments.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

We thank the referees for their positive feedback and constructive comments to our work. Below, we provide a complete point-by-point response to the reviewer's specific suggestions. The corresponding modifications to the text are indicated in red in the word files of the manuscript. In brief, we have added the following data to the manuscript:

Fig. 1a. Re-analysis of THUMPD2 CRAC datasets using peak-calling and feature overlap correction.

Fig. 1f. Extended western blot analysis of proteins co-immunoprecipitated with THUMPD2-His₆-2xFLAG with and without UV crosslinking.

Fig. 3i. Additional replicate northern blots to monitor association of wild-type and catalytic mutant THUMPD2 with the U6 snRNA.

Fig. 4o-p. *In vitro* methylation assays using THUMPD2-TRMT112 with stabilized U6 ISL substrate.

Fig. 5a-b. UpSet analysis of alternative splicing events in wild-type, THUMPD2 KO, LARP7 KD and THUMPD2 KO + LARP7 KD using new mRNA-seq data.

Fig. 5c-e. Sashimi plots showing examples of different types of alternative splicing events regulated by THUMPD2, LARP7 or both.

Fig. 5f. Distribution of significant and not significant inclusion differences for different AS types and conditions.

Fig. 6e. Gradient-based analysis of snRNA levels in snRNP complexes in cells lacking LARP7 or THUMPD2 and LARP7.

Fig. 6f. Quantification of snRNA levels in Fig. 6e.

Fig. 6g. Northern blot analysis of snRNA levels in nuclear extracts from wild-type, THUMPD2 KO, LARP7 KD and THUMPD2 KO + LARP7 KD cells.

Fig. 6i-l. Western blot analysis of PRPF3 (i), PRPF4 (j), PRPF8 (k) and EFTUD2 (l) migration profiles in gradients separating nuclear extracts from wild-type, THUMPD2 KO, LARP7 KD and THUMPD2 KO + LARP7 KD cells.

Supplementary Fig. S1a. Western blots showing recovery of THUMPD2-His₆-2xFLAG during immunoprecipitation.

Supplementary Fig. S1b. Autoradiograph showing RNA fragments crosslinked to THUMPD2-His₆-2xFLAG.

Supplementary Fig. S3f-g. *In vitro* methylation assays using THUMPD2-TRMT112 and U6 ISL stabilized by nucleotide substitutions.

Supplementary Fig. S4a. Updated PCA analysis of new mRNA-seq data.

Supplementary Fig. S4b. Scheme of different types of alternative splicing.

Supplementary Fig. S4c. Updated heatmaps of significant alternative splicing events.

Supplementary Fig. S4d. Analysis of 5' and 3' splice site strength of retained introns.

Supplementary Fig. S4e. Cumulative IR fraction distribution of introns that change pre-mRNA splicing significantly.

Supplementary Fig. S6a. Analysis of U6atac in snRNPs in WT & THUMPD2 KO cells.

Supplementary Fig. S6b. U6atac levels in nuclear extracts from WT & THUMPD2 KO cells.

Supplementary Table S1. Output values of CRAC sequencing read alignments.

Reviewer 1

Reviewer: “In this manuscript, the authors have biochemically and functionally characterized the enzyme complex that m2G-modifies the U6 snRNA at position 72. Although this link has been made before, the authors show that this is the only modification that THUMPD2 generates. THUMPD2 is enriched in Cajal bodies, it associates with the adaptor protein TRMT112 and binds to U6 at later stages of U6 RNP assembly. The authors further use crosslinking experiments to map the interaction of THUMPD2 to U6 and TRMT112. U6 is contacted at multiple regions in the N and the C terminal part of the protein, particularly the THUMP and the MTase domain. TRMT112 was mapped to the C terminus of the protein. They further generated a DNAzyme that allows for the site-specific identification of m2G72. THUMPD2 knock out cells clearly show reduced modification at this position and complementation assays revealed that SAM binding and TRMT112 binding is required for methylation in cells. However, in vitro, the reaction was not efficient and therefore the authors investigated the contribution of associated 2'-O-methylation events. Indeed, pre-2'-O-methylated substrates were efficiently modified by the reconstituted enzyme complex in vitro. Using LARP7 knock down, which globally reduced U6 2'-O-methylation, the authors further show that prior U6 2'-O-methylation is also required in cells. Finally, the authors find that alternative splicing events are affected by THUMPD2 knock out, LARP7 knock down and a stronger effect when both are inactivated simultaneously. This appears to be due to less efficient di- and tri-snRNP formation.

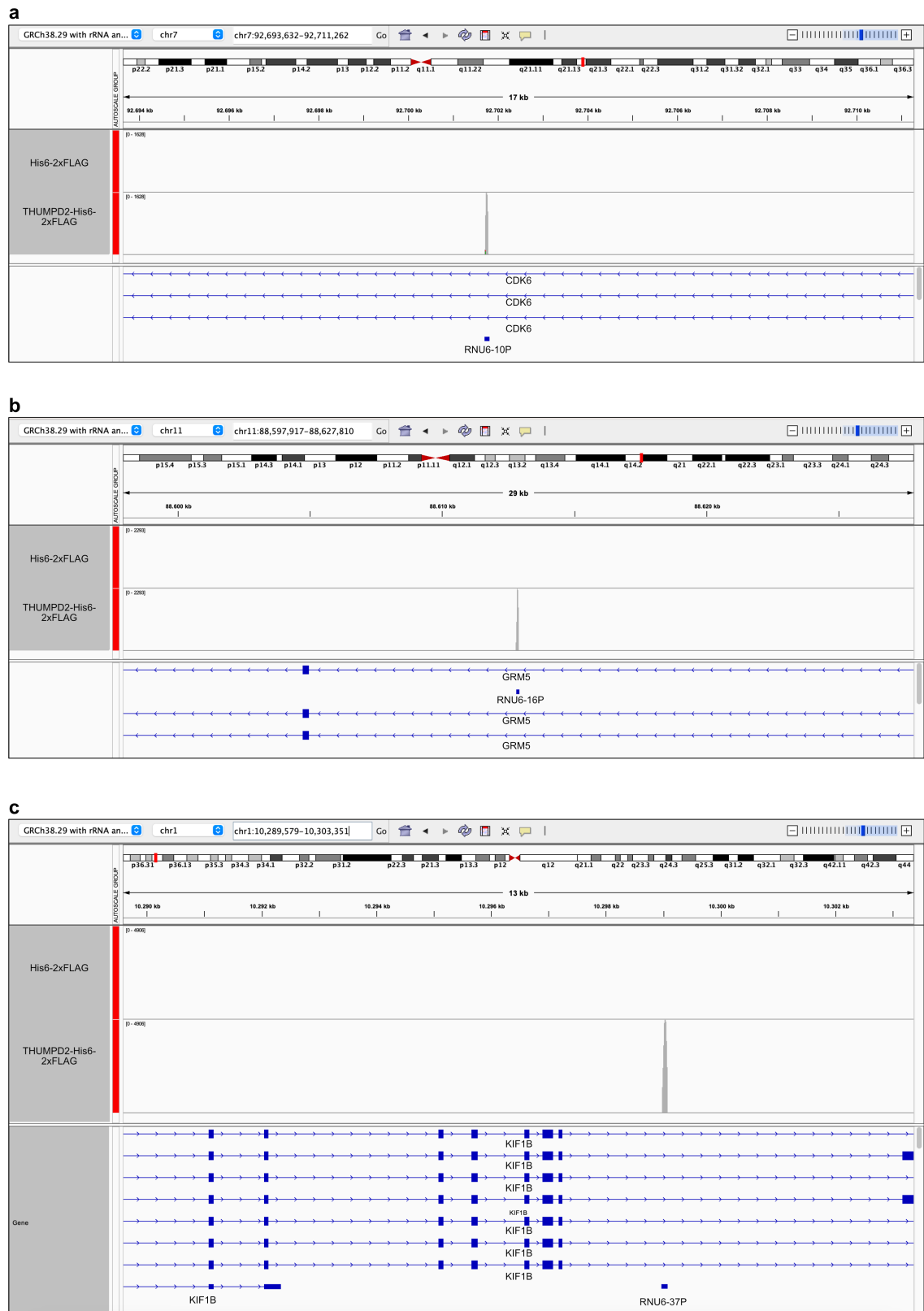
This is a very solid and clear manuscript. All experiments are competently carried out, presented and interpreted. Although the link between THUMPD2 and U6 m2G72 has already been made, the manuscript goes far beyond and presents a thorough characterization of the entire pathway including experiments in cells, generation of new tools such as DNAzymes as well as biochemical in vitro reconstitution of the enzymatic activities. I have only a few comments that could be considered.”

Reply: We thank the reviewer for the positive feedback on our work and manuscript.

Reviewer: “1. Fig. 1A: about 60% of the reads are crosslinked to lncRNAs or mRNAs (which is also seen in the negative control). It is clear that no modification sites besides the one in the ISL of U6 were found. However, THUMPD2 may bind to other transcripts without methylating them. Have some of the mRNAs or lncRNAs been validated to clearly show that this is indeed background noise of the CRAC protocol?”

Reply: Upon closer inspection of the distribution of sequencing reads mapping to protein-coding genes in the THUMPD2 CRAC datasets, we observed that most mapped to introns that contain U6 sequences. This overlap is illustrated by the example genome coverage traces (IGV snapshots) for three mRNAs with high numbers of reads (>800 reads per gene; **Reply Fig. 1**). In the previous bioinformatics analysis, which monitored the normalized (RPM) numbers of reads mapping to different genes, reads mapping to these genes were attributed to mRNAs, whereas they actually reflect reads derived from U6.

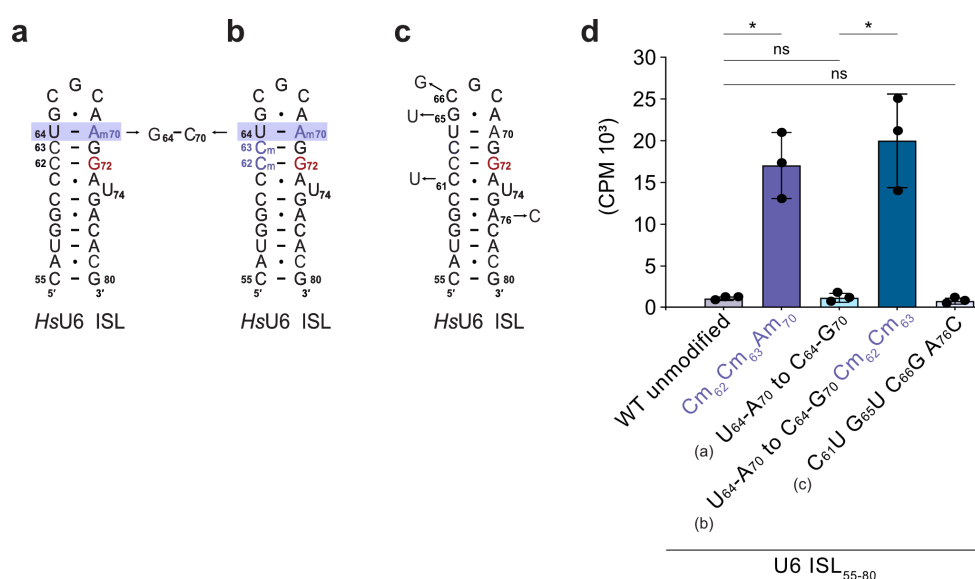
To correct this mis-assignment, we re-analysed the CRAC datasets using peak-calling, and only reads mapping to genomic regions (peaks) significantly enriched in both the THUMPD2 datasets compared to the controls were used for feature count analysis considering “exons” rather than “genes”. For both datasets, this demonstrated that the vast majority of sequencing reads in the THUMPD2 dataset were derived from the U6 snRNA. A degree of variability between the two datasets was observed (80% of the mapped reads significant after peak calling were mapping to U6 in replicate 1 and 61% in replicate 2), likely reflecting technical differences during the library preparation, so instead of averaging the data, the two datasets are rather presented independently in **Fig. 1a** and **Source Data Fig. 1a**.



Reply Figure 1. Mapping profiles of THUMPD2-His₆-2xFLAG and His₆-2xFLAG CRAC data on selected genes. **(a)** Reads mapping to the *CDK6* intron overlapping with *RNU6-10P*. **(b)** Reads mapping to the *GRM5* intron overlapping with *RNU6-16P*. **(c)** Reads mapping to the *KIF1B* intron overlapping with *RNU6-37P*.

Reviewer: “2. Figure 4: 2'-O-methylation of the upper stem should lead to a stronger pairing and thus a more stable stem. To further test this, the first A-U pair (I guess 64) could be swapped to a G-C pair, which might overrule the strict requirement of a 2'-O-methylation at this position. It is also interesting that the system seems to be not optimal since pairing of position 65 generates a much better substrate *in vitro*. What would be the evolutionary selection against such an optimal substrate? This could be further discussed.”

Reply: *In vitro* methylation assays were performed on an RNA substrate in which the U₆₄-A₇₀ pair at the top of the U6 ISL was exchanged for a C₆₄-G₇₀ pair. This oligonucleotide was not methylated by THUMPDP2-TRMT112 under the conditions used (**Reply Fig. 2**). However, similar to the RNA containing the endogenous U₆₄-A₇₀ basepair, addition of 2'-O-methyl groups to C₆₂ and C₆₃ to the substrate containing the C₆₄-G₇₀ pair was sufficient to stimulate THUMPDP2 activity *in vitro* (**Fig. 4f** and **Reply Fig. 2**). To further explore the influence of the stability of the U6 ISL on m²G₇₂ installation by THUMPDP2-TRMT112, *in vitro* methylation assays were performed using a version of the U6 ISL containing four nucleotide substitutions stabilizing the stem (C₆₁U, G₆₅U, C₆₆G and A₇₆C). These changes create two A-U and two G-C basepairs where there are non-conventional basepairs in the wild-type sequence. This oligonucleotide was also not a substrate of THUMPDP2-TRMT112 *in vitro* (**Fig. 1**).



Reply Figure 2. *In vitro* methylation assay with THUMPDP2-TRMT112 and U6 ISL substrates containing nucleotide exchanges and modified/unmodified nucleotides. (a-c) Schematic views of the RNA substrates used for *in vitro* methylation assays. (d) *In vitro* methylation assays were performed using purified His₆-ZZ-BtTHUMPDP2-BtTRMT112 protein complex, [³H]-SAM as methyl group donor and the specific RNA substrates indicated in (a-c). Tritium incorporated in the substrate RNAs was quantified and bar plots show averaged counts per minute (CPM) of n=3 independent experiments with error bars representing the standard deviation. Significance was calculated using unpaired t-tests. *p < 0.05 and ns = not significant.

We conclude, therefore, that the stability of the stem does not overcome the requirement for the Cm₆₂ and Cm₆₃ 2'-O-methylations, but that certain mutations stabilizing the upper part of the stem increase *in vitro* methylation of U6-G₇₂ in the presence of Cm₆₂ and Cm₆₃. The *in vitro* methylation data for the maximally stabilized U6 ISL are presented in the revised manuscript in **Fig. 4p**.

The evolutionary selection against a U6 ISL with a more stable stem is likely a result of the necessity for U6 to assume multiple conformations during its assembly and integration into spliceosomes. In the U6 snRNP, nucleotides 55-80 fold into an intramolecular stem-loop, whereas in the U4-U6 di-snRNP, this helix is completely unwound to basepair with U4 and, in the active spliceosome, this sequence folds back into an intramolecular helix when U6 is base-paired with U2.

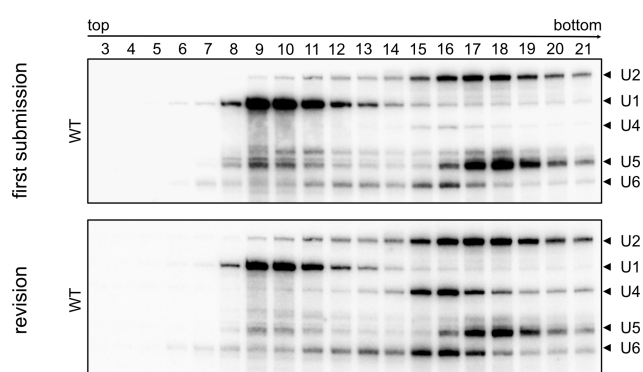
Mutational analysis of U6 transcripts used in experiments with human nuclear extracts show that some nucleotide substitutions that (over-)stabilize the U6 ISL can be detrimental for U4-U6 interaction, snRNP assembly and splicing. For instance, A₇₃G substitution in the upper part of the U6 ISL stabilizes the C₆₁-A₇₃ pair immediately upstream of the bulged U₇₄ and was shown to reduce U4-U6 interaction and spliceosome assembly (PMID: 8330741). Similar experiments also reveal that A₇₆C, which stabilizes the G₅₉-A₇₆ pair close to the catalytic U₇₄ leads to loss of most of the U6-U4 interaction *in vitro*, significantly reducing spliceosome assembly and splicing activity (PMID: 7632731). Additionally, an A₅₆G substitution that stabilizes the A₅₆-C₇₉ pair in the lower stem region of the U6 ISL also reduces U6-U4 interaction (PMID: 7632731). In yeast, it has also been shown that stabilizing the upper loop of the U6 ISL via substitution of C₇₂ (equivalent to C₆₆ in the human U6 ISL) for A₇₂ contributes to a strong cold-sensitive phenotype, likely arising due to perturbed pre-mRNA splicing (PMID: 31229405). Data from yeast also suggests that hyperstabilizing U6 ISL structures can interfere with the catalytic conformation of the U6 ISL (PMID: 31229405). For example, a C₆₇A substitution that allows base-pairing with U₈₀ (equivalent to U₇₄ in the human U6 ISL) is proposed to inhibit docking of U₈₀ into the catalytic triplex, stabilizing a conformation of the ISL that is not properly functional (PMID: 31229405). Additionally, cells expressing an A₇₉G mutant that stabilizes the C₆₇-A₇₉ pair in close proximity to the catalytic site exhibit a cold-sensitive phenotype, suggesting that formation of a rigid structure around the catalytic site can hinder interactions of U₈₀ with the catalytic triplex (PMID: 31229405).

Furthermore, it has been shown that, across different species, the upper ISL segment contains several conserved C-G pairs, while the lower ISL segment has a higher sequence variability, containing frequent weak and non-canonical base-pairings that confer higher ISL flexibility (PMID: 31229405). The conservation of these features suggests that they ensure an optimal balance between stability and flexibility that allows the required rearrangements of this region during the transitions between the different steps in snRNP assembly and

splicing. This topic is now addressed more extensively in the discussion section of the manuscript.

Reviewer: “3. The gradient shown in Fig. 5 are not fully clear. First, U4 is hardly visible and maybe a different probe could be used. U snRNAs are also well stained by silver-staining protocols and since these are the major RNAs in nuclear lysates, this might be enough to visualize them. The authors could perform the gradients under THUMPD2 and LARP7 depletion conditions. This may enhance the effect and it may become clearer in these assays.”

Reply: The previously poor detection of U4 relative to the other snRNAs did indeed hamper interpretation of the gradients, so the northern blot probe used to detect U4 was changed and the amounts of each probe used were titrated so all five snRNAs can be visualised well and with similar signal intensity. **Reply Fig. 3** shows the difference before and after optimizing snRNA detection in the gradients by northern blotting. All northern blots presented in the updated version of **Fig. 6** (and the replicates shown in **Source Data Fig. 6e**) have been re-probed using the optimized approach.



Reply Figure 3. Optimization of northern blotting to detect snRNAs. Representative northern blots show the detection of snRNAs of the major spliceosome using the previous strategy (upper), and with an alternative U4 probe and titrated levels of each probe (lower) in the same membrane.

The gradient-based analysis of the effect of THUMPD2 absence on snRNP complex assembly and integration into the spliceosome has been extended to include the depletion of LARP7 and the combination of LARP7 KD in the THUMPD2 KO cell line. It is important to note that due to the requirement of prior 2'-O-methylation of C₆₂ and C₆₃ for efficient THUMPD2-TRMT112-mediated methylation (**Fig. 4h-i**), in both these conditions, m²G₇₂ is reduced. Interestingly however, we observe that lack of m²G₇₂ alone and the reduction of 2'-O-methylations together with reduced/absent m²G₇₂ have different effects on U6 snRNA levels in snRNPs. Lack of m²G₇₂ alone leads to accumulation of the U6 snRNA in early

gradient fractions, indicating this modification is important for the progression of U6-containing di-/tri-snRNP complexes into larger spliceosomal complexes, whereas upon depletion of LARP7, a strong reduction in U6 incorporation into snRNPs is observed. This suggests that the U6 2'-O-methylations are functionally important prior to the m²G₇₂, and that phenotypically, the lack of U6 2'-O-methylations caused by LARP7 KD therefore predominates over the absence of m²G in the THUMPD2 KO.

To further consolidate our conclusions about the effect of lack of m²G on spliceosomal complex assembly, we have also probed our gradients for additional di-snRNP and tri-snRNP markers. Both PRPF4 and PRPF3 show marked differences in their migration profiles between WT and THUMPD2 KO, supporting sub-optimal progression of U6 snRNA-containing snRNPs into spliceosomal complexes due to lack of m²G₇₂.

These new data are presented in an updated **Fig. 6** and the effects of the different modifications are discussed in the revised manuscript.

Reviewer 2

Reviewer: “In the manuscript by Bohnsack et al, the authors describe their systematic and structured approach to deciphering the installation of m²G in U6 snRNA and its biological consequences. I am impressed by the variety and at the same time high quality of the methods used by the whole team to answer important questions in the field of spliceosome assembly. The authors have used various and orthogonal methods and convincingly present data for the importance of m²G and its Nm interdependence in U6 snRNA. I see no need for additional experiments. Form my perspective everything is covered. Yet, the Materials and Methods sections is not complete and requires careful reading and addition of information to ensure reproducibility.”

Reply: We thank the reviewer for the positive feedback on our work and the suggestions to improve the manuscript.

Reviewer: “Quantification of U6 snRNA modifications by LC-HRMS” is lacking detail. E.g. the digestion enzymes are not mentioned, columns supplier is not mentioned, which timsTOF was used, flow rate of chromatography, the description of the gradient is unclear. The IP-MS data analysis section is more clear, but still misses some details, e.g. column information is missing, the injection volume, was a trap used?”

Reply: We have added the missing details to the LC-HRMS and IP-MS method sections. Please note that the timsTOF used is a first generation machine, so is only named “timsTOF”, without the additional differentiations of next generation machines.

Reviewer: “Further, I have found the manuscript difficult to read and digest. Therefore, I suggest substantial editing of the main text and the discussion. I have made a collection of instances in which I struggled in hope that the manuscript can be advanced and more easy to parse.”

Reply: We thank the reviewer for highlighting instances where sentences can be improved and have incorporated the suggestions pointed out below. In the light of this comment and to conform to editorial length requirements, the text has been substantially edited. Changes introduced for text compaction, which do not alter the meaning or conclusions, are not highlighted in red.

Reviewer: “Some sentences require English editing. E.g. line 157 ff and line 471 ff.”

Reply: The sentences on lines 157 and 471 have been removed in the context of describing the updated analysis of the CRAC data and the new *in vitro* methylation data.

Reviewer: “line 252: The current title is confusing and I suggest changing it to “Predicted RNA contacts localize to the THUMP and MTase domains” or something similar to emphasise the modelling character of the section”

Reply: The RNA contact sites in THUMPD2 were detected experimentally using UV crosslinking coupled to mass spectrometry to identify amino acids physically associated with RNA nucleotides. As the U6 snRNA is likely the sole target of THUMPD2 (**Fig. 1a**) and UV crosslinking is considered a “zero-distance” approach, the crosslinking sites identified can be considered as relevant RNA contact sites of THUMPD2 with U6. The sub-heading has been updated to “RNA contact sites on THUMPD2 are present in the THUMP and MTase domains”. The modelling approach used to generate the structure on which these contact sites are depicted in **Fig. 3c** is described in detail in the text.

Reviewer: “line 376: The sentence “Bovine proteins were used as HsTHUMPD2 expressed poorly” is unclear. If I understand correctly, the authors expressed the bovine orthologs of THUMPD2 (and TRMT112?) because recombinant expression of the human protein was inefficient. This should be clarified explicitly to avoid misinterpretation.”

Reply: It is indeed the case that human THUMPD2-TRMT112 expressed poorly in *E. coli* and for this reason, the bovine proteins were used instead. This has now been clarified in the text.

Reviewer: “line 385 ff: The sentence spanning lines 384–389 is overly long and difficult to follow. I recommend breaking it into shorter sentences. The same is true for lines 252 ff, 276 ff, 333 ff, 422 ff and 511 ff.”

Reply: The highlighted sentences have been amended to aid readability.

Reviewer: “The discussion is more like a recap or summary. Interpretation is limited and missing or weakly developed. Method limitations (e.g. structural modeling and DNAzyme sensitivity and its indirect character) are not discussed.”

Reply: The discussion section has been revised to describe the methodological approaches in more detail and include further interpretation of the data.

Reviewer: “N2,2-7-trimethylated cap (m3G): this nucleotide is commonly abbreviated as m227G according to modomics”

Reply: The abbreviation used in the Modomics database has been added.

Reviewer: “I am confident that the work presented in this manuscript is a great fit for the broad readership of Nature Communications and I recommend publication after text revision.”

Reply: We thank the reviewer for recommending our work for publication in Nature Communications.

Reviewer 3

Reviewer: “The manuscript by Kleiber et al. titled “2'-O-methylation-dependent installation of N2-methylguanosine in the U6 internal stem loop facilitates efficient spliceosome assembly” investigates how the m72G modification of the U6 snRNA is established by the THUMP2 enzyme and its reliance on the prior installation of 2'O methyl modifications on U6 snRNA by LARP7.

The manuscript has three major claims: 1- THUMP2 associates with U6 snRNA during late stages of its biogenesis; 2- m2G72 methylation requires 2'O methylation of U6 snRNA; and 3- m2G72 and 2'O methylations of U6 affect splicing interdependently. Overall, the manuscript is well written with only minor writing errors and a few missing citations. The figures are well prepared and easy to understand. The materials and methods sections are detailed enough to reproduce the experiment. The manuscript makes a strong case for most of the claims and provides robust experimental data that clearly demonstrate the function of 2'O methylations in THUMP2 activity in vitro. It shows that the m72G modification is likely a late-stage modification in U6 biogenesis. The authors demonstrate that the absence of m72G likely affects spliceosome assembly and that the absence of THUMP2 and LARP7 impacts pre-mRNA splicing. The claims regarding the effects on pre-mRNA splicing could be supported more by additional analysis, as currently the effect of 2'O methylations is only tested in vitro. Nevertheless, overall, I believe this manuscript presents

important data that will help the field better understand THUMPD2 function and overall U6 snRNA biology.”

Reply: We thank the reviewer for the positive feedback on our work and manuscript.

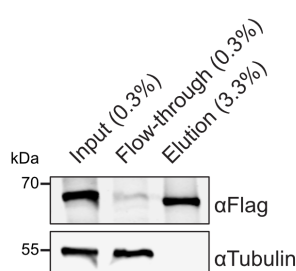
Major comments

Reviewer: “1- Line 149 – specify which human cells are over-expressing and note that they are over-expressing.”

Reply: The cell type used is now mentioned in the Results section as well as the fact that THUMPD2-His₆-2xFLAG is over-expressed.

Reviewer: “2- Line 149, I couldn’t find any evidence or data showing this immunoprecipitations, its efficiency and its specificity in main or supplemental figures. It is important to demonstrate that protein and RNA immunoprecipitation are working effectively and specifically through western blots. Are the authors immunoprecipitating only a small percentage of the tagged protein or most of it? Similarly, how can they show that UV crosslinking is working effectively?”

Reply: A western blot showing the efficiency of the anti-FLAG immunoprecipitation of THUMPD2-His₆-2xFLAG has been included as **Supplementary Fig. S1a**. After quantification of the inputs and eluates in n=3 experiments, an average immunoprecipitation efficiency of 37% was obtained for wild-type THUMPD2-His₆-2xFLAG. Further supporting an adequate recovery of THUMPD2-His₆-2xFLAG in the immunoprecipitation experiments, very little FLAG-tagged protein was left in the flow-through (**Reply Fig. 4**). Uncropped western blots for immunoprecipitations are now provided in **Source Data Fig. 1f** and **Fig. 3i**.

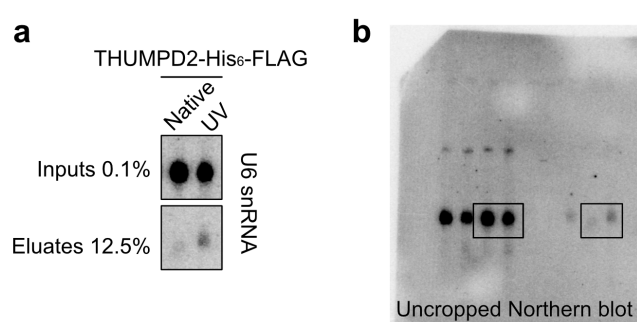


Reply Figure 4. Immunoprecipitation of THUMPD2-His₆-2xFLAG. Input, flow-through and eluate samples from an anti-FLAG immunoprecipitation experiment performed as in Fig. 1f and Fig. 3i were analyzed by western blotting using the indicated antibodies.

It is known that the RNA–protein crosslinking efficiency using UV is naturally very low (PMID: 35581222). Nevertheless, the strong enrichment of the U6 snRNA with THUMPD2-His₆-2xFLAG in CRAC and RNA-IP experiments (**Fig. 1a**, **Supplementary Fig. S1b** and **Fig. 3i**),

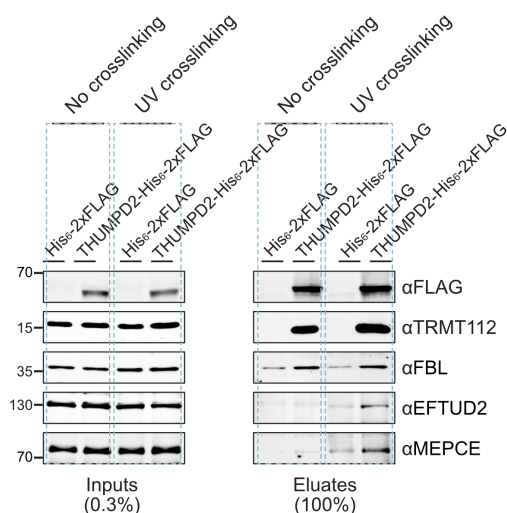
which include a purification step under highly denaturing conditions (6 M guanidine hydrochloride), demonstrates the formation of crosslinked RNA–protein complexes.

The effectiveness of UV crosslinking in stabilizing the THUMP2–U6 snRNA interaction is demonstrated by comparing the amounts of U6 recovered with THUMP2-His₆-2xFLAG in immunoprecipitation experiments performed with and without prior UV crosslinking of cells. In the IP performed without prior crosslinking (native), low amounts of U6 snRNA were recovered whereas after UV crosslinking, the U6 signal obtained by northern blotting was markedly higher (**Reply Fig. 5**). Furthermore, the U6 signal detected in the IP eluate from the crosslinked cells is shifted upwards relative to U6 recovered from the non-crosslinked cells, which is consistent with the formation of covalently crosslinked complexes. These data are now presented in **Supplementary Fig. S1c** of the revised manuscript.



Reply Figure 5. Immunoprecipitation of THUMP2-His₆-2xFLAG from UV crosslinked and non-crosslinked cells. Immunoprecipitation experiment with THUMP2-His₆-2xFLAG (performed as in Fig 1f), showing recovery of the U6 snRNA without (native) or with UV crosslinking (UV) by northern blotting. In (b) the uncropped Northern blot for (a) is shown.

The stabilizing effect of UV crosslinking on THUMP2-containing complexes is further emphasized by the comparison of proteins recovered in immunoprecipitation experiments with and without UV crosslinking as shown in **Fig. 1f**. IP eluates analyzed in parallel on the same membranes (**Reply Fig. 6**, right panel) show that although TRMT112 and FBL are recovered similarly in both IPs, EFTUD2 and MEPCE are more efficiently recovered after UV crosslinking.



Reply Figure 6. Comparison of proteins recovered with THUMPDP2-His₆-2xFLAG in IP experiments with and without prior UV crosslinking. Extracts from cells expressing His₆-2xFLAG or THUMPDP2-His₆-2xFLAG that were left untreated or UV crosslinked were used for anti-FLAG immunoprecipitation experiments. Inputs (left panel) and eluates (right panel) were analyzed by western blotting using the indicated antibodies.

Reviewer: “3- Similarly, Table S1 shows the overall reads mapped to different classes of RNAs or specific snRNAs, but it lacks information on the total reads (unmapped) from each experiment, the percentage mapped, read lengths, and so on.”

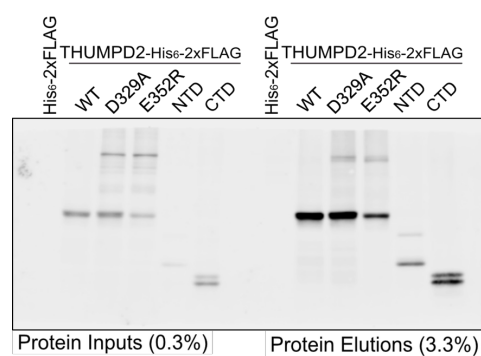
Reply: Parameters including number of input reads, average read length, percentage of uniquely mapped reads are now presented in **Supplementary Table S1**.

Reviewer: “4- Line 274, the main claim of the manual modelling is that it performs better than the Alphafold-generated model. Therefore, can the authors be more specific than “close proximity to THUMPDP2”? And following from this, why isn’t G72 within the active site?”

Reply: As initially indicated in the Material and Methods section, this manual model was obtained by superposing Alphafold models of the THUMPDP2-TRMT112 complex, previously determined structures of the U6 snRNA ISL and a complex between a THUMP domain bound to a double stranded RNA. It can be considered, therefore, as a rigid body model. In this model, G₇₂ is facing the THUMPDP2 active site but is engaged in a Watson-Crick base pair with C₆₂ of the U6 ISL. A simple flipping-out mechanism, commonly observed in structures of double-stranded RNAs (in particular tRNAs and rRNAs) bound to RNA modifying enzymes, could project the G₇₂ base into the THUMPDP2 active site, in an ideal position for N2 of G₇₂ to be opposite the SAM methyl group to be transferred. It was decided not to model this flipped-out G₇₂ base into the THUMPDP2 active site to limit over-modeling. However, the proposed mechanism is described in the revised manuscript.

Reviewer: “5- Figure 3i, similar to an earlier point about IP efficiency, makes it difficult to assess how effective their IP is since the input bands at 0.3% are stronger than the IP at ten times higher loading. The RNA binding data also suggest that THUMPDP2 binds only a very small fraction of U6 at any given time. Is that correct? Does the D329A mutant retain U6 longer?”

Reply: Uncropped western blots of inputs and eluate samples from immunoprecipitations show the relative enrichment of wild-type THUMPDP2-His₆-2xFLAG and the mutant versions and allow the IP efficiency to be determined (**Fig. 3i**, **Source Data Fig. 3i** and **Reply Fig. 7**).



Reply Figure 7. Immunoprecipitation of wild-type THUMPDP2-His₆-2xFLAG and mutant versions. The relative expression levels of wild-type THUMPDP2-His₆-2xFLAG and its derivatives in cells are shown (input) as well as their recovery after anti-FLAG immunoprecipitation (eluate). Protein samples were analyzed by western blotting using an anti-FLAG antibody.

The IP efficiency (%) was calculated in n=3 experiments by quantifying eluate and input bands in the relevant blots (**Reply Table 1**). In all cases, at least a third of the bait protein was recovered in the eluates.

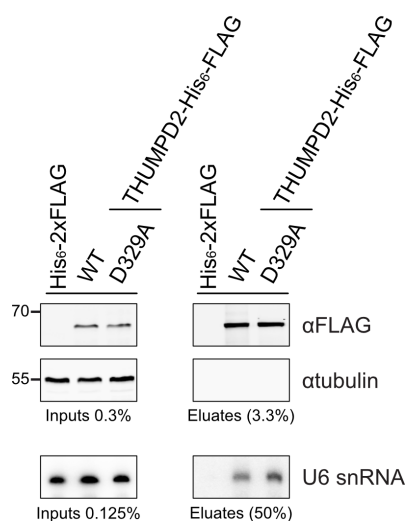
Reply Table 1. Quantification of IP efficiency of wild-type THUMPDP2-His₆-2xFLAG and its derivatives. Values reflect (eluate*0.3)/(input*3.3)*100.

	Rep. #1	Rep. #2	Rep. #3	Average
THUMPDP2	47.0	33.5	32.4	37.6
D329A	40.2	34.9	43.3	39.5
E352R	33.4	32.6	30.8	32.2
NTD	85.0	92.2	33.9	70.4
CTD	33.1	16.3	41.7	30.4

It is indeed the case that only a small proportion of the U6 snRNA is recovered with THUMPDP2-His₆-2x-FLAG. The U6 snRNA has an estimated half-life of 24 h (PMID: 8892983) and our data suggest that THUMPDP2 associates with nascent U6 during its maturation in the nucleoplasm/Cajal bodies (**Fig. 2**), likely dissociating prior to incorporation of the U6 snRNP into di-snRNP complexes. The transient nature of the interaction of THUMPDP2 with the U6 snRNAs is supported by the finding that TRMT112 is the only robust interaction partner of THUMPDP2 identified by IP-MS after crosslinking (**Fig. 1e**). Most of the RNA detected in the

input samples probably represents U6 snRNA incorporated into snRNPs and other spliceosomal complexes whereas only a small proportion of newly synthesized U6 undergoing maturation is captured in association with THUMPD2. The low recovery of U6 snRNA with THUMPD2 is rationalized in the revised manuscript.

Considering the relative recovery of U6 with wild-type THUMPD2 and the catalytic mutant (D329A), it is not possible to accurately quantify the amount of U6 co-precipitated relative to the bait protein as these samples require to be analyzed by northern and western blotting, respectively (i.e. separate gels/membranes). To allow better estimates of the relative enrichment of U6, this experiment has now been performed n=5 times. Considering all data, it appears that slightly higher amounts of U6 are usually recovered with the THUMPD2 mutant compared to the wild-type protein. The data from all experiments are shown in **Source Data Fig. 3i**, and another representative blot (in addition to **Fig. 3i**) is shown below (**Reply Fig. 7**). The relevant section of the text has been amended to reflect this observation.



Reply Figure 7. RNA-IP of U6 with wild-type and catalytically inactive THUMPD2. Extracts from UV crosslinked cells over-expressing His₆-2xFLAG, THUMPD2-His₆-2xFLAG and THUMPD2(D329A)-His₆-2xFLAG were used in anti-FLAG immunoprecipitation experiments. Proteins in inputs (left panel) and eluates (right panel) were analyzed by western blotting using the indicated antibodies (upper), and U6 was detected by northern blotting (lower).

Reviewer: “6- Figure 4 provides a strong argument for the roles of 2’O methylation in G72 methylation by THUMPD2. Additionally, in this figure, the authors demonstrate the use of a novel dnzyme-based method for detecting 2’O methylations. This is really good.”

Reply: We thank the reviewer for appreciating the value of the developed DNAzymes to detect m²G and 2’-O-methylations in a site-specific manner.

Reviewer: “7- In Figure 5, authors make a good start to splicing analysis, but overall this section of the manuscript is weak. It was already known that the absence of THUMPD2 and

LARP7 affects splicing to some extent, and authors' data confirm this, showing there may be a stronger effect when both are missing. However, this does not tell us much about the specific role these modifications play in pre-mRNA splicing or why some genes are affected while others remain unaffected."

Reply: In the initial submission, the analyses of alternative splicing in cells lacking THUMPD2 and/or LARP7 were performed on total RNA-seq datasets (with rRNA depletion). To improve our comparative alternative splicing analyses, poly(A) enrichment and mRNA-seq was now performed to enhance the depth of the sequencing. This increased the number of significant alternative splicing events detected in all conditions, thereby allowing to consolidate the conclusions and probe further features of these datasets. Consistent with the observations that particular 2'-O-methylations in the ISL promote installation of m²G₇₂ (**Fig. 4**), but that lack of U6 2'-O-methylations and m²G₇₂ differently influence snRNP assembly (**Fig. 6**), we detect numerous alternative splicing events that are independently and interdependently affected by lack of THUMPD2, LARP7 or both (**Fig. 5; Supplemental Fig. S4**). The scope of our analysis of alternative splicing extends to include: i) identification of different types of alternative splicing events in all datasets and comparative analysis of those significant in different conditions with an improved visualization using UpSet plots, ii) highlighting three examples, representing the intron retention upon THUMPD2 KO, alternative 3' splice site variation upon LARP7 KD and a skipped exon event that is significantly different from WT, iii) narrowing down types of AS that are specifically affected either upon THUMPD2 KO (intron retention) or LARP7 KD (A3SS), and iv) addressing the combinatorial nature of the two perturbations. Related to i), we modified the analysis for intron retention to include more introns beyond what is possible with rMATS. We quantify intron retention using split reads and exon-intron junction reads for 37885 introns that pass our read count cutoff in all conditions. As expected most introns are well spliced. This is visualized in the new **Supplemental Fig. S4e**. Introns that show a significant increase or decrease in splicing for any of the perturbations tend to be already less efficiently spliced (**Fig. 5c; Supplemental Fig. S4e**). For intron retention, we observe that THUMPD2 KO decreases splicing-efficiency. To visualize all types of AS, we moved away from using Venn diagrams to UpSet plots that better capture the intersections. A3SS, A5SS, MXE and SE are grouped together as they all originate from the AS analysis using rMATS.

We modified the text, figures, tables, as well as the GEO submission to reflect these adjustments.

With these analyses, we extend our previous conclusions, which mainly focused on the increase in significant events for the combined perturbation, by highlighting the importance of LARP7/U6 2'-O-methylation for 3' splice site selection, which is distinct from THUMPD2/m²G₇₂ that affects intron retention in a directed manner. A modest correlation of

5' and 3' SS scores with intron retention was detected (**Supplemental Fig. S4d**), but no sequence features associated with the increase in distal A3SS usage upon LARP7 KD were observed. Nevertheless, the identified distinction between the effects of the different types of modifications is a good basis for further analyses in the future, e.g. with regard to possibly altered co-transcriptional aspects of spliceosome assembly.

Reviewer: “8- Figure 5i-k provides valuable insights into the mechanistic understanding of how these modifications influence spliceosome assembly. However, given its current format and the information provided, it is challenging to determine whether the conclusions are fully supported by the data. For example, in 5i, not only does U6 accumulate in fractions 10-17, but there is also a significant difference in U6 levels. Nevertheless, the authors show in figure s4 that the levels are the same. Additionally, there is little observable change in U4 or U5 levels or fractionation profiles. Furthermore, compared to Mabin et al., who performed similar fractionation, why is free U6 so low (PMID: 34234030)? Lastly, if less U6 is available for tri-snRNP assembly, why is the effect on splicing not more pronounced, impacting all transcripts and introns?”

Reply: The 10-30% glycerol gradients performed under the conditions mentioned in the manuscript (34,000 rpm for 18 h in a SW40 rotor, using 100 mM KCl) allow resolution of snRNPs such that the 12S (U1 snRNP), 17S (U2 snRNP) and 25S (U6:U4.U5 snRNP) complexes are very well separated. However, these gradients do not provide information on the levels of the snRNAs in the spliceosomal complexes (A-complex, B-complex, etc.) as this would require very different conditions. While the overall levels of U6 are not affected by THUMPD2 KO (**Fig. 6g**), the increased U6 levels in fractions 10-17 represent an accumulation of U6 in the complexes migrating in these fractions. Accumulation of U6 in the mono-/di-/tri-snRNP complexes present in fractions 10-17 in the THUMPD2 KO indicates that less U6 progresses into later spliceosomal complexes. Although the accumulation of U6 in these fractions in the THUMPD2 KO is more pronounced than for U4 and U5, the levels of these snRNAs in these fractions also increase. After improving the detection of U4 snRNA by exchanging the northern blotting probe, the accumulation of this snRNA becomes more apparent (**Fig. 6e-f**). Interestingly, we observed that depletion of LARP7 (see point 10) similarly has a stronger effect on the amount of U6 present in these complexes than U4 or U5 (**Fig. 6e-f**). The description of these results has been improved to clarify the conclusions drawn from the blots.

In the study by Mabin et al., the gradients were performed using extracts from 293T cells and more free U6 (see Figure S7D of PMID: 34234030; fractions 5-7) was detected than in our gradients, which were performed with nuclear extracts from HCT116 cells. It is possible that the differences in levels of free U6 arise due to variations between the cell lines. For example,

in Tanackovic et al., similar gradients performed using extracts from lymphoblasts (see Figure 2 of PMID: 21378395) show low amounts of free U6. However, we can not exclude the possibility that lesser U6 incorporated into di-/tri-snRNPs could be explained by differences in the preparation of the extracts or gradients. For comparison, in Yildirim et al., glycerol gradients also performed with extracts from HEK293T cells (see Figure S5A of PMID: 34023904; fractions 5-7) show less free U6, similar to the results we obtained.

As described above, our analyses of alternative splicing in THUMPD2 KO cells in the first submission were based on total RNA-seq (ribosomal RNA depletion). Upon increasing the sequencing depth by performing mRNA-seq (polyA-enrichment), we observe many more alternative splicing events that are significantly affected by loss of THUMPD2 (**Fig. 5**). The accumulation of U6 in di-/tri-snRNP complexes in the absence of m²G₇₂ (**Fig. 6**) suggests that this modification optimizes the maturation of these complexes and their integration into larger splicing assemblies. The mild effects on splicing observed in the THUMPD2 KO are consistent with perturbation of spliceosome assembly kinetics rather than a block in tri-snRNP formation. By contrast, our new data show that LARP7 depletion strongly impairs U6 incorporation into di-/tri-snRNP complexes (see also point 10), and consistent with this, the effects of LARP7 KD on splicing are more pronounced than those arising from THUMPD2 KO. Notably, there are many more alternative splicing events detected in the THUMPD2 KO, but to increase confidence in the reported events, we chose to present only the most robust changes. However, the detection of numerous mild effects on alternative splicing that are below the stringent significance threshold is in line with the proposed, non-essential role of m²G₇₂ in fine-tuning spliceosome function. It is also important to consider the possibility of adaptation in the knockout cell line that may partially compensate for the lack of THUMPD2 and thus contribute to the relatively low number of significant alternative splicing (AS) events in the KO cell line compared to wild-type. It is also worth mentioning that this type of analysis does not provide information on the efficiency of splicing catalysis, i.e. if splicing is slower/faster than normally. We hypothesise that disturbance of snRNP assembly caused by lack of U6-m²G₇₂, may reduce overall splicing efficiency, which could explain the growth defect observed in THUMPD2 KO cells (PMID: 37283053). These topics are now discussed in more detail in the revised manuscript.

Reviewer: “9- Can the authors show whether U6atac is affected in Figure 5i? The expectation is that U6atac should not be affected.”

Reply: Northern blot analysis of U6atac in nuclear extracts shows that its overall level is not affected by THUMPD2 KO (**Supplementary Fig. S6b**). Reprobing of the gradient membranes to detect U6atac, together with U1 snRNA as reference, suggests a minimal effect of THUMPD2 KO on the level of U6atac in mono-/di-/tri-snRNP complexes when

compared to wild-type (**Supplementary Fig. S6a**). Our data show that lack of THUMP2 affects the migration profiles of PRPF3 and PRPF4 in these gradients (**Fig. 6i-j**); as it has previously been observed that mutations in PRPF proteins, which associate with both U6 and U6atac, affect levels of both snRNAs incorporated into snRNPs (PMID: 21378395), we speculate that the mild effects of THUMP2 KO on U6atac may arise due the impact of U6- m^2G_{72} on PRPF3/4 association with snRNP particles. These new data are presented in the revised manuscript as **Supplementary Fig. S6**.

Reviewer: “10- I don’t like asking for additional complex experiments, but authors could consider testing whether spliceosome assembly, specifically of U6, is similarly affected in IARP7 KD cells compared to THUMP2. Currently, the 2’O methylation requirement for m^7g is only evident from in vitro experiments. This could provide in vivo mechanistic insight into whether 2’O methylation can mimic THUMP2 absence. Alternatively, the wording around these results should more clearly indicate that the dependence is only evident in vitro.”

Reply: The requirement for 2’-O-methylations in U6 for THUMP2-TRMT112-mediated installation of m^2G_{72} was demonstrated both *in vitro* (**Fig. 4 e-f**) and in cellular RNAs. In **Fig. 4g-i**, primer extension (**Fig. 4h**) and DNAzyme cleavage assays (**Fig. 4i**, upper and middle panels) confirm a reduction in 2’-O-methylations in U6 upon LARP7 KD. In parallel, we show by DNAzyme cleavage (**Fig 4i**, bottom panel) that the levels of m^2G_{72} are also affected.

We have performed gradients with nuclear extracts prepared from cells depleted of LARP7 in either the wild-type or THUMP2 KO background. Importantly, due to the requirement of prior 2’-O-methylation of C₆₂ and C₆₃ for efficient THUMP2-TRMT112-mediated methylation (**Fig. 4d-i**), in both these conditions, m^2G_{72} is reduced. Strikingly, the effects on snRNP incorporation and assembly are different when comparing loss of only m^2G_{72} (THUMP2 KO) to the reduction of several 2’-O-methylations together with reduced/absent m^2G_{72} (LARP7 KD). Our results indicate that LARP7 KD significantly reduces U6 incorporation into snRNPs, whereas THUMP2 KO leads to accumulation of U6 in di-/tri-snRNPs. It appears, therefore, that the 2’-O-methylations and m^2G_{72} in U6 have distinct functional roles. As the 2’-O-methylations seem necessary at an earlier stage of U6 snRNP assembly than m^2G_{72} , their effect predominates over the absence of m^2G in the combined LARP7 KD + THUMP2 KO. These results are presented in **Fig. 6e-i** and described in the revised manuscript.

Reviewer: “11- I would like to see full-size gels of all Western and Northern blots as a supplementary file to ensure that the data is not selectively presented.”

Reply: Uncropped versions of all replicates of all western and northern blots are presented in the Source Data file provided with the revised manuscript.

Reviewer: “12- I appreciate the authors for providing a detailed materials and methods section.”

Reply: We thank the reviewer for this remark about the Materials and Methods section.

Minor comments

Reviewer: “1- All prime symbols are incorrect. An apostrophe is not a prime symbol, and please ignore my use of one in my review report for ease of writing.”

Reply: All the prime symbols in the manuscript and figures have been corrected.

Reviewer: “2- Line 118 – Citation for the METTL16 being the U6 snRNA m6A methyltransferase should include PMID: 28525753 and PMID: 29262316, and U6 snRNA A43 being involved in splicing should also cite PMID: 36409063 and PMID: 38808663.”

Reply: The missing references are now included.

Reviewer: “3- Line 147 – Is a punctuation mark missing before “so”?”

Reply: A comma has been included to separate the two independent clauses.

Reviewer: “4- In Figure 2e, the GFP panel shows that the cell on the middle left has an incorrectly placed arrow.”

Reply: The misplaced arrow has now been corrected.

Description of Additional supplementary files

Supplementary Data 1: Output parameters of CRAC data mapping and overview of the number of sequencing reads mapping to different classes of RNA in the His6- 2xFLAG and THUMPD2- His6-2xFLAG CRAC datasets.

Supplementary Data 2: Raw data of U6 snRNA nucleoside quantification in U6 snRNA from WT and THUMPD2 KO cells.

Supplementary Data 3: Overview of the number of normalized sequencing reads mapping to the U6 snRNA precursor sequence in the His6-2xFLAG and THUMPD2-His6-2xFLAG 3' RACE datasets.

Supplementary Data 4: Nuclear proteins identified by DIA-MS in the THUMPD2- His6-2xFLAG IP with UVcrosslinking.

Supplementary Data 5: RNA contact sites mapped in THUMPD2 after RNA-protein crosslinking and mass spectrometry.

Supplementary Data 6: Differential gene expression quantified by DEseq2 of WT, THUMPD2 KO, LARP7 KD and the combined THUMPD2 KO & LARP7 KD.

Supplementary Data 7: Significant AS events derived from rMATS analysis.

Supplementary Data 8: Significant AS events derived from IR analysis and IR fractions.

Supplementary Data 9: DNA oligonucleotides used in this study.

Supplementary Data 10: Plasmids used in this study.

Supplementary Data 11: Antibodies used in this study.

Supplementary Data 12: RNA oligonucleotides used in this study.

Supplementary Data 13: siRNAs used in this study.

SUPPLEMENTARY INFORMATION

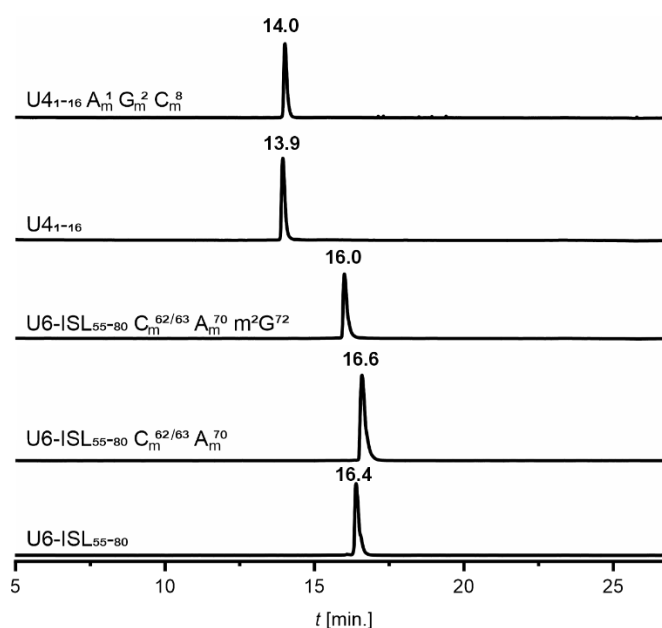
2'-O-methylation-dependent installation of *N*²-methylguanosine in the U6 internal stem loop facilitates efficient spliceosome assembly

Nicole Kleiber, Jonny Petrosyan, Maria Greve, Chairini C. Thomé, Olexandr Dybkov, Laurianne L.E. Tay, Luisa M. Welp, Philipp Hackert, Holger Stark, Markus T. Bohnsack, Henning Urlaub, Lydia Herzelt, Marc Graille, Claudia Höbartner, Katherine E. Bohnsack

SUPPLEMENTARY TEXT

Verification of synthesized RNAs

The purity of the synthesized RNA samples was analyzed by anion-exchange HPLC (Dionex DNAPac PA200, 2 × 250 mm, flow rate 0.5 mL/min, at 60 °C; Solvent A: 25 mM Tris-HCl (pH 8.0), 6 M Urea. Solvent B: 25 mM Tris-HCl (pH 8.0), 6 M Urea, 0.5 M NaClO₄. Gradient: 0-48% solvent B in 12 CV). The resulting anion exchange HPLC-traces (260 nm) are shown below.



To confirm the identity of all synthesized RNA samples, HR-ESI-MS analysis on a Bruker microOTOF-Q-III (negative ion mode, direct injection) was performed. The monoisotopic masses were obtained by charge deconvolution. The results are reported below.

Oligonucleotide	Sum Formula	Mass calculated [Da]	Mass found [Da]
U6-ISL ₅₅₋₈₀	C ₂₄₈ H ₃₁₀ N ₁₀₃ O ₁₇₇ P ₂₅	8336.18686	8336.21252
U6-ISL ₅₅₋₈₀ Cm _{62/63} Am ₇₀	C ₂₅₁ H ₃₁₆ N ₁₀₃ O ₁₇₇ P ₂₅	8378.23381	8378.25405
U6-ISL ₅₅₋₈₀ Cm _{62/63} Am ₇₀ m ² G ₇₂	C ₂₅₂ H ₃₁₈ N ₁₀₃ O ₁₇₇ P ₂₅	8392.24946	8392.24627
U4 ₁₋₁₆	C ₁₅₂ H ₁₉₀ N ₆₁ O ₁₁₁ P ₁₅	5109.76177	5109.73947
U4 ₁₋₁₆ Am ¹ Gm ² Cm ⁸	C ₁₅₅ H ₁₉₆ N ₆₁ O ₁₁₁ P ₁₅	5151.76372	5151.79028

Establishment of methylation-sensitive DNazymes

To label unmodified and modified U6-ISLs with fluorescein on the 3'-end, periodate oxidation was used. The reaction was performed in a total volume of 48 μ L. RNA (5 nmol), sodium phosphate buffer (pH 7.5, 50 mM) and NaIO₄ (20 mM) were incubated for 10 min at 37 °C. After addition of Na₂SO₃ (83 mM), the mixture was incubated for another 5 min at 37 °C. Following addition of Flu-TSC (8 mM), the final mixture was incubated for 2 h at 37 °C. Stop solution was added and the labeled RNA was purified via a 15% PAA gel. After ethanol precipitation, the labeled RNA was used for DNazyme cleavage kinetic assays.

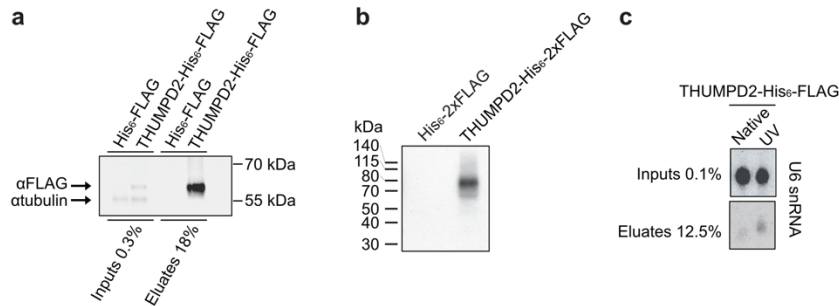
For DNazyme cleavage kinetic assays, 10 pmol of 3'-fluorescein labeled unmodified and modified U6-ISL was annealed to 100 pmol of DNazyme (m²G, Am₇₀ or Cm₆₂/Cm₆₃) in reaction buffer (m²G and Am₇₀: 50 mM HEPES (pH 7.4), 400 mM KCl, 100 mM NaCl; Cm₆₂/Cm₆₃: 50 mM HEPES (pH 7.4), 150 mM NaCl). The reaction was initiated by addition of 5 mM MgCl₂ (m²G), 10 mM MgCl₂ and 10 mM MnCl₂ (Am₇₀) or 1 mM MgCl₂ and 1 mM MnCl₂ (Cm₆₂/Cm₆₃). After incubation at 25 °C, 1 μ L aliquots were quenched at the desired timepoints in 4 μ L of stop solution. Subsequently, half of the reaction mixture was loaded on a 20% PAA gel. After running the gel for 1 h at 25 W, band intensities were visualized and analyzed *via* fluorescence imaging using a ChemiDOCMP with blue epi illumination and an 530/28 nm emission filter. All DNazyme cleavage kinetic assays were performed in triplicates. To obtain k_{obs} the mean value of the cleaved fraction (%) for each time point were plotted against the respective incubation time and fitted according to the equation below where P_{max} = final yield using OriginPro2023b.

$$y = P_{max} * (1 - \exp(-k_{obs} * t))$$

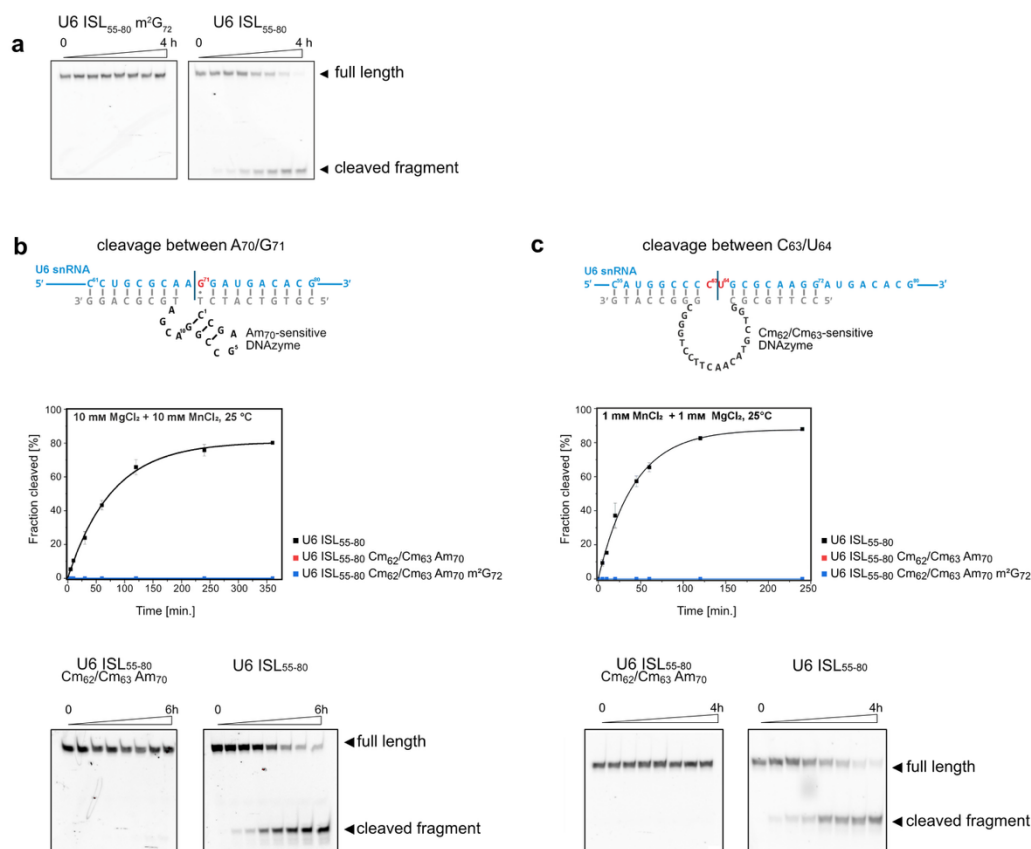
The cleavage site for the Cm₆₂/Cm₆₃ DNazyme was confirmed *via* HR-ESI-MS. The sum formula, calculated and found masses of the U6-ISL fragments after incubation with the Cm₆₂/Cm₆₃ DNazyme for 4 h at 25 °C are shown below.

Oligonucleotide	Sum Formula	Mass calculated [Da]	Mass found [Da]
3'-fragment	C ₁₆₄ H ₂₀₃ N ₇₁ O ₁₁₄ P ₁₆	5485.80774	5485.89581
5'-fragment	C ₈₄ H ₁₀₇ N ₃₂ O ₆₃ P ₉	2850.37967	2850.43181

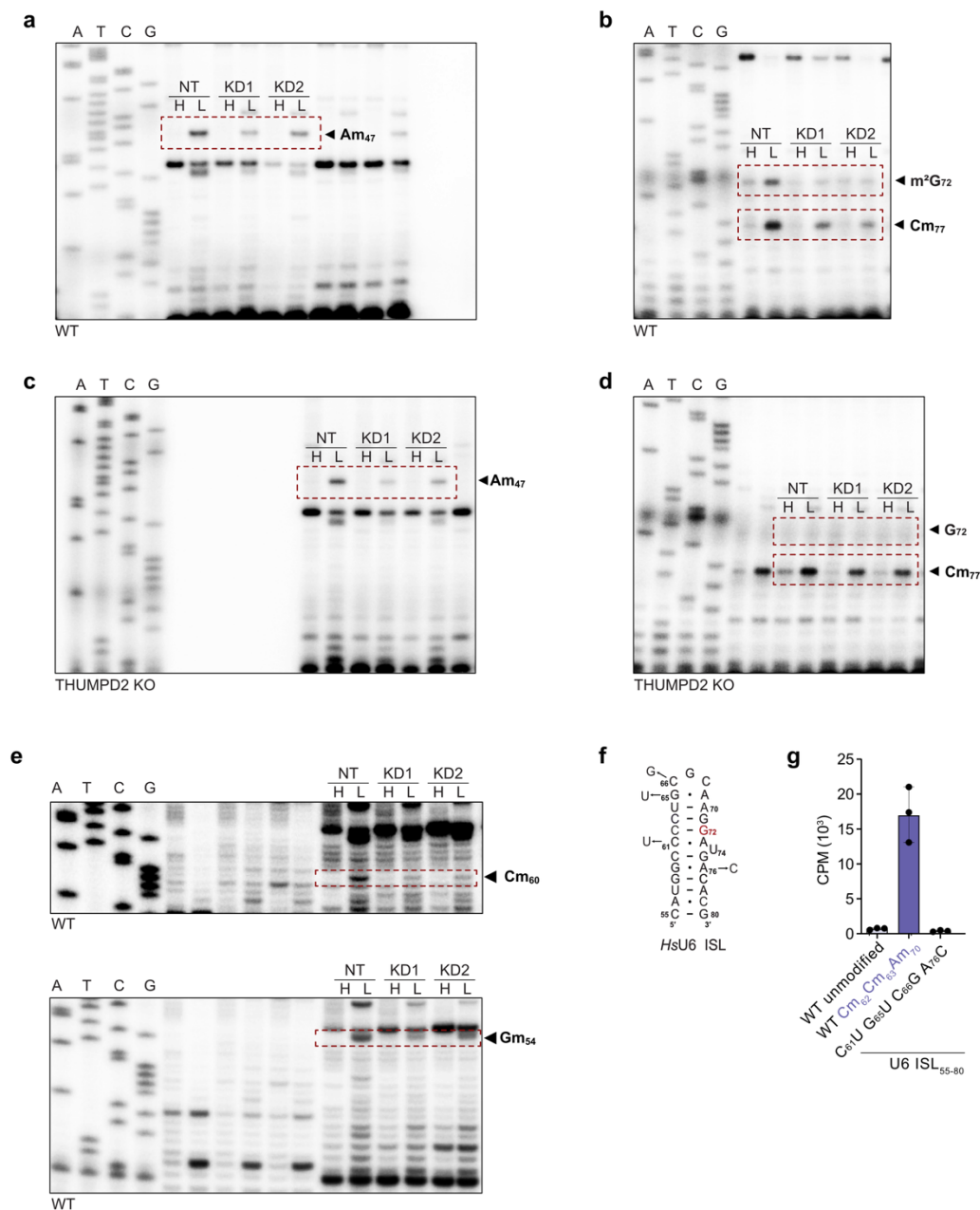
SUPPLEMENTARY FIGURES AND LEGENDS



Supplementary Figure S1. Immunoprecipitation of THUMP2-His₆-2xFLAG and protein–RNA crosslinking (a-b) HEK293 over-expressing His₆-2xFLAG or THUMP2-His₆-2xFLAG were UV-crosslinked and used in n=2 anti-FLAG immunoprecipitation experiments. The FLAG-tagged versions of the proteins were detected by western blotting with an anti-FLAG antibody and tubulin was used as a loading control (a). After partial RNase digestion, the 5' end of recovered RNA fragments were 5'-labelled with [³²P] and adaptors were ligated to the 5' and 3' ends. Protein–RNA complexes were separated by denaturing PAGE, transferred to a nylon membrane and detected by autoradiography (b). **(c)** Extracts from cells expressing THUMP2-His₆-2xFLAG that had been crosslinked (UV) or not (Native) were used for anti-FLAG immunoprecipitation experiments (n=1). RNA was eluted from beads by Proteinase K treatment and RNAs in input and eluate samples were separated by denaturing PAGE and U6 snRNA was detected by northern blotting. Note the increased recovery of U6 after crosslinking and the upwards shift of the U6 signal, demonstrating successful protein–RNA crosslinking.

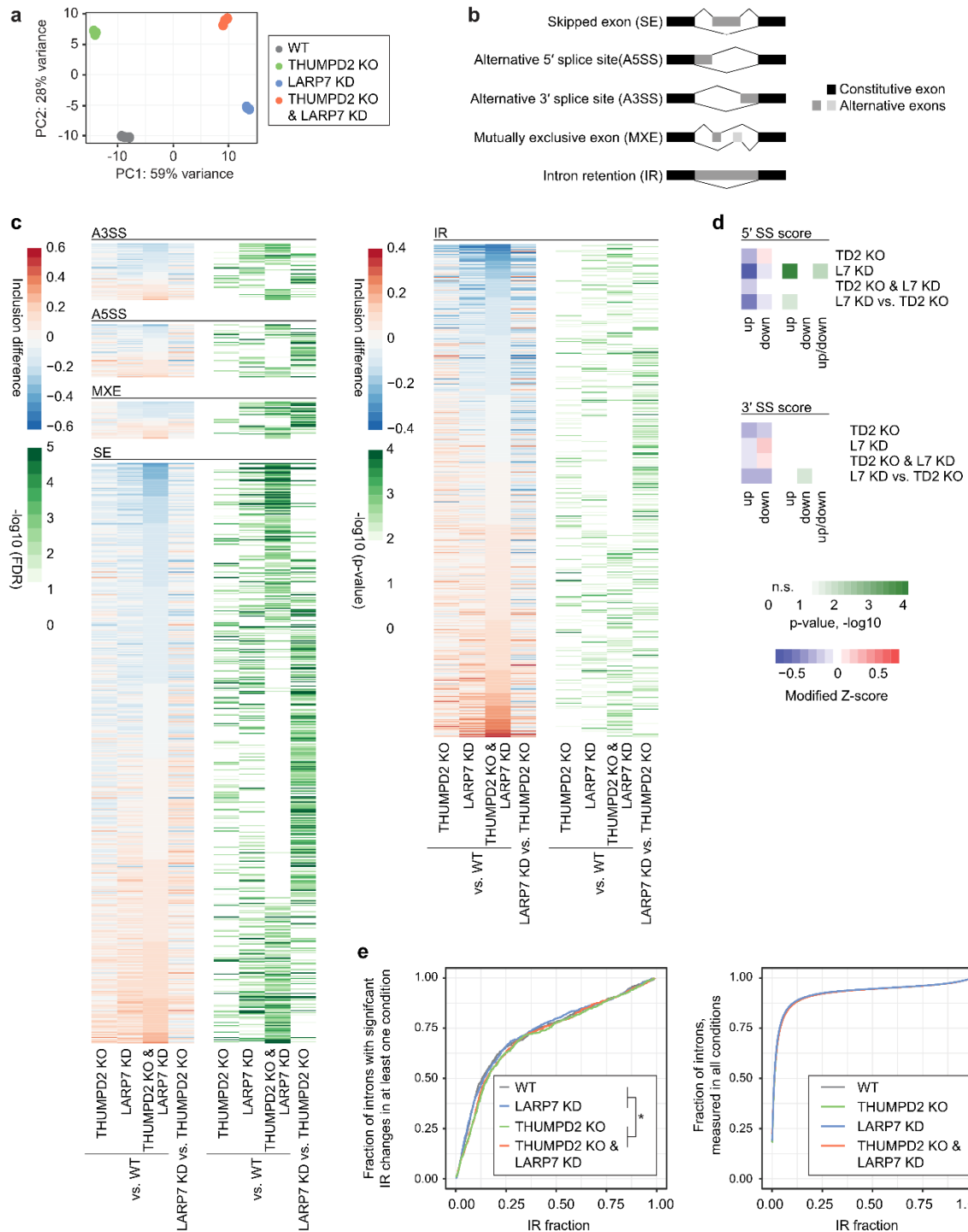


Supplementary Figure S2. Characterization of DNAzymes for the detection of U6 ISL modifications. (a) Synthetic U6 ISL₅₅₋₈₀ m²G-containing (left) or unmodified (right) was treated with an m²G-sensitive DNAzyme that cleaves unmodified RNA. Formation of cleavage fragments was analyzed by denaturing PAGE for the timepoints indicated between 0 and 240 min. Arrows indicate full length RNA oligonucleotide or cleaved product. Representative gel images from n=3 independent experiments. **(b-c)** Schematic view of base-pairing of the Am70- (b, top) or Cm₆₂/Cm₆₃-sensitive DNAzymes (c, top) with U6 snRNA (in blue) to detect the respective 2'-O-methylations. The cleavage site is indicated in the U6 sequence. As in (a), RNA oligonucleotides containing the sequence of the U6 ISL₅₅₋₈₀ unmodified (black line) or containing the respective 2'-O-methylations (red and blue lines, left for Am70 and right for Cm₆₂Cm₆₃) were incubated with the respective DNAzymes and the fraction of cleaved RNA at the times indicated in the plots (middle panels) were determined after quantification of the bands shown in the gels (bottom panels). Arrows indicate full length RNA oligonucleotide or cleaved product. Data from n=3 independent experiments shown as mean ± standard deviation. Source data are provided as a Source Data file.



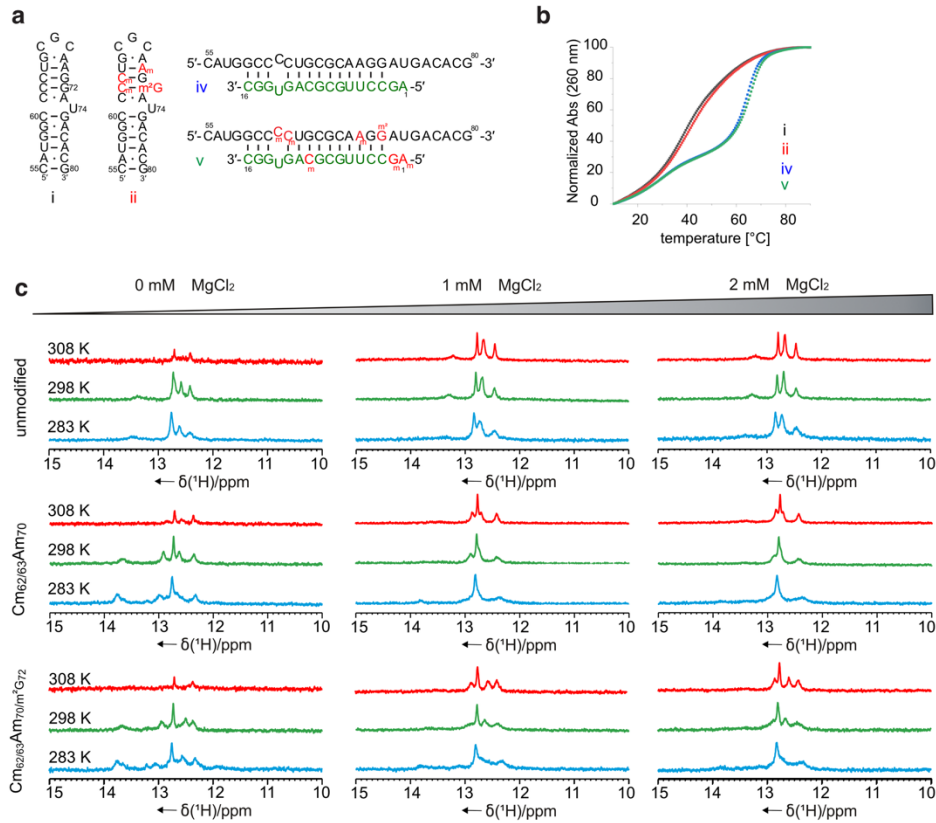
Supplementary Figure S3. LARP7-dependence of 2'-O-methylations in U6 determined by primer extensions and effect of U6 ISL stem stability on THUMPDP2 methylation. (a-d) Uncropped primer extension gels represented in Fig 4h for RNA extracted from HCT116 WT (a, b) and THUMPDP2 KO cells (c, d) after treatment with a non-target siRNA (NT) or siRNAs against LARP7 (KD1 and KD2). H = high dNTP concentration and L = low dNTP concentration. Red boxes highlight the cDNA fragments corresponding to the m^2G_{72} modification (b) and the fragments indicative of the presence of Cm_{77} (b,d) or Am_{47} (a,c). A sequencing ladder was run alongside each set of samples and A, T, C, G indicate ladders prepared using termination mixes containing ddATP, ddTTP, ddCTP and ddGTP, respectively.

Representative images for n=3 experiments. **(e)** RNA from siRNA-treated HCT116 WT cells was used for primer extension assays, as in (a-d) to detect Cm₆₀ (top) and Gm₅₄ (bottom). Representative image for n=2 experiments. **(f)** Schematic view of the human U6 ISL in which four sub-optimal base-pairs closest to G₇₂ were strengthened by nucleotide substitutions (C₆₁U, G₆₅U, C₆₆G and A₇₆C). **(g)** *In vitro* methylation assays were performed using purified His₆-ZZ-BtTHUMPD2-BtTRMT112 protein complex, [³H]-SAM as methyl group donor and the different RNA substrates, including WT unmodified and 2'-O-methylated U6 ISLs and the mutated ISL depicted in (f). Tritium incorporated in the substrate RNAs was quantified and bar plots show averaged counts per minute (CPM) of n=3 independent experiments with error bars representing mean ± standard deviation.

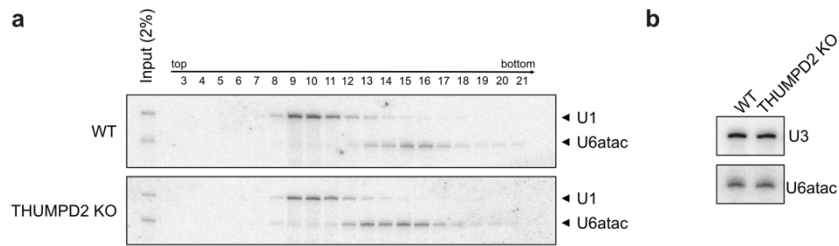


Supplementary Figure S4: Characterization of RNA-seq datasets by principal component analysis and individual alternative splicing events. (a) Principal component analysis based on gene-level expression quantification (top 5000 genes). The three biological replicates are shown as individual dots of the same color, and display good reproducibility. (b) Scheme of different types of alternative splicing. (c) Heatmap representation of alternative splicing events by AS type and sampled condition. Events are shown that change significantly in at least one of the 4 comparisons. Comparisons are made with respect to WT or between

LARP7 KD and THUMPD2 KO (right-most column). Significant AS events (rows) are sorted by increasing inclusion difference in the combined condition. Significant A3SS, A5SS, MXE and SE events are derived from rMATS (FDR < 0.01, absolute inclusion difference > 0.1, median read count in at least one condition > 50). Significant IR events are based on splicing efficiency calculation of annotated introns (p-value < 0.01, absolute inclusion difference > 0.05, median read count in at least one condition > 50). **(d)** Heatmaps showing the change of the median 5'/3' splice site strength in more (up) or less (down) retained introns in the samples indicated. The modified Z-score relative to introns without significant changes and the significance of the changes plotted as $-\log_{10}(\text{p-value})$ from the Wilcoxon rank-sum test are plotted. LARP7 KD = L7 KD, THUMPD2 KO = TD2 KO. **(e)** Cumulative IR fraction distribution of introns that change pre-mRNA splicing significantly in at least one of the 4 comparisons (left) and of all introns passing the read count cutoff (right). p values of comparisons summarized with * were derived with the two-sided Kolmogorov-Smirnov with multiple testing correction using the Bonferroni method. WT against THUMPD2 KO 0.009792; WT against LARP7 KD & THUMPD2 KO 0.025404; LARP7 KD against THUMPD2 KO 0.021108; LARP7 KD against LARP7 KD & THUMPD2 KO 0.043554, and 1 for WT against LARP7 KD and THUMPD2 KO against LARP7 KD & THUMPD2 KO, respectively.



Supplementary Figure S5. Analyses of U4/U6 partial duplex and U6 ISL stability and structure *in vitro*, and snRNA levels in wild-type and THUMPD2 KO cells. (a) Schematic view of the synthetic RNA samples used for biophysical investigations. Modified nucleotides are shown in red. The U6 ISLs (i and ii) are as in Fig. 6a and within the modified and unmodified U4/U6 partial duplexes (iv and v), U6 is shown in black and U4 in green. **(b)** UV thermal melting curves for the samples shown in (a) plotting hyperchromicity at 260 nm. Conditions: 1 μ M RNA, 100 mM NaCl, 10 mM potassium phosphate buffer pH 7.4, T = 10 $^{\circ}$ C – 90 $^{\circ}$ C. One of n=4 reversible heating and cooling curves is shown. **(c)** 1D 1 H-NMR spectra of imino region, 120 μ M RNA in 10 mM phosphate buffer pH 7.4, 90% H₂O/ 10% D₂O at 283 K, 298 K or 308 K in the presence of 0, 1 or 2 mM MgCl₂.

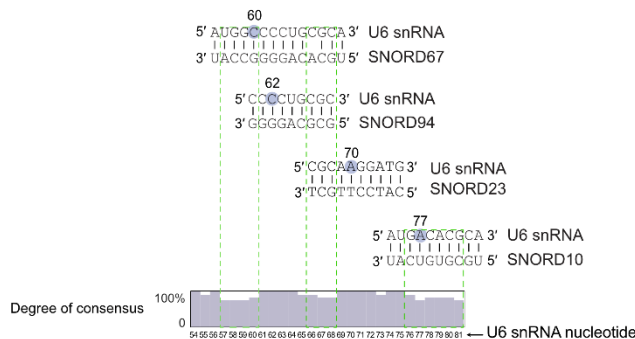


Supplementary Figure S6. Effect of THUMPD2 KO on U6atac levels and assembly into snRNPs. **(a)** Sedimentation of snRNP particles from HCT116 WT and THUMPD2 KO cells in a 10-30% glycerol gradient was analyzed by northern blotting using probes against U1 and U6atac. Representative data from n=3 experiments is shown. **(b)** Northern blot analysis of the levels of U6atac compared to the U3 snoRNA (loading control) in nuclear extracts from wild-type (WT) and THUMPD2 KO cells. Representative data from n=3 experiments is shown.

a

	Type of modification	Modification symbol and position	Enzyme or enzyme/guide RNA(s)	Source (PMID)
Terminal modifications	5' Capping	5' γ -monomethyl	MEPCE	8175647; 17643375
	3' End extension and trimming	3' terminal 2',3'-cyclic phosphate	TUT1 and USB1	TUT1 (U6 TUTase): 9628908; 16790842. USB1 (Mpn1): 8995392; 22899009; 23022480; 23190533
Internal modifications	Pseudouridylations	Ψ 31	Dyskerin / SNORA79 (ACA65)	SNORA79 (ACA65): 16373490
		Ψ 40	Dyskerin / SCARNA23/3(ACA12)	SCARNA23 (ACA12): 15199136 and SCARNA3 (HBI-100): 11387227
		Ψ 86	Dyskerin / SNORA79 (ACA65)	SNORA79 (ACA65): 16373490
	2'-O-methylations	Am47	Fibrillarin / SNORD7	SNORD7 (mgU6-47): 9844635
		Am53	Fibrillarin / SNORD8/SNORD9	SNORD8/9 (mgU6-53): 10490628
		Gm54	Fibrillarin / ?	-
		Cm60	Fibrillarin / SNORD67	SNORD67 (HBI-166): 11387227; 40316533
		Cm62	Fibrillarin / SNORD94	SNORD94 (U94-U106): 14602913
		Cm63	Fibrillarin / ? (possibly SNORD88A/B)	SNORD88A/B: 37792516
		Am70	Fibrillarin / ? (possibly SNORD23)	SNORD23: 40001124
		Cm77	Fibrillarin / SNORD10	SNORD10 (mgU6-77): 9844635
	Base methylations	m ⁶ A ₄₃	METTL16	28525753; 29051200
		m ² G ₇₂	THUMP2	37283053; 38165050; this study

b



Supplementary Figure S7. U6 snRNA modifications and snoRNA base-pairings within the U6 ISL. (a) Summary of the mapped U6 snRNA terminal (5' and 3') and internal modifications (pseudouridylations, 2'-O-methylations and base methylations). Enzymes catalyzing the indicated modifications are listed on the right and where applicable guide RNAs are indicated. Source publications (PMID) are indicated on the right. **(b)** Base-pairings between the U6 ISL and box C/D snoRNAs guiding the 2'-O-methylation of C₆₀, C₆₂, A₇₀ and A₇₇ are shown. 2'-O-methylated nucleotides are highlighted and the position in the U6 sequence is indicated. Regions boxed in a dashed green line correspond to those with lower conservation scores as indicated in the bar plot representing the degree of consensus (0 = no conservation; 100% = maximum percentage identity) from Fig. 4k.