Manuscript 'The metabolic response of pteropods to ocean acidification reflects natural CO2-exposure in oxygen minimum zones' (bg-2011-370) by Maas AE, Wishner KF, Seibel BA.

GENERAL COMMENTS

The authors have carried out a very interesting work in an attempt to test a very exciting hypothesis, the idea that differences in the evolutionary ecology of pteropods examined here (migratory vs. non-migratory species respect to the OMZ) or at least their pre-exposure (or not) to the oxygen minimum zone may underpin their relative vulnerability to OA. The authors go about to test this idea by measuring the metabolic responses (respiration and excretion rates) of pteropod species from the tropical regions of the Pacific Ocean, exposing within few hours upon collection individuals to low pH/hypercapnia conditions following at least 8h acclimation time to 20 °C. The work is indeed important providing novel information on the physiological ecology of warm-water pteropods, an understudied group of pteropods. The work appears overall well organized and conducted. There is some repetition of information between tables and figure (e.g. Table 4 and Figure 3 and 4). However, a major issue with this work is that of the five species examined four are migratory and one species is of the non-migratory kind. It is not possible to conclude on the basis of one single non-migratory species (*Diacria quadridentata*) never found below the mixed layer, that all non-migratory thecosomes will not be able to withstand low pH/high CO_2 conditions. The dataset as it is now cannot be employed to support the idea/statement that the evolutionary ecology or pre-exposure of these groups of pteropods may determine their relative vulnerability to future OA scenario: at best the author could briefly suggest this, but this study (unfortunately) cannot be considered conclusive. It is a true shame, because the authors' idea was (is) truly brilliant, and indeed well above the 'effect test' OA papers which appear to keep accumulating.

I think suggesting to collect more data on more non-migratory species of warm-water pteropods might not help, as the authors may not be in the position to repeat this logistically demanding experiments, thus I have to suggest the authors substantially modify their manuscript focusing on the ecophysiology of warm-water pteropods and avoid inferring on the ecological significance of the evolutionary ecology. In particular, the title has to be changed (as what it states is unsupported), and equally parts of the introduction and discussion relative to the testing of the main hypothesis are changed. Finally, the present manuscript is in parts overly concessive and there is a need to provide more details in the methods sections, and broaden up the use of current OA-related literature. I truly enjoyed reading about the authors work, but I am afraid I cannot recommend this manuscript for publication in *Biogeosciences* in its present form, although a thoroughly revised version could be considered by the Editor if he/she wishes so.

MAJOR COMMENTS

- Page 10298 – line 28 and Page 10299 – line 16. Incubation time varies between 6 and 18h. Why was this not standardize? Could the author provide a rational for not doing so? Finally, and most importantly did the authors included incubation time as a covariate in the analysis (beyond calculating rates by h^{-1}) to control for variation in resting/incubation time?

- Also I suggest strongly the authors do acknowledge (briefly) in the discussion the (potential) limitations of using short term exposures to low pH (6-18h) when inferring on chronic exposure to OA conditions. The authors do actually suggest that their responses documented in migratory thecosome pteropods to short term response to high CO_2 might be a 'best case scenario' (when compared to future chronic exposures). However, it is important that the authors discuss the potential for all pteropods not to have reached a new stable status when metabolic responses were measured. The risk is that the authors are potentially reporting pteropods metabolic rates during 'overshooting'.
- Page 10299 – line 22. The pH levels used (8.32 and 7.96) may not be representative of future OA scenarios, but given global change (included OA) is a regional event, and I do not know the characteristics of the carbonate system of the sea water from the study area it is difficult to say. However, the authors do mention in the Introduction and Discussion about pH < 7.6 and undersaturation with respect to aragonite. At the pH the authors worked there should not be undersaturation for aragonite (or calcite). In turns, it is difficult to talk about OA despite the pH decrease is approximately 0.3 – given OA includes not only a reduction in pH but also in $[CO_3^{2-}]$.
- No details on the carbonate system are reported here. In particular, it appears the authors carried out pH measurements but did not measure TA, DIC or pCO_2. How did the authors verify that the CO_2 concentration in the sea water was 380/400 and 1000ppm for their control and acidified treatments, respectively? Measuring pH is not sufficient, even if the authors used pre-mixed gasses. The work for this aspect does not seem to meet the required by current standards in OA literature. Most importantly, it is not possible to discuss the data not having an idea of what DIC and pCO_2 are, and even more $[CO_3^{2-}]$ Ω_{ara} and Ω_{calc} are.
- Page 10297 – line 6 and Page 10300 – line 24. Remove references Seibel et al. 2011 and Maas et al. 2011 altogether here and elsewhere, as one is submitted and the other one only in preparation. Alternatively provide more information and write 'Seibel/Maas et al. *unpublished*', until these ms are eventually accepted. As a review I cannot access this information at the moment.
- Page 10303 – lines 4 to 8. I strongly suggest the authors include in this section (and make use of) relevant references on the metabolic responses of mollusks (e.g. Gutowska et al. 2008; Rosa and Seibel 2008; Comeau's work; Cumming et al. 2011; Melatunan et al. 2011).
- The interpolation between experimental evidence and field data should be stronger and more analitycal where possible.

DETAILED COMMENTS

- Page 10297 – line 7. Please expand the literature support to other studies too.
- Page 10300 - line 1. Please report the model of the probe and meter you used.
- Page 10301 – line 4 and 10302 – line 10 (and elsewhere in the text). It appears that the authors use the terms 'depression' 'suppression' and 'reduction' interchangeably. If this is the case I suggest they standardize to one terminology, if not I warmly invite the authors in providing strong (even referenced) definitions for each term and use them rigorously.
- Figure 3 and 4 (and text). The international unit for 'hours' is 'h' and not 'hr'. Please change this.
- Figure 1, 2 and 5 are quite small and difficult to read at the moment.