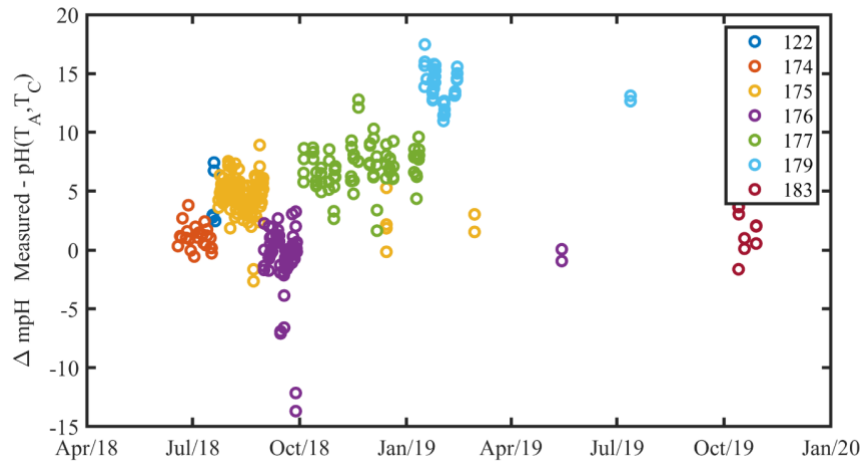**Fig. S1: A timeseries of the residual between measured and calculated CRM pH throughout the experiment. Marker color denotes CRM batch number. There is a clear variability between batches but no systematic drift in the pH system during the experiment.**
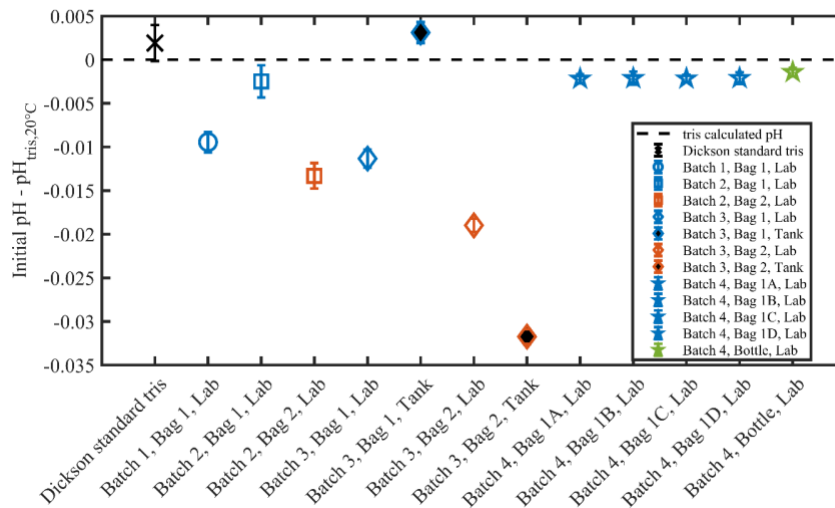

Fig. S2: The initial pH residual of each tris bag or bottle measured in this experiment. The initial pH is reported as a residual from the calculated pH at 20 °C. The initial pH was measured directly for tris batch 4 and extrapolated for tris batches 1-3. Additionally, 2 bottles of Dickson standard tris (show by the black "X") were measured on 12/10/2018. The zero black dashed line is the calculated pH of tris at 20 °C, based upon the measured reagent concentrations (DelValls and Dickson, 1998).
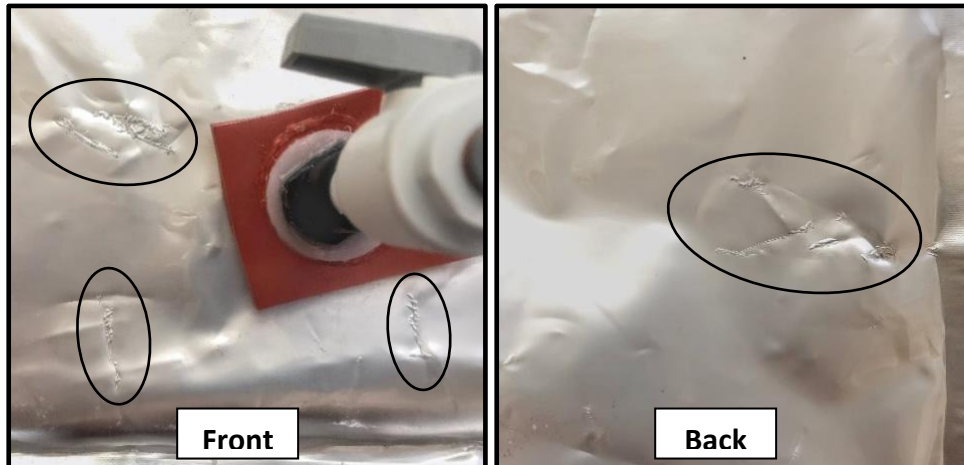
Fig. S3: The ovals indicate marks on the exterior of "Batch 2, Bag 1, Lab". These marks appear to be damage to the interior metallic layer, possibly due to creasing of the bag. These marks were not present on any other bag used in this study.
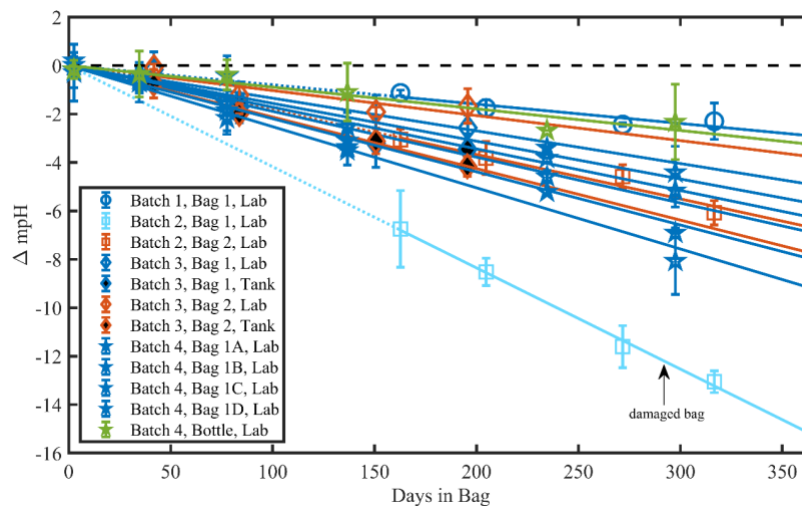

Fig. S4: A composite of all bags from all three tests as the change from the initial pH of tris vs time. Bag type 1 is shown in blue (light blue for the damaged bag of type 1), 2 in orange and bottle in green. Tris batch 1 is depicted as circles, 2 as squares, 3 as diamonds and 4 as stars. Storage location in tank has a black fill and lab symbols have no fill. The solid line is a linear regression starting at the first included pH measurement and ending 365 days after the tris was bagged. The dotted lines are extrapolations back to 0 days stored in bag. See Figure S4 for a regression and confidence interval using all measurements from undamaged bags.