Anonymous Referee #1

Received and published: 2 October 2017

Reicher et al. describe the calibration and operation of a cold-stage for measurement of ice nucleation activity. The instrument used a microfluidic device for generation of monodisperse droplets inside an oil phase. The device is operated in batch mode. Droplets are prepared in a continuous flow system. Then the flow is stopped and the device is placed on a cold-stage/microscope for nucleation experiments. Characterization experiments using pure water, eutectic solutions, and several test dusts are described. Application of the technique to ambient aerosol is demonstrated.

This is a solid paper that describes the application of a microfluidic device for ice nucleation studies. The characterization and validation experiments are of high qual-
ity. Specifically, the temperature characterization and comparison with the various test dusts at water activity of one and less than unity are thorough and convincing. The main conclusion of the paper that “WISDOM is a suitable tool for studying atmospheric ice nucleation, both in homogeneous and heterogeneous immersion freezing” is well justified.

However, the paper should be viewed in perspective of the current literature. At least 10+ similar cold-stages have been built, characterized, and described in the last few years. This is also not the first cold-stage that uses a microfluidic device. In fact the predecessor work by Stan et al. (2009) managed to operate the instrument in continuous flow mode, rather than batch mode, and thus is more technologically advanced than the work here.

In many places the manuscript tries to ‘sell’ or ‘justify’ the microfluidic technique presented here as an advance in technological capabilities. For examples, the microfluidic technique is ‘cheap’, ‘easy to prepare’, ‘easy to operate’, ‘production of droplets is fast’, ‘a range of droplet volumes can be used’, ‘cooling down to homogenous temperature range allows an extensive investigation of atmospheric particles’. A list is provided that suggests that WISDOM solves ‘some critical issues inherent in other currently used instruments’, including ‘fast production of droplets’, high statistical power’ due to ‘fast analysis of thousands of droplets’, ‘droplets are monodisperse and individually analyzed’, ‘the use of oil minimizes possible artefacts’, ‘small droplet volumes decreases freezing artefactsby impurities’, ‘static array opens the possibility to investigate several freezing cycles for the same droplets’, and ‘microfluidics method and the small droplet volumes enable working with small volumes’.

The statements above are either misleading or wrong or have not been demonstrated in this paper (see further below). It indicates that the authors have not critically reflected on the differences and similarity of WISDOM with the current technology and some of the necessary tradeoffs that are being made when designing cold stages. The WISDOM instrument does not demonstrate any aspects that has not been also addressed
in other designs. A blunt assessment of the technique is that it is on par with the current state of cold-stage designs. However, although the technique is certainly valuable (and cool!), it does not represent an advance in either science or technology relative to the published literature. A revised manuscript must better reflect the technique’s ability and limitations relative to other existing techniques.

Issues with justification. Quotes from the manuscript are in bold font.

The microfluidic chips are cheap and easy to prepare and to operate. Production of droplets is fast and a range of droplet volumes can be used within the same microfluidic device.

These statements are highly subjective. Please elaborate. What does cheap mean? The chips are custom made, there is cost for that equipment and labor involved, which is not free. On top the system requires a microscope, liquid flow control, and more complex cooling setup relative to a regular cold-stage. What does fast mean? If cheap and fast are a ‘selling’ argument for the technique, then it should be compared to other cold-stage techniques. The real metric that need to be compared relative to other techniques

(1) Total cost of equipment, total cost per ‘experiment’ in terms of equipment and consumables.

(2) Number of drops and total volume that can be studied in a reasonable ‘experiment’ (e.g. working with a dust sample for one morning). Related to this number is the lowest and highest INP concentrations that can be captured in that experiment. Factors such as channel clogging, or chip damage due to expansion of water upon freezing should be accounted for here fairly.

(3) Ease of use relative to a regular cold-stage technique.

These are the key pieces of information that are needed to weigh whether one should adopt this technique or not (other than personal preference, which may well justify the
A range of droplet volumes can be used

Other cold-stages have worked with the same range of droplet volumes and sample statistics.

Cooling down to homogenous temperature range allows an extensive investigation of atmospheric particles

Furthermore, supercooling is limited due to the presence of impurities, which increases with the volume of the droplet. Hence, to allow comprehensive studies down to the homogeneous region, low volumes are used and generation of these volumes is not trivial and may cause further complications.

The authors should elaborate on the point they are trying to convey here. Presumably, this is meant to convey that a technique like WISDOM is needed? Three points.

First, may authors have accomplished studying homogeneous freezing using small drops successfully in past and present studies. Clearly it is doable and the alluded to ‘complications’ have been solved in some way by these authors.

Second, one might be inclined to believe that going to small drop volume solves some problems. If done correctly, small volumes do allow generation of drops with no impurities and thus studying homogeneous nucleation (as demonstrated widely in the literature). However, the impurities are still present. Simply subdividing the sample into more droplets will not help raising the lower limit of detection of ice nucleation activity for that sample. Specifically, if a sample has $10^5$ nuclei per liter of liquid, studying a large number of small droplets or a smaller number of large droplets will produce the same result. This study has not demonstrated any advance in purifying water. The achievement of pure water freezing has been managed before by others when working with small volumes.

Third, the number of drops studied per experiment here is quite small ( 500). While this
means that homogeneous freezing temperature can be reached, it also means that the
total volume studied is quite small, and thus the lower limit of detection for ice nuclei is
much larger than in devices that use larger droplets.

**WISDOM solved some critical issues inherent in other currently used instruments including**

1. **fast production of droplets minimizes sample sedimentation or other aging process** that may occur in a suspension, leading to higher reliability of the measurements and to a better estimation of the surface area of the material that is exposed in the droplets.

   This has not been shown here. First, it is unclear how fast fast is. No times are given in the paper. Second, it has not been compared to how fast others can perform an experiment. Third, the paper does not demonstrate that this technique is more reliable than others. There is no metric for reliability. This statement is clearly not justified.

2. **The high statistical power that can be achieved easily by fast analysis of thousands of droplets.**

   Again, fast is subjective. The number of drops given here is that approximately 550 forty micron and 120 hundred micron droplets can be monitored per experiment. The duration of an experiment is unclear, but it includes chip production, chip loading, and post processing. So how many experiments can one person do in a week? Is the statistical power really higher than in other studies? For example Hader et al. (2014, ACP) generate and analyze 500-800 drops per experiment in the 80-100 micron size range. Another example Peckhaus et al. (2016, ACP) generate and analyze 1200-1500 drops drops per experiment om 100 micron size range. Both studies also use oil immersion. Does WISDOM have really faster analysis and higher statistical power than these studies? Perhaps so, but it must be proven.

3. **the droplets are monodisperse and individually analyzed, in contrast to other**
emulsion techniques

See above for examples that also do individual analysis.

(4) the use of oil minimizes possible artefacts from droplets’ evaporation, neighbor seeding or vapor . . .

See above cited studies and several other systems that do the same.

(5) the small droplet volumes decreases freezing artefacts by impurities, and in the absence of INPs the water freezes below the homogenous freezing threshold (-37oC), in comparison with instruments that employ droplets with larger volumes which limit the workable temperature range.

First, this point is incorrect as stated above. Working with small droplet volumes allows for study of dilute solutions and ice nuclei that have high concentration in liquid. The high concentrations are the result of the overall low liquid volume that is analyzed. Working with large droplets is a choice to increase the detection limit. Furthermore, many investigators worked with small droplets before, so this is not a distinguishing feature of WISDOM.

(6) using a static array opens the possibility to investigate several freezing cycles for the same droplets

Refreeze experiments have been performed with static droplet arrays since the pioneering work by Vali in the 1960’s. This is not a distinguishing feature of WISDOM.

(7) the microfluidics method and the small droplet volumes enables working with small volumes which is an advantage when working with atmospheric samples.

First, working with small volumes and atmospheric samples has been also done by previous investigators. Second, it is incorrect to cite the small droplet volume as an advantage. It may be an advantage under some circumstances, but not others. Specifically, the poor lower limit of detection for small droplets is a clear disadvantage.
To summarize the critique of this section. The text is prefaced with “WISDOM’s largest advantage is the use of microfluidics technology which solves some critical issues inherent in other currently used instruments”. The issues that WISDOM solves have been solved in some form or another in previous studies. WISDOM shares the advantages and disadvantages of the design choices involving the selection of oil, droplet size, batch mode, and optical detection. Claims regarding high statistical power, ease of use, and cost need to be proven. The paper should give an honest assessment of whether a new investigator should pursue the microfluidics route or the setup or something else. WISDOM and the nanoliter droplet freezing assay have similar control over drop volume, automated drop generation, drop detection, refreeze capabilities, etc. Perhaps telling is that the entire section in the draft manuscript is devoid of any references to the literature to back up their claims.

Other comments

**However, in the atmospheric heterogeneous ice nucleation field, microfluidics techniques are not widely adopted, despite many potential advantages.**

What are these potential advantages? Please be specific. Also there are clearly potential disadvantages to microfluidics. These should be discussed as well.

Figure 2 needs better explanation. The y-axis is not clearly defined. Presumably, Delta GL stands for change in grey level observed during warming or cooling? These quantities should be defined in the caption and text. Is the scale 0-255 or 0-1 or 0-100? What is std mean on the y-axis? Are the curves the population mean, or are they for a single droplet? The thermodynamic prediction for the eutectic melting point for the NaCl and pure water should be added to the graph.

Figure 3. The C and H (presumably cooling and heating) should be explained in the caption. The Delta T is a temperature and should have units of K. The freezing temperature of pure water should be given here. The text states that the delta T is evaluated against an extrapolated temperature at equilibrium conditions. Does that mean that the
equilibrium conditions $T$ for freezing and melting are not constant in the plot?

Section 3.2 provides statistics for the T50 for several devices and repeats for individual devices. However, no spectra are shown. How is the repeatability vis-a-vis early freeze events? The authors should show an overlay of the temperature spectra for all of these samples to convince the reader of the repeatability across the full range of temperatures.

Conclusions → homogenous should be homogeneous