Methods and limitations of ‘clumped’ CO₂ isotope (Δ⁴⁷) analysis by gas-source isotope ratio mass spectrometry


The geochemistry of multiply substituted isotopologues (‘clumped-isotope’ geochemistry) examines the abundances in natural materials of molecules, formula units or moieties that contain more than one rare isotope (e.g. ¹³C¹⁸O¹⁶O, ¹⁸O¹⁸O, ¹⁵N₂, ¹³C¹⁸O¹⁶O²⁻). Such species form the basis of carbonate clumped-isotope thermometry and undergo distinctive fractionations during a variety of natural processes, but initial reports have provided few details of their analysis. In this study, we present detailed data and arguments regarding the theoretical and practical limits of precision, methods of standardization, instrument linearity and related issues for clumped-isotope analysis by dual-inlet gas-source isotope ratio mass spectrometry (IRMS). We demonstrate long-term stability and subtenth per mil precision in ⁴⁷/⁴⁴ ratios for counting systems consisting of a Faraday cup registered through a 10¹² Ω resistor on three Thermo-Finnigan 253 IRMS systems. Based on the analyses of heated CO₂ gases, which have a stochastic distribution of isotopes among possible isotopologues, we document and correct for (1) isotopic exchange among analyte CO₂ molecules and (2) subtle nonlinearity in the relationship between actual and measured ⁴⁷/⁴⁴ ratios. External precisions of ~0.01‰ are routinely achieved for measurements of the mass-47 anomaly (a measure mostly of the abundance anomaly of ¹³C-¹⁸O bonds) and follow counting statistics. The present technical limit to precision intrinsic to our methods and instrumentation is ~5 parts per million (ppm), whereas precisions of measurements of heterogeneous natural materials are more typically ~10 ppm (both 1 s.e.). These correspond to errors in carbonate clumped-isotope thermometry of ±1.2 °C and ±2.4 °C, respectively. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: carbonate thermometry; isotopologues; mass 47; clumped isotopes; precision

Introduction

The geochemistry of multiply substituted isotopologues – colloquially, ‘clumped-isotope’ geochemistry – examines the abundances in natural materials of molecules, formula units or moieties that contain more than one rare isotope; e.g. ¹³C¹⁸O¹⁶O, ¹⁸O¹⁷O, ¹⁵N₂ or ¹³C¹⁸O¹⁶O²⁻.[1] Such species participate in temperature-dependent homogeneous isotope exchange reactions that can be used as paleothermometers that rigorously constrain temperature based on the isotopic composition of one phase alone.[2] The most developed of these is the thermodynamically based carbonate clumped-isotope thermometer, which examines the extent to which ¹³C and ¹⁸O form bonds with each other in the lattices of carbonate minerals.[3,4] More generally, these species undergo distinctive fractionations during a variety of processes (diffusion, photolysis, irreversible reactions, etc.) and potentially provide unique independent constraints on the origins and budgets of natural materials. The best-developed applications of this sort concern the budget of atmospheric CO₂.[5–8]

This article describes the methods and limitations of analyses of multiply substituted isotopologues by gas-source isotope ratio mass spectrometry (IRMS), with particular focus on measurement of CO₂ isotopologues critical for carbonate clumped-isotope thermometry. Our subject has been addressed in the Methods sections and appendices of previous applied studies in this
Mass Spectrometry of Multiply Substituted Isotopologues

The ability to perform high-precision analyses of molecular ions (e.g., CO$_2^+$ or O$_2^+$) is critical for clumped-isotope geochemistry. Clumped-isotope geochemistry provides unique constraints only when one can determine how a given population of rare isotopes is distributed among molecular species; any analysis that disproportionately the analyte into its constituent atoms will destroy that information. Assuming it is possible to avoid fragmentation of the analyte, analyses of naturally occurring multiply substituted isotopologues face two additional hurdles: First, such species typically constitute only 10$^{-5}$ to 10$^{-6}$ of the compound of interest, and in some cases 10$^{-8}$ or less. This makes the clumped isotopic species easily the rarest analytical targets considered by stable isotope geochemistry. Furthermore, many of the processes of greatest interest produce relatively subtle isotopic variations. For example, the excess relative to a stochastic distribution of $^{13}$C$^{18}$O bonds in carbonates is less than 1‰ at room temperature equilibrium and varies by only ~0.005‰ per degree from 0°C to 50°C. Therefore, one must analyze abundance anomalies of multiply substituted isotopologues with precisions of 10$^{-6}$ (thousands of per mil), rivaling the best that have ever been achieved for any mass spectrometric measurement.

To date, all usefully precise measurements of multiply substituted isotopologues in natural materials have been made using a Thermo-Finnigan 253 gas-source IRMS. Such instruments are well suited to the task, as they typically analyze molecular ions and are equipped with an array of Faraday detectors, which permits relatively stable multicolonction measurements. Introduction of samples using a dual-inlet and change-over valve allows rapid, repeated intercomparisons of a sample gas with a standard gas of nominally known isotopic composition, which enables one to correct for most instrumental biases and is key to achieving subthent of per mil precision. Additionally, such instruments have mass resolving powers ($m/Δm$ ~ 200) and abundance sensitivities sufficient to separate cleanly the relatively weak beams of many multiply substituted ions from the much stronger beams for isotopically normal and singly substituted ions.

However, several issues potentially affect measurements of multiply substituted isotopologues using gas-source IRMS systems. First, the mass resolution is insufficient to resolve the many interferences that one encounters when analyzing molecular ions. This is a particular concern for clumped-isotope measurements, where even a ppb-level contaminant that contributes to an analyzed ion beam (e.g., $^{12}$C$^{35}$Cl on $^{13}$C$^{18}$O$^{16}$O) can lead to significant apparent changes in isotopic composition. In addition, the dynamically pumped electron bombardment sources used in such instruments typically yield only one ion per several hundred or more analyte molecules, making it difficult to achieve the desired low counting-statistics errors for analyses of rare isotopic species. As a consequence, we use counting systems consisting of a Faraday cup registered through a 10$^{12}$ Ω resistor to measure ion beams of multiply substituted isotopologues. These have been used for the analysis of hydrogen deuterium (HD) but have not been shown previously to be suitable for analyses that demand subthent per mil precision. Finally, electron bombardment sources fragment analyte molecules and some proportion of those fragments recombines to contribute to the population of detected ions, which may influence the measured proportion of isotopologues.

**Experimental**

**Multiply substituted isotopologues of CO$_2$**

Clumped-isotope analysis uses values of $Δ_i$ to denote the excess of isotopologue $i$ relative to the amount expected for the stochastic distribution of isotopes among isotopologues of a molecule. For the stochastic distribution, the abundance of an isotopologue is the product of the abundances of the isotopes it contains (e.g., $^{13}$C$^{18}$O$^{16}$O$^*$ = 2$^{13}$C$^{18}$O$^{14}$O), where $^*$ denotes the random distribution, and 2 is a symmetry number[20]. The variable $Δ_i$ defines the difference in abundances in terms of per mil deviation from the stochastic distribution:

$$Δ_i = \left( \frac{R_i}{R^*_i} - 1 \right) \times 1000$$

(1)

where $R_i$ is the abundance ratio of the isotopologue of interest relative to the isotopically normal isotopologue, and $R^*_i$ is that ratio in a pool of molecules having the same bulk isotopic composition but a stochastic distribution of isotopologues.[21] Wang et al.[22] derived the relationship between $Δ_i$ and the temperature-dependent equilibrium constant, $K$, for isotope exchange reactions involving one doubly substituted isotopologue. In the case of CO$_2$, for one such exchange reaction involving the doubly substituted isotopologue $^{13}$C$^{18}$O$^{16}$O$^*$ (i.e., $^{13}$C$^{16}$O$_2 + ^{12}$C$^{18}$O$^{16}$O $^{13}$C$^{18}$O$^{16}$O$^* + ^{13}$C$^{17}$O$_2$), the ratio $K/K^*$, where $K^*$ is the equilibrium constant for the stochastic distribution, is related to $Δ_i$ values as follows:

$$-1000 \ln(K/K^*) = Δ_{12^{18}O^{16}} - Δ_{13^{18}O^{16}} - Δ_{13^{17}O_2}$$

(2)

$^{13}$C$^{18}$O$^{16}$O makes up ~97% of mass-47 CO$_2$. Hence, even though existing sector mass spectrometers are unable to distinguish $^{13}$C$^{18}$O$^{16}$O from its isobars, $^{12}$C$^{17}$O$^{18}$O and $^{13}$C$^{17}$O$_2$, a measurement of the $Δ_i$ value calculated as in Eqn (1), but including contributions from all isotopologues having a nominal cardinal mass of 47, should be strongly proportional to abundances of $^{13}$C$^{18}$O$^{16}$O.[22] Thus, from Eqns (1) and (2), the temperature-dependent mass-47 anomaly ($Δ_{47}$) describing the enrichment of $^{13}$C$^{18}$O$^{16}$O in CO$_2$ relative to the stochastic distribution is defined as follows:

$$Δ_{47} = \left[ \left( \frac{R_{47}}{R_{46}} - 1 \right) - \left( \frac{R_{46}}{R_{45}} - 1 \right) - \left( \frac{R_{45}}{R_{44}} - 1 \right) \right] \times 1000$$

(3)

which is calculated as

$$Δ_{47} = \left[ \frac{R_{47}}{2R_{13} × R_{18} + 2R_{17} × R_{18} + R_{13} × (R_{17})^2} - \frac{R_{46}}{2R_{18} + 2R_{13} + R_{17} + (R_{17})^2} - \frac{R_{45}}{R_{13} + 2R_{17} + 1} \right] \times 1000$$
The numerator terms \( R_{47}, R_{46} \) and \( R_{45} \) in Eqn (3) are measured sample abundance ratios relative to mass 44, as derived from measured \( \delta_{47}, \delta_{46} \) and \( \delta_{45} \) for the sample (SA) referenced to a working gas (WG) standard of nominally known isotopic composition \( \delta_{\text{eq}} = \left[ (R_{45}/R_{44}) - 1 \right] \times 1000 \). The denominator terms are the calculated \( R_{47}, R_{46} \) and \( R_{45} \) that would occur in the sample if it had the stochastic distribution, which are based on the sample’s measured \( \delta_{47}, \delta_{46} \) and \( \delta_{45} \) values. As illustrated by the following example, one can calculate \( R_{47}, R_{46} \) and \( R_{45} \) from \( R_{17} \) and \( R_{18} \) (the abundance ratios \( ^{13}C/^{12}C, ^{17}O/^{16}O \) and \( ^{18}O/^{16}O \), respectively, for the sample) because \( R_{13} \) and \( R_{18} \) are derived from the measured \( R_{45} \) and \( R_{46} \) by assuming the stochastic distribution[10,11] and \( R_{17} \) is derived from \( R_{18} \) assuming a specific mass-dependent fractionation between \( R_{17} \) and \( R_{18} \). By definition, \( ^{13}C/^{12}C = [^{13}C]([^{16}O]/^{16}O) \) and \( ^{18}O/^{16}O = [^{16}O]/[^{18}O] + 2[^{12}C]/[^{16}O] \). Assuming the stochastic distribution, \( ^{12}C/^{16}O = 1/1+R_{13} \), \( ^{13}C/^{16}O = R_{13}/[1+R_{13}] \), \( 1^{16}O/^{18}O = 1/(1+R_{17}+R_{18}) \) and \( 1^{13}C/^{16}O = R_{17}/[1+R_{17}+R_{18}] \). Substituting these terms in \( R_{45*} = [45*]^{*}/[44*] \), we find the ratio that would occur in the sample if it had the stochastic distribution: \( R_{45*} = R_{13} + R_{17} \). Similarly, \( R_{46*} = 2R_{18} + 2R_{17}R_{18} + R_{17}^2 \) and \( R_{47*} = 2R_{18} + 2R_{17}R_{18} + R_{13}(R_{17})^2 \). Values of \( R_{48*} \) and \( R_{49*} \) are determined in an analogous fashion to \( R_{45}, R_{46} \) and \( R_{47} \), and can be used to calculate the mass-48 and mass-49 anomalies relative to the stochastic distribution:

\[
\Delta_{48} = \left[ \left( \frac{R_{48}}{R_{48*}} - 1 \right) - 2 \left( \frac{R_{46}}{R_{46*}} - 1 \right) \right] \times 1000
\]

\[
\Delta_{49} = \left[ \left( \frac{R_{49}}{R_{49*}} - 1 \right) - 2 \left( \frac{R_{46}}{R_{46*}} - 1 \right) - \left( \frac{R_{45}}{R_{45*}} - 1 \right) \right] \times 1000
\]

A numerical demonstration of how to calculate \( \Delta_{47}, \Delta_{48} \) and \( \Delta_{49} \) from mass spectrometer data is provided in MATLAB code format in the supporting information.

The fact that \( ^{18}O \) and \( ^{13}C \) values are calculated from measurements of \( R_{45} \) and \( R_{46} \) by assuming the stochastic distribution\(^{10,11}\) has two noteworthy implications. First, the terms in Eqns (3)–(5) containing \( R_{45} \) and \( R_{46} \) in the numerator are equal to unity,\(^{2}\) and thus do not contribute to the \( \Delta \) values. Second, as pointed out by Eiler and Schaeuble,\(^{3}\) Eqns (3)–(5) involve a circularity that might require an iterative calculation to circumvent. However, \( R_{47} \) is so small relative to \( R_{45} \) and \( R_{46} \) that variations in \( \Delta_{47} \) are essentially independent of \( ^{18}O \) and \( ^{13}C \). Indeed, \( \Delta_{47} \) measurements performed at Yale University (using a mass spectrometer identical to MS-I described below) of cylinder CO2 samples representing a 40‰ variation in \( ^{18}O \) and \( ^{13}C \), but equilibrated at room temperature (namely, having the same temperature-dependent \( ^{13}C/^{18}O \) ordering) demonstrate that \( \Delta_{47} \) is independent of \( ^{18}O \) and \( ^{13}C \) (Fig. 1 in the supporting information).

**Carbonate clumped-isotope thermometry**

The most developed application of clumped-isotope geochemistry to date is the carbonate clumped-isotope thermometer. If bonding among isotopes of carbon and oxygen in CO2\(^{2-}\) ionic groups in carbonate minerals were random, the abundance of each carbonate-ion isotopologue would be the stochastic/combinatory product of the abundances of the isotopes of which it is composed.\(^{1}\) However, isotopologue abundances are not equal to a stochastic distribution in the carbonate crystal lattice; ‘clumping’ of heavy isotopes (\(^{13}C, ^{18}O \)) into bonds with each other is thermodynamically favored to an increasing degree as temperature decreases.\(^{1,4}\) As a consequence, the extent to which a carbonate formed in thermodynamic equilibrium is enriched in \(^{13}C/^{18}O \) bonds (relative to the amount expected for a stochastic distribution of isotopes) can provide a measure of the temperature of carbonate growth.

Currently, it is not possible to measure directly abundances of CO2\(^2-\) ionic groups in carbonates with the necessary precision to determine a sample’s \(^{13}C/^{18}O \) bond enrichment. Instead, this enrichment can be inferred by isotopic characterization of CO2 produced by phosphoric acid digestion of carbonate, because the \(^{13}C/^{18}O \) abundance anomaly in product CO2 is proportional to the abundance anomaly in the carbonate mineral.\(^{12}\) The \( \Delta_{47} \) value of CO2 extracted at 25 °C from inorganic, synthetic calcite grown at known temperatures between 0 °C and 50 °C is observed to vary as a function of the growth temperature according to the relation:\(^{23}\)

\[
\Delta_{47} = 0.0592 \times 10^6 \times T^{-2} - 0.02
\]

A least-squares linear regression that takes into account uncertainty in each measurement reported in Ghosh et al.\(^{3}\) yields a slightly modified version of this relation, including 1σ uncertainties on the fit parameters (Fig. 1; see supporting information for details of the regression):

\[
\Delta_{47} = (0.0605 \pm 0.0014) \times 10^6 \times T^{-2} - (0.031 \pm 0.016)
\]

Note this is within uncertainty of Eqn (6). Measured \( \Delta_{47} \) values for CO2 extracted from natural surface and deep water dwelling
Coral reefs grown at known temperatures between ~5 °C and 30 °C[3,13] approximately conform to the trend defined by the synthetic calcite samples, as do results for many other natural carbonates of known or estimated growth temperatures[1,16]. Addition of these data does not significantly change temperature estimates (fit parameters are within 1σ of those listed in Eqn (7)). Hence, for consistency with all previous publications, carbonate growth temperature calculations in this article are performed using the published inorganic carbonate calibration line of Ghosh et al.[13] Although the relationship between growth temperature and Δ47 for fish otoliths is similar in slope to Eqs (6) and (7), it is offset to slightly lower (~0.01–0.02‰) values of Δ47 at any given temperature.[17]

Potential sources of uncertainty in clumped-isotope analysis and temperature estimates

Several potential sources of uncertainty may be encountered when applying clumped-isotope thermometry and other isotopologue techniques to a particular sample. These include unconstrained isotopic fractionations during sample preparation and purification as well as random (i.e. shot noise) errors and any uncharacterized isotopic fractionations that might occur during mass spectrometry. Two mass spectrometric errors or uncertainties are particularly relevant to clumped-isotope analyses in general: (1) subtle but measurable nonlinearity in the relationship between actual and measured 47/44 ratios and (2) isotopic exchange among analyte CO2 molecules associated either with fragmentation and recombination in the source or interaction with the walls of metal capillaries. As detailed below, these phenomena must be corrected for empirically based on analyses of heated CO2 gas, which have a stochastic distribution of isotopes among all possible isotopologues (implying a value of Δ47 ≡ 0‰[5,6]). Uncertainties in these empirical relationships contribute to overall analytical uncertainty. A final source of uncertainty relevant for clumped-isotope thermometry, in particular, is introduced by the empirical calibration of the thermometer itself (Eqn (6)). This uncertainty is systematic rather than random (i.e. it is shared by all unknown samples) and may differ for different classes of carbonates for unknown reasons, such as kinetic isotope effects[18,19] or crystal-structural controls on 13C/18O ‘clumping’.[4]

Carbonate acid digestion, CO2 extraction and purification

Carbon dioxide was extracted from 5 to 12 mg aliquots of carbonate powders by reaction with ~2 ml anhydrous H3PO4 at 25 °C for 8–24 h following the methods of McCrea[20] and Swart et al.[21]. After reaction, conventional vacuum cryogenic purification procedures were used to isolate product CO2 using the glass vacuum apparatus described by Ghosh et al.[3]. The cryogenically purified CO2 was then entrained in He carrier gas flowing at a rate of 3 ml/min and passed through an Agilent Tech 6890N gas chromatograph (GC) column (Supelco-Q-PLOT column with 530 µm internal diameter, 30 m long) held at −10 °C. The GC was baked out at a temperature of 150 °C between samples and at 220 °C for 6 h every 24 h. CO2 exiting the column was cryogenically collected for 40 min in a glass trap immersed in liquid N2. Following the GC step, the He carrier gas was removed while sample CO2 remained condensed in the liquid N2 trap. This procedure does not optimize CO2 collection time (which could be reduced if a multiloo trap were introduced) but purifies CO2 effectively without fractionating or losing sample gas (near 100% collection efficiency). Finally, conventional cryogenic purification procedures were repeated twice before transfer to the mass spectrometer.

Carbon dioxide for heated gas normalization measurements (described below) was prepared in quartz break seals and heated in a muffle furnace to 1000 °C for >2 h and rapidly quenched at room temperature. This procedure previously has been shown to yield CO2 that closely approaches a stochastic distribution of isotopes among all possible isotopologues; i.e. its value of Δ47 ≡ 0‰.[5] Heated gases were purified using conventional cryogenic procedures and GC processing in a manner identical to the preparation of samples before being transferred to the mass spectrometer.

Isotopic analysis

In addition to the Yale University instrument noted above, two dual-inlet gas-source Thermo-Finnigan 253IRMS systems, denoted MS-I and MS-II, were used to measure the isotopic composition of CO2 at the California Institute of Technology. The instruments were equipped with collection systems consisting of both the standard set of three Faraday cups registered through 3 × 108 Ω (for M/z 44), 3 × 1010 Ω (for M/z 45) and 1011 Ω (for M/z 46) resistors and three additional Faraday cups registered through 1012 Ω resistors (for M/z 47, 48 and 49), as described by Eiler and Schaab.[5] Sample sizes were typically approximately 50 µmol with resulting ion beam currents for M/z 44, 45, 46, 47, 48 and 49 of approximately 50 nA, 0.6 nA, 0.2 nA, 2 pA, 0.2 pA and 2 fA, respectively. The working reference gas was a CO2 standard from Oztech (δ18OVPDB = 24.9‰, δ13CVPD = −3.6‰), which was standardized by comparison with CO2 evolved from phosphoric acid reaction with NBS-19 standard. Values for δ13C reported versus VPDB and δ18O reported versus VSMOW were standardized by comparison with the working reference gas. The program Isodat 2.0 was used to calculate δ13C and δ18O from the observed isotopic abundances. For MS-I, isotopic compositions were typically measured for nine acquisitions of 10 cycles each (8 s integration time each for the reference gas and sample sides each cycle), with peak centering, background measurements and pressure balancing between the reference gas and sample sides made before each acquisition. Each acquisition requires 20–30 min, and the total analysis time per sample is typically 3–4 h. Measurements are typically made with an acceleration potential of ca 9.5 kV, an electron energy of ~65–100 V, and with the ‘sulfur window’ (an aperture on the side of the ionization chamber used to control the pressure and residence time of gas in the source) closed. The same configuration was used for MS-II, except that MS-II measurements were made for eight acquisitions consisting of seven cycles (26 s integration time each), and the widths of the Faraday cups on MS-I and MS-II differ slightly. For both machines, the capillary aperture was adjusted to yield a 5 V signal for mass 44 for a bellow pressure of 50 mbar.

Comparison of sample gases to standards of known isotopic composition and state of ordering

The Δ47 value is defined as the difference between the measured R47 value of a sample and the R47 value one would have measured for a sample of that same bulk isotopic composition (i.e. the same δ13C and δ18O values) having the stochastic distribution of isotopologues. This requires that we compare a sample gas to a...
reference gas that has both a known bulk isotopic composition and a known state of ordering (i.e. its own $\Delta_{47}$ value must be independently known). This is challenging because the common reference materials one can use to establish a sample’s bulk composition (e.g. CO$_2$ extracted from NBS-19 carbonate standard) do not have independently known $\Delta_{47}$ values. In addition, the procedure we use to create heated gases with nominally known $\Delta_{47}$ values (i.e. heating for two or more hours in a quartz tube to drive their composition to the stochastic distribution) can change their bulk isotopic compositions (primarily the $\delta^{18}$O values).

Our solution to this problem is as follows: we establish the bulk isotopic composition of a working reference gas (‘WG’) by conventional means (i.e. comparison with recognized interlaboratory standards). Then, we analyze a heated gas (‘HG’) as a sample, using WG as the standard. We calculate the bulk isotopic composition of the heated gas as if both the HG and WG had the stochastic distribution (i.e. using normal ion correction algorithms). This assumption will introduce an error if the working gas has a $\Delta_{47}$ value very far from 0, but it is generally sound for common natural materials (i.e. anything other than synthetic materials that are mixtures of exceptionally isotopically enriched or depleted compounds) because the observed $\Delta_{47}$ range is small. We then calculate $R_i/R_i$ values for observed masses ($i = 45, 46, 47$) as described above. From these, we compute $\Delta_{47}$ relative to the working gas for the HG using Eqn (3).

The process is repeated for the sample: after analyzing a sample (‘SA’) using the working gas as the standard, we calculate the bulk isotopic composition of the sample as if both the sample and WG had the stochastic distribution. We then calculate $R_i/R_i$ values for the observed masses and compute the $\Delta_{47}$ for the sample using Eqn (3). Finally, we calculate the sample’s mass-47 enrichment in excess of the stochastic distribution by subtracting the $\Delta_{47}$ for the heated gas from the $\Delta_{47}$ for the sample: $\Delta_{47}[\text{SA vs. HG}] = \Delta_{47}[\text{SA vs. WG}] – \Delta_{47}[\text{HG vs. WG}]$. (Note that $\Delta_{48}$ and $\Delta_{49}$ are calculated analogously using Eqns (4) and (5).) This final step makes the approximation that $\Delta_{47}$ values can be added and subtracted linearly. Because $\Delta_{47}$ is not a linear function of abundance, this is, strictly speaking, a convenient but incorrect approximation, but for the $\sim 1%$ range in $\Delta_{47}$ values of natural materials it adds no significant errors.

However, the mass spectrometry itself involves a nonlinearity between measured and real $R_i$ values that can introduce significant errors in $\Delta_{47}$ values if it is not carefully documented and corrected.$^{[16,17]}$ We observe linear correlations between the $\Delta_{47}$ values of heated gases measured versus the working gas and their $\delta_{47}$ values measured versus working gas. (We define $\delta_{47}[\text{HG vs. WG}] = (R_{47[\text{HG}]/R_{47[\text{WG}]}} - 1) \times 1000$, a convenient measure of bulk composition. Because $\Delta_{47}$ is a small number, $\delta_{47}$ is approximately equal to $(\delta_{45} + \delta_{49})$, and thus to $(\delta^{13}C + \delta^{18}O)$.) That is, if one analyzes two heated gases, one closely similar in $\delta_{47}$ to the working gas and a second that is substantially different in $\delta_{47}$, two different $\Delta_{47}[\text{HG vs. WG}]$ values will be observed. Note that the working reference gas is the same in these two measurements, and all heated gases have nominal absolute $\Delta_{47}$ values of 0; hence, in the absence of any measurement artifact, the two measured $\Delta_{47}$ values should be indistinguishable. Thus, the observed relationship between $\Delta_{47}$ and $\delta_{47}$ for heated gases reflects a subtle nonlinearity in the relationship between actual $R_i$ values and the measured intensity ratio between the mass-47 and mass-44 ion beams. Although we observe this nonlinearity in heated gas measurements for MS-I, MS-II and the Yale mass spectrometer, it differs in the three machines, varies gradually with time on any given instrument, and was not observed in initial studies using MS-I.$^{[5]}$ Hence, it may not be a universal phenomenon. It is not yet clear whether the ultimate cause of this nonlinearity comes from the performance of the detectors, the resistors through which ion currents are measured or some component of the source or analyzer. Nevertheless, a normalization based on the analysis of heated gases over a range of composition space is straightforward.

If the sample and heated gas have the same $\delta_{47}$, then $\Delta_{47}[\text{SA vs. HG}] = \Delta_{47}[\text{SA vs. WG}] – \Delta_{47}[\text{HG vs. WG}]$. If the sample and heated gas differ in their $\delta_{47}$ values, then $\Delta_{47}[\text{HG vs. WG}]$ corresponding to $\delta_{47}[\text{SA vs. WG}]$ must be estimated by linear regression of measured $\delta_{47}[\text{HG vs. WG}]$ on $\Delta_{47}[\text{HG vs. WG}]$ for the time-period during which the sample was analyzed (see supporting information for details of the regression and error propagation). We find that the slope of the empirically determined line for heated gases in a plot of $\delta_{47}$ versus $\Delta_{47}[\text{HG vs. WG}]$ is stable over time scales of multiple weeks or months, but can in some cases change subsequent to cleaning the source and or replacing the tungsten filament or changing the ion source focusing – presumably reflecting a change in mass spectrometer performance that influences the nonlinearity between actual and observed $R_i$ values (Fig. 2(a–c)).

In the months following our discovery of nonlinearity in $\Delta_{47}$ measurements, the intercepts of lines fit through data for heated gases in plots of $\delta_{47}$ versus $\Delta_{47}$ (‘heated gas lines’) were essentially invariant, even over large ranges in slope.$^{[5]}$ However, subsequent analyses have revealed measurable secular variations in the intercepts of heated gas lines (Fig. 2(d,e)). In order to establish the physical cause of this phenomenon, we conducted experiments in which we varied the residence time of gas in the source by opening and closing the sulfur window. This produced a large variation in the intercept of the heated gas line (Fig. 2(f)), consistent with the hypothesis that a long residence time or high gas pressure in the source promotes fragmentation/recombination reactions that ‘scramble’ isotopes among isotopologues, reducing the contrast in $\Delta_{47}$ between sample and standard gases. If so, then this artifact should be corrected for by multiplying the measured $\Delta_{47}$ values of unknowns by a factor that is proportional to the intercept of the heated gas line. In practice, we scale all data to a fixed intercept of the heated gas line, taken to be $\Delta_{47}[\text{HG vs. WG}] = -0.8453\%$ (i.e. the intercept of the heated gas line used for the calibration study of Ghosh et al.$^{[13]}$). That is, the corrected $\Delta_{47}[\text{SA vs. HG}] = \Delta_{47}[\text{SA vs. HG}]_{\text{compressed}} \times (\Delta_{47}[\text{HG vs. WG}]_{\text{intercept}} - 0.8453)$, where $\Delta_{47}[\text{HG vs. WG}]_{\text{intercept}}$ is the heated gas line intercept for the period during which the sample was analyzed.

In principle, the best way to deal with isotopic ‘scrambling’ in the source would be to open the sulfur window, minimizing residence time in the source. However, the resulting loss of ion intensity dramatically degrades measurement precision. We conclude that the best practical approach to this artifact is a compromise: we accept and correct for a certain amount of isotopic scrambling in exchange for the high precision that comes with elevated source pressure. However, it is possible that other instruments or instrument tuning conditions might require a different approach – one obviously could not accept a condition that resulted in so much ‘scrambling’ that all gases are indistinguishable in $\Delta_{47}$. 


‘Clumped’ isotopes

Precision and Stability in Isotope Ratio Measurements

Internal precision

We compared measurements of \( \Delta_{47} \) conducted using MS-I and MS-II (in continuous use since 2003 and 2008, respectively) to the predictions for \( \Delta_{47} \) precision at the shot-noise limit (Fig. 2). The empirical relationship between \( \Delta_{47} \) (MS-I and MS-II) and \( \Delta_{47} \) (MS-I and MS-II) is approximately linear for 16 V \(^{44}\)V signal heated gas observations (n) from four distinct time intervals in 2007 and 2008. The small circles are individual analyses, and the vertical error bars represent 2 s.e. uncertainty in \( \Delta_{47} \) measurements. The horizontal error bars are of the same magnitude as the vertical error bars. The solid lines represent the least-squares linear regression of data for each time-period, and the shaded gray region indicates the 95% confidence interval. (e) Best-fit lines and 95% confidence intervals (a–d) are superimposed to highlight temporal variations in the heated gas line. Heated gas lines (a), (b) and (c) have approximately the same intercept but different slopes. Heated gas line (d) has a significantly different intercept. (f) The mass-47 heated gas line is shown for different configurations of the sulfur window. The small circles are individual analyses, and the vertical error bars represent 2 s.e. uncertainty in \( \Delta_{47} \) measurements.

a variety of sources, including CO₂ generated by phosphoric acid digestion of carbonates. We compiled the internal precision (i.e. standard error) of isotopic measurements for 21 analyses of CO₂ from the NBS-19 carbonate standard and 172 analyses of different aliquots of 63 different carbonate sample unknowns performed under typical conditions from 2004 to 2008 (Fig. 5 in the supporting information). The average standard error for each \( \Delta_{47} \) measurement (0.009 \( \pm \) 0.003‰) is in agreement with the lower end of the shot-noise limit predicted for analyses of six to nine acquisitions (0.013 – 0.018‰). The average observed standard errors in \( \delta^{18}O \) and \( \delta^{13}C \) measurements of sample unknowns is roughly an order of magnitude better than that observed for \( \Delta_{47} \), although lower-precision outliers up to 0.005‰ exist. The internal precision in \( \Delta_{47} \) measurements is poorly correlated with precision of either \( \delta^{18}O \) or \( \delta^{13}C \) for the standard and sample datasets, with Pearson’s correlation coefficients \( (r) \) of 0.26 – 0.29 for NBS-19 and 0.07 – 0.09 for the samples. Internal precision in \( \delta^{18}O \) and \( \delta^{13}C \) are highly correlated with one another \( (r = 0.78 – 0.83) \), which is not surprising because unlike \( \Delta_{47} \), both \( \delta^{18}O \) and \( \delta^{13}C \) vary strongly with mass-dependent fractionation.[22]

**Long-term signal stability**

As long-term drifts can also affect measurement precision (i.e. in addition to shot noise), we examined the stability and accuracy of our IRMS system by using a modified form of the Allan variance technique, originally developed to characterize ultrastable oscillators.[23] Instead of using time as the independent variable, here we use individual IRMS acquisitions \( x \):

\[
\sigma_{\text{Allan}}^2(x) = \frac{1}{2(N - 1)} \sum_i \left[ (x)_{i+1} - (x)_i \right]^2 \quad (8)
\]

Data are first binned into \( N \) bins (each consisting of one or more acquisitions), and the average of each bin, \( (x)_i \), is taken. The difference between adjacent bins is squared and then summed before being normalized to compute the Allan variance. For example, a data set with six elements \( [x_1, x_2, x_3, x_4, x_5, x_6] \) will produce three values of Allan variance corresponding to bins of one \( ([x_1], [x_2], [x_3], [x_4], [x_5], [x_6]) \), two \( ([x_1, x_2], [x_3, x_4], [x_5, x_6]) \) and three \( ([x_1, x_2, x_3] and [x_4, x_5, x_6]) \) elements each. The maximum bin size is therefore half the total number of elements in a data set.

Plotting Allan variance versus bin size, then, conveniently presents both noise- and drift-related effects on single plot. An example of the \( \Delta_{47} \) Allan variance for a relatively long-duration analysis, consisting of 40 acquisitions of 80 s sample integration each (ten 8 s cycles), made on MS-I is shown in Fig. 3 (these are the same data depicted in Fig. 2(b–d)). An ideal statistical-noise-limited measurement, for which long-term drifts are negligible, will display a power–law relationship \( \sigma^2 = Cx^\alpha \) in the Allan variance plot, where \( \alpha = -0.5 \), reflecting the Poisson statistics governing the precision of the overall measurement. In contrast, a measurement dominated by long-term, correlated random-walk drifts will have \( \alpha = 0.5 \). Consequently, the optimum number of measurements to average together in a given analysis (i.e. the optimum bin size) will occur near where the Allan variance reaches a minimum, i.e. where \( dx/dx \approx 0 \). For our instrument, measurements seem to be dominated by statistical noise for \( N < 10 \) acquisitions, with only a small contribution from long-term drift. The variability of \( \sigma^2 \) appears to increase for measurements made in excess of nine acquisitions (Fig. 3), although we attribute this to the smaller number of bins with greater than nine acquisitions contributing to the Allan variance calculation. More precise values of the Allan variance with these numbers of acquisitions would require a more sustained analytical run by several factors in time. Nevertheless, the ideal precision that should be attainable on a single gas run for measurements that are clearly dominated by statistical noise (nine acquisitions, or 720 s counting time) at the optimal voltage (16 V) is 0.013‰.

**External Precision and Error Propagation**

**Carbonate standards and sample unknowns**

The most straightforward way to quantify external precision in \( \Delta_{47} \) measurements is to determine the reproducibility (standard deviation) of independent analyses of the same material. We compiled the standard deviations in isotope measurements for 92 externally replicated samples (2–11 replicates each) analyzed on MS-I and found that the average external precision in \( \Delta_{47} \) is 0.019\( \pm \)0.015‰, in agreement with the standard deviation of replicate analyses for the NBS-19 carbonate standard (0.022‰; Fig. 5 in the supporting information). The average external precision in \( \Delta_{47} \) is approximately twice the average standard error we observe for each CO₂ gas made by acid digestion of carbonates. In addition, the standard deviation for replicate extractions varies over a wider range than does the standard error for each extraction. This suggests that substantial variance beyond counting statistics can be added to measurements of CO₂ from some carbonate samples due to some combination of sample heterogeneity and analytical artifacts such as uncontrolled fractionations or contaminants.

We suspect that sample heterogeneity causes the failure of many carbonate materials to achieve reproducibility in \( \Delta_{47} \) comparable with counting statistics. An examination of measurements of \( \delta^{18}O \) and \( \delta^{13}C \) in the same gases (Fig. 4) reveals that standard deviations for replicate analyses of \( \delta^{18}O \) and \( \delta^{13}C \) for both sample unknowns and the NBS-19 standard are an order of magnitude larger than the average standard error we observe for each CO₂ gas made by acid digestion of carbonates (0.0014‰ for \( \delta^{18}O \) and 0.0007‰ for \( \delta^{13}C \)), well in excess of the limits imposed by counting statistics. Small errors in \( \delta^{18}O \) and \( \delta^{13}C \) that may be due to sample heterogeneity lead to small changes in calculated \( R_{47}/R_{47}^* \) values. Although this reduces the reproducibility of \( \Delta_{47} \) measurements for samples that are heterogeneous in isotopic composition.
or otherwise affected by analytical artifacts, the variability in $\Delta_{47}$ is smaller than variability in $\delta^{18}O$ and $\delta^{13}C$ because changes in $R_{47}/R_{45}$ are offset by changes in $R_{46}/R_{45}$ and $R_{45}/R_{42}$ (Eqn (3)).

Nevertheless, a large number of samples exhibit standard deviations for replicate analyses of unknowns that are much better than the average and comparable with the limits imposed by counting statistics: 49 of 92 samples exhibit standard deviations in $\Delta_{47}$ for replicate extractions of $\leq 0.010\%$o, and 23 are $\leq 0.005\%$o. These well-replicated samples are nearly twice as prevalent than would occur by random chance if the true analytical error were 0.019%o (the average for the entire population). We conclude that measurements of $\Delta_{47}$ in CO$_2$ extracted from chemically pure (i.e. free or organic matter, sulfides and other sources of contaminant gases) homogeneous carbonates have errors controlled by counting statistics and not significantly degraded by analytical artifacts.

A detailed example of analytical reproducibility for unknown natural samples

We examined replicate analyses of a suite of samples performed over a relatively long interval of time such that the stability of the stochastic reference frame is put to the test. The two biogenic carbonate samples (oyster 95123 and bivalve $anomaia$ sp. 95124) were analyzed previously by Spencer and Patchett[24] and Huntington $et$ $al.$[25] and exhibit no obvious evidence of sample heterogeneity or contamination. Replicate $\Delta_{47}$ measurements of each sample incorporate uncertainties in the heated gas reference frame (Fig. 2), and potential analytical artifacts introduced by carbonate acid digestion, CO$_2$ extraction and purification procedures. The two samples experienced identical temperature conditions during growth and burial and have the same $\Delta_{47}$ value. Thus, reproducibility of their $\Delta_{47}$ measurements also should reflect any noise introduced by sample preparation (cleaning, sampling and powdering).

Figure 4 illustrates the use of heated gas data to normalize $\Delta_{47}[SA$ vs. $WG]$ measurements of unknown samples to a common reference frame and to screen for contaminants. Divergent $\delta_{47}$ and $\Delta_{47}[SA$ vs. $WG]$ results for independent analyses of 95123 and 95124 carried out during the three different time-periods (Fig. 4(a)) collapse to the same $\Delta_{47}[SA$ vs. $HG]$ value within uncertainty (0.003–0.006‰, 1 s.e.) after the heated gas normalization is applied (Fig. 4(b,c)). Eiler and Schauble[5] found that mass-48 and

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Isotope ratio and clumped-isotope thermometry results for two biogenic carbonates. (a) $\Delta_{47}[SA$ vs. $WG]$ is plotted against $\delta_{47}[SA$ vs. $WG]$ for three independent acid digestions each of sample 95123 (white circles with 2 s.e. error bars) and sample 95124 (black circles with 2 s.e. error bars). Each analysis represents 720 s of integration time of a 16 V $44^V$ signal. (b) Heated gas normalization for 95123 and 95124 measurements made over a 5-month period. The heated gas lines for each time-period are taken from Fig. 2, and are labeled with the appropriate range of dates. The dashed arrows illustrate how the best-fit $\Delta_{47}[HG$ vs. $WG]$ value corresponding to the value of $\delta_{47}[SA$ vs. $WG]$ from (a) for each analysis is found using the heated gas line for the time-period during which the analysis was performed. The white and black squares indicate the heated gas normalization factor for three analyses of sample 95123 and 95124, respectively. (c) After the heated gas normalization from (b) is applied (see text), the $\Delta_{47}[SA$ vs. $HG]$ values for all six analyses are indistinguishable. The black line represents a zoomed-in portion of the temperature calibration line shown in Fig. 1(a). The dashed arrow illustrates how the measured $\Delta_{47}$ values are projected onto the calibration line to calculate temperature. (d) Mass-48 measurements (crosses) for the three heated gas lines shown in (b) define the relationship between $\Delta_{48}$ and $\Delta_{47}$ expected for clean samples. The samples do not deviate significantly from the best-fit heated gas mass-48 lines (composed of results for clean CO$_2$ samples), suggesting that the samples are free of hydrocarbon, halocarbon and sulfur contamination. The symbols are larger than the uncertainties.
\end{figure}
mass-49 signals can be sensitive indicators of the presence of hydrocarbons and chlorocarbons in spiked analyte CO2, and Guo and Eiler[26] found evidence that sulfur-bearing contaminants can also lead to substantial interferences on masses 48 and 49 (presumably from species such as $^{32}$S$^{16}$O$^2$). However, we have observed a strong correlation between mass-49 signals and the pressure imbalance between the bellows that regulate the sample and working gas flow into the mass spectrometer source. Values such evidence diminishes the reliability of a measurement.

Clearly could interfere with mass 48 but not 47, we suggest that contaminated samples by noting differences between their indicators of contaminant species; i.e. we can identify clearly contaminated species such as $\delta^47$ do not show this sensitivity to bellows pressure, suggesting that mass-48 measurements are more appropriate indicators of contaminant species; i.e. we can identify clearly contaminated species by noting differences between their $\delta^48 \cdot \Delta^48$ systematics and those exhibited by pure heated gases. As shown in Fig. 4(d), the mass-48 signals for six analyses plot within the range of measured heated gas mass-48 values, suggesting that the samples are free of recognized contaminants. If $\Delta^48[SA \ vs. \ WG]$ for any analysis were significantly higher than $\Delta^48[HG \ vs. \ WG]$, this would provide evidence for impurities that lead to isobaric interferences within the mass range of CO2; although contaminants clearly could interfere with mass 48 but not 47, we suggest that such evidence diminishes the reliability of a measurement.

Although their $\Delta^18O$ values differ (Table 1), $\Delta^47$ values for 95123 and 95124 are indistinguishable. Given the shared history of these samples, we conclude that these six analyses reflect a population of measurements of samples that were indistinguishable in temperature of last equilibration, and thus provide a window into the internal and external precision of the technique. The $\Delta^47[SA \ vs. \ HG]$ results of all six analyses converge by approximately 40 cycles or four acquisitions (320 s integration time), and precision approximately follows counting statistics (Fig. 5). Average observed precision in $\Delta^47$ is 0.0227‰ (1 s.e.) at the level of a single acquisition (80 s integration time) and 0.0088‰ for each analysis (i.e. summing over all acquisitions run on that gas, for a total integration time of 720 s). Average external precision in $\Delta^47$ is 0.0046‰, after integrating data for all three extractions (i.e. a total integration time of 2160 s per sample). The weighted average $\Delta^47$ values for three independent extractions each of 95123 and 95124 are 0.583 ± 0.006‰ and 0.585 ± 0.003‰, and the external error is 0.0039‰ if results from both samples (six analyses) are combined.

Precision in absolute temperature estimates varies as a function of both the temperature of carbonate growth and the uncertainty in $\Delta^47$ (Fig. 6(a); supporting information). For a single analysis of a carbonate sample grown at 20°C (typical ±0.010‰ precision in $\Delta^47$), the uncertainty in absolute temperature is ±2.2°C (1 s.e.). If three replicate measurements are made (typical ±0.005‰ precision in $\Delta^47$), the temperature uncertainty is ∼1.4°C (1 s.e.). Taking advantage of the large number of isotope ratio measurements (in our case, at least 90 cycles for each analysis), we can calculate the 95% confidence limit on temperature estimates by multiplying the standard error of the mean (1 s.e.) given by Fig. 6(a) by Student’s t-factor (1.6)[127]. Although we formally propagate uncertainties using the published calibration data,[13] we note that given the large number of additional data in agreement with the original calibration that have been generated,[1,16,28,29] actual contributions to temperature uncertainty from this calibration are likely to be very small.

The precision of temperature difference ($\Delta T$) estimates is limited by the precision of the least-precise sample $\Delta^47$ measurement and the precision of the slope (i.e. temperature dependence) of the calibration data (Fig. 6(b)). If the absolute temperature estimates for samples A and B are 30.0 ± 2.5°C and 20.0 ± 2.2°C, respectively, (1 s.e. in temperature for typical samples with ±0.010‰ precision in $\Delta^47$ from Fig. 4(a)), the temperature difference ($\Delta T = T_A - T_B$) would be 10.0 ± 3.3°C (Fig. 4(b)), on the order of the root of the sum of the squared uncertainties in $T_A$ and $T_B$. $\Delta T$ precision is slightly worse than the precision of absolute temperatures, but the error is relatively insensitive to the magnitude of the temperature difference Fig. 4(b).

### Discussion

Our analysis confirms that it is possible to measure very subtle variations – routinely with <0.1‰ precision – in the abundance of extremely rare species for clumped-isotope analysis using existing technology (Fig. 4c). For clumped-isotope carbonate thermometry, precision of 1–2°C can be achieved, depending primarily on the integration time of the analysis, and secondarily on the number of calibration data taken into consideration. These isotopologue

### Table 1

Summary of carbon and oxygen isotopic data for two biogenic carbonate samples shown in Figs. s4 and 4. Heated gas-normalized $\Delta^47$ values are reported with uncertainties propagated formally as described in the supporting information. Note that the heated gas normalization contributes little uncertainty because of the large number of independent measurements used to define the best-fit line (Fig. 2).

<table>
<thead>
<tr>
<th>Date (month/day/year)</th>
<th>$\Delta^{13}$C$_{PDB}$ (‰)</th>
<th>$\Delta^{18}$O$_{SMOW}$ (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>$\Delta^{47}$, sample (‰)</th>
<th>T (°C)</th>
<th>$\Delta^1$T (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample 95123</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/07</td>
<td>−1.84</td>
<td>0.872</td>
<td>−0.082</td>
<td>0.0083</td>
<td>12.687</td>
<td>0.0083</td>
<td>−0.661</td>
<td>0.579</td>
<td>0.0085</td>
<td>0.0095</td>
<td></td>
</tr>
<tr>
<td>3/22/07</td>
<td>−1.86</td>
<td>0.396</td>
<td>−0.147</td>
<td>0.0081</td>
<td>12.552</td>
<td>0.0082</td>
<td>−0.742</td>
<td>0.594</td>
<td>0.0082</td>
<td>0.0095</td>
<td></td>
</tr>
<tr>
<td>7/2/07</td>
<td>−1.86</td>
<td>1.041</td>
<td>−0.187</td>
<td>0.0081</td>
<td>12.482</td>
<td>0.0081</td>
<td>−0.764</td>
<td>0.577</td>
<td>0.0083</td>
<td>0.0103</td>
<td></td>
</tr>
<tr>
<td><strong>Sample average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.583</td>
<td>0.0056</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sample 95124</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/07</td>
<td>−0.98</td>
<td>1.499</td>
<td>−0.056</td>
<td>0.0077</td>
<td>14.086</td>
<td>0.0078</td>
<td>−0.643</td>
<td>0.587</td>
<td>0.0079</td>
<td>0.0091</td>
<td></td>
</tr>
<tr>
<td>3/24/07</td>
<td>−1.01</td>
<td>0.811</td>
<td>−0.140</td>
<td>0.0081</td>
<td>14.020</td>
<td>0.0079</td>
<td>−0.730</td>
<td>0.590</td>
<td>0.0082</td>
<td>0.0097</td>
<td></td>
</tr>
<tr>
<td>7/2/07</td>
<td>−1.00</td>
<td>1.474</td>
<td>−0.176</td>
<td>0.0075</td>
<td>13.949</td>
<td>0.0076</td>
<td>−0.755</td>
<td>0.579</td>
<td>0.0078</td>
<td>0.0101</td>
<td></td>
</tr>
<tr>
<td><strong>Sample average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.585</td>
<td>0.0033</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Formation average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.584</td>
<td>0.0029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[a] Without uncertainty from heated gas normalization.
[b] Including uncertainty from heated gas normalization.
measurements approach the shot-noise limit of precision (Fig. s2), indicating that further improvements will require additional methods and/or instrument development that increase the intensity of analyzed ion beams or extend the practical period of analyses of each sample. Although our data compilation indicates that external precision for $\Delta^{47}$ analysis of relatively pure, apparently homogeneous carbonates follows counting statistics to the 0.005‰ level, many other materials fail to reach this level of external precision. In some cases, we can attribute poor precision to sample inhomogeneity, whereas in other cases the cause is not obvious.

Poor mass resolution is an outstanding limitation of the mass spectrometers we used that limits our ability to examine possible contaminants or other artifacts that may be to blame for poor external reproducibility in $\Delta^{47}$ measurements. Poor mass resolution results in an inability to resolve molecular ion interferences that may have a large impact on apparent isotopologue abundances. Nevertheless, it appears that comparing sample mass-48 anomalies with those of clean heated gases may provide a means to screen for certain contaminants that may affect the isotopologue signal of interest.

The most important limitation of current mass spectrometer technology is the trade-off between ion yield per molecule of analyte gas and the occurrence of fragmentation/recombination reactions in the source, or ‘scrambling.’ Our sulfur window experiments (Fig. 2(f)) illustrate that while the efficiency with which a dynamically pumped source ionizes analyte molecules can be improved by increasing the gas pressure, increased pressure increases scrambling. As scrambling influences the measured proportion of isotopologues by driving the analyte toward the stochastic distribution, the dynamically pumped source design fundamentally limits the number of ions that can be produced and counted in a given time interval.

Encouragingly, our analysis indicates reasonable stability and linearity at levels of thousandths of per mil of ion beams of interest collected in a Faraday cup registered through a $10^{12}$ $\Omega$ resistor (Figs. s1, 3). One consequence of stability is that we can integrate small ion currents over very long time-periods.
to improve precision (Fig. 3). However, diminishing returns on precision with time and the possibility that the accuracy of isotopologue measurement may eventually decrease (Fig. s2) suggest that very long counting times (in excess of 720 s) may not be advantageous. Instead, machine time would be spent better on analysis of independently prepared replicates, which enable external precision to be characterized.

Perhaps most importantly, our examples highlight the dependence of precise clumped-isotope analysis on a heated gas reference frame. Unlike standard isotopic analyses, which are referenced to an arbitrary value (i.e. the $\delta^{18}O_{VSMOW}$ value), the heated gas reference for clumped isotopes has a unique physical meaning – by definition $\Delta_{47}$ of 0‰ means a stochastic distribution. The use of such an ‘absolute’ reference frame is as critical for practical considerations (i.e. removing the subtle nonlinearity observed in the relationship between actual $R_{47}$ values and the measured intensity ratio between the mass-47 and 44 ion beams and removing the effects of scale compression due to isotopic ‘scrambling’ in the source) as it is elegant. We recommend analyzing a heated gas each day sample unknowns are analyzed and CO$_2$ extracted from a carbonate standard such as NBS-19 occasionally to establish the $\Delta_{47} = 0$‰ reference frame precisely and monitor possible changes in machine conditions.

Summary and Outlook

Our data compilation indicates that use of a Thermo-Finnigan MAT 253 dual-inlet gas-source IRMS routinely yields subthenth of a per mil precision in $\Delta_{47}$ measurements. For relatively pure, apparently homogeneous carbonates, precision follows counting statistics, and further improvement would require hardware modifications. To the best of our knowledge, our study provides the first detailed examination of the stability and linearity of a counting system consisting of a Faraday cup registered through a $10^{12}$ Ω resistor at levels of thousandths of per mil. We find the stability of response of this system to be sufficient for precise quantitative analyses, but careful standardization is required to correct for subtle nonlinearity. We outline a method by which heated CO$_2$ gases with a stochastic distribution of isotopes among all possible isotopologues ($\Delta_{47} \equiv 0$‰) are used to correct for (1) subtle nonlinearity in the relationship between actual and measured 47/44 ratios and (2) scale compression due to isotopic exchange among analyte CO$_2$ molecules in the source or capillaries.

Measureable improvements in the precision of isotopologue analyses may be possible with instrument and methods developments. Incremental advances in method development could include increasing the ion current by reducing the resistance through which the ion beams are registered or improving the extraction efficiency. We are not optimistic about the use of carrier gas introduction because its effects on exchange among analyte CO$_2$ molecules (‘scrambling’), linearity, stability of ionization and extraction efficiency are unknown. Major advances in precision would require substantial hardware development to achieve higher mass resolution ($m/\Delta m \approx 10^5$), lower current ion counting, an improved ion-counting system, or potentially an improved vacuum system and source design. Until such improvements are realized, subthenth of a per mil precision in $\Delta_{47}$ remains at the limit of mass spectrometer measurements.

Acknowledgements

This work was supported by the National Science Foundation and by the Division of Geological and Planetary Sciences and the Davidow Fund at the California Institute of Technology. H.P.A. thanks the Earth System Center for Stable Isotope Studies of the Yale Institute for Biospheric Studies. The authors also thank John Hayes and an anonymous reviewer for insightful comments that improved this article.

Supporting information

Supporting information may be found in the online version of this article.

References

'Clumped' isotopes


[27] E. Barkan, B. Luz. High-precision measurements of \(^{17}\)O/\(^{16}\)O and \(^{18}\)O/\(^{16}\)O of \(\text{O}_2\) and \(\text{O}_2/\text{Ar}\) ratio in air. Rapid Communications in Mass Spectrometry 2005, 17, 2809.
