

## ***Interactive comment on “Fe(II) stability in seawater” by Mark J. Hopwood et al.***

**Mark J. Hopwood et al.**

mhopwood@geomar.de

Received and published: 21 January 2019

Two reviewers are thanked for detailed comments on the BGD text. Please find responses below (R:).

Meso and micro-cosms were sampled at different locations. The explanation of these experiments is very hard to follow. I am not sure all information is necessary, at least in the main manuscript. A simple table like table 1B is in my view enough.

R: This is obviously a matter on which there are different opinions between different co-authors and different reviewers with some wanting more detail and some wanting less. As a companion text describes the same set of mesocosms (which will be linked on the BGS website), we have reduced the method section further again but are reluctant to remove much more material.

C1

From all these experiments FeII was measured and some samples were taken to observe FeII oxidation rates in the dark. In the end the explanation of different results hardly needs the differences between the experiments, the more reason to move text on the “cosms” to the supplementary information.

R: Yes the ‘chemists’ working on this project would certainly agree with this opinion, but others would argue that the exact setup with respect to light availability, filtering of the water, adding of zooplankton etc are all important details which may have unintended effects on Fe biogeochemistry and thus feel that it is important to include them.

The authors missed one important publication Rijkenberg et al., 2006. In this paper the influence of organic ligands of FeIII and FeII on photo-reduction and oxidation of Fe are studied. In the present text it is assumed to be important, I agree, and the Rijkenberg publication can certainly help here instead of referring to papers where speculation on this subject can be found.

R: We had missed the relevance of the Fe(II)-PPIX work in this paper because the manuscript primary concerns Fe(II) formation rates from Fe(III)-L species, but the comment on Fe(II) production in the dark at the end of the text referred to are indeed very useful and compliment the comments from Reviewer 2 concerning the possibility of ‘dark’ Fe(II) production from a superoxide driven redox cycle. An additional paragraph is added in the discussion under the topic of ‘dark’ Fe(II) production: “Apart from the influence of organic Fe(II) ligands on Fe(II) stability arising from the slower oxidation rates of some complexed Fe(II) species, Fe(II) binding organics may also have a role in the generation of superoxide which is speculated to be a dominant mechanism for the formation of Fe(II) in the dark. Experiments with 65-130 nM of protoporphyrin IX demonstrate increased formation of Fe(II) in the dark with both increasing porphyrin concentration and increasing irradiation of seawater prior to the onset of darkness (Rijkenberg et al., 2006). Whilst the rates of this process are challenging to investigate at the sub-nanomolar porphyrin concentrations expected in natural seawater, the dark formation of Fe(II) mediated by ROS interactions with Fe(II)-organic complexes could

C2

potentially be important in both the diurnal cycling of Fe in the surface ocean and the non-photochemical formation of Fe(II) in the dark of the ocean's interior (Rose 2012). From a mechanistic perspective, it is difficult to establish from the experiments here in whether apparent Fe(II) stability arises from reduced oxidation rates due to Fe(II) complexation, or dark Fe(II) formation via a mechanism, such as that proposed for superoxide, which involves Fe(II)-organic complexes."

One of the reasons that the text is hard to follow is the use of different names for the same locations/experiments (there are 5 locations: the Arctic (Svalbard), the Mediterranean (probably Crete), Patagonia, Gran Canaria and Kiel (mentioned only once and not in methods, but important for the discussion)). Since the setup is so complicated it is even advisable to mention experiments in the same sequence (helping the reader) and not as in the example given changing the sequence in one sentence: "MesoArc than for MesoPat (MesoPat R2 0.0022, gradient  $0.0049 \pm 0.014$ ; MesoArc R." etc The introduction tells the reader that she/he can expect mesocosm experiments in three places. Then section 2.1 starts and experiments were done in 4 places en in the result section experiments in Kiel appear to have happened too. Extra names like the ocean Certain project come out of the blue. First meso is used later without any explanation on page 14 Meso Med, MicroPat MultiPat are used. These last set meso micro and multi are much better suited because they indeed indicate which kind of experiments are meant. Why not use them immediately and define them properly?

R: In hindsight this is confusing to the reader. The reason was we had used the exact terminology adopted by the projects that ran the mesocosms, but this can be confusing as 'MesoPat' was used [within the project that funded it] to refer to the field campaign (which included a mesocosm/multistressor/microcosm in Patagonia). We have therefore adopted a standardized name for each specific experiment e.g. 'MesoPat' refers to the mesocosm in Patagonia, 'MicroPat' refers to the microcosm in Patagonia etc. . . and adopted these throughout the text.

Section 3.5 is very interesting but difficult to read and understand. Half of the text

C3

should be in methods, also be more clear about the Kiel experiments (at least I suspect the spiked Atlantic tests are done in Kiel).

R: As suggested (here and later), we have moved any descriptive material including equations to the method section. Yes the 'Kiel' experiments are those with Atlantic seawater, we now refer only to 'spiked Atlantic seawater experiments' And it is not clear whether the different treatments had influence on the results.

\*\*\*R: This is a key point of the paper which we try to clarify better. We can't address this question (whether the different treatments had influence on the results) because the 'stability' of Fe(II) measured is very sensitive to the Fe(II) concentration at the time the experiment starts (i.e. the time at which seawater is moved into the dark). This is the fundamental finding of the paper (and is raised in some other points below, so I will address it extensively here). Conceptually, in simple terms, there are two Fe(II) 'pools' in seawater when it sits under ambient surface conditions – whether in a mesocosm, or not. A) There is a pool of inorganic Fe(II) which has oxidation rates exactly as predicted by experiments where Fe(II) is spiked into synthetic seawater. B) There is a pool of organic Fe(II) associated with natural organic material akin to ligands which, overall, has a slower oxidation rate than the inorganic pool. Therefore, when a seawater sample is sub-sampled for analysis and moved to the dark (in a bottle, or in opaque tubing flowing into a flow injection analyzer), Pool A decays faster than Pool B. The fraction of the total Fe(II) present as B therefore increases in the seconds-minutes after sample collection. Hence a major experimental problem; even if seawater is pumped straight into a flow injection analyzer using the best available method (a dual-loop FIA system), it experiences 30-60 seconds in the dark prior to analysis. In reality this time is 2-4 minutes due to the time required for subsampling, moving FIA lines, achieving stability with the luminol signal etc. . . During this time the fraction of Fe(II) present as Pool B increases. And because the half-life of pool A is short, the fractional importance of B can increase significantly within minutes of being in the dark. Furthermore, the fractional importance of A and B likely changes on diurnal timescales and over the

C4

experiment duration. Therefore, there is a strong bias towards measuring oxidation rates close to the calculated inorganic rates at times of day when there is a large total Fe(II) pool, and when the decay rate is measured from the exact time at which a sample is moved into the dark. Conversely, there is a strong bias towards measuring apparent Fe(II) stability when the decay rate is measured at a time of day when the total Fe(II) pool is smaller, and when the decay rate is measured from some time after a sample has been moved into the dark. Unfortunately, with any sort of field experiment it's very difficult to measure more than 1 or 2 decay rate experiments simultaneously, so the time of day when an experiment was conducted varies within each experiment set, and it's impossible to produce an experimental setup free from artefacts where the sample time/'move to dark time' is completely identical to the time at which the first sample peak is measured. There is inevitably a delay and unfortunately the delay, even if varying only from 30-60 seconds, can be equivalent to 1-2 Fe(II) half-lives for inorganic Fe(II) species. For these reasons it is not meaningful to look at the difference between treatments within or between mesocosm experiments for the Fe(II) decay rate because the decay rate is biased by the initial Fe(II) concentration for the decay rate experiment and this cannot easily be corrected for.

In the discussion I miss apart from the Rijkenberg paper, discussing the influence of the different sampling and measurement treatments and the different experimental conditions.

Is the difference in time between sampling and analysis discussed, Gran Canaria is different from the others. Can this have had an effect on the results? See also above section 3.5. What is e-microcosm (in Suppl table), this is not explained, still here the largest differences between kmeas and kcal exist.

R: A line is explicitly added to discuss the potential effect of sample acidification. We provide a reference (as in the original text) which describes this in detail. We did not conduct kinetic experiments in Gran Canaria, therefore there is no direct potential for this method -change to affect our main conclusions... "a modification outlined by

C5

Hansard and Landing (2009) which is not thought to significantly affect in-situ Fe(II) concentrations during the short time period between collection and analysis". We also note that in warm seawater there is simply no alternative as the half-life of Fe(II) limits the ability to do any analysis on unamended seawater. 'E-microcosm' was the official project name for the 'MicroPat' experiment (hence comment above, the official project 'names' are difficult to follow, so we have standardized and amended throughout). Yes this experiment shows the highest  $\Delta k$ . But, as noted above (\*\*R) it is very challenging to claim this due a specific biogeochemical phenomena. Detailed comments Sections 2.1 and 2.2 are hard to follow Sentences like "Note that previously a series of experiments in the Mediterranean ('MesoMed') was also included." do not help. If it was previously, why do we bother here.

R: As raised by another reviewer, it is important to note that we attempted these measurements but do not present data (because every single measurement attempted was below detection) to avoid miss-reporting of our findings.

Line 7-9 page3 section 2.1 seem out of place, this has nothing to do with setup and sampling.

R: Moved to results section

Line 13: 10 identical 1000-1500L tanks, 5 tanks got zooplankton. According to table 1A they all received copopods but the addition was different per location.

R: No, table 1A simply lists the variables which were manipulated in each experiment. Table 1B gives the specific treatment for the high zooplankton tanks. We clarify in table B that the treatment is the zooplankton added to the 'high grazing tanks' and not the baseline for all tanks (which was zero, the tanks were filtered through a mesh, and then zooplankton were re-added to 'high' tanks only).

I did not find figure S1, below text and pictures, there is a caption but no schematic figure of the experimental design.

C6

R: There is a table for each experiment matrix. Figure S1 is re-labelled 'Supplementary material' rather than a 'Figure'

Line 26: can bags stand? Are bags mesocosms? In the next line the word tank is used, is this still the same thing?

R: One of the mesocosm experiments used bags (Gran Canaria), two of the experiments used tanks (MesoPat/MesoArc). 'mesocosm' is used widely within the field to refer both to mesocosm experiments (i.e. the whole experiment), but also to each unit within a mesocosm experiment, hence why we try not to use both meanings in the same sentence. We have rephrased throughout now using only the term 'mesocosm experiment' to refer to a whole experiment.

Section 2.2 What is a 10-treatment?

R: an experiment with 10-treatments.

Section 2.3, line 26 after cleaning, what happened with the bottles? Were they stored empty or filled, if so with what.

R: They were stored 'empty'. Now stated explicitly.

Page 7: were the Fell bottles dark plastic?

R: No, to move Fe(II) samples from the experiments to the FIA (which took 1-2 minutes), we opted for transparent containers so the water would remain exposed to ambient light. It was then transferred into a dark box as stated in the text at a time recorded as 'time zero'. We rephrase and clarify this sentence. In Gran Canaria, where the Fe(II) samples were acidified, the bottles were opaque to prevent any further photoproduction of Fe(II) (also clarified in methods).

Line 9: Ocean certain is? It would be so much easier when the normally used names are used here too, meso/micro-Med-Arc-Pat.

R: We now use the terms Meso/micro/multi-Med/Arc/Pat throughout.

C7

Section 2.5: What happened in Kiel? Line 18 tells us what happened in Patagonia and Svalbard (? Pat and ARC, Comau fjord- Kongsfjorden?)

R: we now explicitly add the names (meso/micro-Med-Arc-Pat) when referring to any fieldsite and refer to the laboratory spiked experiments consistently as spiked seawater experiments.

In 3.5 79 experiments are mentioned, that info belongs here. How many per location. Were they kept under ambient temperature, where is the laminar flowhood

R: Details added to the methods section. All of these experiments were in temperature controlled rooms with temperatures as per the collected water for analysis. Page 8 s2.5 re-written accordingly, ... 'All Fe(II) decay experiments were conducted inside the temperate controlled rooms hosting the MultiPat/MultiArc experiments. As such, a constant temperature was maintained throughout these experiments. The FIA instrumentation was arranged with the inflow lines under a laminar flow hood inside the temperature controlled rooms. ...'

3.1 why use the name Svalbard here and not MesoArc.

R: Ammended.

Page 10 line 11, no glucose is mentioned in Table 1

R: No, this was only in the text, now amended.

Page 11 line 9: curiously..Why.is this curious, and why give relations that are not relations, for figures 2 and 4. For figure 4 it is not clear which line belongs to which mesocosm experiment. No idea what the journal guide lines are, but especially as in figure 3the './' is confusing like there is a ratio instead of the unit.

R: Because we can't think of a simple reason why there should be a strong correlation between DFe and TdFe in one mesocosm, but no correlation at all in another, especially when the strongest correlation is for a low organic carbon, high particulate Fe

C8

site. Figures 2 and 4 are removed to save space. Units are changes on the graphs throughout.

Page 13 line 1 as per?

R: There are obviously multiple manuscripts in preparation from this series of meso/multi/micro experiments. One concerns H<sub>2</sub>O<sub>2</sub> (in BGS, the manuscripts will be linked as companion papers so the link will be more obvious).

Page 14: Meso Med, MicroPat MultiPat: have mercy on your reader! Suddenly new abbreviations. However, useful abbreviations that should be used throughout the whole manuscript

R: We have now adopted our own standardized terms for the experiments as per previous comments.

Lines 9-10: Is that to be expected? Reference needed here (also at line 16)

R: Yes, this is discussed in the next paragraph (in the original text). A Reference is added for Fe:P lithogenic material and high TdFe at this fieldsite.

9-16, a lot of different names for the same sites

R: Yes, but each mesocosm obviously has a corresponding fieldsite so we inevitably have to name the experiment and the site. For clarity we annotate the place names with the experiments hosted at that location.

29: chlorophyll a *Italic* Figures 5 are too small. Labels and legend are impossible to read and the sequence in the legend is not logic, one legend might be enough for 5a and 5b. Perhaps make 5 c a separate figure. Be careful with ratio's. The high values are they due to low DFe? Figure 5 c is not mentioned in the text.

R: Amended so the figures display better whilst merged in the text. The legend sequence order is changed. C is shown separately and we now show the concentrations and the ratio (DFe is never particularly low, so no these are not an artefact of sub-

C9

nanomolar DFe concentrations).

Line26: I do not know whether there is a general decline as claimed here, the 1450 microatm perhaps does decrease but the 1300, 700 and 1150 microatm do not, so no general decline here. This figure is not suited to make such statements. An increase between days 20-29 ok.

R: The significance of changes (other than the increase after day 20) is variable between treatments, as this doesn't really affect our conclusions we slim the text accordingly and remove these lines.

Page 15: line 6: what is number 7?

R: Changed to PCO<sub>2</sub> value for this mesocosm.

Lines 6-9: not clear what the authors tell here? Is this still about figure 5b? The sixth mentioned number does not show an increase, not the seventh. They have different CO<sub>2</sub>, haven't they?

R: We clarified this 'number' (above) referred to a specific treatment. It is the 6th number because one mesocosm leaked and was removed from the experiment (this is all clarified by just referring to the treatments by PCO<sub>2</sub> target level throughout rather than arbitrary treatment labels).

Page 16, Line 11: which variation, give reference, do not force the reader to search in another of your papers to find out.

R: The specific variations compared were a dual-loop configuration as described by Croot and Laan (2002) and a pre-concentration method as described by Bowie et al., (2002)

Lines12-14: it depends what you mean with in situ and what you want to do with the k-values. With such an uncertainty one can wonder whether waiting for stabilisation would have been wiser.

C10

R: But then any Fe(II) present would have decayed and we would have to conducted spiked experiments, which wouldn't tell us much about stabilization that we didn't already know. This error is not large considering comparable data in the literature considering what reported 'errors' actually include. Note the exact opposite query is raised by reviewers with Sarthou et al., (2011) BGS. In this excellent 2011 manuscript the authors added small Fe(II) spikes to seawater in order to determine oxidation rates and it was questioned (see comments/discussion with reviewer 2 on that text) whether this approach was meaningful compared to observing in-situ decay rates. Hence our rationale for the setup herein.

Lines 20 onwards: is this what happened in Kiel? Most of this belongs to the method section. Also the equations belong to the method section in my view.

R: Yes, we have moved equations and experiment descriptions to the methods section.

Page 17:I can add a few hypotheses: the aged Atlantic water was probably filtered, in any case no phytoplankton or copepods were present. The added Fe for certain saturated the organic ligands and thus this DFe was in an inorganic form, a colloidal or amorphous Fe-oxide or hydroxide. This is where the equations 2 and 3 were made for: inorganic Fe! So certainly this is other chemistry. I advise to read Rijkenberg et al., 2006 *Geochimica et Cosmochimica Acta* 70 (2006) 2790–2805; Enhancement and inhibition of iron photoreduction by individual ligands in open ocean seawater.

R: Yes, these experiments represent inorganic speciation and this is exactly what we summarize in hypothesis III. Note the experiments concerning ligand saturation in the Rijkenberg 2006 text concern Fe(II) formation from Fe(III) species, a slightly different issue to the Fe(II) decay discussed here.

Page 18:line one, why would low FeII be the most stable?

R: See \*\*\* comment above; because the fraction of this Fe(II) existing in a stable form is likely higher.

C11

Discussion line 18: This can be read that TFe behaves conservatively. Why would DFe-TFe be linear, that is a strange idea. That is assuming all particles have the same properties.

R: In these near-shore waters where a large fraction of Fe comes from a near-point source (e.g. the freshwater outflow into the Pat/Arc fjord sites) it is not an unreasonable statement that all particles have the same properties, note the relative consistency in XRF data, simply because freshwater derived particles account for the vast majority of Fe-rich particles in the water column. On short timescales TdFe does behave conservatively, unlike the rapid removal of DFe in these nearshore environments, a TdFe/S plot is linear showing that the sinking/modification of TdFe takes longer than the residence time of water in these fieldsites.

Page 19: table 4 why is mesopat meso and multistressor so different, this is not discussed. Why is the sequence different, why Svalbard whereas it is Arctic. The different names makes it more difficult to understand.

R: Noting the large standard deviation on both Fe(II)/DFe and DFe/TdFe, it is not clear that they are 'so different'. We clarify that the 'ambient' measurements are at the fieldsite. But as the 'ambient' measurements don't refer to any of the experiments at that fieldsite, they need a separate name. For clarity we sate 'Arctic (Svalbard)'.  
sate

Lines 23-24: why was this not mentioned in the method section?

R: The exact timing of the experiments is shown in the data table appended to the paper and thus we don't feel it necessary to write it out in the methods section. The aim of these experiments was to investigate how Fe(II) decayed, not to produce high resolution Fe(II) time-series across the duration of every mesocosm.

Page 20 line 8-9: thus what is the conclusion?

R: A line is added, 'The ambient concentrations of Fe(II) measured in Patagonia (Comau fjord) and the Arctic (Svalbard, Kongsfjorden) at the mesocosm experiment field-

C12

sites are therefore not necessarily directly comparable to Fe(II) concentrations measured after nutrient addition in the mesocosm experiments.' 4.3: line 16: according to the methods section artificial light was used in micro and multistressor but not in Mesocosm, so why mention artificial light here?

R: The text states 'due to the enclosed HDPE mesocosm design and/or synthetic lighting'. The point being made was that all of the mesocosm/microcosm/multistressor experiments where Fe(II) decay experiments were conducted had low H<sub>2</sub>O<sub>2</sub> concentrations.

Lines 21-25: read Rijkenberg, they saw the influence of ligands on Fe redox, of ligands binding Fe(III) and of a ligand binding Fe(II). That should be added in the discussion here.

R: Photochemical formation of Fe(II) from Fe(III)-L species is not relevant to the discussion here. The specific points concerning dark Fe(II) formation from porphyrin are however interesting and added at the end of this section as per some earlier comments.

Page 21 decay rates in the e-microcosm are different from the calculated k compared to the others, apart from low Fe(II) at t=0? (low Fe(II) occurs also in other experiments) what is e-microcosm, what is different? Could that be an extra reason. Use the work of Rijkenberg et al in the discussion on page 21, they did not assume Fe(II) ligands, they used one in their redox rate experiments.

R: (E-microcosm is the data label for the MicroPat experiment, we have changed this in the text to 'MicroPat' as per our standardized names). There is no obvious experimental difference between the MesoPat/MicroPat/MultiPat experiments that immediately provides an easy explanation for why the largest changes in K should be reported for datapoints from one experiment. There may of course be species-level effects due to the different biological communities at the start of, and throughout, each experiment. Yet, as we note (R\*\*\*), because the discrepancy between measured and calculated K is very sensitive to the Fe(II) concentration at t=0, and because it is not (using the design here) possible to rigorously standardize t=0 so that [Fe(II)] at t=0 is constant, or

C13

to account for the change in Fe(II) concentration and speciation between in-situ conditions and t=0, it is very difficult to deduce any relationships between biogeochemical parameters and the difference in K.

Excel file temp in k, make capital, add start or initial also to the column name for Fe(II). The precision does not warrant the decimals shown with 35% uncertainty. What is an e-microcosm, why are the rates so high here. Add measured to k. No Kiel experiments here?

R: Amended. (See above comment also). 'Kiel' (spiked Fe(II) decay experiments) data is now added to the supplementary file as per the Meso/Micro/Multi data.

---

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2018-439>, 2018.

C14