***Big picture stuff***

**Reviewers:**

**Important correspondence:**

**1.) We don’t want to mess with published data**

**Brent:** What melt data should we use?  My main thought here is that I don't want to draw negative attention to any of our previous melt data...where people begin to question the data we and others have already published using Meltwin. I'm not sure how to handle this. Some possibilities would include (1) using unpublished data but not provide sequences (maybe just refer to them as Duplex 1, not provide raw data, etc.) so we could publish later in a thermo paper (not sure if that's legit), (2) using previously published data if the free energy values are very similar and we're just focusing on smaller error (again...not drawing too much negative attention to previously published work), or (3) collect new data that will not be part of another project (which will take more time and have some cost associated with it). I can see how a direct comparison to Meltwin could be a good thing for MeltR, but I'd hate to do that at the expense of all of the data previously published using Meltwin. This is my major concern. Is there a way to tout the data analysis of MeltR without bringing down Meltwin? Thoughts?

**Phil:** I think we can say only the nicest things about Meltwin. Nothing disparaging at all. At the same time, we can talk about reducing error and improving accuracy going forward. We would be very careful to say that prior data is still valuable.

**Jacob:** I agree with Phil. We say, “Meltwin is great. It enabled a generation of nearest neighbor parameters and all the tools that rely on them. Nobody can get Meltwin anymore and its a black box.” We can leave it at that.

On not drawing negative attention to published data, this is important. Any differences between the analysis below and the published analysis are do to idiosyncrasies. I want to proceed with the data we have. However, this is ultimately, not our data. If we are nearing a complete draft and all authors are not comfortable, I will collect a new data set to use for this paper.

**2.) How do justify baseline trimming versus fitting the whole data set?**

**Brent:** If a lot of the focus is on the baseline trimming tool, we need to be able to convince people (who think like Phil) that there are good reasons to trim.

**Phil:** I think we should allow different options for trimming but my philosophy is the baselines as long as possible but not longer option. This might do something recursive like look for the first five data points away from the transition that fit y=mx+b with random residuals and then lengthen 1 point at a time, stopping when the residual lose randomness.

**Jacob:** There are good reasons to trim. First of all, everyone does it. Second, linear baselines are an approximation and this approximation gets less valid the longer the baseline is. The question is “how to trim?” and everyone has to make their own idiosyncratic rules, which are always hard to justify because every reviewer can say “what about this?”. My feeling is that people would be happy for us to take this out of their hands if we present a sufficiently clever method.

I really liked Phil’s idea for the recursive residual analysis. Unfortunately, it does not work. (1) The baselines of ideal melting curves are never perfectly linear because it is a sigmoid. (2) On real data, this strategy just starts grabbing random parts of the melting curve.

**3.) Meltwin exists and is trusted, how do we justify MeltR**

**Brent:** I'm not sure what you had in mind for the paper. At first, I was thinking about all of the comparisons that could be made to Meltwin. But now, the more I think about it, I'm not sure that is necessary. Not many people use Meltwin, so a point-by-point comparison is probably not needed, as MeltR would be the only software available for this purpose (unless you know of others). But, I could work with Sebastian to come up with a list of limitations related to Meltwin that aren't related to the data analysis (not available for distribution, some features don't work on newer operating systems, etc.) that could be discussed in the paper.

**Phil:** This is more tricky. Again, I think we should cherish Meltwin and the many important papers published with it.

**Jacob:** I want a direct comparison because it will make Meltwin users feel safe and I want to use the word “Meltwin” a lot because it will make our paper pop up on google when somebody does a melting curve, reads a Znosko paper, and starts looking for a copy of Meltwin.

**Introduction to do list**

* Draft

**Results to do list**

* Plot the H and S instead of the H and Tm on Figure 2A-C
* Add the error quantile argument to figure 2D
* Add a plot that more succinctly summarizes percent error between methods to Figure 4
* Collect data on a non-two state folding helix for Figure 5

**Methods to do list**

* Theory section for the baseline trimmer in the manual
* Transfer over to the methods

**Discussion to do list**

* Make a list of points we want to circle back to
* outline

**SI to do list**

* Place most of the BLtrimmer stuff in the SI

**Coding to do list**

* Test easy outlier exclusion protocol to meltR.A?
* Expand description of the outputs in the BLtrimmer help file
* Generalize to any melt data structure
* Deal with memory problem at large numbers of baseline combinations.
* You tube
* Figure out why Tms don’t line up with first derivative for monomolecular melts