

Peer Review File

Manuscript Title: Regulation of hepatic subcellular architecture controls metabolic homeostasis

Redactions – unpublished data

Parts of this Peer Review File have been redacted as indicated to maintain the confidentiality of unpublished data.

Reviewer Reports on the Initial Version:

Referee #1:

This manuscript describes extensive characterization of differences in ER morphology in liver cells and tissue from lean and obese mice. The total amount of ER and the fractions in sheet and tubule form were measured by image analysis of FIB-SEM images, indicating that there was an increase in tubular ER in obese mice. Additional molecular characterization was carried out that was consistent with this observation. Expression of a tubule-enhancing protein (CLIMP-63) reversed this difference and also partially reversed metabolic differences, making a compelling case for the role of ER structure in its function. The results should be of significant interest to the cell biology community.

Since the foundation of the study is the results of the FIB-SEM image analysis, the methods (lines 779-801) should be expanded to include a table with pixel- and object-level classification accuracies (precision, recall and f-1 scores) on test images both for the task of ER segmentation and the task of sheet-tubule discrimination. It would also be valuable to the community for the authors to provide the hand-labeled images used for training and testing.

Minor corrections:

Line 138 "Grand" truth should be changed to "ground" truth

Line 780 Extraneous "ag" after "ariadne.ai"

Robert F. Murphy, Ray & Stephanie Lane Professor of Computational Biology, Carnegie Mellon University

Referee #2:

This study addresses the functional relevance of organelle composition and more specifically the architecture of the endoplasmic reticulum(ER) in hepatocytes in vivo. For this purpose the authors analyzed the number of organelles and the structure of the ER in hepatocytes of lean and obese mice by acquiring large data sets using enhanced FIB-SEM at an isotropic voxel size of 8 nm. Subsequently the organelles in these data sets were automatically segmented using convolutional and U-shaped neuronal networks. ER fine structure was also analyzed by thin section TEM.

The authors detected interesting structural differences between these conditions, e.g. an increase in cell size and amount of lipid droplets in the obese mice as well as a disorganized ER with an enrichment of tubular subdomains and a reduction of parallel ER sheets. EM observations were validated by subcellular fractionation into rough and smooth ER. Important differences between livers from lean and obese mice were also found regarding the enrichment of specific proteins in the respective ER subdomains. Proteins like SEC61 and TRAPa typical for rough ER redistributed to smooth ER in obese mouse liver. Analysis of proteins involved in the determination of ER shape revealed an enrichment of proteins involved in tubular ER formation in liver lysates and ER fractions of obese mice accompanied by elevated expression. Less clear distributed was the ER sheet promoting protein Climp-63 with a higher expression, but decreased abundance in the ER fraction in obese liver lysates. However, subsequent rescue experiments using adenoviral-mediated overexpression of Climp-63 in primary cultured hepatocytes and in vivo in liver of obese mice dramatically altered the ER structure with significant consequences for liver

metabolism.

This is a remarkable study applying an impressive array of methods that provide a large collection of important and valuable data laying a basis for further analyses. It describes a fundamental structure-function relationship by altering a molecular component which is involved in shaping the ER structure in vivo in a relevant pathological mouse model. The study is performed carefully including validation by a number of independent experiments (EM, cell fractionation, immunofluorescence staining, primary cell culture, Western blot including antibody validation and qRT PCR). The visualization of the 3D data in the videos is prize-winning breath taking stunning conveying the content also for the non-expert public in the light of an increasing prevalence of obesity worldwide. In summary this is an important study of interest for a broad readership and appropriate for publication in Nature.

Nevertheless I would like to raise some questions:

The quality of the 3D FIB-SEM data stacks and the derived quantitative data and the comprehensive 3D models and also the TEM data provide the opportunity to possibly extend the analysis.

The 3D data stacks may contain the structural information to look specifically at mitochondria-associated ER membranes (MAMs) implicated in ER stress associated with obesity: Is it possible to visualize and quantify those in the existing data stacks?

If there is any chance to acquire enhanced FIB-SEM data on obese liver with altered ER structure by Climp-63 overexpression that would be fantastic.

Concerning the consequences of Climp-63 overexpression in the liver of lean mice: Extended Data Fig. 7f indicates a striking increase of ER sheets in control mice. Is this quantified? Is this of relevance for liver metabolism in lean mice?

How does the ER structure in obese mice and the Climp-63 induced change in ER structure relate to ER stress and UPR? Apart from the ASGR1 assay which is a read out for correct protein folding in primary cultured hepatocytes, it would be interesting to look at PERK activation or Bip expression in livers of lean and obese mice, but only if the authors feel this is indicated.

Minor points:

How many mice were used for the TEM analysis of the 15 full hepatocytes (Extended data Fig. 4)? This type of analysis offers the opportunity to increase n-numbers with limited effort, e.g. by analysis of 3 hepatocytes in 5 mice. However, the effects are so strong that increasing the group size is probably not needed.

Fig. 3b: Western blot of ER fractions: Why normalized to tubulin? Is this just a mistake in y-axis labeling and calnexin was used for normalization?

Fig 3f and g: The readout of the ASGR1 assay is unclear to me. What exactly is determined by Cluc and Gluc luminescence? Some explanation in the methods section or in the legend of Extended Data Fig. 7 would be helpful.

Fig. 4: Small typo in the legend of Fig. 4b: This is the quantification of the images shown in figure 4a (not 5A), right? Question: From how many animals were the analyzed EM sections taken?

Line 498: "Representative of 3 independent cohorts": This is unclear to me. Does this mean 3 cohorts of mice with n=8 (LacZ) and n=10 (Climp-63)?

Extended Data Fig. 4: What is the unit (y-axis label) in e), f) and g)? In addition it may be helpful to add a small sentence to explain why fed and fasted states were compared.

Extended Data Fig. 6: The main text refers to Extended Data Fig 6f and g, but these subpanels unfortunately don't exist.

Referee #3:

This study by Parlakgöl and Arruda et al. uses state-of-the art FIB-SEM methods to bring to life the architecture of organelles in their native tissue context. The authors observe striking changes to ER shape that vary based on the metabolic state of the organism (lean vs. obese). The authors have pushed FIB-SEM technology to the limit. This is a major technical achievement that required tackling technical and computational obstacles and will inspire future efforts using their tools on a wide array of biological questions. In addition, the interplay between organelle shape and cellular and organismal metabolism dissected in this study will be interesting to a broad audience. I believe this study is worthy of publication with some revisions (described below). Concerns regarding data interpretation are also listed below:

-By altering the level of ER shaping proteins, the authors identify physiological and metabolic consequences of ER shape at both cellular and organismal levels. I have two suggestions to improve and extend the interpretation of these data. First, an alternative strategy to “rescue” ER shape besides Climp63 overexpression should be used. I suggest depletion of Rtn4. If the metabolic results are similar, then one can conclude that shape is important. Second, Climp63 overexpression in obese mice will reduce tubular ER and because the biogenesis of both LDs and mitochondria is regulated by tubular ER contact sites, it would be highly informative if the authors would also map the morphology of mitochondria and abundance of LDs when sheets are expanded- particularly because these organelles are both linked to metabolism as well. Presumably these data are all present in their FIB-SEM reconstructions.

-Ext Data Fig 5, a-e: This experiment can easily be confounded by contamination, as only the weights of the fractions are reported. This limitation is not adequately addressed by EM. The authors should perform immunoblotting on organelle, ribosome, and cytosolic markers on the total (input) and collected fractions to confirm that their enrichment procedure is effective.

-Ext Data Fig 5i: Unlike in panel h, where blot quantification and statistical tests are performed, the conclusion that rough ER proteins “translocated to the tubular ER” (main text, lines 204-205) are drawn from unquantified blots and no statistics. The conclusions drawn are not evident to me, except possibly for Sec61beta, although that is also variation in the lean mouse smooth ER. It is suggested that GRP78 – a soluble luminal protein – is a rough ER protein (not sure why). Can the authors clarify how GRP78 would segregate between smooth and rough ER? Or, is GRP78 better interpreted as an indicator of the total ER content of these fractions (for which there is otherwise no normalization)?

-Fig 3, a-b: If all Climp63 is in the ER, then why are opposite conclusions drawn between the total cell extract (panel a) and ER fraction (panel b)? This leads me to believe that the purifications are variable. It is difficult to assess this experiment as displayed because no evidence is provided that ER is specifically and efficiently enriched in this assay. Additionally, all blot quantifications are labeled as normalized to tubulin. How is this possible for panel b, which should not contain any tubulin after the ER enrichment procedure? The opposite effect is seen for p180, which is reduced in total liver lysates from the obese condition but not significantly changed in the total ER fraction.

-Ext Data Fig 7, c-d: The resolution of the images shown is not sufficient to conclude that ER shape has been altered.

-Fig 3, d-e: This experiment simply does not test what is claimed in the text (lines 264-266), namely, that overexpression of Climp-63 restores Sec61beta localization to ER sheets. Are the authors able to colocalize endogenous Sec61beta with exogenous Myc-tagged Climp-63 in ER sheets? Are the total levels of Sec61beta (by immunoblot) changed?

-Fig 3, f-g: Related to the above point, folding capacity and translocation capacity seem to be used interchangeably within the logical flow of the paper despite being distinct processes. Are the results with the ASGR reporter explained by changes in the levels of translocation machinery or chaperones? Perhaps a correlative argument could be made by blotting for chaperones and translocon subunits in the extracts (lean vs obese; obese control vs Ad-Climp63). This could be done in conjunction with my comments about improving Ext Data Fig 5i above, as it uses a similar readout.

Minor:

-In the main text, it is unclear how many mice are examined (line 87 and following). Is the plural "lean mice" is used to refer to a single mouse? The sample size of the FIB-SEM portion of the study should be clearly explained.

-Figure 1, a and b: What is identified as an ER tubule (panel a)? The ER appears to be predominantly sheets in this region. The inset showing the highest level of zoom is not correctly drawn on the medium zoom inset. Panel b does not contain the high level of zoom inset, precluding comparison of the two conditions.

-Figure 1, i and j: (1) labels are switched in the legend; (2) it is confusing that these very similar plots use % (in i) and ratio (in j); (3) it seems odd to highlight in the main figure the significant difference in ER volume, while Ext Data Fig 2 shows that this is due to a trivial explanation. This manner of presentation risks misinterpretation.

- In the methods section, authors describe culturing Hep1-6 and Cos-7 cells under 10% CO₂, when data-sheet for these cells clarify they use 5% CO₂. Do the authors have a particular reason to use this condition?

-Movie 3: Are lipid droplets modeled inside the nucleus?

-Ext Data Fig 5, f-g, is at odds with its description in the main text (line 190). Which subunit is stained in this experiment: Sec61alpha or -beta?

-Main text, line 236, should refer to Ext Data Fig 6e

-Ext Data Fig 7: Here (and potentially elsewhere) the identity of the fusion proteins should be described accurately. Sec61beta is a tail-anchored protein that must be tagged on its N terminus (GFP-Sec61beta, not Sec61beta-GFP). Likewise, KDEL is an obligatory C-terminal peptide and should be GFP-KDEL. Due to these errors, it is unclear from the figure how you have tagged other proteins like Climp-63.

-Methods and Supplementary Table 2: There are numerous reagents listed that do not appear anywhere in the study that I can find.

Referee #4:

This is an interesting manuscript reporting the acquisition and deconvolution of images obtained from powerful FIB-SEM technology in a collaboration with the HHMI Janiela Research Campus group. The images obtained are novel in 3D details, with the important advantage that cells within intact tissue are evaluated. While Figs 1 and 2 are informative in describing hepatocyte ER structures and effects of obesity, they are not much more informative in conclusions than what can be learned from normal EM techniques. Nonetheless, they add to the many descriptions of obesity effects in the literature, albeit only from one extreme genetic model. And clearly the data is well presented and organized, and the paper well written.

The major problem with the manuscript is that the main conclusion is not properly evaluated and critically tested. The impact of the studies is therefore limited, and advancement of the field incremental. The authors' claim that ER structure determines physiological function is weak on four levels:

1. The ER structural changes observed in Figs 3 and 4 are purported to cause the physiological changes in hepatocyte lipogenesis and glucose output, but this is actually a correlation not a demonstration of causality. The genetic manipulation is assumed to be all due to the ER structural change, yet it is possible and even likely that other direct effects of Climp-63 expression are induced independent of ER structure. There is no data to exclude that problem.

2. Figure 3 shows Climp-63 expression in hepatocytes from obese mice in vitro, and indicates an inhibition in glucose output. This is an inhibitory effect, raising the question of whether it is a result of toxicity vs specific regulation. What is the effect of Climp-63 expression on cells from lean mice? Is glucose output also decreased in "lean" cells? Are other features of the metabolism of these cells also inhibited? What is the effect on mitochondrial respiration, lipogenesis and lipid oxidation? The specificity of this effect is not addressed, nor is a KO strategy imposed, which should show opposite effects on the above metabolic parameters if toxicity is not a factor.

3. Figure 4 again shows only results from obese mice, and lacks evaluation of the control lean mice. Without control, lean mice, other interpretations are possible. Is this a toxicity effect? Is the improved

GTT due to changes in insulin levels? Importantly, does hepatic specific KO of Climp-63 mimic the effects of obesity on ER Structure and metabolic regulation (glucose output and lipogenesis). This is an easy (AAV injection) and important experiment as increases in these metabolic pathways would be expected in the KO, and if observed, would argue in the authors' favor.

4. The notion of ER structure causing metabolic changes is interesting, but unsatisfying unless the mechanism is revealed of how such ER structure is connected to the enzymes and fluxes through the metabolic pathways. There appear to be no measurements of gluconeogenesis or lipogenesis enzyme levels, nor whether they are displaced from ER or other structures. Mitochondria are known to be physically associated with the ER, and this would be a way that ER structure could effect oxidative metabolism (which is connected to lipogenesis and gluconeogenesis). But there is no evaluation of this issue. What is the way in which ER structure could mechanistically regulate such metabolic pathways is central to this study, and if it can be answered, it would negate the criticisms in items 1-3 above, as those issues could be directly tested in the context of such a molecular mechanism.

Author Rebuttals to Initial Comments:

Referee #1

This manuscript describes extensive characterization of differences in ER morphology in liver cells and tissue from lean and obese mice. The total amount of ER and the fractions in sheet and tubule form were measured by image analysis of FIB-SEM images, indicating that there was an increase in tubular ER in obese mice. Additional molecular characterization was carried out that was consistent with this observation. Expression of a tubule-enhancing protein (Climp-63) reversed this difference and also partially reversed metabolic differences, making a compelling case for the role of ER structure in its function. The results should be of significant interest to the cell biology community. Since the foundation of the study is the results of the FIB-SEM image analysis, the methods (lines 779-801) should be expanded to include a table with pixel- and object-level classification accuracies (precision, recall and f-1 scores) on test images both for the task of ER segmentation and the task of sheet-tubule discrimination. It would also be valuable to the community for the authors to provide the hand-labeled images used for training and testing.

We thank the reviewer for the interest in our paper. As requested, we added a new table containing the voxel and object level accuracy, precision, recall, and f-1 scores for organelle segmentation and also for the sheet-tubule re-segmentation. For lipid droplets, mitochondria, nuclear membrane and ER, we selected 3 test volumes of 300×300×50 voxels at random locations. For ER sheet and tubule evaluation, we randomly choose an area of 750×750 pixel² at 4 different locations in each dataset. These values can be found in the **Supplemental Table 2**. The hand-labeled images, raw data and predictions are also uploaded in <https://www.ebi.ac.uk/empirar/database> (Deposition ID: 47483731).

Minor corrections:

Line 138 “Grand” truth should be changed to “ground” truth.

Thank you for pointing out this error, this has now been corrected.

Line 780 Extraneous “ag” after “ariadne.ai”.

Thank you, this has now been corrected.

Referee #2:

This study addresses the functional relevance of organelle composition and more specifically the architecture of the endoplasmic reticulum (ER) in hepatocytes in vivo. For this purpose, the authors analyzed the number of organelles and the structure of the ER in hepatocytes of lean and obese mice by acquiring large data sets using enhanced FIB-SEM at an isotropic voxel size of 8 nm. Subsequently the organelles in these data sets were automatically segmented using convolutional and U-shaped neuronal networks. ER fine structure was also analyzed by thin section TEM. The authors detected interesting structural differences between these conditions, e.g. an increase in cell size and amount of lipid droplets in the obese mice as well as a disorganized ER with an enrichment of tubular subdomains and a reduction of parallel ER sheets. EM observations were validated by subcellular fractionation into rough and smooth ER. Important differences between livers from lean and obese mice were also found regarding the enrichment of specific proteins in the respective ER subdomains. Proteins like SEC61 and TRAP α typical for rough ER redistributed to smooth ER in obese mouse liver. Analysis of proteins involved in the determination of ER shape revealed an enrichment of proteins involved in tubular ER formation in liver lysates and ER fractions of obese mice accompanied by elevated expression. Less clear distributed was the ER sheet promoting protein Climp-63 with a higher expression but decreased abundance in the ER fraction in obese liver lysates. However, subsequent rescue experiments using adenoviral- mediated overexpression of Climp-63 in primary cultured hepatocytes and in vivo in liver of obese mice dramatically altered the ER structure with significant consequences for liver metabolism.

This is a remarkable study applying an impressive array of methods that provide a large collection of important and valuable data laying a basis for further analyses. It describes a fundamental structure-function relationship by altering a molecular component which is involved in shaping the ER structure in vivo in a relevant pathological mouse model. The study is performed carefully including validation by a number of independent experiments (EM, cell fractionation, immunofluorescence staining, primary cell culture, Western blot including antibody validation and qRT PCR). The visualization of the 3D data in the videos is prize-winning breath-taking stunning conveying the content also for the non-expert public in the light of an increasing prevalence of obesity worldwide. In summary this is an important study of interest for a broad readership and appropriate for publication in Nature.

We thank the reviewer for the kind words about our manuscript. We are very grateful and encouraged by the assessment of our studies.

1- Nevertheless, I would like to raise some questions: The quality of the 3D FIB-SEM data stacks and the derived quantitative data and the comprehensive 3D models and also the TEM data provide the opportunity to possibly extent the analysis. The 3D data stacks may contain the structural information to

look specifically at mitochondria-associated ER membranes (MAMs) implicated in ER stress associated with obesity: Is it possible to visualize and quantify those in the existing data stacks?

Thank you for your suggestion. We quantified MAMs (ER and mitochondria junctions 0-16nm apart) content in 2000x2000x400 voxels (820um³) from randomly chosen 3 different cells in our FIB-SEM datasets derived from livers of lean, obese, and obese overexpressing Climp-63 mice. We show here that MAMs are significantly elevated in obese livers compared with the lean counterpart. Climp-63 overexpression in obese livers leads to significant reduction in MAMs. These data are presented below (**Fig. 1, in this letter**) and in **Extended Data Fig.8o**, revised manuscript.

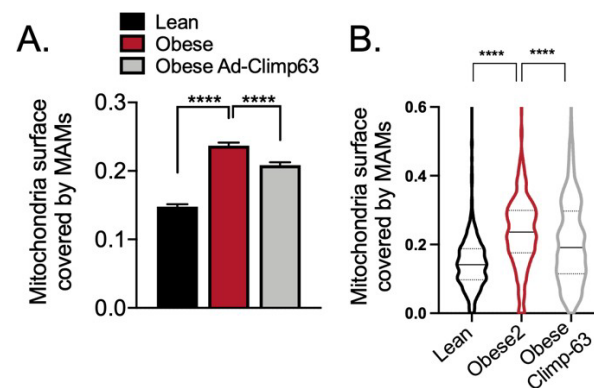


Figure 1: MAM content in hepatocytes from lean, obese, and obese expressing Climp63. Data acquired from FIB-SEM analysis.

2- If there is any chance to acquire enhanced FIB-SEM data on obese liver with altered ER structure by Climp-63 overexpression that would be fantastic.

Thank you for your suggestion. We have now performed the enhanced FIB-SEM analysis of livers from obese mice overexpressing Climp-63. The visuals and quantitative measurements of this dataset are provided in Fig. 4a-c, Extended Data Fig. 8, revised manuscript. Quantitative measures from FIB-SEM data show that Climp63 overexpression in obese mice rescue the ER sheet to tubule ratio to the levels comparable to lean controls (Fig. 2, this letter and Fig. 4b and 4c, revised manuscript). In addition to the FIB-SEM analysis, we also fractionated the rough and smooth ER compartments from lean, obese and obese overexpressing Climp63 livers and see that Climp63 exogenous expression also rescues the rough/smooth ER ratio as validated by this methodology as well (Extended Data Fig. 9a, revised manuscript).

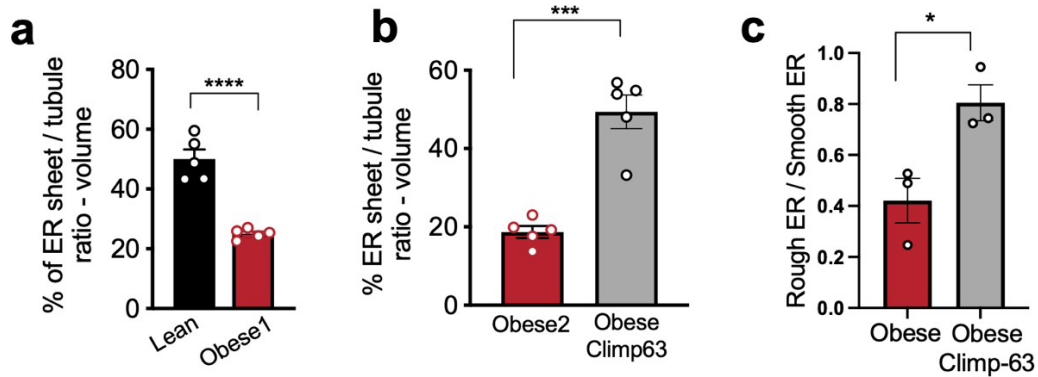


Figure 2: ER sheet to tubule ratio-volume in (a) Lean and obese mouse liver and (b) and those expressing Climp-63. (c) Biochemical fractionation of rough and smooth ER compartments from livers derived from obese mice and obese mice expressing Climp-63.

3- Concerning the consequences of Climp-63 overexpression in the liver of lean mice: Extended Data Fig. 7f indicates a striking increase of ER sheets in control mice. Is this quantified? Is this of relevance for liver metabolism in lean mice?

We have quantified the ER sheets abundance in EM pictures from lean (wild type) and lean exogenously expressing Climp-63. The quantification confirms that ER sheets are significantly increased upon Climp-63 overexpression in lean livers Extended Data Fig. 11a-c, revised manuscript. However, we did not detect any effect of Climp-63 overexpression in metabolic outcomes such as glucose production in primary hepatocytes isolated from lean mice (Fig. 3a, this letter and Extended Data Fig. 7e, revised manuscript). Accordingly, we did not detect any differences in systemic glucose and insulin tolerance when Climp-63 is overexpressed in vivo in lean mice (Fig. 3b and 3c, this letter and Extended Data Figs. 11d, e).

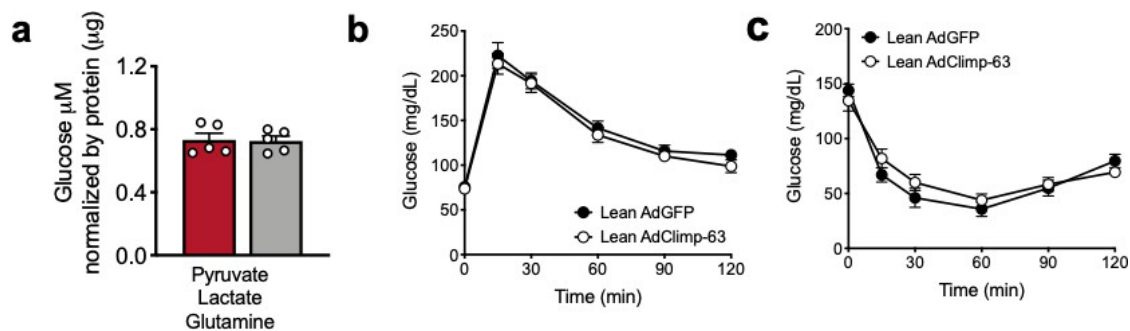


Figure 3: (a) Glucose production assay in primary hepatocytes from lean mice. (b) Glucose tolerance test in lean mice expressing Ad-GFP (control) and Ad-Climp63-myc adenovirus. (c) Insulin tolerance tests in the same mice.

4- How does the ER structure in obese mice and the Climp-63 induced change in ER structure relate to ER stress and UPR? Apart from the ASGR1 assay which is a read out for correct protein folding in primary cultured hepatocytes, it would be interesting to look at PERK activation or Bip expression in livers of lean and obese mice, but only if the authors feel this is indicated.

Thank you for your comment. As suggested, we did measure the expression levels of ATF4 and CHOP, two downstream components of the PERK pathway. We don't have reliable reagents to check for PERK activation directly in our samples. We observed that Climp-63 overexpression in obese liver resulted in decreased expression levels of both CHOP and ATF4 expression, two markers elevated in obesity (Fig. 4, of this letter and Extended Data Fig. 9b, revised manuscript). These results indicates that overexpression of Climp-63 in obese mice reduce ER stress response.

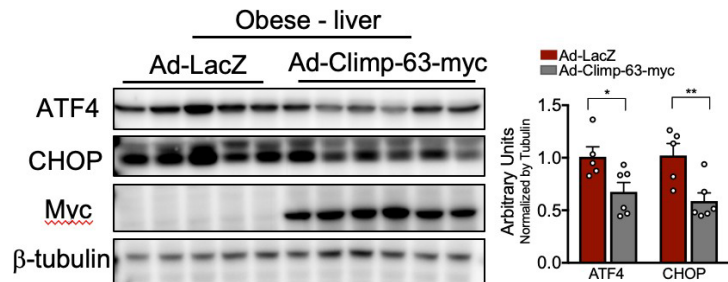


Figure 4: Immunoblotting (left) and quantification (right) of indicated protein in liver lysates from obese mice expressing Ad-LacZ and Ad- Climp-63-myc.

Regarding the regulation of the chaperone BIP (GRP78), we did not detect significant changes in its protein levels neither in total lysates nor in total ER fractions isolated from livers of obese mice overexpressing Climp63 compared to obese control mice (**Extended Data Fig 9d, 9e, revised manuscript**). However, we detected a significant change in BIP's preferential localization between the rough and smooth ER. While in lean mice, BIP abundance is more pronounced in the rough ER, in obese mice, BIP is similarly distributed between rough and smooth ER (**Fig. 5a, this letter and Extended Data Fig 4c, revised manuscript**). Interestingly, Climp-63 over-expression in obese livers shifts this ratio back to values similar to the ones found in lean (wt) mice (**Fig. 5b, this letter and Extended Data Fig. 9c, manuscript**). This was unexpected to us given that BIP is a luminal protein. Other luminal proteins such as Calnexin show evenly distribution between the two different ER fractions.

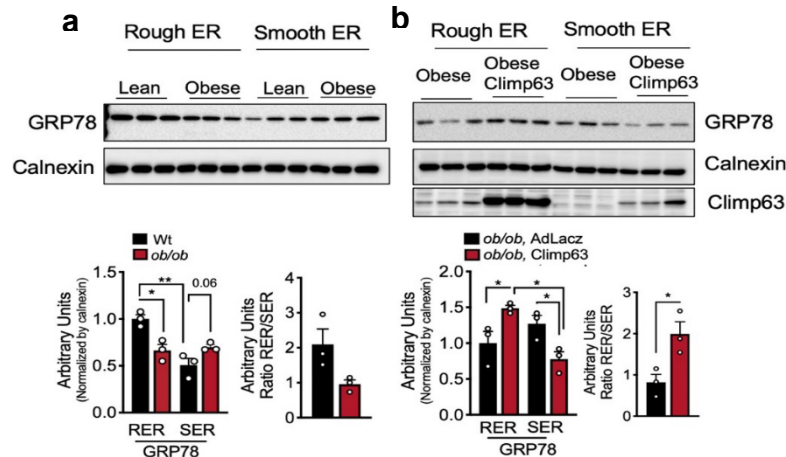


Figure 5: (a) Immunoblot (Upper panel) and quantification (lower panel) of indicated proteins in rough and smooth ER fractions isolated from livers of lean and obese mice. (b) Immunoblot (Upperpanel) and quantification (lower panel) of indicated proteins in rough and smooth ER fractions isolated from livers of obese mice expressing LacZ and Climp-63-myc adenovirus.

Minor points:

1-How many mice were used for the TEM analysis of the 15 full hepatocytes (Extended dataFig. 4)? This type of analysis offers the opportunity to increase n-numbers with limited effort, e.g. by analysis of 3 hepatocytes in 5 mice. However, the effects are so strong that increasingthe group size is probably not needed.

We thank the reviewer for this comment. We agree with the reviewer that this approach offers opportunity to have increased biological replicates. Accordingly, we quantified 20 EM sections from 4 different mice per group. As the reviewer mentioned, the effect is very consistent showinga decrease in ER sheets in obese hepatocytes (**Extended Data Fig. 2, revised manuscript**). We decided to focus our analysis on sections derived from fed mice and removed the quantification ofthe EM sections derived from fasted mice.

2-Fig. 3b: Western blot of ER fractions: Why normalized to tubulin? Is this just a mistake in y-axis labeling and calnexin was used for normalization?

Thank you for noticing this mistake. The results were normalized by Calnexin. The labeling is nowcorrected.

3-Fig 3f and g: The readout of the ASGR1 assay is unclear to me. What exactly is determinedby Cluc and Gluc luminescence? Some explanation in the methods section or in the legend of Extended Data Fig. 7 would be helpful.

We now added more detailed information about this reporter both in the methods section and in the legend of **Extended Data Fig. 9** as follows:

Methods: "In this reporter, ASGR protein lacking its membrane-anchoring domain was fused to luciferase Cluc. Conditions affecting ER folding capacity result in lower levels of secretion of theASGR-Cluc whereas

improved ER folding capacity have the opposite effect. In these assays variations such as expression level, transfection efficiency and cell densities are controlled by normalization of Cluc signal to a constitutively secreted luciferase derived from Gaussia Luciferase, Gluc”.

Legend Ext Fig. 9: “In this reporter, Cluc indicates the folding and secretion of ASGR-Cluc fusionprotein and Gluc, which is constitutively expressed in the construct is used as a normalization of expression level, transfection efficiency and cell densities”.

4-Fig. 4: Small typo in the legend of Fig. 4b: This is the quantification of the images shown in figure 4a (not 5A), right? Question: From how many animals were the analyzed EM sections taken? Line 498: “Representative of 3 independent cohorts”: This is unclear to me. Does this mean 3 cohorts of mice with n=8 (LacZ) and n=10 (Climp-63)?

Thank you for pointing to the typo in Fig. 4b (now, **Fig. 4d, e**, revised manuscript). You are correct, this should have been Fig. 4a instead of 5a. We corrected this label. We also increased the number of analyzed EMs to n=18 (Control) and n=30 (Climp-63) and these images are from 3 different mice per each group. Regarding line 498, yes, it means 3 independent cohorts of study, mice number varying from 8 to 10 per group in each cohort.

5-Extended Data Fig. 4: What is the unit (y-axis label) in e, f) and g)? In addition, it may be helpful to add a small sentence to explain why fed and fasted states were compared.

We updated the units for Fig. 4e, 4f and 4g (now **Extended Data Fig. 2**, revised manuscript) per image scale and added it in the figures. As we mentioned above, we decided to focus our analysis on the fed state and therefore removed the data derived from mice in the fasted state.

6-Extended Data Fig. 6: The main text refers to Extended Data Fig 6f and g, but these subpanels unfortunately don't exist.

Thanks for pointing this out and we apologize for this oversight. This is now included.

Referee #3:

This study by Parlakgöl and Arruda et al. uses state-of-the art FIB-SEM methods to bring to life the architecture of organelles in their native tissue context. The authors observe striking changes to ER shape that vary based on the metabolic state of the organism (lean vs. obese). The authors have pushed FIB-SEM technology to the limit. This is a major technical achievement that required tackling technical and computational obstacles and will inspire future efforts using their tools on a wide array of biological questions. In addition, the interplay between organelle shape and cellular and organismal metabolism dissected in this study will be interesting to a broad audience. I believe this study is worthy of publication with some revisions (described below). Concerns regarding data interpretation are also listed below.

We thank the reviewer for these kind words and interest in our manuscript.

1) By altering the level of ER shaping proteins, the authors identify physiological and metabolic consequences of ER shape at both cellular and organismal levels. I have two suggestions to improve and extend the interpretation of these data. First, an alternative strategy to “rescue” ER shape besides Climp63 overexpression should be used. I suggest depletion of Rtn4. If the metabolic results are similar, then one can conclude that shape is important.

Thank you for this suggestion. We agree with the reviewer that an alternative model to modulate structure would make our conclusion stronger. Following your suggestion, we used an adenovirus expressing shRNA against Rtn4 in livers of obese mice in an attempt to downregulate Rtn4 and increase ER sheets in the hepatocytes (**Fig. 6a, this letter**). Unfortunately, this intervention was not sufficient to increase ER sheets (or reduce tubules) in the obese hepatocytes as shown in our ultrastructural analysis (**Fig. 6b, this letter**). We also didn't see any differences in lipid accumulation in the liver and systemic glucose tolerance (**Fig. 6c and 6d, this letter**). It is possible that a more complete depletion in the liver may result in structural and resulting metabolic improvements, as seen in the deletion model in the "whole body of mice" (PMID: 32738448). However, whether this is related to changes in ER structure in the liver was not explored and these animals are not available for structural analysis.

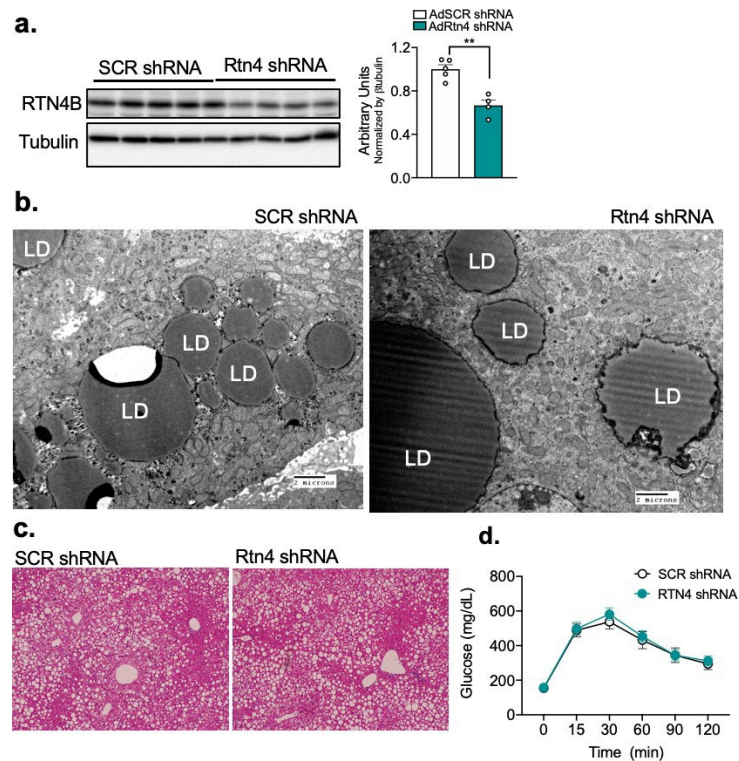


Figure 6: (a) Immunoblot (left panel) and quantification (right panel) of Reticulon 4b (Rtn4b) expression in liver lysates from obese mice expressing Scrambled (SCR) shRNA and shRNA against Rtn4. (b and c) TEM and H&E histological analysis of livers from obese mice expressing SCR shRNA and shRNA against Rtn4. (d) Glucose tolerance test in obese mice expressing SCR shRNA and shRNA against Rtn4.

It is also possible that lack of a phenotype after downregulation of Rtn4 in obese liver could be attributed to the complexity of the tubular ER organization in hepatocytes. Different from cell lines, hepatic tubular ER structure is very complex and there is high level of redundancy in proteins that support tubular structures. As such, it is possible and even likely that the downregulation of a single ER tubule shaping protein would not be enough to change the ER morphology, making loss of function approach challenging (in fact, we have similar experience in another experiment where we targeted Reep5). These results are nevertheless valuable in showing that manipulation of these proteins does not nonspecifically generate metabolic changes and reinforces the concept that structural alterations are required for metabolic regulation in liver.

Given the potential limitations of the loss of function approaches described above, we employed

an additional gain of function strategy as an alternative to understand the relationship of ER shape and metabolism. For that, we took advantage of a construct expressing the Reticulon-homology-domain (Rtn-HD) of the Rtn-4 protein, which was previously developed by Tom Rapoport's group

(PMID: 18442980) and tested in cell lines. Expression of Rtn-HD in livers from lean mice led to a marked enrichment of ER tubules (**Fig. 8a, this letter and Fig. 3c and Extended Data Fig. 6a, revised manuscript**). These livers showed increased levels of lipid droplets and triglycerides (**Fig. 8b and 8c, this letter and Extended Data Fig. 6b, revised manuscript**) compared with livers from lean mice expressing GFP as a control. Interestingly the lipid droplets were presented in areas surrounded by the enriched ER tubules (**Fig. 8a, this letter and Extended Data Fig. 6a, revised manuscript**). We also observed that Rtn-HD overexpression led to increased levels of SCD1, a key enzyme for de novo lipogenesis (**Fig. 8d this letter and Extended Data Fig. 6c**). Thus, changing ER sheet/tubule ratio towards tubule formation in the liver results in alterations in lipid metabolism resembling the ones found in obesity (albeit in a lower degree). In this model, we did not detect loss of ER sheets as observed in obesity and no changes were evident in glucose homeostasis in the lean mice (**Fig. 8e this letter and Extended Data Fig. 6d**).

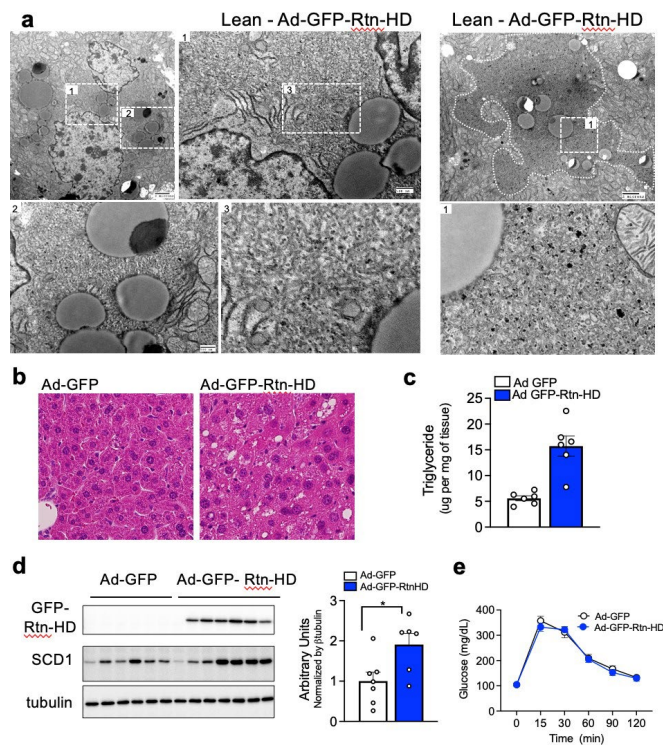


Figure 8: (a and b) TEM and H&E histological analysis of livers from obese mice expressing control(Ad-GFP) and Rtn-HD adenovirus. (c) Triglyceride levels measured in the livers from lean mice expressing control (Ad -GFP) and Rtn-HD adenovirus (d) Immunoblot (left panel) and quantification(right panel) of indicated proteins in liver lysates from lean mice expressing control (Ad-GFP) and

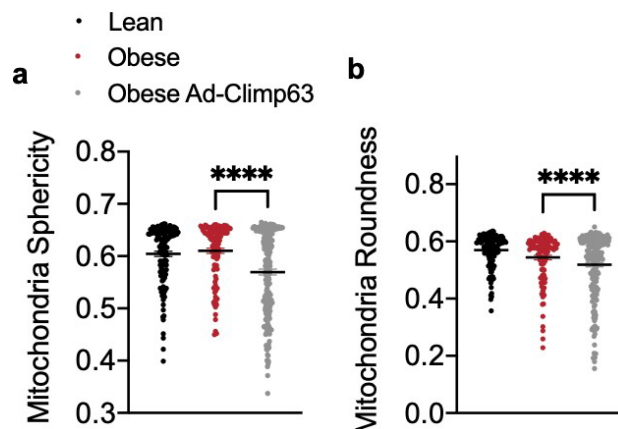
Rtn-HD adenovirus (e) Glucose tolerance test in mice expressing control (Ad-GFP) and Rtn-HD adenovirus.

[REDACTED]

Climp63 expression in obese mice will reduce tubular ER and because the biogenesis of both LDs and mitochondria is regulated by tubular ER contact sites, it would be highly informative if the authors would also map the morphology of mitochondria and abundance of LDs when sheets are expanded- particularly because these organelles are both linked to metabolism as well. Presumably these data are all present in their FIB-SEM reconstructions.

We thank the reviewer for these suggestions. We have determined the total abundance of lipid droplets in the hepatocytes in all our datasets of FIB-SEM. We observed that in lean hepatocytes (5 different cells per volume) the lipid droplet content represents less than 1% of total cell volume. In hepatocytes from obese mice, the abundance of lipid droplets were markedly increased to 30 to 34% (**Fig. 1i, revised manuscript**). We now show in our new FIB-SEM reconstruction from obese mice overexpressing Climp-63 that the lipid droplet content in obese hepatocytes after the induction of ER tubule to sheet conformation decreases to 11% in average (2.5, 9.4, 7.9, 18 and 21% in 5 separate cells). (**Extended data Fig. 8n, revised manuscript**). The changes in lipid droplets and triglycerides by Climp-63 overexpression are also shown in our histological analysis and triglyceride measurements shown in **Fig 4f and 4g, revised manuscript**.

We have also determined the mitochondria morphology in our different FIB-SEM datasets (**Extended Data Fig. 8o-q, revised manuscript**), as suggested. For this quantification, we randomly choose 2000x2000x400 voxels (820um³) from 3 different cells in our FIB-SEM datasets derived from livers of lean, obese, and obese overexpressing Climp-63 mice. From this analysis we observed that Climp-63 overexpression in obese mice led to mitochondria with decreased sphericity and roundness, therefore, more elongated mitochondria. We are performing the quantification of the number and shape mitochondria in the full hepatocytes of all of our data sets (which is very challenging) for a separate study.



- 1) Ext Data Fig 5, a-e: This experiment can easily be confounded by contamination, as only the weights of the fractions are reported. This limitation is not adequately addressed by EM. The authors should perform immunoblotting on organelle, ribosome, and cytosolic markers on the total (input) and collected fractions to confirm that their enrichment procedure is effective.

We thank the reviewer for this suggestion. To evaluate the enrichment of our ER fractions, we performed western blot analysis for different organelle and cytosolic markers in fractions enriched in the ER, mitochondria, and cytosol, in addition to total lysates. As shown in **Extended Data Fig. 3a, revised manuscript**, our ER fractions are significantly devoid of contamination.

2) Ext Data Fig 5i: Unlike in panel h, where blot quantification and statistical tests are performed, the conclusion that rough ER proteins “translocated to the tubular ER” (main

text, lines 204-205) are drawn from unquantified blots and no statistics. The conclusions drawn are not evident to me, except possibly for Sec61beta, although that is also variation in the lean mouse smooth ER. It is suggested that GRP78 – a soluble luminal protein – is a rough ER protein (not sure why). Can the authors clarify how GRP78 would segregate between smooth and rough ER? Or, is GRP78 better interpreted as an indicator of the total ER content of these fractions (for which there is otherwise no normalization)?

Thank you for this comment and we apologize for not providing quantification and loading controls for this set of western blots. We repeated this experiment and performed additional immunoblotting for an array of ER proteins as shown in **Extended Data Fig 4b, revised manuscript**. In this new experiment, we confirmed the finding that GRP78 segregates differently between ER sheets and tubules (**Fig. 10a, this letter and Extended Data Fig 4c, revised manuscript**). While in livers from lean mice, GRP78 is enriched in the ER sheets, in obese the proteins are similarly distributed in ER sheets and tubules. Importantly, the localization of GRP78 preferentially in the rough ER fraction is rescued in liver tissues from obese mice overexpressing Climp63 (**Fig. 10b, this letter and Extended Data Fig 9c, revised manuscript**). Calnexin, another ER luminal protein shows even distribution among our RER and SER fractions. The association of GRP78 with rough ER has been reported before in macrophages (PMID: 29774449).

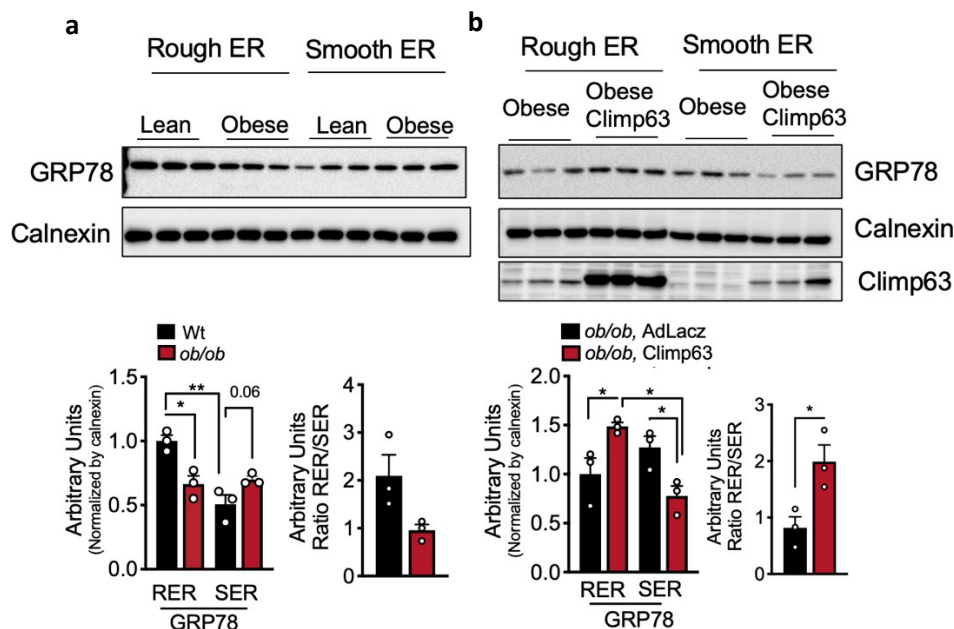


Figure 10: (a) Immunoblot (Upper panel) and quantification (lower panel) of indicated proteins in rough and smooth ER fractions isolated from livers of lean and obese mice. (b) Immunoblot (Upper panel) and quantification (lower panel) of indicated proteins in rough and smooth ER fractions isolated from livers of obese mice expressing LacZ and Climp63-myc adenovirus.

2) Fig 3, a-b: If all Climp63 is in the ER, then why are opposite conclusions drawn between the total cell extract (panel a) and ER fraction (panel b)? This leads me to believe that the purifications are variable. It is difficult to assess this experiment as displayed because no evidence is provided that ER is specifically and efficiently enriched in this assay. Additionally, all blot quantifications are labeled as normalized to tubulin. How is this possible for panel b, which should not contain any tubulin after the ER enrichment procedure? The opposite effect is seen for p180, which is reduced in total liver lysates from the obese condition but not significantly changed in the total ER fraction.

We have repeated the total ER fractionation from lean and obese liver tissues from 5 mice per group. We now provide markers of organelle enrichment in different fractions as shown in **Fig. 11, this letter and Extended Data Fig 5a, revised manuscript**. We detected that p180 is downregulated both in total lysates and total ER fractions consistently. Regarding Climp-63, according to our previous results, no differences in protein levels are found in the total lysates, however Climp-63 is significantly decreased in the total ER fraction. It has been reported that in addition to ER, a small fraction of Climp-63 can be found in the plasma membrane in non-parenchymal cells (PMID: 12913003). These cells have less ER content compared to hepatocytes in the liver. Maybe this can explain the apparent pattern here. Regarding the quantification normalized by tubulin, it was a labeling mistake, and we thank the reviewer for pointing this out. The protein was normalized by Calnexin, and this is corrected now in the revised manuscript.

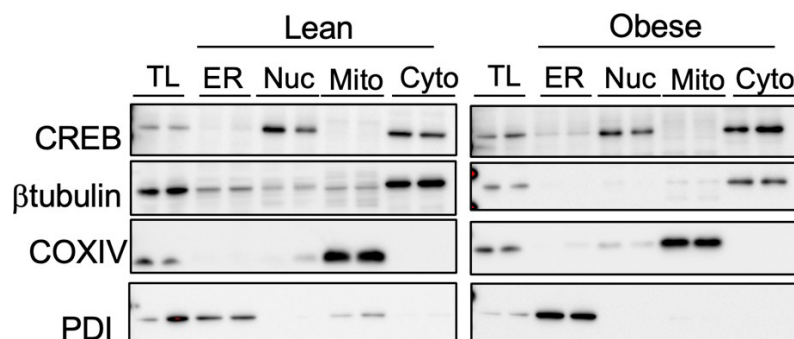


Figure 11: Immunoblotting of indicated proteins in total lysates (TL), endoplasmic reticulum (ER, Nucleus (Nuc), mitochondria (Mito) and cytosol (Cyto) in subcellular preparations from livers from lean and obese.

3) -Ext Data Fig 7, c-d: The resolution of the images shown is not sufficient to conclude that ER shape has been altered.

We now provided better images with monochromatic channels that show the marked proliferation of patches of ER sheets with Climp-63 expression in hepatocytes (**Extended Data Fig. 7c, d, revised manuscript**).

4) -Fig 3, d-e: This experiment simply does not test what is claimed in the text (lines 264-266), namely, that overexpression of Climp-63 restores Sec61beta localization to ER sheets. Are the authors able to colocalize endogenous Sec61beta with exogenous Myc-tagged Climp-63 in ER sheets? Are the total levels of Sec61beta (by immunoblot) changed?

We evaluated the colocalization of exogenous Myc-tagged Climp-63 with endogenous Sec61betaprotein in Extended Data **Fig. 7d, revised manuscript**. Pearson coefficient for co-localization of 2 signals, evaluated by confocal microscopy, ranged from 0.73 to 0.93 (average 0.84) and Mander's coefficient from 0.66 to 0.92 (average 0.82) (n=29 images). The total levels of Sec61beta were not changed, as assessed by western blot analysis.

-Fig 3, f-g: Related to the above point, folding capacity and translocation capacity seem to be used interchangeably within the logical flow of the paper despite being distinct processes. Are the results with the ASGR reporter explained by changes in the levels of translocation machinery or chaperones? Perhaps a correlative argument could be made by blotting for chaperones and translocon subunits in the extracts (lean vs obese; obese control vs Ad- Climp63). This could be done in conjunction with my comments about improving Ext Data Fig 5i above, as it uses a similar readout.

We did not detect changes in the expression levels of the chaperone BIP (GRP78) in total lysates and total ER isolated from livers of obese mice and obese mice expressing Climp-63 (**Extended Data Fig. 9d, e**). However, as mentioned above we detected a significant change in BIP preferential localization between the rough and smooth ER (**Fig. 10a, this letter and Extended Data Fig. 9c, revised manuscript**). Climp-63 overexpression shifts this ratio back to values similar to the lean (wt) mice (**Fig. 10b, this letter and Extended Data Fig. 9c**). This suggests that restoring ER sheets by Climp63 expression, re-localizes chaperone BIP in the same sub-domain (ER sheets) as the translocon complex and ribosomes.

Minor:

1) In the main text, it is unclear how many mice are examined (line 87 and following). Is the plural "lean mice" used to refer to a single mouse? The sample size of the FIB-SEM portion of the study should be clearly explained.

For the FIB-SEM analysis, we utilized one lean mouse, two obese mice and one obese mouse overexpressing Climp63. This is clarified in the manuscript. The volume we imaged corresponds to 24 to 27 full or partial hepatocyte volumes in each dataset. To our knowledge this is the highest volume imaged in the tissue context and involved an extraordinary amount of microscopy and computational work.

2) Figure 1, a and b: What is identified as an ER tubule (panel a)? The ER appears to be predominantly sheets in this region. The inset showing the highest level of zoom is not correctly drawn on the medium zoom inset. Panel b does not contain the high level of zoom inset, precluding comparison of the two conditions.

We removed that inset since the information was not clear.

3) Figure 1, i and j: (1) labels are switched in the legend; (2) it is confusing that these very similar plots use % (in i) and ratio (in j); (3) it seems odd to highlight in the main figure the significant difference in ER volume, while Ext Data Fig 2 shows that this is due to a trivial explanation. This manner of presentation risks misinterpretation.

We corrected the legend for **Fig. 1i and 1j**. We also changed the Y axis of the plots for these graphs. Thanks for the clarification. We moved the previous Ext Data Fig 2 to the main figure in **Fig. 1k** (revised manuscript)

to avoid any misinterpretation. Thanks for this suggestion.

4) In the methods section, authors describe culturing Hep1-6 and Cos-7 cells under 10% CO₂, when datasheet for these cells clarify they use 5% CO₂. Do the authors have a particular reason to use this condition

The cells were cultured in 5% CO₂. We are grateful for the detailed and careful analysis.

5) Movie 3: Are lipid droplets modeled inside the nucleus?

No, this is an annotation mistake.

6) Ext Data Fig 5, f-g, is at odds with its description in the main text (line 190). Which subunit is stained in this experiment: Sec61alpha or -beta?

We stained Sec61  and corrected this in the text.

7) Main text, line 236, should refer to Ext Data Fig 6e.

We thank the reviews and now corrected this mistake in the revised manuscript.

8) Ext Data Fig 7: Here (and potentially elsewhere) the identity of the fusion proteins should be described accurately. Sec61beta is a tail-anchored protein that must be tagged on its N terminus (GFP-Sec61beta, not Sec61beta-GFP). Likewise, KDEL is an obligatory C-terminal peptide and should be GFP-KDEL. Due to these errors, it is unclear from the figure how you have tagged other proteins like Climp-63.

The tagged Sec61-beta plasmid was obtained from Tom Rapoport's lab (Addgene: 15108). We corrected the naming of tagged proteins.

9) Methods and Supplementary Table 2: There are numerous reagents listed that do not appear anywhere in the study that I can find.

We went over the list and updated it. Thank you for the input.

Referee #4 (Remarks to the Author):

This is an interesting manuscript reporting the acquisition and deconvolution of images obtained from powerful FIB-SEM technology in a collaboration with the HHMI Janelia Research Campus group. The images obtained are novel in 3D details, with the important advantage that cells within intact tissue are evaluated. While Figs 1 and 2 are informative in describing hepatocyte ER structures and effects of obesity, they are not much more informative in conclusions than what can be learned from normal EM techniques. Nonetheless, they add to the many descriptions of obesity effects in the literature, albeit only from one extreme genetic model. And clearly the data is well presented and organized, and the paper well written.

The major problem with the manuscript is that the main conclusion is not properly evaluated and critically tested. The impact of the studies is therefore limited, and advancement of the field

incremental. The authors' claim that ER structure determines physiological function is weak on four levels:

1. The ER structural changes observed in Figs 3 and 4 are purported to cause the physiological changes in hepatocyte lipogenesis and glucose output, but this is actually a correlation not a demonstration of causality. The genetic manipulation is assumed to be all due to the ER structural change, yet it is possible and even likely that other direct effects of Climp-63 expression are induced independent of ER structure. There is no data to exclude that problem.

We thank the reviewer for these remarks. To date, the vast majority of the published papers show the primary function of Climp63 on the regulation of ER sheets with very limited other functions identified to date. Having said that, it is possible that functions of Climp-63 other than the regulation of ER structure, could have an implication on the metabolic phenotype described here. In the absence of any other known function or relevance to metabolism, dissecting this point experimentally is challenging. However, we did take several and extensive additional efforts to link structure to metabolic function using alternative strategies and generated strong data that these are relevant for metabolism. Taken together, the evidence generated from multiple structural interventions support our conclusions in the revised manuscript in a much stronger manner, as detailed below.

In order to further interrogate the relationship between ER architecture and metabolic regulation we tested whether the manipulation of ER structure towards tubule formation in a lean mouse would have a metabolic implication which resemble the metabolic alterations found in obese liver. For that we used a construct expressing the Reticulon homology domain (Rtn-HD) of the Rtn-4 protein, which was previously tested in cell lines (PMID: 18442980).

As shown in **Fig. 12a, this letter and Fig. 3c and Extended Data Fig. 6a, revised manuscript**, expression of Rtn-HD in livers from lean mice led to a marked enrichment of ER tubules. These livers showed increased levels of lipid droplets and triglycerides (**Fig. 12b and c, this letter and Extended Data Fig. 6b**) compared with livers from lean mice expressing GFP as a control. Interestingly the lipid droplets were presented in areas surrounded by the enriched ER tubules (**Fig. 12a this letter and Fig. 3c and Extended Data Fig. 6a**). We also observed that Rtn-HD overexpression led to increased levels of SCD1, a key enzyme for de novo lipogenesis (**Fig. 12d, this letter and Extended Data Fig. 6c**). Thus, changing ER sheet/tubule ratio towards tubule formation in the liver results in alterations in lipid metabolism resembling the ones found in obesity (albeit in a lower degree). In this model, we did not detect loss of ER sheets as observed in obesity and no changes were evident in glucose homeostasis in the lean mice (**Extended Data Fig. 6d, revised manuscript**).

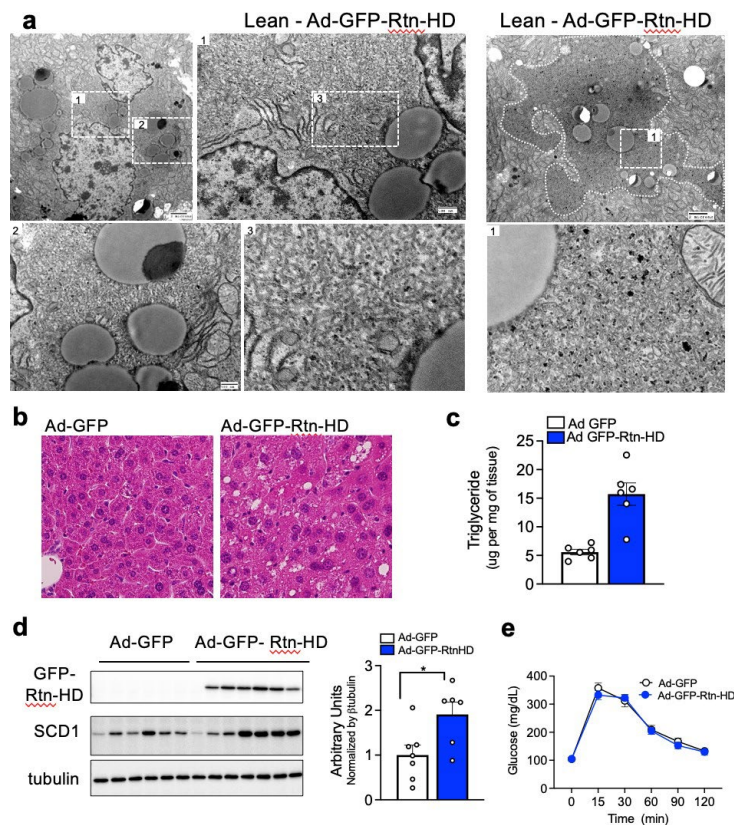


Figure 12: (a and b) TEM and H&E histological analysis of livers from obese mice expressing control(Ad-GFP) and Rtn-HD adenovirus. (c) Triglyceride levels measured in the livers from lean mice expressing control (Ad -GFP) and Rtn-HD adenovirus (d) Immunoblot (left panel) and quantification(right panel) of indicated proteins in liver lysates from lean mice expressing control (Ad-GFP) and Rtn-HD adenovirus (e) Glucose tolerance test in mice expressing control (Ad-GFP) and Rtn-HD adenovirus.

[REDACTED]

2. Figure 3 shows Climp-63 expression in hepatocytes from obese mice in vitro and indicates an inhibition in glucose output. This is an inhibitory effect, raising the question of whether it is a result of toxicity vs specific regulation. What is the effect of Climp-63 expression on cells from lean mice? Is glucose output also decreased in “lean” cells? Are other features of the metabolism of these cells also inhibited?

We thank the reviewer for these comments and suggestion. We did perform these experiments as suggested. In the lean animals, Climp-63 overexpression in hepatocytes did not result in any difference in metabolic parameters such as glucose production in vitro and glucose and insulin tolerance test in vivo (**Fig. 14 a-c, this letter and Extended Data Fig. 7e and 11d, e, revised manuscript**). These results are important in addressing the potential toxicity concern expressed by the reviewer.

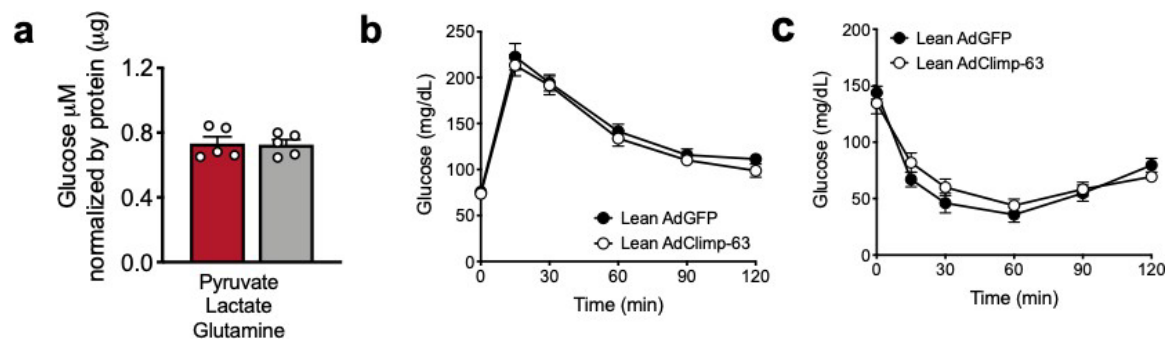


Figure 14: (a) Glucose production assay in primary hepatocytes from lean mice. (b) Glucose tolerance test in lean mice expressing control (Ad-GFP) and Ad-Climp63-myc adenovirus. (c) Insulin tolerance test in the same mice.

3. What is the effect on mitochondrial respiration, lipogenesis and lipid oxidation? The specificity of this effect is not addressed nor is a KO strategy imposed, which should show opposite effects on the above metabolic parameters if toxicity is not a factor.

We thank the reviewer for these comments and suggestions. We performed multiple experiments to address these questions and obtained very insightful results. First, we detected that Climp-63 overexpression led to significant decrease in de novo lipogenesis rates measured both in isolated primary hepatocytes (**Fig. 15a, this letter and Fig. 3h, revised manuscript**) and in the liver in vivo (**Fig. 15b, this letter and Fig. 4h, revised manuscript**). The decreased de novo lipogenesis levels were accompanied by a dramatic reduction in the expression of the ER protein SCD1 in liver tissues overexpressing Climp-63, a key limiting step in the de-novo lipogenesis pathway (**Fig. 15c, this letter and Extended Data Fig. 10a, revised manuscript**). Together with the decreased de novo lipogenesis, we also detected that gain of function of Climp-63 led to increased levels of fatty acid oxidation in primary hepatocytes from obese mice as measured by ^{14}C labeled CO_2 release after loading the cells with ^{14}C palmitic acid (**Fig. 15d, this letter and Fig. 4i, revised manuscript**).

We do not have access to a knockout model for Climp63 and performing this experiment was not possible within a reasonable time frame. Furthermore, it's been shown in the original paper that revealed Climp63 as a key regulator of ER sheets, that deletion of Climp63 in cells does not result in the disappearance of ER sheets and does not significantly affect the ER sheet to tubule ratio (Shibata, Rapoport, et al., Cell, 2010).

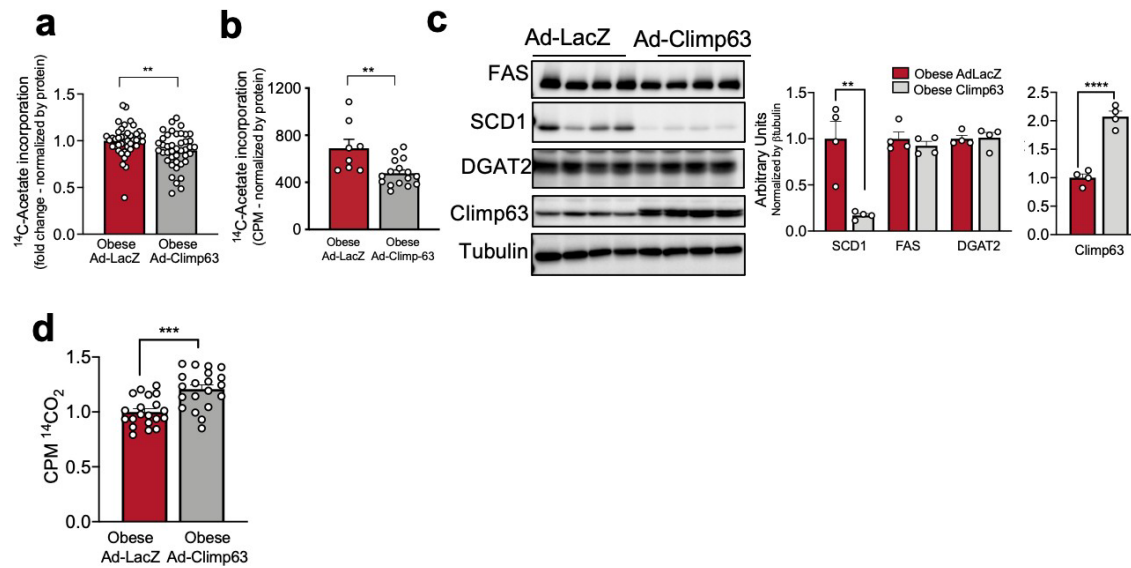


Figure 15: (a) in vitro de novo lipogenesis in primary hepatocytes from obese mice expressing Ad-LacZ and Ad-Climp-63. (b) in vivo de novo lipogenesis in obese mice expressing Ad-LacZ and Ad-Climp-63. (c) Immunoblotting of indicated proteins in total lysates from obese mice expressing Ad-LacZ and Ad-Climp-63. (d) Fatty acid oxidation assay in primary hepatocytes from obese mice expressing Ad-LacZ and Ad-Climp-63.

3. Figure 4 again shows only results from obese mice and lacks evaluation of the control lean mice. Without control, lean mice, other interpretations are possible. Is this a toxicity effect?

We addressed this issue in our revised manuscript. Overexpression of Climp-63 in lean hepatocytes did not result in metabolic outcomes such as reduction in glucose production in primary hepatocytes isolated from lean mice (**Fig. 14a, this letter and Extended Data Fig. 7e, revised manuscript**). We also did not detect any differences in systemic glucose and insulin tolerance when Climp-63 is overexpressed in vivo in lean mice (**Fig. 14b and c, this letter and Extended Data Fig. 11d, e, revised manuscript**).

4. Is the improved GTT due to changes in insulin levels? Importantly, does hepatic specific KO of Climp-63 mimic the effects of obesity on ER Structure and metabolic regulation (glucose output and lipogenesis). This is an easy (AAV injection) and important experiment as increases in these metabolic pathways would be expected in the KO, and if observed, would argue in the authors' favor.

Insulin levels were not changed in obese mice expressing Climp-63 compared with the controls (**Fig. 16a, this letter**). However, in the newly performed experiments, we saw that response to insulin was significantly improved with Climp-63 expression in obesity, assessed by in vivo insulin signaling experiment (**Fig 16b, this letter and Fig. 4j, revised manuscript**).

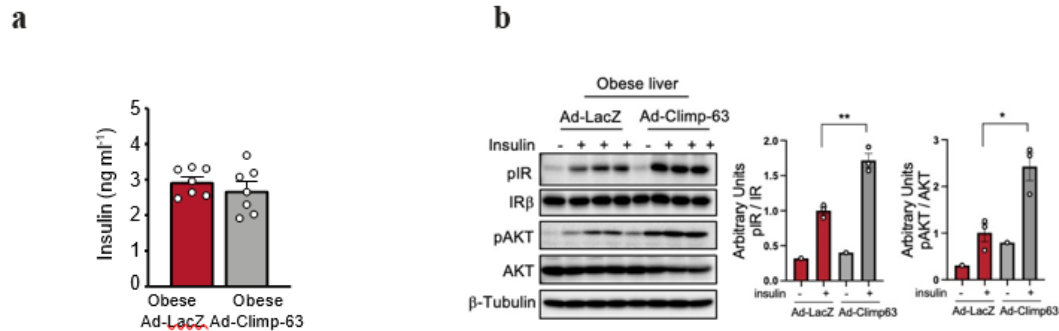


Figure 16: (a) Insulin levels in obese mice expressing LacZ and Ad-Climp-63 measured by Elisa assay.

(b) In vivo insulin signaling, determined by phosphorylation levels of insulin receptor (at Tyr972) and AKT (at Thr308).

4. The notion of ER structure causing metabolic changes is interesting, but unsatisfying unless the mechanism is revealed of how such ER structure is connected to the enzymes and fluxes through the metabolic pathways. There appear to be no measurements of gluconeogenesis or lipogenesis enzyme levels, nor whether they are displaced from ER or other structures.

In the revised manuscript, we provided additional information regarding these questions. We observed that increasing the ER sheet/tubule ratio by Climp-63 exogenous expression led to decreased de novo lipogenesis. This is accompanied by a drastic decrease in SCD1 expression (**Fig. 17a, this letter and Extended Data Fig. 10, revised manuscript**). SCD1 is an ER membrane protein markedly upregulated in the liver in obese conditions in both RER and SER (**Fig. 17b, this letter**). As shown in **Fig. 17c, this letter**, SCD1 is downregulated in both ER subfractions upon Climp-63 overexpression, being this effect is more pronounced in the tubular ER. These data suggest that ER sheet proliferation leads to decreased de novo lipogenesis.

We also observed that GRP78 segregates differently between ER sheets and tubules (**Fig. 10a, this letter and Extended Data Fig 4c, revised manuscript**). While in livers from lean mice, GRP78 is enriched in the ER sheets, in obese the proteins are similarly distributed in ER sheets and tubules. Importantly, the localization of GRP78 preferentially in the rough ER fraction is rescued in liver tissues from obese mice overexpressing Climp63 (**Fig. 10b, this letter and Extended Data Fig 9c, revised manuscript**). Calnexin, another ER luminal protein shows even distribution among our RER and SER fractions.

Thus, one clear consequence of changes in ER sub-domains in hepatocytes is the alteration of proper representation and localization of its resident proteins.

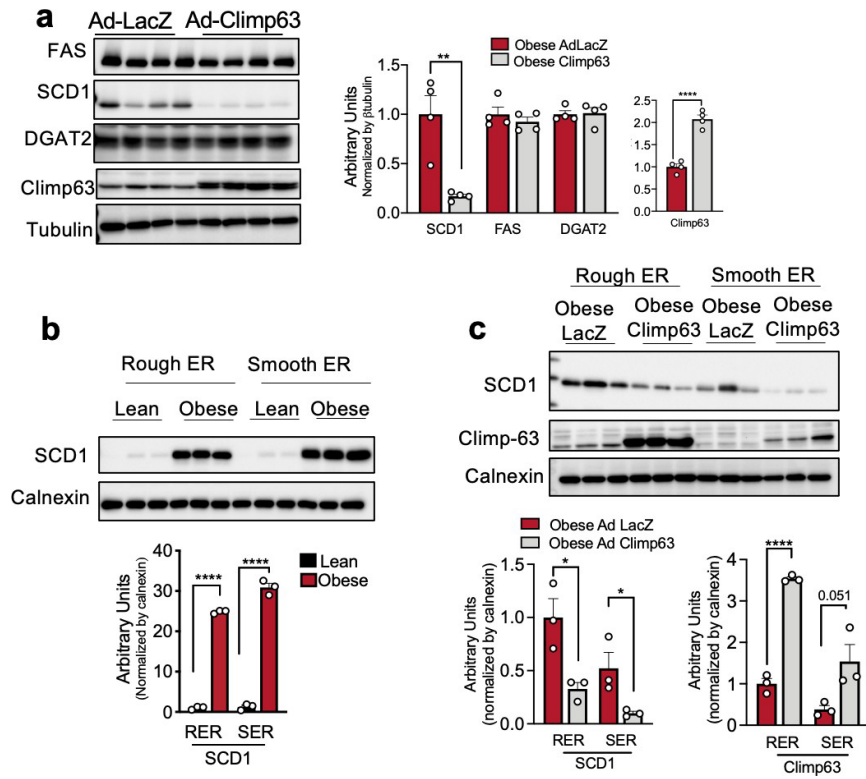


Figure 17: (a) Immunoblot of indicated proteins in total lysates from obese mice expressing Ad- LacZ and Ad-Climp-63. **(b)** Immunoblot and quantification (lower panel) of indicated proteins in rough and smooth ER fractions isolated from livers of lean and obese mice **(b)** Immunoblot (and quantification (lower panel) of indicated proteins in rough and smooth ER fractions isolated from livers of obese mice expressing Ad-LacZ and Ad-Climp-63-myc.

5. Mitochondria are known to be physically associated with the ER, and this would be a way that ER structure could affect oxidative metabolism (which is connected to lipogenesis and gluconeogenesis). But there is no evaluation of this issue. What is the way in which ER structure could mechanistically regulate such metabolic pathways is central to this study, and if it can be answered, it would negate the criticisms in items 1-3 above, as those issues could be directly tested in the context of such a molecular mechanism?

Thank you for this insight. The tight contacts between mitochondria and ER are preferentially formed by smooth ER tubules and mitochondria fission occurs at these tight contact sites. Changing the ER tubule content by shifting the ER toward ER sheets could impact these structures and their functions. We evaluated MAMs, mitochondria morphology and de novo lipid synthesis and oxidation. We quantified MAMs (ER and mitochondria junctions 0-16nm apart) content in 2000x2000x400 voxels (820um³) from 3 randomly chosen different cells in our FIB-SEM datasets derived from livers of lean, obese, and obese overexpressing Climp-63 mice. We show here that MAMs are significantly elevated in obese livers compared with the lean counterpart. Climp-63 overexpression in obese livers leads to significant reduction in MAMs (Fig. 18a-b, this letter and Extended Data Fig. 8o, revised manuscript). We also detected that mitochondria were more elongated (lower sphericity and roundness) in Climp-63 expressing tissue Fig. 18c-d, this letter and Extended Data Fig. 8p-q, revised manuscript).

It's been shown that mitochondria in the elongated state demonstrate higher efficiency in fatty acid oxidation (PMID: 23562075). In fact, we measured fatty acid oxidation by ^{14}C labeled palmitic acid tracing and showed that overexpression of Climp-63 significantly increased fatty acid oxidation in primary hepatocytes. We also showed that Climp-63 overexpression led to decreased de novo lipogenesis and marked reduction of SCD1 expression. Taken together, these extensive new experiments provide substantial insights into how the changes in ER structure impact lipid and glucose metabolism in the liver and impact insulin sensitivity.

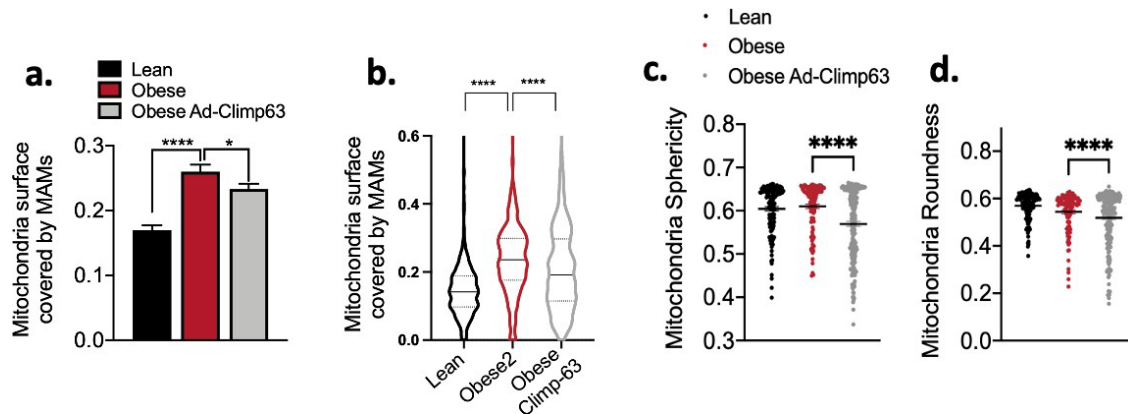


Figure 18: MAM content in hepatocytes from lean, obese and obese expressing Climp63.

We believe with these new insights, and numerous additional in vitro and in vivo experiments, our extensively revised manuscript presents a much stronger resolution of structural regulation of subcellular organelles in obesity and the role in hepatic and systemic metabolic responses. We thank the reviewers for their careful analysis, comments, and constructive criticisms.

Reviewer Reports on the First Revision:

Referee #1:

The authors have tried to address my previous comments but there are problems with the revisions. The manuscript now includes references to Supplementary Tables 1 and 2, but no Supplementary Table 1 is provided and the file that is linked as Supplementary Table 2 is just a spreadsheet. The spreadsheet includes accuracy, precision, recall and F1 scores but little information is provided on how these have been generated, either in this file or in the manuscript. What architecture was used for the deep learning models? Was a single classifier trained for each organelle using a mixture of lean and obese images? Or were separate classifiers trained for lean and obese? The text box in the spreadsheet mentions that the evaluation was performed on downsampled images. Was the analysis in the manuscript also done using downsampled images? The data on ER sheets and tubules is presumably calculated pixel-wise but that is not stated. Again, whether a separate classifier was trained for lean and obese was not mentioned. This is important because if separate classifiers were trained that could incorporate unconscious bias by the investigators about how these differ between lean and obese.

The manuscript mentions Jaccard index evaluation in Supplementary Table 2 but that is not present in the spreadsheet.

The manuscript refers to an EMPIAR deposition containing the training and testing data but searching EMPIAR for 47483731 does not find anything.

The source code branch on github has not been updated since the original submission so presumably any changes made for the revised manuscript are missing.

Referee #2:

In my view the authors have addressed all my questions, concerns and suggestions, invested significant effort and added valuable information to the manuscript. I am also delighted that a data set of a Climp-63 overexpressing obese mouse liver was obtained and added to figure 4. I feel that also the other reviewer's comments were adequately addressed, therefore I recommend publication of this revised manuscript.

Wiebke Möbius

Referee #3:

As before, this study by Parlakgöl and Arruda et al. uses state-of-the art FIB-SEM methods to bring to life the architecture of organelles in their native tissue context. The authors observe striking changes to ER shape that vary based on the metabolic state of the organism (lean vs. obese). The authors have pushed FIB-SEM technology to the limit. This is a major technical achievement that required tackling technical and computational obstacles and will inspire future efforts using their tools on a wide array of biological questions and I can appreciate their goal to link organelle shape with organismal metabolism. Most of my concerns in my original review were addressed however two rather large ones remain.

-A major conclusion drawn is a connection between ER shape and metabolism. However, the way this is tested was to compare Climp63 overexpressing cells with controls. Although Climp63 does expand sheets, this is a tricky protein because its depletion has a minimal effect on ER shape. Thus, I requested that the authors change ER shape by some other method. There Rtn4 depletion was ineffective (clear from the western) so they chose instead to overexpress Rtn4 RHD to make more tubules and then aimed to score a similar phenotype. I know well that the Rtn4 RHD does not drive tubule formation without its cytoplasmic domain, it just localizes to tubules (see the rapoport papers referenced), only full length Rtn4a can drive tubule formation, this overexpression construct is a poor choice for this experiment and a cause for concern if it gives a reassuring result.

-I am not a mouse person, but n=1, n=2, n=1 mice over the 3 experimental conditions seems low to me. I appreciate the huge amount of technical work that went into this, but is this number of mice enough to make broad conclusions?

Referee #4:

The paper has been substantially expanded with new data and improved, however, perhaps the most important experiment that was requested was not performed. A Climp63 depletion would not require a KO mouse, as liver knockdowns can be easily achieved by injection of shRNA or ASO compounds for liver targeting. As over-expressions can lead to non-physiological effects, KO is the preferred method to test function.

Author Rebuttals to First Revision:**Referee #1 (Remarks to the Author):**

1- The authors have tried to address my previous comments but there are problems with the revisions. The manuscript now includes references to Supplementary Tables 1 and 2, but no Supplementary Table 1 is provided and the file that is linked as Supplementary Table 2 is just a spreadsheet.

We are sorry for this difficulty. The 'Supplementary Table 1', which corresponds to "dataset volume information", was embedded in the manuscript (main text, Word and PDF file) and the 'Supplementary Table 2' (segmentation evaluation) was provided as a separate Excel document. We now added all the supplementary tables as separate pages in the Word file and PDF as per Nature's submission guidelines. Again, we apologize for any inconvenience that it might have caused.

2- The spreadsheet includes accuracy, precision, recall and F1 scores but little information is provided on how these have been generated, either in this file or in the manuscript.

We apologize for the lack of clarity and limited information. Following the reviewer's feedback, we now provide additional information on how the precision, recall and F1 scores were generated. The full description of methods used for evaluation of the segmentation are added in the "**Methods**" section, sub-item "**Evaluation of FIB-SEM segmentation**".

The evaluation of organelle segmentation was performed at the same resolution as the segmentation. Raw data, manual annotation data and predictions are provided as tiffstacks (8bit or 32bit, compressed, packaged in a .zip folder). These data can be downloaded at EMPIAR database (please see **Q8** below for login credentials):

https://www.ebi.ac.uk/empair/deposition/id/47483731/open_metadata/

For lipid droplets, mitochondria, nuclear membrane, and ER, we selected 3 test volumes (150×150×25 voxels) at random locations. These volumes were not used for training. We manually segmented the contained organelles in these volumes and calculated object and voxel-based scores for all classes by comparing the manual segmentation with the automated segmentation in these test volumes. Object based scores were only calculated for mitochondria and lipid droplets. This is because each mitochondrion or lipid droplet is a biological unit and there are multiple of these objects, whereas ER is actually one single continuous membrane (single object) that spans the whole cell volume (**Supplementary Table 2**).

The evaluation of ER sheet and tubules segmentation was done pixel-wise. For that, we randomly choose an area of 750×750 pixel² at 4 different locations in each dataset that were used for training (**New Supplementary Table 3**).

For the voxel-based evaluation, TP (True Positive), FP (False Positive), FN (False Negative) and TN (True Negative) are evaluated on a per voxel level by comparing the manually annotated test samples with the automatic segmentation.

TP_{vx}: A voxel that is labeled in both the automatic and manual segmentation of the test volume.

FP_{vx}: A voxel that is labeled in the automatic segmentation but not in the manual segmentation of the test volume.

FN_{vx}: A voxel that is not labeled in the automatic segmentation but that is labelled in the manual segmentation of the test volume.

TN_{vx}: A voxel that is not labeled in either the automatic nor the manual segmentation of the test volume.

For the Object-based evaluation:

TPO_{object}: An object that is labeled in both the automatic and manual segmentation of the test volume with a shift tolerance of 5 px and minimal overlap of 5 %.

FPO_{object}: An object that is labeled in the automatic segmentation but not in the manual segmentation of the test volume with a shift tolerance of 5 px and minimal overlap of 5 %.

FNO_{object}: An object that is not labeled in the automatic segmentation but that is labelled in the manual segmentation of the test volume with a shift tolerance of 5 px and minimal overlap of 5 %.

Recall, precision, and F1-score were calculated as follows:

Precision: The fraction of relevant instances among the total retrieved instances.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Recall (sensitivity): The fraction of relevant instances that were retrieved.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

The **F1-score** is the harmonic mean of precision and recall.

$$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

For ER sheet and tubule segmentation, we also calculated accuracy, which is defined as a ratio of correctly predicted observation to the total observations.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN}).$$

3- What architecture was used for the deep learning models?

A customized convolutional neural network (CNN) architecture based on 3D U-Net was used for the deep learning-based automated segmentation methods. This was developed by ariadne.ai. This information is added in the “**Methods**” section, sub-item “**FIB-SEM segmentation**”.

4- Was a single classifier trained for each organelle using a mixture of lean and obese images? Or were separate classifiers trained for lean and obese?

For general organelle segmentation (mitochondria, lipid droplet, nuclear and cellular membranes and endoplasmic reticulum), each dataset was segmented using pre-trained models, which were developed by ariadne.ai based on segmentation of other similar datasets, and then individually refined and optimized. We would like to emphasize here that for general organelle segmentation, we only aimed to segment the 3D appearances of the structures but did not classify them into lean or obese nor healthy or diseased states. The accuracy of the resulting segmentation has been controlled and validated using human annotated data that was not used for model training. We think that in this case, training a single classifier that generalizes to both obese and lean data would suffer from class-imbalance issues given the intrinsic difference of these samples, which would be difficult to balance.

However, for the ER sheet and tubule sub-segmentation, where we classify the ER channel into two separate classes (e.g. sheets or tubules), we trained a U-Net architecture using a mixture of lean and obese images (single classifier). This analysis gave us the ER sheet/tubule ratio (one of the main foundations of this manuscript).

All the information described above is incorporated in **Methods** section.

5- The text box in the spreadsheet mentions that the evaluation was performed on downsampled images. Was the analysis in the manuscript also done using downsampled images?

The general organelle segmentation and evaluation were performed at the same resolution, 2x downsampled images.

Regarding the analysis in the manuscript, we mostly used the relative ratios of bulk voxel abundance of each class (e.g. total ER voxels divided by total cell voxels, etc). This was calculated at the original raw data resolution (not downsampled).

6- The data on ER sheets and tubules is presumably calculated pixel-wise but that is not stated. Again, whether a separate classifier was trained for lean and obese was not mentioned. This is important because if separate classifiers were trained that could incorporate unconscious bias by the investigators about how these differ between lean and obese.

Yes, the evaluation of ER sheets and tubules was done pixel-wise. We have stated this information in the “Methods” section, sub-item “Evaluation of FIB-SEM segmentation”. and in the new Supplemental Table 3. Regarding the training, we agree with the reviewer and appreciate this insightful comment. For this analysis, we have used a mixture of lean and obese ER examples as training data (single classifier) to re-segment the ER sheets and tubules.

7- The manuscript mentions Jaccard index evaluation in Supplementary Table 2 but that is not present in the spreadsheet.

We apologize for this mistake. We provided precision, recall and F-1 scores for evaluation per the reviewer’s request. We did not include Jaccard index and we removed this sentence from the manuscript.

8- The manuscript refers to an EMPIAR deposition containing the training and testing data but searching EMPIAR for 47483731 does not find anything.

We are sorry for this miscommunication. The full raw data, segmentation and evaluation data have been uploaded to EMPIAR on August 23rd, 2021, but it is set to be made public upon initial publication of the paper. The data can be accessed by the reviewer in the link below with the credentials provided. The data can be downloaded from the “Associate image sets with the data” page. It might take few minutes for the directory structure to be generated, but once done, the reviewer should be able to download the data via Aspera. For easier access we are providing below a Dropbox link where all the raw, segmentation and evaluation of the FIB-SEM imaging data can also be accessed.

EMPIAR login credentials:

username: review_9f801800

password: 32b360ae

data link: https://www.ebi.ac.uk/empiar/deposition/id/47483731/open_metadata/ (Deposition ID: 47483731, Accession code: EMPIAR-10791)

Dropbox link with all the data:

<https://www.dropbox.com/sh/c4u53y7v1laqcyc/AACeVQwgnjioCZejq5ip4BJAa?dl=0>

9- The source code branch on github has not been updated since the original submission so presumably any changes made for the revised manuscript are missing.

We have not made any changes in the source code. During the revision process, in terms of automated image analysis, we only acquired and analyzed the 1857_Obese_Climp63 dataset, which is essentially a dataset from an obese liver. Therefore, we used the same U-Net architecture that was previously pre-trained with mixed lean and obese samples for ER sheet and tubule segmentation of this dataset.

Referee #2 (Remarks to the Author):

In my view the authors have addressed all my questions, concerns and suggestions, invested significant effort and added valuable information to the manuscript. I am also delighted that a data set of a Climp63 overexpressing obese mouse liver was obtained and added to figure 4. I feel that also the other reviewer's comments were adequately addressed, therefore I recommend publication of this revised manuscript.

We thank the reviewer for all the insightful suggestions and for recommending our manuscript for publication.

Referee #3 (Remarks to the Author):

As before, this study by Parlakg  l and Arruda et al. uses state-of-the-art FIB-SEM methods to bring to life the architecture of organelles in their native tissue context. The authors observe striking changes to ER shape that vary based on the metabolic state of the organism (lean vs. obese). The authors have pushed FIB-SEM technology to the limit. This is a major technical achievement that required tackling technical and computational obstacles and will inspire future efforts using their tools on a wide array of biological questions and I can appreciate their goal to link organelle shape with organismal metabolism.

We thank the reviewer for this valuable feedback.

1- Most of my concerns in my original review were addressed however two rather large ones remain. -A major conclusion drawn is a connection between ER shape and metabolism. However, the way this is tested was to compare Climp63 overexpressing cells with controls. Although Climp63 does expand sheets, this is a tricky protein because its depletion has a minimal effect on ER shape. Thus, I requested that the authors change ER shape by some other method. Their Rtn4 depletion was ineffective (clear from the western) so they chose instead to overexpress Rtn4 RHD to make more tubules and then aimed to score a similar phenotype. I know well that the Rtn4 RHD does not drive tubule formation without its cytoplasmic domain, it just localizes to tubules (see the Rapoport papers referenced), only full length Rtn4a can drive tubule formation, this overexpression construct is a poor choice for this experiment and a cause for concern if it gives an assuring result.

We thank the reviewer for this feedback and insight. As the reviewer mentioned, this construct is described in the 2008 JBC paper (PMID: 18442980), where the authors investigated the tubular ER structure of yeast (*S. cerevisiae*) and monkey kidney (COS-7) cell line. In this paper, the authors show that *"the conserved reticulon homology domain (RHD) containing the two hydrophobic segments is sufficient for reticulon oligomerization; this domain alone causes slow lateral mobility in the membrane and forms homo-oligomers"* therefore showing that the GFP-RTN-HD is sufficient for tubule localization and stabilization. Similar findings were reported by Zurek et al, Traffic, 2011 (PMID: 20955502). In another study (PMID: 24558039), it was shown that in Rtn4 knockout MEF cells, the exogenous expression of the

RHD of the Rtn4 is sufficient to induce and restore ER tubulation and to rescue the ER morphology defects in RTN4-KO MEF cells compared to controls (Fig. 11B). Therefore, the effects of RTN4-HD exogenous expression on ER tubules may be context dependent.

In our case, we overexpressed the RTN4 homology domain (GFP-RTN4-HD) in mouse liver with an adenovirus which generates stable expression for longer periods of time (14days). This is a much longer duration compared with transient transfection in cell lines. In our ultrastructural analysis, we clearly identified that the expression of the GFP-RTN4-HD induced an enhanced network of smooth ER tubules, associated with lipid droplets and devoid of any other organelle (**Fig.1 of this letter**). To date, we have acquired hundreds of TEM images from healthy lean mice liver tissue in fed state and have never observed such an example. It's important to note that in our samples, the overexpression of GFP-RTN4-HD did not eliminate the ER sheets (As we have stated in the manuscript - line 237). These results suggest that the presence of exogenous GFP-RTN4-HD possibly stabilized the ER tubules rather than promote the sheet to tubule conversion.

We know (from published and our un-published data) that daily physiological cycles such as fasting and feeding regulates the relative abundance of ER sheet/tubule ratio in hepatocytes. Therefore, the accumulation of the ER tubules might not be directly caused by GFP-RTN4-HD expression but could be an accumulation of ER tubules over time. The increased number of tubules (which we observe in our TEMs) is our benchmark for this experiment to proceed and test the metabolic consequences of this phenomena. We cannot exclude, however, that the increased tubular network we observed in the hepatocytes result from a secondary effect of RHD expression. If it is appropriate, we can correct this language accordingly in the manuscript.

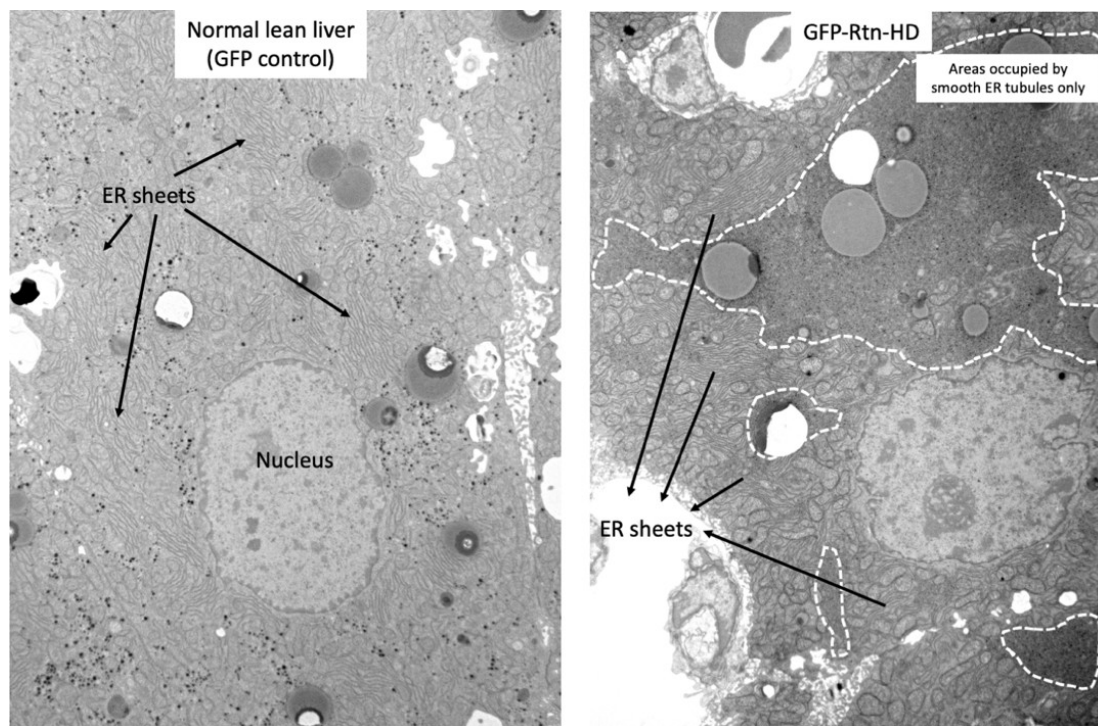


Fig.1. TEM images from lean mice expressing either control or GFP-RTN-HD constructs.

To further explore this point, we generated and tested two additional animal models to induce ER tubule formation in hepatocytes, an adenovirus expressing GFP-Reep5 (GFP-DP1) and adenovirus expressing the full length RTN4A as the reviewer requested. As it can be seen in **Fig.2A of this letter**, the overexpression of GFP-Reep5 also led to enrichment of ER tubules without eliminating parallel organized ER sheets. Metabolically, this approach led to increased lipid droplet accumulation (**Fig. 2B of this letter**), higher

triglyceride content (**Fig. 2C of this letter**) and higher SCD1 expression (**Fig. 2D of this letter**), the same phenotype seen with RHD expression. We have included these data in

Extended Data Figure 6 of the revised manuscript. We believe, these results increase confidence in our interpretation of the link between structure and function.

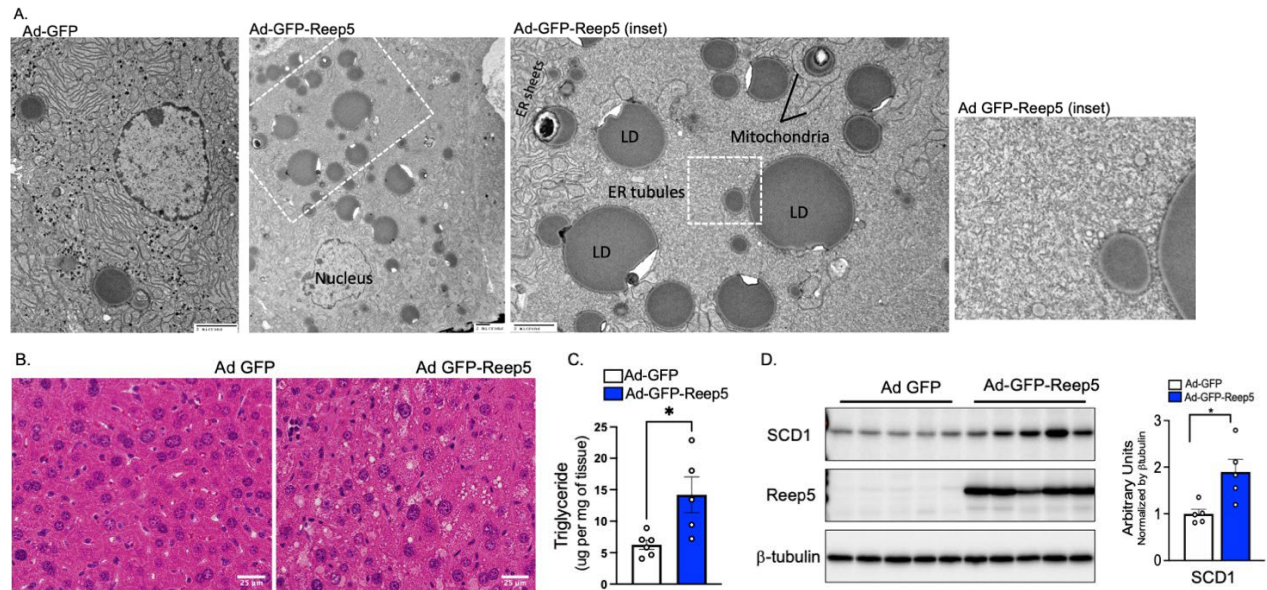


Fig.2. TEM images from lean mice expressing GFP-Reep5 construct (top panel) and metabolic evaluation of this condition (bottom panel).

Similar results were also obtained by the overexpression of the full length RTN4A in livers of lean mice (**Fig. 3 and Fig. 4 of this letter**). However, different from RHD and Reep5, overexpression of full length RTN4A also led to an increase in the number of autophagosomes in the cells, indicating the induction of autophagy in non-fasted state. In fact, in the literature it's been shown that different reticulon isoforms can act as ER-phagy receptors in addition to their ability to regulate tubular ER shape. Additionally, other effects of full length RTN4A (independent of its ER shaping function) have also been reported, such as effects on AKT and tyrosine kinase signaling (**PMID: 30078441, 27353365**). While upon RTN4A expression, we did observe the same metabolic outcomes; higher triglyceride content, increased lipid droplet accumulation and higher SCD1 expression, it is our view that focusing on the homology domain helped to eliminate the question of shape-independent effects of RTN4. Therefore, in our opinion, the results obtained with the RHD of the Rtn4 and further supported by another model through expression of Reep5 can be more directly attributed to the shaping of the ER tubules in liver tissue in vivo and the consequent metabolic alterations.

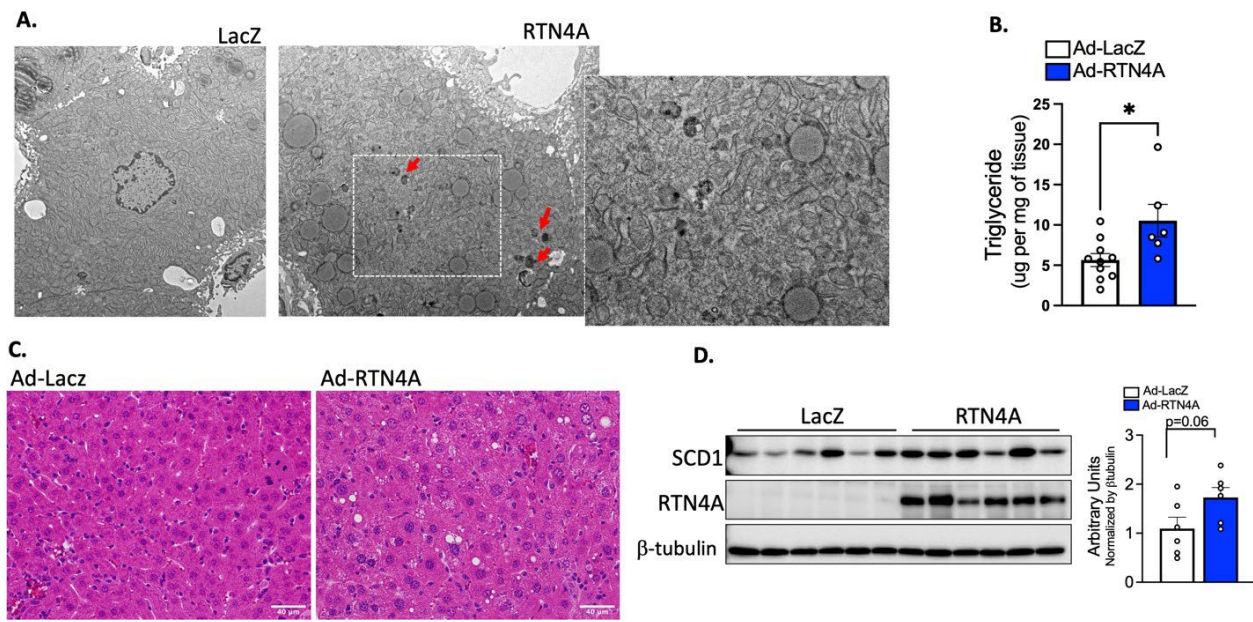


Fig.3. a, TEM image from lean mouse expressing RTN4A, exhibiting increased ER tubules, autophagosomes (red arrows) and ERphagy (top left). b-d, Metabolic effects of RTN4A expression in the livers of lean mice.

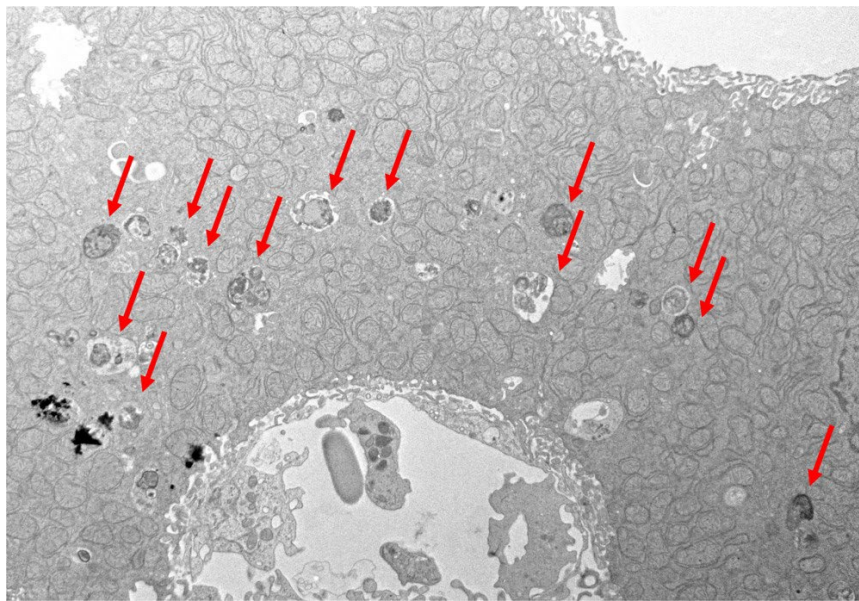


Fig.4. TEM image from lean mouse expressing RTN4A, exhibiting increased autophagosomes (red arrows) and ERphagy.

2- I am not a mouse person, but n=1, n=2, n=1 mice over the 3 experimental conditions seems low to me. I appreciate the huge amount of technical work that went into this, but is this number of mice enough to make broad conclusions?

We thank the reviewer for their insight. We performed FIB-SEM analysis of 1 or 2 mice per condition, but we analyzed 20 different whole cells as “biological units” providing, for the first

time, the ultrastructure details of many organelles of entire hepatocytes in their native tissue environment. Our raw data includes 92 partial or full hepatocyte volumes in total. As the reviewer acknowledged, this work involved an immense amount of human labor, computational work and microscopy hours. For example, it took roughly 2 months for the generation of each of the datasets and another 3-4 months for high precision automated segmentation, data processing and analysis, all done with the state-of-the-art imaging and computational tools, which are associated with high costs. We hope this process will become faster, less expensive, and more accessible in the near future to allow analysis of more "n"s, in terms of animal replicates. We are working towards this goal.

However, we would like to clarify that for each of our conclusions, we supported the FIB-SEM data with complementary methodologies utilizing at least 3 different mice per condition. For example, for the quantification of the parallel organization of ER sheets, we supported the FIB-SEM data performing 2D TEM, n=20 liver sections from 4 different mice per condition. We supported the finding of changes in smooth ER versus rough ER utilizing biochemical fractionation performed in n=6 mice per condition. We additionally, performed fluorescence analysis where we evaluate the distribution of endogenous SEC61b, an indication of ER sheets in 3-5 different mice per condition. We believe, we provide enough evidence to support our claim that obesity leads to striking remodeling of ER with the massive decrease of rough ER sheets and increased smooth ER tubules.

Additionally, we performed all our physiological in vivo tests, such as GTT, ITT in groups of mice ranging from n=7-10 mice per genetic condition, which were repeated in at least 3 separate occasions. All the number of replications and "n"s are stated in each figure legend or in the text as required.

Referee #4 (Remarks to the Author):

The paper has been substantially expanded with new data and improved, however, perhaps the most important experiment that was requested was not performed. A Climp63 depletion would not require a KO mouse, as liver knockdowns can be easily achieved by injection of shRNA or ASO compounds for liver targeting. As over-expressions can lead to non-physiological effects, KO is the preferred method to test function.

We thank the reviewer for the valuable feedback. The reviewer suggested using loss of function of Climp-63 in lean mouse as a strategy to decrease ER sheets and increase the relative abundance of ER tubules in healthy liver. However, as we mentioned in our point-by-point letter sent to the reviewers and also stated in the 3rd Reviewer's comment, it is well established in the literature that, surprisingly, knockdown or depletion of Climp-63 has a minimal effect on ER shape (PMID: 21111237). Therefore, in this specific case, loss of function approach cannot mirror overexpression. For this reason, we now used overexpression of multiple tubular ER shaping proteins (RHD, Reep5 and RTN4A) in order to provide additional in vivo evidence to support the proposed link between subcellular architecture and metabolic outcomes, we see in obese livers in lean condition. With the addition of alternate models that shape tubular ER and yield similar functional outcomes, we believe the evidence supports the link between structure and metabolic regulation. We hope that the addition of multiple additional animal models, all pointing to the same conclusion would be able to address the concerns of the reviewer.

Reviewer Reports on the Second Revision:

Referee #1:

The revised manuscript addresses all of the concerns I had raised in my previous reviews.

Robert F. Murphy, Professor of Computational Biology Emeritus, Carnegie Mellon University

Referee #3:

This response to my referee comment about the Rtn4 RHD expression construct is not at all correct:

"As the reviewer mentioned, this construct is described in the 2008 JBC paper (PMID: 18442980), where the authors investigated the tubular ER structure of yeast (*S. cerevisiae*) and monkey kidney (COS- 7) cell line. In this paper, the authors show that "the conserved reticulon homology domain (RHD) containing the two hydrophobic segments is sufficient for reticulon oligomerization; this domain alone causes slow lateral mobility in the membrane and forms homo- oligomers" therefore showing that the GFP-RTN-HD is sufficient for tubule localization and stabilization. Similar findings were reported by Zurek et al, Traffic, 2011 (PMID: 20955502).

While both these papers showed that the Rtn4 RHD is immobilized, NEITHER showed that the RHD alone can propagate ER tubules and so referencing them as such is a significant misrepresentation. Thus the 2008 JBC Reference (and 2011 paper for that matter) should not be used within the results section to support the claim/rationale that the Rtn4 RHD makes tubules.

So, my original concern with these RHD data remain because the Rtn4a RHD does not change the tubularity of the ER, and a single section EM tomograph does not show tubularity.

I like other aspects of the paper and appreciate the challenge of such a large scale EM study.

Referee #4:

The authors have responded with some further revisions, but, related to my registered concern the authors suggest that it is apparently not possible to do the loss of function experiment that is suggested. This is unfortunate, as gain of function maneuvers are notorious for producing non-physiological results, and yet only gain of function results are reported here on this issue. The authors should clearly state this caveat in their conclusions in additional text, and clearly caution in the abstract that the major hypothesis is not yet established by loss of function studies.

Author Rebuttals to Second Revision:

Referee #1 (Remarks to the Author):

The revised manuscript addresses all of the concerns I had raised in my previous reviews.

We thank the reviewer for all the insightful suggestions and for recommending our manuscript for publication.

Referee #3 (Remarks to the Author):

This response to my referee comment about the Rtn4 RHD expression construct is not at all correct. While both of the mentioned papers showed that the Rtn4 RHD is immobilized, NEITHER showed that the RHD alone can propagate ER tubules and so referencing them as such is a significant misrepresentation. Thus the 2008 JBC Reference (and 2011 paper for that matter) should not be used within the results section to support the claim/rationale that the Rtn4 RHD makes tubules. So, my original concern with these RHD data remain because the Rtn4a RHD does not change the tubularity of the ER, and a single section EM tomograph does not show tubularity.

I like other aspects of the paper and appreciate the challenge of such a large-scale EM study.

We apologize for the lack of clarity in our previous “Response to Reviewers” letter. As recommended by the reviewer, we are not using the above references to support the claim/rationale that the Rtn4-HD makes tubules. We removed the references from the manuscript. In the re-revised version of the manuscript, we simplified the interpretation of our data regarding the use of RTN4-HD and added these data to the Extended Fig 6. We also added to this figure the data obtained by the use of full length RTN4A. Additionally, we added the data obtained by the overexpression of Reep5 to Figure 3. We thank the reviewer for their insightful feedback.

Referee #4 (Remarks to the Author):

The authors have responded with some further revisions, but, related to my registered concern the authors suggest that it is apparently not possible to do the loss of function experiment that is suggested. This is unfortunate, as gain of function maneuvers are notorious for producing non-physiological results, and yet only gain of function results are reported here on this issue. The authors should clearly state this caveat in their conclusions in additional text, and clearly caution in the abstract that the major hypothesis is not yet established by loss of function studies.

We thank the reviewer for their point. In the re-revised manuscript, we added the following statement to our results section “However, we recognize that these conclusions are based on gain of function models and may require additional work to substantiate.”. We also added the sentence “As the opportunity emerges to overcome the redundancies and regulate ER structure genetically with loss of function models, further studies may provide additional insights.” to our conclusions and discussions.

Reviewers were consulted about the authors’ rebuttal and were satisfied with the responses at this stage.