Characterisation of new reference materials IAEA–610, IAEA–611 and IAEA–612 aimed at the VPDB δ 13 C scale realisation with small uncertainty

Article in Rapid Communications in Mass Spectrometry - December 2020
DOI: 10.1002/rcm.9014

8 CITATIONS
228 READS

5 authors, including:

Sergey Assonov
Formerly - at International Atomic Energy Agency (IAEA)
86 PUBLICATIONS 1,467 CITATIONS

Ales Fajgelj
Home
104 PUBLICATIONS 2,263 CITATIONS

Jean-François Hélie
Université du Québec à Montréal
84 PUBLICATIONS 1,356 CITATIONS

Colin E. Allison
The Commonwealth Scientific and Industrial Research Organisation
78 PUBLICATIONS 3,839 CITATIONS

Some of the authors of this publication are also working on these related projects:

SIPCLIP project View project

Origin of CO2 in mexican geothermal fields of Los Azufres, Los Humeros and Cerro Prieto View project
Characterisation of new reference materials IAEA-610, IAEA-611 and IAEA-612 aimed at the VPDB δ^{13}C scale realisation with small uncertainty

Sergey Assonov\textsuperscript{1} | Ales Fajgelj\textsuperscript{1} | Jean-François Hélie\textsuperscript{2} | Colin Allison\textsuperscript{3} | Manfred Gröning\textsuperscript{1}

\textsuperscript{1}International Atomic Energy Agency (IAEA), Wagramerstrasse 5, Vienna, 1400, Austria
\textsuperscript{2}Université du Québec à Montréal, Montréal, QC, H2X 3Y7, Canada
\textsuperscript{3}CSIRO, GPO Box 1700, Canberra, ACT 2601, Australia

Correspondence
S. Assonov, Formerly at the International Atomic Energy Agency (IAEA), Wagramerstrasse 5, Vienna 1400, Austria.
Email: assonov_sergey@yahoo.com

Rationale: LSVEC, the second anchor Reference Material (RM) for the VPDB δ^{13}C scale realisation, was introduced in 2006. In 2015, its δ^{13}C value was found to be drifting and, in 2017, its use as an RM for δ^{13}C was officially discontinued by IUPAC. New RMs of low uncertainty are needed. This paper describes the preparation and characterisation of IAEA-610, IAEA-611 and IAEA-612 (calcium carbonate, of chemical origin) which shall serve as a set of RMs aimed at anchoring the VPDB scale at negative δ^{13}C values.

Methods: The preparation and characterisation of IAEA-610, IAEA-611 and IAEA-612 were performed by addressing the contemporary technical requirements for RM production and characterisation (ISO Guide 35:2017). The three RMs were produced in large quantities, and the first batch was sealed into ampoules (0.5 g) to ensure the integrity of the RM during storage; additional batches were sealed for long-term storage. The most accurate method of CO\textsubscript{2} preparation and stable isotope measurements was used, namely carbonate-H\textsubscript{3}PO\textsubscript{4} reaction under well-controlled conditions combined with well-tested stable isotope ratio mass spectrometry.

Results: The assigned values of δ^{13}C and associated uncertainties are based on a large number of analyses (\sim 10 mg aliquots) performed at IAEA and address all the known uncertainty components. For aliquots down to \sim 100 μg, the δ^{13}C uncertainty is increased. The uncertainty components considered are as follows: (i) material homogeneity, (ii) value assignment against IAEA-603, (iii) potential storage effects, (iv) effect of the 17O correction, and (v) mass spectrometer linearity and cross-contamination memory in the ion source.

Conclusions: The new RMs IAEA-610, IAEA-611 and IAEA-612 have been characterised on the VPDB δ^{13}C scale in a mutually consistent way. The use of three RMs will allow a consistent realisation of the VPDB δ^{13}C scale with small uncertainty to be established, and to reach metrological compatibility of measurement results over several decades.

1 \hspace{1cm} INTRODUCTION

LSVEC, the second anchor Reference Material (RM) for the VPDB δ^{13}C scale realisation, was introduced in 2006\textsuperscript{1} to synchronise calibrations at negative δ^{13}C values. When a drift in the LSVEC δ^{13}C value was reported in 2015,\textsuperscript{2} that material was no longer suitable for use as a high-level RM at negative δ^{13}C with a low uncertainty, as required for the VPDB scale realisation (see, e.g., \textsuperscript{3,4}). The
International Atomic Energy Agency (IAEA, Vienna, Austria) meeting on stable isotope RMs held in 2016\(^6\) stressed the urgent need to develop new RMs with very low uncertainty that follow the contemporary requirements for RM production and characterisation principles outlined in ISO Guide 35 (we refer to the 2017 edition\(^5\)). These principles stress the modern uncertainty requirements, and assigning values based on the metrological traceability principle.\(^6\) It was also a requirement to present a clear hierarchical structure of RMs on the VPDB scale, verify the compatibility of existing RMs and, correspondingly, present a self-consistent revision of the VPDB $\delta^{13}C$ scale.\(^3\) The importance of self-consistency in the value assignment of modern RMs has been stressed recently.\(^7,8\)

In 2017, the 19th World Meteorological Organization (WMO)/IAEA Meeting on Greenhouse Gas Measurement Techniques (GGMT-2017) highlighted these needs and requested that “potentially a range of materials may be introduced to help identify drift” and that the new RMs “must be stable, homogenous, sufficiently abundant, and cover the $\delta^{13}C$ range of interests” (cited from GAW Report No. 2429). The GGMT meetings over many years have set the long-term compatibility targets at $\pm 0.010\%$ ($k = 1$) for $\delta^{13}C$ values of atmospheric CO$_2$, ranging from $-9.5$ to $-7.5\%$ and at $\pm 0.020\%$ ($k = 1$) for the $\delta^{13}C$ value of atmospheric methane, ranging from $-51$ to $-46\%$.\(^9-12\) Achieving these targets is not possible without fit-for purpose RMs covering the $\delta^{13}C$ range, with their uncertainty, or at least the uncertainty component due to RM inhomogeneity and long-term stability, being smaller than the targets cited above. No such RMs were available in 2017.

To address the above requirements, in 2018 the IAEA started preparation of three new Ca-carbonate RMs aimed at the VPDB $\delta^{13}C$ scale realisation with high accuracy and small uncertainties\(^\dagger\). The preparation and value characterisation for new RMs are based on the approaches outlined in ISO Guide 35:2017, in particular using Clause 9.3.2 for Certified Reference Materials (CRMs). The IAEA follows the term of CRM specified in VIM 3\(^{12}\) in the following: (i) necessary documentation released according to ISO Guide 31, and (ii) procedures for the production and certification corresponding to ISO Standard 17034 and ISO Guide 35 (the latest editions were applied in this work\(^5,13,14\)). This concept of CRM implies a specific process driven by the producer of the RM. Here we cite from van der Veen et al\(^\dagger\): "Certification is the whole process of obtaining the property values and their uncertainty, which includes homogeneity testing, stability testing, and characterization." Development of this approach at the IAEA and its application for new stable isotope RMs are presented in section 5.

As a result of the project, new RMs IAEA-610, IAEA-611 and IAEA-612 (collectively referred to as IAEA-610/612) have been made available at the IAEA. Their values were assigned in a mutually consistent way, each with uncertainty less than 0.015\% ($k = 1$) in the $\delta^{13}C$ range down to $-36.7\%$. Consistency with RMs on the VPDB scale of 2006\(^1\) was successfully demonstrated using NIST CO$_2$ RMs 8562–8564.

Together with the primary RM IAEA-603, the new RMs IAEA-610/612 can be used for realisation of the VPDB $\delta^{13}C$ scale with small uncertainty including the multi-point realisation. Values of the four RMs were assigned using the same approach, in a self-consistent way. Advantages of the multi-point realisation of the VPDB $\delta^{13}C$ scale using these RMs are discussed in detail in a companion paper.\(^16\)

This manuscript describes the preparation and characterisation of new RMs including the mass spectrometry analyses and related tests and method validation. The paper is targeted at metrologists, RM producers, the greenhouse gas observation community and numerous readers interested in $\delta^{13}C$ determinations with high accuracy. The VPDB scale history and explanation of the concept of the multi-point VPDB scale realisation based on the new RMs and its use as scale anchors are described in a companion paper.\(^16\)

Note regarding USGS44: This project was not established to simply develop a replacement material for LSVEC. USGS (Reston, VA, USA) has started characterising RM USGS44 (undergoing distribution at USGS since 2019\(^\dagger\)) for negative $\delta^{13}C$ values, aimed at the two-point normalisation of $\delta^{13}C$ results produced by TCEA combustion and aiming to replace LSVEC. We did not include this material USGS44 in the RM development process at the IAEA for several reasons. Since ISO Guide 35:2017 considers RM homogeneity and stability assessment as obligatory prerequisites for the value characterisation, and the homogeneity and stability of USGS44 were not assessed at the moment of RM IAEA-610/612 value characterisation, measurements of USGS44 were not included in the RM value characterisation at the IAEA. In addition, the TCEA combustion method was used in the USGS44 study but not in our study. When the final revision of the current manuscript was submitted, the report of the USGS44 study became available online as an accepted article.\(^17\)

Although JFH was involved in both studies, he did not share any information, values and concepts applied between one study and the other.

2 | SELECTION AND PREPARATION OF MATERIALS

2.1 | Selection of materials

The selection of suitable starting materials is critical. The materials are required to have different $\delta^{13}C$ values, be homogeneous and be available in sufficient quantity for a large number of identical units to be produced. For decades, the use of geological carbonates was the preferred option for RMs for several reasons. First, such carbonates are compatible with the carbonate-acid reaction widely used in geosciences since the 1950s (e.g.,\(^18,19\)), and CO$_2$ released in this way can be used for both $\delta^{13}C$ and $\delta^{18}O$ calibrations with high accuracy. Second, carbonate RMs are easier to produce and store in a large

\(^{1}\)Covering the $\delta^{13}C$ range from atmospheric CO$_2$ at $\sim -8\%$ down to $-30\%$ (biosphere processes) was set as the highest priority; further scale extrapolation to more negative $\delta^{13}C$ values (methane-in-air calibration mixtures with $\delta^{13}C$ of $\sim -50\%$) was considered a joint task for laboratories to pursue.

\(^{2}\)Targeting high accuracy in $\delta^{18}O$ values was considered to be beyond the scope of the IAEA project.
quantity than pure CO₂ RMs. A unit of 0.5 g carbonate RM contains from 50 to >1000 aliquots (depending on analytical instrumentation), whereas a single CO₂ ampoule can be used only once. Third, the carbonate-acid reaction gives better accuracy in δ¹³C of CO₂ produced than other methods, e.g., combustion of RMs of other matrices.

RMs prepared in the past and based on geological carbonates include NBS18, NBS19, IAEA-CO-1, IAEA-CO-820 and, recently, IAEA-603. Geological carbonates are known to be more resistant to alterations than chemical carbonates and are fully compatible with the instrumentation and methods used in geosciences. However, finding mineralogically pure geological carbonates of desired δ¹³C is known to be difficult, and testing and preparation steps usually take several years. Geological carbonates are also known to be often inhomogeneous at the grain-to-grain level, which may hinder their use at small aliquot sizes.

In contrast, chemically produced Ca-carbonates are typically much more homogenous at low aliquot size and they cover a large range in δ¹³C. Thus, the use of carefully selected and well-preserved chemically precipitated commercial Ca-carbonates as new RMs is justified.

### 2.2 Feasibility study and material preparation

The feasibility study surveyed the isotopic composition and physical properties of a number of chemically precipitated commercial Ca-carbonates. X-ray diffraction (XRD) was used to confirm the crystalline form of calcite and the absence of aragonite. This is important for two reasons; first, ISO Guide 35:2017 and the value characterisation approach that we have adopted require the new RMs to have the same chemical composition as the primary RM IAEA-603 which is calcite, and, second, calcite and aragonite have different ¹⁸O-isotope fractionation in the reaction carbonate → H₃PO₄ which could lead to problems with ensuring the consistency of measurements.

As a result of the feasibility study, three commercially available Ca-carbonate chemical products were selected (grain size at < 50 μm, purity of >99.5% or higher). The materials were homogenised batchwise in a tumbling machine and the first batch of each RM was then sealed in 2-mL borosilicate glass ampoules (0.5 g). A total of 3000–4000 ampoules of each RM (Table 1) were produced, in a similar way to IAEA-603. Additional batches were stored in hermetically closed 5-L glass bulbs and reserved for RM production in the future.

### 3 METHODS IN USE AT THE IAEA

To address all the strict requirements outlined in section 1, the current work was designed to fulfil the technical requirements for RM characterisation laid out in ISO Guide 35:2017, Clause 9.3.2, adapted to this study as follows:

1. the new RMs are of the same chemical matrix of Ca-carbonate as the primary RM IAEA-603 (see above),
2. the analyte CO₂ gas preparation is based on the well-understood carbonate-H₃PO₄ reaction, the same as used for the primary RM IAEA-603 (see below),
3. a well-understood and well-tested method of CO₂ mass spectrometry (see below) must be used.

The instruments and methods in use at the IAEA are the same as described and used in the IAEA-603 characterisation. They include the carbonate-CO₂ preparation by the carbonate-H₃PO₄ reaction and the well-tested instrumentation as used in the RM IAEA-603 study (see also the supporting information). The RM sampling schemes used in this work for the homogeneity and value characterisation are also very similar to the schemes used in the IAEA-603 project. Following the requirements for calibration laboratories laid out by ISO Standard 17025 and for the characterisation of RMs under ISO Guide 35:2017, we have developed and performed critical specific tests to the mass spectrometry as described below. In addition, a comprehensive method validation (see below) was performed.

The identical treatment principle is shown to be crucial in RM value assignment and, as we did for the IAEA-603 characterisation and value assignment, we followed this principle at all steps: homogenisation, preparation, sampling strategy, carbonate-acid reaction, sample CO₂ transfer and gas introduction, measurements and referencing schemes, raw data treatment, and the treatment of uncertainty components.

### Table 1 Number of ampoules produced for each RM and additional batches reserved

<table>
<thead>
<tr>
<th>RM</th>
<th>Number of glass ampoules sealed as the first batch</th>
<th>Number of batches reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAEA-603</td>
<td>5200</td>
<td>4</td>
</tr>
<tr>
<td>IAEA-610</td>
<td>3000</td>
<td>2</td>
</tr>
<tr>
<td>IAEA-611</td>
<td>3750</td>
<td>2</td>
</tr>
<tr>
<td>IAEA-612</td>
<td>4100</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 2 Number of aliquots of each RM included in the four preparation sequences followed by successful IRMS analyses during the characterisation study

<table>
<thead>
<tr>
<th>Sequence</th>
<th>IAEA-603</th>
<th>IAEA-610</th>
<th>IAEA-611</th>
<th>IAEA-612</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>15</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

* Aliquots of another material were included.
* Six aliquots loaded, one sample lost.
### 3.1 | Mass spectrometry at the IAEA

The CO₂ isotopic composition was measured in dual-inlet (DI) mode on a MAT 253™ isotope ratio mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). CO₂ portions produced from carbonate aliquots were introduced via a custom-made 10-position manifold operated by the MAT 253™ operating ISODAT software. In this way, the 20 CO₂ portions produced in each carbonate-H₃PO₄ preparation sequence were measured in the same way, through software control, in sub-batches of 10 against the same working reference-CO₂ (WR-CO₂) gases filled in the MAT 253™ reference bellow (WR-CO₂ refilled before starting each sequence of 20).

To ensure isotopic compositions of WR-CO₂ as close as possible to those of samples, two different WR-CO₂ gases were used in the homogeneity study on each particular RM (δ¹³C close to −31 ‰ and −9 ‰, respectively). In addition, one more WR-CO₂ gas (δ¹³C ~ −3 ‰ which is closer to IAEA-603) was used for value characterisation measurements on IAEA-610/612 (and also on NIST RMs 8562–8564) against IAEA-603. The reproducibility of our WR-CO₂ refilling system was confirmed previously: during this study, some additional and more critical tests were performed as follows:

- No dependence of δ⁴⁵R–δ⁴⁶R of the WR-CO₂ due to gas consumption in the reference bellow (“bleeding” effect) was observed (see supporting information). This effect is known to result in a change in the δ⁴⁵R–δ⁴⁶R of WR-CO₂ over a single day; under non-optimised conditions, correction in δ¹³C (the latter value calculated from δ⁴⁵R) is found to account for 0.03 ‰ over 10 h (Figure 1 in Brand et al.

- Stability of the WR-CO₂ refills over many days (e.g., our internal reference scheme used in the homogeneity assessment) can be demonstrated by the raw δ⁴⁵R–δ⁴⁶R, as obtained against the WR-CO₂ on IAEA-610/612 in the homogeneity study (time lapse up to 3 weeks).

The results of the homogeneity tests on IAEA-610/612 performed in this study can be used to demonstrate the two above effects. First, data scatter of Standard Deviation (SD) δ¹³C = ±0.006‰ was demonstrated for each daily analytical sequence (no corrections applied), each corresponding to a single WR-CO₂ refill (see below Section 4.1 and Table 3, and Figure S1, supporting information). If a significant bleeding effect took place, observing such small SD values for each analytical sequence would not be possible.

The MAT 253™ standard performance tests and the ion source linearity tests were performed regularly, and they all indicated reliable and high accuracy performance (see below, and more details in the supporting information). The best reproducibility of our MAT 253™ instrument observed at SD(δ¹³C, n = 16) = ±0.006‰ and SD(δ¹⁸O, n = 16) = ±0.015 ‰ was demonstrated on 16 aliquots of the same CO₂ refilled in the sample bellow (without WR-CO₂ refill); these values are much better than MAT 253™ specifications.

One more potential effect is the formation of undesirable¹²C¹⁶O₂⁺ ions contributing to m/z 45. H₂O levels were monitored on m/z 18 in all sample runs, and values were found at the same low values as detected on pure WR-CO₂ which corresponds to ¹⁸O⁻ fragment ions on pure CO₂. In addition, the variable ion source conductance (VIS) was kept open to reduce the cross-contamination (see below); its opening further decreases formation of undesirable¹²C¹⁶O₂⁺ ions.

### 3.2 | Specific mass spectrometry tests

Cross-contamination in the ion source reduces the measured raw δ⁴⁵R–δ⁴⁶R value differences which is crucial in assigning δ¹³C values of RMs correctly. In this work, ion source parameters were applied to reduce the cross-contamination; namely, the VIS was kept fully open, the emission current was reduced and the acceleration voltage was set at 6 kV.

The cross-contamination factor can be quantified as follows:

\[
\eta = (\delta_{\text{true}} - \delta_{\text{meas}}) / (\delta_{\text{true}} + C \delta_{\text{meas}})
\]

In the acceptance protocol template, MAT 253 specifications have been specified at SD(δ¹³C, n = 10) ≤ 0.0125 ‰ and SD(δ¹⁸O, n = 10) ≤ 0.025 ‰, to be demonstrated without sample-CO₂ and WR-CO₂ refills. We also refer to the official MAT 253 brochure which provides typical performance parameters.

---

<table>
<thead>
<tr>
<th>RM</th>
<th>SD(δ¹³C), ‰</th>
<th>SD(δ¹⁸O), ‰</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAEA-610</td>
<td>±0.0060 (n = 80)</td>
<td>± 0.040 (n = 80)</td>
<td>Four analytical sequences, no outliers. Values of the 1st sequence were shift-corrected, see Figure 1. Exclusion of data from the 1st sequence does not change the SD for n = 60.</td>
</tr>
<tr>
<td>IAEA-611</td>
<td>± 0.0059 (n = 66)</td>
<td>± 0.043 (n = 66)</td>
<td>3.5 analytical sequences, no outliers, some gas samples lost.</td>
</tr>
<tr>
<td>IAEA-612</td>
<td>± 0.0063 (n = 73)</td>
<td>± 0.058 (n = 73)</td>
<td>Four analytical sequences, no outliers, some gas samples lost.</td>
</tr>
</tbody>
</table>

Notes:

- For standard error of the mean, we use the abbreviation SE.

---

In the acceptance protocol template, MAT 253 specifications have been specified at SD(δ¹³C, n = 10) ≤ 0.0125 ‰ and SD(δ¹⁸O, n = 10) ≤ 0.025 ‰, to be demonstrated without sample-CO₂ and WR-CO₂ refills. We also refer to the official MAT 253 brochure which provides typical performance parameters.
where $\delta^{45,46}_{\text{R_meas}}$ and $\delta^{45,46}_{\text{R_true}}$ are, respectively, the measured and true values. The respective corrected values were calculated following:

$$
\delta_{\text{true}} = \delta_{\text{R_meas}} / (1 - 2 * \eta * \delta_{\text{R_meas}})
$$

The method of pairwise CO2 measurements developed for the MAT 252 mass spectrometer (Thermo Fisher Scientific) was shown to give the best estimation of the $\eta_{\text{RMS}}$-factor. The method is based on the IAEA-235 and other methods proposed. Note, the MAT 252 and the MAT 253 have the same ion source. The method of “changeover response” proposed in Ghosh et al results in an apparently low, underestimated value of $\eta_{\text{RMS}}$. These tests with isotopically enriched CO2 may potentially give rise to not well recognised memory effects.

In the present study, the cross-contamination was monitored by analysing isotopically different CO2 gases (large $\delta^{45,46}$ difference from values of WR-CO2), with the “idle time” (time between change-over and the start of signal integration) varying from 16 s to typical analysis to 300 s as suggested in the literature. A few tests with “idle time” much longer than 300 s demonstrated that the values obtained at 300 s can safely be considered as “true” values. The values of $\eta_{\text{RMS}}$ and $\eta_{\text{RBM}}$ observed in this work are among the lowest values reported to date which translate to small corrections to uncorrected results.

### 3.3 Design of the RM homogeneity study

The homogeneity study was designed following the approach and the methods applied in the IAEA-603 characterisation. The ISO Guide 35:2017 requirements were considered in terms of selection of ampoules, the number of aliquots to be analysed and reasons for potential variability. Based on the fact that the starting materials of new RMs were well homogenised, and knowing that the analytical uncertainty (preparation and measurements) at the IAEA is close to the $\delta^{13}C$ variability expected as a consequence of potential between-unit heterogeneity, the tests were focused on the between-unit variability.

The homogeneity study was carried out by sampling 30 ampoules from the final products (3000 to 4000 units, Table 1) of each RM. Following ISO Guide 35:2017, this is a statistically representative sampling of the units produced, with the coverage factor close to $k = 2$. Two aliquots of $\sim$10 mg from each selected ampoule were taken for analysis. In addition, two ampoules of each RM were taken as controls, and two aliquots from two controls were included in each analytical sequence. In this way, the in-unit homogeneity was also assessed. Assessing the in-unit heterogeneity by analysing multiple aliquots from numerous units (suggested $n = 3$ for within-unit sampling) was considered not possible due to the extra-large number of analyses required. The isotope ratio mass spectrometry (IRMS) measurements were performed on the same day as the end of the preparation of each batch.

Measurements in the homogeneity study and the RM value characterisation required different approaches. In the homogeneity study, it was important to detect the smallest value variance among numerous aliquots of each RM, to be analysed under fully identical conditions (“repeatability conditions” in VIM3). In contrast to the highest accuracy required for the value characterisation, the homogeneity assessment tolerates some uncorrected biases as these tend to be instrumental in nature and do not affect the homogeneity assessment of the sample.

Since the isotope value difference between the sample CO2 under analysis and the WR-CO2 magnifies the uncertainty (seen as potential value fluctuations) due to the cross-contamination effect (see above and Table S2, supporting information), the value of the WR-CO2 that we use should be as close to the values of the sample CO2 as possible. Two WR-CO2 gases were used in the homogeneity study; one with $\delta^{13}C \sim \pm 9$‰ for IAEA-610, and one with $\delta^{13}C \sim \pm 31$‰ for IAEA-611 and IAEA-612. The two WR-CO2 gases provided an internal reference frame to determine $\delta^{13}C$ variability in the homogeneity tests only. Since these tests were not intended to be used for the RM value characterisation, at this step, the uncertainties due to cross-contamination corrections and other potential non-linearity effects do not need to be included.

### 3.4 Design of the RM value characterisation

The value characterisation following ISO Guide 35:2017, Clause 9.3.2.5 implies that aliquots of new RMs (based on a RM batch produced) shall be measured directly against the RM used for calibration (here, the primary RM IAEA-603). Here, representative sampling is crucial. Since the $\delta^{13}C$ inhomogeneity assessment of $\pm 0.009$‰ ($k = 1$) dominates in the IAEA-603 uncertainty budget, five randomly selected IAEA-603 ampoules were used. At this point in the study, IAEA-610, IAEA-611 and IAEA-612 were already found to be homogenous in $\delta^{13}C$ to within SD $= \pm 0.006$‰ (see below, Section 4.1 and Table 3). Since no deviations were found for the control ampoules of each RM, the two controls of each new RM used in the homogeneity study were also taken for the value characterisation.

We applied the sample loading scheme used in the IAEA-603 characterisation (Figure 3 in Assonov et al.), a few aliquots of each new RM were bracketed between aliquots of IAEA-603 used for calibrations. Four preparation sequences in total were performed (Section 3.4, Table 2); in order to randomise potential between-sequence effects, samples of each RM were included in 2–3 analytical sequences. The MAT 253 measurements were performed on the same day as the end of the preparation of each sample batch. In this way, the raw data of the new RMs were directly linked against the primary RM IAEA-603.

All CO2 portions produced were measured against a WR-CO2 with $\delta^{13}C \sim \pm 3.1$‰ which is close to the isotope composition of
IAEA-603. In contrast to the homogeneity study, where measurement reproducibility is critical and some biases may still be tolerated, it is well recognised that (cited from van der Veen et al.15) “Correctness of analytical results is an obvious prerequisite for characterisation.” Since the $\delta^{13}C$ values of the new RMs have large differences from that of IAEA-603, establishing the cross-contamination correction and its fluctuations (uncertainty) is critical. The cross-contamination was regularly monitored during the value characterisation, and the same correction factor was applied to all the raw data sets; the uncertainty related to this correction was included in the uncertainty budgets. All other possible effects affecting correctness of $\delta^{13}C$ values were quantified (see Table S1, supporting information), and the corresponding uncertainty components were also included. We note that this very approach (analysis of samples of different compositions against the same WR-CO$_2$ on the same day, then correcting all the raw data for the cross-contamination) was applied in the characterisation of NIST CO$_2$ RMs 8562–8564 in 2004.27

3.5 | Data treatment and $\delta^{13}$C calculations

The data treatment was performed consistently in terms of the raw $\delta^{45}$R-$\delta^{46}$R values obtained on the RM samples and propagated to the VPDB-CO$_2$ scale zero (see details in Brand et al.40). In the characterisation study, the values were rescaled through the raw $\delta^{45}$R-$\delta^{46}$R of IAEA-603-CO$_2$ analysed in the same analytical sequences, and then the cross-contamination correction was applied. In the homogeneity study, the $\delta^{45}$R-$\delta^{46}$R values measured on RMs against WR-CO$_2$ were rescaled to the averaged $\delta^{45}$R-$\delta^{46}$R of IAEA-603-CO$_2$ obtained in the characterisation study; at this step relative variations were of importance only, and potential biases may be tolerated.

To ensure the high accuracy of $\delta^{13}$C values, a consistent use of the $^{17}$O correction is crucial.9,40-43 All the calculations were performed based on the numerical set from the latest IUPAC recommendation (Brand et al.40), by solving equations for $^{45}$R-$^{46}$R numerically without approximations. At all steps, assigning and knowing exactly the $\delta^{13}$C-$\delta^{18}$O values of WR-CO$_2$ are not required.

3.6 | Design of the method validation based on NIST CO$_2$ RMs 8562–8564

The methods at the IAEA include the optimised carbonate-H$_3$PO$_4$ reaction, and well-tested mass spectrometry including tests aimed to evaluate the cross-contamination correction, potential non-linearity and other instrumental effects (see supporting information). Numerous critical tests performed demonstrate a robust performance of the MAT 253 instrument. In addition, validation of the methods in use was made by (i) analysing NIST CO$_2$ RMs 8562–8564 along with the new RMs, and (ii) analysing the new RMs IAEA-610/612 at Geotop-UQAM, Canada (see below).

NIST CO$_2$ RMs 8562–8564 were selected for the method validation as they (i) cover the $\delta^{13}$C range of interest, (ii) do not require additional preparations, and (iii) have been shown to demonstrate ampoule-to-ampoule scatter in $\delta^{13}$C to within $\pm$0.007‰, $k = 1.45$. The RMs 8562–8564 value characterisation in 2004 involved seven laboratories with raw data corrected for cross-contamination.27 In 2006, the $\delta^{13}$C values obtained in the 2004 study were re-evaluated, in order to be consistent with the VPDB $\delta^{13}$C scale realisation based on LSVEC.1 The inclusion of the carbonate

![FIGURE 1](image)
RMs analysed in 2004\(^2\) and re-evaluated in 2006 into our study was not possible; only NBS18 and IAEA-CO-8 were available in 2019 (and both have \(\delta^{13}C\) values close to \(-5\) \(\%\), not helping with the value assignment at negative \(\delta^{13}C\)). The IAEA-CO-9 supply was exhausted, and the use of LSVEC has been discontinued due to its \(\delta^{13}C\) value drifts and enlarged uncertainty\(^2\)–\(^4\).

In the present study, NIST CO\(_2\) RMs 8562–8564 were analysed against the same WR-CO\(_2\) as we used in the RM value characterisation and following the same approach, including the cross-contamination correction and consideration of potential uncertainty due to other effects.

4 | RESULTS

4.1 | Results of the RM homogeneity study at the IAEA

Based on the raw \(\delta^{45}R-\delta^{46}R\) data of three RMs obtained in the homogeneity study against the two WR-CO\(_2\) gases in use, we calculated \(\delta^{13}C-\delta^{18}O\) values. This was done by using these two WR-CO\(_2\) gases as the internal reference frame and the link to the third WR-CO\(_2\) gas used in the current work for measurements during the characterisation study. The \(\delta^{13}C-\delta^{18}O\) values calculated at this step (Figures 1, 2 and 3) were used to assess the RM homogeneity and potential trends in ampoule filling process only, and not used in the final value assignment.

All three RMs demonstrated inhomogeneity at \(SD(\delta^{13}C) = \pm 0.006 \%\) (Table 3), as assessed on aliquots sampled from a representative selection of ampoules. Some sub-sets of IAEA-611 and IAEA-612 samples were lost, resulting in lower numbers (\(n = 66\) and 73) than the \(n = 80\) on IAEA-610. The number of successfully analysed aliquots for each RM (Table 3, \(n = 80\), \(n = 66\) and \(n = 73\), respectively) is statistically significant to evaluate the RM inhomogeneity.

The only correction was applied to \(\delta^{13}C\) values of IAEA-610 obtained in the first analytical sequence (\(n = 20\)); these values were corrected for a small shift of 0.016 \(\%\) in the \(\delta^{13}C\) value. Pooling data from other sequences did not require any shift-correction. Since the 2\(^{nd}\) and following sequences were separated by 18 calendar days.

![FIGURE 2](image2) \(\delta^{13}C-\delta^{18}O\) values of IAEA-611 obtained in the homogeneity study. Values obtained on \(~10\) mg aliquots (\(n = 66\)). The outlier in \(\delta^{13}C\) was not removed; its exclusion from the data pool insignificantly reduces SD from 0.0059\(\%\) to 0.0052\(\%\).

![FIGURE 3](image3) \(\delta^{13}C-\delta^{18}O\) values of IAEA-612 obtained in the homogeneity study. Values obtained on \(~10\) mg aliquots (\(n = 73\)).
from the 1st sequence (Figure 1) and the same values SD(δ13C) = ±0.006‰ were found for each of the four IAEA-610 sequences, this shift-correction is justified. Alternatively, by discarding data from the 1st sequence and pooling data from the 2nd, 3rd and 4th sequences (Figure 1), one obtains the same SD(n = 60) = ±0.006 ‰ as obtained with the shift-correction. For IAEA-611 and IAEA-612, no shift-corrections were needed; the data were pooled as they were obtained against WR-CO2.

For δ18O values, the SD(δ18O) variances observed in the homogeneity study (Table 3) were found to be somewhat larger than the best performance of the carbonate-H₃PO₄ preparation as observed in the IAEA-603 homogeneity study (SD(δ18O, n = 143) = ±0.036‰). Similar to the IAEA-603 study, the δ18O variances found in the present work appeared to be limited by the preparation rather than the performance of the MAT 253 instrument (SD(δ18O, n = 16) = ±0.015‰). This is probably due to less control on the reaction temperature than in the IAEA-603 study, as powdered carbonates violently react with H₃PO₄ whereas crystalline ones do not. This may affect the actual acid temperature and thus the δ18O value of the CO2 produced. Additional effects due to more moisture being adsorbed on powder RMs than on crystalline IAEA-603 cannot be excluded. The δ18O scatter does not affect the intended use of RMs for the high accuracy δ13C calibrations.

The “true” RM inhomogeneity for each RM shall be estimated by subtracting the method performance and analytical uncertainty from the data variance observed in the homogeneity study. For δ13C, the variances observed on all three RMs (Table 3) were found to be exactly the same as the best analytical performance of our MAT 253, as SD(δ13C, n = 16) = ±0.006 ‰ observed on refilling pure CO2 in the sample bellow, without using carbonate-acid preparations (tests described earlier). Since subtracting the preparation and analytical performance from the δ13C variances found in the homogeneity study was not possible, the actual inhomogeneity in δ13C may be better than the SD(δ13C) = ±0.006‰ found for all three RMs. We also note that the SD(δ13C) = ±0.006‰ observed on each of three RMs is smaller than the IAEA-603 inhomogeneity of SD(δ13C) = ±0.009‰.

### 4.2 Results of RM value characterisation against IAEA-603

The characterisation study was performed over four analytical sequences (Section 3.4, Table 2) and included measurements on CO2 with contrasting isotope compositions. Since the new RMs are aimed at the highest accuracy in δ13C, and potential drifts in raw δ45R values of the WR-CO2 over several days of the characterisation study cannot be fully excluded (for example, a small drift of WR-CO2 was observed in the IAEA-610 homogeneity tests, Figure 1), two scenarios linking the raw δ45R data obtained on the new RMs over four sequences were tested.

In the first scenario, the raw δ45R values of IAEA-610, IAEA-611 and IAEA-612 were linked to the averages of the raw IAEA-603 δ45R values obtained on each analytical sequence, and then all values were corrected for the cross-contamination. The scaled δ45R values obtained this way were pooled for each RM. By pooling data obtained in a single day, the first scenario was not sensitive to potential day-to-day drift of WR-CO2.

In the second scenario, the raw IAEA-603 δ45R values obtained against the WR-CO2 in four analytical sequences (n = 30) were used to estimate a linear trend against time. The raw δ45R data of IAEA-610, IAEA-611 and IAEA-612 were then scaled to the IAEA-603 trend (according to the exact timing of each run), and the re-scaled values were then corrected for the cross-contamination. By applying a small linear drift of WR-CO2 over four sequences in the second scenario, we were able to pool all n = 30 aliquots of IAEA-603.

The δ45R averages for IAEA-610, IAEA-611 and IAEA-612 (rescaled to IAEA-603) evaluated using each of the two scenarios were found to agree well within their k = 1 uncertainties (Table 4). Thereafter, and for simplicity reasons, δ45R values after the 1st scenario were taken for the final value assignment.

Since the raw data scatter in δ45R was found to be much larger than that in δ45R and no trend was observed, the first scenario was applied for δ45R data pooling. The δ45R scatter on each RM was found to be similar to that in the homogeneity study (see Figures 1, 2 and 3).

To eliminate an apparent correlation between δ45R and δ45R deviations due to 17O contribution to m/z 45, the variances of δ13C and δ18O in the characterisation study were estimated based on the δ18O values calculated for all aliquots under analysis (see below, Table 5).

### 4.3 Estimation of RM storage effects and potential alterations

Isotopic exchange with air-moisture and air-CO2 is known to be a major reason for changes in the δ13C and δ18O values of carbonate RMs over time. Following the approach developed for IAEA-603, the three RMs were hermetically sealed in 2-mL glass ampoules, completely preventing exposure to air and related alterations.

<table>
<thead>
<tr>
<th>RM</th>
<th>Raw δ45R values (%o, rescaled against IAEA-603-CO2 averages obtained for each analytical daily sequence)</th>
<th>Raw δ46R values (%o, rescaled against IAEA-603-CO2 trend estimated for IAEA-603 data of four analytical days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAEA-610</td>
<td>−11.358 ± 0.007</td>
<td>−11.357 ± 0.008</td>
</tr>
<tr>
<td>IAEA-611</td>
<td>−31.065 ± 0.012</td>
<td>−31.067 ± 0.012</td>
</tr>
<tr>
<td>IAEA-612</td>
<td>−36.864 ± 0.014</td>
<td>−36.868 ± 0.015</td>
</tr>
</tbody>
</table>
**TABLE 5** Value assignment to IAEA-610, IAEA-611 and IAEA-612 and the uncertainty components (k = 1) associated with the values assigned. (Since new RMs target at high accuracy in $\delta^{13}$C, the values for $\delta^{13}$C are given with four digits after the decimal point)

<table>
<thead>
<tr>
<th>Uncertainty components</th>
<th>$\delta^{13}$C, ‰ VPDB, uncertainties at k = 1</th>
<th>$\delta^{18}$O, ‰ VPDB, uncertainties at k = 1</th>
<th>$\delta^{13}$C, ‰ VPDB, uncertainties at k = 1</th>
<th>$\delta^{18}$O, ‰ VPDB, uncertainties at k = 1</th>
<th>$\delta^{13}$C, ‰ VPDB, uncertainties at k = 1</th>
<th>$\delta^{18}$O, ‰ VPDB, uncertainties at k = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAEA-610</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homogeneity assessment.</td>
<td>0.0060 n = 80</td>
<td>0.0396 n = 80</td>
<td>0.0059 n = 66</td>
<td>0.0427 n = 66</td>
<td>0.0063 n = 73</td>
<td>0.0586 n = 73</td>
</tr>
<tr>
<td>Characterisation against IAEA-603.</td>
<td>0.0020 (IAEA-610, n = 15)</td>
<td>0.0089 (IAEA-610, n = 15)</td>
<td>0.0036 (IAEA-611, n = 14)</td>
<td>0.0058 (IAEA-611, n = 14)</td>
<td>0.0042 (IAEA-612, n = 14)</td>
<td>0.0094 (IAEA-612, n = 14)</td>
</tr>
<tr>
<td>IAEA-603 certificate uncertainties taken excluding the component related to IAEA-603 inhomogeneity (1)</td>
<td>0.0022 (IAEA-603, n = 23)</td>
<td>0.0059 (IAEA-603, n = 23)</td>
<td>0.0033 (IAEA-603, n = 22)</td>
<td>0.0074 (IAEA-603, n = 22)</td>
<td>0.0051 (IAEA-603, n = 14)</td>
<td>0.0101 (IAEA-603, n = 14)</td>
</tr>
<tr>
<td>IAEA-603 homogeneity</td>
<td>no outliers</td>
<td>no outliers</td>
<td>no outliers</td>
<td>no outliers</td>
<td>no outliers</td>
<td>no outliers</td>
</tr>
<tr>
<td>Correction due to cross-contamination</td>
<td>0.0026</td>
<td>0.0085</td>
<td>0.0070</td>
<td>0.0001</td>
<td>0.0082</td>
<td>0.0051</td>
</tr>
<tr>
<td>Tailing of m/z 44 to m/z 45</td>
<td>0.0010</td>
<td>n/a</td>
<td>0.0028</td>
<td>n/a</td>
<td>0.0034</td>
<td>n/a</td>
</tr>
<tr>
<td>Amplifier’s linearity</td>
<td>0.0021</td>
<td>n/a</td>
<td>0.0021</td>
<td>n/a</td>
<td>0.0021</td>
<td>n/a</td>
</tr>
<tr>
<td>Stability assessment (2)</td>
<td>0.0048</td>
<td>n/a</td>
<td>0.0002</td>
<td>n/a</td>
<td>0.0015</td>
<td>n/a</td>
</tr>
<tr>
<td>Maximal effect due to $^{17}$O correction</td>
<td>0.0058</td>
<td>0.0058</td>
<td>0.0058</td>
<td>0.0058</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigned values and uncertainty (3)</td>
<td>$-9.1092 \pm 0.0116$</td>
<td>$-18.8344 \pm 0.0442$</td>
<td>$-30.7948 \pm 0.0132$</td>
<td>$-42.2432 \pm 0.0459$</td>
<td>$-36.7223 \pm 0.0148$</td>
<td>$-12.0792 \pm 0.0620$</td>
</tr>
</tbody>
</table>

1. The IAEA-603 homogeneity is addressed by sampling a representative number of aliquots (n = 30) taken from five ampoules in this study.
2. Assuming that CO₂ remaining in ampoules has a $\delta^{13}$C = −30 ‰.
3. Combined as independent uncertainty components following GUM. The values and combined uncertainties were calculated from the raw isotopic data obtained by the MAT 253 instrument by using off-line data treatment that includes the necessary $^{17}$O isotopic correction which was applied without approximations.
Similar to the IAEA-603 study, the CO₂ amounts in ampoules which were not flushed with argon and then sealed of were measured, and a maximum value of 2.6E-2 cc STP was observed (lower amounts are expected to be in ampoules flushed with argon, as was done for real RMs). The estimations for potential storage effects, which are maximum values, are based on the assumption that the CO₂ came from the flame with a δ¹³C of -30‰ and was potentially fully trapped by the carbonate. These maximum estimations are very conservative estimates and were nonetheless introduced in the uncertainty budget (see below, Table 5).

Because these RMs were not designed for high accuracy δ¹⁸O measurements, and because powder carbonate RMs may absorb local moisture after ampoule opening, no estimations for potential change in δ¹⁸O values were made.

### 4.4 Uncertainty components associated with the δ¹³C and δ¹⁸O value assignment

The uncertainty budget for δ¹³C and δ¹⁸O values assigned to the new RMs included the following individual components (Table 5):

- Uncertainty obtained in the homogeneity study on ~10 mg aliquots. This component was taken as the SD. As the method and analytical uncertainties were not separated (not subtracted) from the δ¹³C inhomogeneity assessment, the actual inhomogeneity in δ¹³C may be better.
- Uncertainty due to RM characterisation against IAEA-603. Average values obtained on each RM were taken with the corresponding Standard Error (SE) and then combined with the SE obtained on IAEA-603 analytical data for each daily analytical sequence.
- Uncertainty of the IAEA-603 certificate values, taken without the homogeneity term. This term was addressed by analysing 30 aliquots of IAEA-603 taken from five ampoules.
- Uncertainty related to raw data correction for the cross-contamination. This was applied to the δ⁴⁵R value in δ⁴⁶R values scaled to IAEA-603 (gross δ⁴⁵R and δ⁴⁶R values) and then included in the δ¹³C and δ¹⁸O uncertainty budget.
- The uncertainty due to amplifier linearity and m/z 44 peak tailing to m/z 45. Corrections for the tailing effect** are known to have been considered and applied for high accuracy determinations on older instruments (e.g. Meijer & Li).
- Uncertainty of the ¹⁷O correction estimated from ¹⁷O-excess deviations (from the reference relationship with λ = 0.528) measured in the new carbonate RMs (E. Barkan, personal communication, 2020).

**The peak tailing is typically a consequence of the energy spread of the ions leaving the ion source. The impact is most noticeable for m/z 44 contribution to m/z 45 due to its ~100-times greater abundance. It is particularly noticeable in some older mass spectrometers and is seldom considered, let alone quantified. The double-focusing mass spectrometer that was used for decades at USGS has an electric sector that reduces the energy spread and minimises the peak tailing.

### TABLE 6 Values obtained at Geotop-UQAM on aliquot size of ~100 μg by acid reaction at 90 °C. Values obtained at the IAEA on ~10 mg aliquots at 25°C are listed for comparison.

<table>
<thead>
<tr>
<th>RM</th>
<th>δ¹³C, ‰ VPDB, obtained at Geotop</th>
<th>δ¹³C, ‰ VPDB, obtained at IAEA (Table 5)</th>
<th>δ¹⁸O, ‰ VPDB, obtained at Geotop</th>
<th>δ¹⁸O, ‰ VPDB, obtained at IAEA (Table 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAEA-610</td>
<td>± 0.03</td>
<td>± 0.03</td>
<td>± 0.04</td>
<td>± 0.04</td>
</tr>
<tr>
<td>IAEA-611</td>
<td>± 0.03</td>
<td>± 0.03</td>
<td>± 0.05</td>
<td>± 0.05</td>
</tr>
<tr>
<td>IAEA-612</td>
<td>± 0.02</td>
<td>± 0.02</td>
<td>± 0.03</td>
<td>± 0.03</td>
</tr>
</tbody>
</table>

Note: These RMs are not intended for high accuracy in δ¹⁸O. The δ¹⁸O values for IAEA-611 and IAEA-612 overlap within the k = 1 reported uncertainty.
• The uncertainty due to long-term RM stability (storage effect), taken as maximal shifts in δ\(^{13}\)C, if CO\(_2\) remaining in the ampoules were completely adsorbed by carbonate materials (Section 4.3).

Some uncertainty components were estimated in δ\(^{13}\)C-δ\(^{18}\)O variances (e.g., inhomogeneity assessment), whereas others were estimated based on raw δ\(^{45}\)R-δ\(^{46}\)R values (e.g., correction for cross-contamination). In the latter case, translation of δ\(^{45}\)R-δ\(^{46}\)R uncertainty to the δ\(^{13}\)C uncertainty followed the linear appreciation formula for δ\(^{13}\)C calculations (Equation 19 in Brand et al\(^{40}\)). By using this formula, we also estimated the effect of uncertainty in the \(^{17}\)R/\(^{18}\)R numerical ratio\(^{40}\) and the δ\(^{46}\)R analytical uncertainty on the calculated δ\(^{13}\)C value. The maximal effect did not exceed 0.001 ‰ and was included along with the effect of 17O-excess.

Potential biases in δ\(^{13}\)C due to 17O-excess require explanations. The 17O-excess values measured on the new carbonate RMs (data by E. Barkan, personal communication, 2020) were found to be within 0.08 ‰ of the 17O-excess reported on IAEA-603; the latter is reported to deviate by 0.012 ± 0.008 ‰ (k = 1) from the value of NBS19.\(^{23,49}\) The 17O-excess values were considered as a potential uncertainty of the 17O correction of the new RMs and were included as an uncertainty component (in its maximal estimate) for δ\(^{13}\)C values calculated in this work (Table 5).

All the uncertainty components were considered as independent and thus combined in quadrature in accordance with GUM.\(^{50}\) Uncertainty propagation (e.g., in corrections after Equation (2)) and combining uncertainties were verified by using the NIST web-based software application for combining the measurement uncertainty components by the Monte Carlo method.\(^{51}\)

### 4.5 Analyses on NIST CO\(_2\) RMs 8562–8564 at the IAEA (method validation)

NIST CO\(_2\) RMs 8562–8564 were analysed in the same way as CO\(_2\) gases released from the new carbonate RMs, namely by applying the same gas introduction, the same raw data treatment and calibration against IAEA-603. CO\(_2\) from each RM glass ampoule/tube was split into 2–3 portions and analysed separately. A good agreement between results obtained on two-three portions from each ampoule was observed, thus confirming no fractionation effects in CO\(_2\) transfer from ampoules; the averages (over portions from each ampoule) were then pooled for data analysis (the first row in Table 7).

### 4.6 Confirmation measurements at Geotop

Samples from the ampouled RMs were analysed at Geotop-UQAM (Université du Québec à Montréal), Montreal, Canada, by using an Isoprime isotope ratio mass spectrometer (Micromass; now Elemental UK Ltd, Cheadle, UK) equipped with an automated carbonate preparation system (MultiCarb; now Elemental UK Ltd). The

### Table 7 Values obtained on NIST CO\(_2\) RMs 8562–8564 in this study and published values,\(^{1,27,38,54–56}\) published values are based on the same 17O correction as used in this study. Combined uncertainties (k = 2) are given for consistency with NIST certificates; numbers of digits after the decimal point correspond to the original publications

<table>
<thead>
<tr>
<th>Value, publication</th>
<th>RM 8562</th>
<th>RM 8563</th>
<th>RM 8564</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ(^{13})C, values in ‰ VPDB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>δ(^{13})C, obtained in this study(^{1})</td>
<td>−3.736 ± 0.034 ‰ (n = 9**)</td>
<td>−41.556 ± 0.038 ‰ (n = 7**)</td>
<td>−10.466 ± 0.032 ‰ (n = 3**)</td>
</tr>
<tr>
<td>Homogeneity in δ(^{13})C, at SD level (this study)</td>
<td>± 0.008 (n = 9**)</td>
<td>± 0.007 (n = 7**)</td>
<td>± 0.006 (n = 3**)</td>
</tr>
<tr>
<td>δ(^{13})C, reported by NIST in 2004 (from Table 10D, uncertainties are based on Table 8(^{27}))</td>
<td>−3.72 ± 0.06</td>
<td>−41.57 ± 0.08</td>
<td>−10.45 ± 0.06</td>
</tr>
<tr>
<td>δ(^{13})C, re-normalised to LSVEC in 2006(^{1}) (values from Table S8 in supplementary material to Coplen et al(^{1}))</td>
<td>−3.716 ± 0.076</td>
<td>−41.585 ± 0.116</td>
<td>−10.452 ± 0.079</td>
</tr>
<tr>
<td>δ(^{13})C, obtained at IRMM in 2009(^{18}) (‘‘)</td>
<td>−3.74 ± 0.03 (n = 3**)</td>
<td>−41.55 ± 0.03 (n = 2**)</td>
<td>−10.46 ± 0.03 (n = 4**)</td>
</tr>
<tr>
<td>δ(^{13})C, NIST certificate updates in 2018(^{52–55}) (‘‘)</td>
<td>−3.72 ± 0.07</td>
<td>−41.59 ± 0.09</td>
<td>−10.45 ± 0.07</td>
</tr>
<tr>
<td>δ(^{18})O, values in ‰ VPDB-CO(_2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>δ(^{18})O, obtained in this study(^{1})</td>
<td>−18.449 ± 0.086 ‰ (n = 9**)</td>
<td>−33.596 ± 0.126 ‰ (n = 7**)</td>
<td>−9.916 ± 0.080 ‰ (n = 3**)</td>
</tr>
<tr>
<td>δ(^{18})O, obtained at IRMM(^{38}) (‘‘)</td>
<td>−18.62 ± 0.03 (n = 3’’)</td>
<td>−33.75 ± 0.05 (n = 2’’)</td>
<td>−10.09 ± 0.04 (n = 4’’)</td>
</tr>
<tr>
<td>δ(^{18})O, reported by NIST in 2004 (Table 8 in(^{27})) and in NIST certificate update in 2018(^{52–55}) (‘‘)</td>
<td>−18.49 ± 0.44</td>
<td>−33.52 ± 0.48</td>
<td>−10.09 ± 0.40</td>
</tr>
</tbody>
</table>

\(^{1}\)The combined uncertainties (k = 2) include analytical uncertainties observed on respective NIST RMs and on IAEA-603-CO\(_2\) in the current work, the uncertainty due to raw data correction for the cross-contamination effect and the IAEA-603 assigned (certificate) uncertainty taken without the inhomogeneity term.

\(^{2}\)Number of ampoules of each RM analysed.

\(^{3}\)Includes the calibration uncertainty at IRMM.

\(^{4}\)NIST SP260-149 (in Table 8\(^{27}\)) provided values on the VPDB-CO\(_2\) scale. Certificate δ\(^{18}\)O values given on the VPDB scale\(^{54–56}\), if rescaled to the VPDB-CO\(_2\) scale, with a factor of 1.01025, result in the same values as NIST SP260-149.
carbonate-H3PO4 reaction was performed at 90°C, at an aliquot size of ~100 μg. IAEA-603 was analysed under the same conditions in each daily sequence; as additional control, aliquots of laboratory standards have been included.

The raw data treatment in terms of δ^{13}C-δ^{18}O values including the 17O correction recommended by IUPAC was applied; namely, all the raw δ^{46}R data obtained on the new RMs against WR-CO2 were rescaled to the raw data average of IAEA-603 (as measured in the same daily sequence) and the δ^{13}C-δ^{18}O values were then calculated. The correction for the cross-contamination was applied to the δ^{13}C-δ^{18}O values calculated (Table 6).

To evaluate the cross-contamination effect, the same CO2 as used for tests at the IAEA (δ^{13}C ≈ −68‰) was analysed at Geotop-UQAM, by running the Isoprime under the same conditions. A δ^{46}R value of (2.29 ± 0.24)E-3 was estimated; the value of δ^{46}R was assumed to be the same. Being estimated in this way, the δ^{46}R factor accounts for both the cross-contamination effect and a potential contribution of m/z 44 tailing to the m/z 45 signal. The latter cannot be excluded on a low radius mass spectrometer such as the Isoprime. (Similarly, a correction for tailing effects was applied for high accuracy measurements.)

The values obtained at Geotop-UQAM on aliquot sizes of ~100 μg are given in Table 6. The results obtained on IAEA-610 are used to estimate inhomogeneity only, as these values demonstrated a correlation between raw δ^{46}R-δ^{18}O values and calculated δ^{13}C-δ^{18}O values, not given in Table 6, which implies incomplete reaction and uncontrolled bias(es) due to isotope fractionation. As a potential reason, an electrostatic charge effect was observed during the days of the IAEA-610 analytical sequences which resulted in the material spreading in the reaction vials. This probably has resulted in incomplete reaction, as was confirmed by the CO2 pressure in the reaction vessels being lower than expected. However, a compact cluster of δ^{13}C-δ^{18}O was observed and allowed us to estimate the inhomogeneity at ~100 μg.

The results obtained at Geotop-UQAM (Table 6) demonstrate two aspects:

1. They confirm a good homogeneity of all three new RMs at ~100 μg aliquots, for both δ^{13}C and δ^{18}O values. The homogeneity of the IAEA-610 material is also confirmed; even with incomplete reaction, one would not observe such small SD values (Table 6, the first row for IAEA-610) on inhomogeneous material.

2. The values agree well for IAEA-611 and IAEA-612 as obtained on ~100 μg aliquots (Table 6), with the values obtained at the IAEA on ~10 mg aliquots (the last line in Table 5). The discrepancies are within the stated uncertainties (k = 1) except for IAEA-610 for which the values are biased (see above, values not included in Table 6).

We note that the Geotop-UQAM data were based on a representative number of IAEA-603 aliquots taken for calibration. However, only a very limited number of the cross-contamination tests was performed. Thus, the correction for this effect substantially contributes to the final uncertainty.

5 | DISCUSSION

5.1 | Values assignment scheme

The RM value assignment followed the approach outlined in ISO Guide 35:2017, Clause 9.3.2, Characterisation by value transfer from a reference material to a closely matched candidate reference material using a single measurement procedure performed by one laboratory. In this case, we cite from ISO Guide 35:2017: “values are assigned ... by directly comparing results on the candidate CRM with those on an already characterised and closely matched CRM (the “primary CRM”). The RM characterisation scheme utilising values obtained at a single expert laboratory and supported by other means is accepted at NIST; a very similar scheme was applied in the IAEA-603 characterisation.”

We stress that the IAEA uses the term for CRMs as described in VIM 3, in particular in terms of (i) supporting documentation released according to ISO Guide 31, and (ii) procedures for the production and certification corresponding to ISO Standard 17034 and ISO Guide 35. In this respect we stress (cited from van der Veen et al., who contributed to developing concepts used in ISO Guide 35): “Certification is the whole process of obtaining the property values and their uncertainty, which includes homogeneity testing, stability testing, and characterisation”. This RM characterisation shall be performed on the RM batch produced.

The project on RMs IAEA-610/612 fulfils all the technical requirements of ISO Guide 35 and requirements for CRMs listed above. The new RMs are of the same Ca-carbonate matrix as the primary RM IAEA-603, the analysis was based on the well-tested CO2 preparation by the carbonate-H3PO4 reaction and by using the instrumentation used for the IAEA-603 characterisation. The homogeneity testing and value assignment were based on representative sampling of IAEA-610/612 and representative sampling of IAEA-603; in the present study, 262 aliquots of the new RMs and 30 aliquots of IAEA-603 were analysed.

Table 5 lists the uncertainty components included and provides the combined standard uncertainties (k = 1). The latter includes components due to the homogeneity assessment, the stability assessment, the value characterisation, the method-related corrections (contributions by relevant corrections and MAT 253 tests) and the IAEA-603 certificate uncertainty. This characterisation of a RM batch implies linking average values obtained on the new RMs and the one used for calibrations (here on IAEA-610/612 and IAEA-603) with corresponding Standard Errors (often named as Standard Deviation of the Mean, the uncertainty of averaged values).

Strictly speaking, these are of the same chemical matrix but not the same crystalline grain-size. This fact may have some significance, e.g., one should prevent static electricity and adapt the acid reaction time.
Since the uncertainty components are assumed as uncorrelated and independent, these can be combined in a squared form.\textsuperscript{15,39,50}

The uncertainty combined in this way provides an estimate of how well a single (or a few) random aliquot from a single (randomly selected) ampoule represents the value assigned to the RM in question. This concept developed by van der Veen et al.\textsuperscript{15} and later used in ISO Guide 35:2017,\textsuperscript{5} implies that the end-users will (as a rule) use only one (or only a few) aliquot at a time.

The same concept was used in IAEA-603 uncertainty assignment,\textsuperscript{23} and its certificate uncertainty is intended for use on a single aliquot (as a rule – a few) analysed at the user laboratory. Since the current study includes 30 aliquots of IAEA-603 sampled from five randomly selected ampoules, thereby providing an assessment of its homogeneity, the uncertainty budget (Table 5) includes the IAEA-603 certificate uncertainty, excluding the component related to IAEA-603 inhomogeneity assessment\textsuperscript{23}; the IAEA-603 inhomogeneity is directly addressed in the current study by sampling and analysing a representative number of aliquots (n = 30) in this study.

We stress that the combined uncertainties of RMs IAEA-610/612 (Table 5) provide robust conservative estimations since the components related to inhomogeneity assessment in $\delta^{13}\text{C}$, potential $\delta^{13}\text{C}$ shifts during storage, the effect of potential MAT 253 non-linearity and peak-tailing, and also the component related to the $^{17}\text{O}$ correction, were taken as the maximum estimates. The inhomogeneity components are at maximum, since we were not able to subtract the analytical performance as typically required (see above and van der Veen et al\textsuperscript{39}).

The lower panel in Figure 4 compares the uncertainty components for IAEA-610/612 (grouped in four groups) with those of the new primary RM IAEA-603.\textsuperscript{23} As we expected based on the nature of the material (section 2.1), RMs IAEA-610/612 appear to be more homogenous than RM IAEA-603 which is a carbonate of

![Figure 4](image_url)

**FIGURE 4** Upper panel: Uncertainty in $\delta^{13}\text{C}$ for new RMs IAEA-610/612 (black) and old RMs (grey) available at the IAEA in 2020. The latter include NIST CO$_2$ RMs 8562–8564 (archived) and NBS18; all these were re-characterised in 2006. The uncertainty of old RMs does not include the homogeneity and stability assessment.

Lower panel: Uncertainty components in $\delta^{13}\text{C}$ for new RMs IAEA-610/612 grouped in four groups (see text and Table 5). The uncertainty of these RMs (presented in the upper panel) is the quadrature combination of the uncertainties presented in this panel.
geological origin and contains 1-2 grains (out of 100) of slightly different isotope composition.

Uncertainties associated with the values of four new RMs at the IAEA (IAEA-603 and IAEA-610/612), are found to be substantially smaller than those of the carbonate and CO₂ RMs of 2006 (we refer to the Excel spreadsheet in the supplementary material to Coplen et al). In particular, the uncertainties of RMs in 2006 included the value assignment only, and did not include the assessments of inhomogeneity, potential storage effects, mass spectrometer effects and instrumental corrections.

5.2 Validation of methods at the IAEA

Three levels of method and value validation are included:

- Critical mass spectrometer tests including the cross-contamination, detector linearity and peak tailing. These tests already imply a reliable performance of our MAT 253 instrument.
- Independent measurements of the new RMs performed at Geotop-UQAM (see above). These tests demonstrate agreement of values, despite acid preparations being done there at 90°C on small aliquots.
- Results obtained at the IAEA on NIST RMs 8562-8564. These tests provide several confirmations and will be discussed in more detail below.

The values obtained on NIST CO₂ RMs 8562-8564 are found to agree with the δ¹³C values assigned by NIST within specified uncertainties (Table 7), thus verifying the consistency of the methods in use in terms of CO₂ mass spectrometry, cross-contamination correction and CO₂ extraction from the primary carbonate RM and positioning on the VPDB scale. Since δ¹³C consistency is of paramount importance, Table 7 compares our results with the published values which are based on the use of the same ¹⁷O correction of 2003, or its equivalent. In our calculations, the analytical uncertainties (analyses in the current work) on RMs 8562-8564 and IAEA-603-CO₂ were conservatively taken as SD (not SE); other uncertainty components were taken in the same way as discussed for new RMs (Table 5) excluding the uncertainty due to the ¹⁷O correction. The latter was not included in the RMs 8562-8564 value assignment in 2004.

As values of RMs 8562-8564 have been revised several times, and because the δ¹⁸O uncertainties that we obtain are smaller than those reported earlier, some explanations are needed. First, we note that the values published by NIST in 2004 were based on the raw data corrected for the cross-contamination. Second, in 2006, Coplen et al applied the LSVEC-based normalisation to the values published by NIST in 2004 with the aim of correcting all RM values for the cross-contamination. This LSVEC-based normalisation effectively duplicated the cross-contamination correction applied to the raw data by NIST in 2004. Other than being unnecessary, this might have reduced the self-consistency of the values of RMs 8562-8564, at least due to not including the LSVEC analytical uncertainty (here we refer to the supplementary data file in Coplen et al). Third, the update of NIST certificates in 2018 is based on values obtained in 2004 and re-normalised in Coplen et al, without reporting any other measurements after 2004 (we refer to the reference lists in 54-56). A few measurements performed at NIST in 2018 were not included in the NIST certificates’ update in 2018 (we refer to the reference list in the latest NIST certificates 54-56). To summarise, we consider the values published in 2004 (Table 10D in Verkouteren and Klinedinst) as being the best estimates in terms of self-consistency compared with re-evaluations in 2006 and 2018.

Another self-consistent set of values for RMs 8562-8564 was obtained at IRMM in 2009 using a well-tested MAT 252 mass spectrometer. IRMM reported very low cross-contamination factors of ¹⁸O = 1.9E-4 and ¹³C = 4.3E-4 which are very close to the values of ¹⁸O = (2.2 ± 1.1)E-4 and ¹³C = (5.8 ± 2.6)E-4 found in this work (see Table S2, supporting information). The IRMM values for RMs 8562-8564 agree to within 0.010‰ with those that we obtained for all three RMs (Table 7).

In 2005, Ghosh et al reported δ¹³C(n = 2) = −41.640 ± 0.013‰ (k = 1) for RM8563, obtained in DI mode on pure CO₂ (2 ampoules of RM 8563; value in Table 3 in Ghosh et al). The uncertainty from Table 7.17. Because RMs 8562 and 8564 were not analysed, it is dangerous to compare the scale expansion/stretching and also exclude potential biasing effects of the value of RM8563 reported in Ghosh et al.

We stress that the difference between δ¹³C values obtained on RMs 8562 and RMs 8564 in each study (Table 7) provides an internal quality control. This difference is sensitive to the cross-contamination correction and absence of non-linear effects, but is insensitive to α(n) potential bias in exact δ¹³C positioning on the VPDB δ¹³C scale. In this respect, our data agree within k = 1 uncertainties with the NIST-2004 values and IRMM values; the agreement with the LSVEC-normalised values (Coplen et al and NIST certificates in 2018) is still within k = 1 uncertainties, despite potentially reduced self-consistency of re-normalised values in 2006.

The δ¹⁸O values obtained in this study (Table 7) agree with the NIST reported values within the (relatively large) uncertainty assigned by NIST. It is known that local air moisture can stay adsorbed on the inner walls of the ampoule-cracker and outside the ampoules after their exposure to the atmosphere, thus modifying the δ¹⁸O value of CO₂. This may explain some δ¹⁸O discrepancies which, being larger than our uncertainties (Table 7), are still within the uncertainties assigned by NIST. Since the δ¹⁸O values of new RMs are for information only (see below), addressing all potential effects in δ¹⁸O was beyond the scope of this project. The effects of δ¹⁸O on the high accuracy δ¹³C determinations can be excluded, as (i) water levels were monitored in all mass spectrometer runs and no elevated values were detected; (ii) the ¹⁷O correction was based on the actual δ¹⁸O value measured on CO₂.
To summarise, the agreement of our results on NIST CO2 RMs 8562–8564 with the values published from several studies (Table 7) demonstrates the following: (i) consistency of our $\delta^{13}C$ calibrations on the VPDB scale by carbonate–H$_3$PO$_4$ reaction with the new primary RM IAEA-603, (ii) consistency of $\delta^{13}C$ mass spectrometry and the cross-contamination correction applied, and (iii) agreement of the new RMs with CO$_2$ RMs on the VPDB scale of 2004 and 2006. The latter implies agreement with all other carbonate RMs analysed in 2004.

We note that the VPDB $\delta^{13}C$ scale is based on carbonate RMs. Since combustible RMs were re-characterised in 2006 based on LSVEC which was included in the NIST 2004 study along with RMs 8562–8564, our agreement with RMs 8562–8564 indirectly implies an agreement (within realistic uncertainty) with the RMs re-characterised in 2006. The stability of combustible RMs after 2006 and evaluation of relevant uncertainty components deserve a separate study.

5.3 Metrological traceability of values of new RMs

The property values assigned to IAEA-610, IAEA-611 and IAEA-612 are traceable to the VPDB scale for $\delta^{13}C$ and $\delta^{18}O$ values (including the VPDB-CO$_2$ scale for $\delta^{18}O$) by direct measurements against the primary RM IAEA-603. IAEA-603 is currently used as the primary RM on the VPDB scale (replacement for NBS19); no other higher calibrator on the scale is available.

Measurements were carried out on CO$_2$ prepared by the carbonate–H$_3$PO$_4$ reaction under standard conditions and at 25 °C by applying the data treatment algorithm and $^{17}O$ correction recommended by IUPAC exactly, without approximations. The homogeneity tests and the value characterisation were performed on a statistically representative number of aliquots, both for each new RM and for IAEA-603.

The combined uncertainty (Table 5) includes all the uncertainty components described above. This way of uncertainty estimation guarantees that values obtained on a random aliquot of the new RMs, taken from any ampoule, should represent the values assigned within the uncertainty. As the analytical uncertainty was not subtracted from the homogeneity assessment, the actual material inhomogeneity may be better.

5.4 Status of IAEA-610, IAEA-611 and IAEA-612

IAEA-610/612 have been prepared to address the need for reliable and stable-over-decades RMs, and to be suitable for $\delta^{13}C$ calibrations with low uncertainty. Given that none of the $\delta^{13}C$ RMs characterised in 2004 and 2006, and available in 2018, cover a large range in $\delta^{13}C$ with low uncertainty (see Figure 4, upper panel), another goal in preparing new RMs was to re-establish the VPDB $\delta^{13}C$ scale aiming at two-point calibration/normalisation with low uncertainty and coherent RM value assignment.

IAEA-610/612 are of the same Ca-carbonate matrix as the primary RM IAEA-603. Commutable with the same methods used for $\delta^{13}C$ calibrations, they cover a $\delta^{13}C$ range from +2.46 to −36.72 ‰, and have values assigned in a mutually consistent way. Their technical characteristics (large number of identical units, stability guaranteed by sealed glass ampoules and material homogeneity) fulfil the critical requirements formulated for scale-anchoring RMs.

The analytical methods in use were validated by measurements on NIST CO2 RMs 8562–8654. The values of these RMs present the best VPDB value characterisation still available today. The $\delta^{13}C$ values obtained at the IAEA are in agreement with the published values within the k = 1 uncertainties (see above), thus validating the $\delta^{13}C$ mass spectrometry at the IAEA, including all corrections. In addition, measurements on IAEA-610/612 performed at Geotop-UQAM demonstrated material homogeneity at ~100 µg aliquots and measured values in agreement with those obtained at the IAEA.

Based on the high accuracy values and combined uncertainty assignment with three digits after the decimal point, and the agreement with NIST CO$_2$ RMs on the VPDB $\delta^{13}C$ scale realisation of 2006, we propose that the new RMs be given the status of scale-anchors on the VPDB $\delta^{13}C$ scale, aiming at $\delta^{13}C$ data normalisation and in characterising secondary RMs (see companion paper). The agreement observed in the present work with $\delta^{13}C$ values of NIST CO$_2$ RMs 8562–8654 also indirectly implies the consistency of new RMs IAEA-610/612 with the $\delta^{13}C$ values of combustible RMs characterised on the VPDB scale of 2006. Associated uncertainty components including the after-2006 stability of RMs deserve a separate study.

Multipoint scale-anchoring on the VPDB $\delta^{13}C$ scale, as based on RMs IAEA-603 and IAEA-610/612, is discussed to bring an added value in terms of the scale sustainability and verification of RM value drifts by cross-measurements in the future.

5.5 On the practical use of IAEA-610, IAEA-611 and IAEA-612

We can see several ways in which a laboratory may wish to use the four available RMs (IAEA-603 and IAEA-610/612). First, a laboratory may wish to perform a two-point “calibration” over a specific range. This is a scenario similar to the previous use of LSVEC-based two-point data normalisation. However, with the four new RMs a smaller range could be selected for the calibration, although the use of IAEA-603 is recommended to maintain the best possible link.

Second, if a laboratory wishes to make measurements using only a single RM to link their measurements to the VPDB scale (e.g., to measure some variability in a narrow range), they have a choice of materials. The combined uncertainty assigned to each of the RMs helps to establish the link. The other two RMs could be used for quality control purposes.

Third, a “multi-point calibration line”, similar to that used for atmospheric trace gas concentrations, could be constructed. While the value of this could be discussed separately, it may help
laboratories establish their own “calibration” materials (lab standards) and provide a more robust link to more negative or positive $\delta^{13}C$ values. Regarding recommendations on the two-point data normalisation and the multi-point calibration approach, we refer to numerous publications available to date (e.g., 57–63).

Fourth, since RMs IAEA-610/612 were characterised in a mutually consistent way, their use can help to identify any drift in the $\delta^{13}C$ values of RMs and also help to verify the $^{17}O$ correction schemes at user laboratories.

When not in use, RMs IAEA-610/612 should be stored in desiccators, under dry conditions at room temperature. The aliquot size should be selected according to the methods and instrumentation in use (aliquots of 10 mg and 100 μg were tested in this work) and ampoules (i.e., storage vials with plastic caps) returned to storage as soon as sampled to avoid possible problems. Since IAEA-610/612 may be sensitive to static electricity, special precautions should be taken during a sample transfer. In particular, IAEA-610 was found to be more vulnerable to this; thus, controlling its quantitative transfer to reaction vessels and monitoring its reaction yield (at 100%) are recommended.

We have assigned $\delta^{18}O$ values for information purposes only; the use of these RMs as high-level calibrators for $\delta^{18}O$ measurements is not foreseen due to three factors. First, the new RMs are fine powders and react vigorously with $H_3PO_4$; the temperature of this exothermic reaction is more difficult to control than the reaction of geological carbonates. Second, these powders may absorb local moisture that could change their measured $\delta^{18}O$ values. Third, no data scaling to the $\delta^{18}O$ VSMOW/SLAP scale by means of CO2-water equilibration (as in 27) has been performed in this study.

5.6 | Disclaimer about RM IAEA-610/612 values stated in this paper

The IAEA-610/612 values and associated uncertainties stated in this paper may be updated if more information and/or improvements of analytical techniques become available in the future. Users of these RMs should ensure that the Reference Sheet in their possession is current, and should cite the version and date of the Reference Sheet used in all publications. The current version may be found in the IAEA’s Reference Materials web-page catalogue: https://nucleus.iaea.org/sites/ReferenceMaterials/SitePages/Home.aspx.

6 | SUMMARY

1. Three new Ca-carbonate RMs, IAEA-610, IAEA-611 and IAEA-612 (IAEA-610/612), have been developed at the IAEA to address the needs for RMs with small uncertainty (e.g., 3.9) and to enable self-consistent multi-point realisation for the VPDB scale. This follows the discontinuation in 2017 of LSVEC as a RM for $\delta^{13}C$ after its value was reported to be drifting.2 Batches of the new RMs were prepared in sealed glass ampoules (3000 to 4000 ampoules in each batch of each RM). Additional material of each RM was reserved for preparation of future supplies.

2. The $\delta^{13}C$ values assigned to RMs IAEA-610/612 (Table 8) are traceable to the primary RM IAEA-603. A RM characterisation scheme, following the technical requirements of ISO Guide 35:2017, was based on representative sampling of multiple ampoules of each RM produced and focused on providing realistic uncertainty estimates that include homogeneity and stability assessment. The associated uncertainties, evaluated for ~10 mg and ~100 μg aliquots (Table 8), represent how well a single (or few) aliquot, taken from a randomly selected RM ampoule, represents its assigned value. These uncertainty estimates are made in a conservative way and based on maximal estimates of RM inhomogeneity assessment, instrumental corrections and the $^{17}O$ correction. Since the new RMs were not intended for high accuracy $\delta^{18}O$ calibrations, $\delta^{18}O$ values are specified for information only.

3. The $\delta^{13}C$ values obtained in this work on a representative number of NIST CO2 RM units agree with the values published by NIST within the stated uncertainties. This agreement demonstrates the consistency of the assigned $\delta^{13}C$ values of the new RMs IAEA-610/612 with the VPDB $\delta^{13}C$ scale realisation of 2006 which was based on RMs NBS19 and LSVEC11. Since the VPDB scale is based on carbonate RMs, and because all carbonate and CO2 RMs were characterised in a self-consistent way in 2004 and no new measurements have been performed since then, the latter agreement indirectly implies an agreement with the values of combustible RMs re-characterised in 2006 by Coplen at al11; a detailed study on these RMs is still needed.

4. Together with IAEA-603, new RMs IAEA-610/612 provide a self-consistent RM set with coherently assigned $\delta^{13}C$ values of high accuracy. IAEA-610/612 can be used to scale anchor RMs for $\delta^{13}C$ calibrations and two-point data normalisation, including value assignment for new secondary RMs. Discussion on the VPDB $\delta^{13}C$ scale realisation principles and related concepts can be found in a

<table>
<thead>
<tr>
<th>TABLE 8</th>
<th>$\delta^{13}C$ values assigned to the new RMs in 2020(*) and $\delta^{18}O$ values given for information purposes. Uncertainties are given with k = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM</td>
<td>$\delta^{13}C$, % VPDB, k = 1 uncertainties given at different aliquot sizes</td>
</tr>
<tr>
<td>IAEA-610</td>
<td>$-9.109 \pm 0.012 (~10 mg)$</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.03 (~100 μg)$</td>
</tr>
<tr>
<td>IAEA-611</td>
<td>$-30.795 \pm 0.013 (~10 mg)$</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.04 (~100 μg)$</td>
</tr>
<tr>
<td>IAEA-612</td>
<td>$-36.722 \pm 0.015 (~10 mg)$</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.03 (~100 μg)$</td>
</tr>
</tbody>
</table>

*For the latest updates, users should refer to the current Reference Sheet at the IAEA webpage https://nucleus.iaea.org/sites/ReferenceMaterials/SitePages/Home.aspx.

LSVEC $\delta^{13}C$ value was obtained by measurements against NBS19, see text.

For two-point data normalisation, see publications referenced in the text.
REFERENCES


29. Installation Acceptance Form MAT 253 Rev 1.2. ThermoFisher Scientific, Bremen, Germany (official document provided to owners of MAT 253 instruments during installation); 2014.


36. Assonov S, Taylor P, Brenninkmeijer CAM. A system for high-quality CO$_2$ isotope analyses of air samples collected by the CARIBIC Airbus

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of this article.

SUPPORTING INFORMATION for paper https://doi.org/10.1002/rcm.9014

“Characterisation of new reference materials IAEA-610, IAEA-611 and IAEA-612 aimed at the VPDB δ13C scale realisation with small uncertainty”,
by Sergey Assonov, Ales Fajgelj, Jean-François Hélie, Colin Allison and Manfred Gröning

For literature citations 1-63, see main article; for citations 64–67, see end of this document.

S1. Abbreviations, and acronyms, in use

CRM - certified reference material. VIM 3\textsuperscript{12} describes a CRM as a, “reference material, accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceabilities, using valid procedures.”

DI(-mode) – dual inlet gas introduction to mass-spectrometer,

GGMT – Greenhouse Gas Measurement Techniques, the abbreviation used for WMO/IAEA meetings on Greenhouse Gas Measurement Techniques.

GHG – greenhouse gases.

IAEA-603 – reference material (marble Ca-carbonate, mineral calcite) released in 2016\textsuperscript{23,46} by the IAEA in order to replace NBS19 (supply exhausted).

IRMM - Institute for Reference Materials and Measurements, located in Geel, Belgium; one of institutes belonging to the Joint Research Centre (JRC) of European Commission. The current acronym is JRC-Geel.

IRMS – Isotope Ratio Mass Spectrometry.

IT principle – Identical Treatment principle. All samples and RMs used for calibration shall follow the same preparation, the same gas introduction and the same data treatment.

LSVEC – reference material (in the form of Li$_2$CO$_3$) at negative δ$_{13}$C aimed for δ$_{13}$C normalization, introduced this function in 2006\textsuperscript{1}. In 2015 it was found as not suitable for this purpose and in 2016 was removed from sales at the IAEA. In 2017, IUPAC officially discontinued LSVEC as RM for δ$_{13}$C\textsuperscript{4}.

NBS19 - the reference material (marble Ca-carbonate) with δ$_{13}$C= 1.95 ‰ and δ$_{18}$O= -2.20 ‰ used for the realisation of the VPDB scale and VPDB-CO$_2$ scale\textsuperscript{64}, by means of NBS19-CO$_2$ gas produced by reaction with H$_3$PO$_4$ under specified standard conditions. In 2016 a replacement for NBS19, IAEA-603 was released by the IAEA\textsuperscript{23,46}.

PDB – Pee Dee Belemnite, which was used by H.Craig as a reference to report relative δ$_{13}$C and δ$_{18}$O measurement results\textsuperscript{18,19}.

RM – reference material.

SD – Standard Deviation.

SE – Standard Error, often known as Standard Error of the Mean (SEM), describes the uncertainty related to the mean value.

VPDB – Vienna PDB, the international conventional reference scale used to express δ$_{13}$C and δ$_{18}$O measurement results relative to the (non-existing) artefact VPDB. The VPDB δ$_{13}$C and δ$_{18}$O scale is realised through the metrological reference material IAEA-603 (recently replaced NBS19), see above. The VPDB δ$_{18}$O-scale is used to express δ$_{18}$O of carbonates only.
VPDB-CO₂ - the international reference scale used to express δ¹⁸O measurement results of CO₂ gases, including CO₂-in-air. Expressing δ¹⁸O of CO₂-in-air on the "VPDB-carbonate" scale is not accepted. Note that δ¹⁸O values on the VPDB or VPDB-CO₂ scales are not the same; values of δ¹³C are expressed either on VPDB or VPDB-CO₂ scale which is the same for δ¹³C.

WMO – World Meteorological Organisation

WR-CO₂ – working reference CO₂ filled in the dual inlet of a mass-spectrometer.

S2. Analytical methods and equipment at the IAEA

The methods and equipment in use at the IAEA for the current project are the same as used in the RM IAEA-603 characterisation. The multi-purpose preparation system (Scientific Solutions, NZ, produced in 1997, this small company no longer exists) was used for the carbonate-H₂PO₄ acid reaction. Carbonate ~10 mg aliquots were reacted with ~2.0 mL H₃PO₄ (1.925-1.930 g/cm³ density) at 25.0±0.1°C for ~20 to 24 hours while shaking. The importance of the specific acid density has been discussed in several publications (e.g., and references therein). As the value transfer to the new RMs includes analyses of IAEA 603, we use the same acid density and other reaction conditions as used for the IAEA-603 value assignment.

These reaction conditions mainly correspond to the carbonate-H₂PO₄ reaction used for the NBS19 value assignment and RM value assignment study led by at the National Institute of Standards and Technology (NIST) in 2004, with the difference that the acid density was selected following. Under these conditions, the carbonates are fully dissolved. CO₂ portions released were cryo-dried at -80 °C and transferred into glass transfer-vessels with pneumatic valves.

S3. Mass-spectrometry and standard tests on MAT 253

Measurements at the IAEA were performed on a mass-spectrometer MAT 253™ (Thermo Fisher Scientific) operated in DI mode with oil-free fore-vacuum pumping, the performance of the MAT 253 has been previously described. When assigning values for high level RMs over a wide δ¹³C range, the mass-spectrometry effects potentially decreasing (“shrinking”) value differences between sample and WR-CO₂ shall be considered. The related standard tests included the following:

- Mass-spectrum baseline was regularly monitored and no background peaks were observed.
- Baseline around peaks of m/z 45 and m/z 46 was monitored, and no negative baseline as described in some of the latest publications (e.g.,) was observed.
- Measurements were performed in DI mode, with m/z 44 signals of both sample-CO₂ and WR-CO₂ adjusted to approximately 4500 mV. Over the measurement cycle, intensity drops were small with the intensity mismatch at the end of the measurement no more than ~200 mV.
- Linearity tests were performed that demonstrated the system was linear to within ±0.02 ‰/V indicating that linearity effects could be neglected. CO₂ produced from the new RMs and IAEA-603 were introduced into the same bellow and capillary, and both experienced the same signal drops during analysis. In this way, any linearity effect was eliminated.
The linearity test also implies the linearity of amplifiers and high ohmic resistors. The “best” linearity test values were observed within linearity ($\delta^{45}R$) SD$\leq \pm 0.01 \%_/V$ in the broad range of $m/z$ 44 signals. Without high linearity of amplifiers and resistors, one cannot not observe the high linearity of ion beam ratios.

Presence (absence) of $H_2O$ and air-leaks. Signals at $m/z$ 18 and $m/z$ 40 were monitored in all analytical sequences, and no elevated $H_2O$ (compared to pure WR-CO$_2$ gas) or air-leaks were detected.

Other standard tests implemented in the ISODAT software were performed upon filament replacement and ion source retuning, as follows:

- Tailing of the high intensity peak $m/z$ 44 to the low intensity peak $m/z$ 45 (Abundance test). The value observed at 1.0E-6 is within specifications and guarantees a negligible contribution to $m/z$ 45. For $m/z$ 44 signal at 4500 mV, this tailing at the centre of $m/z$ 45 was observed at $\sim$0.35-0.40 mV. For the $\delta^{45}R$ difference of $-38.6 \%$ between IAEA-612 and IAEA-603, the maximum bias in $\delta^{45}R$ was estimated at $-0.0035 \%$. For IAEA-610 and IAEA-611, the potential biases are lower, proportional to the $\delta^{45}R$ difference.

An estimation of the related uncertainty components affecting raw $\delta^{45}R$ values is given below in Table 1S.

Table S1

<table>
<thead>
<tr>
<th>Parameter tested</th>
<th>Observed value</th>
<th>Maximal effect on raw $\delta^{45}R$, estimated at $\delta^{45}R$ = $-38.6 %$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity in $\delta^{45}R$ response due to the ion source tuning</td>
<td>Within $\pm 0.02 %_/V m/z$ 44</td>
<td>No effect. Effect is cancelled out due to the use of the Identical Treatment: CO$_2$ produced from the new RMs and from IAEA-603 followed the same gas introduction and signal decrease during measurements.</td>
</tr>
<tr>
<td>Amplifier’s linearity in $\delta^{45}R$</td>
<td>The lowest Linearity test values within $\pm 0.01 %_/V m/z$ 44, observed several times</td>
<td>$\pm 0.0010 %$ Mismatch of $m/z$ 44 signals up to $\sim$250 mV at the end of 12 changeovers. The average mismatch is taken at 200 mV.</td>
</tr>
<tr>
<td>The tailing contribution of $m/z$ 44 peak up to 0.35-0.40 mV to $m/z$ 45</td>
<td>Abundance test, value of $\sim$1.0E-6</td>
<td>$\sim 0.0035 %$ for $\delta^{45}R$=38.6 % (value of IAEA-612 against IAEA-603), lower effect for a smaller difference in $\delta^{45}R$</td>
</tr>
</tbody>
</table>
Table 2S summarises the cross-contamination factors observed in this work and published values. Examples of correction magnitude for raw \( \delta^{13}R - \delta^{16}R \) for CO2 gasses measured in this study (Section 2.2) are given.

**Table S2**
The cross-contamination factors observed in this work and published values. Examples of correction magnitude for raw \( \delta^{13}R - \delta^{16}R \) are given.

<table>
<thead>
<tr>
<th>Instrument type and method used to estimate the cross-contamination</th>
<th>( \eta_{E5} )</th>
<th>Example of ( \delta^{13}R ), ( % )</th>
<th>Corrected ( \delta^{13}R ), ( % )</th>
<th>( \eta_{E6} )</th>
<th>Example of ( \delta^{16}R ), ( % )</th>
<th>Corrected ( \delta^{16}R ), ( % )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT 253 used in this work, pairwise measurements on CO2 with a large isotope difference</td>
<td>(2.17±1.07) E-04</td>
<td>-40.000</td>
<td>-40.017 ±0.008 E-04</td>
<td>(5.78±2.58) E-04</td>
<td>-35.000</td>
<td>-35.040 ±0.018</td>
</tr>
<tr>
<td>MAT 253 at NIST, pairwise measurements of NIST RMs CO2 8562 vs 8563 (^{32})</td>
<td>1.50E-03</td>
<td>-40.000</td>
<td>-40.118 2.70E-03</td>
<td>-35.000</td>
<td>-35.187</td>
<td></td>
</tr>
<tr>
<td>Maximum value reported in NIST study (Table 3 in (^{25})), instruments not specified</td>
<td>3.85E-03</td>
<td>-40.000</td>
<td>-40.304 5.16E-03</td>
<td>-35.000</td>
<td>-35.359</td>
<td></td>
</tr>
<tr>
<td>MAT252 at NIST, value before ion source cleaning / upgrade, VISC closed (Table 1 in (^{16}))</td>
<td>1.07E-02</td>
<td>-40.000</td>
<td>-40.857 2.18E-02</td>
<td>-35.000</td>
<td>-36.566</td>
<td></td>
</tr>
<tr>
<td>MAT252 at IRMM, pairwise measurements on CO2 with a large isotope difference (^{16})</td>
<td>1.9E-04</td>
<td>-40.000</td>
<td>-40.015 4.3E-04</td>
<td>-35.000</td>
<td>-35.030</td>
<td></td>
</tr>
<tr>
<td>MAT 253 at MPI-BGC, two sets of values reported for 2018 and 2019 (Figure 2, in Qi et al. (^{17}))</td>
<td>(4.6 ±5.0) E-04 and (3.2 ±5.0) E-04</td>
<td>-40.000</td>
<td>40.036 ±0.039 and -40.025 ±0.039 E-04</td>
<td>(9.7±10.0) E-04 and (7.8±10.0) E-04</td>
<td>-35.000</td>
<td>-35.067 ±0.069 and -35.054 ± 0.069</td>
</tr>
</tbody>
</table>

**S4. Reproducibility of WR-CO2 refills and absence of “bleeding effect”**
To use a WR-CO2 as an internal reference frame used to pool data, the reproducibility of the WR-CO2 refills is critical. Another undesirable effect is a potential dependence of \( \delta^{13}R - \delta^{16}R \) of WR-CO2 due to gas consumption in the reference bellow (“bleeding effect”). Such an effect may result in a visible change of \( \delta^{13}R - \delta^{16}R \) of the WR-CO2 through time which would seriously limit the ability to pool data, both in the homogeneity tests and in the value characterisation, unless corrected for. In the case of a poorly optimised mass-spectrometer operation, a drift of \( \delta^{13}C \) of WR-CO2 over 10 hrs may account for 0.03 \% (Figure 1 in \(^{28}\)). This effect diminishes the analytical uncertainty and shall be eliminated.

First, the data obtained in the homogeneity tests can be used to demonstrate the \( \delta^{13}R \) reproducibility of the WR-CO2 refills. The raw /uncorrected data for IAEA-610/612 (Figures 1, 2 and 3 in main text) show that the data clusters (value averages) presented zero trends in time. If a significant bleeding of drift effect took place, observing such small SD(\( \delta^{13}C \))=±0.006 \% values for each analytical sequence and for the data pooled over several sequences (Table 3 in main text) would not be possible.

For IAEA-610, only the 1\(^{st}\) analytical sequence, separated by 18 calendar days from the 2\(^{nd}\), 3\(^{rd}\) and 4\(^{th}\) sequences, demonstrated a small deviation in \( \delta^{13}C \) (Figure 1 in main text). Since we were interested in the data scatter in a sequence, the shift-correction estimated at 0.016 \% was applied; this shift is comparable to the MAT 253 daily specification of SD(\( \delta^{13}C \), n=10)≤0.0125

4
% (to be demonstrated without any sample CO₂ and WR-CO₂ refills) and much smaller than the bleeding effect of 0.03 % over 10 hrs as cited above (Figure 1 in [28]).

The δ¹⁸O reproducibility of WR-CO₂ refills is less critical, the absence of any effect is demonstrated by Figures 1-3 (in main text).

Second, one can zoom-in to the uncorrected data obtained in the homogeneity tests. 16 analytical sequences were performed in total, and none demonstrates any bleeding trends for δ¹³C as plotted in time (Figure S1, Supplementary Information).

Third, the uncorrected δ⁴⁵R- δ⁴⁶R raw data obtained on NIST RMs 8562-8564 CO₂ (these were analysed alone with new RMs in the value characterisation) against WR-CO₂ also demonstrate zero trend (Figure S2, Supplementary Information).

S5. Data treatment
A consistent data treatment was applied at all steps. The δ¹³C and δ¹⁸O values of the new RMs were calculated without knowing exactly the δ¹³C and δ¹⁸O values of the WR-CO₂ in use. Following the IAEA and International Union of Pure and Applied Chemistry (IUPAC) recommendations [40,43], the raw δ⁴⁵R-δ⁴⁶R values obtained on the samples against the WR-CO₂ were scaled (in terms of raw δ⁴⁵R-δ⁴⁶R values) to the VPDB-CO₂ scale-zero by using δ⁴⁵R-δ⁴⁶R values obtained on the primary RM IAEA-603 used as the calibrator.

The ¹⁷O correction was applied based on the IUPAC recommended set of numerical values (Table 2 in [40]), by solving all equations without approximations.

S5. On IAEA-603 analytical uncertainty and its uncertainty associated with the certificate values.
First, in the RM value assignment one links the average values measured on new RMs to the average value obtained by measurements on IAEA-603; the associated analytical uncertainties are taken as Standard Errors. We note that another term used for “Standard Error” is “Standard Deviation of the Mean”, this describes its as the uncertainty related to the mean value.

Second, in uncertainty budget (Table 5, main text) we take the IAEA-603 certificate uncertainty without its homogeneity term for the reasons as following. The IAEA utilises the concepts of ISO Guide 35:2017 [5] which implies that the uncertainty associated with the RM value describes how well a(ny) single aliquot (of specified size) taken from a randomly selected ampoule represents the value assigned to RM (more details in the main text and in [5,15,39]); this concept was applied to IAEA-603 [23,46]. Given that the IAEA-603 certificate uncertainty specifies its single aliquot, and given that in this study we analysed 30 aliquots of IAEA-603 in total (taken from 5 randomly selected ampoules), the uncertainty component related to IAEA-603 inhomogeneity is included in the 30 aliquots of analysed. The averaged value better represents the value than a single aliquot. In this way, this uncertainty component is not taken twice.

Third, the combined uncertainties (Table 5, main text) provide the maximal estimates which are larger than the actual uncertainty. This is because the analytical uncertainty related to the methods in use (CO₂ preparation by carbonate-acid reaction followed by IRMS analysis) was not subtracted from the data variance observed in the homogeneity assessment of IAEA-603 [23] and of IAEA-610/612 (the present work); such subtraction is typically a part for the RM value assignment concept [15,39] which is used by ISO Guide 35:2017 [5].
Figure S1

Upper panel:
Relative scatter of $\delta^{13}$C values of IAEA-610 obtained in four analytical sequences (left to right) during the homogeneity study against the same WR-CO2 used as internal reference calibration frame. Values of the 1st sequence were shift-corrected for 0.016 ‰, see text.

Plots are similar to Figure 1 (main text) and demonstrate detailed zooming-in of the time axis. Unity of time code corresponds to one calendar day.
Figure S1 (continued).

Middle panel:
Relative scatter of $\delta^{13}$C values of IAEA-611 obtained in four analytical sequences (left to right) during the homogeneity study against the same WR-CO$_2$ used as internal reference calibration frame.

Plots are similar to Figure 2 (main text) and demonstrate detailed zooming-in of the time axis. Unity of time code corresponds to one calendar day.
Figure S1 (continued).

Lower panel:
Relative scatter of $\delta^{13}$C values of IAEA-612 obtained in four analytical sequences (left to right) during the homogeneity study against the same WR-CO$_2$ used as internal reference calibration frame.

Plots are similar to Figure 3 (main text) and demonstrate detailed zooming-in of the time axis. Unity of time code corresponds to one calendar day.
Raw $\delta^{45}$R-$\delta^{46}$R values obtained on NIST RM 8562 and RM 8563 against WR-CO$_2$ (the same as used in the value assignment) demonstrate a stability of WR-CO$_2$ refills over 10 days. As error bars, MAT 253 daily-specification values are given with k=2.

Unity of time code corresponds to one calendar day.
REFERENCES:


