molecular
systems
biology

Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium

Vera van Noort, Jan Seebacher, Samuel Bader, Shabaz Mohammed, Ivana Vonkova, Matthew J Betts, Sebastian Kühner, Runjun Kumar, Tobias Maier, Martina O'Flaherty, Vladimir Rybin, Arne Schmeisky, Eva Yus, Jörg Stülke, Luis Serrano, Robert B Russell, Albert JR Heck, Peer Bork and Anne-Claude Gavin

Corresponding authors: Peer Bork and Anne-Claude Gavin, European Molecular Biology Laboratory

---

Transaction Report:

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

1st Editorial Decision   08 August 2011

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees whom we asked to evaluate your manuscript. As you will see from the reports below, the referees clearly find the topic of your study of potential interest. However, they raise substantial concerns on the work, which should be convincingly addressed in a major revision of the present work.

Beside the comments related to the need to clarify the text and explain the MS methodology in more details, the main concern expressed both by reviewer #1 and #3 refers to the acetylation dataset. Thus, it would be very informative to analyze the acetylation and phosphorylation profiles in an acetyl-transferase mutant strain. This would provide deeper insights into the cross-talk between the two PTM systems and it would consolidate the acetylation dataset, which is particularly important in view of the rather serious issues regarding the reproducibility of the reported acetylation profiles.

Please reformat the manuscript to the MSB style and include a detailed Materials & Methods section. We would also kindly ask you to deposit your MS primary datasets in one of the major public databases (see http://www.nature.com/msb/authors) and include a link/hashcode/accession in Materials & Methods.

*** PLEASE NOTE *** As part of the EMBO Publications transparent editorial process initiative (see our Editorial at http://www.nature.com/msb/journal/v6/n1/full/msb201072.html), Molecular

---

Systems Biology will publish online a Review Process File to accompany accepted manuscripts. When preparing your letter of response, please be aware that in the event of acceptance, your cover letter/point-by-point document will be included as part of this File, which will be available to the scientific community. More information about this initiative is available in our Instructions to Authors. If you have any questions about this initiative, please contact the editorial office msb@embo.org.

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favourable.

Yours sincerely,

Editor
Molecular Systems Biology

http://www.nature.com/msb

Referee reports

Reviewer #1 (Remarks to the Author):

Review of "Phosphorylation is intertwined with different layers of post-transcriptional regulation in a genome-reduced bacterium".

This study uses MS/MS to provide a global perspective of the connections between phosphorylation and protein abundance/acetylation patterns in Mycoplasma pneumoniae. These authors published another study on the M. pneumoniae phosphoproteome using the same mutants several months ago (Schmidl et al, 2010 Mol Cell Proteomics). However, the protein abundance and lysine acetylation data provides a wealth of new results to warrant publication, pending the following clarifications (especially with respect to reproducibility of the lysine acetylation data).

Manuscript format:
The manuscript should be reformatted to meet MSB guidelines. It lacks a Methods and Materials section. Also, references are presented by authors in the text but are listed numerically in the order they appear in this reference section, making it cumbersome to find the references.

Figures:
Fig 1C: The table portion of this figure reports the same data as supplementary table 4. The plot is a cartoon of how peptide abundance changes were normalized for MS intensity. As the method for quantifying peptide abundance ratios is not discussed in the text and simply references another paper (Mortensen et al, 2010) in the supp info, this cartoon should be removed.

Fig 1D: This figure should be removed as it is not referenced in the text and appears to be a simplified cartoon of Fig 5.

Fig 2A. What does "% coverage" on the ordinate axis mean? Mean percent coverage of the proteins with peptide identified by mass spec? Fraction of the proteins in each COG identified by mass spec?

Fig 4. This figure shows strong global decreases in lysine acetylation in the pknB mutant and corresponding increases in the hprK mutant. However, this direct connection between kinase inactivation and lysine acetylation is not discussed in the text. Why?

Supp Info:
Mass spec data is reported in the text and in Fig 1B to include both biological and technical replicates. However, data in the supplementary information for abundance (data set 1) phosphorylation (data set 2) and lysine acetylation (data set 3) report only mean changes without

standard deviations or data from the replicates. Please report data from the replicates so the reader can assess reproducibility.

Questions about phosphorylation results:
Schmidl et al, 2010 identified 63 phosphosites by comparing wild-type to the same mutants strains, compared to 93 sites in this study. However, only 30 sites are common to both studies. What caused these differences? Do the authors believe that the actual phosphoproteome is the union or intersection of these datasets?

I was surprised to see that 14% of kinase-regulated proteins showed higher phosphorylation in the kinase mutants. How do the authors explain this 'inverted directionality'? Does deletion of the normal kinase cause the other kinase to overcompensate?

Phosphorylated peptides were separated by strong cation exchange (SCX). Other studies (a few refs below) have improved upon this method by including a second phosphopeptide enrichment by titanium dioxide (TiO2) or immobilized metal affinity chromatography(IMAC). Why wasn't a second phosphopeptide enrichment used in this study?

Olsen et al, 2010 "Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis" Science Signalling.

Villen and Gygi 2008 "The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry" Nature Protocols.

Questions about acetylation results:
This study reports 719 lysine acetylation sites, showing this PTM affects >30% of proteins and is 8 times more frequent than phosphorylation. Indeed, the remarkably high level of acetylation and its connection to phosphorylation is one of the main conclusions of the study. However, this is remarkably high level of lysine acetylation requires some clarification.

1.Supplementary Fig 2D shows that the reproducibility of the acetylation data between technical replicates was only r2=0.13. Further, Supplementary Fig 2E plots biological replicates for protein abundance and phosphorylation, but omits acetylation. The manuscript text should mention the low reproducibility between technical replicates. Fig 1B suggests that biological replicates were run for lysine acetylation. Please show the biological replicates of lysine acetylation data in Supplementary fig 2E.

2.Protein abundance and phosphopeptide data were filtered to an FDR of 1% using a target-decoy approach. In contrast, the FDR for the acetylation data was estimated from C-terminally acetylated lysines. Does this mean that different FDR methods were used for the datasets? If so, how many acetylation sites are found when the target-decoy approach used for protein abundance and phosphorylation is applied to the acetylation data?

3.The phosphorylation data was supported using kinase/phosphatase mutants. Given the high acetylation levels, why is there no discussion in the text about how this may be occurring other than a single mention of the N-acetyltransferase Mpn114? Moreover, why wasn't a similar approach using a mutant of Mpn114 used to support the acetylation data?

Reviewer #2 (Remarks to the Author):

In this work "Phosphorylation is intertwined with different layers of post-transcriptional regulation in a genome-reduced bacterium" Noort et al. have attempted to understand the interplay between protein phosphorylation and lysine acetylation. It is an important question in the field of protein PTMs. They have utilized Mycoplasma pneumoniae, a prokaryote with one of the most reduced genome and which has only two kinases and one phosphatase. The advantage of this system is that it enables them to do combinatorial perturbations of the kinases and phosphatase and observe the effects on total phosphorylation and lysine acetylation at an organismal level. However since it's an endocellular parasite (using the host systems genetic machinery) it is not clear how these perturbations translate into biological phenotype of the organism. This study points out to some interesting observations and it has got good data set but it fails to convince on the broader claims

made by the authors. The overall aim and goal of the study seems to warrant a publication in MSB but lot of assertions by the authors need to be substantiated with real evidence and the manuscript has to be re-written.

Points of note:

1. Their title and abstract seem to be slightly misleading there is not much evidence and discussion about post-transcriptional regulation. The definition of their post-transcriptional regulation seems not to be clear (or did the authors mean post-translational?).

2. The authors have not clearly explained and discussed how they account for the differences in the proteome abundance without any changes in the transcriptional levels. Without clearly indicating and investigating what those mechanisms might be leading to this discordant observation between protein abundance changes and no-change in transcription level their argument seems to be more of a hypothesis than a premise of a paper. Their interpretation that the phosphorylational perturbations in chaperone proteins may lead to protein abundance changes lacks validity and has to be proved.

3. They don't talk about Tyrosine phosphorylation, though Fig. 2b indicates there is some Tyrosine phosphorylation observed.

4. On page 3 the authors say that they combine genetics and high resolution quantitative mass spectrometry to measure the global effect of serial kinase and phosphatase deletions on proteome abundance, phosphorylation and lysine acetylation. The authors could have just mentioned about the 3 mutants and discussed their MS strategy in detail giving more credence to their work. More importantly they should explain how quantitative was their MS approach and what are the advantages of the chosen labeling techniques. In short, the strength and the weakness of the paper lies in the quantitative MS approach, to make their arguments appealing about the quantitation and the conclusion they should discuss the MS approach in more detail and convince the readers that they were done with utmost care. The word serial kinase and phosphatase deletions are also misleading, there were just 3 mutants.

5. In their discussion the authors have not discussed the biological variation between the 3 differently labelled wild type samples. It is difficult to assess what they report in their results were whether normal biological variation or over and above biological variation.

6. "All fractions were analyzed by high resolution Mass spectrometry" on page 4 suggests next to nothing. Since the paper is heavily reliant on the MS quantitation the authors should explain why and how they did this.

7. "Out of 564 proteins 460 were quantified" again is not a helpful statement to understand why the remaining proteins were discarded.

8. In fig 1c the authors have observed that there were 93 phospho-sites on 92 peptides. Does this imply that identified a single doubly phosphorylated peptide?

9. Fig 2a is not very informative, categorizing proteins into metabolic, cell signaling and information storage and processes pathways doesn't establish that there were no biases in the MS identifications. There would be so many over laps in each category.

10. The assertion that lysine acetylation is almost 8 times as frequent as phosphorylation could be biased. The authors claim that the median sequence coverage was only 43% indicating that half of the peptides were not identified by MS. This finding may also be explained by the differences in the experimental protocols used to purify phospho- and acetyl-containing peptides. The former were purified by ion exchange chromatography, while the latter were purified using an antibody. A more accurate comparison of differential abundance of these two modifications should involve a shared reference for normalizing differences in enrichment affinity, such as a phospho- and acetyl-modified peptide standard.

11. The statement "Of the 45 relative abundances in the three knockouts 39 were in agreement with the MS data" is not clear.

12. On page 7 the authors claim that "these results show that perturbations of phosphorylation network in M pneumoniae significantly affects protein abundance and turn over, acting at both transcriptional but also post-translational levels", but only 39 out of 447 (9%) proteins seem to show differences in abundance levels.

13. Out of 67 phospho-sites the authors identify that 16 (23.9%) phospho-sites were not affected by deletions of the kinases or phosphatases. This is a significant number, almost 1/4th the phospho-sites, the authors have not discussed this clearly.

14. The authors conclude that because "an alternative lysine could be frequently found within a window of three amino acids upstream or downstream of the original site (Fig. 2d), this suggests that for some lysine acetylation sites, the exact position may not be so critical to maintain function." In the absence of direct evidence that specific modifications are redundant, this conclusion is not supported by the results. Clusters of multiple acetylated residues may be functionally significant, as a result of cooperativity, specific interaction partners, etc.

15. When the authors mention "levels of occupancy" do they mean that the phosphorylation event is present but not to that extent of the wild type. If so then it is not made clear how do they get phosphorylated in the first place when their kinases are mutated.

16. The discussion in the section 'Structural support for a molecular barcode in prokaryotes' is not entirely clear. The authors talk about multiply modified proteins being in the hub of signaling networks, which is a known observation. Then they give evidence for protein-protein interaction surfaces being the places of modifications, and then they talk about gel filtration chromatography of two proteins. The arguments are not cohesive and it is difficult to even interpret the context.

Minor points:

1. The authors suggest that there is no large scale study that consistently investigated the direct modulation of one PTM by another (page 3). But there have been studies where large scale protein ubiquitination and protein phosphorylation have been investigated.

2. Page 4 first paragraph "to gather insides into the mechanistic of prokaryote ...." should be "to gather insights" and re-worded. The sentence seems to be little confusing.

3. There is no reference for the statement that serine/threonine phosphorylation are ancient PTMs. Histidine and Aspartate two component signaling system was one of the ancient PTMs.

4. In page 7, the sentence "Phosphosites which levels of occupancy have been exclusively affected by deletion...." should be re-written. It is not clear.

5. Their western blot figures are not shown and they don't indicate how they validated the western blot information.

Reviewer #3 (Remarks to the Author):

The authors report a quantitative protomic analysis of Mycoplasma pneumoniae, focusing on the interaction between protein phosphorylation (S/T/Y) and protein acetylation (K). The experimental setup for mass spectrometry proteomics is sound, and data appear to be of good quality. These results are further supported by some transcriptomics, Western analysis and so on, which is very good. However, I feel some analyses on the acetylation influence on phosphorylation is required, and I have some reserves about the interpretation of some data.

Major remarks:

1. The title states that phosphorylation is intertwined with different layers of PTMs, but the manuscript is focused almost exclusively on the interplay between phosphorylation and acetylation. The title should reflect this.

2. The experimental setup of the manuscript is strangely uni-directional, it probes the influence of phosphorylation on acetylation, i.e. focuses on mutants of two kinases and a phosphatase. However, this interplay may be bi-directional. The reader cannot help but wonder which enzyme(s) acetylate(s) all these proteins in M. pneumoniae, what happens if they are knocked out? How is acetylation affected, and what about phosphorylation?

3. On page 8 the authors argue that some P-sites are affected by both PknB and HprK KOs. They conclude that this suggests that these kinases act sequentially. How so? How does this result preclude them from acting in parallel (overlapping specificities)? If they act sequentially, can the authors show which acts first and which second, and does one activate the other?

4. I am not at all comfortable with the "molecular barcode" metaphor. A barcode is attached to an item, does not modify it in any meaningful sense and is read by an independant detector. PTMs often modify proteins' chemical properties and behavior profoundly, and are most of the time not "read" by a detector. SH2 domains might be likened to a detector, but these do not exist in bacteria, as mentioned by the authors. Please reconsider.

5. At the end of the "introduction" and "conclusions" the authors argue that the existence of interplay between phosphorylation/acetylation in both eukaryotes and prokaryotes supports the idea that it evolved before they diverged. If that is so, what is the evidence presented by the authors to exclude convergent evolution of phosphorylation/acetylation systems and their interplay?

6. In figure 6, elution profiles of some proteins isolated from WT and kinase mutants are used as evidence that phosphorylation impacts protein-protein interactions. With EF-Tu and RpsB the profile stops before the protein is eluted, where is the other part of the elution peak (as seen for GroEL)? Are there more peaks? Why are there no molecular wieght standards on the profile? They are needed to estimate the size of the WT complex. Differences in elution profiles could be due to problems with protein folding/stability/aggregation and not necessarily protein-protein interactions (could be checked by CD). This is not enough to support the conclusions.

Minor remarks:

1. The manuscript would gain a lot by language proofreading. For example, the results section starts with the sentence: "To gather insides into the mechanistic of..." One can only guess that by "insides" they meant "insights". "Mechanistic" is an adjective that appears to be used as a noun. This, and many other small incidents are disruptive for the flow of the manuscript and should be corrected.

2. On page 2: "Protein phosphorylation is catalyzed by kinases and phosphatases..." Protein phosphatases do not catalyze protein phosphorylation.

3. M. Pneumonie is described as a self-replicating organism. Are there organisms that cannot self-replicate?

4. On page 5 it seems to me that the authors consider the lysine residues, found to be acetylated in M. pneumoniae, and CONSERVED in eukaryotes, to be conserved acetylation sites. This is only true if they were also detected as ACETYLATED in eukaryotes. Please correct me if I al missinterpreting what I've read.

1st Revision - authors' response   06 December 2011

Reviewer #1 (Remarks to the Author):

This study uses MS/MS to provide a global perspective of the connections between phosphorylation and protein abundance/acetylation patterns in Mycoplasma pneumoniae. These authors published another study on the M. pneumoniae phosphoproteome using the same mutants several months ago

(Schmidl et al, 2010 Mol Cell Proteomics). However, the protein abundance and lysine acetylation data provides a wealth of new results to warrant publication, pending the following clarifications (especially with respect to reproducibility of the lysine acetylation data).
Manuscript format:
The manuscript should be reformatted to meet MSB guidelines. It lacks a Methods and Materials section. Also, references are presented by authors in the text but are listed numerically in the order they appear in this reference section, making it cumbersome to find the references.

We apologize for the inconvenience. We reformatted to meet MSB guidelines. We have now added a Materials and Methods section page 15 and have reformatted of the bibliography.

Figures:
Fig 1C: The table portion of this figure reports the same data as supplementary table 4. The plot is a cartoon of how peptide abundance changes were normalized for MS intensity. As the method for quantifying peptide abundance ratios is not discussed in the text and simply references another paper (Mortensen et al, 2010) in the supp info, this cartoon should be removed.

We agree with the reviewer that Table S4 and Figure 1C are to a great extend redundant. The numbers being displayed provide an important overview on the dataset and we decided keep Figure 1C and to remove Supplementary Table S4.
Regarding the methods for the quantification of peptides, we believe this is one of the strengths of the present analysis. In contrast to previous work reporting quantification of PTMs (for example, Bodenmiller *et al.* 2010, *Sci. Signal.* 3:rs4 and Choudhary *et al.* 2009, *Science* 325:834), here the changes in protein phosphorylation and lysine acetylation have been normalized for possible changes in protein abundance. We have now included in the main text, pages 5-6, a description of the mass spectrometry and quantification approaches we performed including the normalization procedure. We now refer to the important table in Figure 1C in the main text.

Fig 1D: This figure should be removed as it is not referenced in the text and appears to be a simplified cartoon of Fig 5.

We followed this reviewer's suggestion and now removed panel D from Figure 1.

Fig 2A. What does "% coverage" on the ordinate axis mean? Mean percent coverage of the proteins with peptide identified by mass spec? Fraction of the proteins in each COG identified by mass spec?

We clarified this point and changed the coordinate axis to "% of proteins in a COG class".

Fig 4. This figure shows strong global decreases in lysine acetylation in the pknB mutant and corresponding increases in the hprK mutant. However, this direct connection between kinase inactivation and lysine acetylation is not discussed in the text. Why?

We thank the reviewer for this interesting observation that we now mention page 12, paragraph 1.

Supp Info:
Mass spec data is reported in the text and in Fig 1B to include both biological and technical replicates. However, data in the supplementary information for abundance (data set 1) phosphorylation (data set 2) and lysine acetylation (data set 3) report only mean changes without standard deviations or data from the replicates. Please report data from the replicates so the reader can assess reproducibility.

The reported data includes fold-changes as well as p-values. The p-values are derived from the entire datasets and not only the technical and biological duplicates. The p-values are derived based on the variation in all peptides with similar intensities using quantile regression. Indeed a log2 fold change of 1 for peptide with low intensity is much less reliable than for a peptide with a high intensity. The p-values are therefore much more informative than the fold changes and standard deviations based on two or four replicates of only one single peptide. This is a well established and accepted procedure that has been adopted for the quantification of a variety of large datasets (Anders & Huber (2010) Genome Biology 11:R106).

Only changes that were statistically significant and that could be further confirmed by visual inspection of the XICs (chromatograms) were considered for further analysis. We calculated reproducibility post-analysis for up- and down-regulation between replicates; its reaches 85%, 93%, 47% for protein, phosphopeptides and acetyl lysine, respectively (Supplementary Figure 1F; before filtering). Most importantly, however, for the final datasets (Supplementary Tables S1-S3), all changes in abundances that were not reproducible between the technical duplicates were excluded (Supplementary Figure 1F; after filtering). We included reproducibility as an intrinsic criterion for the identification of regulated proteins, phosphosites, and lysine-acetylated sites in the final tables (Supplementary Tables S1, S2 and S3). This is now described page 5 (last paragraph) and page 6 (paragraph 1).

Questions about phosphorylation results:
Schmidl et al, 2010 identified 63 phosphosites by comparing wild-type to the same mutants strains, compared to 93 sites in this study. However, only 30 sites are common to both studies. What caused these differences? Do the authors believe that the actual phosphoproteome is the union or intersection of these datasets?

There must be a slight confusion. Schmidl *et al.* 2010 identified 63 phosphorylated proteins, but only 16 phosphosites (on 14 proteins). Schmidl *et al.* identified phosphorylated proteins by staining 2D gels with a phospho-dye, ProQ Diamond. The stained spots were then cut from the 2D gel and proteins were subsequently identified by mass spectrometry. All proteins identified within a ProQ Diamond stained spot were considered phosphorylated, even though some unphosphorylated proteins might have co-migrated. Such 2D gel based methods are known to have two main limitations: limited sensitivity of the dye ProQ Diamond and low accuracy (because of co-migrating proteins and limited specificity of the dye ProQ Diamond). This is why shotgun approaches, such as the one we used here, are generally preferred. De facto, only proteins for which the phosphopeptide have been characterized represent solid and unambiguous phosphorylation events. In the case of the work by Schmidl *et al.* this represents 14 proteins in total: our work recapitulates 10; four represent novel proteins that we missed in our analysis.
At peptide level, Schmidl *et al.* 2010 identified 11 additional peptides that are not in our current dataset. This could be explained by the use of different separation/fractionation methods (2D gels or SCX). Also Schmidl *et al.* 2010 used less stringent peptide probability scores for phosphopeptide identification. We mention the overlap with Schmidl *et al.* 2010 page 6, paragraph 2.

I was surprised to see that 14% of kinase-regulated proteins showed higher phosphorylation in the kinase mutants. How do the authors explain this 'inverted directionality'? Does deletion of the normal kinase cause the other kinase to overcompensate?

There could be several explanations, the most likely being the existence of complex regulatory events whereby kinases and phosphatases might directly or indirectly regulate each other's activities for a subset of substrates. We clarify on page 10, paragraph 2.

Phosphorylated peptides were separated by strong cation exchange (SCX). Other studies (a few refs below) have improved upon this method by including a second phosphopeptide enrichment by titanium dioxide (TiO2) or immobilized metal affinity chromatography(IMAC). Why wasn't a second phosphopeptide enrichment used in this study?
Olsen et al, 2010 "Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis" Science Signalling.
Villen and Gygi 2008 "The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry" Nature Protocols.

The choice of analytical strategies really depends on the complexity of the sample being analyzed. The methods mentioned by the reviewer are indeed well suited for complex eukaryotic proteomes (20,000 genes, many alternative splice variants, etc) where extensive and multidimensional fractionation steps are required, but certainly overwhelming for such a simple organism as *Mycoplasma pneumoniae* (691 genes). Consistent with this view, in four pilot experiments, phosphopeptide enrichment using TiO2 in addition to SCX fractionation, led to identification of only two more peptides (data not shown). We mention this observation page 6, paragraph 2.

Questions about acetylation results:

This study reports 719 lysine acetylation sites, showing this PTM affects >30% of proteins and is 8 times more frequent than phosphorylation. Indeed, the remarkably high level of acetylation and its connection to phosphorylation is one of the main conclusions of the study. However, this is remarkably high level of lysine acetylation requires some clarification.
1.Supplementary Fig 2D shows that the reproducibility of the acetylation data between technical replicates was only r2=0.13. Further, Supplementary Fig 2E plots biological replicates for protein abundance and phosphorylation, but omits acetylation. The manuscript text should mention the low reproducibility between technical replicates. Fig 1B suggests that biological replicates were run for lysine acetylation. Please show the biological replicates of lysine acetylation data in Supplementary fig 2E.

The datasets imply two levels of analysis. One level is the identification of proteins, peptides and PTMs. The second level is the quantification of these.
On the first level, the biological reproducibility is high and varies from 70% to 98% (Supplementary Figure 1F). For comparison Mann *et al.* report in human cell lines, a maximum reproducibility of 57% for the identification of acetyl-lysine (Choudhary *et al.* 2009, *Science* 325:834). To clarify and address this important point we now include an extra panel (F) in Supplementary Figure S1 that describes reproducibility at the level of proteins, phosphopeptides and lysine acetylated peptides identifications. This benchmark clearly demonstrates that the datasets of identified protein, phosphosites sites and lysine acetylation sites are reproducible and of high quality.
The quantification is the second level of analysis. Supplementary Figures 2D and 2E (now renumbered Supplementary Figures 1D and 1E) show the correlations between the quantitative measurements in the technical or biological replicates. These are measured directly from the MS data before statistical analysis and filtering; so the figures don't represent reproducibility of regulation. We now clarify this in the legend of Supplementary Figure 1. The figures show the experimental variation we had to deal with during the analysis to filter out the experimental noise. For each peptide, the statistical significance of the observed change in abundance was computed using quantile regression. The test provides a p-value based on the variation in the normalized ratios observed for all peptides with similar intensities. Only changes that were statistically significant and that could be further confirmed by visual inspection of the XICs (chromatograms) were considered for further analysis. We also set stringent thresholds and only changes higher than ~2.8x (for proteins and phosphopeptides) and 4x (for lysine-acetylated peptides) were further considered. The reproducibility post-analysis for up- and down-regulations between technical replicates reaches 85%, 93%, 47% for protein, phosphopeptides and acetyl lysine, respectively (Supplementary Figure 1F; before filtering). Most importantly, however, for the final datasets (Supplementary Tables S1-S3), all changes in abundances that were not reproducible between the technical duplicates were excluded (Supplementary Figure 1F; after filtering). We included reproducibility as an intrinsic criterion for the identification of regulated proteins, phosphosites, and lysine-acetylated sites in the final tables (Supplementary Tables S1, S2 and S3). This is now described page 5 (last paragraph) and page 6 (paragraph 1).
The quantification of acetylated-lysines were done on the same lysates as can be seen from the Figure 1B, we thus have no biological duplicates for the lysines-acetylated peptides; this is why we did not include in Supplementary Figure 1. We report biological reproducibility for changes in protein abundance and phosphorylation, 63% and 90%, respectively.

2. Protein abundance and phosphopeptide data were filtered to an FDR of 1% using a target-decoy approach. In contrast, the FDR for the acetylation data was estimated from C-terminally acetylated lysines. Does this mean that different FDR methods were used for the datasets? If so, how many acetylation sites are found when the target-decoy approach used for protein abundance and phosphorylation is applied to the acetylation data?

We apologize for the lack of clarity. All peptide identifications were filtered to FDR of 1% using a target-decoy approach and this also includes the acetylation datasets. As one of our main messages relies on the quantification of lysine acetylation, we wanted to be careful and estimated FDR based on an **additional** criterion, i.e. C-terminally acetylated lysines. We have now clarified this point, pages 21 (last paragraph) and 23 (paragraph 1) of the Material and Methods.

3. The phosphorylation data was supported using kinase/phosphatase mutants. Given the high acetylation levels, why is there no discussion in the text about how this may be occurring other than

a single mention of the N-acetyltransferase Mpn114? Moreover, why wasn't a similar approach using a mutant of Mpn114 used to support the acetylation data?

Two putative lysine acetyl transferases have been predicted in *M. pneumoniae*, Mpn027 and Mpn114. To date no deacetylases have been identified. We performed additional experiments to address this reviewer's question. We analyzed the *M. pneumoniae* strains with a mutation in Mpn027 or Mpn114. As expected this affected the level of lysine acetylation in both strains. We additionally observed significant changes in the level of phosphorylation consistent with the view that the cross-talk goes both directions. We included a paragraph describing these observations page 12 (last paragraph) and page 13 (first paragraph) and a Supplementary Figure 8.

Reviewer #2 (Remarks to the Author):

In this work "Phosphorylation is intertwined with different layers of post-transcriptional regulation in a genome-reduced bacterium" Noort et al. have attempted to understand the interplay between protein phosphorylation and lysine acetylation. It is an important question in the field of protein PTMs. They have utilized Mycoplasma pneumoniae, a prokaryote with one of the most reduced genome and which has only two kinases and one phosphatase. The advantage of this system is that it enables them to do combinatorial perturbations of the kinases and phosphatase and observe the effects on total phosphorylation and lysine acetylation at an organismal level. However since it's an endocellular parasite (using the host systems genetic machinery) it is not clear how these perturbations translate into biological phenotype of the organism. This study points out to some interesting observations and it has got good data set but it fails to convince on the broader claims made by the authors
The overall aim and goal of the study seems to warrant a publication in MSB but lot of assertions by the authors need to be substantiated with real evidence and the manuscript has to be re-written. Points of note:
1. Their title and abstract seem to be slightly misleading there is not much evidence and discussion about post-transcriptional regulation. The definition of their post-transcriptional regulation seems not to be clear (or did the authors mean post-translational?).

According to this reviewer's suggestion we changed the title to "Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium."

2. The authors have not clearly explained and discussed how they account for the differences in the proteome abundance without any changes in the transcriptional levels. Without clearly indicating and investigating what those mechanisms might be leading to this discordant observation between protein abundance changes and no-change in transcription level their argument seems to be more of a hypothesis than a premise of a paper. Their interpretation that the phosphorylational perturbations in chaperone proteins may lead to protein abundance changes lacks validity and has to be proved.

Other groups reported similar observations both in pro- and eukaryotes (Khositseth et al. 2011, Mol Cell Proteomics 10:M110.004036; Schmidl et al. 2010, Infect Immun 78:184). Several mechanisms have been proposed to account for this effect, including changes in the rates of translation, in protein stability. The question clearly goes beyond the scope of this manuscript. We agree that changes in chaperone protein activity or specificity represent only one out of many other hypotheses and this is why we broaden the discussion pages 14 (last paragraph) and 15 (paragraph 1) to include additional possibilities "Many components of the translational machinery, including ribosomal proteins, tRNA synthetases, translation initiation and elongation factors, and chaperones were affected by the perturbation of the phosphorylation network, which might account for observed changes in protein abundance that are not obviously the result of transcriptional regulation."

3. They don't talk about Tyrosine phosphorylation, though Fig. 2b indicates there is some Tyrosine phosphorylation observed.

We now mention page 6, paragraph 2, the tyrosine phosphorylation events we observed and refer to previous proteomics analysis in bacteria that also reported tyrosine phosphorylation. The enzymes or mechanisms responsible for tyrosine phosphorylation in bacteria are unknown.

4. On page 3 the authors say that they combine genetics and high resolution quantitative mass spectrometry to measure the global effect of serial kinase and phosphatase deletions on proteome abundance, phosphorylation and lysine acetylation. The authors could have just mentioned about the 3 mutants and discussed their MS strategy in detail giving more credence to their work. More importantly they should explain how quantitative was their MS approach and what are the advantages of the chosen labeling techniques. In short, the strength and the weakness of the paper lies in the quantitative MS approach, to make their arguments appealing about the quantitation and the conclusion they should discuss the MS approach in more detail and convince the readers that they were done with utmost care. The word serial kinase and phosphatase deletions are also misleading, there were just 3 mutants.

We rephrased this sentence (now page 5) and just mentioned we deleted the predicted kinases and phosphatases. We also included a new paragraph, page 5 (last paragraph) and page 6 (paragraph 1), describing how we performed the quantitative mass spectrometry experiments and the data analysis. We additionally included a paragraph, page 6 (paragraph 1) describing data reproducibility. Finally, we moved the Material and Methods section from the Supplementary Information to the main text page 15. We hope the new version addresses now this reviewers concern and convinces the reviewer that the work was done with utmost care.

5. In their discussion the authors have not discussed the biological variation between the 3 differently labelled wild type samples. It is difficult to assess what they report in their results were whether normal biological variation or over and above biological variation.

We included a description of reproducibility page 6, paragraph 1. We show that for the identification of proteins and PTMs, biological reproducibility is high and varies from 70% to 98% (described in a new Supplementary Figure 1F). For comparison Mann *et al.* report in human cell lines, a maximum reproducibility of 57% for the identification of acetyl-lysine (Choudhary *et al.* 2009, *Science* 325:834). This benchmark clearly demonstrates that our datasets of identified proteins, phosphosites sites and lysine acetylation sites are reproducible and of high quality.
For the quantification we measured the statistical significance of the observed change in abundance using quantile regression. The test provides a p-value based on the variation in the normalized ratios observed for all peptides with similar intensities. Only changes that were statistically significant and that could be further confirmed by visual inspection of the XICs (chromatograms) were considered for further analysis. We also set stringent thresholds and only changes higher than ~2.8x (for proteins and phosphopeptides) and 4x (for lysine-acetylated peptides) were further considered. The reproducibility post-analysis for up- and down-regulation between technical replicates reaches 85%, 93%, 47% for protein, phosphopeptides and acetyl lysine, respectively (Supplementary Figure 1F; before filtering). Most importantly, however, for the final datasets (Supplementary Tables S1-S3), all changes in abundances that were not reproducible between the technical duplicates were excluded (Supplementary Figure 1F; after filtering). We clarify that for the quantification of proteins and PTMs, we used technical reproducibility as a filtering criterion for the identification of regulated protein, phosphosites, and lysine acetylated sites in the final tables (Supplementary Tables S1, S2 and S3) page 5 (last paragraph) and page 6 (paragraph 1). We also report biological reproducibility for changes in protein abundance and phosphorylation, 63% and 90%, respectively.

6. "All fractions were analyzed by high resolution Mass spectrometry" on page 4 suggests next to nothing. Since the paper is heavily reliant on the MS quantitation the authors should explain why and how they did this.

We now replace "All fractions were analyzed by high resolution Mass spectrometry" with "All fractions were analyzed by nano LC-LTQ-Orbitrap (Thermo, San Jose, CA)". Following this reviewer suggestion, we also included a detailed section describing the MS quantification, page 5 (last paragraph) and page 6 (first paragraph). The Materials and Methods section has now been moved to the main text (page 15).

7. "Out of 564 proteins 460 were quantified" again is not a helpful statement to understand why the remaining proteins were discarded.

The other proteins were not discarded; we simply did not have enough data for quantification. We apologize for the lack of clarity, but this is a well known phenomena: quantification requires more stringent criteria than simple protein or peptide identifications. As a consequence, MS quantification cannot always be achieved for all MS identifications. For example, in some cases proteins were not identified with enough peptides to allow unambiguous and accurate quantification (we require at least three); this means the protein was identified but could not be quantified. In other cases, the quantification software MSQuant could not identify a clear and clean XIC peak for given peptides, e.g. there were other peptides co-eluting with the peptides of interest, precluding reliable quantification. We clarify this point page 6, paragraph 2.

8. In fig 1c the authors have observed that there were 93 phospho-sites on 92 peptides. Does this imply that identified a single doubly phosphorylated peptide?

We recognize that the reason for this apparent discrepancy was somewhat unclear. As can be read in Supplementary Table S2, there are indeed two peptides with two phosphosites and two overlapping peptides that cover one single phosphosite. We have added this in the figure legend of the Figure 1 page 36.

9. Fig 2a is not very informative, categorizing proteins into metabolic, cell signaling and information storage and processes pathways doesn't establish that there were no biases in the MS identifications. There would be so many over laps in each category.

We have now rephrased this section page 7 (paragraph 2).

10. The assertion that lysine acetylation is almost 8 times as frequent as phosphorylation could be biased. The authors claim that the median sequence coverage was only 43% indicating that half of the peptides were not identified by MS. This finding may also be explained by the differences in the experimental protocols used to purify phospho- and acetyl-containing peptides. The former were purified by ion exchange chromatography, while the latter were purified using an antibody. A more accurate comparison of differential abundance of these two modifications should involve a shared reference for normalizing differences in enrichment affinity, such as a phospho- and acetyl-modified peptide standard.

The fact that the use of alternative methods to enrich for phosphorylation, such as TiO2 or 2D gels, led to only two and 11 additional phosphopeptides, respectively) suggests that the phosphoproteome reported here is probably comprehensive. We believe that the general message, lysine acetylation is a frequent PTM, at least as frequent as phosphorylation, holds true.
This being said, we agree with the reviewer that it is difficult to quantitatively compare phosphorylation and lysine acetylation given the fact that they have been identified following different enrichment methods, which effectiveness may largely vary between individual peptides. Also they may behave quite differently in MS experiments. We have now thus toned down this somewhat strong assessment page 7 (paragraph 2) and we now say that "The results show that in *M. pneumoniae*, but likely also in other prokaryotes, lysine acetylation is a very common modification, being apparently at least as frequent as phosphorylation".

11. The statement "Of the 45 relative abundances in the three knockouts 39 were in agreement with the MS data" is not clear.

We now clarify page 8, paragraph 3: "Of the 45 abundances measured by western blotting in the three knock-outs, 39 (86.6%) showed up- and down regulation consistent with the previous MS data (Supplementary Figure 5)."

12. On page 7 the authors claim that "these results show that perturbations of phosphorylation network in M pneumoniae significantly affects protein abundance and turn over, acting at both transcriptional but also post-translational levels", but only 39 out of 447 (9%) proteins seem to show differences in abundance levels.

We toned down this strong statement page 9, paragraph 1.

13. Out of 67 phospho-sites the authors identify that 16 (23.9%) phospho-sites were not affected by deletions of the kinases or phosphatases. This is a significant number, almost 1/4th the phospho-sites, the authors have not discussed this clearly.

We now address this point more clearly page 9, paragraph 2: "They might represent compensatory mechanisms, whereby deletion of one kinase may cause the other kinase to compensate. Alternatively they may account for HprK-, PknB- and PrpC-independent phosphorylation events, including autophosphorylation of some metabolic enzymes {Jolly, 2000 #619} or metabolic intermediates observed at catalytically active sites. For example, the constitutive phosphoserine (S64) in the ATP-binding site of the guanylate kinase (Gmk|Mpn246) might represent a metabolic intermediate, since in the *Escherichia coli* structure the equivalent serine lies very close, though not obviously bound, to both phosphate and sulfate groups."

14. The authors conclude that because "an alternative lysine could be frequently found within a window of three amino acids upstream or downstream of the original site (Fig. 2d), this suggests that for some lysine acetylation sites, the exact position may not be so critical to maintain function." In the absence of direct evidence that specific modifications are redundant, this conclusion is not supported by the results. Clusters of multiple acetylated residues may be functionally significant, as a result of cooperativity, specific interaction partners, etc.

We apologize for being unclear in this statement; we have not found clusters of acetylated lysines. We considered aligned positions where the *M. pneumoniae* acetylated lysine is not conserved in another species, i.e. in the other species there is no lysine at this position. Then, in this other species there is a lysine within three amino acids upstream or downstream from the site aligned to the *M. pneumoniae* site. We have changed the sentence page 8, first paragraph, to reflect this. "However, **when the acetylated lysine was not conserved,** an alternative lysine could be frequently found **in other species** within a window of three amino acids, one upstream or one downstream of the original **aligned** site (Figure 2D)".

15. When the authors mention "levels of occupancy" do they mean that the phosphorylation event is present but not to that extent of the wild type. If so then it is not made clear how do they get phosphorylated in the first place when their kinases are mutated.

We apologize for the lack of clarity. What we meant is that the impact of protein kinases and phosphatases deletion largely depends on the extent of their substrate phosphorylation before perturbation, i.e. in the wild type cells. Indeed for sites fully phosphorylated in the wild type, knocking down the phosphatase won't lead to further increase; such sites can only be affected by kinase knock-outs. Reciprocally for sites kept mainly dephosphorylated in the wild type, knocking down the kinase won't give measurable decrease; such sites are only be affected by phosphatase deletion. We thought to use the abundance profiles across the different knock-out strains to derive information on the phosphorylation stoichiometries of the different substrates. We have now rephrased this paragraph page 9 (paragraph 2).

16. The discussion in the section 'Structural support for a molecular barcode in prokaryotes' is not entirely clear. The authors talk about multiply modified proteins being in the hub of signaling networks, which is a known observation. Then they give evidence for protein-protein interaction surfaces being the places of modifications, and then they talk about gel filtration chromatography of two proteins. The arguments are not cohesive and it is difficult to even interpret the context.

We have extensively rewritten this section pages 13 and 14. We now quote the work from Matthias Mann's lab (PMID: 19608861) who showed that proteins being part of protein complexes are more frequently lysine acetylated. We further refine this observation and show that PTMs often occur at interaction interfaces and in multifunctional proteins. We have improved the method, and have rerun the analysis. We include additional examples that illustrate how modifications can affect the oligomerisation state of proteins in a modified Figure 6. For the chaperone GroS, we also provide a more cohesive example, bringing experimental support (sucrose gradient) to the structural models predicting that S29 phosphorylation could interfere with GroS oligomerization.

Minor points:

1. The authors suggest that there is no large scale study that consistently investigated the direct modulation of one PTM by another (page 3). But there have been studies where large scale protein ubiquitination and protein phosphorylation have been investigated.

We have found indeed previous reports on the impact of SUMOylation on protein phosphorylation. We apologize for this omission and have now included the reference (PMID:21685386) page 4, first paragraph.

2. Page 4 first paragraph "to gather insides into the mechanistic of prokaryote ...." should be "to gather insights" and re-worded. The sentence seems to be little confusing.

We have now corrected this typo (now page 5 first paragraph).

3. There is no reference for the statement that serine/theronine phosphorylation are ancient PTMs. Histidine and Aspartate two component signaling system was one of the ancient PTMs.

We added the references to the statement, page 7, paragraph 3, that serine/threonine phosphorylation and lysine acetylation are ancient PTMs.

4. In page 7, the sentence "Phosphosites which levels of occupancy have been exclusively affected by deletion...." should be re-written. It is not clear.

We clarified the sentence now on page 9, paragraph 2.

5. Their western blot figures are not shown and they don't indicate how they validated the western blot information.

We do not understand what the reviewer means. All western blot figures have been made available as Figures or supplementary Figures:
i) the changes in protein lysine acetylation upon PknB, HprK or PrpC knock-out are in the Supplementary Figure 2.
ii) the changes in protein abundance upon PknB, HprK or PrpC knock-out are in the Supplementary Figure 5.
iii) the changes in the elution profile of GroS and RplA are in Figure 6 and Supplementary Figure 9.

Reviewer #3 (Remarks to the Author):

The authors report a quantitative proteomic analysis of Mycoplasma pneumoniae, focusing on the interaction between protein phosphorylation (S/T/Y) and protein acetylation (K). The experimental setup for mass spectrometry proteomics is sound, and data appear to be of good quality. These results are further supported by some transcriptomics, Western analysis and so on, which is very good. However, I feel some analyses on the acetylation influence on phosphorylation is required, and I have some reserves about the interpretation of some data.
Major remarks:
1. The title states that phosphorylation is intertwined with different layers of PTMs, but the manuscript is focused almost exclusively on the interplay between phosphorylation and acetylation. The title should reflect this.

According to this reviewer's suggestion we changed the title to "Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium."

2. The experimental setup of the manuscript is strangely uni-directional, it probes the influence of phosphorylation on acetylation, i.e. focuses on mutants of two kinases and a phosphatase. However, this interplay may be bi-directional. The reader cannot help but wonder which enzyme(s) acetylate(s) all these proteins in M. pneumoniae, what happens if they are knocked out? How is acetylation affected, and what about phosphorylation?

Two putative lysine acetyl transferases have been predicted in *M. pneumoniae*, Mpn027 and Mpn114. To date no deacetylases have been identified. We performed additional experiments to address this reviewer's question. We analyzed the *M. pneumoniae* strains with mutation in Mpn027 and Mpn114. As expected this affected the level of many lysine acetylation sites. We additionally observed significant changes in the level of phosphorylation consistent with the view that the cross-talk goes both directions. We included a paragraph describing these observations page 12, last paragraph and page 13, first paragraph. The results are described in a new Supplementary Figure 8.

3. On page 8 the authors argue that some P-sites are affected by both PknB and HprK KOs. They conclude that this suggests that these kinases act sequentially. How so? How does this result preclude them from acting in parallel (overlapping specificities)? If they act sequentially, can the authors show which acts first and which second, and does one activate the other?

Kinase A modulates kinase B which phosphorylates substrate C; knock-out of either kinase will affect substrate C phosphorylation. This being said, we agree with the reviewer the two hypotheses hold and we in fact already captured this point our next sentence. We realize we were probably not clear enough and now rephrase this part page 10, paragraph 2: "A series of phosphosites were also found to be affected by both PknB and HprK knock-outs (17.6% of all regulated sites)(Figure 3 and Supplementary Table S2), suggesting the two kinases act sequentially. Alternatively, for some of the substrates, the two kinases might have redundant specificities."
We cannot say which kinase acts first or second from this type of genetic data; for example, kinetic studies upon, say, chemical knock-outs, would be required which is beyond the scope of this work.

4. I am not at all comfortable with the "molecular barcode" metaphor. A barcode is attached to an item, does not modify it in any meaningful sense and is read by an independant detector. PTMs often modify proteins' chemical properties and behavior profoundly, and are most of the time not "read" by a detector. SH2 domains might be likened to a detector, but these do not exist in bacteria, as mentioned by the authors. Please reconsider.

The molecular barcode metaphor has been proposed by many others. It represents a broadly accepted concept in eukaryotes: some PTMs affect protein-protein interaction (writer-reader-eraser metaphor). As pointed by this reviewer, this certainly represents only one of the many (other) functions of PTMs (molecular switches, etc). In *M. pneumoniae*, we observe that phosphorylation and lysine acetylation happen frequently in interaction interfaces, suggesting a role in the regulation of protein interaction. We present structural and experimental support for this hypothesis (page 13 "Structural support for a molecular barcode in prokaryotes") and would like to propose that mechanisms reminiscent of the eukaryotic molecular barcode (may be a more ancestral one) also exists in prokaryotes. The "readers" are not limited to series of SH2, 14-3-3, bromo, etc domains but may more generally imply interaction interfaces.
We nevertheless followed this reviewer's suggestion and toned down our statement. We replaced page 13 (paragraph 3) the title "Structural support for a molecular barcode in prokaryotes" by "PTMs target interaction interfaces, altering protein oligomerization states". We now speculate on the existence of a possible molecular bar code in prokaryote at the end of the paragraph 1, page 14.

5. At the end of the "introduction" and "conclusions" the authors argue that the existence of interplay between phosphorylation/acetylation in both eukaryotes and prokaryotes supports the idea that it evolved before they diverged. If that is so, what is the evidence presented by the authors to exclude convergent evolution of phosphorylation/acetylation systems and their interplay?

We report the conservation level of phosphorylation and acetylation sites and observe a higher level of conservation than for unmodified residues. The most parsimonious explanation is that they existed before divergence of eukaryotes and prokaryotes, but of course this reviewer is right and we cannot exclude convergent evolution. We have thus now toned down our statements page 4, last paragraph, and page 15, paragraph 2.

6. In figure 6, elution profiles of some proteins isolated from WT and kinase mutants are used as evidence that phosphorylation impacts protein-protein interactions. With EF-Tu and RpsB the profile stops before the protein is eluted, where is the other part of the elution peak (as seen for GroEL)? Are there more peaks? Why are there no molecular wieght standards on the profile? They are needed to estimate the size of the WT complex. Differences in elution profiles could be due to

problems with protein folding/stability/aggregation and not necessarily protein-protein interactions (could be checked by CD). This is not enough to support the conclusions.

We have extensively rewritten this section page 13 and 14. We have improved the structure-based predictions, and have rerun the analysis. We have modified the Figure 6 to include additional examples that illustrate how modifications can affect the oligomerization state of proteins. For the experimental validation, we chose a better example, the chaperone GroS, for which we have directly tied the structural models, predicting that S29 phosphorylation could interfere with GroS oligomerization, to experimental support (sucrose gradients).
To answer this reviewer's points, we now provide elution profiles covering all fractions. We also added molecular weight standards. We provide results based on sucrose gradient experiments run in (biological) triplicates– and we report statistical significance for the sedimentation profiles of GroS in the PrpC mutant strain (new Figure 6). As CD is not feasible because it requires large amounts of highly purified proteins and we work here with endogenously expressed proteins (to avoid artefacts of over-expression), we thought to confirm the sedimentation profiles using different separation methods, elution/retention during gel filtration (data not shown).

Minor remarks:
1. The manuscript would gain a lot by language proofreading. For example, the results section starts with the sentence: "To gather insides into the mechanistic of..." One can only guess that by "insides" they meant "insights". "Mechanistic" is an adjective that appears to be used as a noun. This, and many other small incidents are disruptive for the flow of the manuscript and should be corrected.

We apologize for the inconvenience; the manuscript has been now extensively proofread.

2. On page 2: "Protein phosphorylation is catalyzed by kinases and phosphatases..." Protein phosphatases do not catalyze protein phosphorylation.

We corrected the sentence, page 3, first paragraph: "Protein phosphorylation is regulated by a variety of kinases and phosphatases which are themselves regulated by phosphorylation within complex networks."

3. M. Pneumonie is described as a self-replicating organism. Are there organisms that cannot self-replicate?

The phrasing is correct and there are indeed organisms such as endosymbionts that are unable to self-replicate, i.e. replicate without the help of the host cellular machinery.

4. On page 5 it seems to me that the authors consider the lysine residues, found to be acetylated in M. pneumoniae, and CONSERVED in eukaryotes, to be conserved acetylation sites. This is only true if they were also detected as ACETYLATED in eukaryotes. Please correct me if I al missinterpreting what I've read.

Indeed we analyzed the site conservation and not the modification conservation in over 300 species and compared this to other residues in the same proteins. In the absence of comprehensive proteomic datasets covering many species, this is a well accepted approximation of PTM site conservation and functionality (PMID:21791702).

2nd Editorial Decision   **20 January 2012**

Thank you again for submitting your work to Molecular Systems Biology. First of all, I would like to apologize for the delay in getting back to you, which was due to the Christmas break. We have now heard back from the three referees who accepted to evaluate the study. As you will see, reviewer #1 and #3 are fully supportive of the work. Reviewer #2 acknowledges that the study has been significantly improved but still raises several points that deserve clarification. We would thus

ask you to revise the text to address these remaining issues in a minor revision of the present manuscript.

In particular, the number of independent biological replicates should be clearly stated and the statement that "the biological reproducibility for the identification of proteins, phosphopeptides and lysine-acetylated peptides varied between 70% and 98%" should be clarified in the case of lysine-acetylated peptides. In view of the comments of reviewer #2 on the limited number of biological replicates for the acetylation data, we asked our Advisory Board for advice. In this case, the recommendation was positive and in favor of publication since the issue is related to the identification of modified peptides and the identifications is seen as credible.

Please merge supplementary text and supplementary figures in a single PDF file that starts with a table of content (see instructions at http://www.nature.com/msb/authors).

Data availability: our attempt to download the data from Tranche failed .Please make sure that the full dataset is available without restriction.

Thank you for submitting this paper to Molecular Systems Biology.

Best wishes,

Editor
Molecular Systems Biology

http://www.nature.com/msb

Referee reports:

Reviewer #1 (Remarks to the Author):

Thank you for sending the revised version of van Noort et al, MSB-11-3089R. I have read the manuscript and am satisfied with the author's responses, particularly with respect to the technical aspects of the dataset. The N-acetyltransferase mutants were a nice addition to the manuscript; I would recommend mentioning these analyses in the abstract along with the phosphorylation mutants.

Reviewer #2 (Remarks to the Author):

The revised manuscript by van Noort et al. entitled "Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium" has been significantly improved when compared to the original version. However, before this manuscript can be accepted, I suggest addressing/revising the following issues:

Figure 1 and corresponding text: It should be highlighted that the authors only have biological repeats for the wild type and the phosphorylation analysis of the PknB mutant, i.e. all other results (phosphorylation of HPrK and PrpC mutants, all acetylation data) are the results of single biological samples (= independent cultures). In the case of the acetylation, there are only technical repeat data for the PknB mutant. I leave it up to the editor to decide whether this is sufficient for MSB or not. Nevertheless, I would request that the limited availability of repeat data is clearly mentioned in the text.

Page 7 ("To assess quality further we randomly selected 11 proteins from the lysine acetylome dataset and independently confirmed nine lysine acetylations by immunostaining with an anti-acetyl-lysine antibody..."): What is with the other two acetylated proteins? Their acetylation could not be confirmed?! Please clarify.

Throughout the text: the authors like to claim comprehensiveness for their analysis. This is a pretty strong claim given the fact that they only used trypsin which results in average in only 43% sequence coverage, i.e. data for more than 50% of each protein are missing. Thus, I would strongly

suggest reconsidering the use of the term 'comprehensive'.

Page 11 ("Matching unmodified peptides..."): This term is not completely clear (is it the unmodified cognate peptide?). Please clarify.

Page 18: Which Orbitrap was used for the initial analysis? Classic? XL? Please specify.

Page 21 ("Nevertheless, to reduce the false positive rate, the spectra of modified peptides with a low Mascot score were manually inspected and modified peptides with a Mascot score lower than 10 were removed from the dataset."): Reading this statement, I got a little bit scared. Does that mean that peptides with Mascot scores close 10 survived the FDR analysis? I have a hard time accepting peptides with such low Mascot score. Please clarify (and mentioned the score cut-offs that were determined for the different datasets and types of peptides). In this context it would be helpful to list the Mascot scores for the identified peptides in the supplementary tables (it should not be required, to download and open the Scaffold output file in order to obtain such essential information).

Figure 3: It is not completely clear to me what the 'white' proteins are, i.e. those that are not kinase, kinase & phosphatase or phosphatase regulated. Please clarify.

Reviewer #3 (Remarks to the Author):

The authors have constructively responded to all my queries. This included producing a substantial amount of new data, notably concerning the influence of acetylation on phosphorylation. This paper should be accepted for publication.

2nd Revision - authors' response   26 January 2012

Reviewer #1 (Remarks to the Author):
Thank you for sending the revised version of van Noort et al, MSB-11-3089R. I have read the manuscript and am satisfied with the author's responses, particularly with respect to the technical aspects of the dataset. The N- acetyltransferase mutants were a nice addition to the manuscript; I would recommend mentioning these analyses in the abstract along with the phosphorylation mutants.

We thank the reviewer for this constructive suggestion and have added one sentence to the abstract on page 2.

Reviewer #2 (Remarks to the Author):
The revised manuscript by van Noort et al. entitled "Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium" has been significantly improved when compared to the original version. However, before this manuscript can be accepted, I suggest addressing/revising the following issues:
Figure 1 and corresponding text: It should be highlighted that the authors only have biological repeats for the wild type and the phosphorylation analysis of the PknB mutant, i.e. all other results (phosphorylation of HPrK and PrpC mutants, all acetylation data) are the results of single biological samples (= independent cultures). In the case of the acetylation, there are only technical repeat data for the PknB mutant. I leave it up to the editor to decide whether this is sufficient for MSB or not. Nevertheless, I would request that the limited availability of repeat data is clearly mentioned in the text.

We are aware of this limitation of our study and now mention this explicitly in the ms text page 6 paragraph 1.
"
Page 7 ("To assess quality further we randomly selected 11 proteins from the lysine acetylome dataset and independently confirmed nine lysine acetylations by immunostaining with an anti-

acetyl-lysine antibody..."): What is with the other two acetylated proteins? Their acetylation could not be confirmed?! Please clarify.

As can be seen in supplementary Figure 2, for two proteins acetylation could not be confirmed by western blot. We cannot completely exclude the possibility of False Positives. However, Western Blot is also not without artefacts. There could be several problems, inaccessibility of the acetylated lysine to the antibody, problems caused by overexpression of the protein, sensitivity of the assay itself. We have now added explicitly the following sentence in paragraph 1 on page 7:
"Reciprocally, deletion of the two putative N-acetyltransferases affects protein phosphorylation, confirming cross- talk between the two PTMs."
"The biological reproducibility for the identification of proteins and phosphopeptides (between the four analyzed mixtures) were 98% and 86%, respectively. The reproducibility for the identification of lysine-acetylated peptides (two mixtures) is 70%. We also measured reproducibility for the quantification as follows: 1) technical replicates with reverse labeling were included for all phosphorylation measurements (all four strains), whereas for lysine acetylation measurements technical duplicates were present in the wild type and PknB mutants, and 2) biological replicates were included for the quantification of proteins and phosphopeptides in the wild type and the PknB mutant (Figure 1).
, two could not be confirmed using this method" ".

Throughout the text: the authors like to claim comprehensiveness for their analysis. This is a pretty strong claim given the fact that they only used trypsin which results in average in only 43% sequence coverage, i.e. data for more than 50% of each protein are missing. Thus, I would strongly suggest reconsidering the use of the term 'comprehensive'.

The reviewer is right to point out that the proteome is not covered to 100%. With current technology our study is among the most comprehensive analyses available. We have toned done the message of comprehensiveness throughout the ms and changed all occurrences.

Page 6 last paragraph: " "The phosphorylation dataset is extensive,"
Page 7 paragraph 2: " "Taken together, the dataset is amongst the most comprehensive analyses"
Page 14 paragraph 2: "Our extensive analysis of protein abundance"

Page 11 ("Matching unmodified peptides..."): This term is not completely clear (is it the unmodified cognate peptide?). Please clarify.

This is indeed the unmodified peptide with the same sequence as the modified peptide. We modified the sentence to make this clear.

Page 11, paragraph 1: " Corresponding unmodified peptides   "

Page 18: Which Orbitrap was used for the initial analysis? Classic? XL? Please specify.

The Orbitrap Classic was used for analysis of the SCX fractions. Page 18 paragraph 3: "LC-LTQ-Orbitrap   Classic".

Page 21 ("Nevertheless, to reduce the false positive rate, the spectra of modified peptides with a low Mascot score were manually inspected and modified peptides with a Mascot score lower than 10 were removed from the dataset."): Reading this statement, I got a little bit scared. Does that mean that peptides with Mascot scores close 10 survived the FDR analysis? I have a hard time accepting peptides with such low Mascot score. Please clarify (and mentioned the score cut-offs that were determined for the different datasets and types of peptides). In this context it would be helpful to list the Mascot scores for the identified peptides in the supplementary tables (it should not be required, to download and open the Scaffold output file in order to obtain such essential information).

For the FDR filtering not only the MASCOT score is taken into account, but also 17 other parameters. MASCOT percolator presents a significant advance over using MASCOT scores alone in terms of sensitivity and accuracy. Indeed some MASCOT scores might be low even though other parameters would give confidence to the identification. The most important of those is the delta ppm

(how precise does the observed peptide mass match the calculated mass). Further confidence in identifications is gained if a modified peptide has been identified more than once and with more than one isotopic label. We have now added these parameters for identifications of modified peptides to the supplementary tables S2 and S3, so that interested researchers can judge the PTMs themselves.

Figure 3: It is not completely clear to me what the 'white' proteins are, i.e. those that are not kinase, kinase & phosphatase or phosphatase regulated. Please clarify.

We are sorry to not have made this completely clear. The phosposites that are in white have not been quantified in all samples, thus we cannot unambiguously tell in which category they belong. For the top sites that are not in a colored box, they are at least kinase regulated but may also be phosphatase regulated, vice versa for the lower sites. We have changed the figure legend to clarify this.
Page 37 paragraph 1: "Phosphosites outside the box represent some kinase-regulated phosphosites that could not be measured in the phosphatase knock-out"

Reviewer #3 (Remarks to the Author):
The authors have constructively responded to all my queries. This included producing a substantial amount of new data, notably concerning the influence of acetylation on phosphorylation. This paper should be accepted for publication.