Both review texts:

REVIEW 1:

Dear Daniela,

I have reviewed the manuscript titled “Effect of chemical abrasion of zircon on SHRIMP U/Pb, δ18O, Trace element, and LA-ICPMS trace element and Lu-Hf isotopic analyses” by Kooymans et al. This manuscript presents a major effort to report U-Pb, O and Lu-Hf isotopic systems and various trace elements including OH of zircon, to show if any effect is observed by chemical abrasion (CA) treatment.

It is indeed a very interesting subject, one that I’d like to see published in the near future, but not in the current form. My biggest concern is that the manuscript could not focus on solving the main problems (which were not clearly defined), but instead distractedly explained too many things and presented data that are not essential. In my opinion, the main argument in this paper (in the other word, the most novel component of this manuscript that zirconologists must be interested in) is that CA could produce poorer surface condition on zircon that possibly affect the sputtering process and change U and Pb ion emissions, resulting in scattered SIMS U-Pb data for moderately damaged zircon.

The introduction is rewritten to more clearly state what the paper is about.

The authors did not discuss this well enough but continued discussing other isotopic/elemental data and MPV zircon core data. The other data unchanged by CA are interesting, but unfortunately distracting in the current manuscript, together with the poor writing.

The overall quality of the writing makes the manuscript challenging to navigate. Even the introduction lacks a clear focus on a main purpose, and the motivation and aims of the study are not distinctly described. The manuscript could be significantly shortened by removing repetitive and unnecessary sections.

Done.

Many paragraphs consist of only a sentence or two without a clear main idea, indicating a need for substantial improvement in paragraph construction. The overall structural organization is notably lacking, contributing to a distracted and unfocused narrative. A more focused and concise writing style would immensely enhance the manuscript.

Furthermore, many sentences in the manuscript appear as guesses or the authors' opinions without thorough discussion or supporting references or evidence. Oftentimes, conclusions, interpretations, or the significance of these statements are missing entirely. The manuscript lacks a clear logic build-up that is necessary for constructing a reasonable and convincing argument. Throughout my reading, I had to keep asking – "WHY?" or "SO WHAT?". I had to spend much more time than necessary to make a guess what the authors meant. I still do not comprehend many sections and may have misunderstood some parts.

Agreed, and rewritten.

I strongly recommend a comprehensive rewrite of the manuscript, with a narrowed focus on SIMS U-Pb data (possibly with O). Considering my suggestion, non-SIMS sections are not much commented below. Expecting a thorough rewriting, I may not provide detailed comments on every aspect I have noted.

Please consider my review anonymous.

Specific comments

McKanna et al (2023, GChron) conducted a comprehensive surface study of CA-zircon, and mentioned "sponge-like texture" in CA-treated high-damage zircon samples. They also reported that "acid …regularly accesses crystal cores to dissolve … interior zones", implying that the zircon core also can get the sponge texture if it is highly damaged. That is readily critical to SIMS analysis, which is a very good motivation of this manuscript. McKanna's work is a key paper to introduce the aims of this study.

Addressed in discussion.

Scattered U-Pb data observed from CA-treated MPV zircon rims unquestionably raise concerns about the reliability of SIMS U-Pb data for CA-zircon and necessity of CA in any SIMS work. Two hypotheses were introduced to explain it, but they were only briefly addressed in a sentence or two, L603-610, requiring further elaboration and explanation.

Addressed in discussion rewrite. Obviously further work is required.

Additional session to re-run the zircon grains where the scattered U-Pb data was a good idea, although would have preferred a more comprehensive double-check, including all the same grains rather than just those deviating from the average U-Pb dates – so that you can properly see if the data are less scattered with >0.05 probability. And the conclusion that sputtering difference on CA-zircon surface (nanospongeform - not visible at all with high mag SE?) is reasonable. But again, I'd like to see more comprehensive discussion about this conclusion.

The argument that CA of moderately-strongly damaged zircon yields poorer SIMS U-Pb data should be consistently maintained throughout the manuscript, not trying to interpret CA inherited zircon data.

Both the discussion and the introduction have been expanded to address these issues.

Without considering the possibility of worse SIMS U-Pb data after CA-treatment more, the authors just kept discussing about the MPV core ages in the geological context, which is insignificant because the MPC core age data are highly likely less reliable than MPU. Need discussion if the CA reaching depth into zircon is throughout the core area too (yes, according to McKanna et al 2023). The probability density diagrams of CA-treated and untreated MPV cores – the latter looks a lot sharper, which may be consistent with the possibility of MPV CA cores provided more scattered U-Pb data, same as on the rims. + I'd like to see the diagram of the source rock if you can specify it.

We have added concordia diagrams for the complete core + rim MPU and MPC data plotted on the same plot for better comparison.

I also found the higher ratio of the magmatic age (430 Ma) out of total number of core analyses potentially misleading and prone to misinterpretation. 6 vs 18 looks obvious difference, but I believe it may be due to a target selection bias for analysis.

Target bias would work the other way, as we were looking for inherited cores, and they are easier to see in MPC

As clearly mentioned and illustrated in the CL figures, CA treated MPV zircons have obvious etch channels. It is highly possible that the authors selected better-looking zircons, avoiding visible cracks, inclusions and possibly etch channels too (even unconsciously). As the authors mentioned, the 430 Ma-cores have the least evident etch channels around, making them more likely chosen for anlaysis compared to those surrounded channels (possibly older). What is the ratio of measured cores that show obvious etch channels? - I could identify max 10 cores analysed in the CL image. Additionally, I cannot understand "survivor bias" (L589) explanation.

Those above are the main comments to the SIMS U-Pb data. Some others follow:

For sample selection: No indication of density or magnitude of damage of zircon? It reads like the authors tempted to use U content as a potential indicator in the section 1.2, but the connection between U content and the level of damage is not explained. In the similar regards, it is not clear why MP volcanic zircon was chosen for this study. Clarifying the rationale behind choosing this specific type of zircon would enhance the reader's understanding of the study's context and objectives.

The MPV were chosen as the simplest 2 component system (sedimentary core + igneous rim) we had in our collection. This is addressed in the introduction under 2.1 Sample Selection.

U-Pb internal/external error: you always should report the U-Pb dates with external error propagated, as the published dates and ages are going to be compared to the other ages. Not sure why the authors consider reporting the internal errors only, all the explanations about which make the manuscript at least 1-2 pages longer.

Internal errors are correct when comparing data from the same session with each other, to look for subtle instrumental differences. External errors are correct when comparing results to reference values. We do both in this paper, which is why we take the time to explain them.

Untreated TEM2 was used as a primary standard for U-Pb (with two other untreated zircon references) for running both CA-treated or untreated unknowns. This kind of test should be done under the same condition, so it is a reasonable question that any other readers might have. What if the U-Pb calibrations are different between CA-treated and untreated zircon? You showed TEM2 has same U-Pb calib curve slope for both preparations in the results, but it was supposed unknown during the analytical strategy stage (and QGNG and OG1 show different calibration slopes, although the authors consider it trivial; I'd like to see the calibration curves illustrated in the manuscript). If it was deliberate to see how untreated standard affects to CA-treated unknown zircon U-Pb calibration, it should be mentioned.

We have removed calibration slope discussion at the suggestion of the editor, as it is not the main point. We present data using both T2C (figure 4; table 2b) and T2U (figure 3; table 2a) as the primary reference material, so that the differences can be compared.

CommonPb correction using 204Pb. L265: "…as 204Pb overcounts were within uncertainty of zero…" …? What else would you do if it is above zero then? And why don't you use 207Pb-corrected 6/38 age for MPV zircon dates?

Introduction rewrite specifies that we are looking for consistent data treatment for all samples, not the most precise treatment for each sample depending on its age.

Many unnecessary – examples:

L413: "…did not run as smoothly…" unless you are going to point something out for the less smooth session, it is super unnecessary. Every session has its own condition, and you cannot compare them all the time. L271: Truly unnecessary. Calibration slope could be different even session to session on a same instrument.

We have removed calibration slope discussion as tangential.

Many unclear – examples: L24, "…the analyses of chemically abraded materials show excess scatter" OF WHAT?

Rewritten.

L107: "These volcanic zircon rims are also lower in U content than S-type granite rims, which often go metamict…" S-type granite "zircon" rim does not have constant U concentration. Which specific S-type granite do you mean? And what is the U content of the MP volcanic zircon? How do you say something is lower than something without showing their data? What is the U concentration threshold to go metamict? Any reference for that?

We are referring to LFB Silurian igneous rocks- e.g. the most likely intrusive equivalents to the MPV. Rewritten for clarity.

L149: "We also include four new aliquots for OGC…" Why? Then they are newly measured not from the literature data? And the new data should be properly reported.

Intro rewritten to explain that this is the missing aliquot level data from Bodorkos et al. 2009 which is finally being provided.

Result-discussion mixed up: L287-9, that is one of the aim of this study, not method

L502: "… reduces the uncertainty in the reference ages by 140-290%..." According to Table 1, ratio uncertainty except 91U (why is that?) is certainly smaller including tracer, but not age uncertainties. How come? Age uncertainties are actually larger including tracer.

Table 1 was misprinted. Fixed.

L541: "…, which is consistent with our SHRIMP U-Pb data" How and what exactly?

Rewritten for clarity.

L549-551: Not sure.. Fig 2 shows QNGN and OG1 have different calibration slopes between U-C

Removed as tangential.

L552-554: self-calibrated? What about the other standards data using the low-sloped 91500 calibration?

Removed as tangential.

L556-560: that is a pure guess. If you want to say so, you need more examples of low Hf & REE zircon data. Do not argue anything without supporting evidence.

Removed as tangential.

L572-578: Why S-type granite only? Slow(er) crystallization is a general condition to form granite than volcanic rock. And common Pb is not only from Pb-rich inclusion. "For comparison…": comparison of what and what? Why the granites from Bodorkos et al (2015)? Apparently same

igneous ages (~430 Ma), so are they potential co-genetic bodies? There is no explanation. "… statistically significant common Pb": which means…? "…raw 207Pb/206Pb ratios…": raw ratios? Same as total 207/206? 207, 206Pb are mostly radiogenic Pb in zircon anyway? "…unusually high common Pb contents" I think it is not uncommon to see common Pb? Apart from many question marks about every single sentence in this paragraph, I do not see why this full paragraph is necessary.

Rewritten for clarity

L582-3: show the data. Compare the size of uncertainties or etc.

Figures added to paper.

L608-9: Uh… probably MPV zircon is more damaged than the reference zircons?

Could be.

L613: total 206Pb… why?

To constrain potential diffusion gradients (which were not observed).

L617-: I don't get what you try to explain in this paragraph.

Removed.

L627-633: "This is consistent with…" "consistent" is to compare to the others' arguments or consensus. To me, this paragraph is a guess, with no other supporting arguments. From the rim data (MPU, MPC), I am convinced that the CA method weakens the SIMS U-Pb validity and think the MPC core dating is less reliable due to the CA-induced surface damage (nanospongeform – according to the authors' description, L607).

We have rewritten much of the U-Pb introduction and discussion to explain why this is generally not the case.

L650-: That is bizarre to publish. I strongly suggest the authors to re-run a session for Temora2 using a same batch. What if the huge d18O difference after CA is true?

We would love to as well, but we will not be able to do this until at least 2025 due to scheduling commitments, and this work is already 7 years old…

L654: MPV zircon scattered d18O may indicate that the source rock/melt was not in equilibrium. Not all S-type granites show scattered zircon d18O!

Rewritten to compare to Silurian Lachlan Fold Belt samples, not worldwide S-type granites (See Vogt et al. (2023) for chemically abraded Variscian S-types).

L657-8: what kind of evidence do you need?

Rim d18O does not seem to correlate with core d18O.

L658: Ti thermometry? Reference? Did you perform the measurement?

Reference is given in the table caption and the methods, and measurement results are in the tables on a sample mean bassi, and the supplementary data on a spot-by-spot basis.

L662: "… not all … volcanic zircon cores are detrital" They never are.

Within the Silurian LFB they are nigh ubiquitous.

L668-: this paragraph should to go the result section

Fair.

L677-8: I don't see significant difference in Fig 7 except OG1, which look plotted wrong

Supplementary data table and methods corrected to explain.

Technical corrections

Supplemantary tables:

Check all the table format (not only for the suppl tables)

For what?

All negative values should be corrected to zero or "-"

We disagree, due to issues arising from biasing averages to positive values.

Th/U not a ratio of isotopes (232Th/238U) but total Th/U estimate

Reasonable effective number and decimal places

Oxygen data table only show internal SE (standard error), which should be replaced by fully propagated uncertainties. (or if you argue that no external reproducibility is not necessarily considered because only within-session data are compared, it should be explained in the text)

Figures:

Fig 1: it is useful to have some indications to figure our easily which cores are syn-eruptic and which "cracks" are the etch channels.

Full annotation is in the supplementary figures.

Fig 4: readers would like to see the Concordia plots for all standard/samples.

Done.

Fig 11: CA treated zircon data should be included too.

Done.

Other comments

It does not read a formal scientific manuscript, rather oral speech script or a personal journal in many aspects – use of informal words (best example is SHRIMPing), use of subject adjectives without giving numbers, lots of mistakes, sentences scientifically incorrect…

Lots of mistakes: typos, wrong capital letters, use of spaces, inconsistent expression (U-Pb vs U/Pb, use of acronyms), etc

Always make it clear what uncertainty you indicate – 1se, 1sd, 1s (65% conf) or 2s/ts (95% conf) etc.

Checked.

SHRIMP – a specific brand name, but a method. Change it to SIMS except where specific SHRIMP IIe and SI are described.

Checked.

Scientifically incorrect or text-data-plot not matching:

L89, 207/206Pb ratio is almost constant: not necessarily especially for old ones (in that regards, I actually would like to see Pb/Pb data too especially for OG1)

Shown in new plots.

L91: "… SHRIMP can produce 7/6 age with 2‰ precision (really meant permil not percent?)" It cannot be a general comment as 7/6 age precision depends on the zircon age and Pb concentration.

Checked.

L98: "…old enough for decent counting statistics…" Again, not necessarily. Counting statistics of U-Pb dating depends on U concentration too.

Sure, but high U zircons come with their own problems, which are beyond the scope of this study. Try Magee et al. (2017) for discussion of the hi-U effect..

L512: "a piezo stage, automated analyses" why should they be conditions to get better precision?

Better X axis reproducibility during automated analysis.

L513: "…better than the 1-3% value…" Better? you see the exactly that range from TEM2 and 91500 in Table 2 (and larger for the other standards)?

Table two shows an external 95% confidence interval of 1.35 to 1.8 Ma, out of 417 Ma = 0.43%. This is less than 1%. It shows an external 95% confidence envelope for 91500 of 3.6-5 Ma, out of 1063 Ma = 0.47%. These numbers are less than one percent.

L535: you do not have DR12 and 13 figures; if they are the ones from Magee et al., those fig numbers are not necessary or you need to make it clear.

Clarified.

L554: "For a difference in slope of 0.5, this would yield ages 0.2% older…" Calibration slope does not necessarily increase or decrease ages, as it is depending on where the unknown data sit compared to the standard's calib curve.

Removed for clarity.

Fig 7E: OH/18O values of OGC are all >0.1 in the suppl table, but they are plotted mostly <0.1 in thei figure.

Supplement fixed to show both raw and corrected data, and methods updated to explain correction.

Fig 12: The repeated session is 210046? It is 170124 session in the fig

Clarified.

REVIEW 2

This manuscript is dealing with the topic which all geochronologist would like to know the answer. Careful design of whole experiment as well as precise analysis using various instrument (TIMS, SHRIMP, LA-ICP) should have a huge contribution and implication on the area of geochronology and geology. This result could be cited many times in future and be possibly mentioned in the textbook, too.

However, the manuscript itself is not very straightforward for reader to follow and understand. Most materials of tables and figures are not ready to be published yet. Significant revision in both text, table, and figures are necessary in text and table and figures for the next step. (I agree many of things which referee #1 pointed out and will not repeat that here.)

1. Since authors are trying many experiment, especially for comparisons 1) between untreated and chemically abrasion and 2) among four well-known reference zircons and more, the summarized graphic including procedure and results would be effective way to make clear the output of this research.

We tried making such a graphic, but it wasn't very clear, so we didn't submit it.

2. For zircon from Mount Painter Volcanics, cathodoluminescence imaging (figure 1) and probability density diagram (figure 5) are not enough. Additional concordia diagram, table for all values, and CL imaging with a higher magnification including the beam spot after SHRIMP analysis will be necessary for the argument.

Additional concordia diagrams, with consistent format, are provided for all SHRIMP (and some TIMS) U-Pb data.

3. The slope of SHRIMP calibration in the diagram of ln(Pb/U) vs. ln(UO/U): it will be better to present all dataset on (Pb/U) vs. ln(UO/U daigram with figure 2. The argument in the line 554-555 is not valid because the slope itself cannot change the date of each spot analysis. The combination of slope AND UO/U value of each spot can be affected the calibrated date.

This part of the manuscript is a distraction, so we removed that part of the manuscript.

4. Figure 3: last diagram should be T2C not 91U. and what does green line in the middle means?

Fixed. The green line in the central value as output by Isoplot.

5. Figure 6 and figure 10: dashed line (green and blue) is not well recognized.

We are not sure how to fix this while maintaining a colourblind-friendly palette.

6. Figure 11: what is the reason to show both A and B? Moreover, the upper intercept of age of figure 11-B seems to be meanlingless.

We have replotted this data in a format consistent with the other plots.

7. Figure 4, 11, and 12: No consistent format of all concordia plots. Especially Figure 12 are too confused diagram and it is very hard to get the point.

Fixed.