Response to Reviewer #1

The authors are grateful to the reviewer for the important comments that helped the authors to clarify the manuscript.

In black and bold are quotations that the reviewer cites from the manuscript, the reviewer comments are in black and the authors' reply in blue. Changes made in the text in dark blue bold and italic:

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. 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 route of microfluidics).

Authors' Reply: The descriptions "cheap and easy" are referred to the microfluidic chips (or devices) and not to the entire WISDOM system. However, the authors understand the reviewer's concern and in order to avoid any subjective declarations this statement is removed from the manuscript.

In more detail:

(1) The authors agree that a statement about the microfluidic system cost is subjective and will differ from lab to lab. However, in the manuscript we only claim that the microfluidic devices themselves are cheap and do not refer to the entire system. Because cost is subjective, and we could not find this information in any other paper from the relevant fields, we could not provide a comparison of the costs with different cold stages. Hence, we removed this statement from the text, from the introduction and from the summary. We evaluate our devices at 1 euro per each device, the cost consists of the cost of a microscope glass slide and the cost of a PDMS layer on top. We do not expect this to change much from place to place.

(2) Fast: Time to produce the droplets is about 30-40 seconds and then about 10 seconds to disconnect the device from the tubes and to transfer it to the cooling stage.
We cannot provide any specific information about clogging of the devices, as we did not study this in detail. Clogging will depend on the particles' properties and concentration and maybe with the flows used. Hence if channels are clogged or damaged (normally rare and we do recycle each device for a few experiments) we do not use the device since it is unknown how this affects the concentration of the material inside the droplets or the change in the

temperature equilibration in the device if the channels are damaged. In section 3.6 we state where we discuss the disadvantages of this system.

(3) The authors agree here too, a statement about ease of use in comparison with other cold stages is subjective. In the quoted statement it was claimed that the preparation of the microfluidic device is easy and not that WISDOM is easier than the other cold stages.

The authors would like to clarify that the cooling stage is not more complex relative to the other stages as the reviewer pointed out and similar cooling stages are in use within the ice nucleation community. We did not find in the text where it was presented as complex.

A range of droplet volumes can be used Other cold-stages have worked with the same range of droplet volumes and sample statistics

The statement above introduces the versatility of the microfluidic device to study various volumes. For example, device with 100-µm diameter can be used to produce droplets with any diameter that is smaller than 100 microns.

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 105 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.

Authors' Reply: The authors use this statement to explain what were the considerations to employ low volumes of droplets.

(1) In the statement above, it was not stated that homogeneous freezing in low volumes was not studied successfully before and figure 5 and figure 6 present some of these studies which are used as benchmarks for WISDOM.

(2) With all due respect, the authors do not agree here with the referee. How the active sites are distributed as a function of temperature is critical here. It is correct that ice nucleation of materials can be studied well by using higher volumes of droplets, if the nucleation sites are active in temperatures that are warmer than the freezing of the impurities that are present in the sample (normally between ~-20 to -25C). A good example for this is snomax. In the atmosphere, supercooling is observed down to -37C and mineral dust have also active sites in colder regions than -25C. Therefore, low volumes are more sensitive to the active sites that are less rare and that are activated between -25C and down to the homogenous region. This was our consideration. In contrast, WISDOM is not sensitive to the rare active sites if they are active at warmer temperatures. In some cases, it is impossible to increase the surface area is (if there is not enough sample or if the high particle concentration disables the droplets production), then it will be possible to use a larger sample of droplets. This important point is now added to section 3.6:

"….because of the small droplets' volumes, there is less chance to characterize rare active sites."

(3) 500 droplets are normally not small for the concentrations and the materials that we use. The considerations regarding the effect of overall volume are given in (2) above.

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.

<u>Authors' Reply:</u>

(1) We agree that "fast" is subjective - removed from the text.
(2) High statistical power - removed from the text.

(3) the droplets are monodispersed and individually analyzed, in contrast to other emulsion techniques See above for examples that also use individual analysis.

Authors' Reply: In Hader et al (2014) the droplets are not monodispersed and in Peckhaus et al (2016) do not have an emulsion. The authors refer here to few other studies that used emulsion in the DSC for example and recorded the average freezing behavior. This meant to clarify that there are some more parameters as frozen fraction and freezing rate that can be obtained with individual droplets, which cannot be obtained by bulk methods. We removed the sentence "in contrast with other emulsion techniques".

(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.

Authors' Reply: As was explained earlier, these are advantages in WISDOM that the authors find important and maybe this adds to the variability between the different techniques. It is important clarify that it was not stated that this is the only cold stage that has these properties.

(5) the small droplet volumes decrease freezing artefacts by impurities, and in the absence of INPs the water freezes below the homogenous freezing threshold (-37°C), in comparison with instruments that employ droplets with larger volumes which limit the workable temperature range. This point is incorrect as stated above. Working with small droplet volumes allows studying 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 work with small droplets before, so this is not a distinguishing feature of WISDOM.

Authors' Reply: (5) Similar points are explained before.

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

Authors' Reply: (6) The authors did not claim that this is a distinguishing feature of WISDOM. We only described other possible experiments in 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.

Authors' Reply: (7) We explained the considerations for choosing the small volume in the previous replies. It is not claimed that it was not done before, and we clearly discuss the advantages and disadvantages of the volume range chosen. In any case, microfluidics allows the use of higher volumes. The authors of this manuscript made every effort to describe the system in an objective way, and to cover existing literature.

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.

Authors' Reply: These points are clearly stated in section 3.6, the authors corrected section 3.6 in order to represent better the advantages and disadvantages of WISDOM, considering the points that the reviewer raised:

"WISDOM uses of microfluidics technology which solves some critical issues inherent in other currently used instruments: (1) good control of size number of monodisperse droplets, (2) fast production of hundreds of nearly monodisperse droplets minimizes sample sedimentation or agglomeration that may occur in a suspension, leading to a good estimation of the surface area of the suspended material. Moreover, several droplet diameters can be employed in the same device without its modification, (3) good statistics achieved by individual analysis of thousands of droplets, (4) monodisperse droplets individually analyzed, in contrast to some emulsion techniques as the Differential Scanning Calorimeter (DSC) experiments, allow to obtain the frozen fraction at each temperature, and to achieve detailed information about active sites and freezing rates, (5) the use of oil minimizes possible artefacts from droplets' evaporation, neighbor seeding or vapor transfer due to the Wegener–Bergeron–Findeisen processes, (6) the small droplets' volume decreases freezing artefacts by impurities, allows to reach the homogenous freezing threshold (-37ºC), (7) possible investigation of several freezing cycles for the same droplets, (8) the microfluidics method and the small droplet volumes enable working with small sample volumes which can be an advantage when working with atmospheric samples.

WISDOM has a very accurate temperature calibration that spans a wide temperature range, using the eutectic freezing method. WISDOM most resembles the instrument used by in Edd et al. (2009). However, it seems that issues with temperature calibration in Edd et al. (2009) led to a temperature offset, and hence different freezing rates. Stan et al. (2009) achieved better temperature accuracy and high statistics. However, the freezing experiment was conducted in a flow mode, which is more complicated than in the WISDOM setup and requires complicated modeling. In addition, the cooling rates that were used were very fast, which induces additional errors. Riechers et al. (2013) had high temperature accuracy as

they also used a DSC. However, they had to collect the droplets from the device as there was no static array option and this may add further complication and contamination.

It is noted also that use of microfluidics technology has a few disadvantages. These may include: (1) oil may interact with some of the analyzed particles, possibly leading to biased data, (2) the microchannels are susceptible to clogging, (3) it is not possible to perform any post analysis to the droplets content after the experiment, (4) because of the small droplets' volumes, and there is less chance to characterize rare active sites".

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.

<u>**Authors' Reply:**</u> More detailed explanation is added to the text in section 2.3:

"The optical brightness of a droplet changes during a phase transition (freezing or melting) due to the different interaction of light with the liquid and the solids. For phase transition detection, an in-house image processing LabVIEW program monitors automatically the optical brightness change. The program detects the droplets using a spherical shape criterion and sets a square surrounding the droplet that defines an array of pixels that are attributed to that specific droplet. A change in the optical brightness is represented by the gray level value of the image's pixels, ranging from 0 to 255. Freezing is calculated per movie frame and is defined as the subtraction of the brightness mean value for each droplet in two consecutive frames (ΔGL), thus allowing derivation of freezing rates. At the beginning of the analysis, the first 15 frames are used to identify the noise level of the signal by calculating its standard deviation (std(ΔGL)). The program then searches for the maximal freezing signal that is also greater than 5 times the noise level. The temperature associated with this freezing signal is assigned as the freezing temperature for that droplet.

In this algorithm, the program can distinguish successfully between a phase transition event and noise that arises from the camera signal, droplet movement or any other interruption. Figure 2 presents a spectral analysis for different types of phase transitions observed in WISDOM. Since WISDOM operates in transmission microscopy mode, the light is scattered more efficiently by ice crystals in comparison with a liquid droplet and a freezing event involves droplet darkening and a negative signal. Example for the negative signal of a freezing event of a single droplet can be seen in Figure 2a. In comparison, during melting, the droplet becomes brighter until all the crystals melt, and the signal is positive. In Figure 2b+c the analysis of melting signal and eutectic melting signal are presented for the whole frame".

And to the figure 2 caption:

"Spectra of different phase transition events as observed in WISDOM. a) freezing, b) eutectic melting, and c) melting onset and clear point (liquefaction) are the mean of all sampled droplets in a single experiment. The phase transition is defined optically by the brightness

information obtained by the gray level of the image pixels. Std(ΔGL) describes the standard error of the difference in mean GL for two consecutive frames. At the beginning of the experiment the noise level is studied and freezing or melting is detected only if std(ΔGL) is as least 5 times greater than the noise std level. Freezing and melting examples are for pure water droplets and the eutectic melting example is for aqueous solution droplets of NaCl. Eutectic melting point of NaCl and pure water melting point are marked by the black line in b and c, correspondingly. In all cases the droplets diameter is 100 µm."

melting points were added to the graph as well.

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?

Authors' Reply: Explanation is added to figures 3. The extrapolation results in equilibration are the values that we calibrate against.

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.

Authors' Reply: Per the Reviewer's request, a graph of the full range is now added to the appendixes.

Conclusions → homogenous should be homogeneous.

Corrected in the text, thank you!