Research Article

Clumped isotope measurements of small carbonate samples using a high-efficiency dual-reservoir technique

Sierra V. Petersen*,† and Daniel P. Schrag

Harvard University, Department of Earth and Planetary Sciences, 20 Oxford Street, Cambridge, MA 02138, USA

RATIONALE: The measurement of multiply substituted isotopologues of CO_2 derived from carbonate has allowed the reconstruction of paleotemperatures from a single phase ($CaCO_3$), circumventing uncertainty inherent in other isotopic paleothermometers. Traditional analytical techniques require relatively large amounts of carbonate (3–8 mg per replicate), which limits the applicability of the clumped isotope proxy to certain geological materials such as marine microfossils, commonly used for paleoclimate reconstructions.

METHODS: Clumped isotope ratio measurements of small samples were made on a new, high-efficiency, dual-reservoir sample-preparation inlet system attached to a Thermo-Finnigan MAT 253 mass spectrometer. Sample gas produced on the inlet is introduced from a 10 mL reservoir directly into the source via a capillary. Reference gas fills an identical 10 mL reservoir installed between the reference bellows and the capillary. The gas pressures in the two reservoirs are initially balanced, and are allowed to decrease together over the run.

RESULTS: Carbonate samples from 1 mg to 2.6 mg produced Δ_{47} values equivalent to those from the traditional two-bellows method with identical single-sample precision (1 SE = 0.005–0.015‰) and external standard error (SE = 0.006–0.015‰, n = 4–6). The size of sample needed to achieve good precision is controlled by the sensitivity of the mass spectrometer and the size of the fixed reservoirs and adjacent U-trap installed on our inlet.

CONCLUSIONS: The high-precision clumped isotope measurements of small aliquots of carbonate obtained in this method allows for the application of this proxy to a wider range of geological sample materials, such as marine microfossils, that until now have been nearly impossible to analyze given sample size limitation. Copyright © 2014 John Wiley & Sons, Ltd.

The measurement of multiply substituted isotopologues of CO_2 derived from carbonate materials has allowed for the reconstruction of paleotemperatures in a variety of geologic settings. By deriving a temperature from a single phase ($CaCO_3$), the clumped isotope paleothermometer circumvents the uncertainty inherent to traditional paleothermometers that require information about the isotopic composition of additional phases (H_2O). The carbonate clumped isotope paleothermometer is based on the temperature-dependent ordering of the heavy isotopes ^{13}C and ^{18}O within the carbonate lattice.[1] At colder temperatures, these heavy isotopes 'clump' to produce mass-63 $CaCO_3$ ($Ca^{13}C^{18}O^{16}O_2$) at a level above that expected by a random (stochastic) distribution of these isotopes. When the carbonate is converted into CO_2 via acid digestion, this mass-63 anomaly manifests itself as excess mass-47 CO_2, denoted by the quantity Δ_{47} (see Eqn. (1)).[1] A Δ_{47} value of zero indicates a fully random distribution of isotopes. Over normal Earth

surface temperatures, Δ_{47} varies between about 0.55 and 0.8 in the absolute reference frame.[2] Equation (1) defines Δ_{47} where the $R^x = {}^xCO_2/{}^{44}CO_2$ and R^{x*} is the corresponding stochastic distribution ratio.

$$\Delta_{47} = \left[\left(\frac{R^{47}}{R^{47*}}-1\right)-\left(\frac{R^{46}}{R^{46*}}-1\right)-\left(\frac{R^{45}}{R^{45*}}-1\right)\right]\times 1000 \qquad (1)$$

The temperature dependence of Δ_{47} has been demonstrated for a variety of carbonate materials including synthetic, biogenic, and inorganic carbonates,[1,3,4] as well as by theoretical calculations.[5,6] Although this is still a new proxy, researchers have already used it to determine the paleo-altitude of growing mountains,[7] reconstruct hydrological and ecological conditions in Africa during the time of early humans,[8] and measure the body temperature of dinosaurs,[9] along with many other applications.

Traditional mass spectrometry practices used to measure stable isotope ratios ($\delta^{18}O$ and $\delta^{13}C$ values) require micrograms of carbonate material for a single measurement. This facilitates the creation of high-resolution records made up of many measurements of small (<1 mg) carbonate samples such as foraminifera. In comparison, the clumped isotope technique requires 5–15 mg of carbonate per replicate,[3] limiting the possible applications of this proxy. In a few cases, this sample size limit has been reduced. Zaarur et al.[10] adjusted the capillary cross section and flow rate and

---

* *Correspondence to:* S. V. Petersen, University of Michigan, Department of Earth and Environmental Sciences, 1100 N. University Ave, Ann Arbor, MI 48109, USA.
E-mail: sierravp@umich.edu

† Present address: University of Michigan, Department of Earth and Environmental Sciences, 1100 N. University Ave, Ann Arbor, MI 48109, USA.

installed a new bellows potentiometer, and were able to measure samples of 3–4 mg per replicate. Using a Kiel device to measure tiny aliquots of carbonate (0.2 mg) for a few minutes each and averaging the data from 6–10 aliquots (equivalent to 1.2–2 mg), Schmid and Bernasconi[11] were able to calculate a Δ_{47} value with a precision of 0.015–0.040‰ (1 SE), while at the same time producing a high-resolution stable isotope record. By averaging together 5–13 of these runs (equivalent to 6–26 mg), they were able to achieve an external precision of 0.005–0.010‰ (1 SE). The total mass of carbonate required for this is similar to ~4 replicates at 3–4 mg (equivalent to 12–16 mg), and both methods are an improvement over traditional sample requirements (>24 mg for 3 replicates).[3] Meckler et al.[12] showed that with additional corrections an external error of 0.007–0.009‰ (1 SE) could be achieved with 4.5–6 mg of total sample.

The clumped isotope proxy has the potential to be very useful in the field of paleoceanography due to its ability to separate the influences of temperature and the isotopic composition of seawater on the $\delta^{18}O$ values of marine carbonates.[1] Foraminifera, a commonly used sample material for paleoceanographic studies, have been shown to follow the same temperature-Δ_{47} relationship as other biogenic carbonates.[13,14] However, current analytical techniques make it difficult to acquire enough sample material for replicate analysis of foraminifera due to their small size. A few studies have successfully measured the clumped isotope composition of foraminifera using the traditional large-sample methods[13,15] and the Kiel device method described above.[14] New methods to reduce sample size requirements will make this proxy more widely accessible as a tool for paleoceanography.

Here we present a new method of measuring individual aliquots as small as 1 mg. A high-efficiency dual-reservoir inlet system allows analysis of smaller samples by reducing the 'wasted' gas remaining in the bellows and the sample vial. Gas pressures decrease slowly from a fixed-volume sample reservoir and from an identical reference reservoir installed between the reference bellows and the change-over block. In this study we test the sample size limit of this inlet configuration by measuring carbonate standards from 1.0 mg to 2.6 mg and demonstrate internal precision of 0.005–0.015‰ (1 SE) and external standard error of 0.006–0.015‰ (1 SE) for 4–6 replicates, in line with the traditional dual-bellows configuration (1 SE = 0.002–0.018‰).[4,10,16,17]

EXPERIMENTAL

Samples and standards

To calibrate the newly constructed sample-preparation inlet, carbonate standards of known composition were measured repeatedly at a range of sizes. Two high-temperature carbonates (CM2, NBS19) with similar Δ_{47} values, but different $\delta^{13}C$ and $\delta^{18}O$ values, were measured. One low-temperature coral (RTG) with a higher Δ_{47} value was measured for comparison. All isotopic compositions are reported relative to V-PDB. All errors on Δ_{47} reported in this section are external standard errors (1 SE) calculated on many measurements, with the number of measurements in parentheses.

- CM2 – An in-house Carrara Marble standard with isotopic composition $\delta^{13}C = 2.29$‰ and $\delta^{18}O = -1.77$‰.[18] Previous analysis of this standard at Harvard using the traditional

two-bellows method and large (8 mg) sample sizes yielded a Δ_{47} value of 0.385 ± 0.005‰ (n = 40)[2] in the absolute reference frame.[2]

- RTG – A coral specimen from Raratonga, used as a cooler-temperature in-house standard, with isotopic composition $\delta^{13}C = -2.20$‰ and $\delta^{18}O = -4.11$‰.[18] Limited previous analyses of this standard in the two-bellows configuration produced a Δ_{47} value of 0.720 ± 0.007‰ (n = 11).[18]

- NBS-19 – An IAEA Carrara Marble standard with isotopic composition $\delta^{13}C = 1.95$‰ and $\delta^{18}O = -2.20$‰. Previous analyses at Harvard produced a Δ_{47} value of 0.373 ± 0.007‰ (n = 7)[2] in the absolute reference frame.[2] Other labs found similar values [$\Delta_{47} = 0.399 \pm 0.005$‰ (n = 12) at Johns Hopkins (Baltimore, MD, USA),[2] 0.404 ± 0.006‰ (n = 40) at Yale (New Haven, CT, USA),[2] and 0.373 ± 0.004‰ (n = 20) and 0.359 ± 0.004‰ (n = 19) at 25°C and 90°C at Goethe University (Frankfurt, Germany)[16]].

To correct the raw Δ_{47} data to the absolute reference frame,[2] heated and equilibrated gases were measured through the same sample-preparation inlet between carbonate sample runs. A large number of heated and equilibrated gases (10–15) were run at the beginning of each measurement period to establish the calibration lines. During the measurement period, a gas standard was run every 1–2 days (every 3–8 samples). To prepare heated gases, aliquots of four gases (2 tank gases and reacted CM2 and RTG) of distinct composition were transferred into quartz tubes, which were heated to 1000°C for 2 h. This procedure randomizes the isotope distribution to produce a near-stochastic arrangement, which we measure to correct for a number of mass spectrometer source effects.[19] To prepare equilibrated gases, aliquots of the same four gases were transferred into Pyrex tubes containing ~1 mL of deionized water. The tubes were placed in water baths held at 10°C and 35°C and allowed to equilibrate over a minimum of 2 days. Each tube was removed immediately before analysis and the gas was extracted within minutes, before the CO_2 could equilibrate to room temperature (see description below).

Preparation of samples

Samples were prepared through a newly constructed high-efficiency dual-reservoir sample-preparation inlet (Fig. 1). This inlet follows the procedure outlined by Dennis and Schrag[4] for creating and cleaning CO_2. Gas was introduced into the inlet in two different ways. For carbonate samples and standards, CO_2 was created by reaction with anhydrous phosphoric acid held at 90°C in a common acid bath and continuously frozen into a large U-trap submerged in liquid nitrogen (LN_2). The reaction time was 6 min, and it extended well beyond the point when visible bubble formation ceased. On its way to the U-trap, the gas passed through a trap held at −80°C to remove any trace amounts of water. For gas standards (heated gases and equilibrated gases), CO_2 was introduced into the inlet via a cracker. As with the reacted carbonate samples, the CO_2 passed through a −80°C trap to remove water and was frozen into the large U-trap on the far side. For equilibrated gases, within 2–3 min of being removed from the water bath (10 or 35°C), the base of the Pyrex tube was submerged in LN_2, freezing both the water and the CO_2 and inhibiting the equilibration of the CO_2 with
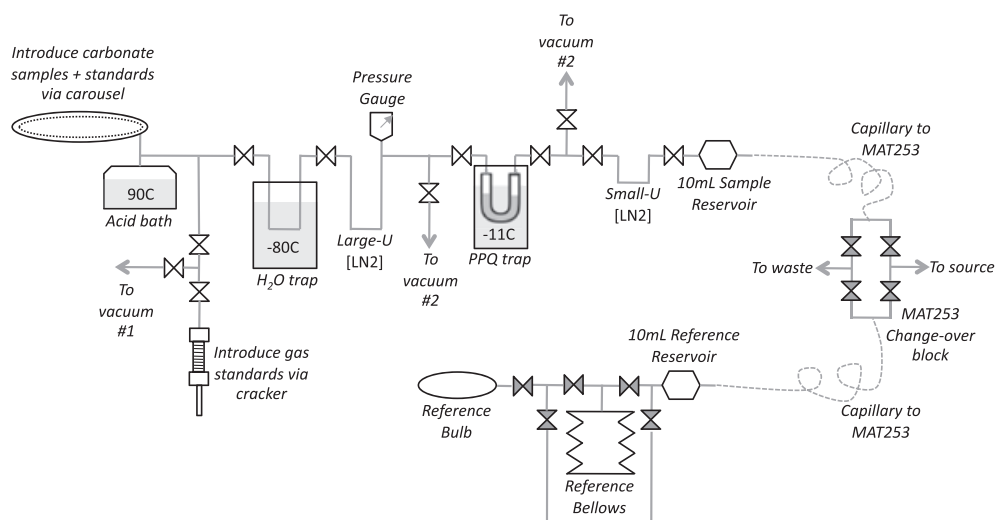
Figure 1. Diagram of the new, high-efficiency, dual-reservoir sample-preparation inlet. Double triangles represent valves. Shaded valves are controlled by the MAT 253, white valves are part of the inlet. Further details on where the 10 mL reservoir attaches to the mass spectrometer can be found in Supplementary Fig. S1 (Supporting Information).

water at room temperature. The LN_2 was then replaced by a $-80°C$ trap before the cracking step to release the CO_2 while the water stayed frozen, and was kept at that temperature while the CO_2 was transferred to the U-trap. This whole process took less than 10–15 min and the exposure of CO_2 to water at room temperature was short enough to avoid the significant re-equilibration observed in other experiments.[20] For all sample types, once the freezing step was complete, the CO_2 was allowed to warm up to room temperature within the large U-trap and the volume of gas created or transferred was roughly determined by an analog pressure gauge.

To remove trace contaminants, the gas was then passed through a Pyrex U-trap (outer diameter ½") packed with Porapak Q (PPQ) material held at -10 to $-12°C$ by immersion in cooled ethanol. Gas was frozen on the far side into a small U-trap (o.d. = ¼", internal volume ~6 mL) immersed in LN_2. During this step, the pressure measured on the gauge on the large U-trap gradually decreased, demonstrating that the gas was leaving the large U-trap. When the pressure on the gauge neared baseline and stopped decreasing (~4–7 min depending on sample size), this step was deemed to be complete. The small U-trap was closed off and the clean CO_2 allowed to warm up to room temperature. Finally, the gas was expanded from the small U-trap into the 10 mL sample reservoir and allowed to equilibrate for 3 min. This completed the sample preparation procedure, which in total took ~30–40 min per sample. While one sample was being analyzed on the mass spectrometer, and before the next sample was processed, the PPQ trap was baked for 20–35 min at ~150°C to remove any collected contaminants.

Mass spectrometry

While the CO_2 was equilibrating between the small U-trap and the 10 mL sample reservoir, the gas was introduced directly from the reservoir into the source. A more precise determination of yield was estimated at this point (compared with the rough estimation from the inlet pressure gauge) using the initial beam intensity and the pressure reading from the vacuum gauge of the Finnigan MAT 253 isotope ratio mass spectrometer (Thermo Scientific, Bremen, Germany) (Fig. 1). Reference gas from the bellows filled an identical 10 mL reference reservoir installed between the bellows and the change-over block (Fig. 1, and Supplementary Fig. S1, Supporting Information). The reference bellows were manually adjusted until the intensities of the two m/z 47 beams were balanced (on average to within 55 mV, and within 12 mV for samples with initial m/z $47 < 1000$ mV). The reference and sample reservoirs were then closed off from the bellows and the small U-trap, respectively, so that the volume from which gas entered the source was identical on the sample and the reference sides. These reservoirs remained closed for the entire run, unlike in the dual-bellows method, where the reservoirs are replenished at the beginning of each acquisition during the pressure adjustment phase.
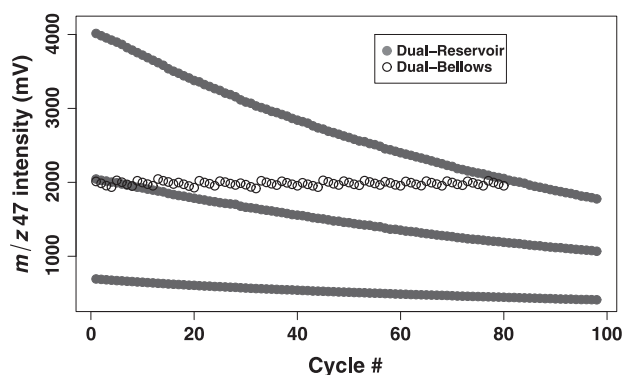
In order to perfectly balance these volumes, 87 clean glass beads (3 mm diameter, ~1.2 mL total volume, Supplementary Fig. S1, Supporting Information) were placed permanently in the sample-side reservoir. These beads are necessary to balance the difference between the internal volumes of the adjacent MAT 253 and inlet valves. The inlet valve has a larger internal volume, so the volume of the sample reservoir needs to be decreased accordingly. The specific number of beads was determined by gradually adding beads to the sample side until the beam intensities decreased at the same rate. Any minor offsets between the initial gas pressures set by manual bellows adjustment are eliminated during the course of the run as the reservoir with the higher gas pressure decreases more quickly and eventually matches the other reservoir closely.

The MAT 253 at Harvard is equipped with five Faraday cups with resistors of 3×10^7 Ω, 3×10^9 Ω, and 1×10^{10} Ω for masses 44 through 46 and 1×10^{12} Ω for masses 47 and 48. The capillaries on the mass spectrometer have been changed from the factory-fitted stainless-steel variety to a deactivated fused-silica capillary (~1 m in length, 110 μm i.d.) to prevent

S. V. Petersen and D. P. Schrag

Rapid
Communications in
Mass Spectrometry

the exchange of CO_2 and H_2O within the capillaries (Supplementary Fig. S1, Supporting Information). One sample run lasts about 2 h and 20 min and comprises seven acquisitions of 14 cycles each, with an integration time of 26 s/cycle and an idle time of 12 s, equivalent to 2548 s of total integration time on each sample. The raw voltage data is processed as outlined by Huntington et al.[19] to obtain raw Δ_{47} values. The carbonate unknowns are then corrected to the absolute reference frame using the heated and equilibrated gas data,[2] followed by an additional Δ_{48} correction described below.

In the traditional two-bellows measurement configuration, the beam intensities on both the sample and the reference side are set to a target value at the beginning of each acquisition (e.g. m/z 47 = 2 V or 8 V for the Harvard instrument[4,18]). Over the course of one acquisition, the sample and reference beam intensities decrease somewhat (the amount depends on how many cycles per acquisition and the starting beam intensity), but are returned to the target value at the start of the next acquisition by compressing both the bellows (Fig. 2). In this setup, all the cycles are performed at, or closely below, the target voltage (Fig. 2). By staying near the target voltage, issues of nonlinearities in the source are avoided, and there is no risk of gas fractionating as it decreases to a very low pressure. A 10 mL reservoir was installed between the bellows and the capillary on both the reference and the sample side of the mass spectrometer to increase the volume from which the gas enters the source (Supplementary Fig. S1, Supporting Information), therefore reducing the rate at which the gas pressures (and beam intensities) decrease.[4] At m/z 47 = 2 V, $\Delta_{47\text{-raw}}$ can be measured to a precision of 0.005–0.010‰ (1 SE) for a single sample run,[4] reflecting beam stability and shot noise limits.[21]

In our dual-reservoir configuration, each cycle is measured at progressively lower beam intensity (Fig. 2). Over the course of a 2-h 20-min run, the m/z 47 beam intensity decreases by 40–60%, at a rate proportional to the gas pressure in the reservoirs (Fig. 2). While this gas pressure decreases, the beam intensities on the sample and reference sides remain balanced. The ratios between the beams also

remain constant, so the calculated isotope ratios do not show a trend with beam intensity. Despite the beam intensity changing significantly over the run, this configuration can achieve similar precision to the dual-bellows configuration for a single measurement, demonstrating isotope-ratio stability over a large voltage range. For a single sample run in this measurement configuration, $\Delta_{47\text{-raw}}$ can be measured to a precision of 0.005‰ to 0.016‰ (1 SE), depending on the sample size, consistent with the shot noise limit[21] (see Supporting Information for further discussion).
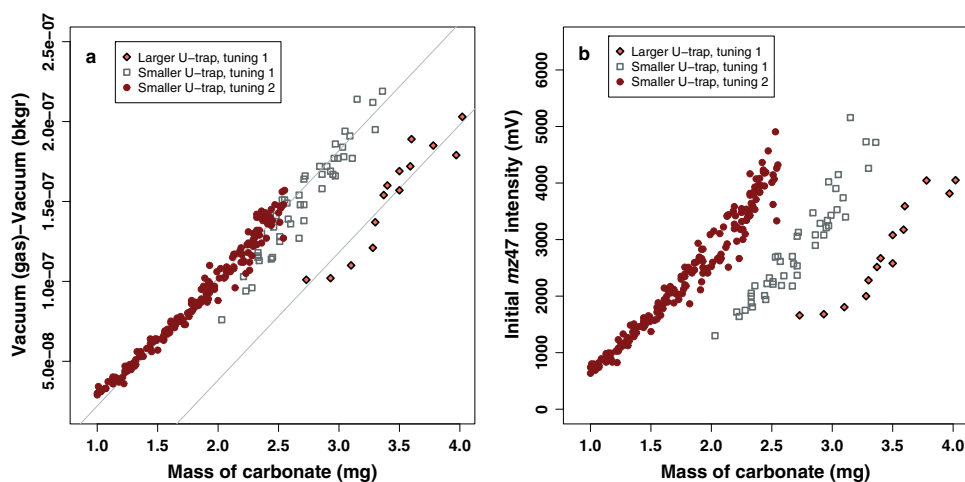
RESULTS

Demonstration of yield at small sample sizes

The sample yield was measured by calculating the increase in pressure recorded by the source vacuum gauge after introducing the sample gas into the source. This yield estimate represents the volume of gas reaching the source per milligram of carbonate reacted, and includes both yield from the acid digestion step and any loss of gas that may have occurred through the cleaning process. The data shows a clear linear relationship between this pressure increase and the mass of carbonate reacted (Fig. 3(a)), as is expected from the ideal gas law. In June 2013, installation of a smaller U-trap in front of the sample reservoir decreased the volume of 'wasted' gas left in the U-trap and increased the pressure in the source for the same mass of carbonate reacted (Fig. 3(a)). Even at the smallest sample sizes, the linear relationship was maintained, indicating that a good yield was being achieved. This corroborates the information from the inlet pressure gauge, which indicates that nearly all the gas is transferred away from the large U-trap during the PPQ cleaning step.

The sample size can also be compared with the initial m/z 47 beam intensity (Fig. 3(b)). For comparison with other instruments, the m/z 44 beam intensity is about 5/8 of the m/z 47 intensity. A source tuning after a power outage in early October 2013 strongly increased the sensitivity (mV beam intensity/mol of gas) of the mass spectrometer, which allowed us to decrease our sample size further, while maintaining the same level of precision (Fig. 3(b), smaller U-trap, tuning 2). This tuning was performed using the Isodat autofocus routine (Thermo Scientific). The increase in sensitivity came mainly from a decrease in the Extraction parameter, and corresponded to a significant decrease in the slope of the gas calibration lines (0.009 to 0.006) but minimal changes in the intercepts. After this shift in sensitivity, sample sizes ranging from 1.0 to 2.5 mg produced initial m/z 47 beam intensities between 650 and 5000 mV.

Diagnosing an unknown fractionation

We observed in all our data a strong correlation between Δ_{48} and Δ_{47}, both raw ($\Delta_{47\text{-raw}}$) and fully corrected to the absolute reference frame[2] ($\Delta_{47\text{-RFAC}}$), with Δ_{47} increasing by ~0.05‰ for every 1‰ increase in Δ_{48} (Fig. 4). This relationship was present in carbonate standards and heated gases, and the slope of this relationship was similar among all sample types (Supplementary Table S1, Supporting Information). The slope was also nearly unchanged before and after the correction to

Figure 2. m/z 47 beam intensity (mV) over the course of a sample run, shown for one run in the dual-bellows configuration (20 acquisitions × 4 cycles, target voltage 2 V on m/z 47) and three runs in the dual-reservoir configuration with high, medium, and low starting voltages for m/z 47 (7 acquisitions × 14 cycles, changing m/z 47). m/z 47 beam intensity decreases by 55%, 48%, and 41% for the three cases, respectively.

2374

wileyonlinelibrary.com/journal/rcm   Copyright © 2014 John Wiley & Sons, Ltd.   Rapid Commun. Mass Spectrom. 2014, 28, 2371–2381

Figure 3. Indicators of sample yield. (a) Mass vs source vacuum pressure above background. Lines are fitted through all three data sets and show a consistent yield relationship. (b) Mass vs initial m/z 47 beam intensity. Three time periods are plotted. 05/13–06/13 - Larger U-trap near the sample reservoir, 1st source tuning configuration; 06/13–09/13 - Smaller U-trap near the sample reservoir, 1st source tuning configuration; 10/13–03/14 - Smaller U-trap near the sample reservoir, 2nd source tuning configuration.
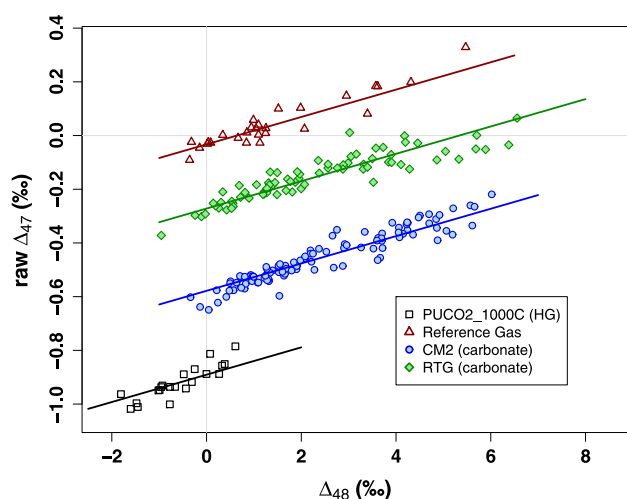
Figure 4. Correlation between Δ_{48} and $\Delta_{47\text{-raw}}$ shown for two carbonate standards, one heated gas, and the reference gas run as a sample, all run from September 2013 to March 2014. Lines are the result of a group regression to find a common slope (0.0509) and individual intercepts. Error bars are smaller than the symbols. Separate regressions of each sample type are found in Supplementary Table S1 (Supporting Information).

the absolute reference frame (Supplementary Fig. S2 and Table S2, Supporting Information), and across significant changes in source tuning (Supplementary Fig. S3 and Table S3, Supporting Information). We observed no significant correlation between Δ_{48} and $\delta^{18}O$ or $\delta^{13}C$ (Supplementary Fig. S4, Supporting Information), or between sample size and $\delta^{18}O$ or $\delta^{13}C$ (Supplementary Fig. S5, Supporting Information). Sample size seemed to have an influence on the magnitude of the fractionation, with the smallest carbonate samples often (but not always) having higher Δ_{48} and Δ_{47} values (Supplementary Fig. S6, Supporting

Information). The slope of the fractionation was mildly dependent on the average Δ_{47} value of the sample type, with the slope getting slightly steeper (0.04 to 0.08) as the degree of clumping decreased (Supplementary Table S1, Supporting Information).

Previously, high Δ_{48} values were interpreted as being caused by contamination by hydrocarbons, chlorocarbons, or sulfur compounds, which produce a mass interference on mass-47 and mass-48.[22] Conventionally, samples should plot within the envelope of calibration gas points in δ^{48} vs Δ_{48} space. Points outside this range are thought to be contaminated and would be thrown out in a typical study.[19] In our data, we observed Δ_{48} values both within and outside the envelope of the calibration gas data (see Fig. 5), for samples that we consider to be clean, such as reference gas run through the inlet (Fig. 4). These high Δ_{48} values have corresponding high Δ_{47} values offset from the known value (zero in the case of reference gas run against itself) and in line with the relationship described above. The result is that we observed a larger range in Δ_{48} values for 'clean' samples than previously deemed acceptable.

We do not know the cause of this fractionation, but we can rule out some potential explanations. We see the relationship between Δ_{47} and Δ_{48} in heated gases originating from multiple tanks and in the reference gas run as a sample, all of which are pure CO_2. This suggests that our fractionation is not due to sample contamination. The fractionation is also not produced during the acid-digestion step because it is observed in standard gases that do not get reacted (and are measured both with and without an acid bath attached to the inlet). Variations in yield (Fig. 3) also do not cause the fractionation because the residuals on this yield do not correlate with Δ_{48} (Supplementary Fig. S7, Supporting Information).

One possible explanation is that this fractionation is produced in the PPQ cleaning step. Previous studies measuring extremely small samples (<15 µmol CO_2 equivalent to <1.5 mg $CaCO_3$) observed a fractionation in Δ_{47} associated with the gas chromatography cleaning step, producing Δ_{47} values up to
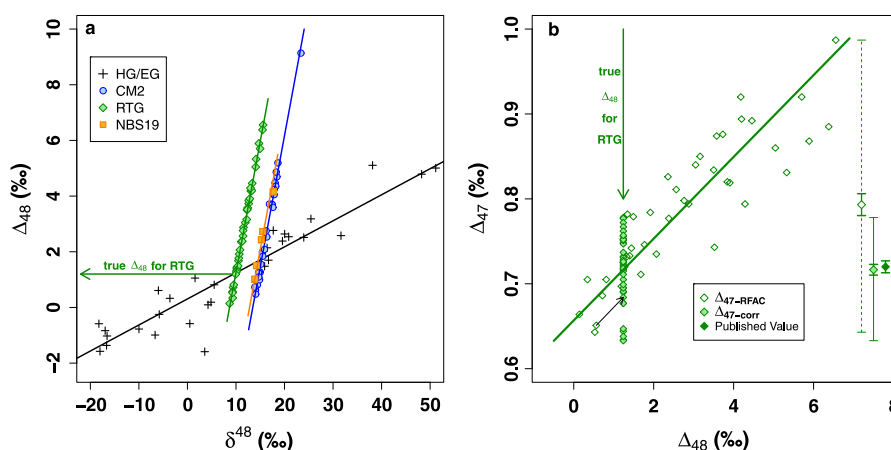
Figure 5. Diagram of correction method for RTG data points from February to March 2014. (a) δ_{48} vs Δ_{48} for three carbonate sample types and for heated and equilibrated gases (HG/EG). The line fitted to the HG/EG data (black) gives $Slope_{HG/EG}$ and $Int_{HG/EG}$ and the lines fitted together to the carbonate data (colored) give one $Slope_{CARB48}$ and individual Int_{CARB48} values. (b) Δ_{48} vs Δ_{47} for RTG data corrected ($\Delta_{47\text{-corr}}$) and uncorrected ($\Delta_{47\text{-RF/AC}}$) for the Δ_{48} fractionation. In both panels, the green arrow denotes the 'true' Δ_{48} value. The small black arrow indicates the manner in which points are corrected back to the 'true' Δ_{48} value along the slope. Bars on the right indicate the full range of points, corrected and uncorrected. Large points and error bars on the right indicate the mean and standard error (1 SE) of the corrected and uncorrected data sets. The published value $(0.720 \pm 0.007\text{‰})$[18] is plotted for comparison.

0.2‰ higher than in larger samples.[23] This increase in Δ_{47} is of the same order of magnitude as the fractionation observed in our data (Fig. 4). However, the authors of that study observed no correlation between Δ_{47} and either mass-48 or mass-49 excesses.[23]

We observe different relationships between Δ_{48} and Δ_{47} when gas does or does not pass through the PPQ trap. Reference gas introduced directly into the small U-trap (therefore bypassing the PPQ step) did not show a strong relationship in Δ_{48} vs Δ_{47} (Supplementary Fig. S8, Supporting Information). However, reference gas frozen directly into the small U-trap (instead of expanded) shows a fairly strong correlation in Δ_{48} vs Δ_{47}, but with a slope almost twice as steep as when the same gas passed through the whole inlet (Supplementary Fig. S8, Supporting Information). Both the amount of gas frozen into the U-trap and the voltage at which the gas was run did not correlate with the magnitude of the fractionation, although none were as small as our smallest samples (Supplementary Fig. S9, Supporting Information). Replacing the PPQ material did not change the slope (Supplementary Fig. S3, Supporting Information). The duration of baking the PPQ trap before passing a new sample through was observed to have some influence (longer baking = lower Δ_{48}), but this effect was inconsistent.

The sample size dependence of this fractionation points to a possible influence of the pressure-dependent negative baselines (PBLs) observed in many other studies,[24–26] which would have the strongest influence on samples run at the lowest signal intensity. We did not measure the PBLs in this study, but the correction to the absolute reference frame implicitly takes the PBL effect into account if reference gases and samples are run at the same voltage. In this study, the calibration gases spanned a range of sizes that generally overlapped with the range of sample sizes measured

(Supplementary Fig. S10, Supporting Information). Some of our smallest samples are outside this range. However, looking at a subset of samples and calibration gases run over a narrow voltage range (initial m/z 47 = 3300–3800 mV), we still observe the same relationship, both before and after correction to the absolute reference frame (Supplementary Fig. S2 and Table S2, Supporting Information). This suggests that the size (and running voltage) of the reference and sample gases is not causing the observed fractionation.

Heated gases show a smaller range of fractionation than the carbonate samples (Fig. 4). The volume of CO_2 in each aliquot of heated or equilibrated gas was often larger than the volume of CO_2 produced by a typical carbonate sample (especially compared with our smallest samples) due to the method of preparation of aliquots from the reference tanks. After passing through our sample preparation inlet (described above), the heated and equilibrated gases were chopped down to a volume comparable with the carbonate samples so they could be run at similar beam intensities (Supplementary Fig. S10, Supporting Information). This was done by expanding the gas into the T-junction adjacent to the small U-trap (Fig. 1), or by closing off the reservoir, evacuating the small U-trap, and expanding the gas back into the U-trap. The smaller magnitude of fractionation in the heated gases suggests that the volume of gas passing through the PPQ trap may play a role in setting the magnitude of fractionation.

Another possibility is that this fractionation is caused by our dual-reservoir configuration and the way in which the beam intensities decrease throughout the run. However, Δ_{48} and Δ_{47} show little trend over the 2-h 20-min run, and the Δ_{48} value, whether low or high, is identifiable in the first few cycles, which are no different from the first few cycles of a traditional two-bellows analysis. Regardless of the cause of the fractionation, it is consistent, well defined, and can be corrected for.

Correcting for a fractionation in Δ_{47} and Δ_{48}

The observed relationship between increased Δ_{48} and Δ_{47} causes an undesirable, large scatter in the Δ_{47} values of measured carbonate samples and standards. If a 'true' Δ_{48} value is known for each sample type (e.g. CM2, RTG), the Δ_{47} data can be simply corrected back to this value along the observed slope. In the past, the 'true' (uncontaminated) Δ_{48} value has been determined in relation to the behavior of heated and equilibrated gases in δ^{48} vs Δ_{48} space.[19] Heated and equilibrated gases show a positive linear relationship between δ^{48} and Δ_{48} values, and replicates of a single carbonate sample form an intersecting line of steeper slope (see Fig. 5(a)). The deviation between the Δ_{48} of an individual replicate and the line determined by the heated/equilibrated gases is denoted $\Delta(\Delta_{48})$ and is calculated based on Eqn. (2):

$$\Delta(\Delta_{48}) = \Delta_{48} - \left(\delta^{48} * Slope_{HG/EG} + Int_{HG/EG} \right) \qquad (2)$$

where $Slope_{HG/EG}$ and $Int_{HG/EG}$ are the slope and intercept of the line fitted to all the heated and equilibrated gas points in δ^{48} vs Δ_{48} space (Fig. 5(a), black points/line). The 'true' Δ_{48} value for each carbonate sample type (CM2, RTG, unknown) is then determined to be the value at which $\Delta(\Delta_{48}) = 0$ (i.e. the point of intersection between the sample line and the heated/equilibrated gas line). This can be calculated for each sample type using Eqn. (3):

$$\text{'true' } \Delta_{48} = \frac{\left(Int_{CARB48} * Slope_{HG/EG} - Int_{HG/EG} * Slope_{CARB48} \right)}{\left(Slope_{HG/EG} - Slope_{CARB48} \right)} \qquad (3)$$

where $Slope_{CARB48}$ and Int_{CARB48} are the slope and intercept of the line fitted to all the data of a single carbonate sample type in δ^{48} vs Δ_{48} space (Fig. 5(a), colored points/lines). $Slope_{CARB48}$ and Int_{CARB48} can be determined individually for each sample type by fitting a line to all replicates of that sample. The individual δ^{48} vs Δ_{48} slopes for carbonates agree with each other within error when sample types have enough replicates (Supplementary Table S5, Supporting Information) suggesting that all carbonates are following the same slope and the lines are parallel. Therefore, data sets of all carbonate sample types can be fitted simultaneously for a common slope and individual intercepts. This is the same as the mathematical procedure performed to solve for the slope and intercepts of the heated and equilibrated gases in δ^{47} vs Δ_{47} space when correcting data to the absolute reference frame.[2] By solving all data sets together, this decreases the influence of any individual outlying point, and improves the fit to a sample for which you have few replicates (e.g. NBS19). Slopes and intercepts resulting from the group fit are the same as those from individual fits for CM2 and RTG within error, but not for NBS19, which has only 2–5 replicates per measurement period (See Supplementary Table S5, Supporting Information). The close agreement between slopes of different sample types, and the benefits when measuring unknowns with few replicates, lead us to prefer the group fit for this step. (See the Supporting Information for an in-depth discussion of group fit vs individual fit.) The δ^{48} vs Δ_{48} lines in this study have been calculated using a group fit of CM2, RTG, and NBS19 together. The resulting single $Slope_{CARB48}$ and multiple Int_{CARB48} values are inserted into Eqn. (3) to calculate the 'true' Δ_{48}.

Once the 'true' Δ_{48} value is determined for each sample type, the scattered Δ_{47} data can be corrected to this value using Eqn. (4), following the observed relationship between Δ_{47} and Δ_{48} (see Fig. 5(b)):

$$\Delta_{47\text{-corr}} = \Delta_{47\text{-RF/AC}} - (\Delta_{48} - \text{'true' } \Delta_{48}) * Slope_{CARB47} \qquad (4)$$

where $\Delta_{47\text{-RF/AC}}$ is the Δ_{47} value corrected to the absolute reference frame and adjusted for the acid-digestion fractionation.[2] We choose to apply this correction to the data already corrected into the absolute reference frame since any nonlinearities caused by source effects should be removed in the $\Delta_{47\text{-RF/AC}}$ data. However, the slopes of the carbonate lines in Δ_{48} vs Δ_{47} space ($Slope_{CARB47}$) are very similar for the reference frame-corrected and raw data (Supplementary Fig. S2 and Tables S2 and S3, Supporting Information), so the choice of order of operations does not have a large effect (mean difference of 0.001‰).

The $Slope_{CARB47}$ can be found by fitting replicates of a single sample type in Δ_{48} vs Δ_{47} space, or by fitting data sets of all carbonate sample types simultaneously for a common slope. When correcting unknown data with a small number of replicates (e.g. NBS19), it is preferable to perform the group fit again (see discussion in the Supporting Information). When correcting samples or standards with a large (>10–15) number of replicates, the individual fit should accurately capture the $Slope_{CARB47}$ for that sample type. For carbonate samples, $Slope_{CARB47}$ is roughly equal to 0.05 and this value does not change much through time (Supplementary Fig. S3, Supporting Information). We do observe a slight but consistent difference between the slopes of CM2 and RTG (Supplementary Table S3, Supporting Information), which may be Δ_{47}-dependent. To preserve the difference in slope between CM2 and RTG while still getting the benefits of the group fit for NBS19, we chose to fit RTG individually, and CM2 and NBS19 together for this step. The small number of NBS19 points have a minimal effect on the group fit, making the resulting slope identical within error to the CM2 individual fit. A comparison of data corrected using different fit procedures is shown in the Supporting Information. In future studies of unknowns, it would be preferable to have multiple replicates of each sample that can be corrected together (i.e. run in the same measurement period) to constrain $Slope_{CARB47}$ for each sample individually. Where this is not possible, samples of similar Δ_{47} should be fitted together. This can either use a data set of all unknowns, where at least a few samples have multiple replicates, or it could include a standard of similar Δ_{47} composition, run many times. Wherever possible, the behavior of a new unknown sample type should be independently constrained to help determine which standard is appropriate for a group fit. Group fits with a standard material following a different slope can result in corrected data that is skewed too high or too low.

Corrected carbonate data

Applying this correction dramatically reduces the scatter in our Δ_{47} data and shifts the mean value to within error of the published values (see Table 1). A full comparison of corrected and uncorrected data can be found in Supplementary Fig. S11 and Table S4 (Supporting Information). For CM2, the

Table 1. Summary of all carbonate samples of sizes 1.0–2.6 mg measured from September 2013 to March 2014 compared with published values. The $\delta^{13}C$ and corrected Δ_{47} values compare well with published values. The $\delta^{18}O$ values are consistently 0.15–0.2‰ too light. Future analysis of unknown carbonates will be corrected for this fixed offset. All listed errors are 1 standard error (1 SE) estimates on the mean of all points of that sample type, with the number of replicates listed in parentheses

Sample	$\delta^{13}C$ (‰)	$\delta^{18}O$ (‰)	$\Delta_{47\text{-RFAC}}$ (‰)	$\Delta_{47\text{-corr}}$ (‰)
CM2 (n = 108)	**2.22 ± 0.004**	**−1.95 ± 0.008**	0.429 ± 0.010	**0.378 ± 0.003**
CM2 published[2]	2.29 ± 0.006	−1.77 ± 0.010		0.385 ± 0.005 (n = 40)
RTG (n = 76)	**−2.19 ± 0.006**	**−4.34 ± 0.007**	0.773 ± 0.010	**0.723 ± 0.004**
RTG published[18]	−2.2 ± 0.062	−4.11 ± 0.019		0.720 ± 0.007 (n = 11)
NBS19 (n = 9)	**1.95 ± 0.019**	**−2.34 ± 0.037**	0.372 ± 0.021	**0.366 ± 0.007**
NBS19 published[2]	1.95	−2.2		0.373 ± 0.007 (n = 7)

correction shifts the mean Δ_{47} from 0.429 ± 0.010‰ to 0.378 ± 0.003‰ (n = 108). For RTG, the correction shifts the mean Δ_{47} from 0.773 ± 0.010‰ to 0.723 ± 0.004‰ (n = 76). These values compare nicely with the published values of 0.385 ± 0.005‰ (n = 40) for CM2 and 0.720 ± 0.007‰ (n = 11) for RTG.[2] The offset between the uncorrected and corrected mean values reflects the number of points in a certain data set above or below the 'true' Δ_{48} value. This is not a fixed value and would not be expected to be the same for two similar standards (CM2 and NBS19) or even two subsets of the same standard (such as the two subsets of NBS19 in Table 1). The $\delta^{13}C$ values agree well with published data, whereas the $\delta^{18}O$ values are consistently too light by 0.15–0.2‰ from known values (Table 1). When measuring unknowns in the future, stable isotope data of unknown samples will be corrected for this systematic offset using carbonate standards, as has been done in other studies.[27]

Figure 6 shows all the corrected data for CM2 and RTG. The data is separated into 0.1 mg bins and samples in each bin are treated as replicates of one sample. In each bin, the average and standard error are computed (Fig. 6, filled symbols). For both CM2 and RTG, there is no systematic increase in the scatter of individual points or in the external standard error of the binned groups as the sample size is reduced. The RTG data points show more scatter than the CM2 data points ($\Delta_{47\text{-corr}}$ 1 s.d. = 0.036 vs 0.031, all replicates). This might be explained by increased heterogeneity of the sample material between replicates. The RTG coral material was more coarsely ground than the CM2 marble, and shallow-water corals are known to be heterogeneous in Δ_{47}.[3,28] In addition, fewer replicates of RTG were run in many of the mass bins than of CM2 (Supplementary Table S4, Supporting Information), leading to larger standard errors for the same mass bins.

A smaller number of replicates of NBS19 were run, producing a mean Δ_{47} value of 0.366 ± 0.007‰ (n = 9). This compares well with the value of 0.373 ± 0.005‰ calculated from previous measurements of this standard at Harvard.[2] Replicates of NBS19 fall into two mass groups – a larger group (2.3–2.5 mg) and a smaller group (1.2–1.4 mg). The means of these two groups (0.365 ± 0.011‰ (n = 5) and 0.367 ± 0.010‰ (n = 4), respectively) were within error of each other and of the published value (see Table 2).

The errors (1 SE) on the raw Δ_{47} values of individual samples increase as the sample size decreases following the shot noise limit[21] (see Supporting Information). At 1 mg, the error bars on the $\Delta_{47\text{-raw}}$ values for individual replicates are 0.014–0.016‰ for both CM2 and RTG. At 1.5 mg, the same error bars are 0.009–0.011‰. When these raw 1 SE errors are propagated through the reference frame and Δ_{48} corrections, they become significantly larger and the influence of sample size disappears (see Supporting Information for details of the error propagation calculations). The average error increases from 0.009‰ to 0.034‰. The majority of this increase comes from the reference frame correction, which is responsible for 0.019‰ of the 0.025‰ increase, whereas the Δ_{48} correction contributes only 0.007‰ on average. The Δ_{48} correction had the largest impact during measurement period #1, where it contributed a 0.013‰ increase to the average error, due to the limited number of carbonate samples run during this period (n = 15), causing greater uncertainty in the carbonate slopes and intercepts needed for the correction. The reference frame correction contribution was smallest in measurement period #2 (only 0.006‰), where the most standard gases were run. Other measurement periods (#1, #3) were cut short due to mechanical issues (e.g. power outages) before the desirable number of gas standards were run, resulting in larger uncertainties in the heated and equilibrated gas lines and the empirical transfer function during these measurement periods.

DISCUSSION

Precision and sample size

The benefit of this technique is the ability to measure Δ_{47} on small aliquots of carbonate. All applications of clumped isotopes require at least 3 replicates per unknown to reduce the uncertainty on the mean. For geological applications, errors of the order of 0.010‰ are needed to obtain meaningful results. To improve the overall sample requirements (combined mass of all replicates), it is necessary to balance the sample size per replicate and the number of replicates needed to achieve the necessary precision.

In applying this method, we recommend analyzing 5–6 replicates of mass 1.2–1.4 mg per unknown (equivalent to 6–8.4 mg of $CaCO_3$) to minimize the total amount of sample material needed while achieving acceptable precision. Our measurements generally show 0.007–0.010‰ external errors in this mass range. For CM2, mass bins in this range give standard errors of 0.007–0.008‰ for 5 or 6 replicates. The
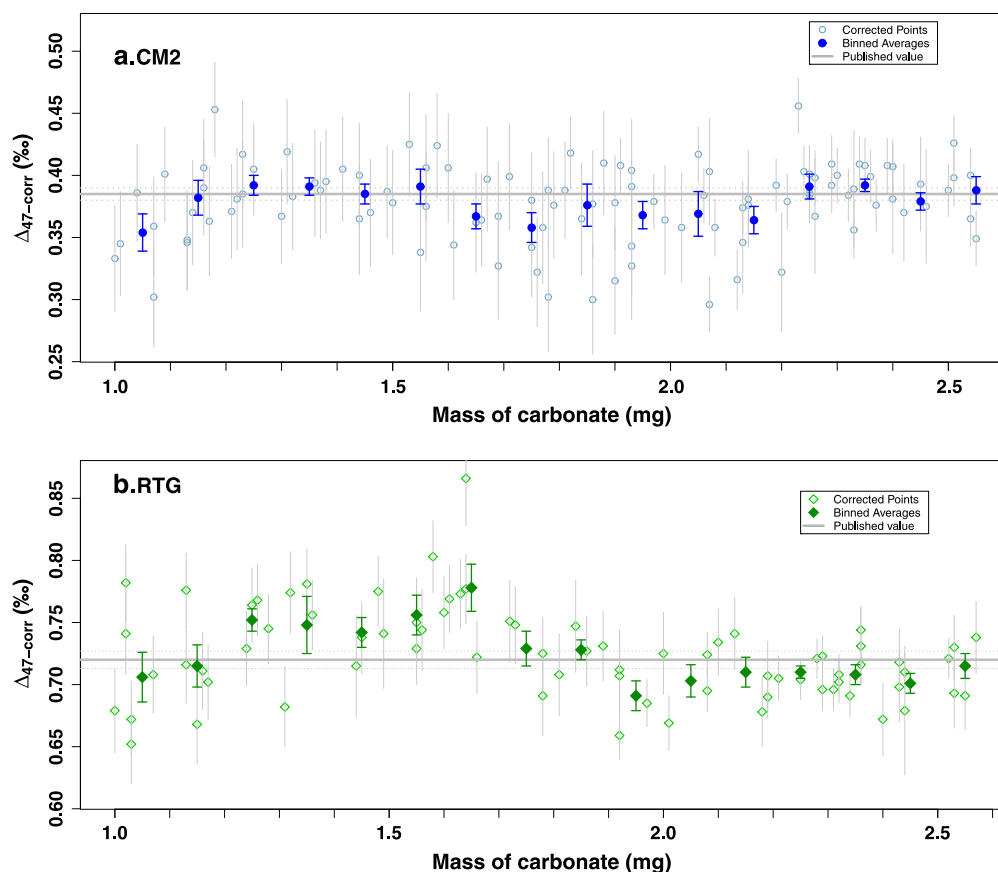
Figure 6. (a) Fully corrected $\Delta_{47\text{-corr}}$ vs Mass for all CM2 points. Individual replicates are shown with error bars representing the original shot noise error fully propagated through all correction steps (see Supporting Information). Binned averages (0.1 mg bins) are shown in filled symbols. Grey horizontal lines indicate published value and error. (b) Same as (a) but for RTG.

Table 2. Raw data used in the correction of NBS19 data points measured from September 2013 to March 2014. Errors on $\Delta_{47\text{-raw}}$ are 1 SE on all cycles measured for an individual run. Errors on $\Delta_{47\text{-RF/AC}}$ and $\Delta_{47\text{-corr}}$ have taken the raw error and propagated it through the reference frame and Δ_{48} correction, respectively. The average errors (1 SE) on the δ^{47}, δ^{48}, and Δ_{48} values are 0.011‰, 0.053‰, and 0.052‰, respectively. The errors reported on the mean values represent external 1 SE of that sample group

Date	Sample	Mass (mg)	δ^{47} (‰)	$\Delta_{47\text{-raw}}$ (‰)	$\Delta_{47\text{-RF/AC}}$ (‰)	δ^{48} (‰)	Δ_{48} (‰)	$\Delta_{47\text{-corr}}$ (‰)
12/10/13	NBS19	2.32	17.317	-0.587 ± 0.008	0.307 ± 0.017	12.926	0.235	0.379 ± 0.022
12/13/13	NBS19	2.41	17.494	-0.547 ± 0.007	0.346 ± 0.016	14.026	1.147	0.366 ± 0.021
01/13/14	NBS19	2.37	17.463	-0.564 ± 0.006	0.329 ± 0.040	14.220	1.407	0.327 ± 0.045
02/05/14	NBS19	2.47	17.491	-0.546 ± 0.006	0.348 ± 0.039	13.564	0.658	0.390 ± 0.044
03/25/14	NBS19	*1.23*	17.497	-0.449 ± 0.012	0.414 ± 0.036	15.537	2.729	0.356 ± 0.039
03/25/14	NBS19	*1.36*	17.401	-0.484 ± 0.010	0.379 ± 0.037	14.217	1.502	0.387 ± 0.039
03/27/14	NBS19	2.32	17.499	-0.553 ± 0.006	0.327 ± 0.037	13.889	1.022	0.363 ± 0.039
03/27/14	NBS19	*1.27*	17.480	-0.477 ± 0.011	0.385 ± 0.037	15.229	2.431	0.343 ± 0.040
03/27/14	NBS19	*1.33*	17.102	-0.345 ± 0.010	0.517 ± 0.034	17.634	4.151	0.380 ± 0.037
ALL				$\mathbf{-0.504 \pm 0.025}$	$\mathbf{0.372 \pm 0.021}$			$\mathbf{0.366 \pm 0.007}$
Group 1 (1.2–1.4 mg, n = 4)				-0.439 ± 0.032	0.331 ± 0.007			0.367 ± 0.010
Group 2 (2.3–2.5 mg, n = 5)				-0.555 ± 0.009	0.424 ± 0.032			0.365 ± 0.011
Published value								*0.373 ± 0.007 (n = 7)*

smaller NBS19 mass bin (1.2–1.4 mg) has a standard error of 0.010‰ for 4 replicates. For RTG, the 1.2–1.3 mg mass bin has a standard error of 0.009‰ for 4 replicates. The one exception is the larger RTG mass bin (1.3–1.4 mg), which has a higher error (0.023‰), but only 4 replicates. The addition of 1–2 more replicates would improve the error in these mass bins. Even with a larger number of replicates, this is about half the mass required for the best existing traditional techniques (~4 replicates at 3–4 mg[10]) and is similar to the lower limit of measurements made with the Kiel device technique[11,12] (6–26 mg).

Detecting contaminated samples

One issue that arises is how to detect contaminated samples if a perfectly clean sample can have a high Δ_{48} value due to this unexplained fractionation. We suggest that contamination would be unlikely to affect Δ_{47} and Δ_{48} in exactly the same ratio as our fractionation. Contaminated samples would therefore deviate from the observed relationship, probably in the direction of higher Δ_{48}. In this study, 7 samples deviated significantly from the fractionation line in the direction of high Δ_{48} (Supplementary Fig. S12, Supporting Information), and we believe them to be contaminated. These few points have been excluded from all the plots and values shown here.

In practice, a threshold or envelope around the fractionation line should be set and points outside this range should be discarded, similar to how the heated gas line in δ^{48} vs Δ_{48} space has been used previously. When measuring unknowns, the residuals of the unknown data around an unknown's group fit line can be compared with the residuals of carbonate standards. Scatter in excess of the typical carbonate standard (especially in the high Δ_{48} direction) should be deemed contamination. This method of determining contamination does, however, carry the risk of failing to eliminate some 'mildly contaminated' samples. Inclusion of these samples would result in lower Δ_{47} and temperatures that were too hot.

A lower limit on sample size

We have shown that for sample sizes as low as 1 mg, the average of multiple replicates can faithfully reproduce the mean value and give standard errors within an acceptable range (0.005–0.010‰), similar to that achieved by traditional methods.[4,10,16,17]

The inlet sample preparation procedure does not seem to limit the sample size that can be run because for all sample sizes we obtain good yields consistent with the ideal gas law (Fig. 3). Instead, we are limited by the beam intensities at which the sample is run, which are intimately linked to the shot noise limit and the maximum achievable precision on each replicate. As voltages drop below ~500 mV on m/z 47, we observe less stability in the isotope ratios through the run, despite obtaining correct mean values when many replicates of this size are measured. This may indicate the lower limit at which the mass spectrometer can operate. The gas pressure (and therefore beam intensity) produced by a given mass is affected by a few parameters: (1) the size of the small U-trap; (2) the sensitivity of the MAT 253; and (3) the size of the sample and reference reservoirs.

Installing a smaller U-trap in front of the sample reservoir decreased the sample size necessary to achieve an internal precision of 0.010‰ and 0.007‰ from 3.1 mg to 1.6 mg and

from 3.7 mg to 2.4 mg, respectively. Currently, the internal volume of the small U-trap is about 6 mL. If this were decreased further to 3 mL, it would further decrease the above sample sizes to 1.4 mg and 2.1 mg, respectively. In an ideal case, the volume of the U-trap would be as small as possible while still allowing the sample to freeze and expand without fractionating.

After tuning the source using the Isodat autotune function, we observed an increase in average sensitivity from 1.94×10^{10} to 2.66×10^{10} mV on m/z 47 per mbar source vacuum pressure. This is equivalent to an increase from m/z 47 = ~2900 mV to ~4100 mV at ~35 mbar pressure in the bellows. This caused the sample size necessary to achieve an internal precision of 0.010‰ and 0.007‰ to decrease further from 1.6 mg to 1.4 mg and from 2.4 mg to 2.0 mg, respectively. An instrument with higher sensitivity could measure even smaller samples and achieve the same level of precision.

Changing the size of the sample (and reference) reservoir(s) could also increase the signal from a given volume of sample gas, by the same mechanism as decreasing the size of the small U-trap. However, decreasing the reservoir size also increases the rate at which gas depletes. The 10 mL reservoir used in this study is the smallest available from Swagelok (Supplementary Fig. S1, Supporting Information), but this technique should work equally well for a smaller reservoir. For very small reservoir sizes, the gas may deplete away before the 2-h run completes. In other methods using microvolumes, either the gas was not run for a very long time[11,12] or additional corrections for fractionation in stable isotopes were necessary.[29] We selected the 10 mL reservoir size to balance the sample yield and the depletion rate.

Our high-efficiency dual-reservoir configuration, in its current state, can handle samples as small as 1 mg. At this low limit, we begin to see more instabilities in the isotope ratios. A similar inlet with a smaller U-trap, a slightly smaller reservoir, or higher sensitivity of the mass spectrometer, could push the sample size even lower using this measurement technique.

CONCLUSIONS

This study demonstrates that sample sizes as small as 1 mg can be measured using a newly constructed, high-efficiency, dual-reservoir inlet, and 5–6 replicates of 1.2–1.4 mg can reliably produce external standard errors in the acceptable range (0.007–0.010‰). By eliminating the sample bellows and installing matching fixed reservoirs from which gas gradually enters the source, the volume of 'wasted' gas typically left in a sample vial or bellows is reduced and the yield is increased. With some adjustments, this measurement configuration could be successful for even smaller sample sizes. This achievement will facilitate the application of the clumped isotope proxy to sample materials for which it is difficult to acquire a large amount of material, such as foraminifera, and will expand the use of this proxy in fields such as paleoceanography.

REFERENCES

[1] J. M. Eiler. Paleoclimate reconstruction using carbonate clumped isotope thermometry. *Quaternary Sci. Rev.* **2011**, *30*, 3575.

[2] K. J. Dennis, H. P. Affek, B. H. Passey, D. P. Schrag, J. M. Eiler. Defining an absolute reference frame for 'clumped' isotope studies of CO_2. *Geochim. Cosmochim. Acta* **2011**, *75*, 7117.

[3] P. Ghosh, J. Adkins, H. P. Affek, B. Balta, W. Guo, E. A. Schauble, D. P. Schrag, J. M. Eiler. ^{13}C-^{18}O bonds in carbonate minerals: A new kind of paleothermometer. *Geochim. Cosmochim. Acta* **2006**, *70*, 1439.

[4] K. J. Dennis, D. P. Schrag. Clumped isotope thermometry of carbonatites as an indicator of diagenetic alteration. *Geochim. Cosmochim. Acta* **2010**, *74*, 2736.

[5] E. A. Schauble, P. Ghosh, J. M. Eiler. Preferential formation of ^{13}C-^{18}O bonds in carbonate minerals, estimated using first-principle lattice dynamics. *Geochim. Cosmochim. Acta* **2006**, *70*, 2510.

[6] W. Guo, J. L. Mosenfelder, W. A. Goddard III, J. M. Eiler. Isotopic fractionations associated with phosphoric acid digestion of carbonate minerals: Insights from first-principles theoretical modeling and clumped isotope measurements. *Geochim. Cosmochim. Acta* **2009**, *73*, 7203.

[7] P. Ghosh, C. N. Garzione, J. M. Eiler. Rapid uplift of the Altiplano revealed through ^{13}C-^{18}O bonds in paleosol carbonates. *Science* **2006**, *311*, 511.

[8] B. H. Passey, N. E. Levin, T. E. Cerling, F. H. Brown, J. M. Eiler. High-temperature environments of human evolution in East Africa based on bond ordering in paleosol carbonates. *Proc. Natl. Acad. Sci.* **2010**, *107*, 11245.

[9] R. A. Eagle, T. Tutken, T. S. Martin, A. K. Tripati, H. C. Fricke, M. Connely, R. L. Cifelli, J. M. Eiler. Dinosaur body temperature determined from isotopic (^{13}C-^{18}O) ordering in fossil biominerals. *Science* **2011**, *333*, 443.

[10] S. Zaarur, G. Olack, H. P. Affek. Paleo-environmental implication of clumped isotopes in land snail shells. *Geochim. Cosmochim. Acta* **2011**, *75*, 6859.

[11] T. W. Schmid, S. M. Bernasconi. An automated method for 'clumped-isotope' measurements on small carbonate samples. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1955.

[12] A. N. Meckler, M. Ziegler, M. I. Millan, S. F. M. Breitenbach, S. M. Bernasconi. Long-term performance of the Kiel carbonate device with a new correction scheme for clumped isotope measurements, *Rapid Commun. Mass Spectrom.* **2014**, *28*, 1705.

[13] A. K. Tripati, R. A. Eagle, N. Thiagarajan, A. C. Gagnon, H. Bauch, P. R. Halloran, J. M. Eiler. ^{13}C-^{18}O isotope signatures and 'clumped isotope' thermometry in foraminifera and coccoliths. *Geochim. Cosmochim. Acta* **2010**, *74*, 5697.

[14] A.-L. Grauel, T. W. Schmid, B. Hu, C. Bergami, L. Capotondi, L. Zhou, S. M. Bernasconi. Calibration and application of the 'clumped isotope' thermometer to foraminifera for high-resolution climate reconstructions. *Geochim. Cosmochim. Acta* **2013**, *108*, 125.

[15] A. K. Tripati, S. Sahany, D. Pittman, R. A. Eagle, J. D. Neelin, J. L. Mitchell, L. Beaufort. Modern and glacial tropical snowlines controlled by sea surface temperature and atmospheric mixing. *Nat. Geosci.* **2014**, *7*, 205.

[16] U. Wacker, J. Feibig, B. R. Schoene. Clumped isotope analysis of carbonates: comparison of two different acid digestion techniques. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 1631.

[17] N. Thiagarajan, J. Adkins, J. M. Eiler. Carbonate clumped isotope thermometry of deep-sea corals and implications for vital effects. *Geochim. Cosmochim. Acta* **2011**, *75*, 4416.

[18] K. J. Dennis. Clumped isotope thermometry and its application to Earth's history, *PhD thesis*, Harvard University, Cambridge, MA, USA, **2011**.

[19] K. W. Huntington, J. M. Eiler, H. P. Affek, W. Guo, M. Bonifacie, L. Y. Yeung, N. Thiagarajan, B. H. Passey, A. K. Tripati, M. Daeron, R. Came. Methods and limitations of 'clumped' CO_2 isotope (Δ_{47}) analysis by gas-source isotope ratio mass spectrometry. *J. Mass Spectrom.* **2009**, *44*, 1318.

[20] H. P. Affek. Clumped isotopic equilibrium and the rate of isotope exchange between CO_2 and water. *Am. J. Sci.* **2013**, *313*, 309.

[21] D. A. Merritt, J. M. Hayes. Factors controlling precision and accuracy in isotope-ratio-monitoring mass spectrometry. *Anal. Chem.* **1994**, *66*, 2336.

[22] J. M. Eiler, E. A. Schauble. $^{18}O^{13}C^{16}O$ in Earth's atmosphere. *Geochim. Cosmochim. Acta* **2004**, *68*, 4767.

[23] W. Guo, J. M. Eiler. Temperatures of aqueous alteration and evidence for methane generation on the parent bodies of the CM chondrites. *Geochim. Cosmochim. Acta* **2007**, *71*, 5565.

[24] B. He, G. A. Olack, A. S. Colman. Pressure baseline correction and high-precision CO_2 clumped-isotope (Δ_{47}) measurements in bellows and micro-volume modes. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 2837.

[25] L. Y. Yeung, E. D. Young, E. A. Schauble. Measurements of $^{18}O^{18}O$ and $^{17}O^{18}O$ in the atmosphere and the role of isotope-exchange reactions. *J. Geophys. Res.* **2012**, *117*, D18306.

[26] S. M. Bernasconi, B. Hu, U. Wacker, J. Fiebig, S. F. M. Breitenbach, T. Rutz. Background effects on Faraday collectors in gas-source mass spectrometry and implications for clumped isotope measurements. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 603.

[27] G. D. Price, B. H. Passey. Dynamic polar climates in a greenhouse world: Evidence from clumped isotope thermometry of Early Cretaceous belemnites. *Geology* **2013**, *41*, 923.

[28] C. Saenger, H. P. Affek, T. Felis, N. Thiagarajan, J. M. Lough, M. Holbomb. Carbonate clumped isotope variability in shallow water corals: Temperature dependence and growth-related vital effects. *Geochim. Cosmochim. Acta* **2012**, *99*, 224.

[29] I. Halevy, W. W. Fischer, J. M. Eiler. Carbonates in the Martian meteorite Allan hills 84001 formed at $18 \pm 4°C$ in a near-surface aqueous environment. *Proc. Natl. Acad. Sci.* **2011**, *108*, 16985.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.

wileyonlinelibrary.com/journal/rcm