We'd like to thank the reviewer for their constructive feedback. Initially, we also shared the same concern that droplets colliding in the manner produced by our wind tunnel produced a different physical process to that of the Low and List experiments. Unfortunately, we did not have fine control over drop sizes and so it's not possible for us to directly experimentally compare with the Low and List drop-pair sizes either. However, we've re-examined the collisional kinetic energy (CKE) of our drop collisions using the equation presented in the Low and List paper and compared them to the values expected if the collision had occurred in nature at their terminal velocities. We found that the values are similar (within an order of magnitude), to the limits of the sensitivity our measurements allowed. Thus we are probably consistent with those collisions occurring in nature. Differences between our observations and those of others are likely due to larger drops being used. Significant variation in oscillation behaviour and lengths of time before breakup after a collision-coalescence event, despite some similar pre-collision conditions, leads us to conclude that for larger drops, the direct importance of collisional kinetic energy to the resulting breakup drop-size distribution is in question. Once the drop configured itself where it became clear a given breakup mode will occur, the resulting fragments for that breakup type followed a similar distribution on average (see the videos at http://youtu.be/3lxOFufnQZg). Our observations suggest that either collisional kinetic energy or earlier drop interaction history may not necessarily be directly important to the resulting post-breakup fragment distributions in our experimental setup, and that they may only be indirectly important through possibly influencing the eventual breakup type. Only the breakup type itself seems to directly determine the resulting drop-size distribution. Whether this is true for other types of experimental approach is open to debate and further study. This is now made considerably clearer in the revised manuscript.

We have also now included comparisons with the predictions of fragment distributions for 4 and 6 mm drops using the McFarquhar (2004) parameterisations.

We have updated the introduction section to recognise that 3-peak distributions are in question, and that 2 peaks are more common, as discussed in McFarquhar 2010—thanks for this.

All minor comments listed have been addressed in the manuscript. Some comments:-

>> Page 11746, line 24. What is the definition of a significant change? Is a statistical test applied?

No, this means that the 2 values on the y-axis of the figure where the lines plateau is the same in Ai and Aii.

>> Page 11747, line 2. What does slight mean?

Yes, this was not clear. The sentence has been restructured.

>>Page 11750, line 29. 25 breakup events is very few. By contrast, Low and List observed approximately 1,000 breakup events.

We agree; however, we can estimate the sample size required to estimate the results at some level of significance by using the z-statistic:

$$z = \frac{E\sqrt{n}}{s}$$

Where E is the margin of error, n the sample size and s the standard deviation. Rearranging the formula for n

$$n = \left(\frac{z \times s}{E}\right)^2$$

If we assign a significance level of 10% or z=1.28 (so we are 90% confident in our results) and a margin or error, E, of less than 5 drops, and also using our data, the standard deviation on the number of drops for the bag break ups is s=~20 drops using the formula above to estimate n gives n=~29, if we choose 25% significance we get n=~7 samples need to be taken for adequate stats. Smaller standard deviations were seen in each bin so stratifying the data by bin produced more favourable statistics.