REVIEWER COMMENTS ON bg-2023-138 by Delacroix et al.

This study reports on the analysis of the effects of the exposure of phytoplankton to ocean liming by using brucite ($Mg(OH)_2$ and $Ca(OH)_2$) mimicking the initial concentrations following a dispersion scenario from a ship. Three experiments were done: 1) exposure of the marine chlorophyte *Tetraselmis suecica* to simulated dispersion of the minerals from an allegedly moving ship; 2) growth rate inhibition of *Skeletonema costatum* exposed to Mg(OH) by using a marine algal growth inhibition test (called Whole Effluent Toxicity test i.e. WET)*;* 3) exposure of a natural plankton community from the Oslo fjord to increasing concentrations $Mg(OH)_2$.

The authors conclude that $Mg(OH)_2$ is a suitable mineral for OAE application. This was supported by: 1) high *T. suecica* mortality at high $Ca(OH)_2$ concentrations during the first hour of a supposedly discharge from a moving ship *vs.* no mortality to $Mg(OH)_2$ discharge. For this, green cell viability fluorescence stains were used; 2) Cell numbers of *S. costatum* assessed by fluorescence (666nm?) showed apparently no differences between LOEC and NOEC for $Mg(OH)_2$ while EC50 was 2 and 3-fold higher concentration regarding to LOEC and NOEC respectively; 3) there was a significant difference in algal survival between the low and high $Mg(OH)_2$ concentrations three days of exposure of diatoms, dinoflagellates and other unspecified organisms from a natural phytoplankton community.

The outcome of these experiments seems to be that bioassays based on initial local and temporary discharge simulation from alkaline mineral dispersion from ships, demonstrated that $Mg(OH)_2$ resulted in lower biological impacts on marine microalgae when compared to $Ca(OH)_2$.

Unfortunately, I am afraid I cannot support publication in BG since in my opinion, there are important flaws regarding the experimental design, data interpretation and overstated conclusions, that prevent this work to meet the required quality to be published in BG.

The arguments in which my recommendation is based are the following:

MAJOR COMENTS

1. One major concern is that authors use the computational models initially , to set-up the background and aims for the subsequent experiments. However, I see there is a huge gap in between the computational models output and what it was simulated in the lab experiments. Especially, regarding the carbonate system, speciation, pH and precipitation.

Has not pH increase undesirably fast? And so, it is possible that uncontrolled CaCO3 precipitation could lower the CO2 sequestration efficiency of the approach? Theoretically Ca(OH)2, and (Mg(OH)2) should dissolve rapidly in the ocean surface , but I am not sure this the case. I have the impression that the ratios dissolution /precipitation were not controlled to check for reactivity and spontaneous

precipitation in seawater. In short, were these hydroxides well dissolved? In addition, and regarding the commercially sourced material, was the carbonate content measured before the experiment to check for carbonation? I mean, how can you limit carbonation being present within the hydroxides?

I have serious difficulties concealing the results obtained from de lab experiments with the just commented above. Could your results be generalised, and compared to the simulated computer models and extrapolated to general conclusions that can have relevance for the application of OAE in the real-world? I am not sure…

2. I have the impression that the use of the double staining method with Fluorescein Diacetate (FDA) and 5-chloromethylfluorescein diacetate (CMFDA) have been used lacking rigour (just my opinion and experience).

It is important to note that fluorescent stains must be validated in each different experiment before use by analysing the optimal dye loading concentration and loading kinetics for each specific monoculture and /or phytoplankton community, to avoid sub-optimal fluorescence or saturated fluorescence reaching the laser detectors of flow cytometers or epifluorescence microscopes. This essential step seems to have been omitted previously to the start of the experiment. The concentration used by the authors in this experiment is 2 orders of magnitude higher (2.5 mM final concertation) than working concentrations widely used in different marine coastal and open oceanic waters , as well as in lab cultures by well stablished protocols and SOPs (please check references at the end of the review). Attending to this, the lack of mortality can be an artefact due to the spill over of green fluorescence due to excess oversaturated signal, which commonly occurs when the fluorescent dye concentration has not been customised for every cell type.
For example, for CMFDA long-term staining (more than about 3 days) or the use of rapidly dividing cells, 5–25 µM dye is required. Less dye (0.5–5 µM) is usually needed for shorter experiments, such as viability assays in cultures and about 20 µM in natural populations (always final concentrations) as it is the case in this study. To maintain normal cellular physiology and reduce potential artefacts, as already mentioned, the dye concentration must be kept as low as possible. The effects of overloading may not be apparent, hence, to check for this, a cell death stain must be used in combination in the same set of aliquots containing the cells aim of study (please check references at the end of the review).

In addition, the CMFDA fluorescent probe is well retained in living cells through several generations. The probe is transferred to daughter cells but are not transferred to adjacent cells in a population. Cells loaded with the CMFDA fluorescent probes display fluorescence for at least 72 hours and exhibit ideal tracking dye properties—they are stable, nontoxic at working concentrations well retained in cells, and brightly fluorescent at physiological pH. Therefore, assessing cell viability with this fluorescent probe is not entirely accurate since daughter cells can be metabolically non-viable and yet, show green fluorescence. In this case, cell viability is overestimated (please check references at the end of the review).

Regarding FDA I have similar concerns except for the transfer of dye to the daughter cells that is not the case with FDA.

3. I must also comment that I miss the detailed staining protocol for cells and fluorescent microscopy quantification. How were cells harvested: by centrifugation and aspirate the supernatant? By filtration? Were they resuspended in pre-warmed or RT working Solution ? Gently or vortexed? For how long were cells incubated with the dyes? Were cells centrifuged to remove the excess dye working solution? Was culture media added and the labelled cell dispensed onto slide or into a chamber-wells for imaging? For how long were they imaged using the appropriate emission and excitation filters under the scope? The point being, if someone is to reproduce your experiment, not sure that would be possible with the insufficient information provided…

The kinetics and loading curves set-up for the optimal dye concentration and time of incubation should be provided in supplemental material, or at least mention in the text that they are available to reviewers in case they would like to check on them (as it is my case for example)

4. More details on negative and positive controls choice would be desirable. Were general procedural negative controls were done?

5. I was wondering why in the case of *Tetraselmis* a flow cytometer was not used… this would have most likely had produced more accurate results as compared to microscope observations. Regarding the scope is not clear how the % of viable cells was calculated nor how representative your sample was. What was N? how many fields of view (FOVs) were counted per slide or well? In each of the independent cultures? If only one slide/independent culture counted, seems not enough to me. Not clear either which statistical analyses has been carried out for this? Epifluorescence (or any microscopically quantification) can lead to artefactual data unless N is large enough (not sure this is the case, and clarification is needed), or other intercalibrating method is used to contrast with numbers, such as flow cytometry.

6. Another question that is confusing to me is the lack of standardised methods for measuring cell performance. It is not clear enough why fluorescent probes were used with *Tetraselmis*, but not for Skeletonema nor for the natural community. Moreover, for Skeletonema fluorescence, authors do not specify which fluorescence was measured? Red 666nm? In which device was this measured? Green with probes? Again, in which device? Along the same line, the rationale for the natural community analysis it is not well understood, nor how the % of viable cells was also calculated. Table 3 and 4 are difficult to understand because it cannot be discriminated to which functional group each % belong, therefore, not sure I see the point for this.

I'd like to point out that the methodology description does not suffice to understand how this experiment was performed. The ms. does not have an easy thread-line to be followed. Could it perhaps be better organised?

7. The statistical approach used, does not seem appropriate. First, does data distribution meet the requirements for parametrical tests? Assuming so, T-tests do not capture the variability of the system you might have. Hence, most likely, significant differences are not well resolved, questioning the results. I would suggest that a one-way ANOVA would be adequate since you have 5 levels of concentrations. Also, in those cases in which time is a continuous variable and not an end-point variable, the right approach could be a split-plot ANOVA in which the fixed factor would be the concentration and the repeated measures factor would be time.

8. In my opinion the discussion is shallow. It does not get deep insight on explaining the data, nor debating them. Not enough quality for a discussion I'm afraid. The same applies to the conclusions, which are merely descriptive.

MINOR COMMENTS

Ln. 139. The statement "were repeated three times for each alkaline mineral" to what this exactly refers? Repeated when? How would your N then vary?

Ln. 214. "A preliminary study was made to verify the microalgal growth in this modified media". Why is this not shown?

Ln.221. What do you envision would have occurred if instead of applying the minerals to log-phase cells , they would have been added in lag-phase (in which many cells are in natural conditions at sea)? Would have you expected any growth? This particularity of the growth curve should have been tested too. The outcome can be surprising and non-acclimation i.e. death, shall be considered as a very likely possibility, questioning the application of these chemicals in the ocean.

Ln. 236. "25 L grab-sample from the surface water of Oslofjord
was directly used for the test or a 2 L subsample was mixed to 2 L of 60 m deep seawater from Oslofjord". You are mixing surface cells acclimated at a probable high irradiance with water at 60m depth and submitting them to a few umols photons m-2s-1. This might for sure have consequences in cell viability, just because metabolism can change due to the new light scenario. Metabolic activity is reflected by viability cell stains since they are dependent on esterases inside the cell.

Ln 282. Perhaps this table would be best in supplementary material?

<u>References:</u>

1. McIntyre, H. L. & Cullen, J. J. 2016. Classification of phytoplankton cells as live or dead using the vital stains fluorescein diacetate and 5-chloromethylfluorescein diacetate.J. Phycol. 52, 572–589.
2. Peperzak, L. & Brussaard, C. P. D. 2011. Flow cytometric applicability of fluorescent vitality probes on phytoplankton. J. Phy- col. 47:692–702.

3. Segovia, M., & Berges, J. A. (2009). Inhibition of caspase-like activities prevents the appearance of reactive oxygen species and dark-induced apoptosis in the unicellular chlorophyte Dunaliella tertiolecta. Journal of Phycology, 45(5), 1116–1126.
4. Segovia, M., Lorenzo, M. R., Maldonado, M. T., Larsen, A., Berger, S. A., Tsagaraki, T. M., & Egge, J. K. (2017). Iron availability modulates the effects of future CO2 levels within the marine planktonic food web. Marine Ecology Progress Series, 565, 17-33.
5. Sobrino, C., Segovia, M., Neale, P. J., Mercado, J. M., García-Gómez, C., Kulk, G., & Ruan, Z. (2014). Effect of CO2, nutrients and light on coastal plankton: Physiological responses. Aquatic Biology, 22, 77-93
6. Steinberg, M. K., Lemieux, E. J. & Drake, L. A. 2011. Determining the viability of marine protists using a combination of vital, fluorescent stains. Mar. Biol. 158:1431–7.
7. Veldhuis, M. J. W., Cucci, T. L. & Sieracki, M. E. 1997. Cellular DNA content of marine phytoplankton using two new fluorochromes: taxonomic and ecological implications. J. Phycol. 33:527–41.
8. Veldhuis, M. J. W., Kraay, G. W. & Timmermans, K. R. 2001. Cell death in phytoplankton: correlation between changes in membrane permeability, photosynthetic activity, pigmentation and growth. Eur. J. Phycol. 36:167–77.