Instrument inter-comparison of glyoxal, methyl glyoxal and NO_2 under simulated atmospheric conditions

R. Thalman, M. T. Baeza-Romero, S. M. Ball, E. Borrás, M. J. S. Daniels, I. C. A. Goodall, S. B. Henry, T. Karl, F. N. Keutsch, S. Kim, J. Mak, P. S. Monks, A. Muñoz, J. Orlando, S. Peppe, A. R. Rickard, M. Ródenas, P. Sánchez, R. Seco, L. Su, G. Tyndall, M. Vázquez, T. Vera, E. Waxman, and R. Volkamer

Response to Referees

Response to C. Chan Miller

Thalman et al. use results from two chamber studies (NCAR and EUPHORE) to compare the relative performance of nine instruments used to measure glyoxal and/or methylglyoxal. Glyoxal is measured by four main techniques: Visible Absorption Spectroscopy, Laser Induced Phosphoresence (LIP), Fourier Transform Infra-red spectroscopy (FTIR) and Solid Phase Microphase Extraction (SPME). Proton Transfer Reaction Mass Spectroscopy (PTR-MS) is additionally employed to measure methylglyoxal.

Thalman et al. showed good consistency between visible/IR reference cross sections and PTR-MS ion-molecule rate constants based on good agreement between the chamber measurement techniques that employed each separate calibration factor. They then compared seven instruments at EUPHORE to pure glyoxal/methylglyoxal mixtures, o-xylene and isoprene photoxidation experiments and ambient air. They found larger variability for measurement techniques that require external calibration or offline sampling. By systematically varying NO_2 in the chamber, they were able to show cavity enhanced visible instruments are moderately impacted by NO_2 interference, likely through wavelength dependent pathlength distortions. Finally they experimentally determined instrument limits of detection based on flushed chamber measurements.

Overall this paper fits well within the scope of AMT. As the authors note, the measurement techniques for detecting α-dicarbonyls have emerged very recently, providing ample need for an intercomparison. Their paper will be a useful resource for both instrument developers and users of their data. I recommend publication after the following comments are addressed.

We thank C. Chan Miller for his very detailed review of this manuscript. We feel that the comments have helped improve the use of the manuscript as a resource for the atmospheric chemistry community interested in a-dicarbonyls in polluted and pristine environments.

Major Comments

• **Section 2.1.6:** How frequently was the CRDS calibration step employed for the Mad-LIP instrument? Should we expect drift in the calibration response factor from mirror alignment variations over the course of a single chamber experiment?

For the first part of the campaign the instrument was calibrated about every two days. In the latter half of the campaign, the instrument was calibrated less often. Thus it is hard to judge how much the calibration factor/alignment may have drifted over the course of a single experiment.

• **Section 3.2.1:** Do you know the cause of the second cluster of BBCEAS points offset from the 1:1 line Figure 3b (x-axis > 1.5 ppbv)?

These points are errant points left in the plot from experiment 8b (long-term dilution) and have been removed from the plot. These points are not included in the correlation fits and do not alter the values in the table.

• **Section 3.2.1:** How well do the Mad-LIP/CE-DOAS slopes agree close to the start of the experiment? It would be interesting to know how well the Mad-LIP response is close to CRDS calibration to separate this from drift associated with mirror alignment changes.

See comment above (frequency of calibrations), which explains why this is difficult to address. The new single-pass design, and more frequent calibrations of Mad-LIP should eliminate this issue.

• **Section 3.2.3:** The poor agreement between BBCEAS and CE-DOAS for methylglyoxal given the excellent agreement for stronger absorbers is interesting, given the similarity of the techniques. Past studies have shown that weak absorbers are highly sensitive to the spectral window chosen for analysis. The reported windows analysed by CE-DOAS and BBCEAS are 435-465 nm and 430-486 nm respectively, which could account for this discrepancy. Have the authors investigated sensitivity of fit window position to methylglyoxal?

We would not categorize the agreement between BBCEAS and CE-DOAS for methyl glyoxal as poor. For E2 both instruments agree within 1% with R^2 = 0.9987. This level of agreement suggests that the difference in the fit window does not present a fundamental limitation. The agreement is less good for mixed compound experiments (E3, E6 and E7), but still better compared to the other techniques. Three of the four experiments show agreement within better 10%.

We note that both instruments appear 'similar' only at first sight. In practice, the number of light path foldings inside the optical cavities (number of reflections on mirrors) is very different (about 4 higher for CE-DOAS). As a result, different criteria apply to optimize the fit window for each instrument. For example, the wider BBCEAS fit window requires a higher polynomial order (typically 6, see Section 2.1.4), and leverages the differential and absolute cross-section to quantify methyl glyoxal. The CE-DOAS fit window is optimized primarily to measure the differential (structured) absorption, and dependence on broad-band absorption is minimized by fitting a lower polynomial order in a smaller fit window. Sensitivity studies can hence not isolate 'fit window' without also varying the polynomial order, and reproducing joint settings would result in sub-optimal settings for at least one instrument. That despite very different, independently optimized fit settings (two research teams have coordinated only regarding the use of cross-sections) the BBCEAS and CE-DOAS show the best agreement among all techniques for methyl glyoxal we interpret to indicate robustness of retrievals. A further optimization of fit-window/polynomial order would benefit from a technique that measures methyl glyoxal reliably without interferences, and at a level of accuracy that is much better than 10%. Such a technique is currently missing, as is also stated in the text.

• **Section 4.1:** I could not quite follow how the authors arrived at the recommended line strength values at the end of the section from the previous discussion. Could this be clarified?

The recommend value is based on the comparison of six independent sources of calibration: (1-4) IR spectra from NCAR, PNNL (Profeta et al., 2011), (Raber, 1992), and (Talukdar et al., 2011), (5) the UV-vis

cross-section (Meller et al. 1991), and (6) PTR calibrations from ion-molecule rate constants as cited in the paper. The following has been added to Section 4.1: "Based on a careful comparison of the available spectra, we recommend the following integrated infrared cross section values for use in future studies: $(2.86 \pm 0.14) \times 10^{-17}$ cm molecule^{-1} between 1600 and 1800 cm^{-1} (average of NCAR, PNNL (Profeta et al. 2011) and NOAA (Talukdar et al. 2011) spectra); $(9.9 \pm 0.5) \times 10^{-18}$ cm molecule^{-1} between 2780 and 2880 cm^{-1} (average of PNNL and NOAA spectra). The recommended values include an estimated overall uncertainty of 5%."

• **Section 4.2:** Supplement, Line 81 - I was not quite sure how the 15% fit error for SCDgly was derived for the high NO_2 case. Was it estimated from the fit residuals? How was the 10 ppbv threshold determined for when the NO_2 pathlength distortion becomes significant? In principle it seems that the path length variations caused by the stronger absorber can be modelled. Is there also a possibility of other spectral effects at high concentrations (e.g. Raman "filling in" of NO_2 absorption structures)?

In the high NO_2 case the fit error for the SCD_{gly} is determined from the fitting algorithm error, which increases due to higher RMS due to reduced photon count, and systematic residual structures. The 10 ppbv limit for NO_2 was found (for CE-DOAS) based on the growth of the residual from a single NO_2 cross-section fit, as well as the extinction due to NO_2, which becomes limiting with respect to the mirror loss and Rayleigh scattering at about this concentration (Eq. 1).

While Raman 'filling-in' of NO_2 absorption structures can in principle add spectral effects, it can be ruled out here because the cavity acts as an effective filter. Raman scattering (as Rayleigh scattering) is in all directions, but only the fraction of Raman scattered photons that exactly scatter along the optical axis of the cavity have a chance of reaching the detector. Since Raman scattering is a fraction of the Rayleigh scattered light, the Raman intensity further decreases as NO_2 absorption becomes limiting to overall extinction (at high NO_2). Spectral effects at high NO_2 concentration are difficult to model, because they are primarily related to imperfections in the knowledge about instrument line shape asymmetry. The NO_2 residual fitted in experiments N3, E9 and E10 (see Section 2.1.3) represents an 'effective sum' of imperfections that is empirically determined from observing pure NO_2 gas at high concentrations. High NO_2 does not pose a fundamental limitation to retrievals, and retrievals benefit if pure NO_2 spectra can be recorded.

• **Section 4.4:** A potential application of the α-dicarbonyl measurements is constraining high and low NOx VOC oxidation pathways, making the NO_2 interference very relevant. It seems that the dominant effect tested was pathlength distortion at very high NO_x concentrations. Are you able to bound the associated interference at lower NO_x levels (0-10ppbv) where reference cross section uncertainties are likely the dominant driver of interference?

We agree, and have added a new Figure S5 and S6 to assess the effect of lower NO_2 at ambient glyoxal concentrations. The literature cross-section does not pose a measurable limitation in CE-DOAS, and results are within error also for BBCEAS. Section 4.4 has been expanded with a discussion of the new Supplementary Figures.

• **Section 4.5:** Many previous modeling studies have compared simulation results with glyoxal/methylglyoxal measurements derived from chemical derivitisation techniques such as DNPH-HPLC. The results of comparisons with SPME seemed highly variable relative to the spectroscopic techniques. In this paper, this was attributed to manual manipulation of samples. Does this problem apply more generally to the other chemical techniques?

The physical manipulation referred to here is the timing of the SPME fiber exposure in the chamber (as well as insertion) and the extraction. Previous derivatization techniques would also have similar challenges, but they are not unsurmountable. Just as further standardization of the procedure following this campaign improved the reliability of the measurement as noted with the accompanying reference, this should not necessarily call into question previous studies using derivitization.

• **Section 4.5** Instrument performance was generally assessed at relatively high glyoxal concentrations (up to 15 ppbv). Ambient levels (outside of heavily polluted regions) tend to fall in the range of 0-200 pptv. Do you have a sense of how much confidence we should have for measurements in this range based on the results of this study?

Yes, we do. We have added a new Section 4.6 'Comparison of atmospheric glyoxal concentrations' and a new Table 5 to compare ambient glyoxal concentrations (generally below 300-500 pptv). For this concentration range CE-DOAS, BBCEAS and Mad-LIP provide meaningful error bars; any lower threshold concentration would have limited the number of instruments further. Generally we find a lower overall variability in the slopes of correlations between low concentration data as in the high concentration experiments, with R^2 between 0.8-0.98, and intercepts generally below 25 pptv (lower than LOD, compare Table 4).

• **Section 5:** Could the authors clarify the basis of the conclusion for a lack of water vapour interference in point 5. It was not mentioned in the discussion section. Also it is not immediately obvious if it could be deduced based on the ambient air measurements in experiment E6, since the glyoxal concentrations could also be driven by many other confounding factors due to the uncontrolled nature of the experimental conditions.

We have added a new Section 4.5 'Interferences from H_2O' to the discussion Section to make our reasoning transparent. An extract reads: "The effect of water on glyoxal retrievals depends on the absolute amount of water present in the gas-phase. This was investigated systematically during experiments E6 and E7. The humidity during E6 (50% RH at 303 K) corresponds to 2.1 $\%_{v/v}$ H_2O, or 15.2 g m^{-3} absolute humidity. This is 5.5 and 1.6 times higher than the absolute humidity of 2.76 and 9.65 g m^{-3} at 50% RH at 275K and 295K that is characteristic of the arctic-, and mid-latitude troposphere; and somewhat lower than humidity in tropical air (26.9 g m^{-3}; 80% RH at 305 K). E6 and E7 provide a meaningful assessment of the effect of overlapping H_2O spectral features at visible wavelengths. In principle also other unknown factors in the ambient air studied during E6 could affect the retrievals; however, for the absorption techniques H_2O is the primary and the only known factor. In particular, NO_2, O_3, and aerosols can be ruled out to influence the glyoxal retrievals during E6 significantly (see Table 3, Sects. 3.2.3, 3.2.4, 4.3, and 4.4). Tables 3 and 5 do not provide evidence that would suggest a significant specific influence of H_2O absorption on glyoxal retrievals."

Editorial Comments

• **Figure 4:** The right hand side y-axis is missing a label

This correction has been made.

• **Figure 4 and 6:** I think a definition of what the error bars are is needed in the captions.

Captions have been added to the figure describing the meaning of the error bars.

• **Figure S2:** The CE-DOAS points are a mixture of squares and circles. Is there a reason for this?

This was a formatting error in the figure and has been corrected.