THE EMBO JOURNAL

Widespread displacement of DNA- and RNA-binding factors underlies toxicity of arginine-rich cell-penetrating peptides

Vanesa Lafarga, Oleksandra Sirozh, Irene Díaz-López, Antonio Galarreta, Misaru Hisaoka, Eduardo Zarzuela, Jasminka Boskovic, Bogdan Jovanovic, Rafael Fernandez-Leiro, Jaime Muñoz, Georg Stoecklin, Iván Ventoso, and Oscar Fernandez-Capetillo
DOI: 10.15252/embj.2019103311

Corresponding author(s): Oscar Fernandez-Capetillo (ofernandez@cnio.es)

Review Timeline:

Submission Date:	26th Aug 19
Editorial Decision:	15th Oct 19
Revision Received:	2nd Mar 21
Editorial Decision:	9th Apr 21
Revision Received:	13th Apr 21
Accepted:	20th Apr 21

Editor: Stefanie Boehm

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

1st Editorial Decision 15th Oct 2019

15th Oct 2019

Re: EMBOJ-2019-103311
Displacement of DNA- and RNA-binding factors mediates toxicity of Arg-rich cell-penetrating peptides

Dear Dr. Fernandez-Capetillo,

Thank you for submitting your study on the basis for arginine-rich cell penetrating peptide cellular toxicity for consideration by The EMBO Journal. Please apologize the delay in getting back to you with this decision, which was also due to the slower referee responses during the summer vacation period. We have now received four referee reports on your study, which are included below for your information.

As you will see, the reviewers express an overall interest in the study and the proposed model, however also raise several major concerns that would need to be addressed in a revised version. In particular, the referees are not convinced that the physiological relevance of the findings has been sufficiently demonstrated. First, they find that key experiments would also need to be repeated in a relevant cell type. Secondly, the proposed oligonucleotide-based therapeutic approach should also be tested in a disease-relevant context. Third, the referees are concerned about the concentration of the peptides used in the experiments compared to the concentrations found in a human disease setting, and this aspect should at least be carefully discussed. In addition, referees #2, #3, and #4 note that the experimental controls have to be more explicitly described and additional controls included (ref# 2-2, ref# 3-1, ref#4- 2). Furthermore, all referees have specific concerns regarding the interpretation and presentation of data in Figure 3, as well as raising several more technical issues.

Should you be able to address these concerns, then we would be happy to consider the study further for publication. Therefore we would now like to invite you to prepare and submit a revised manuscript. Please note that EMBO Journal policy allows only a single round of major revision, therefore it is important to clarify all key concerns at this stage.

Please feel free to contact me should you have any further questions regarding the revision. I look forward to receiving your revised manuscript.

Kind regards,

Stefanie Boehm

Stefanie Boehm, PhD
Editor
The EMBO Journal

Referee #1:

Using synthetic arginine-rich dipeptide repeat proteins that have been linked to amyotrophic-lateral sclerosis and frontotemporal dementia, the authors show in this study that their presence leads to a generalized
displacement of RNA- and DNA-binding proteins from chromatin and mRNA.
The authors show that any reaction involving nucleic acids such as RNA transcription, translation, splicing and degradation or DNA replication and repair are impaired by the presence of these DPRs, similarly to what is occurring for other cell-penetrating peptides (CPPs), such as Protamine or the TAT peptide. The conclusions drawn are that widespread coating of nucleic acids and consequent displacement of RNA- and DNA-binding factors from chromatin and mRNA accounts for the toxicity of arginine-rich CPPs, including those that have been recently associated to the onset of ALS.

Overall, this is a well-executed study which provides enough evidence of a potential unifying mechanism to explain the widespread effect of arginine-rich peptides on nucleic acid homeostasis, and provides some initial proof-of-principle concepts as to how this knowledge can be exploited for therapeutic purposes. Considering the role that these R-rich dipeptides might have in certain neurodegenerative diseases linked to RAN-translation mechanism, I believe this study will have a medium impact in the field and amongst the audience of EMBO J. The authors came short of providing the evidence that the toxic phenotype of the ALS-linked arginine rich dipeptides, in ALS/FTD-relevant cells like motor or cortical neurons, can be rescued by their suggested strategy. In addition, the concentrations of the synthetic R-rich dipeptides used throughout the experiments here appears to be well above the concentration of these dipeptides in neurons in the human disease. If this cannot be ascertained, at least the authors should consider to discuss the concentrations used here in relations to the potential contribution of these arginine-rich peptides to pathogenic mechanisms in humans.

Beside the suggested rescue experiment, which I believe will strongly support their conclusions and will considerably elevate the impact of this study, there are some other concerns that I believe should be addressed prior to allow publication:

1- The authors need to better clarify how the qPCR data were normalized and how the procedure was done in general. Given the ability of PR to affect the assay the finer details matter here.
2- Figure 1E. There appears to be a major accumulation of Coilin in the cytoplasmic compartment. If that is a real phenotype it's not mentioned in the paper. Alternatively, if that is an artifact of uneven imaging intensities the authors may want to choose a better representative image.
3- In figure 3 panel D, PR is forming an obvious ring around the nucleus that disappears with the addition of ssDNA. This is not addressed in the body of the paper. The authors stated that- "Importantly, RNA or DNA oligonucleotides did not prevent the entry of the (PR)20 into the nucleus or its accumulation at nucleoli (Fig. 3D and Fig. S3B). On the contrary, the presence of the DPR led to the entry of ssDNA and, to a lesser extent, ssRNA, into cells, and their accumulation at nucleoli (Fig. 3E and Fig. S3C)". The image suggests that ssDNA greatly increased the entry of PR20 into the nucleus as the nuclear ring is absent in the bottom panel.
4- The pattern of PR localization in panels 3 D and E are very different. What's going on? Both panels have the same dosage of PR, same cell line, same time point. Yet no PR ring in the absence of ssDNA in panel E. Also in panel E, PR is significantly more granular.
5- Figure 1 Panel F, It looks like the Y axis is off. The average is below 100%
6- Figure 2 Panel E, Is the data for hour 10 significant? Perhaps, the authors should list the p value

in the figure legend.

Referee #2:

In the manuscript titled „Displacement of DNA- and RNA-binding factors mediates toxicity of Arg-rich cell-penetrating peptides", Lafarga et al. describe a series of experiments that show how the arginine-rich C9orf72-associated peptide (PR)20 and the cell-penetrating peptide (CPP) Protamine inhibit different reactions involving RNA and DNA. They use synthetic PR20 in molecular biological in vitro assays and cells treated with the same peptides. The presence of free nucleic acids can rescue the CPP induced impairment of replication/transcription/translation processes, as well as cell viability. Their main conclusion is a general displacement of DNA and RNA binding proteins after treatment and internalization of PR20. The findings are of general interest and can influence our view on different disease mechanisms. The manuscript is well written and data are well presented. Here are some points of criticism.

Major points:
- Biochemical characterization of PR20, GR20, and Protamine to show oligomerization and truncation products (e.g. by Western lot, non/semi-denaturing PAGE, SEC)
- Missing controls for unspecific peptide effects: In all experiments, treatment of cells and addition of protein to in vitro reactions has to be performed with control peptides of same concentration (e.g. scrambled peptides of positive, negative and neutral net charges ... or BSA, negatively and positively charged cellular proteins)
- Do the authors know the concentration of internalized PR20 peptides and can compare what was reported for C9orf72 ALS-cases?
- A Direct comparison of the affinity of RNA to PR20 vs Protamine vs Histone vs stress granule protein (affinity assay e.g. ITC, SPR, or others) would be helpful to estimate the cellular impact
- Most cell experiments were done by treating the cells with 7.5-10uM PR20; this concentration was shown to be cell toxic (Fig S2C) and the observed effects could thus be due to general apoptotic reactions. This has to be corrected by repeating the cell assays with non-toxic 5uM PR20 in the medium, or by showing that the observed effects are independent of cell death/apoptosis
- Figure 3B: ssDNA seems to increase viability of cells treated with PR20. Please show that the ssDNA actually binds PR20 peptides. Is PR20 pre-incubated with ssDNA and then added to the cells, or are they added simultaneously, or sequentially? This will be important for the suggested treatment of CPP effect by oligonucleotides or similar reagents.
- Arginine-rich protein domains have also been shown to be drivers of liquid phase separation in ALS-associated RNA-binding proteins, including C9orf72 DPRs, and other proteins. A potential connection to the observed effects on RNA and DNA processes should be discussed in the manuscript.

Minor points:
- It seems weird that the authors start the entire first chapter of their manuscript with data in the Supplemental Material, instead of starting with Figure 1; furthermore, the data in Figure S1 - on PR20-induced ribosome changes and binding - are very interesting but somewhat out of context, and they seem not directly relevant for the rest of the manuscript. They offer a nice example though on how PR20 could disable ribosome function and hence explain the general translation decrease by PR20.I suggest to move the data presented in Figure S1 to the end of the manuscript to show the impact on ribosomes as an example for what can go wrong in the presence of PR20 in the cell. The manuscript text could then start with Figure 1 and the general effects that PR20 has

on RNA reactions; in this case I would certainly suggest to show an image of PR20 treated cells in Figure 1 so that the reader sees the uptake and cellular distribution of PR20.
- Does PR20 also bind single nucleotides?
- Cell toxicity assays are missing for Protamine and GR20, and control peptide(s)
- RNA-splicing is impaired after treatment with Protamine; it would be good to show splicing data also for PR20 and GR20 and PK20. This is of interested because other ALS-related proteins like FUS and TDP-43 have been shown to alter splicing of certain mRNAs as well.
- Figure 1D: Does the RNA retained in the nucleus colocalize with PR20?
- Figure 2E: please show representative images of H2a.X in cells
- Figure S1A: Network model of ribosomal proteins does not contribute to the manuscript and should be deleted, especially since there is no comparison between PR20 treated and untreated cells.
- Figure S2B: what is the control condition the data from PR20 treated cells are normalized to?
- Figure S1D: It is unclear how the spectrum was obtained; what technique? SEC? Also, there is a reduction of 60S ribosomal units for PR20 treated cells, which is not mentioned. Please explain in the text and/or Figure legend.

Referee #3:

Summary:

In this manuscript, Lafarga et al. present an investigation into the mechanisms of cellular toxicity associated with arginine-rich cell-penetrating peptides (CPPs). Focusing on one such molecule, the (PR)n dipeptide repeat protein that has been previously linked to C9ORF72-ALS, the authors hypothesize that arginine-rich CPPs non-specifically bind (or "coat") nucleic acids within a cell and lead to general disruption of protein:nucleic acid interactions and resulting reactions such as transcription, translation, splicing, etc. Through a combination of various in vitro and cellular models, the authors show that the PR20 peptide binds to single- and double-stranded DNA/RNA molecules with similar affinities and is capable of interfering with reactions involving both RNA (such as reverse transcription, RNase-mediated RNA degradation, translation, viral RNA replication, etc) and DNA (such as PCR, DNA replication, DNA repair, etc) - many of which have been previously implicated in the pathogenesis of C9ORF72-ALS. The authors then proceed to show that some of these reactions that are impaired by the presence of PR20 peptides can be rescued by decoy or "scavenging" non-coding RNA/ssDNA oligonucleotides, including a reduction in cellular toxicity produced by PR20, without affecting its nucleolar localization in cells. In a final set of experiments, the authors then explored whether other CPPs would recapitulate many of the effects observed with the PR20 peptide, focusing here on the sperm-specific, arginine-rich polypeptide protamine. Similar to PR peptides, protamine was found to bind DNA and RNA with similar affinities and also interfered with both DNA and RNA-based reactions in vitro. Protamine was also demonstrated to mimic the nucleolar accumulation of PR20 peptides when administered in cells and interrupt nucleic-acid-based reactions such as transcription, translation and splicing. Subsequent proteomic analysis of both chromatin-bound and RNA-bound proteins in cells exposed to either PR20 or protamine uncovered a similar set of displaced factors between the two treatments, including ALS-linked proteins such as TDP-43 and FUS. In summation, the authors propose that displacement of DNA-/RNA-binding proteins in the presence of arginine-rich dipeptide repeat proteins may explain the general defects in nucleic acid processing and metabolism that are observed in C9ORF72-ALS. While the results presented in this manuscript are interesting and valuable to the field, there are numerous questions:

Major concerns/questions:

One major concern is the general lack of proper controls within experiments throughout the paper. While many of the experiments are compared to a "control" condition, it is not clearly explained what said control condition consists of (i.e. mock treatment, untreated, etc). Furthermore, more proper controls may consist of treatment with non-arginine-rich peptides -- the easiest examples being other ALS-linked DPRs like (PA)n, (GA)n, etc, but PK peptides used in Figure S2 could suffice.
Another question relates to binding specificity of (PR)n peptides. It was not clear what sequences of ss/dsDNA and ss/dsRNA were being used in various experiments and why they were chosen. The authors should determine whether PR peptides bind nucleic acids of different sequences with similar affinity (suggesting non-specific, "generalized" binding) or whether they may have higher affinities for certain sequences/structures as this may have important implications for especially vulnerable transcripts/genomic regions in C9ORF72-ALS. Furthermore, CLIP-based techniques could be utilized to determine whether nucleic acid "coating" truly occurs or whether there may be some consensus sequence to which PR peptides bind preferentially.
The authors mention instability of (GR)n peptides as a limitation preventing the investigation of GR-repeat peptides in tandem with PR-repeat peptides. Given the previously reported toxicity of GR peptides, as well as the arginine content suggested to play a central role toxicity of CPPs like PR20, it would be very valuable to the field to investigate whether GR-repeat proteins recapitulate many of the effects observed with PR20 and protamine. Mammalian expression vectors encoding these proteins are commonly used in the field and could be utilized for intracellular experiments.
Given the previously reported effects of dipeptide repeat protein length on various pathogenic processes and toxicity (for example: Wen et al. (2014), Mizielinska et al. (2014), Callister et al. (2016)), it would be worth investigating whether (PR)n dipeptides of longer (and more pathophysiologically-relevant) lengths produce more severe effects on the various DNA- and RNA-based reactions observed with PR20 peptides. Furthermore, it would be interesting to determine whether nucleic acid binding affinity (be it more non-specific or sequence/structure-specific) would similarly change as a function of DPR length.
In Figure 3, the authors show a rescue of cellular toxicity associated with PR peptides through treatment with ssDNA oligonucleotides in U2OS cells. Considering the degeneration occurring in ALS seems to be predominantly neuronal, the authors should consider testing whether similar short oligonucleotide-based therapies could be used to "scavenge" DPRs in neuronal cell lines (most ideally C9ORF72-ALS patient iPSC-derived motor neurons).

Minor concerns/questions:

In general, more clear explanations of methods, controls and experiments would be useful to readers. It is often unclear what experiments exactly consist of (i.e. concentrations of peptide treatments, control treatments, etc) which may hinder reproducibility in the future.
The authors show that ssDNA oligonucleotides are seemingly localized to nucleoli by binding to PR peptides. Does this occur with endogenous transcripts in cells exposed to PR treatment? Could this be another mechanism of toxicity?
In Figure 3A, the authors use noncoding oligonucleotides to prevent DPRs from interfering with in vitro translation reactions. Would this same effect be achieved by any negatively-charged molecule (through nonspecific electrostatic interactions)? Polyanion molecules such as heparin, dextran, etc could be examples of such controls.

Referee #4:

Arginine rich cell penetrating peptides are known to be highly toxic to cells. Lafarga et al seek to determine the mechanism by which these CPPs, including the C9 ALS associated DPR, Poly(PR) are toxic. Using artificial PR20 peptides, the authors show that arginine CPPs may exert toxicity by binding to RNAs and DNAs and displacing RBPs. While the study is of potential interest to a broad readership, there is little novelty presented and only minimal mechanistic advances in our understanding of toxic DPRs presented. Additionally, there are concerns over the quality and interpretation of some of the data presented due to the lack of proper controls and the biological relevance of the findings is lacking. Therefore, substantial revisions should be made prior to consideration for publication.

Major Concerns:
1. The authors seemingly use a different cell type/cell line for almost every experiment presented (at least in their initial figures). Due to biological differences amongst cell types, in order to properly evaluate the data presented, a singular cell line should be used for all core experiments to substantiate claims. In addition, where appropriate, a technical explanation should be provided as to why there is an inconsistency amongst cell lines used for each experiment.
2. Throughout the manuscript (especially figures 1 and 2), there is no explanation of what the "control" is. This must be provided. It is unclear if appropriate controls are being used without proper labeling or explanation. For example, are the controls an HA tag or simple cells without PR? Additionally, the authors use a PK20 control for a single experiment. A more relevant control (for C9 disease biology) would be PA20 and should be used for all core experiments to substantiate claims related to arginine rich CPP toxicity and potential mechanisms.
3. There is little novelty involved in this study. Previous publications have demonstrated that Poly(PR) interacts with ribosomal proteins, colocalizes with heterochromatin, and its toxicity can be mitigated by RNA (Lee et al 2016, Zhang et al 2019, Boeynaems et al 2017). The authors begin to provide mechanistic/biological relevance to these phenomenon; however, the data presented are quite preliminary. Like many other studies, the authors rely a completely artificial overexpression system. They should demonstrate that their findings can be recapitulated in real patient cells with endogenous levels of PR and GR. Furthermore, the authors propose that PR binds to 5'UTRs and using a proteomics approach that in the presence of PR, binding of disease relevant RNA BPs (TDP-43, FUS) to RNAs is reduced. TDP-43 is known to bind to 3' UTRs of target RNAs and therefore, the authors should discuss how PR binding to 5' UTRs could displace TDP-43 binding to 3' UTRs. In addition, to advance the biological relevance of their findings, the authors should address whether PR binding to the 5' UTRs of RNAs is 1. Direct and 2. Occurs for disease relevant RNAs that are known to be pathologically dysregulated in C9 ALS/FTD.

Minor Concerns:
1. For all images where it is not already, panels showing PR distribution need to be shown to correspond PR localization to the phenotypes presented.
2. There is a lot of inconsistency within data panels shown. For example, in figure 3, why does the distribution of PR look very different between panel D and E? In figure 1, the DAPI signal is drastically different between panels C and E.
3. The authors use in vitro translation of an exogenous luciferase RNA reporter to provide evidence that translation is reduced when PR is expressed in cells. However, a much more realistic and meaningful readout would be to puromycin incorporation instead of a purely OE exogenous system.

Referee #1:

Using synthetic arginine-rich dipeptide repeat proteins that have been linked to amyotrophic-lateral sclerosis and frontotemporal dementia, the authors show in this study that their presence leads to a generalized displacement of RNA- and DNA-binding proteins from chromatin and mRNA.

The authors show that any reaction involving nucleic acids such as RNA transcription, translation, splicing and degradation or DNA replication and repair are impaired by the presence of these DPRs, similarly to what is occurring for other cell-penetrating peptides (CPPs), such as Protamine or the TAT peptide. The conclusions drawn are that widespread coating of nucleic acids and consequent displacement of RNA- and DNA-binding factors from chromatin and mRNA accounts for the toxicity of arginine-rich CPPs, including those that have been recently associated to the onset of ALS.

Overall, this is a well-executed study which provides enough evidence of a potential unifying mechanism to explain the widespread effect of arginine-rich peptides on nucleic acid homeostasis, and provides some initial proof-of-principle concepts as to how this knowledge can be exploited for therapeutic purposes. Considering the role that these R-rich dipeptides might have in certain neurodegenerative diseases linked to RAN-translation mechanism, I believe this study will have a medium impact in the field and amongst the audience of EMBO J. The authors came short of providing the evidence that the toxic phenotype of the ALS-linked arginine rich dipeptides, in ALS/FTD-relevant cells like motor or cortical neurons, can be rescued by their suggested strategy. In addition, the concentrations of the synthetic R-rich dipeptides used throughout the experiments here appears to be well above the concentration of these dipeptides in neurons in the human disease. If this cannot be ascertained, at least the authors should consider to discuss the concentrations used here in relations to the potential contribution of these arginine-rich peptides to pathogenic mechanisms in humans.

We thank the reviewer for his overall appreciation of our work and for stating that "overall, this is a well-executed study" and that we "provide enough evidence of a potential unifying mechanism to explain the widespread effects of arginine-rich peptides on nucleic acid homeostasis". That, and only that, was the purpose of our work.

One important aspect that shall clarify with the reviewer is that, while we have used (PR) dipeptide repeats as a model of arginine-rich CPPs, our work is not focused on ALS. Instead, our aim was to provide the scientific community with a model that explains the toxicity, in general, of arginine-rich cell-penetrating polypeptides. While this phenomenon is somewhat new to the ALS-community, the toxicity of polyarginines has been known for decades and is present in multiple biological contexts.

Our model postulates that the affinity of oligoarginines for nucleic acids is so high that it leads to a general displacement of RNA- or DNA-binding proteins from chromatin and mRNA. As far as we can see, this simple model has never been proposed before

and provides a straight-forward explanation as to why so many reactions involving nucleic acids are affected by arginines-rich CPPs.

As an example that the implications of our model are not restricted to ALS, we also show that the effects of poly(PR) peptides in displacing proteins from RNA and chromatin are equally observed when cells are treated with the sperm-specific protein protamine. Noteworthy, the physiological function of protamine is actually to displace histones from sperm chromatin due to its high affinity for DNA provided by its high density of arginines. As a final anecdote in this regard, after submitting our work I discovered works from 30 years ago reporting that pigs in their intestine produce a toxic Arginine-rich small peptide named PR-39 that behaves an antibiotic! Interestingly, in these early papers it was already noted that pig PR-39 peptides interfere with reactions that involve nucleic acids such as translation or replication (Boman et al Infect Immun 1993). Thus, even though the ALS community is now excited about the relevance of arginine-rich peptides in C9ORF72 pathology, this is only one area their toxicity is relevant, and our work provides a general model to explain this phenomenon.

All of the above being said, following the reviewers´ suggestion we have now repeated our rescue experiments in two new cell types, mouse motoneurons derived from either primary mESC or from NSC34 cells. We now show that, in both cell types, non-coding oligonucleotides rescue the toxicity of arginine-rich peptides. Moreover, and following a suggestion from reviewer #3, we now also show that the toxicity of arginine-rich CPPs in motoneurons is also alleviated by the polyanion Heparin, therefore providing 2 examples of rescue in more relevant cell-types. This second experiment is particularly interesting as protamine is clinically used as an antidote of Heparin. Even though these rescue experiments work in vitro, I want to be crystal clear in that by no means we are proposing a new therapy for ALS or for any other disease, and we are making that clear in our discussion. I am a basic scientist, well aware of the distance between our findings and the clinic. We simply conducted these rescue experiments to further illustrate that the toxicity of oligo-arginines is related to their generalized nucleic acid binding properties. I hope you share that, collectively, all the complementary approaches provided in our manuscript are sufficient to make that case.

Finally, in what regards to the concentration of the peptides used in our study, out of fairness I think is actually unclear how much polyPR peptide, and for how long, is present in cells from ALS patients. While DPR peptides have been observed in patients´ cells, nobody has done a proper biochemical quantification of the amount of peptide, which in addition is likely to change from patient to patient and during lifetime. In any case, I cannot foresee how the different dosage should affect the fundamental biochemical properties of these peptides regarding their affinity for nucleic acids. Any model for a given pathology has, inevitably, limitations. I have dedicated most of my scientific career to the field of DNA repair, where how cells repair DNA breaks has been solved by exposing them to extremely high doses of radiation or to very high concentrations of genotoxic chemicals. The severity of these insults was obviously supraphysiological, but how cells respond to one or 100 breaks does not substantially

differ at the molecular level, and these studies have been instrumental to help us understand DNA damage responses.

Beside the suggested rescue experiment, which I believe will strongly support their conclusions and will considerably elevate the impact of this study, there are some other concerns that I believe should be addressed prior to allow publication:

1- The authors need to better clarify how the qPCR data were normalized and how the procedure was done in general. Given the ability of PR to affect the assay the finer details matter here.

This is now clearly explained in the Materials and Methods section of the current version.

2- Figure 1E. There appears to be a major accumulation of Coilin in the cytoplasmic compartment. If that is a real phenotype it's not mentioned in the paper. Alternatively, if that is an artifact of uneven imaging intensities the authors may want to choose a better representative image.

While in all of our experiments arginine-rich peptides clearly reduced the number of Cajal Bodies, not in all cases the treatment led to the appearance of cytoplasmic Coilin. We now provide better representative images to illustrate what is most often seen. To further address this point, we have now performed additional immunofluorescence experiments with an independent marker of Cajal bodies (SMN). Consistent with our original claim, the presence of $(PR)_{20}$ peptides significantly decreases the number of Cajal bodies with either marker.

3- In figure 3 panel D, PR is forming an obvious ring around the nucleus that disappears with the addition of ssDNA. This is not addressed in the body of the paper. The authors stated that- "Importantly, RNA or DNA oligonucleotides did not prevent the entry of the (PR)20 into the nucleus or its accumulation at nucleoli (Fig. 3D and Fig. S3B). On the contrary, the presence of the DPR led to the entry of ssDNA and, to a lesser extent, ssRNA, into cells, and their accumulation at nucleoli (Fig. 3E and Fig. S3C)". The image suggests that ssDNA greatly increased the entry of PR20 into the nucleus as the nuclear ring is absent in the bottom panel.

The reviewer is right in that this could be an interesting observation, but truth is that this ring pattern is actually very rare; thanks for spotting it. Consistent with the intracellular distribution of (PR) DPRs presented in other reports, these peptides accumulate mostly at nucleoli (we believe this simply reflects the abundance of rRNA) although cytoplasmic staining can also be detected (normally masked by the very high intensity of the nucleolar signal).

We have now included a better representative figure to illustrate the most frequent distribution of (PR)20 peptides. In any case, and regardless of what a few selected images might represent, we want to note that these data are in fact quantified from hundreds of cells by High-Content Microscopy and the statements we make in our

manuscript are based on these quantifications. These quantifications were actually provided in our supplementary panels (Figures S3A,B in the current version).

The point of this experiment was just to illustrate that the presence of oligonucleotides does not prevent the entry of (PR)n into cells, so that the rescue cannot simply be explained by an inability of the peptides to enter into cells. To our surprise, not only this did not happen, but actually (PR)20 peptides served as a "transfection reagent" and led to internalization of the oligonucleotide! Following this path, I later discovered that, in fact, argine-rich peptides such as protamine or the arginine-rich fraction of histone extracts were the first transfection reagents ever reported in the history of biomedical research (Ryser and Hanckock 1965; Smull and Ludwig, 1962). Thus, once again, while the interest in arginine cell-penetrating peptides has recently raised due to its potential links to ALS, their use as carriers of DNA or proteins into cells has been studied for decades.

Finally, to further strengthen this dataset, we now also show that (PR)n peptides also enable the cellular uptake of ssRNA (Figure 6E).

4- The pattern of PR localization in panels 3 D and E are very different. What's going on? Both panels have the same dosage of PR, same cell line, same time point. Yet no PR ring in the absence of ssDNA in panel E. Also in panel E, PR is significantly more granular.

The reason that PR patterns looks different in 3D and 3E is because each pannel was done using a different immunofluorescence protocol. In 3E, in order to simultaneously detect UBF, we used a protocol that allows for a better detection of chromatin-bound factors and that is based on the prior extraction of cytoplasmic and nucleoplasmic proteins. In contrast, the images in 3E arise from a standard fixation protocol that preserves cytoplasmic proteins. This is why the images in 3D have almost no cytoplasmic signal.

For the sake of the discussion, and as to what is the true distribution of polyPR in cells, this peptide simply accumulates where nucleic acids are. While in many papers polyPR is shown as nucleolar, we believe this simply reflects the very high density of RNA at nucleoli, which takes most of the signal. However, at higher exposure times a specific cytoplasmic signal is invariably detected and has in fact been reported in previous publications. The reality is that polyPR peptides are present throughout the cell, including in the cytoplasm.

5- Figure 1 Panel F, It looks like the Y axis is off. The average is below 100%

Thanks for spotting it. This was a mistake and the graph has been corrected.

6- Figure 2 Panel E, Is the data for hour 10 significant? Perhaps, the authors should list the p value in the figure legend.

It is significant, and the p-value is now provided.

Referee #2:

In the manuscript titled „Displacement of DNA- and RNA-binding factors mediates toxicity of Arg-rich cell-penetrating peptides", Lafarga et al. describe a series of experiments that show how the arginine-rich C9orf72-associated peptide (PR)20 and the cell-penetrating peptide (CPP) Protamine inhibit different reactions involving RNA and DNA. They use synthetic PR20 in molecular biological in vitro assays and cells treated with the same peptides. The presence of free nucleic acids can rescue the CPP induced impairment of replication/transcription/translation processes, as well as cell viability. Their main conclusion is a general displacement of DNA and RNA binding proteins after treatment and internalization of PR20. The findings are of general interest and can influence our view on different disease mechanisms. The manuscript is well written and data are well presented. Here are some points of criticism.

We thank the reviewer for his/her very nice words on our work.

Major points:

- Biochemical characterization of PR20, GR20, and Protamine to show oligomerization and truncation products (e.g. by Western lot, non/semi-denaturing PAGE, SEC)

While we thank the reviewer for his/her comment, we feel that the biochemical characterization of oligo-arginine oligomerization is somewhat outside of the scope of our manuscript. In addition, this has been thoroughly addressed in prior reports (eg. Lin et al Cell 2016; Boeynaemns et al Mol Cell 2017) where it was shown that these peptides tend to aggregate and/or phase separate (something we also see in our own EMSA assays).

- Missing controls for unspecific peptide effects: In all experiments, treatment of cells and addition of protein to in vitro reactions has to be performed with control peptides of same concentration (e.g. scrambled peptides of positive, negative and neutral net charges ... or BSA, negatively and positively charged cellular proteins)

In the revised version of the manuscript, we have included numerous controls both in vitro and in cells (including BSA, $(PK)_{20}$, $(GR)_{20}$ and $(GA)_{20}$). Importantly, all of the relevant effects that we report (toxicity, transfection capability etc...) are only observed with arginine-rich peptides and not with any of the controls.

- Do the authors know the concentration of internalized PR20 peptides and can compare what was reported for C9orf72 ALS-cases?

Out of fairness I think is actually unclear how much arginine-rich peptides, and for how long, are present in cells from ALS patients. While DPR peptides have been observed in patients´ cells, nobody has done a proper biochemical quantification of the amount of peptide, which in addition is likely to change from patient to patient and during lifetime. In any case, I cannot foresee how the different dosage should affect the fundamental biochemical properties of these peptides regarding their affinity for nucleic acids. Any model for a given pathology has, inevitably, limitations. I have

dedicated most of my scientific career to the field of DNA repair, where how cells repair DNA breaks has been solved by exposing them to extremely high doses of radiation or to very high concentrations of genotoxic chemicals. The severity of these insults was obviously supraphysiological, but how cells respond to one or 100 breaks does not substantially differ at the molecular level, and these studies have been instrumental to help us understand DNA damage responses.

On a final note, I shall make it clear to the reviewer that while we used (PR)20 peptides as a model, our work is not just focused on ALS. While this is new to the ALS community, the toxicity of arginine-rich CPPs is something that has been known for decades and is relevant in various contexts. As a final anecdote in this regard, after submitting our work I discovered works from 30 years ago reporting that pigs in their intestine produce a toxic Arginine-rich small peptide named PR-39 that behaves an antibiotic! Interestingly, in these early papers it was already noted that pig PR-39 peptides interfere with reactions that involve nucleic acids such as translation or replication (Boman et al Infect Immun 1993). Thus, even though the ALS community is now excited about the relevance of arginine-rich peptides in C9ORF72 pathology, this is only one area their toxicity is relevant, and our work provides a general model to explain this phenomenon.

- A Direct comparison of the affinity of RNA to PR20 vs Protamine vs Histone vs stress granule protein (affinity assay e.g. ITC, SPR, or others) would be helpful to estimate the cellular impact

Again, while we appreciate the suggestion from the reviewer, we believe the molecular characterization of arginine-rich peptides has been addressed before in other papers and lies somewhat aside from the scope of our MS, which is more focused on the consequences of PR peptide exposure at the cell level. We have nevertheless performed EMSA assays with (PR)20 and protamine and, in vitro, both have similar dynamics.

In this regard, we want to note that the actual biological function of protamine is to displace histones from DNA during spermatogenesis, which provides biological proof of the extremely high affinity of oligoarginines for nucleic acids and strong support to our model.

- Most cell experiments were done by treating the cells with 7.5-10uM PR20; this concentration was shown to be cell toxic (Fig S2C) and the observed effects could thus be due to general apoptotic reactions. This has to be corrected by repeating the cell assays with non-toxic 5uM PR20 in the medium, or by showing that the observed effects are independent of cell death/apoptosis

Many of our experiments are done in vitro where this does not apply (eg. effects on translation, transcription or RNA degradation). As for the *in cellulo* experiments, these done at times at which we see no obvious cell-death of the inhibitor. For instance, all the proteomics was done after 90´and the effects on DNA replication (4h) and repair (1h) were also evaluated at short times. In our experience, we only see substantial toxicity after more than 2 days of exposure, and no experiment was done at that time.

In the new version of the manuscript, we have also included experiments in motoneurons, where we used lower doses of the peptides and observed limited toxicity. All the conclusions hold regardless of the cell line, dose or time of the experiments used.

- Figure 3B: ssDNA seems to increase viability of cells treated with PR20. Please show that the ssDNA actually binds PR20 peptides. Is PR20 pre-incubated with ssDNA and then added to the cells, or are they added simultaneously, or sequentially? This will be important for the suggested treatment of CPP effect by oligonucleotides or similar reagents.

PR peptides indeed directly bind to ssDNA (and also ssRNA, dsDNA and dsRNA) as was already shown in our EMSA assays. In our in vitro assays, nucleic acids are added simultaneously to the peptides. In vivo, however, the peptides are added to subsequently (to cells already containing nucleic acids).

- Arginine-rich protein domains have also been shown to be drivers of liquid phase separation in ALS-associated RNA-binding proteins, including C9orf72 DPRs, and other proteins. A potential connection to the observed effects on RNA and DNA processes should be discussed in the manuscript.

We thank the reviewer for the suggestion and now include a discussion about the possible impact of phase separation in our studies.

Minor points:

- It seems weird that the authors start the entire first chapter of their manuscript with data in the Supplemental Material, instead of starting with Figure 1; furthermore, the data in Figure S1 - on PR20-induced ribosome changes and binding - are very interesting but somewhat out of context, and they seem not directly relevant for the rest of the manuscript. They offer a nice example though on how PR20 could disable ribosome function and hence explain the general translation decrease by PR20.I suggest to move the data presented in Figure S1 to the end of the manuscript to show the impact on ribosomes as an example for what can go wrong in the presence of PR20 in the cell. The manuscript text could then start with Figure 1 and the general effects that PR20 has on RNA reactions; in this case I would certainly suggest to show an image of PR20 treated cells in Figure 1 so that the reader sees the uptake and cellular distribution of PR20.

There is a historical reason as to why the manuscript starts with the work on ribosomes, this being that -truthfully- our initial work trying to understand how (PR) dipeptide repeats affected translation was what led us to hypothesize that the effects could be due to a generalized coating of nucleic acids. The manuscript simply reflects the true chronological order in which this scientific work and its ideas evolved.

- Does PR20 also bind single nucleotides?

Yes it does, and we thank the reviewer for suggesting this possibility since it can indeed have important functional implications which are now discussed in the manuscript (Figure 3G,H). This finding has never been reported before in the literature, is relevant, and is now presented in the main figures of the current version of the manuscript. Thanks again for the suggestion.

- Cell toxicity assays are missing for Protamine and GR20, and control peptide(s)

We have now performed an exhaustive quantification of toxicity including these and other peptides, including in more relevant cell types such as motoneurons. In addition, we have performed those assays in the presence of heparin or ssDNA, to further illustrate the rescue effect.

- RNA-splicing is impaired after treatment with Protamine; it would be good to show splicing data also for PR20 and GR20 and PK20. This is of interested because other ALS-related proteins like FUS and TDP-43 have been shown to alter splicing of certain mRNAs as well.

We have now analyzed the impact of splicing of the suggested peptides, and also of GA. As shown in the new Figure 4E, only arginine rich PR20 and GR20 trigger abnormal splicing, while GA20 or PK20 do not.

Just for the sake of the discussion with the reviewer, I find it interesting that FUS and TDP43 are proteins with high affinity for RNA, and we also wonder as to what extent their impact on splicing could be related to a generalized impact on RNA binding dynamics rather than an actual role on splicing reactions.

- Figure 1D: Does the RNA retained in the nucleus colocalize with PR20?

rRNA does, but mRNA is excluded from nucleoli in all conditions, even in the presence of PR20.

- Figure 2E: please show representative images of H2a.X in cells

Done, images are now included.

- Figure S1A: Network model of ribosomal proteins does not contribute to the manuscript and should be deleted, especially since there is no comparison between PR20 treated and untreated cells.

This panel was simply shown to illustrate the fact that our ribosome purification protocol worked fine. We still believe this could be informative to illustrate the enrichment obtained but have left it out as requested.

- Figure S2B: what is the control condition the data from PR20 treated cells are normalized to?

This is now clearly explained in the Materials and Methods section of the current version.

- Figure S1D: It is unclear how the spectrum was obtained; what technique? SEC? Also, there is a reduction of 60S ribosomal units for PR20 treated cells, which is not mentioned. Please explain in the text and/or Figure legend.

All this information is clearly explained in the Materials and Methods section on ribosome profiling.

Referee #3:

In this manuscript, Lafarga et al. present an investigation into the mechanisms of cellular toxicity associated with arginine-rich cell-penetrating peptides (CPPs). Focusing on one such molecule, the (PR)n dipeptide repeat protein that has been previously linked to C9ORF72-ALS, the authors hypothesize that arginine-rich CPPs non-specifically bind (or "coat") nucleic acids within a cell and lead to general disruption of protein:nucleic acid interactions and resulting reactions such as transcription, translation, splicing, etc. Through a combination of various in vitro and cellular models, the authors show that the PR20 peptide binds to single- and double-stranded DNA/RNA molecules with similar affinities and is capable of interfering with reactions involving both RNA (such as reverse transcription, RNase-mediated RNA degradation, translation, viral RNA replication, etc) and DNA (such as PCR, DNA replication, DNA repair, etc) - many of which have been previously implicated in the pathogenesis of C9ORF72-ALS. The authors then proceed to show that some of these reactions that are impaired by the presence of PR20 peptides can be rescued by decoy or "scavenging" non-coding RNA/ssDNA oligonucleotides, including a reduction in cellular toxicity produced by PR20, without affecting its nucleolar localization in cells. In a final set of experiments, the authors then explored whether other CPPs would recapitulate many of the effects observed with the PR20 peptide, focusing here on the sperm-specific, arginine-rich polypeptide protamine. Similar to PR peptides, protamine was found to bind DNA and RNA with similar affinities and also interfered with both DNA and RNA-based reactions in vitro. Protamine was also demonstrated to mimic the nucleolar accumulation of PR20 peptides when administered in cells and interrupt nucleic-acid-based reactions such as transcription, translation and splicing. Subsequent proteomic analysis of both chromatin-bound and RNA-bound proteins in cells exposed to either PR20 or protamine uncovered a similar set of displaced factors between the two treatments, including ALS-linked proteins such as TDP-43 and FUS. In summation, the authors propose that displacement of DNA-/RNA-binding proteins in the presence of arginine-rich dipeptide repeat proteins may explain the general defects in nucleic acid processing and metabolism that are observed in C9ORF72-ALS. While the results presented in this manuscript are interesting and valuable to the field, there are numerous questions:

We thank the reviewer for his/her assessment and for finding that our results are interesting and valuable to the field.

Before moving to the specific comments, I would want to first clarify with this reviewer that while we used PR20 peptides as a model of a CPP, our work is not focused on ALS. While this is somewhat new to the ALS community, the toxicity of oligo-arginines is something that has been known for decades and is relevant in various contexts. For instance, arg-rich histone fractions and peptides were the first transfection factor ever discovered in biomedical research (Ryser and Hanckock 1965; Smull and Ludwig, 1962), although their toxicities were noted early on and have limited their clinical development. In addition, while protamine is clinically used as an antidote of heparin, its toxicity is driving the search for alternative heparin antidotes (Sokolowska et al., 2016). As a final anecdote in this regard, after submitting our work I discovered works from 30 years ago reporting that pigs in their intestine produce a toxic Arginine-rich

small peptide named PR-39 that behaves an antibiotic! Interestingly, in these early papers it was already noted that pig PR-39 peptides interfere with reactions that involve nucleic acids such as translation or replication (Boman et al Infect Immun 1993). Thus, even though the ALS community is now excited about the relevance of arginine-rich peptides in C9ORF72 pathology, this is only one area their toxicity is relevant, and our work provides a general model to explain this phenomenon.

Major concerns/questions:

One major concern is the general lack of proper controls within experiments throughout the paper. While many of the experiments are compared to a "control" condition, it is not clearly explained what said control condition consists of (i.e. mock treatment, untreated, etc). Furthermore, more proper controls may consist of treatment with non-arginine-rich peptides -- the easiest examples being other ALS-linked DPRs like (PA)n, (GA)n, etc, but PK peptides used in Figure S2 could suffice.

In the current version we have added numerous controls and additional peptides (including (GA)n, (PK)n, BSA and (GR)n); both in in vitro and in vivo experiments. The results with all these controls support that the effects we describe are exclusively observed with arginine-rich peptides.

Of note, the results with (PK)20 peptides were particularly interesting. While these peptides also enter into cells, they do not accumulate at nucleoli nor have any obvious effect in facilitating the transfer of nucleic acids into cells. This is consistent with all the body of literature showing that arginines have much more affinity to nucleic acids than lysines.

Another question relates to binding specificity of (PR)n peptides. It was not clear what sequences of ss/dsDNA and ss/dsRNA were being used in various experiments and why they were chosen. The authors should determine whether PR peptides bind nucleic acids of different sequences with similar affinity (suggesting non-specific, "generalized" binding) or whether they may have higher affinities for certain sequences/structures as this may have important implications for especially vulnerable transcripts/genomic regions in C9ORF72-ALS. Furthermore, CLIP-based techniques could be utilized to determine whether nucleic acid "coating" truly occurs or whether there may be some consensus sequence to which PR peptides bind preferentially.

The fact that arginine-rich peptides bind to nucleic acids irrespectively of its sequence has been reported before (e.g. Kanekura et al Hum Mol Genet 2016), and we also have failed to see any distinct effect of the sequence on the different oligonucleotides (ssDNA/ssRNA etc) that we have used in our study. Furthermore, I believe that our proteomic study clearly shows that the presence of arginine-rich CPPs leads to a generalized displacement of all RBPs and chromatin-binding proteins, arguing against any sequence-specific bias.

Note that the model that we are proposing here is very much in line with the actual biological function of protamine, an endogenous arginine-rich peptide, which

collectively displaces all histones from DNA during spermatogenesis with no specific sequence bias.

The authors mention instability of (GR)n peptides as a limitation preventing the investigation of GR-repeat peptides in tandem with PR-repeat peptides. Given the previously reported toxicity of GR peptides, as well as the arginine content suggested to play a central role toxicity of CPPs like PR20, it would be very valuable to the field to investigate whether GR-repeat proteins recapitulate many of the effects observed with PR20 and protamine. Mammalian expression vectors encoding these proteins are commonly used in the field and could be utilized for intracellular experiments.

GR peptides were already noted to be unstable on the early work of Kwon et al in Science (2014).

Nevertheless, and as mentioned above, we have now been able to figure out conditions to work with GR20, and in fact the current version of the manuscript also contains data for GR20, GA20 and PK20 peptides. The summary is that the effects reported are exclusively observed with arginine-containing peptides.

Given the previously reported effects of dipeptide repeat protein length on various pathogenic processes and toxicity (for example: Wen et al. (2014), Mizielinska et al. (2014), Callister et al. (2016)), it would be worth investigating whether (PR)n dipeptides of longer (and more pathophysiologically-relevant) lengths produce more severe effects on the various DNA- and RNA-based reactions observed with PR20 peptides. Furthermore, it would be interesting to determine whether nucleic acid binding affinity (be it more non-specific or sequence/structure-specific) would similarly change as a function of DPR length.

As noted by the reviewer, there is already substantial literature indicating that repeat length influences the effects of PR DPRs. Consistent with this, in our hands (PR)n dipeptides shorter than 20 units (e.g. $(PR)_5$ or $(PR)_{10}$) do not exert cellular toxicity in our experiments, in contrast to $(PR)_{20}$.

Of note, DPR length is not the only variable as others such as concentration or duration of the exposure to the peptide will ultimately determine the biological effects of arginine rich DPRs.

In Figure 3, the authors show a rescue of cellular toxicity associated with PR peptides through treatment with ssDNA oligonucleotides in U2OS cells. Considering the degeneration occurring in ALS seems to be predominantly neuronal, the authors should consider testing whether similar short oligonucleotide-based therapies could be used to "scavenge" DPRs in neuronal cell lines (most ideally C9ORF72-ALS patient iPSC-derived motor neurons).

Before addressing this point, and as I mentioned above, I should once again state that while we used (PR)20 peptides as a model of a toxic CPP, our work is not focused on ALS. While this is somewhat new to the ALS community, the toxicity of arginine-rich

peptides is something that has been known for decades and is relevant in various contexts. We hope our work provides a model to address them in general.

Nevertheless, and following the reviewers request, we have now conducted rescue experiments in two new cell types, mouse motoneurons derived from either primary mESC or from NSC34 cells. We now show that, in both cell types, non-coding oligonucleotides rescue the toxicity of (PR)20 peptides. Moreover, and following a nice suggestion from this reviewer in his/her next comment, we now also show that the toxicity of the poly(PR) peptides in motoneurons is also alleviated by the polyanion Heparin, therefore providing 2 examples of rescue in more relevant cell-types.

Even though these rescue experiments work nicely in vitro, I want to be crystal clear in that by no means we are proposing a new therapy for ALS or for any other disease. I am a basic scientist, well aware of the distance between our findings and the clinic. We simply conducted these experiments to further illustrate that the toxicity of oligo-arginines is related to their generalized nucleic acid binding properties, and we make this clear in our discussion. I hope the reviewer shares that, collectively, all the complementary approaches provided in our manuscript are sufficient to at least consider this model worthy of consideration.

Minor concerns/questions:

In general, more clear explanations of methods, controls and experiments would be useful to readers. It is often unclear what experiments exactly consist of (i.e. concentrations of peptide treatments, control treatments, etc) which may hinder reproducibility in the future.

We have now extended on these and other details both in the Figure Legends and the Materials and Methods section of the current version.

The authors show that ssDNA oligonucleotides are seemingly localized to nucleoli by binding to PR peptides. Does this occur with endogenous transcripts in cells exposed to PR treatment? Could this be another mechanism of toxicity?

The presence of PR leads to an accumulation of nascent RNA at nucleoli, which would support this possibility. In addition, we now also show that PR peptides also drive the accumulation of ssRNA oligonucleotides at nucleoli. However, we believe that the nascent RNA that accumulates at nucleoli is rRNA as polyA+ mRNA is actually excluded from nucleoli with or without PR peptides (see Figure 1D).

In Figure 3A, the authors use noncoding oligonucleotides to prevent DPRs from interfering with in vitro translation reactions. Would this same effect be achieved by any negatively-charged molecule (through nonspecific electrostatic interactions)? Polyanion molecules such as heparin, dextran, etc could be examples of such controls.

Thanks for this suggestion. We now show that in fact even the toxicity of arginine rich CPPs is rescued by the presence of heparin. However, I don´t think that this applies

to "any negatively-charged molecule" and there was a basis to focus on heparin. Heparin is used as a clinical antidote for protamine, which is an extremely arginine-rich small peptide. It is nice to see that the "antidote" effect that heparine has on protamine can be extended to other arginine-rich CPPs such as PR or GR dipeptides.

For the sake of the discussion with this reviewer, I shall note that in contrast to what we see with oligonucleotides, the presence of heparin does partially reduce the amount of PR peptides that enter into cells. This might be related to the well-established fact that the internalization of arginine-rich CPPs occurs through the binding to heparanated receptors on the cell membrane (see, for instance (Fuchs et al Biochemistry 2004) or (Borrelli et al Molecules 2018). Thus, this strategy would only be relevant to reduce the reported cell-to-cell transmission of PR peptides, which I think is fair to say is yet unclear as to how much it contributes to the pathology.

Referee #4:

Arginine rich cell penetrating peptides are known to be highly toxic to cells. Lafarga et al seek to determine the mechanism by which these CPPs, including the C9 ALS associated DPR, Poly(PR) are toxic. Using artificial PR20 peptides, the authors show that arginine CPPs may exert toxicity by binding to RNAs and DNAs and displacing RBPs. While the study is of potential interest to a broad readership, there is little novelty presented and only minimal mechanistic advances in our understanding of toxic DPRs presented. Additionally, there are concerns over the quality and interpretation of some of the data presented due to the lack of proper controls and the biological relevance of the findings is lacking. Therefore, substantial revisions should be made prior to consideration for publication.

This comment certainly gives away what to expect from the rest of the review. We are particularly surprised by the categoric assessment made by the reviewer in that our study lacks novelty and ask him or her to tell us where precisely we could find another report that proposes that the mechanism of toxicity of arginine-rich peptides is due to a generalized effect of these peptides in displacing proteins from nucleic acids (RNA or DNA). Nevertheless, this reviewer also dislikes the quality of virtually all our experiments and argues that our interpretations are also incorrect.

Having faced similar reviews in the past, I am aware that there is little we could do would convince this reviewer that our work is meritorious to be shared with the scientific community. We below nevertheless provide answers to all the comments made.

Major Concerns:

1. The authors seemingly use a different cell type/cell line for almost every experiment presented (at least in their initial figures). Due to biological differences amongst cell types, in order to properly evaluate the data presented, a singular cell line should be used for all core experiments to substantiate claims. In addition, where appropriate, a technical explanation should be provided as to why there is an inconsistency amongst cell lines used for each experiment.

Arginine-rich peptides are toxic for all cells. In our study, we mostly used U2OS cells as these are a widely used model in cell biology, and in fact were the ones used by Kwon et al (Science 2014) in the original study reporting the effects of polyPR peptides in cellular toxicity. As requested by this and other reviewers, we now also provide new figures using mouse motoneurons derived from primary mESC or NSC34 cells, and all of the results that we reported hold in these cell types.

I think is important to clarify with this reviewer that while we used polyPR peptides as a model of a CPP, our work is actually not focused on ALS. While this is new to the ALS community, the toxicity of oligo-arginines is something that has been known for decades and is relevant in various contexts. For instance, arg-rich histone fractions and peptides were the first transfection factor ever discovered in biomedical research (Ryser and Hanckock 1965; Smull and Ludwig, 1962), although their toxicities were

noted early on and have limited their clinical development. In addition, while PROTAMINE is clinically used as an antidote of heparin, its toxicity is driving the search for alternative heparin antidotes (Sokolowska et al., 2016). As a final anecdote in this regard, after submitting our work I discovered works from 30 years ago reporting that pigs in their intestine produce a toxic Arginine-rich small peptide named PR-39 that behaves an antibiotic! Interestingly, in these early papers it was already noted that pig PR-39 peptides interfere with reactions that involve nucleic acids such as translation or replication (Boman et al Infect Immun 1993). Thus, even though the ALS community is now excited about the relevance of arginine-rich peptides in C9ORF72 pathology, this is only one area their toxicity is relevant, and our work provides a general model to explain this phenomenon.

2. Throughout the manuscript (especially figures 1 and 2), there is no explanation of what the "control" is. This must be provided. It is unclear if appropriate controls are being used without proper labeling or explanation. For example, are the controls an HA tag or simple cells without PR? Additionally, the authors use a PK20 control for a single experiment. A more relevant control (for C9 disease biology) would be PA20 and should be used for all core experiments to substantiate claims related to arginine rich CPP toxicity and potential mechanisms.

The controls defined in the previous version always referred to cells not treated with the indicated peptides. In addition, in the current version of the paper we have added numerous control peptides to our in vitro and in vivo experiments, including GR20, PROTAMINE, GA20 and PK20, all of which support that it is the presence of arginines what determines the biological effects of these DPRs on nucleic acid metabolism.

3. There is little novelty involved in this study. Previous publications have demonstrated that Poly(PR) interacts with ribosomal proteins, colocalizes with heterochromatin, and its toxicity can be mitigated by RNA (Lee et al 2016, Zhang et al 2019, Boeynaems et al 2017). The authors begin to provide mechanistic/biological relevance to these phenomenon; however, the data presented are quite preliminary. Like many other studies, the authors rely a completely artificial overexpression system. They should demonstrate that their findings can be recapitulated in real patient cells with endogenous levels of PR and GR. Furthermore, the authors propose that PR binds to 5'UTRs and using a proteomics approach that in the presence of PR, binding of disease relevant RNA BPs (TDP-43, FUS) to RNAs is reduced. TDP-43 is known to bind to 3' UTRs of target RNAs and therefore, the authors should discuss how PR binding to 5' UTRs could displace TDP-43 binding to 3' UTRs. In addition, to advance the biological relevance of their findings, the authors should address whether PR binding to the 5' UTRs of RNAs is 1. Direct and 2. Occurs for disease relevant RNAs that are known to be pathologically dysregulated in C9 ALS/FTD.

The reviewer here makes a wrong interpretation of our own claims. At no place did we indicate that we believe that PR20 binding occurs primarily at 5´UTRs. To the contrary, we actually show that PR20 peptides or protamine, both lead to a generalized and unbiassed displacement of proteins from RNA and DNA. Bear in mind that the biological function of protamine, a naturally occurring arginine-rich CPP, is to displace histones from DNA during spermatogenesis, which occurs indiscriminately throughout

the genome. Our work postulates that the effects reported for arginine-rich peptides would be similar to what would happen if protamine were to be accidentally expressed in somatic cells.

As for the novelty statement, and as mentioned above, we ask the reviewer to direct us to any other report where this model of generalized nucleic-acid coating and the consequent generalized displacement of DNA- and RNA-binding factors has been proposed to explain why arginine-rich peptides are toxic and have so promiscuous effects in impairing biological reactions using nucleic acids.

Minor Concerns:

1. For all images where it is not already, panels showing PR distribution need to be shown to correspond PR localization to the phenotypes presented.

We have now added new images of the PR distribution (see the response to #2 below) at all relevant experiments.

2. There is a lot of inconsistency within data panels shown. For example, in figure 3, why does the distribution of PR look very different between panel D and E? In figure 1, the DAPI signal is drastically different between panels C and E.

The reason that PR patterns looks different in 3D and 3E is because each pannel was done using a different immunofluorescence protocol. In 3E, in order to simultaneously detect UBF, we used a protocol that allows for a better detection of chromatin-bound factors and that is based on the prior extraction of cytoplasmic and nucleoplasmic proteins. In contrast, the images in 3E arise from a standard fixation protocol that preserves cytoplasmic proteins. This is why the images in 3D have almost no cytoplasmic signal.

As to what is the true distribution of polyPR in cells, this peptide simply accumulates where RNA is. While in many papers polyPR is shown as nucleolar, we believe that this simply reflects the very high density of RNA at nucleoli, which takes most of the signal. However, at higher exposure times a specific cytoplasmic signal is invariably detected and has in fact been reported in previous publications. The reality is that polyPR peptides are present throughout the cell, including in the cytoplasm. These peptides accumulate wherever nucleic acids are.

Finally, I shall mention that images in the manuscript are simply provided as means of illustration as cell-to-cell variability is intrinsic to any microscopy datasets. I myself take the relevance of individually provided images in any manuscript with caution and understand them as means of illustration. No claims in our work are made based on the images and this is why in our laboratory any imaging-data relies on the use of unbiassed High-Content Mediated analyses of imaging data, which provides a rigorous quantification pipeline and minimizes the impact of outliers.

3. The authors use in vitro translation of an exogenous luciferase RNA reporter to provide evidence that translation is reduced when PR is expressed in cells. However, a much more realistic and meaningful readout would be to puromycin incorporation instead of a purely OE exogenous system.

We used the in vitro system to make the point that PR20 peptides can directly interfere with the translation reaction. The effect of PR20 peptides in lowering translation in cells by measuring the incorporation of puromycin derivatives has been reported numerous times, but as requested we now also show this effect in our paper (e.g. Figures 4F and 6B).

9th Apr 2021

Re: EMBOJ-2019-103311R
Displacement of DNA- and RNA-binding factors mediates toxicity of Arg-rich cell-penetrating peptides

Dear Dr. Fernandez-Capetillo,

Thank you for submitting your revised manuscript. Please also excuse the delay in communicating this decision to you, which was due to delayed referee reports on account of the current pandemic and delays during the pre-acceptance checks. We have now however received comments from two of the initial referees, which are included below for your information. As these referees now also support publication, I would now like to ask you to address a number of editorial issues that are listed in detail below. Please make any changes to the manuscript text in the attached document only using the "track changes" option. Once these remaining issues are resolved, we will be happy to formally accept the manuscript for publication.

Referee #1:

In this revised manuscript the authors addressed all the concerns I had about the first submission. The authors provided clarifications and added new experiments in support of their conclusion. I believe this study is as good as it gets. It highlights some general mechanisms of toxicity by the R-rich peptides with a focus on dipeptides toxicity that might be relevant for the pathogenesis of C9orf72-linked ALS. The novelty is not there as these mechanisms have long been theorized and extrapolated from other observations. However, the quality and rigor of the data are high and the experiments are now well-controlled and executed.

Referee #2:

In the revised version of the manuscript, Lafarga et al. address most concerns experimentally and/or verbally through text changes.
I have no major concerns anymore for publishing the work in EMBO J.

Thank you again for submitting the final revised version of your manuscript. I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Oscar Fernandez-Capetillo
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2019-103311

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

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

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

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

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

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	N/A
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	N/A
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No data were excluded.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	Randomization did not apply as samples were independent.
For animal studies, include a statement about randomization even if no randomization was used.	N/A
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	All microscopy datasets were analyzed automatically through High-Through Microscopy for rigorous quantifications. In addition, key experiments were reproduced independently by independent authors.
4.b. For animal studies, include a statement about blinding even if no blinding was done	N/A
5. For every figure, are statistical tests justified as appropriate?	They are indcated and defined where appropiate.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes
Is there an estimate of variation within each group of data?	Yes, every graph provides SEM values.

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

Is the variance similar between the groups that are being statistically compared?	Yes it is.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Catalog numbers are provided for all antibodies, from which specificity and references can be obtained.: HA RocheCat#11867423001 UBF1Santa CruzCat#sc-13125 CoilinLamond´s Laboratory204.11 gH2AXMilliporeCat#05-636 TUBB3 Biolegend Cat#801202 SMN Santa Cruz Cat#sc-32313 Anti-Mouse IgG-488Life TechnologiesCat#A11029 Anti-Mouse IgG-555Life TechnologiesCat#A21422 Anti-Rabbit IgG-488Life TechnologiesCat#A11008 Anti-Rabbit IgG-555Life TechnologiesCat#A31572
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Done. U2OS, BHK-21, RPE and NSC34 cells were acquired from ATCC as the parental HeLa from which stable transfectants were made. Primary mES cell lines were generated by the transgenic unit of CNIO. All cell lines are regularly tested for mycoplasma contamination.

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

D- Animal Models

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

E- Human Subjects

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

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Done. A "Data availability" section has been added at the end of our Materials & Methods which reads: "Correspondence and request of materials should be addressed to O.F. Mass spectrometry proteomics datasets associated to this work are available in the following databases: •Impact of oligoarginine peptides in ribosome composition and chromatin proteomes: PRIDE PXD010555 (http://www.ebi.ac.uk/pride/archive/projects/PXD010555) •Impact of oligoarginine peptides in RNA-binding proteomes: PRIDE PXD014085 (http://www.ebi.ac.uk/pride/archive/projects/PXD014085)".
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	N/A
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	N/A
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	N/A

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No