Dear Drs Bruley, Gaëtan, Benzerara, Duprat and other co-authors.

Thank you very much for submitting the aforementioned M/S for evaluation to PCI-Microbiology. In the first place, I want to apologize for the delay of my response. I have had enormous problems in finding reviewers, among the more than 20 invited I only succeeded in obtaining one report, which to me appears very valuable and helpful. To advance the process I decided to do the second review myself, despite the fact that I only master this area only partially and I am not a specialist in transcriptomics. Nevertheless, I hope that you will find my report useful.

For the time-being, albeit promising, I cannot yet recommend this paper and thus advice formally "This preprint merits a revision" (considering the 3 possibilities for PCI recommendors). Hence, I ask you to prepare a carefully revised version. Please carefully consider the points raised by the reviewers and take them into account for the preparation of the revised version. I highly appreciate if you can prepare a reply where you pointwise address the issues raised by the reviewers and indicate how you have accommodated them in the revised version. I am looking forward to receiving the revised version with the reply in due course. Upon receipt, I will contact the other reviewer for additional assessment and also prepare my own assessment.

Thank you very much Dr. De Wit for handling this editorial process and reviewing our manuscript. We understand that it is particularly difficult to find reviewers, probably because people may have little time left but also possibly because this work is at the forefront of several fields and not "mainstream" so to speak. Thereafter we respond point by point to all the review comments and indicate the changes that have been made to address all of them. We do have a version of the revised manuscript with tracked changes if this is useful. It is on bioarchiv in the supplementary files

Reviewer Rutger De Wit

1) *This study analysis the transcriptome of* Microcystis aeruginosa *PCC7806 cultured under a L/D cycle and describes a clear diel rhythm for the gene expression of ccyA, a gene documented in the literature for a putative role in intracellular $CaCO_3$ precipitation in cyanobacteria, i.e. intracellular amorphous calcium carbonates (iACC). Hence a roughly 2.5 variation in abundance has been observed with highest values of* ccyA *expression at the end of the dark period. Nevertheless, the exact role of* ccyA *remains unknown and the documentation of its diel rhythm assumes that it can be used as a marker gene to learn about the function of this intracellular $CaCO_3$ precipitation (iACC) in cyanobacteria. As the authors write it could be linked to CCM, intracellular pH buffering, and creating "ballast" for regulating buoyancy and floatability. In their Introduction and Discussion the authors neglect that such functions for iACC would imply a necessity of mechanisms for the dissolution of iACC in concert with its precipitation; fine-tuning of both resulting in homeostasis or cyclic temporal patterns of increasing decreasing iACC.*

Thank you for this comment. We now mention in the introduction and in the discussion that there might be processes involved in the controlled dissolution of iACC but that this has to be evidenced and studied in the future.

To obtain more indications about the role of iACC precipitation, the authors have studied the transcriptomes very largely, to detect gene-expressions with temporal patterns that positively or negatively (anticorrelate) with ccyA. A particular interest was put on neighboring genes (both upstream and downstream) to detect a possible operon comprising ccyA with other genes.

Very interestingly they discovered that some genes coding for Ca^{2+}/H^{+} antiporter systems, which occur as neighbors showed transcripts with abundances that correlate with the ccyA. Hence, the existence of an operon with ccyA and these Ca^{2+}/H^{+} antiporter systems seems plausible. It can be considered as the other main finding of this study.

2) *Strikingly, most CCM transcripts are higher during daytime, while some others are higher during night time. This may appear confusing (in the Discussion the authors explain it by CCM being coupled to photosynthesis and therefore more related to daytime), but is perhaps more interesting than suggested. We realize that CCM comprises a multitude of inorganic carbon uptake systems differing by use of substrate (CO_2 or HCO_3^-) and different affinities and maximum uptake rates, that can operate in parallel and can be complementary. It could thus be envisioned that some of them are better suited to accumulate inorganic carbon during night time, and why not we can hypothesize that such nighttime CCM operates in concert with the iACC precipitation to create a temporal stock of sequestered inorganic carbon that could be liberated during daytime.*

Thank you. We now mention this hypothesis of transporters with distinct affinities that may operate over day or night and stress that this should be further investigated by future studies.

3) *In the Introduction the authors mention that diel patterns are controlled by Circadian rhythms, but this is not necessarily the case as the control of gene expression may also be based on sensing physiological conditions as e.g. pH, amount of storage compounds etc. A first indication whether a Circadian clock is involved can be obtained by transferring the culture to continuous light conditions (the cycle should then be maintained for several roughly 24 h periods). Has it been studied whether this strain PCC7806 has a Circadian clock? For another strain of Microcystis aeruginosa PCC 7820 this has been documented by Huang et al (Huang, J., Wang, J. & Xu, H. The circadian rhythms of photosynthesis, ATP content and cell division in Microcystis aeruginosa PCC7820. Acta Physiol Plant* ***36****, 3315–3323 (2014). https://doi.org/10.1007/s11738-014-1699-1) and the authors may consider citing this paper and any others that could support the occurrence and role of a Circadian clock in PCC7806. The authors could perhaps even check if genes for KaiA3 and KaiB3-KaiC3 or analogues are expressed in this strain PCC7806.*

The genes involved in the circadian cycle were previously studied in the PCC 7806 strain by Straub et al. (2011) under a 24-hour light/dark cycle. The study found that while the *kaiA* gene showed no significant variations during the cycle, the transcription patterns of the *kaiB* and *kaiC* genes, as well as the *sasA* gene encoding the two-component sensor histidine kinase (a KaiC-interacting protein), exhibited significant changes. These findings suggested that light is not the sole factor triggering the transcription of genes involved in photosynthesis and respiration; instead, their transcription may also be regulated by an endogenous circadian clock. This is now mentioned in the introduction and we also cite Huang et al 2014.

4) *-A most general important point: At least, in the bioRxiv preprint, the quality of the Figures is generally low, particularly of Figs. 4 and 6, and also 3. Please can you provide better quality Figures with more attractive lay-out and easier to read. This will be very important to improve the understanding for the readers and thus the impact of*

your paper.

It seems that the conversion to the pdf available in bioRxiv was detrimental to the figure quality. We have now resaved Figures 3, 4 and 6 at 5000 dpi.

5) *-Line 113: Methods: In the Methods, first point to mention is the species and strain used (with its reference allowing to obtain the strain), before describing the culture conditions.*
   This was modified accordingly

6) *-Line 115: (50 mmol photons.m-2.s-1 appears excessively high (20 to 25 times Zenith intensities outside in the Tropics), I think it should be (50 µmol photons.m-2.s-1 (micromolar), which is low but not surprising when using artificial light.*

Absolutely! Sorry this was a font change mistake. This has now been corrected

7) *-Line 163: rephrase "Raw RNA-seq reads (available online, see reference Raw transcriptomics data)" by "Raw RNA-seq reads (available online, see section Data, scripts, code, and supplementary information availability).*

This has been rephrased accordingly

8) *-Upon publication you should make these data available (relieve the private constraint).*

   Data will be released upon publication, the data will be under the accession number GSE255450

9) *-Line 210-211 : Each replicate (i.e. three independent cultures) at a single time showed only minor variations between them, at least along axis 1 (accounting for almost 50% of the variance), and were distinctly separated from the replicates at other time points." – should be reformulated as follows: "Among the triplicates (i.e. three independent cultures for each time-point) only minor variations were observed, at least along axis 1, and their values were clearly separated from the samplings at other time points."*

Thank you, this has been modified accordingly.

10) *-Line 240: replace "Figure 2 - Abundance profile of ccyA transcripts during a day/night cycle" by "Figure 2 – Time course of the abundance of ccyA transcripts during a day/night cycle"*

This was modified accordingly

11) *-Please use terminology consistently. Note that diurnal = during daytime (i.e. L period) as opposed to nocturnal (during night or D period). Diel = variation during the entire 24 h cycle comprising both L (day) and D (night) periods.*

This was corrected and diurnal was replaced by diel.

12) *-To prevent confusion, please consistently use either the term "log base 2" or the term*

"binary logarithm" but not both (personally I prefer "log base 2", which is more commonly used and clear). You may also use the mathematical formulation "$\log_2 (x)$"

Binary logarithm has been replaced by log base 2.

Reviewer 1

This manuscript describes laboratory experiments with a strain of toxic bloom forming cyanobacterium, M. aeruginosa PCC 7806 in order to clarify the molecular basis for internal calcium carbonate accumulation. The study focuses on rRNA transcriptomics, particularly that of the expression of ccyA gene and its upstream and downstream neighbors in a light-dark incubation. Based on their observations, the authors show a diel expression pattern, correlation with CCMs. They also use Foldseek to assign functions to hypothetical proteins, encoded for by neighboring genes on the same DNA strand. If assumptions are correct, this indicates that specific transporters and specific carbon concentrating mechanisms are activated (note to the authors: you may be aware of this but do not mention in the manuscript that many CCM exist in cyanobacteria; Kupriyanova et al. 2011) when the calcium concentrating gene is also turned on (at night). Although the authors are speculative in their interpretation of some observations, they also provide novel insights based on elegant experiments, especially the description of neighboring gene functions.

We have now added a reference to Kurpiyanova et al 2023 to mention that there are multiple mechanisms involved in CCM (we did not find the 2011 reference).

This reviewer found a few issues somewhat puzzling:

-The observations are clear and interpretations somewhat speculative but persuasive, this reviewer is not convinced that identical general expression patterns of genes would be found in species that do not contain the ccyA gene. The absence of a proper control is weakness of this manuscript.

Indeed, we do not have a dataset on a strain not forming intracellular ACC. However, we do not hold any conclusion comparing cyanobacteria forming vs not-forming iACC and therefore we do not feel that such a control is so much needed. Moreover, another species not hosting *ccyA* would have several other genes distinct from the ones in PCC 7806 and overall it would be difficult to consider it as a real control. A more reliable control would be a mutant of 7806 with a deleted *ccyA* gene but unfortunately this mutant does not currently exist to our knowledge and making a mutant of PCC 7806 or another *Microcystis* is not an amenable task and a challenge. We now mention these points.

-Growth of the strain was carried out in full strength BG-11, something generally prevented in carbonate precipitating experiments due to artefacts cause by e.g., high phosphate concentrations (Rivadeneyra et al. 2006, 2010).

It is true that high phosphate concentrations such as those in BG-11 (~180 µmol.L-1) may prevent the precipitation of extracellular carbonates and/or induce that of Ca-phosphates as shown by Rivadeneyra et al. However, it has been shown multiple times that this does not prevent at all the precipitation of intracellular carbonates. All past experiments were performed in BG-11 (Cam et al., 2017; De Wever et al., 2019). We have now added two sentences about this point in the method section.

-If energy is required for "biomineralization" (lines 79-80), why is the strongest gene

expression for ccyA observed at the end of the dark period?

Straub et al. (2011) mentioned that the metabolism of *M. aeruginosa* is compartmentalized between the light period, during which carbon uptake, photosynthesis and the reductive pentose phosphate pathway lead to the synthesis of glycogen, and the dark period, during which glycogen degradation, the oxidative pentose phosphate pathway, the TCA branched pathway and ammonium uptake promote amino acid biosynthesis". We now underscore this point so that it does not appear surprising to the reader that expression of ccyA is the strongest at the end of the dark period.

-Although the Gaëtan et al (2022) shows illustrations of calcium inclusions in some but not all Microcyctis spp. obtained from several lakes in France and Spain, it is still not clarified which fraction of a natural bloom has calcium concentrating capabilities. Furthermore, by far the majority of lakes around the globe have calcium concentrations far lower than those found in full-strength BG-11 (Weyhenmeyer et al. 2009).

It is true that it is still not clear which fraction of a natural bloom has Ca concentrating capabilities and this remains an interesting research line for the future as we now mention it. About the Ca concentrations in lakes Weyhenmeyer et al. 2009 showed that for lakes with a pH above 7.4, more than 50% of them have a Ca concentration above that of BG11. But we agree that there are still many lakes with a Ca concentration below that of BG11. Yet, a comparison with (batch) culture conditions seems more complicate than that. Indeed, the relative volume of a lake compared with the volume of cells is very high, much higher than in a dense culture in a closed system as an Erlenmeyer. So eventually, it may be the quantity of Ca in a lake and not its concentration that matters. As an example, Gloeomargarita, the first cyanobacterium shown to form iACC, was found in Lake Alchichica where Ca concentration is about 50 µM, lower than in BG11. We have added a comment about this point in the revised manuscript.

-The authors refer to an earlier paper from the same group that hypothesizes that calcium carbonate (r = 2.71 g.cm^{-3}) would be part of buoyancy regulation. This do not make much sense, as it would increase the sinking rate, and thus removal from the photic zone. In fact, Gu et al. (2020, 2023) demonstrated that the presence of calcium (in concentrations below but close to that of full strength BG11) induce massive exopolymeric substance production. There is no report in this manuscript if exopolymers were formed.

Yet, whatever the benefit cells can get from it, there is no question that the density of an ACC inclusion is higher than that of a cell and therefore the formation of iACC increases cell density. We note that buoyancy is not just controlled by cell density but also by additional parameters such as cell aggregation. For example, Gu et al. (2020) showed that Ca induces EPS production which can increase buoyancy. The relative contribution of these opposing parameters on the buoyancy should be studied in the future. We mention this point in the revised manuscript.

-Walter et al. (2016) demonstrated a role of calcium in regulating carbonic anhydrase in Anabaena sp., perhaps a complicating factor if this exists in the Microcystis strain used in this study? On the same topic, but a different issue is the impact of the culture conditions on gene expression. Why did the authors choose 11 hr dark/13 hr light (which, by the way, is not diurnal as stated in line 101, but diel, as correctly stated in the title of the manuscript).

We now cite Walter et al. (2016) and underline that extracellular calcium concentrations should also be considered in future studies. A 13-hour light period corresponds to what is

observed in September at our latitude, a month during which *Microcystis* proliferates in numerous ecosystems. This is now mentioned in the methods section. Finally, diurnal was replaced by diel.

-Why would a Na^{+}-expelling antiporter (ApNhaP) typically found in halotolerant cyanobacteria be of functional importance to Ca^{2+} accumulation?

The apnhaP gene encodes a Na^{+}/H^{+} antiporter that also functions as a Ca^{2+}/H^{+} antiporter at alkaline pH (Waditee et al., 2001). It is homologous to genes found in non-halotolerant cyanobacteria such as *Synechocystis* or *Pseudomonas*. This is now mentioned in the revised manuscript.

-It was not clear to this reviewer what "normalized counts" used, e.g., in Figs 1 and 7 refer to. Normalized to what? The use of statistics was useful and thorough in this manuscript, especially since some observations seem barely significant (Table 1, the day night transcriptome numbers).

This is explained in the Methods section. The estimated transcript abundances were normalized by Salmon according to the transcript size, genome size (number of CDS) and sample size (number of reads). This is now specified in the captions of Figs 1 and 7.

-I suggest that the authors use be more specific when making certain statements (e.g., line 22 "many phylogenetically diverse"; lines 46, 66 how "widespread?, 67 "several" – how many exactly; line 83 " one third of the publicly released genomes" but how many are published?, etc.) and tone down some other statements (the presumed importance in (bio)geochemical cyles; the cell may accumulate but then lyse after a bloom, so it would merely be another transport mechanism; line 63 "massively sequester"; line 66 " widespread"; line 78 potential "environmental importance, etc.). The definition of "biomineralization" in line 46 is ambivalent and weak – directly/indirectly? Induce, produce…?

We removed "many" on line 22 (the number of these cyanobacteria is constantly changing and the latest one is unpublished so it does not make sense to give a number here). We replaced "widespread" by found. Again, we do not know of a definitive number of taxa able to biomineralize or this would be a poorly accurate number. We specify the number of genomes in "one third…". If cells do accumulate Ca even transiently, this impacts the geochemical cycle at a certain timescale but anyhow, we toned down the importance in biogeochemical cycles based on the fact that there are still unknowns as mentioned above. We removed massively (although studies have shown that in an open system they do significantly change the Sr and Ba concentrations. "widespread" was removed. We disagree about the definition of biomineralization. This is the definition in the broadest sense. Then, we agree that there are several types of biomineralizations based on how they work, that can be direct or indirect, induced or controlled... This is specified later on.

-The authors introduce in line 36, abstract and in later in the text the "CoBaHMA" domain. For those of us who have read the Benzerara 2022 paper may recall the introduction of this abbreviation, but given the importance the authors give to the potential function and thus importance to the calcium concentrating mechanism, the may consider providing a brief explanation.

We removed the term coBaHMA in the abstract. And we now provide a brief explanation about the CoBaHMA domain when we first mention it.