

RC1: ANONYMOUS REFEREE #1

The study by Bednarsek et al utilizes available data from the ocean acidification literature to evaluate how marine calcifiers could respond to ocean alkalinity enhancement (OAE). The analysis takes a statistical approach. The key concern is their use of the TA:DIC metric, which is considered indicative for carbonate chemistry changes induced by OAE due to its correlation with Omega (concerns detailed below). The value and conceptual basis of TA:DIC is currently unclear or possibly not valid. Therefore the authors would need to use another metric or provide a much better justification for the use of TA:DIC that is found on more than a correlation with Omega. I also have several other comments that warrant attention.

1.

- *The link between TA:DIC and OAE is perhaps not valid or at least not sufficiently well justified. In OA research, TA:DIC only changed because of increasing DIC. This, however, leads to other changes of relevant carbonate chemistry parameters (e.g. pH, CO₂, HCO₃⁻, CO₃²⁻) than a change in TA. There is no plausible explanation why the ratio of TA:DIC is a valid metric for OAE. The one argument made (control of TA:DIC on Omega) is not convincing so far because Omega itself may not be such a relevant metric for biotic calcium carbonate precipitation (although of course the relevant metric for abiotic precipitation/dissolution). If the authors cannot show that TA:DIC is indicative for OAE then the analysis of biological responses relative to this ratio is also not sound. Thus, demonstrating the conceptual validity of TA:DIC (much beyond a correlation with Omega) is crucial.*
- *For the derivation of TA:DIC, it also needs to be considered with what state of OAE the concept correlates. Before equilibration with atmospheric CO₂ or after?*
- *One way forward could be correlation analyses of TA:DIC (where TA is left constant and DIC is varied, representative for OA), ideally using the data of the studies synthesised here. It could then be explored if TA:DIC is a useful metric for a specific transient state of OAE (e.g. unequilibrated or equilibrated with the atmosphere). However, if TA:DIC does not reflect carbonate chemistry changes of OAE more comprehensively (much beyond an Omega correlation) then the authors should use another metric to correlate their biological responses with.*
- *Another question in this context is: If TA:DIC is used as a proxy for Omega, why wasn't Omega used in the first place?*
- *Finally, ratios such as TA:DIC do not consider absolute concentrations, which is another potential weakness for a metric that has not been derived from physiological theory. This is particularly problematic because most of the data is sourced from ocean acidification research, which is looking at the other direction of the pH scale.*

Response: We appreciate the reviewer raising these points, which we aim to address in a combined way, because it will help us to target them much more comprehensively in the resubmitted manuscript. We indeed believe that we have not described these issues sufficiently in the previously submitted version of the manuscript and that we need to provide details and rationale for introducing the TA:DIC ratio and its

application to the experimental data. We provide a more structured background and reasoning on the importance of introducing the TA:DIC ratio and indicate where aragonite saturation state (Ω_{ar}) falls short when conducting OAE experimental work. We explain how the TA:DIC ratio helps to simplify the carbonate system while conducting experimental work. We will add an extensive description of this in the resubmitted manuscript (we will create an additional Methods/Appendix section) where we can address this efficiently and comprehensively, particularly for those readers who need these guidelines when conducting experimental work.

For the background: whenever various OAE compounds are added to seawater so as to take up CO_2 from the atmosphere, the seawater carbonate chemistry changes in a multifaceted way affecting all acid-base species present in the seawater, resulting in all four of the commonly measured parameters (TA; DIC; pH_T ; pCO_2) changing simultaneously. Understanding of such changes is absolutely essential for those biological experimentalists who are conducting biological assessments to assess potential OAE effects. However, assessment of the changes in the carbonate chemistry induced by OAE is neither intuitive nor straightforward; it requires detailed carbonate chemistry calculations. Nor is it clear as to what constitutes an appropriate “independent variable” that one can use to summarize the state of acid-base systems in a particular seawater when planning or reporting such biological responses. It is convenient when considering seawater acid-base chemistry to describe the system in terms of two of the four parameters noted above (as well as the salinity, temperature, and pressure), though strictly that is only true for a system where the both the total quantities of non- CO_2 acid-base systems, and the various equilibrium constants themselves can be inferred from S, T, & p alone, otherwise an additional piece of information is needed for each additional acid-base pair.

The approach we took was first to assume that such a restricted system was a reasonable approximation, and then to assume that TA and DIC were known for the various seawaters involved. (This pair has the added convenience that, when expressed as amount contents (moles per kilogram of seawater)), they are separately conservative to mixing, *i.e.*, the initial composition obtained by mixing a seawater (with known TA and DIC) with an OAE compound (whose chemical composition is itself known) can be estimated provided that the mixing ratio is known. We then chose to use the ratio of TA:DIC as a potential independent variable (an examination of the plots in Figure 1 below shows that the various isolines representing pH_T , pCO_2 , and Ω_{ar} are approximately straight lines, *i.e.*, implying they correspond to a single value of the ratio TA:DIC).

Figure 1 shows the changes in the carbonate chemistry system inherent upon NaOH (black line) and Na_2CO_3 (dotted line) additions. Using such graphs of TA vs DIC (appropriate to a single salinity, temperature, and pressure) makes it easier to estimate compositional changes resulting from different alkaline additions in experimental settings. It is important to note that this estimate represents the initial state of the OAE conditions and is not representative of equilibrium conditions with respect to uptake of CO_2 from the atmosphere. That process would imply a further increase in DIC until the estimated pCO_2 equaled that of the atmosphere itself.

When we add NaOH, we increase TA only, and when we add Na₂CO₃ we increase TA and DIC at a 2:1 ratio (Figure 1a). With such additions, we then clearly understand how Ω_{ar} (Figure 1b) and pCO₂ change (Figure 1c), and how much of a change is required to bring the system back to equilibrium with respect to the atmosphere. We should also point out that if TA < 1000 $\mu\text{mol/kg}$ and DIC < 500 $\mu\text{mol/kg}$, the isolines for Ω_{ar} are no longer straight, however, such conditions are rare in the ocean and thus not widely applicable.

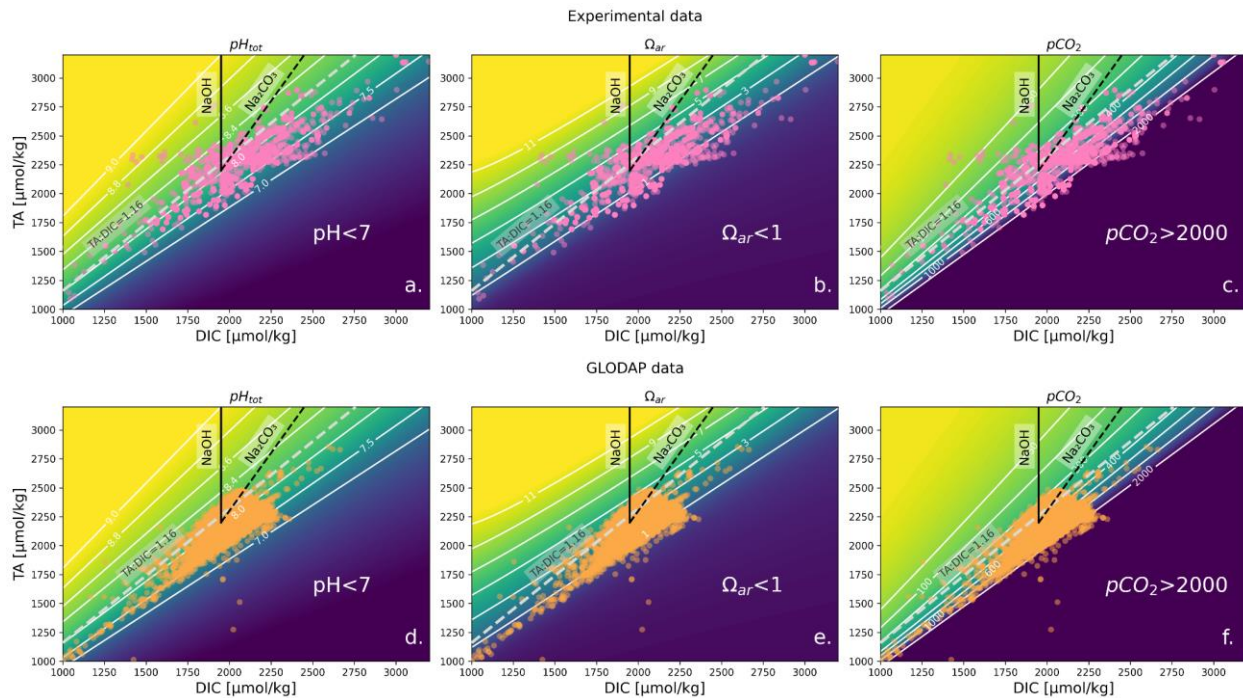


Figure 1: The effect of changes in TA and DIC on the properties of seawater ($S = 34.7$, $T = 20^\circ\text{C}$, $[\text{SiO}_2] = 50 \mu\text{mol/kg}$, $[\text{PO}_4^{3-}] = 0.5 \mu\text{mol/kg}$, $\text{TA} = 2200 \mu\text{mol/kg}$, $\text{DIC} = 1950 \mu\text{mol/kg}$), based on figures from [Schulz et al. \(2023\)](#). Pink dots (figures a-c) represent experimental TA and DIC data used in our meta-analysis, orange dots (figures d-f) represent GLODAP data for surface waters (0-50m depth). Subfigures show pH_T , aragonite saturation state and pCO_2 . Calculations were carried out with the Python version of CO2SYS ([Humphreys et al., 2022](#)) using the stoichiometric dissociation constants for carbonic acid from [Sulpis et al. \(2020\)](#), for sulfuric acid by [Dickson et al. \(1990\)](#) and the total boron value from [Uppström \(1974\)](#). The solid black line indicates the effect of adding NaOH and the dashed black line indicates the effect of adding Na₂CO₃. This pair of lines can be translated on these plots so that its initial position moves elsewhere on these figures, to visualize different initial conditions for TA and for DIC (all other parameters used: S , T , p , Si_T , remain the same, as do the various equilibrium constants).

In addition, TA:DIC is also an excellent proxy for the carbonate ion concentration in such plots. From a number of biological experimental studies, we know that carbonate ion concentration is the appropriate driver of the calcification process for many calcifying groups, although not all, rather than aragonite saturation state (Ω_{ar}), which is proportional to carbonate ion concentration provided that the amount of calcium in the solution does not change. In that way, the TA:DIC ratio perhaps represents the calcification process better than Ω_{ar} . Furthermore, using TA:DIC could also work for species in which other parameters drive the calcification, e.g. bicarbonate in autotrophic organisms, aragonite saturation state in bivalves

and H^+ flux in foraminifera. In that way, we encompass all the parameters that would otherwise influence the carbonate system and come up with a more straightforward way to express the experimental conditions, which would then enable easier comparisons among the experiments.

2. *The methods require a much more thorough description of what has been done. Some important steps are insufficiently clarified (see specific comments below).*

Response: We have extensively expanded the method section and the description of all important steps in the manuscript.

3. *What was the rationale for using GLODAP data to derive correlations between TA:DIC and Omega? Wouldn't it have been more reasonable to use data from the individual studies. Furthermore, how were temperature (salinity) differences between studies taken into account, which also affect Omega but not TA:DIC.*

Response: As suggested, we have additionally re-created Figure 8 from the manuscript where we plotted experimental data and derived thresholds. The graph shows comparable patterns of the TA:DIC vs Ω_{ar} correlation as also observed on the graph with GLODAP data. We show both images displayed in this response. We agree that this fits the context of deriving and using experimental data and explaining the implications. However, the graph with the experimental data shows a much greater spread of experimental data; this might be due to multiple causes (e.g. the use of wrong TA and DIC coupled values or the impact of temperature and salinity). We note that for each TA and DIC datapoint, the salinity and temperature specific to that data point were used to compute the saturation state of aragonite. This saturation state is plotted against TA:DIC in Figure 2. As such, displaying the TA:DIC relationship and threshold in the GLODAP context still remains valuable, also for the demonstration how the thresholds can be derived from the experimental data and be extrapolated in the field data across much larger TA:DIC gradients to demonstrate global distribution. The reasoning for extrapolating data and thresholds across the regional scales is based on the fact that we could apply the thresholds even for the regions for which we do not have sufficient or reliable data or experimental coverage, allowing to make the inferences about the OAE impact even in those regions. As such, we will continue using the graph with GLODAP data and thresholds, but with removed experimental data in the main part of the document, while we can add the graph with experimental data in the Supplementary material.

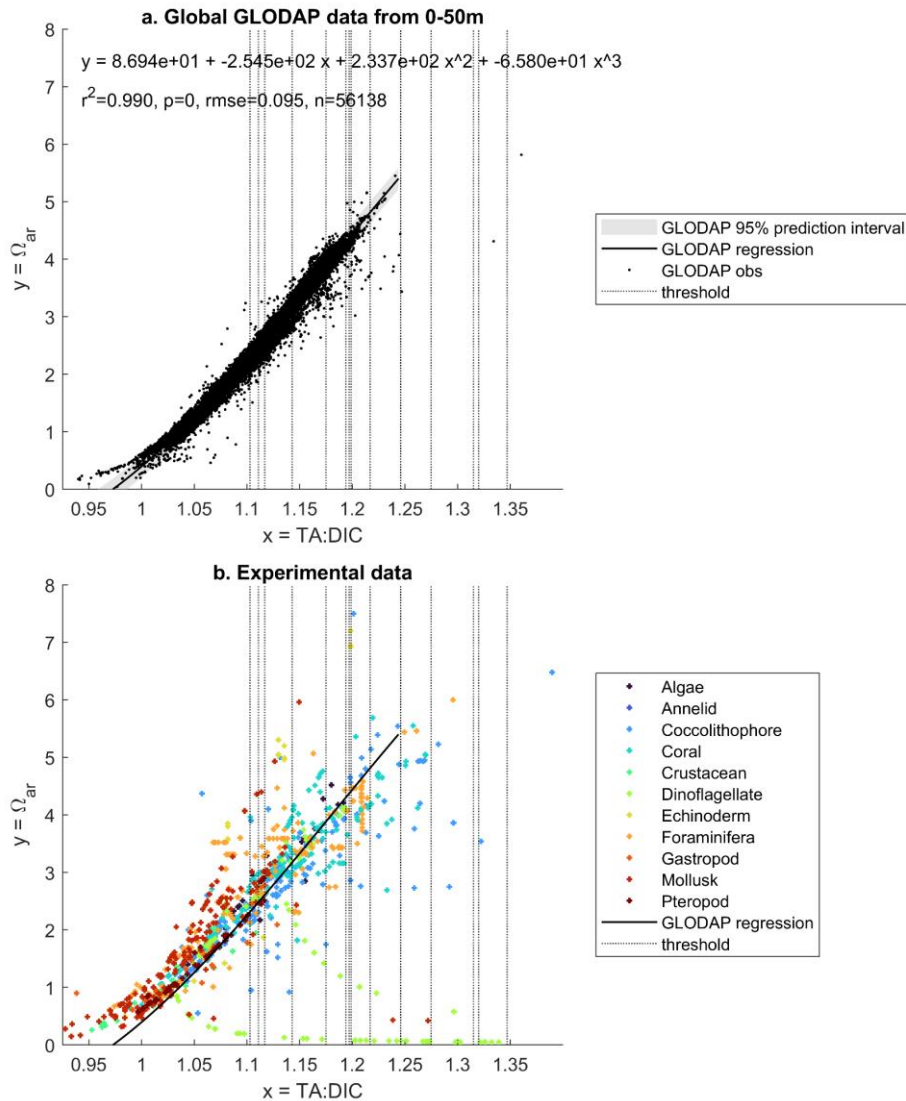


Figure 2: On top is the GLODAP data, showing the correlation between TA:DIC and Ω_{ar} . Below this is our experimental TA:DIC plotted against Ω_{ar} , with the latter being computed using experimental TA and DIC, as well as the temperature and salinity specific to each data point. The GLODAP regression line is plotted on top to show that the trend is the same, but the experimental data has more scatter.

4. The statistical approach is enigmatic to me. I am unsure if it is insufficiently described or I am lacking statistical knowledge (quite possibly the latter). Hopefully the other reviewer has better understanding and provide a robust review.

Response: [Ries et al. \(2009\)](#) use a similar method, where they chose ‘the regression analysis that yielded the lowest square root of the mean squared error (RMSE) for a given species and that was statistically significant ($p \leq 0.05$)’. When applying this method to our data, parabolic and exponential regressions were always favored over linear regressions. When examining these regressions, we found that choosing the

best fit based on the lowest p-value yielded better fits, as this method prevents overfitting to noise in the data. Therefore, we chose final responses based on the lowest p-value, and not the RMSE. According to reviewer #2, the statistical analysis seems fair.

Other comments:

Line 24: So far OAE has no relevance for climate change mitigation.

Reword to: 'OAE is gaining prominence in its ability to mitigate climate change and ocean acidification.'

Line 30: success or social license?

Changed to: 'social license and success'.

Line 40: No parabolic responses?

Not sure what the reviewer means here, the parabolic response is mentioned.

Line 43: What does realistic refer to here? That the conditions in most lab studies are not realistic wrt perturbation magnitude? Also, what would be a realistic perturbation. (I agree with your point but think this is not thoroughly backed up at this stage).

This sentence has been removed to avoid confusion.

Line 46: It is totally unclear at this stage what the TA:Ca ratio is a proxy for and why it is helpful. This must either be explained or taken out of the abstract.

We have removed this from the abstract.

Line 48: Unclear what framework you are referring to. TA:Ca? Needs specification.

Added: 'framework based on TA, Ca and calcification rates that...'

Line 69: CDRs is a weird plural. Removals?? Do you mean marine CDR methods?

'CDRs' -> 'CDR'

Line 78: Some OAE methods are well beyond concept stage (see Eisaman et al., 2023)

Changed to 'despite mostly being in the concept stage', and Eisaman et al. (2023) added as a source.

Line 100: adaptation or acclimation? (I think you mean the latter).

'Adaptation' changed to 'acclimation'

Line 109: Alkaline or "higher pH". Strictly speaking OA still investigated alkaline conditions wrt pH.

'Alkaline' -> 'higher pH'

Line 121: Unclear how a systematic framework should help here. Vague term.

Changed to: 'This study aims at systematically evaluating species responses under OAE influence, placing them into a framework of categories based on calcification rate responses.'

Line 139: The sentence implies that massive applications will happen anyway, in which case the environmental assessment before would implicitly have no influence on whether they are implemented.

Changed to: 'delineate what experiments are most urgently needed to fill knowledge gaps before massive OAE field implementation can be considered.'

Line 153: Based on what criterion were studies selected? Were all studies selected that were found by browsing? Or the first X hits?

Added: 'For several functional groups data was easy to find (algae, corals, foraminifera, mollusks and dinoflagellates), so no new studies were added after 10 to 15 studies were found. Around five studies were found for the coccolithophore, crustacean, echinoderm and gastropod groups. Only one study was found for both pteropods and annelids.'

Line 153: Were temperature differences between treatments in the OA studies considered?

For CO2SYS calculations, a temperature of 20°C and the average salinity per species was used. To calculate the effect of this choice on our results, we computed CO2SYS calculations up until an addition of 500 µmol/kg for both NaOH and Na₂CO₃ for the maximum and minimum temperature per species, as well as the maximum and minimum salinity per species and compared this to the results for T = 20°C and average salinity. The largest difference for each combination of maximum/minimum salinity and maximum/minimum temperature compared to our original results for an NaOH addition of 500 µmol/kg. pH had an average uncertainty of 0.12 and a max of 0.35. For Ω_{ar} , the average uncertainty was 0.30, with a max of 0.12.

Line 242: Unclear how NaOH was added to the TA:DIC ratio.

We believe this is thoroughly explained in lines 216 - 228.

Line 242: Unclear why a parabolic response is by default a negative response. This requires additional justification.

It is classified as a negative responder because upon a certain amount of alkalinity enhancement calcification rate will decrease. Rewritten as: 'Negative responders: species with predicted linear negative, parabolic and threshold negative response in calcification rate upon (a certain amount of) TA addition.'

Figure 2 is unclear. Is this the entire dataset? Or a specific subset of data from various species. It is also unclear if each datapoint is a treatment level from an individual study.

Legend updated to include green dots, labeled as 'Experimental data', and caption updated:

Figure 2: Conceptual diagram to show how experimental data (green dots), predicted values at various additions of alkalinity (stars), the regression line and prediction error margins are fitted for a given species a) linear positive; b) linear negative; c) parabolic; d) exponential for threshold positive; e) exponential for threshold negative. The uncertainty interval indicates four standard deviations. The red line indicates zero net dissolution (calcification rate is equal to 0; dissolution rate = calcification rate).

Line 270: The "not strong" correlation between Omega and TA:DIC basically underscores that this metric is not representative of OAE.

This line is saying the correlation is strong between Ω_{ar} and TA:DIC, and not strong for pH.

Line 308: 98 or 96?

This has been updated: due to new studies being added it is now 99 studies.

Line 373: How do you convert a response observed as TA:DIC ratio into a delta_concentration above which thresholds are reached?

We have the absolute concentrations for TA and DIC, not just the ratio. For each addition of NaOH or Na₂CO₃, we calculate the new TA, DIC, pH and Ω_{ar} using CO2SYS. TA and DIC changes are due to a simple mass balance, and are not affected by salinity or temperature. However, for pH and Ω_{ar} these are affected by temperature and salinity. To calculate threshold pH and Ω_{ar} we normalize for a temperature of 20°C, and use the average salinity per species. This is because most experiments were done at constant salinities, but at varying temperatures.

Line 408: Several decades is exaggerated.

'Several' has been removed.

Line 412: while...while

Changed to 'which will'.

Line 446: An interesting question would be if the results here are consistent with predictions for species where predictions are possible due to mechanistic understanding.

We thank the reviewer for this fundamental question that is based on the premise that we could predict calcification responses to perturbations in marine habitats if the mechanistic understanding of the unless calcification driver(s) were available. The premise is that the mechanistic relationships with identified carbonate chemistry driver(s) are available for species, which would make predicting calcification rate under various OAE scenarios feasible. Unfortunately, for most of the species, we still must rely on empirical, single-parameter relationships, for example saturation state, bicarbonate ion concentration, to substrate-to-inhibitor ratio (SIR) (i.e. the bicarbonate ion to hydrogen ion concentration ratio). In addition, more studies have shown that using a single parameter has significant limitations and also generates inaccurate changes in calcification to environmental changes, which was recently comprehensively elaborated in the study by [Ninokawa et al. \(2024\)](#), and supported by findings by [Li et al. \(2023\)](#). Both of the studies emphasize that using one parameter only insufficiently explains the calcification process, and as such, at least two parameters have to be taken into account for more accurate calcification predictions. Ultimately, it is important to note, that as a scientific community, we do not yet have an ultimate consensus on the carbonate chemistry drivers, and especially not on getting a generalizable pattern across the groups and more research efforts need to be dedicated to this.

To address the reviewer's comment, we used a few established mechanistic relationships of calcification for coccolithophores, bivalves and corals. We used these correlations and examined it against the experimental data compiled in our meta-analyses. We used experimental TA and DIC data to calculate the parameters represented in the mechanistic response.

For *Emiliana huxleyi*, we used the experimental TA and DIC data to calculate the $[\text{HCO}_3^-]$, $[\text{H}^+]$ and $[\text{CO}_2]$ concentrations. Using the mechanistic rate equation from [Bach et al. \(2015\)](#) and the sensitivity parameters for *Emiliana huxleyi* in table 1, we calculated and plotted the mechanistic rate. We applied linear, polynomial (second-order) and exponential regressions and chose the best fit based on the lowest p-value (same method as for our experimental calcification rate data regressions). Like the mechanistic rate regression, our experimental calcification rate also shows a parabolic relationship for *Emiliana huxleyi*.

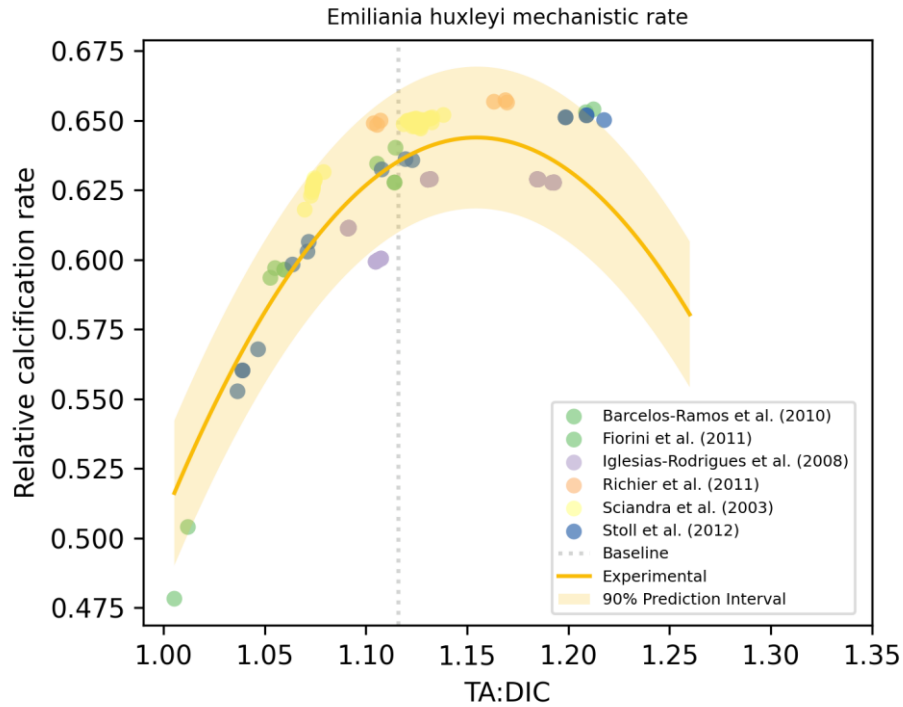


Figure 3: Mechanistic rate equation and parameters ($a = 9.56e-1$, $b = 7.04e-4$ mol/kg, $c = 2.1e6$ kg/mol, $d = 8.27e6$ kg/mol) taken from [Bach et al. \(2015\)](#) and fitted using experimental data for *E. huxleyi* (used data from the studies indicated in legend).

We also note that when using the proposed mechanistic relationship from [Bach et al. \(2015\)](#) for another coccolithophore species *Calcidiscus leptoporus*, the experimental values across the studies did not align with the proposed mechanistic relationship for this specific species; instead, a neutral relationship was obtained using experimental data. Given species difference in mechanistically explaining calcification rate between *Emiliana huxleyi* and *Calcidiscus leptoporus*, this reveals the fact that such relationships are likely very specific and dependent on a lot of parameters, with one equation not covered for different species from diverse regional settings.

Previously, some studies supported the use of the substrate-to-inhibitor ratio (SIR), i.e. the bicarbonate ion to hydrogen ion concentration ratio to mechanistically explain the calcification in marine calcifiers ([Roleda et al., 2012](#); [Cyronak et al., 2015](#); [Fassbender et al. 2016](#)). This relationship, which is tightly related to Ω_{ar} , attempted to better capture the calcification compared to Ω , however, in its recent form as a one-component parameter, it was evaluated as insufficient ([Ninokawa et al., 2024](#)). We have tested data from the experiments involving the mollusk, coral and coccolithophore groups against the SIR ratio.

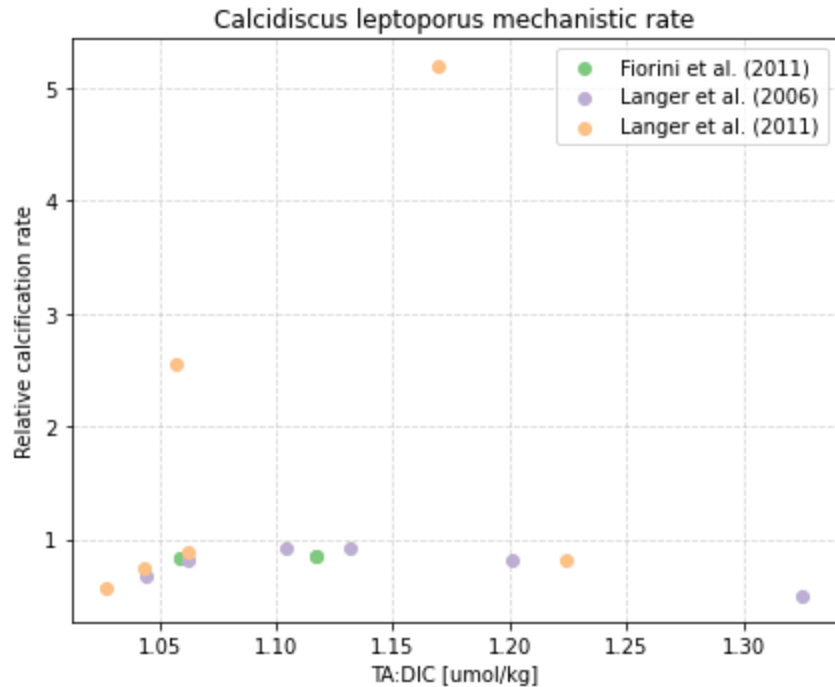


Figure 4: Mechanistic rate equation and parameters ($a = 2.31e5$, $b = 3.55e2$ mol/kg, $c = 2.19e5$ kg/mol, $d = 3.76e7$ kg/mol) taken from [Bach et al. \(2015\)](#) and calculated using experimental data for ‘*Calcidiscus leptoporus*’ (used data from the studies indicated in legend).

For the coccolithophore group, the experimental rate regressions cannot be explained using SIR mechanisms; most of the correlation predicts an insignificant response (p -value = 0.2; figure XY:). Reasons for these discrepancies could potentially be that SIR might insufficiently include various biological processes (e.g. how carbon is provisioned or the ability to regulate calcifying fluid pH), as well as salinity and temperature variations.

For mollusks, half of the mechanistic rate regressions based on the SIR agreed with the experimental calcification rate regressions, the other half does not agree, especially for the studies with experimental conditions of $\Omega_{ar} > 1$, as $\Omega_{ar} > 1$ was reported not to be an appropriate indicator of the calcifying environment (as suggested by [Bach et al. \(2015\)](#)).

For corals, using SIR regressions, the majority of coral species ($n = 21$) were classified as having a linear positive mechanistic relationship. When comparing this to our experimental rate regressions, we only found agreements with the mechanistic regressions in 8 out of 21 species.

As such, we observe that using SIR relationships to successfully describe calcification was limited to only a few species and there are no generalizable patterns that could be applicable across multiple groups, with our findings agreeing with [Ninokawa et al., 2024](#).

Line 466: The study by Albright did not show higher calcification rates but higher net calcification in a reef, which according to the authors could be due to reduced dissolution of the reef platform.

The experiment by Albright has been removed as validation.

Line 553: Unclear how the framework would be able to establish baseline conditions, which vary in time and space. If a new framework is proposed here then it should be spelled out much more thoroughly. It is currently very vague

We have attempted to clarify this framework by elaborating further: 'should create a framework in which responses are predicted and categorized, establish initial baseline conditions, identify suitable risk analyses...'