Dissecting the energy metabolism in Mycoplasma pneumoniae through genome-scale metabolic modeling

Judith A.H. Wodke, Jacek Puchalka, Maria Lluch-Senar, Josep Marcos, Eva Yus, Miguel Godinho, Ricardo Gutierrez-Gallego, Vitor Martins dos Santo, Luis Serrano, Edda Klip, Tobias Maier

Corresponding author: Tobias Maier, CRG Centre de Regulacio Genomica

Review timeline:

- Submission date: 11 June 2012
- Editorial Decision: 13 July 2012
- Revision received: 06 November 2012
- Editorial Decision: 09 January 2013
- Revision received: 04 February 2013
- Accepted: 20 February 2013

Editor: Andrew Hufton / Thomas Lemberger

Transaction Report:

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

1st Editorial Decision 13 July 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who agreed to evaluate your manuscript. As you will see from the reports below, the referees find the topic of your study of potential interest. They raise, however, substantial concerns on your work, which, I am afraid to say, preclude its publication in its present form.

While the reviewers raise a series of important technical concerns, the most substantial issue is the concern raised by Reviewer #1 that the most interesting new biological claims, particularly the mechanistic explanation for the gap between predicted and experimental growth rates, remained weakly supported at present. This reviewer clearly felt that the "ATPase-maintenance" hypothesis required direct experimental validation before this work would be appropriate for publication at Molecular Systems Biology. The other two reviewers, with deep expertise in flux balance analysis, also had some concerns that seem material to this claim, and felt that this prediction and other conclusions needed to be rigorously evaluated for their sensitivity to changes in the biomass composition, and their dependence on the somewhat simplified calculations of proteome and transcriptome production costs. Any revised manuscript will need to address these concerns conclusively, possibly with new supporting experimental data.

The first reviewer also felt that much of the content of the related manuscript (MSB-12-3826) could and should be folded into this manuscript to provide a more complete picture of the model building effort and the supporting experimental data.

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 favorable.

------------------------------------------------------

Referee reports:

Reviewer #1 (Remarks to the Author):

Review of: metabolic modeling and quantitative biomass and energy balancing in Mycoplasma pneumonia (Wodke et al.)

The authors present a new genome-scale metabolic reconstruction of M. pneumoniae, based primarily on a wiring diagram presented in a previous study. They augment this metabolic reconstruction with a number of manual curations based on literature-derived observations and correcting for in silico simulations that don't correspond to actual growth experiments. They then run a number of new growth experiments in which they measure the concentrations of several compounds over time, and they fit the carbohydrate consumption and also fermentative product formation of the cells in batch culture. The reconstruction is constrained at different phases of growth based on the data collected, and various predictions are made about which genes are active or non-active during different phases. One of their principle findings is that actual growth rate is much lower than that predicted by the models (given the substrate uptakes that were measured), and they investigate factors that might contribute. They hypothesize that the strongest factor contributing to the high cellular maintenance term (which fills the gap between the model predicted vs. measured growth rate) is a large need to maintain a proton gradient.

The study is straightforward and appears generally well designed (with a few notable exceptions, as will be described). The methods are not particularly novel, but their combination and application in the manner done in the paper (setting uptake and secretion fluxes based on measured values at different parts of the growth curve in order to learn more about the actual physiological states & conditions of the cells) is noteworthy and could yield valuable insight. This, however, as it stands now, is not in our opinion enough to justify publication of this work in MSB. However, There is one key scientific insight that comes from the analysis that we do feel could make this paper worthy of publication in MSB in our minds: this is the observation of the gap between predicted and measured growth rates, and the subsequent deduction that the most likely factor filling that gap is a high cost to maintain the cellular proton gradient. This key result, however, is not actually validated in the study, and in fact is presented quite misleadingly as if it had been validated (unless we missed or misinterpreted something). This is done in multiple ways:

(1) On page 16, it states "The ATPase was found to be the major energy consumer using about 57 - 80% of the total generated energy in order to maintain a favorable proton gradient across the membrane and the intracellular pH"

(2) Later on the same page, it says "Therefore, at later growth stages most available cellular energy gets diverted towards pH maintenance, eventually leading to growth arrest."

(3) Figure S3F (which, if this effect is proven, should be a key figure in the main text) is slipped into a larger figure (S3) in the supplement which shows **experimentally measured** values for all plots EXCEPT S3F (this one only shows the *calculated* values, as far as we can tell, with an unclear figure legend and no further explanation or reference (and it's actually not clear if this figure is related to the ATPase activity (and thus related to the ATPase story), or ATP generation in the cell - this should be made clear, but either way, it should also be moved to a different figure or explained sufficiently that it is clear whether it is in vitro or in silico data). As the ATPase story is the key and most scientifically impressive insight from the study, this figure and result should be explained more carefully and presented accordingly.

(4) The only experiment done to show that this effect is real, which involves re-buffering the
medium in late growth phase as presented in figure S6, is (a) tucked away in the supplement and (b) is in fact not a validation of the ATPases being dominant in the maintenance cost, it is only suggestive. It is possible, for instance, that low pH simply shuts down cellular processes for other reasons (e.g., signaling pathways or protein deactivation at low pH) -- Additionally, what was the final pH at re-buffering? This should be reported, as it is important for interpretation of whether this effect is significant enough to be believable.

(5) Figure 3C shows the ATPase activity as if it is a measured result, not as if it is a hypothesis. At this point, we do not recommend the paper for acceptance. However, we would support publication if further experiments are done that validate the ATPase-maintenance hypothesis (such as by directly measuring the activity or expression of ATPases).

Here are a list of MAJOR REVISIONS:

(1) Experimentally validate the ATPase story, as explained above.

(2) The figures need considerable work in order to be comprehensible. Figure 1 is interesting looking, but confusing and not comprehensible (e.g., what does it mean to have 'alternative sugar sources' as an output of step 2?). Figure 2 mentions 'moonlighting enzymes', which are never discussed in the paper. Figure 4A perhaps belongs more appropriately in the supplement, unless key features (e.g., pathways) can be highlighted that give biological insight (and its colors are not defined). Figure 5A has no labels and the whole figure is confusing and difficult to interpret - key features that point to a biological insight should be highlighted, or the figure not shown. In Figure 6, some of the gene names are covered. Also, in general, it is meaningless to most readers to see the gene ORF codes - you should list the gene names in all figures that will be in the main text.

(3) Much of the model-building work relies on unpublished data from another submitted manuscript. With so much of the model building effort folded into another paper, it makes the full elucidation of the ATPase story (or another strong biological insight, if experiments prove that hypothesis wrong) ever more critical for suitability of this paper at MSB, because it is difficult to judge how much of the model building (and the work as a whole) is actually unique to this effort.

(4) It is unclear if the authors tried to actually increase flux through the ATPase in the model to compensate for the missing maintenance energy, and if so how this would change the flux results?

(5) When determining which fluxes are up/down regulated between conditions, KOs etc. - how did they consider multiple optima? (sampling? FVA?).

(6) The authors provided no sensitivity analysis of the robustness of the results with respect to the biomass composition (See for instance Feist et al. MSB 2007). This is particularly called for due to the inclusion of several end-products in "small arbitrary quantities" (lines 140-143).

(7) It is unclear how the authors determined the #s in S5 which say 'in silico result'; how did they quantify these? There is no clear explanation given.

And some MINOR REVISIONS:

(1) Usually the convention for naming models is "i_initialsofbuilder_#ofgenes". Please check that you've followed the convention here, as you list the # of *enzymes*, not genes.

(2) Please elucidate the methods used to constrain the models based on flux constraints more clearly (mentioned on page 6, "As a general strategy, we minimized the possible number of flux...")

(3) Please check the titles of the supplementary tables in the excel sheet: at least one of them (table S3) is messed up.

(4) What insight (if any) to draw from the clustering analysis (figure 5) is unclear - please look over the explanation and tighten the biological explanation.

(6) Rolfsson et al. 2011 is not the correct reference for the human metabolic reconstruction. This should be Duarte et al. PNAS 2007.

(7) There is an incoherent sentence on lines 163-165 (is it a misplaced comma?).

(8) line 438: a pval should be given for the enrichment, properly corrected for multiple hypotheses (since the authors likely checked all pathways).

(9) Table 2: The characters are not displayed correctly in the pdf and appear as empty boxes.

Reviewer #2 (Remarks to the Author):

Wodke et al. present a comprehensive review of the flux-balanced metabolic network reconstruction process for M. pneumoniae, an interesting model species due to its minimal genome size. Their work builds on the manually curated network by Yus et al. (Science, 2009) through various stages; first by improving the annotation of reaction reversibility and the addition transport reactions, second by incorporating a detailed experimentally determined biomass composition, and third by comparing predictions at various reconstruction stages with experimental results. While iterative model refinement based on experimental data is the standard procedure for the reconstruction of constraint-based metabolic models and it has been described many times, the authors describe three specific examples where such process leads to improved reaction and annotations in the network. The end result is a model with high predictive power in terms of the effects of gene knockouts under the conditions assayed. The work relies heavily on the vast amount of experimental data available for this species and the group's own results from the accompanying paper. Applying constraints to metabolite exchange fluxes and growth rates from experimental data at different time points, along with detailed biomass measurements, allowed the authors to analyze two processes in detail: ATP energy utilization resulting in the discovery of a large contribution to the in non-growth-associated maintenance-, and a global picture of the dynamics of metabolic fluxes during exponential growth. This portion of the work is an interesting demonstration of the flux balance analysis framework providing insight into relevant metabolic processes.

I think the manuscript is worthy of publication and have a few comments presented below.

1. As described above, iterative metabolic network reconstruction based on model predictions and their experimental validation is a well described process. While the authors use this process to describe interesting features of the M. pneumoniae metabolism, such chronological description takes a lot of space and distracts from the most interesting findings. I would suggest that technical details of the reconstruction process are minimized and moved to the methods or the supplementary materials.

2. While it is interesting that the combination of experimental data with the high-quality metabolic model allows a dynamic analysis of metabolic flux changes, it is important to understand how these results vary (depend) on alternative solutions for a given FBA problem. The authors should investigate the variability of the flux values once the experimental constraints are put in place. For example, does the FBA degeneracy significantly affect the energy (ATP) partition to different functions? Are the predicted flux changes mostly robust to possible alternative optima?

3. There is no description of the colors in figure 4A. Are these the same as in figure 5? A color scale should be provided for this figure as well.

4. What are the different slices in figure 5A? The meaning of each color should be described in the figure as it is not obvious even from the main text.

Reviewer #3 (Remarks to the Author):

Maier et al. report the reconstruction of the metabolism of M. pneumoniae from its genome annotation and a published metabolic map. This reconstruction was extended and refined using experimental data, which this group submitted as a parallel manuscript to Molecular Systems Biology. This data revealed contradictions with initial model predictions, which led subsequently to re-annotation of ORFs. The manuscript benefits from the quantitative data that this group has
generated in the past. In particular, the energy balance analysis is very interesting. Overall, this is a well-written manuscript and the presented analysis will be interesting to the systems biology community.

The reviewer has the following concerns, which the authors should address to make this manuscript suitable for publication in Molecular Systems Biology.

- In the abstract, the authors state that they "analyzed ... the metabolic network ... in unprecedented detail, integrating different large-scale data sets". This is an overstatement of the presented work. The authors performed characterizations of the metabolic network, which are more or less standard in the field (e.g., biomass production under different medium environments, single and double deletion analysis, etc.). To the reviewer's knowledge, the energy balance analysis has not been performed as such before. In terms of large-scale data integration, the authors used overall protein production information and mRNA/protein halflife data to set constraints on some model reactions. However, 'integration of large-scale data' implies the imposition of constraints on many network reactions based on these data. Again, setting constraints based on experimental data as well as the integration of omics data as modeling constraints has been done before by others. The authors are encouraged to tone-down this statement.

- The authors base their energy balance analysis on the assumption that the model is missing ATP consuming reactions and thus the in silico biomass rate is much higher than experimentally reported. This is probably the case, but it is also possible that some of the higher computed biomass production is due missing biomass constituents from the biomass equation or incorrect fractional contribution of the constitutents (e.g., for the cofactors). A sensitivity analysis could be insightful.

- The authors report the computed percentage of energy cost for protein production etc during exponential growth, however, cellular composition changes with growth rate and medium composition, and thus, transcriptome and proteome composition will change and consequently, the associated energy requirements. However, the authors represent the energy cost and precursor requirements for proteome and transcriptome with one static reaction. An alternative would be to formulate condition-specific reactions. At the minimum, the authors should consider to discuss this limitation of their analysis.

- In the quantitative and dynamic growth simulations section, the authors state that they applied a hybrid, quasi-dynamic modeling approach. If the reviewer understands the approach correctly, then the authors back calculated uptake rates from measured concentration changes of compounds in the medium, and subsequently applied these uptake rates as model constraints. Again, this is a standard procedure to obtain and apply constraints to FBA models and to reproduce growth performance under different growth conditions. The discussion section should be modified accordingly as well.

- Furthermore, in the quantitative and dynamic growth simulations section, the authors discuss in great detail the number of reactions that change or do not change with the biomass production in the simulation based on their FBA results. However, given the information provided in the method section, the authors performed a standard FBA optimization, and thus cannot exclude the existence of alternative flux distributions having the same optimal value. Thus, analyzing only one flux distribution provides only limited insight and results cannot be generalized as presented in the manuscript. The same is the true for the analysis of fluxes associated with single gene deletion events (section Prediction of gene essentiality).

- The authors observed a high correlation between flux direction changes and protein abundance changes. It would be also interesting to compare the flux direction changes with transcript abundance changes.

- The authors state that "A comparison of maintenance energy costs between M. pneumoniae and E. coli revealed fundamental differences in their energy sink reactions, suggesting individual and characteristic energy expense profiles for different bacteria." This statement implies that the authors did such analysis but it is not reported in the result section nor is there any reference to Table S12 (which contains the comparison) in this manuscript.

- The authors state twice in the method section that RNA and DNA elongation reactions cannot be represented within a stoichiometric matrix. This is incorrect. In principle every biochemical transformation can be represented as a stoichiometric reaction, thus also polymerization reactions. In fact, stoichiometric matrices have been generated and analyzed using FBA-based methods for signaling (e.g. PMID: 15240442, PMID: 22363624), transcription and translation (PMID: 19282977), and regulation (e.g. PMID: 19503608). The authors should correct their statements accordingly.
- Line 68 - This is not the correct reference. It should be PMID: 17267599.
- The reviewer suggests including this sentence from the SI to the main text (Line 392): "In case of a contradiction between prediction and transposon study, we screened an M. pneumoniae transposon library for respective mutants (figure S5). Thus, we could confirm five model predictions that were not supported by the M. genitalium study alone, raising the prediction accuracy to 91.6% and the specificity to 97.96%.” for clarity.
- Line 578 - "atoms” should be "elements"
- Line 621 - "The FBA is” should be replaced with "The FBA solution is"
- For the calculation of the ATP flux, which formula was used? What is the assumed cellular weight? Which unit was used for the FBA calculation? The reviewer could not find units reported for the flux calculation in the SI tables either.
- Line 619 - The relationship between doubling time and growth rate is tdoub = ln(2)/ov.
- Line 678 - Is there a section missing?

1st Revision - authors’ response 06 November 2012

We would like to thanks the reviewers for their insightful comments on our manuscript. For this revised version, we addressed all points raised and documented the changes in detail below, in line with the comments of the referees.

Major improvements of the manuscript include:
1. We experimentally confirmed the energy maintenance hypothesis and added a new figure
4.
2. We integrated large sections of the Maier et al. manuscript (MSB-12-3826) into this version (i.e. 13C tracer experiments, experiments relevant for the biomass composition)
3. We extended our in silico analysis of the data, now including a sensitivity analysis as well as a flux variability analysis.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Review of: metabolic modeling and quantitative biomass and energy balancing in Mycoplasma pneumonia (Wodke et al.)

The authors present a new genome-scale metabolic reconstruction of M. pneumoniae, based primarily on a wiring diagram presented in a previous study. They augment this metabolic reconstruction with a number of manual curations based on literature-derived observations and correcting for in silico simulations that don’t correspond to actual growth experiments. They then run a number of new growth experiments in which they measure the concentrations of several compounds over time, and they fit the carbohydrate consumption and also fermentative product formation of the cells in batch culture. The reconstruction is constrained at different phases of growth based on the data collected, and various predictions are made about which genes are active or non-active during different phases. One of their principle findings is that actual growth rate is much lower than that predicted by the models (given the substrate uptakes that were measured), and they investigate factors that might contribute. They hypothesize that the strongest factor contributing to the high cellular maintenance term (which fills the gap between the model predicted vs. measured growth rate) is a large need to maintain a proton gradient.

The study is straightforward and appears generally well designed (with a few notable exceptions, as will be described). The methods are not particularly novel, but their combination and application in the manner done in the paper (setting uptake and secretion fluxes based on measured values at different parts of the growth curve in order to learn more about the actual physiological states & conditions of the cells) is noteworthy and could yield valuable insight. This, however, as it stands now, is not in our opinion enough to justify publication of this work in MSB. However, There is one key scientific insight that comes from the analysis that we do feel could make this paper worthy of publication in MSB in our minds: this is the observation of the gap between predicted and measured
growth rates, and the subsequent deduction that the most likely factor filling that gap is a high cost to maintain the cellular proton gradient. This key result, however, is not actually validated in the study, and in fact is presented quite misleadingly as if it had been validated (unless we missed or misinterpreted something). This is done in multiple ways:

We agree that this is a key finding of the study. We carried out additional experiments showing that more energy is diverted to maintenance tasks when the cellular proton gradient is challenged by artificially changing the medium pH. The results confirm our energy balance calculation. They are presented in a new figure 4. We added a corresponding section to the main text of the manuscript. More details are given in the corresponding sections below.

(1) On page 16, it states "The ATPase was found to be the major energy consumer using about 57 - 80% of the total generated energy in order to maintain a favorable proton gradient across the membrane and the intracellular pH"

To distinguish between experimental and theoretical results we now describe the putative ATP consumption by the ATPase as follows:

‘The combination of in silico calculations taking into account the ATPase protein copy number in M. pneumoniae and catalytic rates from other organisms reported in literature, as well as the in vivo analysis of growth in medium with different pH suggest that the ATPase uses about 57 – 80% (growth stage depending) of the total generated energy in order to maintain a favorable proton gradient across the membrane and the intracellular pH.’ (page 18, lines 548ff)

(2) Later on the same page, it says "Therefore, at later growth stages most available cellular energy gets diverted towards pH maintenance, eventually leading to growth arrest."

Changed to ‘Therefore, in consistency with the abundance of the core components of the ATPase protein complex (Maier et al, 2011) and our experimental results for M. pneumoniae growth under pH stress (figure 4A-B), we propose that at later growth stages most of the available energy is diverted towards cellular maintenance, eventually leading to growth arrest, and that the ATPase is the major energy sink of M. pneumoniae in batch culture growth.’ (page 19, lines 559ff).

(3) Figure S3F (which, if this effect is proven, should be a key figure in the main text) is slipped into a larger figure (S3) in the supplement which shows **experimentally measured** values for all plots EXCEPT S3F (this one only shows the *calculated* values, as far as we can tell, with an unclear figure legend and no further explanation or reference (and it's actually not clear if this figure is related to the ATPase activity (and thus related to the ATPase story), or ATP generation in the cell - this should be made clear, but either way, it should also be moved to a different figure or explained sufficiently that it is clear whether it is in vitro or in silico data). As the ATPase story is the key and most scientifically impressive insight from the study, this figure and result should be explained more carefully and presented accordingly.

Figure S3F is now part of figure 3, which in addition shows the comparison of in silico and in vivo doubling times (A), the energy usage proposed by the model (B), the defined and calculated expenses (C). The fitting for the maintenance energy is included in figure 3B, where ATP usage by biomass production and maintenance functions is shown. Also we included an additional main figure (now figure 4) for the experimental results supporting the energy predictions of the model. This figure includes the new results showing that M. pneumoniae at low pH values takes up glucose in comparable amounts but do not grow (protein increase) as fast as at higher pH, as expected if the ATPase is a major energy sink in batch culture growth (figure 4A-B), and the pH-re-buffering experiment already included in the first submission (figure 4C).

(4) The only experiment done to show that this effect is real, which involves re-buffering the medium in late growth phase as presented in figure S6, is (a) tucked away in the supplement and (b) is in fact not a validation of the ATPases being dominant in the maintenance cost, it is only
suggestive. It is possible, for instance, that low pH simply shuts down cellular processes for other reasons (e.g., signaling pathways or protein deactivation at low pH) — Additionally, what was the final pH at re-buffering? This should be reported, as it is important for interpretation of whether this effect is significant enough to be believable.

We have moved the old figure S6 to the new main figure 4 (figure 4C). In the figure we added the pH curve for normal batch culture growth and added information about the re-buffering time and the final pH at re-buffering (figure 4A, pH6 re-buffered to pH7.7).

We conducted additional experiments comparing growth and glucose consumption of *M. pneumoniae* in fresh medium at different pH (figure 4A-B). We measured biomass production at protein level, and consumption of glucose as a function of produced energy (figure 4A). These experiments show that biomass production is decoupled from energy metabolism, particularly at low pH, where no increase of biomass can be observed while cells keep metabolizing glucose to carry out maintenance tasks.

(5) Figure 3C shows the ATPase activity as if it is a measured result, not as if it is a hypothesis. At this point, we do not recommend the paper for acceptance. However, we would support publication if further experiments are done that validate the ATPase-maintenance hypothesis (such as by directly measuring the activity or expression of ATPases).

Figure 3C combines the *in silico* determined expenses with the results from the upper boundary calculations for i) DNA maintenance, ii) protein-folding, iii) post-translational modifications, and iv) the ATPase in order to show their proposed respective contributions to the cellular energy balance. All these calculations were based on experimentally measured protein abundances measured in *M. pneumoniae* (Maier et al., 2011). For the ATPase the beta subunit (smallest experimental error for biological replica) was used to determine the copy number of the ATPase in *M. pneumoniae* as well as catalytic rates reported in the literature. As mentioned above we have now conducted an experiment to show how decreasing pH results in increased decoupling of energy production and growth (biomass production), supporting that the ATPase as a major player in energy use (figure 4A-B).

Here are a list of MAJOR REVISIONS:

(1) Experimentally validate the ATPase story, as explained above.

We conducted an experiment, challenging the ATPase function by artificially changing the medium pH in order to determine the correlation between external pH and energy consumption by the ATPase as well as biomass production (figure 4A-B).

(2) The figures need considerable work in order to be comprehensible. Figure 1 is interesting looking, but confusing and not comprehensible (e.g., what does it mean to have 'alternative sugar sources' as an output of step 2?).

We improved figure 1 in order to display the workflow more clearly. We also improved the readability of all figures by adjusting fonts, styles, and legends. Furthermore, we corrected the captions of all figures and added more information to describe the respective results shown (see below).

Figure 2 mentions 'moonlighting enzymes', which are never discussed in the paper.

Moonlighting enzymes are multifunctional proteins as defined in Yus *et al*, 2009, where also the first metabolic map has been published. However, since there is no direct relation to this manuscript, we removed this eventually misleading information from figure 2.

Figure 4A perhaps belongs more appropriately in the supplement, unless key features (e.g., pathways) can be highlighted that give biological insight (and its colors are not defined).
We included information on pathways into this figure (now figure 5A) clustering the reactions according to their pathways and added the missing legend.

Figure 5A has no labels and the whole figure is confusing and difficult to interpret - key features that point to a biological insight should be highlighted, or the figure not shown.

We included the missing labels (now figure 7A) and describe them in the figure caption. Furthermore, we changed figure 7B, now just showing the overall percentages of flux changes. The biological insight obtained from figure 7 is now more clearly discussed in the main text (section ‘Prediction of gene essentiality’, 1st paragraph, page 15, lines 446ff & 3rd paragraph, page 16, lines 469ff).

In Figure 6, some of the gene names are covered. Also, in general, it is meaningless to most readers to see the gene ORF codes - you should list the gene names in all figures that will be in the main text.

We agree with the reviewer, however, in *M. pneumoniae* quite a large fraction of genes (38%) has only a MPN ID but no gene name. To allow consistency in the figures and tables, we used the MPN IDs throughout the entire manuscript. For an easy identification of the respective gene names, we included an additional table relating MPN IDs and gene names (supplementary table S14). In addition, for better a visibility of the IDs in the figure of the double knock-out results (now figure 8), we increased the fonts and separated the different genes a bit more to avoid overlapping names.

(3) Much of the model-building work relies on unpublished data from another submitted manuscript. With so much of the model building effort folded into another paper, it makes the full elucidation of the ATPase story (or another strong biological insight, if experiments prove that hypothesis wrong) ever more critical for suitability of this paper at MSB, because it is difficult to judge how much of the model building (and the work as a whole) is actually unique to this effort.

The major fraction of the new experimental data used for the model building process and the correction of the network structure are now included in this manuscript. In detail, we included the information about the fatty acid composition (basis for 20% of the biomass; SI, page 3f) and the 13C-glucose tracer experiments (validating the network structure and the flux predictions, main text page 5, lines 135ff, page 14, lines 398ff, & SI, page 10). The quantification of the remaining free cellular metabolites only account for 1.5% of the total biomass and hence does not play a significant role in the model construction process.

(4) It is unclear if the authors tried to actually increase flux through the ATPase in the model to compensate for the missing maintenance energy, and if so how this would change the flux results?

Indeed we simulated increased flux through the ATPase reaction and found that it results in increased doubling times and reduced maintenance costs due to the additional ATP consumption associated with the proton transport, further supporting our energy predictions.

(5) When determining which fluxes are up/down regulated between conditions, KOIs etc. - how did they consider multiple optima? (sampling? FVA?).

We performed a flux variability analysis and describe the respective results in our manuscript (page 13, lines 366ff, table S10).

(6) The authors provided no sensitivity analysis of the robustness of the results with respect to the biomass composition (See for instance Feist et al. MSB 2007). This is particularly called for due to the inclusion of several end-products in "small arbitrary quantities" (lines 140-143).

We carried out a sensitivity analysis for different biomass components to analyze the impact of the cellular macromolecules (protein, lipids, mRNA) and reaction cofactors
on the *in silico* growth rate and, alternatively, the maintenance energy. To this end we varied the protein, mRNA, and lipids fractions according to their contributions to mycoplasma cells reported by Razin et al., 1963 (figure S4). In addition, we changed the abundances of all cofactors from 0.1 to 10-fold with respect to the biomass composition defined for *M. pneumoniae* and included a paragraph in the results section (page 6, lines 159ff, SI, page 7f, table S4).

(7) It is unclear how the authors determined the #'s in S5 which say 'in silico result'; how did they quantify these? There is no clear explanation given.

The numbers in table S5 with source ‘*in silico* result’ have been quantified based on the biochemical reactions involved in the synthesis of the respective cellular macromolecules. To this end fluxes and stoichiometric information on ATP have been integrated. For the total ATP produced the ATP produced net in glycolysis + the ATP produced net in arginine fermentation have been summed up. We now indicate this more clearly in the legend of table S5.

And some MINOR REVISIONS:

(1) Usually the convention for naming models is "i_initialsofbuilder_#ofgenes". Please check that you've followed the convention here, as you list the # of *enzymes*, not genes.

We changed this.

(2) Please elucidate the methods used to constrain the models based on flux constraints more clearly (mentioned on page 6, "As a general strategy, we minimized the possible number of flux..")

We included a more detailed description of the different defined constraints in the SI, section ‘Definition of Flux Constraints’ (SI, page 8f).

(3) Please check the titles of the supplementary tables in the excel sheet: at least one of them (table S3) is messed up.

The names of tables S2 and S3 have been changed (they had been switched).

(4) What insight (if any) to draw from the clustering analysis (figure 5) is unclear - please look over the explanation and tighten the biological explanation.

We removed the clustering from the figure and the text and also changed the respective figure (now figure 7B) further by replacing the ‘heat map’ for the flux changes by a pie plot showing the fractions of all reactions that change according to the defined categories.


We fixed this.

(6) Rolffsson et al. 2011 is not the correct reference for the human metabolic reconstruction. This should be Duarte et al. PNAS 2007.

The reference of Duarte *et al.*, 2007 has been added in the respective section of the text.

(7) There is an incoherent sentence on lines 163-165 (is it a misplaced comma?).

Yes, the comma was wrong.

(8) line 438: a pval should be given for the enrichment, properly corrected for multiple hypotheses (since the authors likely checked all pathways).
We calculated p-values for the 0-hypothesis that the genes of certain pathways are not enriched amongst those involved in synthetic lethal interactions (as compared to all analyzed metabolic genes) and included them in the text (we used a threshold of 0.001 to reject the 0-hypothesis). (page 17, line 503) However, since the subset of genes involved in synthetic lethal interactions is very small to properly apply statistics (only 27 genes in total, belonging to 5 different pathways), we did not correct the p-values further for multiple hypothesis but only calculated a p-value for each pathway associated to genes involved in synthetic lethal interactions.

(9) Table 2: The characters are not displayed correctly in the pdf and appear as empty boxes.

This was a formatting error during the submission process, we hope the characters are now displayed correctly.

Reviewer #2 (Remarks to the Author):

Wodke et al. present a comprehensive review of the flux-balanced metabolic network reconstruction process for M. pneumoniae, an interesting model species due to its minimal genome size. Their work builds on the manually curated network by Yus et al. (Science, 2009) through various stages; first by improving the annotation of reaction reversibility and the addition transport reactions, second by incorporating a detailed experimentally determined biomass composition, and third by comparing predictions at various reconstruction stages with experimental results. While iterative model refinement based on experimental data is the standard procedure for the reconstruction of constraint-based metabolic models and it has been described many times, the authors describe three specific examples where such process leads to improved reaction and annotations in the network. The end result is a model with high predictive power in terms of the effects of gene knockouts under the conditions assayed. The work relies heavily on the vast amount of experimental data available for this species and the group's own results from the accompanying paper. Applying constraints to metabolite exchange fluxes and growth rates from experimental data at different time points, along with detailed biomass measurements, allowed the authors to analyze two processes in detail: ATP energy utilization - resulting in the discovery of a large contribution to the non-growth-associated maintenance-, and a global picture of the dynamics of metabolic fluxes during exponential growth. This portion of the work is an interesting demonstration of the flux balance analysis framework providing insight into relevant metabolic processes.

I think the manuscript is worthy of publication and have a few comments presented below.

1. As described above, iterative metabolic network reconstruction based on model predictions and their experimental validation is a well described process. While the authors use this process to describe interesting features of the M. pneumoniae metabolism, such chronological description takes a lot of space and distracts from the most interesting findings. I would suggest that technical details of the reconstruction process are minimized and moved to the methods or the supplementary materials.

   We agree with the reviewer that the focus of the paper should be set on the biological findings and tried to move even more technical details to the supplement, for example the details of the network refinement process. However, we still include a description of the basic reconstruction process in the main text since it aids understanding, and influences the obtained results as well as the drawn conclusions.

2. While it is interesting that the combination of experimental data with the high-quality metabolic model allows a dynamic analysis of metabolic flux changes, it is important to understand how these results vary (depend) on alternative solutions for a given FBA problem. The authors should investigate the variability of the flux values once the experimental constraints are put in place. For example, does the FBA degeneracy significantly affect the energy (ATP) partition to different functions? Are the predicted flux changes mostly robust to possible alternative optima?

   We carried out a flux variability analysis (FVA). As expected from the linearity of the network and the absence of many alternative metabolic routes known from other
organisms, we found that only a very limited number of optima exists in case that more than one is found for a specified set of constraints and most of the reactions that allow variable flux are associated either directly or indirectly (e.g. using NTs as cofactors) associated to nucleotide metabolism. We describe our FVA results on page 13, lines 366ff (table S10).

3. There is no description of the colors in figure 4A. Are these the same as in figure 5? A color scale should be provided for this figure as well.

   We apologize, a color legend has now been included in the figure (now figure 5A) as well as additional information on pathways.

4. What are the different slices in figure 5A? The meaning of each color should be described in the figure as it is not obvious even from the main text.

   We apologize and included the respective information in the figure and the legend (now figure 7A).

Reviewer #3 (Remarks to the Author):

Maier et al. report the reconstruction of the metabolism of *M. pneumoniae* from its genome annotation and a published metabolic map. This reconstruction was extended and refined using experimental data, which this group submitted as a parallel manuscript to Molecular Systems Biology. This data revealed contradictions with initial model predictions, which led subsequently to re-annotation of ORFs. The manuscript benefits from the quantitative data that this group has generated in the past. In particular, the energy balance analysis is very interesting. Overall, this is a well-written manuscript and the presented analysis will be interesting to the systems biology community.

The reviewer has the following concerns, which the authors should address to make this manuscript suitable for publication in Molecular Systems Biology.

- In the abstract, the authors state that they "analyzed ... the metabolic network ... in unprecedented detail, integrating different large-scale data sets". This is an overstatement of the presented work. The authors performed characterizations of the metabolic network, which are more or less standard in the field (e.g., biomass production under different medium environments, single and double deletion analysis, etc.). To the reviewer's knowledge, the energy balance analysis has not been performed as such before. In terms of large-scale data integration, the authors used overall protein production information and mRNA/protein halflife data to set constraints on some model reactions. However, 'integration of large-scale data' implies the imposition of constraints on many network reactions based on these data. Again, setting constraints based on experimental data as well as the integration of omics data as modeling constraints has been done before by others. The authors are encouraged to tone-down this statement.

   We thank the reviewer for his opinion and toned down the statement. Now it reads: “Here we analyzed and characterized the metabolic network of *M. pneumoniae* in high detail, integrating data from different -omics analyses into a constraint-based model.” However, to clarify this further, when writing “…in unprecedented detail…” we were not referring to the characterization of metabolic networks in general but to that of *M. pneumoniae*, which indeed had not been analyzed to this extent before.

- The authors base their energy balance analysis on the assumption that the model is missing ATP consuming reactions and thus the in silico biomass rate is much higher than experimentally reported. This is probably the case, but it is also possible that some of the higher computed biomass production is due missing biomass constituents from the biomass equation or incorrect fractional contribution of the constituents (e.g., for the cofactors). A sensitivity analysis could be insightful.

   We conducted a sensitivity analysis in order to test the influence of the qualitatively included cofactors on the total growth rate and found this impact to be negligible.
We included a respective section in the SI and refer to it in the Model generation section (page 6, lines 159ff).

- The authors report the computed percentage of energy cost for protein production etc during exponential growth, however, cellular composition changes with growth rate and medium composition, and thus, transcriptome and proteome composition will change and consequently, the associated energy requirements. However, the authors represent the energy cost and precursor requirements for proteome and transcriptome with one static reaction. An alternative would be to formulate condition-specific reactions. At the minimum, the authors should consider to discuss this limitation of their analysis.

We are grateful for this comment and we tried to explain our definition of the biomass more clearly (several rules applied are listed in the supplementary material). While we completely agree that the overall composition of proteome and RNA change along with changing environmental conditions and growth time, it has been shown that in *M. pneumoniae* neither the overall amount of RNAs nor proteins change significantly during the exponential growth phase (i.e. from 24 – 60 hours after inoculation) (Güell *et al.*, 2011, Maier *et al.*, 2011, Yus *et al.*, 2012). For example, for the total number of amino acids bound in the proteome a change of only about 10% is observed along a four days batch culture growth experiment, which is below the error involved in protein quantification by mass spectrometry (see figure below). Since energetically there is no difference in taking up one or another amino acid or one or the other nucleobase, for simplicity reasons we decided to not use condition-specific reactions.

Further supporting the low variability of the proteome during the exponential growth phase, we show here a histogram for the abundances of amino acids in the proteome for the different days representing the different growth phases of *M. pneumoniae* in batch culture growth: lag phase (day0), exponential growth phase (day2), and stationary phase (day4) (based on data from Maier *et al.*, 2011):

- In the quantitative and dynamic growth simulations section, the authors state that they applied a hybrid, quasi-dynamic modeling approach. If the reviewer understands the approach correctly, then the authors back calculated uptake rates from measured concentration changes of compounds in the
medium, and subsequently applied these uptake rates as model constraints. Again, this is a standard procedure to obtain and apply constraints to FBA models and to reproduce growth performance under different growth conditions. The discussion section should be modified accordingly as well.

We toned down the title of the section and also the respective sections in the results and discussion to not raise wrong expectations. However, we are not aware of previous works where back calculations from experimental data are combined with in silico determined maintenance costs, thus allowing to access discrete points in time for which no explicit data is available.

- Furthermore, in the quantitative and dynamic growth simulations section, the authors discuss in great detail the number of reactions that change or do not change with the biomass production in the simulation based on their FBA results. However, given the information provided in the method section, the authors performed a standard FBA optimization, and thus cannot exclude the existence of alternative flux distributions having the same optimal value. Thus, analyzing only one flux distribution provides only limited insight and results cannot be generalized as presented in the manuscript. The same is the true for the analysis of fluxes associated with single gene deletion events (section Prediction of gene essentiality).

We tested for multiple optima but found that based on the linearity of the network and the few branching metabolites the existence of alternative (multiple) optima is quite limited and mainly associated to nucleotide metabolism (page 13, lines 366ff). Based on this low variability (table S10) and the fact that not all possible combinations of changed fluxes would also lead to optimal solutions, we conclude that the analysis of the reproducibly predicted optimal flux distributions as well under normal growth conditions as in knock-out simulations (less reactions in the network normally do not lead to higher variability) provides valid insight into the metabolic behavior of M. pneumoniae and that alternative optima only will include minor changes to the overall flux distributions. Also, we discuss the production of dADP and dGDP in nucleotide metabolism as an example for energetically equal alternative routes in Mpn of which only a few exist (page 13, lines 369ff).

- The authors observed a high correlation between flux direction changes and protein abundance changes. It would be also interesting to compare the flux direction changes with transcript abundance changes.

We compared our flux changes directly to protein changes since it had been shown previously that in M. pneumoniae mRNA and protein levels do only correlate moderately (pearson coefficient 0.53, Maier et al, 2011).

- The authors state that "A comparison of maintenance energy costs between M. pneumoniae and E. coli revealed fundamental differences in their energy sink reactions, suggesting individual and characteristic energy expense profiles for different bacteria." This statement implies that the authors did such analysis but it is not reported in the result section nor is there any reference to Table S12 (which contains the comparison) in this manuscript.

We now report the results of the comparative analysis (now table S7) in the correct section of the manuscript. (section Cellular Energy Balance, page 11, lines 315ff) as well as in the discussion (page 19, lines 569ff).

- The authors state twice in the method section that RNA and DNA elongation reactions cannot be represented within a stoichiometric matrix. This is incorrect. In principle every biochemical transformation can be represented as a stoichiometric reaction, thus also polymerization reactions. In fact, stoichiometric matrices have been generated and analyzed using FBA-based methods for signaling (e.g. PMID: 15240442, PMID: 22363624), transcription and translation (PMID: 19282977), and regulation (e.g. PMID: 19503608). The authors should correct their statements accordingly.

We corrected the statements about the inability of representing elongation reactions with a stoichiometric matrix to “… to cope with reactions that can be represented in
stoichiometric models only with considerable impact on the model complexity, such as for example RNA and DNA elongation reactions (see below).” (page 22, lines 664ff) and “To overcome missing experimental information and to restrict the complexity of the model, a number of approximations have been made which are fully listed in the SI.” (page 23, line 702ff).

- Line 68 - This is not the correct reference. It should be PMID: 17267599.

  We added the original reference for the human metabolic model: Duarte et al, 2007.

- The reviewer suggests including this sentence from the SI to the main text (Line 392): "In case of a contradiction between prediction and transposon study, we screened an M. pneumoniae transposon library for respective mutants (figure S5). Thus, we could confirm five model predictions that were not supported by the M. genitalium study alone, raising the prediction accuracy to 91.6% and the specificity to 97.96%.” for clarity.

  As suggested, we moved this sentence to the main text for better understanding (page 16, lines 460ff) and rewrote the SI accordingly.

- Line 578 - "atoms" should be "elements"

  Thanks, we changed this.

- Line 621 - "The FBA is" should be replaced with "The FBA solution is"

  We agree and did so.

- For the calculation of the ATP flux, which formula was used? What is the assumed cellular weight? Which unit was used for the FBA calculation? The reviewer could not find units reported for the flux calculation in the SI tables either.

  In the used modeling platform ToBiN, all fluxes are given in mmol*(g(cells))^{-1}*h^{-1}. For the calculation of the ATP flux we added up the net ATP produced in glycolysis and arginine fermentation: \( \text{flux(production(lactate))}*2 + \text{flux(production(acetate))}*4 + \text{flux(production(ornithine))} = \text{flux(production(ATP))} \). We included this additional information in material and methods, Energy Calculations (new methods subsection, page 24, lines 735ff).

- Line 619 - The relationship between doubling time and growth rate is \( t_{doub} = \ln(2)/ov \).

  We agree with the reviewer that this is the common formula to relate the objective value to the doubling time in an exponentially increasing population. However, in the used modeling platform the number of cells is not growing but kept constant at 1 g of cells. In this case the relation between doubling time and objective value of the FBA is as stated in our paper – \( t_{doub} = 1/ov \). We added this information to the respective methods section, Growth Simulations (page 24, lines 716ff).

- Line 678 - Is there a section missing?

  Sorry, when formatting the manuscript a wrong heading was placed there.
ideally as soon as possible. If we do not receive the revised manuscript within this time period, the file might be closed and any subsequent resubmission would be treated as a new manuscript. Please use the Manuscript Number (above) in all correspondence.

Thank you for submitting this paper to Molecular Systems Biology.

Referee reports:

Reviewer #1 (Remarks to the Author):

We feel that the authors have sufficiently addressed our concerns from the last revision.

One anomaly in the paper which must be addressed:

In Figure 4, how can one of the bars be negative if they represent protein production?

In addition, we have a few minor comments:

1. Lines 188-190: I there a misplaced comma before "verified"? What does the sentence mean? How did the author confirm "a wide range of experimental data on mRNA expression and for the design of a defined medium"? A reference to supplementary material specifying the tests that were run would be appropriate.
2. Lines 241-243: how did the authors incorporate a constraint on the "total protein increase"? If this is not a standard method - do the results reproduce without it?
3. Lines 317-322: the authors discuss their GAM/NGAM ratio vis-a-vis the ratios in the model of Feist et al. However, GAM/NGAM estimations varied considerably between different revisions of the E. coli model. This is in part due to the the effects of the different experimental settings from which data was gleaned. See the discussion in the reviewing process of the recent Orth et al. paper in MSB 2011: http://www.nature.com/msb/journal/v7/n1/extref/msb201165-s4.pdf.
   We think that:
   a) the considerable variance in those estimations should be mentioned in the discussion, and
   b) the discussion should consider the effect of the experimental settings on the result concerning NGAM/GAM ratio in M. pneumoniae
4. Lines 325-327: "It is important to note that even during the exponential growth phase 326 78% to 89% of the available energy is not directly used for the production of cell building 327 blocks but for cellular homeostasis. --- Is this sentence referring again to the M. genitalium study of Karr et al.? It is not clear from the context because of the interceding discussion of M. mobile. In addition - is this an in-silico result? If so, it should be explicitly said.
5. Lines 343-344: The authors write that they fitted a logarithmic curve, but doesn't the figure show a linear fitting?
6. Lines 346-347: what are the "respectively determined constraints" for the simulated rich medium? A supplementary note is in order here.
7. Lines 346-354: how did the authors calculate the difference between flux distributions? FVA? difference between representative FBA solutions. If the latter is true, an explicit note should be made that the multiple optima are of no concern in this model, as the authors explain in the cover letter.
8. Line 422: the comma after excluded is redundant
9. Lines 471 - 475: again, is this a FVA analysis?
10. Line 744: the sentence seems incomplete or grammatically-incorrect
11. Please fix the spelling mistakes in Fig. S6 (enzymes, enzymes)
12. One of the subfigures of Fig S7 seems to be missing
1. Referee #1 makes several comments and suggestions, which should be addressed with the suitable amendments and clarifications. See below.

2. Please provide the quantitative experimental data from figure 4 as figure 'source files' (http://www.nature.com/msb/authors/index.html#a3.4.3) and include the 13C tracing dataset in supplementary information.

   The experimental raw data for figures 4, S2, and S3 have been added as one numerical data sheet (dataSheet) to the supplementary tables.

3. Please provide your model in SBML format in supplementary information. We would encourage you to also deposit it to one of the appropriate public databases (eg Biomodels).

   We added MIRIAM annotation to our model in sbml format and apart from including it in the submission we also provide it online for download at (http://nin.crg.es/serranolab/mycomap/; user name: mycomap, password: bicha987). Furthermore, we uploaded the model to the BioModels database, it has the ID MODEL1301290000. We added this information at the beginning of Material and Methods.

Reviewer #1 (Remarks to the Author):

We feel that the authors have sufficiently addressed our concerns from the last revision.

One anomaly in the paper which must be addressed:

In Figure 4, how can one of the bars be negative if they represent protein production? This is not an anomaly but a real result showing a decrease in the total protein content. It can be explained by the fact that under high pH stress, as well as at the end of a four days batch culture growth experiment, M. pneumoniae cells possibly start dying. This does not happen to cells grown under more favorable conditions. However, to make this clear, we now specify this in the figure legend (page 31, lines 938/939).

In addition, we have a few minor comments:

1. Lines 188-190: There is a misplaced comma before "verified"? What does the sentence mean? How did the author confirm "a wide range of experimental data on mRNA expression and for the design of a defined medium"? A reference to supplementary material specifying the tests that were run would be appropriate.

   We removed the misplaced comma and rewrote the sentence, more specifically pointing out experimental data used for the evaluation of model predictions (page 7, lines 185ff). In addition, we added a paragraph in the supplementary information (SI) that details our work (SI, page 10).

2. Lines 241-243: how did the authors incorporate a constraint on the "total protein increase"? If this is not a standard method do the results reproduce without it?

   We did not set a constraint on the protein production reaction but fixed the objective value according to the doubling times obtained from the in vivo protein measurements. In the text we now specify this further (page 9, lines 238-240) and also added a sentence to the end of the constraint section in the SI (page 10).

3. Lines 317-322: the authors discuss their GAM/NGAM ratio vis-a-vis the ratios in the model of Feist et al. However, GAM/NGAM estimations varied considerably between different revisions of the E. coli model. This is in part due to the the effects of the different experimental settings from which data was gleaned. See the discussion in the reviewing process of the recent Orth et al. paper in MSB 2011: http://www.nature.com/msb/journal/v7/n1/extref/msb201165-s4.pdf.

   We think that:
a) the considerable variance in those estimations should be mentioned in the discussion, and

b) the discussion should consider the effect of the experimental settings on the result concerning NGAM/GAM ratio in M. pneumoniae

We agree with the reviewer that GAM/NGAM ratios highly depend on the experimental conditions, and we show that a primary indicator for this ratio is the doubling time. Following the suggestions, we rewrote the respective part of the discussion as follows: “Despite it has been shown that GAM/NGAM estimations show considerable variance depending on the experimental data used (Varma & Palsson, 1994b; Feist et al, 2007; Orth et al, 2011), our finding is in agreement with recent results from the whole-cell model in M. genitalium (Karr et al, 2012). Furthermore, we showed that the doubling time has major impact on the ratio between GAM and NGAM since assuming different doubling times while providing the same amount of nutrients leads to a significant alteration of this ratio (table S7).” (page 19, lines 565ff).

4. lines 325-327: "It is important to note that even during the exponential growth phase 78% to 89% of the available energy is not directly used for the production of cell building blocks but for cellular homeostasis. --- Is this sentence referring again to the M. genitalium study of Karr et al.? It is not clear from the context because of the interceding discussion of M. mobile. In addition - is this an in-silico result? If so, it should be explicitly said.

The respective sentence now reads “It is important to note that in M. pneumoniae in silico even during the exponential growth phase 78% to 89% of the available energy is not directly used for the production of cell building blocks but for cellular homeostasis.” (page 11, lines 322ff).

5. lines 343-344: The authors write that they fitted a logarithmic curve, but doesn't the figure show a linear fitting?

We apologize for the confusion, but we used a polynomial fitting (blue) (not a logarithmic one) because it resulted in an R² value of 1, slightly better than the linear one (red), and we aimed to calculate the constraints as exact as possible based on the available data. However, the difference to a linear fitting is really small (hardly visible in the small plot) and does not affect significantly the results. In the text the logarithmic has been replaced by polynomial (page 12, line 339).

6. lines 346-347: what are the "respectively determined constraints" for the simulated rich medium? A supplementary note is in order here.

We explain how the constraints are calculated in the SI (page 8f) and the respective constraints are listed in table S5. We reference the respective supplementary material now better in the main text (page 12, line 343).

7. lines 346-354: how did the authors calculate the difference between flux distributions? FVA? difference between representative FBA solutions. If the latter is true, an explicit note should be made that the multiple optima are of no concern in this model, as the authors explain in the cover letter.

We added a sentence explaining the usage of representative flux distributions referring to the FVA: “To this end, we used representative flux distributions instead of considering multiple optima since a
flux variability analysis (FVA) showed only negligible variability in the majority of the reactions (table S10, see below)” (page 12, lines 345ff) and also that we used representative flux distributions when analyzing the predicted mutant phenotypes (page 16, line 471).

8. line 422: the comma after excluded is redundant
   removed

9. lines 471 - 475: again, is this a FVA analysis?
   Yes, and as indicated the variability for the fluxes we observed when allowing up to 0.01% less optimal solutions (as explained in the methods) is given in table S10.

10. line 744: the sentence seems incomplete or grammatically-incorrect
    The sentence has been changed into “In order to know the production per *M. pneumoniae* cell, we converted the obtained fluxes into molecules*cell⁻¹*second⁻¹.” (page 24, lines 742ff).

11. Please fix the spelling mistakes in Fig. S6 (enzymes, enzymes)
    done

12. One of the subfigures of Fig S7 seems to be missing
    We apologize, this figure legend was wrong. The information about the statistical analysis had been moved to table 3 but the legend was not corrected up to now.