

Final Report on CCQM-K167: Carbon Isotope Delta Measurements of Vanillin

Michelle M.G. Chartrand^{*1}, Ian Chubchenko², Philip J.H. Dunn³, Bruno Carius Garrido⁴, Lu Hai⁵, Fong-Ha Liu⁶, Nives Ogrinc⁷, Adnan Şimşek⁸, Maíra Fasciotti⁴, Heidi Goenaga-Infante³, Sarah Hill³, Anna Kolobova², Dmitriy Malinovsky³, Jeffrey P. Merrick⁶, Eliane Cristina Pires do Rego⁴, Doris Potočnik⁷, Thays Vieira da Costa Monteiro⁴, Wagner Wollinger⁴, Li Xiao⁵, Blaza Toman⁹, Zoltan Mester¹, and Juris Meija¹

¹National Research Council Canada, Metrology, 1200 Montreal Rd., Ottawa, K1A 0R6, Canada

²D. I. Mendeleev Institute for Metrology, 190005, 19 Moskovsky pr., St. Petersburg, Russia

³LGC Limited, Queen's Road, Teddington, Middlesex TW11 0LY, UK

⁴National Institute of Metrology, Quality and Technology, Avenida Nossa Senhora das Graças, 50 – Xerém – Duque de Caxias – RJ, 25250-020, Brazil

⁵National Institute of Metrology, No. 18 Bei San Huan Dong Lu, Chaoyang District, Beijing 100029, PR China

⁶National Measurement Institute Australia, 105 Delhi Rd., North Ryde, NSW 2113, Australia

⁷Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

⁸TÜBİTAK Ulusal Metroloji Enstitüsü, Barış mah. Dr. Zeki Acar cad. No:1, 41470 Gebze/KOCAELİ, Turkey

⁹National Institute of Standards and Technology, 100 Bureau Dr., Gaithersburg, MD 20899, USA

*Correspondence to: M.M.G. Chartrand, Michelle.Chartrand@nrc-cnrc.gc.ca

Abstract

The CCQM Isotope Ratio Working Group (IRWG) determined that an additional key comparison of carbon isotope delta measurements was required to capture the progress of this field. Vials containing 0.25 mg of vanillin were prepared at NRC, and evaluated for bottle-to-bottle homogeneity prior to distribution to the eight participating institutes. Participants were able to choose any suitable method and reference materials for carbon isotope delta measurements, and report a carbon isotope delta value and the associated uncertainty, and analysis details. To determine the key comparison reference value (KCRV) and its associated uncertainty, the NRC in collaboration with the NIST, developed a multivariate Bayesian random laboratory effects model, which also incorporates the uncertainty due to bottle-to-bottle homogeneity and any correlations between the reported results that arise due to the use of common reference materials. The KCRV for this study was determined to be -25.833 ‰ relative to the VPDB, with associated uncertainty of 0.028 ‰, and expanded uncertainty of 0.056 ‰ ($k=2$). All the results reported by the participants in this study were considered equivalent to the KCRV.

Table of Contents

1.0 Introduction and Justification	4
2.0 Details of the Key Comparison	5
2.1 Sample Preparation	5
2.2 Sample Evaluation Prior to Distribution	5
2.3 Instructions Provided to Institutes	8
2.4 Schedule	9
2.5 Participating Institutes	9
3.0 Reported $\delta_{VPDB}^{(13C)}$ Measurement Parameters	10
3.1 Analysis Technique and Instrumentation used	10
3.2 Metrics to Optimize Performance	10
3.2.1 System Background Test	10
3.2.2 Stability Test	11
3.2.3 Linearity Test	11
3.2.4 Other Tests	11
3.2.5 Sample Preparation and Analysis Metrics	11
3.3 Measurement Sequence Details	12
3.4 Calibration and Traceability to the VPDB Scale	12
3.4.1 Quality Control Samples	14
3.5 Post-Measurement Analysis – Applied Corrections	14
3.5.1 ^{17}O Correction	15
3.5.2 Blank Correction	15
3.5.3 Drift Correction	15
3.5.4 Linearity Correction	15
3.5.5 Memory (or Carry-Over) Correction	16
4.0 Reported $\delta_{VPDB}^{(13C)}$ Vanillin Results and Uncertainties	16
4.1 Uncertainty budget	17
5.0 KCRV Calculation and Associated Uncertainty	19
5.1 KCRV – Choosing a statistical model	19
5.2 Considerations in the development of the Bayesian REM to determine the KCRV and associated uncertainty	20
5.2.1 Inclusion of Uncertainty due to Homogeneity	20

5.2.2 Correlations Between the Reported Results	21
5.2.3 Correlations Between the Reference Materials	22
5.3 Calculation of the KCRV	23
5.4 Degrees of Equivalence.....	26
6.0 CMC Claim.....	27
7.0 References.....	28
Appendix 1. Registration	31
Appendix 2. Measurement Protocol Sent to Participants	32
Appendix 3. Submission of Results.....	35

1.0 Introduction and Justification

Verification of the authenticity of food items is essential to ensure the quality and safety of food products. Vanillin extracted from vanilla beans follows a lengthy and expensive process, and the price can be hundreds-fold more expensive than vanillin synthesized from inexpensive, petrol-based precursors [1, 2]. In recent years, vanillin synthesized from rice or wood has become increasingly popular, as consumers wish to move towards the consumption of “natural” products [1, 2].

Carbon isotope delta, $\delta_{VPDB}(^{13}C)$, measurements are routinely used to determine the authenticity of food products by determining the source of the food product, and detecting adulteration of food products via the addition of unreported, and often lower-cost, additives [2]. Carbon isotope delta measurements in a sample are usually expressed on the VPDB scale as:

$$\delta_{VPDB}(^{13}C, \text{ sample}) = (r_{\text{sample}} / r_{VPDB}) - 1$$

where, r is the carbon isotope amount ratio [3] in the sample and standard, respectively, with $r = n(^{13}C)/n(^{12}C)$, and VPDB denotes the Vienna Pee Dee belemnite, which is a virtual material. In practice, the ion intensities of mass 44, 45 and 46 are measured, and the ratios of these ion intensities are used to calculate $\delta_{VPDB}(^{13}C)$. Since 2006, the VPDB carbon isotope delta scale has been realized using two reference materials (RMs), NBS 19 calcium carbonate and LSVEC lithium carbonate, which have consensus values of $\delta(^{13}C, \text{ NBS 19}) = +1.95 \text{ ‰}$ exactly and $\delta(^{13}C, \text{ LSVEC}) = -46.6 \text{ ‰}$ exactly [4]. Recent studies have noted that the carbon isotope ratio of LSVEC is not stable over time [5, 6]. Accordingly, the International Atomic Energy Agency (IAEA) [6] and IUPAC [7] no longer recommends the use of LSVEC for normalization of isotopic measurements, but $\delta_{VPDB}(^{13}C)$ measurements must still be normalized to the VPDB scale, preferably using at least two appropriate RMs selected by the analysts [8]. When reporting $\delta_{VPDB}(^{13}C)$ measurements, it is recommended to also report the $\delta_{VPDB}(^{13}C)$ values and the associated uncertainties of the RMs used to normalize samples [9].

Isotope delta measurements are currently not traceable at the required uncertainty level to the International System of Units (SI), and a traceability exception to use conventional delta scales has been approved by the International Committee for Weights and Measures (Comité International des Poids et Mesures, CIPM) [10]. Within the Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM), there have been a handful of key comparisons and pilot studies involving isotope delta measurements. The most recent key comparison for $\delta_{VPDB}(^{13}C)$ measurements (CCQM-K140) was conducted in the Inorganic Analysis Working Group (IAWG), measuring honey [11]. The October 2018 meeting of the Isotope Ratio Working Group (IRWG) in Ottawa, Canada, determined that additional Key Comparisons of light stable isotope measurements is needed. This vanillin study, CCQM-K167, is the first key comparison for carbon isotope delta $\delta_{VPDB}(^{13}C)$ measurements initiated by the IRWG. A parallel pilot study, CCQM-P211, was simultaneously conducted, and those results are reported elsewhere.

2.0 Details of the Key Comparison

2.1 Sample Preparation

Vanillin, an easily combustible and forensically relevant sample, was chosen for this key comparison. A 2 kg tub of high-purity vanillin (> 0.99 g/g) was purchased from Sigma Aldrich Canada. An aliquot of this vanillin was sieved using US Standard 30, 40 and 50 mesh sieves, and the fraction that fell through the 50 mesh sieve was collected. 100 vials (2 mL glass “shorty” vials, with polycarbonate lined caps) were prepared using the fraction that fell through the US Standard 50 mesh sieve, with each vial containing between 200 and 250 mg of vanillin. Vials were stored in a drybox at room temperature until analyzed, or distributed to participants.

2.2 Sample Evaluation Prior to Distribution

Prior to distribution of vanillin samples to the study participants, 20 vials of vanillin were used to evaluate homogeneity and short term stability [12]. To assess short term stability, an isochronous experiment was performed (Figure 1). Three vials of vanillin were subdivided into three aliquots, A, B and C, and stored in a glass vial with a tightly closed lid. Aliquots A were stored for 14 days in a drybox with Dririte[®] desiccant at room temperature (20 °C), aliquots B were stored in a 40 °C oven for 7 days, then moved to a drybox with Dririte[®] desiccant at room temperature for 7 days, and aliquots C were stored in an oven at 40 °C for 14 days. The temperatures 40 °C and 20 °C could represent typical environments during shipping and recommended storage.

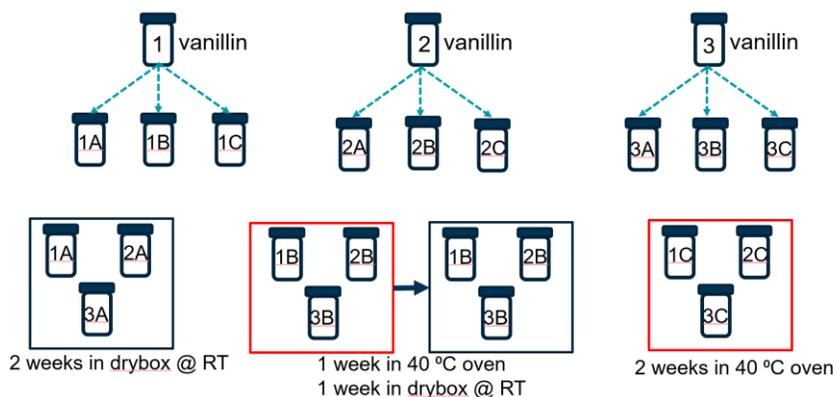


Figure 1. Scheme for assessing short term stability of vanillin.

After 14 days, four aliquots of ~ 665 μg were weighed from each sample, and were measured at NRC. Five RMs, with their assumed $\delta_{\text{VPDB}}(^{13}\text{C})$ values \pm their standard uncertainty, were used for normalization of the vanillin (Figure 2): IAEA-CH-6, -10.45 ± 0.04 ‰ [13]; USGS65, -20.29 ± 0.04 ‰ [14]; IAEA-600, -27.77 ± 0.04 ‰ [8]; NBS 22, -30.03 ± 0.05 ‰ [8]; and USGS61, -35.05 ± 0.04 ‰ [15].

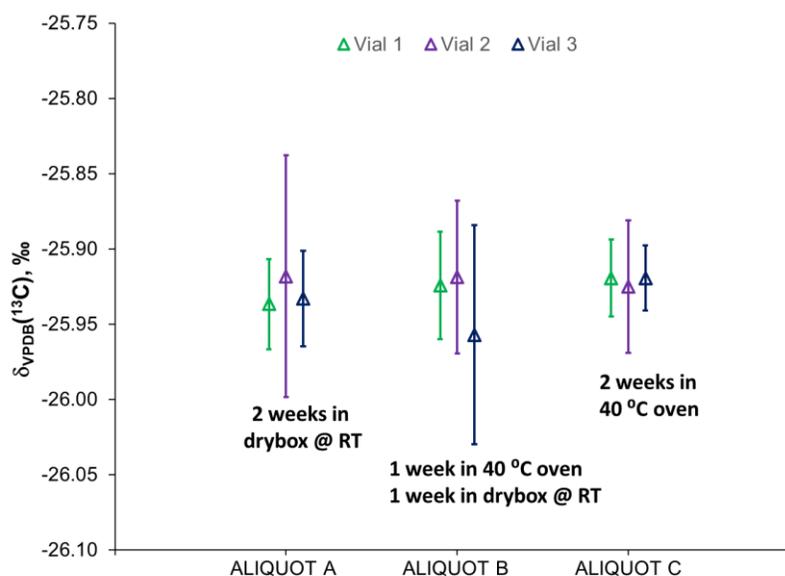


Figure 2. Results of the short term stability for vanillin: symbol is the average of the four measurements, and the vertical lines represent one standard deviation of the four replicate measurements.

The average $\delta_{\text{VPDB}}(^{13}\text{C})$ values of all nine samples were similar, and all 36 individual measurements ranged between -25.80 ‰ and -26.04 ‰. There is no correlation between temperature and the average $\delta_{\text{VPDB}}(^{13}\text{C})$ values for vanillin, demonstrating that vanillin is stable under the conditions outlined in this experiment, and fit for purpose.

The 17 remaining vials of vanillin were used for CCQM-K167/P211. Bottle to bottle homogeneity was assessed at NRC for each of the 17 vials. Three to five aliquots of ~ 665 μg from each of the 17 vials were weighed and measured for $\delta_{\text{VPDB}}(^{13}\text{C})$ at NRC over three different measurement sequences (Figure 3), normalizing to IAEA-CH-6, USGS65, IAEA-600, NBS 22 and USGS61, with $\delta_{\text{VPDB}}(^{13}\text{C})$ values and standard uncertainties given above. Out of the 17 vials, 9 were distributed to CCQM-K167 participants, 4 to CCQM-P211 participants, and 4 were not distributed.

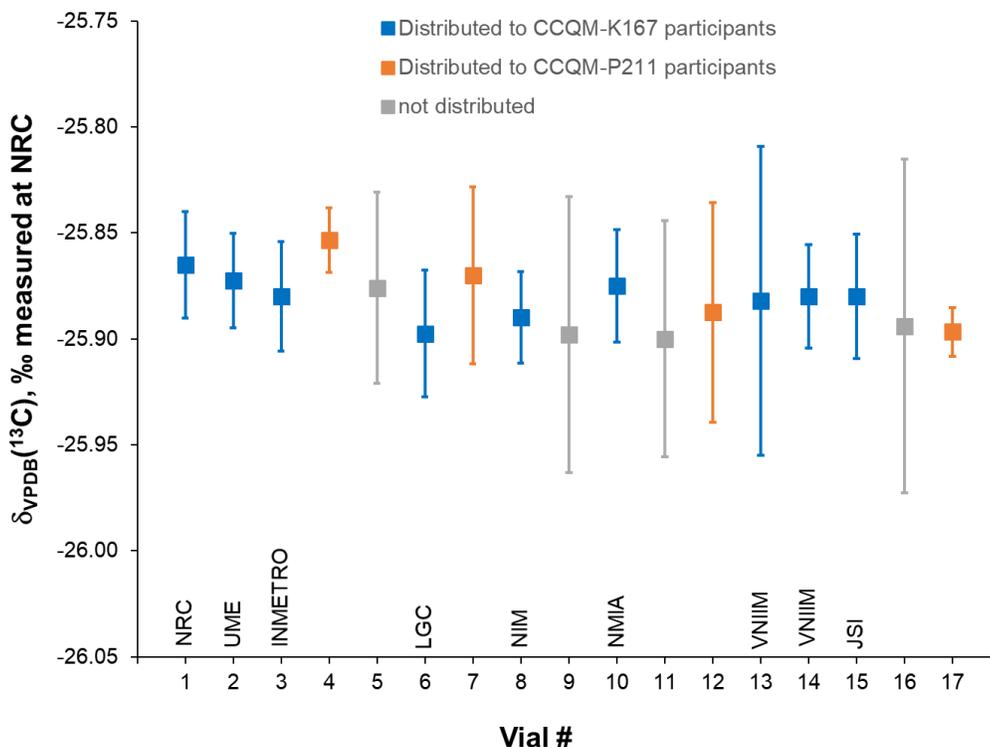


Figure 3. NRC measured $\delta_{VPDB}(^{13}C)$ values for all 17 vials available for CCQM-K167/P211. Nine vials (blue) were distributed to eight K167 participants, and four vials (orange) were distributed to the four P211 participants. Colored squares are the average ($N = 3$ to 5) $\delta_{VPDB}(^{13}C)$ values, and the vertical lines are 1 standard deviation on replicate measurements.

Homogeneity was evaluated using Bayesian ANOVA, with weekly-informative standard prior distributions of all parameter estimates. Homogeneity was determined for two sets of data (Table 1): 1) for all 17 vials, and 2) on the 9 vials distributed to CCQM-K167.

Table 1. Homogeneity estimates from the full set of 17 vials and the 9 vials used in CCQM-K167.

Parameter	ALL 17 VIALS		9 CCQM-K167 VIALS	
	u , ‰	95 % CI, ‰	u , ‰	95 % CI, ‰
Homogeneity	0.004	0.000 to 0.012	0.004	0.000 to 0.014
Measurement	0.042	0.036 to 0.048	0.034	0.028 to 0.042

For the 17 vials, the uncertainty due to homogeneity was $u_{hom} = 0.004$ ‰, which is small compared to the estimated measurement uncertainty (0.042 ‰), demonstrating that this set of 17 vials is adequately homogeneous. Comparatively, for the set of 9 vials used for the CCQM-K167 study, the uncertainty due to homogeneity was identical, 0.004 ‰, demonstrating that this method for evaluating the homogeneity of the vanillin sample (with respect to its carbon isotopic composition) is robust, as there was no significant difference in the uncertainty between the two sets of vanillin samples.

Although the bottle-to-bottle homogeneity for the 9 vials is presented in Table 1 as a single value (0.004 ‰, the median of the distribution), the homogeneity uncertainty is best represented as a gamma distribution, with a shape of 1.26, and a rate of 236 (Figure 4).

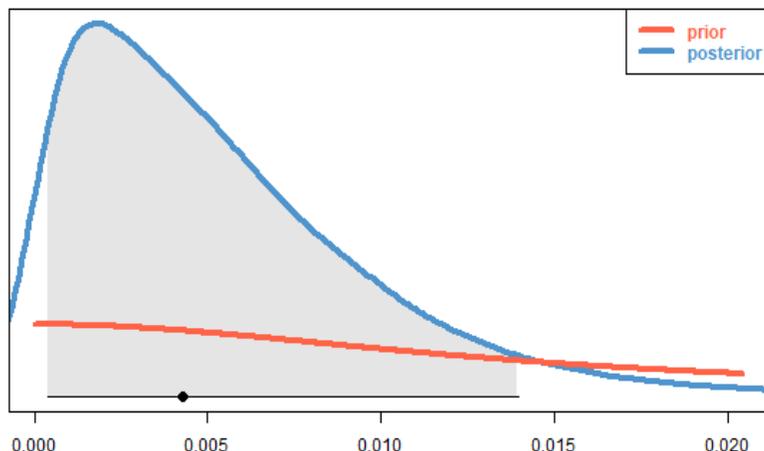


Figure 4. Homogeneity estimate for the 9 vials used in CCQM-K167 can be summarized with a gamma distribution with shape = 1.26 and rate = 236, whose median is 0.004 ‰.

2.3 Instructions Provided to Institutes

The measurand was carbon isotope delta relative to the Vienna Peedee belemnite, $\delta_{VPDB}^{(13C)}$ where VPDB is a virtual material defined by reference material NBS 19 such that $\delta_{VPDB}^{(13C, NBS19)} = +1.95$ ‰ exactly [16]. For the vanillin sample, the target $\delta_{VPDB}^{(13C)}$ value is -20 ‰ to -40 ‰.

Samples were distributed by courier to the participants, with each laboratory participating in the key comparison receiving one sample vial of vanillin, except for VNIIM, who requested 2 samples. Institutes were asked to inform the coordinator immediately if the test sample has been compromised in any way, or if it arrives in questionable condition. It was recommended to keep the vial containing vanillin tightly capped and stored in a cool dry place out of direct sunlight when not in use.

Participants were welcome to select any suitable method for $\delta_{VPDB}^{(13C)}$ measurements, and choose any appropriate $\delta_{VPDB}^{(13C)}$ reference material(s) to realize the VPDB scale, or perform absolute carbon isotope ratio measurement and report it on the VPDB scale. Due to the issues surrounding the long term stability of LSVEC, the use of LSVEC was not acceptable [7].

The study protocol required a detailed description of the instrumentation, calibration methods including traceability to the VPDB scale, and analysis details. The request for a description of the carbon isotope delta analysis details was purposefully kept vague, as we did not want to influence a laboratory's operations by suggesting procedures they

“should” employ. The protocol requested the ^{17}O correction be performed using the procedure recommended by IUPAC [17].

Participants were asked to report a final single $\delta_{\text{VPDB}}(^{13}\text{C})$ value for their vanillin measurements, standard uncertainty, and expanded uncertainty including the coverage factor. Additionally, it was requested that at least five replicates of vanillin be measured and reported; as participants may use common $\delta_{\text{VPDB}}(^{13}\text{C})$ RMs, individual measurements may be required to establish the correlation between laboratory results. A comprehensive uncertainty budget should be reported with the results.

2.4 Schedule

Due to the COVID-19 pandemic, the original proposed schedule was altered (Table 2). In particular, the deadline to report results was extended for 3 additional months, and the results of the study were discussed over a series of three meetings in September and October 2020.

Table 2. Schedule for CCQM-K167.

Event	Original Deadline	Altered Deadline
Call for participation	October 2019	n/a
Registration	November 26, 2019	n/a
Ship materials	December 1, 2019	December 9, 2019; all samples received by December 31, 2019
Report results	March 31, 2020	June 30, 2020
Distribute draft report	April 15, 2020	July 7, 2020
Discussion of results	IRWG meeting April 2020	Series of 3 meetings: September 16, 2020 October 1, 2020 October 15, 2020
Distribute Draft A	May 30, 2020	December 14, 2020

2.5 Participating Institutes

In total, 8 institutes participated in CCQM-K167 (Table 3). Compared to CCQM-K140, there was an increase from 5 to 8 institutes participating in this key comparison.

Table 3. Participants in the key comparison study, CCQM-K167.

Institute	Abbreviation	Country	Contact
National Institute of Metrology, Quality and Technology	INMETRO	Brazil	Bruno Garrido
Jožef Stefan Institute	JSI	Slovenia	Nives Ogrinc
LGC Ltd	LGC	UK	Philip Dunn
National Institute of Metrology China	NIM	China	Lu Hai
National Measurement Institute Australia	NMIA	Australia	Fong Liu
National Research Council Canada	NRC	Canada	Michelle Chartrand
TÜBİTAK UME	UME	Turkey	Adnan Şimşek
D.I.Mendeleev Institute for Metrology	VNIIM	Russia	Ian Chubchenko

3.0 Reported $\delta_{VPDB}(^{13}C)$ Measurement Parameters

Participants could choose any suitable method for $\delta_{VPDB}(^{13}C)$ measurements. The following information was requested from participants:

- 1) Analysis technique (e.g. EA-IRMS, CRDS, off-line) and instrumentation used (brand and manufacturer)
- 2) Metrics employed to ensure optimal instrument performance
- 3) Measurement sequence (e.g. frequency of calibrant measurements, etc.)
- 4) Correction methods applied such as ^{17}O , blank, drift, etc., for each analysis

3.1 Analysis Technique and Instrumentation used

VNIIM measured $\delta_{VPDB}(^{13}C)$ values of vanillin using a Picarro combustion module with a Picarro G2131i cavity ring down spectrometer, while the remaining seven participants used an elemental analyzer (EA) interfaced with an isotope ratio mass spectrometer. Several types of EAs were used, including a Flash HT 2000, Vario PYRO Cube, and Vario EL III, and various IRMS instruments including Delta Plus XP, Delta V Plus, Delta V Advantage, IsoPrime 100, and MAT 253. No participants chose to analyze vanillin using off-line techniques with dual inlet, or perform absolute carbon measurements with subsequent conversion to isotope delta.

3.2 Metrics to Optimize Performance

Every laboratory has a standard operating procedure to ensure their instrumentation is performing well. These metrics would comprise of any pre-analysis checks to ensure instrument was in acceptable working order, including but not limited to, system background tests, stability tests, and linearity tests [9], and any sample preparation techniques.

Table 4. Typical pre-analysis checks performed by participants prior to $\delta_{VPDB}(^{13}C)$ measurements of vanillin.

Institute	Gas Background	Stability	Linearity using CO ₂ working gas
INMETRO	yes	yes	no
JSI	no	yes	yes
LGC	yes	yes	yes*
NIM	yes	yes	no
NMIA	yes	yes	yes*
NRC	yes	yes	yes
UME	yes	yes	no
VNIIM	yes	yes	yes

*LGC and NMIA assessed linearity using various masses of samples in addition to the CO₂ working gas

3.2.1 System Background Test

Seven of the participants reported monitoring system background levels prior to analysis (Table 4). The purpose to monitor the amount of gases typically found in air (e.g. CO₂, Ar, O₂, N₂, CO, and H₂O) is twofold. First, it provides a quality control measure for the carrier gas, He, and second, it can be an indicator of a leak. Laboratories who routinely perform these background tests typically see small variations between He cylinders, but

large increases in one or more gases may indicate a He cylinder of inferior quality (if the He tank was recently changed), or a leak in the instrumentation.

3.2.2 Stability Test

A zero enrichment test is a sequence consisting of repeated CO₂ working gas pulses at a constant pressure, and is also called a stability test. This test was performed by all laboratories (Table 4), and is used to determine if the instrument is stable by measuring repeated CO₂ pulses, and determining if the standard deviation on these repeated measurements are within an acceptable parameter.

3.2.3 Linearity Test

For $\delta_{VPDB}^{(13C)}$ measurements, the pressure (or concentration) of CO₂ in the source may affect the $\delta_{VPDB}^{(13C)}$ value of the CO₂ gas. To quantify this phenomenon, a linearity test where the pressure of the CO₂ is incrementally increased over several CO₂ gas pulses, was performed by 5 laboratories (Table 4). If there are significant differences in the peak heights (or peak areas) between the samples, RMs and /or the reference pulse(s) of the CO₂ working gas, then a linearity correction may be applied (See Section 3.5.4). LGC and NMIA also performed a linearity test using increasing amounts of a long-term QC sample, and the RM USGS40, respectively. INMETRO ensured the reference pulse intensity matched the intensity of the samples, thus minimizing any potential linearity effects.

3.2.4 Other Tests

Several other tests to evaluate instrument performance were reported. JSI and LGC assessed the peak center, and LGC and NRC performed leak tightness checks on the EA-IRMS system. NIM reported following the verification guideline (JJF1158-2006) prior to each measurement sequence, which included assessing sensitivity and peak shape, in addition to blank and repeatability evaluations.

On the CM-CRDS system, VNIIM performed an oxygen flow rate test, where the oxygen was measured and adjusted, if necessary, to ensure a consistent oxygen flow rate for all measurement sequences. In addition, VNIIM also performed two tests to verify the combustion efficiency. First, an empty space (i.e. a blank with no tin capsule) was analyzed after a sample to ensure complete combustion of the sample. Second, two analysis sequences were performed using two different reactor temperatures: 980 °C to 1100 °C. The results showed no difference in the CO₂ signals between the different temperatures, showing that the combustion at 980 °C was complete.

3.2.5 Sample Preparation and Analysis Metrics

Several laboratories report sample preparation metrics. To minimize any linearity effects, LGC, NMIA, NRC and VNIIM report target weighing of the samples and RMs to analyze the same amount of carbon in each sample. To minimize any potential carry-over, NMIA analyzed a blank tin capsule between each group of samples. NMIA also reports all standards and samples were prepared and weighed in an identical manner, on the same day as analysis when feasible.

3.3 Measurement Sequence Details

For the measurement sequence, all participants reported performing several determinations of RMs and the vanillin sample within a sequence. Six of the eight participants reported performing multiple sequences over multiple days, with two participants reporting a single sequence (Table 5). The total number of measurements ranging from 10 to 30 reported values, with two participants reporting single values that are averages of several measurements. The range of the average mass of vanillin weighed for $\delta_{VPDB}^{(13C)}$ measurements was 69 μg to 827 μg . Analysis of the reported $\delta_{VPDB}^{(13C)}$ values vs mass of vanillin weighed did not show any obvious trend relating these two parameters.

Table 5. Details of the measurement space for CCQM-K167 $\delta_{VPDB}^{(13C)}$ measurements of vanillin.

Institute	Number of Sequences	Total Number of Reported Measurements	Average mass of Vanillin (μg)
INMETRO	2	12	176
JSI	1	17	470
LGC	5	5*	247
NIM	2	5**	175
NMIA	6	30	672
NRC	2	25	664
UME	1	10	69
VNIIM	9	19	827

* an average $\delta_{VPDB}^{(13C)}$ value was reported per sequence, and a total of 69 individual measurements were considered

** average $\delta_{VPDB}^{(13C)}$ values from several replicate measurements

3.4 Calibration and Traceability to the VPDB Scale

The reported $\delta_{VPDB}^{(13C)}$ value and standard uncertainty for the reference materials used by all participants for vanillin normalization are shown in Table 6. Seven of the eight participants used multi-point calibration to normalize the vanillin sample, and one participant used a single point calibration. Five laboratories reported using the same RM(s) for all measurement sequences. To test the delta scale calibration, VNIIM, LGC and NIM reported using different sets of RMs for normalization of the vanillin sample over multiple sequences.

Table 6. $\delta_{VPDB}(^{13}\text{C})$ values and associated standard uncertainties for reference materials used for normalization of vanillin by CCQM-K167 participants.

RM	$\delta_{VPDB}(^{13}\text{C}),$ ‰	$u,$ ‰	JSI	VNIIM [§]	NRC	LGC [†]	UME	NMIA	INMETRO	NIM [‡]
LGC 1712	+12.55	0.05				•				
USGS66	-0.67	0.04				•				
IAEA-CH-6	-10.45	0.04			•	•	•	•		•
BCR-657	-10.76	0.04								•
USGS65	-20.29	0.04	•		•					
UME 1312	-24.02	0.34		•						
IAEA-CH-3	-24.72	0.04		•						•
USGS40	-26.39	0.04				•				•
IAEA-600	-27.77	0.04		•	•					•
NBS 22	-30.03	0.05			•		•		•	•
IAEA-CH-7	-32.15	0.05		•				•		•
USGS61	-35.05	0.04			•					
USGS64	-40.81	0.04	•			•				
LGC 1711	-42.13	0.13				•				

[†]multi-point normalization with 4 RMs; USGS40 and IAEA-CH-6 were common to all analytical sequences, with either USGS64 and USGS66, or LGC 1711 and LGC 1712 as the remaining 2 RMs

[‡]vanillin was normalized using multipoint (4 RMs and 2 RMs) and single point calibration

[§]vanillin was normalized using 2 RMs, either IAEA-600 and IAEA-CH-3, or IAEA-CH-7 and UME 1312

Not all participants reported the same $\delta_{VPDB}(^{13}\text{C})$ values or uncertainties for common RMs used for vanillin normalization (Table 7). As there are several sources for these values from different suppliers (i.e. websites, certificates, publications), and these sources may not always agree, there can be some confusion regarding the most “up to date” $\delta_{VPDB}(^{13}\text{C})$ value for these materials. As an example, for USGS40, the certificate from USGS reports 0.04 ‰ for the combined uncertainty, while the NIST certificate reports 0.09 ‰ as the expanded uncertainty at 95% confidence. If one assumes that the coverage factor on the expanded uncertainty is 2, then the standard uncertainty would be 0.045 ‰, which rounds to 0.05 ‰. Hence, the recommended practice is to report the $\delta_{VPDB}(^{13}\text{C})$ values and uncertainties used in data normalization so that the reported values can then be recalculated as RM $\delta_{VPDB}(^{13}\text{C})$ values or uncertainties are updated [4, 9].

Table 7. Differences in reported $\delta_{VPDB}(^{13}\text{C})$ values and associated uncertainties for reference materials used for normalization of vanillin in this study.

Reference Material	$\delta_{VPDB}(^{13}\text{C}), \text{‰}$	Std $u, (\text{‰})$	Institutes with different reported values
USGS40	-26.39	0.04	LGC: $u = 0.05$ NIM: $u = 0.09$
IAEA-CH-6	-10.45	0.04	NRC: $u = 0.03$ NIM: $u = 0.07$
IAEA-CH-7	-32.15	0.05	NIM: $u = 0.1$
NBS22	-30.03	0.05	NRC: $u = 0.04$ NIM: $u = 0.09$
UME 1312	-24.02	0.34	VNIIM: $\delta_{VPDB}(^{13}\text{C}) = -24.095$ VNIIM: $u = 0.039$

3.4.1 Quality Control Samples

Quality control samples, or samples with a known $\delta_{VPDB}(^{13}\text{C})$ value, are included in a measurement sequence to verify the accuracy of the calibration, but are not used themselves in the calibration. All participants reported the inclusion of QC materials within their measurement sequences. The QC materials varied between the laboratories. VNIIM and NRC used the internal QC materials acetanilide and nicotinamide, respectively. International RMs were also employed as QC samples: IAEA-600 (JSI, NIM, and NMIA), and NMIA also included USGS24, USGS40, USGS73, NBS 22 and IAEA-601. LGC was the only participant to include two exact matrix matched (vanillin) QC materials.

3.5 Post-Measurement Analysis – Applied Corrections

Post-analysis data treatment may include the following corrections: ^{17}O correction, blank, instrument drift, linearity and memory [9, 18]. The correction procedures evaluated and applied are summarized in Table 8.

Table 8. Post-analysis treatment of $\delta_{VPDB}(^{13}\text{C})$ sequences. NA is the correction was not assessed.

Institute	^{17}O Correction	Blank Correction	Drift Correction	Linearity Correction	Memory Correction
INMETRO	IUPAC	no	no	no	no
JSI	Craig	no	no	no	no
LGC	IUPAC	yes	no	no	no
NIM	IUPAC	yes	yes	NA	NA
NMIA	IUPAC	no	no	no	no
NRC	IUPAC	no	yes/no*	no	NA
UME	SSH	yes	NA	NA	NA
VNIIM	NA	no	yes	yes	no

*one of the two sequences was drift corrected

3.5.1 ¹⁷O Correction

For $\delta_{VPDB}^{(13C)}$ measurements, the mass to charge ratios $m/z = 44, 45, \text{ and } 46$ are measured from CO_2^+ , corresponding to isotopologues $^{12}\text{C}^{16}\text{O}^{16}\text{O}$, $^{13}\text{C}^{16}\text{O}^{16}\text{O} + ^{12}\text{C}^{17}\text{O}^{16}\text{O}$, and predominately (99.8 %) $^{12}\text{C}^{18}\text{O}^{16}\text{O}$, respectively. To account for the $^{12}\text{C}^{17}\text{O}^{16}\text{O}$ contribution to the m/z 45 signal, a ^{17}O correction is performed on the data. This isobaric interference is well-documented e.g. [17, 19, 20]. Several methods have been developed to account for the ^{17}O contribution, and it is beyond the scope of this report to go into detail regarding the approach taken for each method. In the instructions to participants, it was stated that the ^{17}O correction should be done using the procedure recommended by IUPAC [17]. Five of the eight participating laboratories reported using the IUPAC recommended procedure; JSI applied the Craig correction [19], and UME the Santrock, Studley, and Hayes correction (SSH)[20]. VNIIM performed measurements using CRDS, and a ^{17}O correction is not applicable for this technique (Table 8).

3.5.2 Blank Correction

For EA-IRMS and CM-CRDS analysis, samples are typically weighed into tin capsules, and then introduced into the instrumentation for analysis. NIM reported using silver capsules rather than tin. To account for any contribution to the CO_2 signal from the capsule itself, a blank correction may be applied to the data. This typically involves employing a mass balance equation, where the contribution of the blank and sample are quantified as an amount (either via peak amplitude, or peak areas) with a corresponding $\delta_{VPDB}^{(13C)}$ value (the un-normalized isotope delta value, also known as the raw value). Blank corrections were evaluated by all participants, and deemed necessary to apply by three participants (Table 8).

3.5.3 Drift Correction

Drift corrections may need to be applied if the measured signal on either the IRMS or the CRDS changes over time in a systematic way. Drift is typically evaluated by measuring a sample with a known $\delta_{VPDB}^{(13C)}$ value, usually a RM, over time. Frequently, drift corrections involve linear models, although polynomial or spline-based drift corrections may also be applied. Seven participants evaluated their sequences for drift, and drift corrections were applied by three participants (Table 8).

3.5.4 Linearity Correction

For $\delta_{VPDB}^{(13C)}$ measurements, the pressure (or concentration) of CO_2 in the source may affect the $\delta_{VPDB}^{(13C)}$ value of the CO_2 gas. Linearity of the instrument is typically assessed prior to the commencement of a measurement sequence. If the difference in peak heights (or peak areas) of the samples and RMs are substantially different between each other, or between the reference pulse(s) of the CO_2 working gas, then a linearity correction may be applied. For instances where the mass of carbon in each tin capsule is well controlled and nearly identical, then a linearity correction is likely unnecessary. Linearity was assessed by six participants, and only VNIIM using CRDS applied a linearity correction to their measurement sequences.

3.5.5 Memory (or Carry-Over) Correction

In cases where quantitative conversion of the sample to CO₂ was not achieved, there may be some carry-over, or memory, of the previous sample contributing to the $\delta_{VPDB}^{(13C)}$ measurement of a subsequent sample, and a mass balance approach may correct for this carry-over. Carry-over was assessed by five participant laboratories, but a correction was deemed unnecessary (Table 8).

4.0 Reported $\delta_{VPDB}^{(13C)}$ Vanillin Results and Uncertainties

The reported $\delta_{VPDB}^{(13C)}$ value and uncertainty for the vanillin sample from participants in CCQM-K167 are reported in Table 9, and represented graphically in Figure 5.

Table 9. CCQM-K167 reported $\delta_{VPDB}^{(13C)}$ values and uncertainties for vanillin.

Institute	$\delta_{VPDB}^{(13C)}$, ‰	Standard Uncertainty (‰)	Expanded Uncertainty (‰)	Coverage Factor, <i>k</i>
INMETRO	-25.96	0.056	0.11	2
JSI	-25.87	0.06	0.11	2
NRC	-25.86	0.03	0.06	2
NMIA	-25.833	0.041	0.083	2.01
UME	-25.82	0.08	0.16	2
NIM	-25.812	0.048	0.096	2
VNIIM	-25.81	0.04	0.09	2
LGC	-25.72	0.05	0.11	2.00*

*rounded from 1.9955

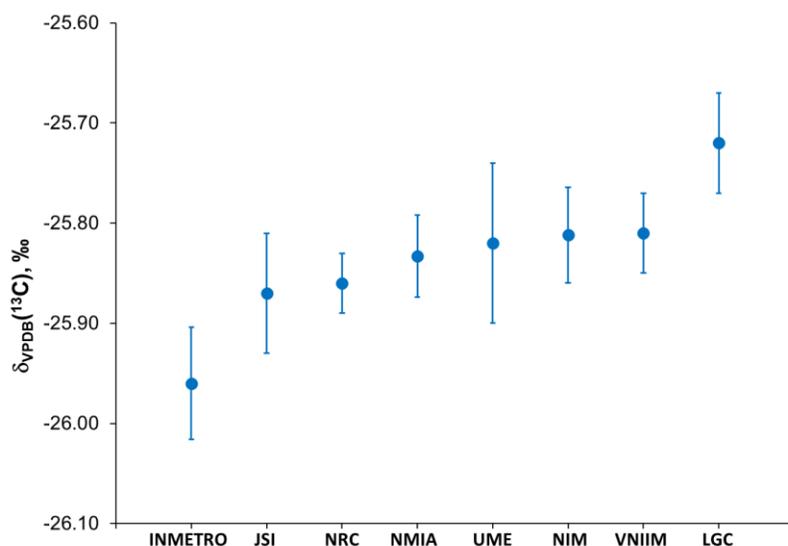


Figure 5. Reported vanillin $\delta_{VPDB}^{(13C)}$ values and standard uncertainties (solid error bars) for participants in CCQM-K167.

NMIA and NIM both reported $\delta_{VPDB}(^{13}\text{C})$ values to three decimal places, with the remaining participants reporting to two decimal places. The total range of reported $\delta_{VPDB}(^{13}\text{C})$ values for vanillin was small, only 0.24 ‰ separated the most positive and most negative reported result. The reported standard uncertainties ranged from 0.03 ‰ to 0.08 ‰. Each laboratory reported a different method for determining the uncertainty on their vanillin $\delta_{VPDB}(^{13}\text{C})$ measurements, and all appeared to be reasonable for this analysis.

4.1 Uncertainty budget

In addition to reporting the uncertainty of their measurements, participants were requested to provide an uncertainty budget (Table 10), which estimates or quantifies the uncertainty contribution from several sources to the final reported uncertainty.

There were several factors that were common to most of the uncertainty budgets: the repeatability (SD) of the sample measurements, the repeatability (SD) of the RMs measurements used for normalization, and the associated uncertainties on the $\delta_{VPDB}(^{13}\text{C})$ values of those RMs. All laboratories evaluated the uncertainty contributions, and the largest factor was the uncertainty of the RMs themselves, followed by the repeatability of the sample and/or RM measurements. Other factors that were considered include the uncertainty due to scale calibration, working gas (WG) peak variation for sample and RM measurements, blank measurement values, coherence of the RMs, and ^{17}O correction methods.

Table 10. Reported uncertainty budgets for vanillin $\delta_{VPDB}(^{13}\text{C})$ measurements. SD (standard deviation); WG (working gas).

Institute	Uncertainty Component	Contribution	% Variance Contribution
INMETRO	Certificate u for NBS 22		46.7
	Repeatability of sample		26.1
	^{17}O correction – sample		9.3
	^{17}O correction – RM		6.9
	Repeatability of NBS 22 (normalization)		3.7
	WG – sample		3.7
	WG – RM		3.6
LGC	Assigned values of normalization RMs		45
	SD of mean across all replicates		29
	WG peak variation for sample (vanillin)		22
	WG peak variation for normalization RMs		4
	Isotopic composition of WG		<1
	Absolute ratios of VPDB and value for λ		<1
	WG peak variation for blank		<1
UME	Certificate u for USGS40 (WG calibration)	0.045	32 [†]
	Certificate u for NBS 22 (normalization)	0.045	32 [†]
	Certificate u for sucrose (normalization)	0.035	20 [†]
	Repeatability of NBS 22 (normalization)	0.022	8 [†]
	Repeatability of sample	0.018	5 [†]
	Repeatability of sucrose (normalization)	0.012	2 [†]
	Repeatability of WG	0.005	<1 [†]
VNIIM	Certified values of RMs	0.04	90 [†]
	Reproducibility (SD of the mean)	0.01	5 [†]
	Measured values of blank	0.01	5 [†]
NMIA	Method precision		1.0
	u from calibration standards		82.5
	u relating to measurement reproducibility		6.0
	u from scale calibration (including ^{17}O correction)		10.4
JSI	u from calibration curve (normalization)		93
	u from repeatability		5
	Certificate u of RMs for normalization		2
NIM	SD on 5 individual measurements	0.0196	16 [†]
	Certificate u of USGS40 (closest value to the final value of the sample)	0.045	84 [†]
	u from linearity of the calibration curve	0.00	0 [†]
NRC	u of RMs used for normalization		~90
	Coherence of RMs		<1
	Repeatability of RM measurements		<5
	Repeatability of sample RM measurements		<5

[†]% contribution was calculated by NRC using the participant's reported contributions.

5.0 KCRV Calculation and Associated Uncertainty

5.1 KCRV – Choosing a statistical model

The key comparison reference value (KCRV) is determined from the reported CCQM-K167 $\delta_{VPDB}(^{13}\text{C})$ values and uncertainties, and may be calculated using a variety of statistical models. Among several common models, the arithmetic mean utilizes the reported $\delta_{VPDB}(^{13}\text{C})$ values, but ignores the reported uncertainties. In contrast, the uncertainty weighted mean trusts the reported uncertainties while ignoring the observed dispersion between the laboratory results. Random effects models (REM) incorporate all the provided information while also allowing for the uncertainty component that models the differences between the laboratory results, otherwise known as the dark uncertainty.

To evaluate which REM model to use, the decision guide outlined in document Decision Guide for selecting CCQM Model [21] from the CCQM organic analysis working group (OAWG) was followed. The following flowchart (Figure 6) provides guidance on which model would be best suited to the dataset at hand. There has been interest expressed in the IRWG to develop a decision flowchart similar to the one used here from the OAWG, and considering the decision flowchart from the inorganic analysis working group (IAWG) [22], to be used in future IRWG key comparisons or pilot studies.

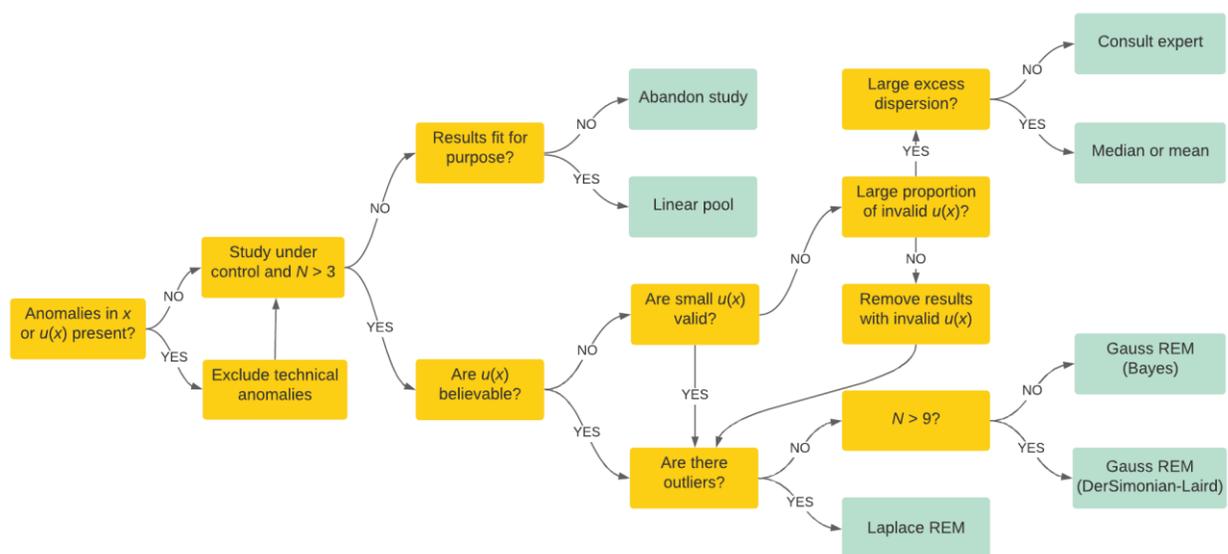


Figure 6. Flowchart based on the decision flowchart from the OAWG working group to adopt an appropriate model to determine the KCRV for CCQM-K167.

The rationale for following the flowchart is outlined below.

1. Are there any anomalies in x or $u(x)$ present? NO

In our opinion, there were no issues with the reported $\delta_{VPDB}(^{13}\text{C})$ values and uncertainties of vanillin, and there were no technical outliers in the data set. The methods reported by

participants seemed valid, and no participants requested to withdraw their data on a technical basis.

2. Is the study under control and $N > 3$? YES

In our opinion, there were no technical conditions that might undermine the study, and no issues with the sample were reported by the participants. The homogeneity and stability of the vanillin were both deemed fit for purpose, and there were no significant differences between the reported results. Also, N is ≥ 4 , as there were 8 participants.

3. Are $u(x)$ believable? YES

Reported uncertainties ranged from 0.03 ‰ to 0.08 ‰, which is reasonable for $\delta_{VPDB}(^{13}\text{C})$ measurements. However, not all participants calculated uncertainty using the same methods, which may account for part of the difference in reported uncertainties between laboratories.

4. Are there outliers? NO

A Student's t -test was performed, and at the 99 % confidence level (as per document CCQM 13-22 [23]), there were no statistical outliers in the CCQM-K167 data set.

5. Is $N > 9$? NO

There were 8 participants.

This pathway in the flowchart leads to the suggestion to employ Gaussian REM (Bayesian), and compare it to the results of the DerSimonian-Laird method. Fitting the REM to the results using Bayesian method or the DerSimonian-Laird method provided nearly identical results for the KCRV and its associated uncertainty.

5.2 Considerations in the development of the Bayesian REM to determine the KCRV and associated uncertainty

5.2.1 Inclusion of Uncertainty due to Homogeneity

Although the uncertainty due to bottle-to-bottle homogeneity was small, it was agreed by all participants to include this source of uncertainty in the overall calculation of the KCRV. There is no specific guidance provided in document CCQM13-22 [23] on how to incorporate uncertainty due to bottle-to-bottle homogeneity into the overall uncertainty on the KCRV.

A naïve approach would be to add the homogeneity uncertainty component to the consensus value,

$$u(\text{KCRV}) = \sqrt{u(\text{consensus})^2 + u(\text{hom})^2},$$

or add it to each of the participant's reported uncertainty,

$$u(\text{participant}) = \sqrt{u(\text{reported})^2 + u(\text{hom})^2}$$

prior to fitting the results to the random effects model. As the homogeneity component affects the true value of isotope delta in the materials analyzed by each laboratory, it is more appropriate, albeit more complex, to incorporate the effect of homogeneity into the overall statistical model. We consider the following observation equation:

$$x_i \sim \text{normal}(\text{mean} = \mu, \text{var} = u^2_{\text{lab}} + u^2_{\text{hom}} + u^2(x_i))$$

where the reported laboratory result ($i = 1 \dots 8$) is affected by three sources of uncertainty: due to measurement, due to homogeneity of the sample, and due to random laboratory effects whose magnitude can only be uncovered from the interlaboratory study. Keeping with the above likelihood function, the following Bayesian statistical model was adopted:

Likelihood: $x_i \sim \text{normal}(\text{mean}=L_i, \text{sd}=u(x_i))$
 Random laboratory effect: $L_i \sim \text{normal}(\text{mean}=\mu, \text{sd}=u_{\text{total}})$

Prior distributions on the model parameters:

KCRV: $\mu \sim \text{normal}(\text{mean}=-25.0, \text{sd}=25.0)$
 Homogeneity effect: $u_{\text{hom}} \sim \text{gamma}(\text{shape}=1.26, \text{rate}=236)$ with $u_{\text{hom}} > 0$
 Total random effect: $u_{\text{total}} \sim \text{hcauchy}(\text{mean}=0.0, \text{scale}=1.0)$ with $u_{\text{total}} > u_{\text{hom}}$

The model was fit to data with Markov-chain Monte Carlo using JAGS [24] and STAN [25] which are implemented in packages rjags and rstan in the R software.

5.2.2 Correlations Between the Reported Results

Several of the participants used common calibrants (Table 6) to measure the $\delta_{\text{VPDB}}(^{13}\text{C})$ of vanillin. As the uncertainty of the reference material may be a substantial constituent in the overall uncertainty, as evidenced in the reported uncertainty budgets for several participants, the correlations between the reported $\delta_{\text{VPDB}}(^{13}\text{C})$ vanillin measurement results have to be evaluated [26].

We have evaluated the magnitude of correlations between the reported results by recreating the calibration plots using Monte Carlo simulations as described in Chartrand *et. al.* [12]. From the choice of calibrators by the laboratories, we expect the highest correlations to be between the results of NRC, UME, NIM, and INMETRO, as they all involve NBS 22. Indeed, we observe this, in general, with the highest correlation between the UME and INMETRO results (Figure 7).

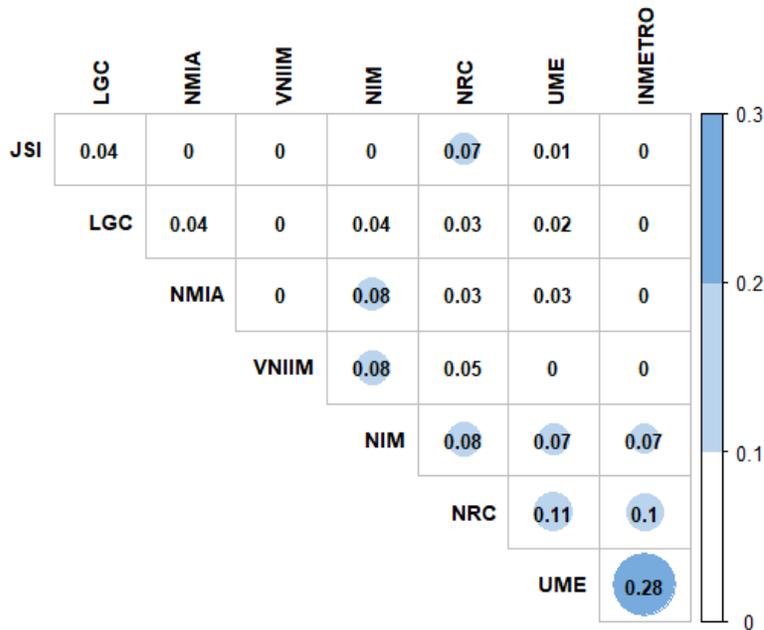


Figure 7. The estimated magnitude of correlation between the reported results for CCQM-K167.

A multivariate model can be used to account for correlations between the reported results by appropriately modifying the likelihood function:

Likelihood:
$$x_i \sim \text{multi_normal}(\text{mean}=L_i, \text{cov}=\text{cov}(x_i))$$

Here, the covariance matrix of the reported results is obtained by combining the reported uncertainties with the estimated correlations. Typically, positive correlations between the participant results will result in increased uncertainty of the consensus value. However, given the low magnitude of the correlations in this study, the effect of the correlations is insignificant on the consensus value:

Univariate model (correlations ignored): $\delta_{KCRV} = -25.834(26) \text{ ‰}$
 Multivariate model (correlations included): $\delta_{KCRV} = -25.833(28) \text{ ‰}$

5.2.3 Correlations Between the Reference Materials

Similar to the correlations between the reported results, the reference materials themselves can be viewed as related. This is because the value assignment for most international reference materials includes a small set of shared high-quality reference materials such as the NBS 22 or IAEA-CH-6 [27].

The extent of correlations between the RMs is not known and is arduous to evaluate retrospectively. Consider a scenario whereby all the reference materials used in this study share half of their respective uncertainties with a common source (Scenario 3, Table 11). This, in turn, results in an increased correlation between the reported values, with an average correlation coefficient of 0.20 and a slightly increased uncertainty of the resulting consensus value.

Table 11. The effect of correlations on the uncertainty of the KCRV.

Scenario	Correlations applied	$u(\delta_{\text{KCRV}})$, ‰
1	- No correlation between laboratories considered - No correlation between RMs	0.026
2	- Correlation between laboratories considered - No correlation between RMs	0.028
3	- Correlation between laboratories considered - 25 % correlation between all RMs (likely)	0.032
4	- Correlation between laboratories considered - 50 % correlation between all RMs (unlikely)	0.034
5	- Correlation between laboratories considered - 75 % correlation between all RMs (unrealistic)	0.037

The effect of correlations between the RMs used on the consensus value and its uncertainty was explored by entertaining scenarios from no correlations at all up to large correlations (Table 11). As the extent of correlations between the RMs increased, the uncertainty on the KCRV also increased, by ~ 40 % from no correlations to large correlations.

While the effect of correlations between the RMs used in this study cannot be dismissed in principle, imposing a magnitude of correlations that could be seen as reasonable (Scenario 3) leads to marginal increase in the uncertainty of the consensus value. Hence, we defer to the Scenario 2 until the topic of RMs correlations is better understood.

5.3 Calculation of the KCRV

In collaboration with Blaza Toman (NIST), a multivariate Bayesian REM was developed to include correlations between the reported $\delta_{\text{VPDB}}(^{13}\text{C})$ measurements and the uncertainty due to homogeneity of the material to determine the KCRV and its associated uncertainty. The statistical model was implemented in BUGS, JAGS, and STAN and below is the STAN model:

```

data {
  vector[8] delta;          // isotope delta values
  matrix[8,8] delta_cov;   // covariance matrix
}
parameters {
  real mu;                  // KCRV
  real<lower=0> u_hom;      // Uncertainty due to homogeneity
  real<lower=u_hom> tau;    // Combined (total) dark uncertainty
  vector[8] L;             // Laboratory effects
}
model {
  // priors on model parameters

```

```

mu ~ normal(-25.0, 25.0);
u_hom ~ gamma(1.26, 236.0); // Data derived from homogeneity study
tau ~ cauchy(0.0, 1.0);
// likelihood
delta ~ multi_normal(L, delta_cov);
L ~ normal(mu, tau);
}
generated quantities{
vector[8] delta_p;
vector[8] DoE;
// Prediction of the reported values
delta_p = multi_normal_rng(rep_vector(mu, 8), delta_cov + diag_matrix(rep_vector(tau^2, 8)) );
// Degrees-of-equivalence
DoE = delta - delta_p;
}

```

The covariance matrix of the reported results was constructed from the reported uncertainties and the estimated correlations. The KCRV and associated uncertainty using the Bayesian REM are presented in Table 12, and graphically using the Bayesian REM in Figure 8.

Table 12. KCRV and associated uncertainty determined using Bayesian REM (STAN).

δ_{KCRV} , ‰	Standard uncertainty, ‰	Expanded uncertainty, ‰	Coverage factor, <i>k</i>
-25.833	0.028	0.056	2

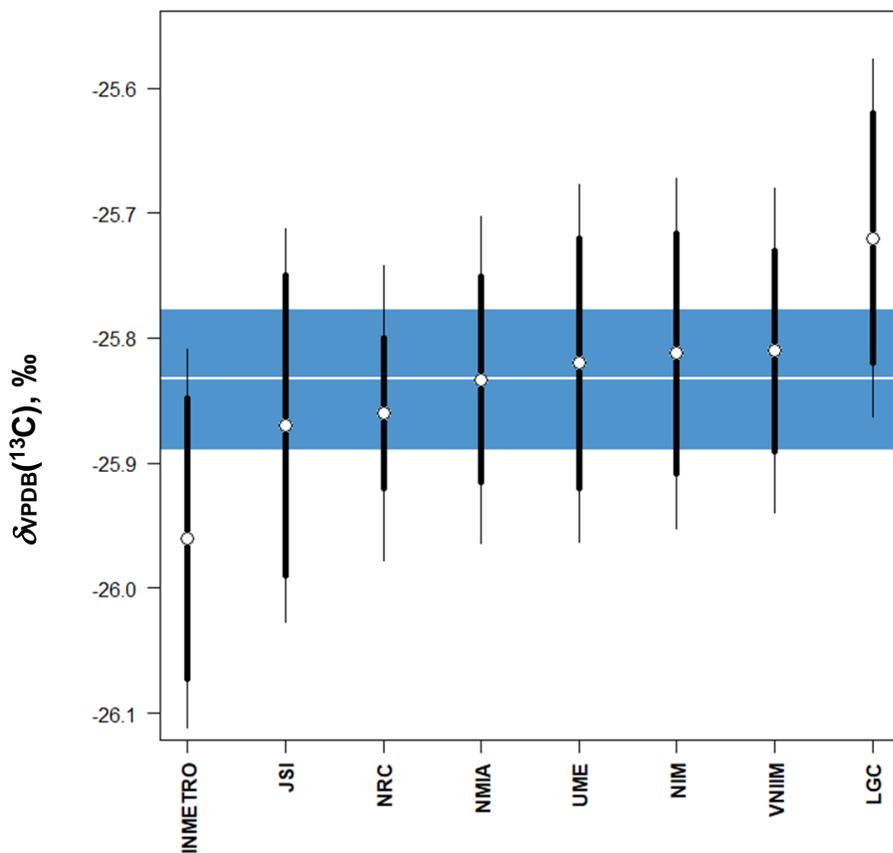


Figure 8. KCRV and associated uncertainty determined using Bayesian REM. The white horizontal line is the KCRV (-25.833 ‰), and the blue box in the expanded uncertainty (0.056 ‰, $k=2$). White dots and thick error bars are the participants' reported $\delta_{VPDB}(^{13}C)$ vanillin measurement results and expanded uncertainties, respectively, and the thin error bars are the expanded uncertainty including the dark uncertainty (0.051 ‰).

The dark uncertainty, or the component of the uncertainty arising due to the spread of results in this data set, was determined to be 0.051 ‰ for $k=1$. This value is larger than the majority of the individual laboratories' reported standard uncertainties, and can suggest the laboratories have underestimated their reported uncertainties. The addition of the dark uncertainty to each laboratory's reported uncertainty increases that uncertainty between 19 % and 97 % (Table 13).

$$u(\text{total}) = \sqrt{u(x)^2 + u(\text{dark})^2}$$

Table 13. CCQM-K167 reported standard uncertainties including the dark uncertainty.

Institute	$u(x)$, ‰	$u(\text{dark})$, ‰	$u(\text{total})$, ‰	% increase in $u(x)$
INMETRO	0.056	0.051	0.076	35
JSI	0.06	0.051	0.079	31
NRC	0.03	0.051	0.059	97
NMIA	0.041	0.051	0.065	60
UME	0.08	0.051	0.095	19
NIM	0.048	0.051	0.070	46
VNIIM	0.04	0.051	0.065	62
LGC	0.05	0.051	0.071	43

5.4 Degrees of Equivalence

For each participant, the degree-of-equivalence, DoE, is the difference between the reported value and the value predicted by the consensus model for that laboratory:

$$\text{DoE}[i] = \delta_{\text{reported}}[i] - \delta_{\text{predicted}}[i]$$

The uncertainty of the DoE for each laboratory is obtained from fitting the random effects model to the results using the Bayesian method which are presented in Table 14 and Figure 9. The $\delta_{VPDB}(^{13}\text{C})$ results are considered to be equivalent to the KCRV if the expanded uncertainty interval ($U_{95}\text{DoE}$) encompasses zero (Figure 9). For CCQM-K167, all reported results are considered equivalent to the KCRV.

Table 14. CCQM-K167 degrees of equivalence for $k=1$ (U_{DoE}) and $k=2$ ($U_{95}\text{DoE}$).

Institute	$\delta_{VPDB}(^{13}\text{C})$, ‰	$u(x)$, ‰	$U(x)$, ‰	DoE	U_{DoE}	$U_{95}\text{DoE}$
INMETRO	-25.96	0.056	0.11	-0.127	0.089	0.178
JSI	-25.87	0.06	0.11	-0.037	0.090	0.180
NRC	-25.86	0.03	0.06	-0.027	0.073	0.146
NMIA	-25.833	0.041	0.083	0.000	0.078	0.156
UME	-25.82	0.08	0.16	0.013	0.085	0.170
NIM	-25.812	0.048	0.096	0.021	0.083	0.166
VNIIM	-25.81	0.04	0.09	0.023	0.077	0.154
LGC	-25.72	0.05	0.11	0.113	0.083	0.166

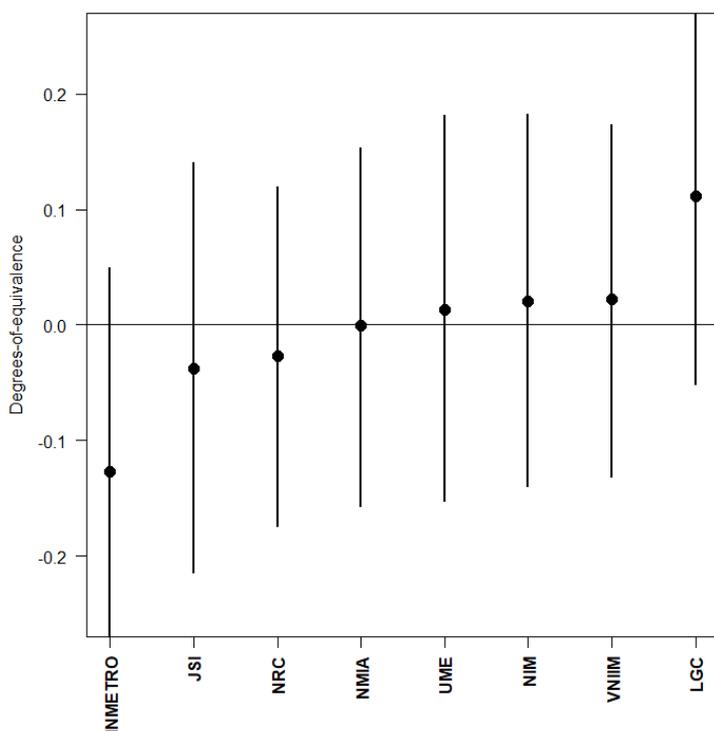


Figure 9. CCQM-K167 degrees of equivalence (DoE). Circles are the calculated degrees of equivalence for each participant, and the error bars are the expanded uncertainty ($k=2$) of the uncertainty on the determined degree of equivalence. Laboratories with DoE and its expanded uncertainty crossing the zero value are considered to be equivalent with the KCRV.

6.0 CMC Claim

Participation in CCQM-K167 will support measurement claims for bulk $\delta_{VPDB}(^{13}\text{C})$ measurements of easily combustible materials within the interval of $\delta_{VPDB}(^{13}\text{C}) \approx -50\text{‰}$ to $+40\text{‰}$, typically realized using a multipoint calibration, and underpinned by internationally recognized reference materials measured on the VPDB scale. The reported $\delta_{VPDB}(^{13}\text{C})$ measurement uncertainty is the best demonstrable uncertainty for the method employed. Difficult-to-combust materials can be supported where quantitative conversion of C to CO_2 is demonstrated, but may result in larger uncertainties.

7.0 References

1. Bomgardner, M. (2016). The Problem with Vanilla. *Chemical & Engineering News* September 14, 2016.
2. Hansen, A.-M. S., Fromberg, A., and Frandsen, H. L. (2014). Authenticity and Traceability of Vanilla Flavors by Analysis of Stable Isotopes of Carbon and Hydrogen. *Journal of Agricultural and Food Chemistry* 62(42): 10326-10331. DOI: 10.1021/jf503055k
3. Coplen, T. B. (2011). Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results. *Rapid Communications in Mass Spectrometry* 25(17): 2538-2560. DOI: 10.1002/rcm.5129
4. Coplen, T. B., Brand, W. A., Gehre, M., Gröning, M., Meijer, H. A. J., Toman, B., and Verkouteren, R. M. (2006). New Guidelines for $\delta^{13}\text{C}$ Measurements. *Analytical Chemistry* 78(7): 2439-2441. DOI: 10.1021/ac052027c
5. Qi, H., Coplen, T. B., Mroczkowski, S. J., Brand, W. A., Brandes, L., Geilmann, H., and Schimmelmann, A. (2016). A new organic reference material, l-glutamic acid, USGS41a, for $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ measurements – a replacement for USGS41. *Rapid Communications in Mass Spectrometry* 30(7): 859-866. DOI: 10.1002/rcm.7510
6. Assonov, S. (2018). Summary and recommendations from the International Atomic Energy Agency Technical Meeting on the Development of Stable Isotope Reference Products (21–25 November 2016). *Rapid Communications in Mass Spectrometry* 32(10): 827-830. DOI: 10.1002/rcm.8102
7. IUPAC (2018). Standard Atomic Weights of 14 Chemical Elements Revised. *Chemistry International* 40(4): 23-24. DOI: 10.1515/ci-2018-0409
8. Brand, W. A., Coplen, T. B., Vogl, J., Rosner, M., and Prohaska, T. (2014). Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). *Pure and Applied Chemistry* 86(3): 425-467. DOI: 10.1515/pac-2013-1023
9. Dunn, P. J. H., and Carter, J. F., eds. (2018). Good practice guide for isotope ratio mass spectrometry, 2nd Edition. FIRMS. ISBN 978-0-948926-33-4.
10. Traceability Exception: Delta Value Isotope Ratio Measurements. In Decision CIPM/104.26. <https://www.bipm.org/utis/common/documents/CIPM-MRA/Traceability-Exception-QM1.pdf>. Accessed on November 25, 2020.
11. Dunn, P. J. H., Goenaga-Infante, H., Goren, A. C., Şimşek, A., Bilsel, M., Ogrinc, N., Armishaw, P., and Hai, L. (2017). CCQM-K140: carbon stable isotope ratio delta values in honey. *Metrologia* 54(1A): 08005-08005. DOI: 10.1088/0026-1394/54/1a/08005

12. Chartrand, M. M. G., Meija, J., Kumkrong, P., and Mester, Z. (2019). Three certified sugar reference materials for carbon isotope delta measurements. *Rapid Communications in Mass Spectrometry* 33(3): 272-280. DOI: 10.1002/rcm.8357
13. International Atomic Energy Agency. NBS 22, IAEA-CH-3, IAEA-CH-6, IAEA-CH-7, USGS24 reference sheet issue date: 3 August 2007.
14. U.S. Geological Survey. (2019). Report of Stable Isotopic Composition Reference Materials USGS64, USGS65, USGS66 (Carbon and Nitrogen Isotopes in Glycine). <https://www.usgs.gov/media/files/rsil-report-stable-isotopic-comp-reference-materials-usgs61-62-63>. Accessed on November 10, 2021.
15. U.S. Geological Survey. (2019). Report of Stable Isotopic Composition Reference Materials USGS61, USGS62 and USGS63 (Hydrogen, Carbon and Nitrogen Isotopes in Caffeine). <https://www.usgs.gov/media/files/rsil-report-stable-isotopic-comp-reference-materials-usgs64-65-66>. Accessed on on November 10, 2021.
16. Coplen, T. B. (1995). Discontinuance of SMOW and PDB. *Nature* 375: 285.
17. Brand, W. A., Assonov, S. S., and Coplen, T. B. (2010). Correction for the 17O interference in $\delta(13C)$ measurements when analyzing CO₂ with stable isotope mass spectrometry (IUPAC Technical Report). *Pure and Applied Chemistry* 82: 1719-1733. DOI: 10.1351/PAC-REP-09-01-05
18. Werner, R. A., and Brand, W. A. (2001). Referencing strategies and techniques in stable isotope ratio analysis. *Rapid Communications in Mass Spectrometry* 15(7): 501-519. DOI: 10.1002/rcm.258
19. Craig, H. (1957). Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide *Geochimica et Cosmochimica Acta* 12: 133-149.
20. Santrock, J., Studley, S. A., and Hayes, J. M. (1985). Isotopic Analyses Based on the Mass Spectrum of Carbon Dioxide. *Analytical Chemistry* 57: 1444-1448.
21. Nelson, M., Meija, J., and Toman, B. (2019). Decision Guide for Selecting CCQM Model 1 key comparison reference value estimation procedures, Version 1:2019-09-23.
22. Possolo, A., Koepke, A., Newton, D., and Winchester, M. R. (2021). Decision Tree for Key Comparisons. *Journal of Research of National Institute of Standards and Technology* 126: 126007. DOI: 10.6028/jres.126.007
23. CCQM (2013). CCQM-1322: Guidance Note: Estimation of a consensus KCRV and associated Degrees of Equivalence. Version 10, Date 2013-04-12.

24. rjags: Bayesian Graphical Models using MCMC. <https://CRAN.R-project.org/package=rjags>.
25. Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Bentacourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). STAN: A probabilistic programming language. *Journal of Statistical Software* 76(1): 35345. DOI: 10.18637/jss.v076.i01
26. Meija, J., and Chartrand, M. M. G. (2017). Uncertainty evaluation in normalization of isotope delta measurements results against international reference materials. *Analytical and Bioanalytical Chemistry* 410(3): 1061-1069. DOI: 10.1007/s00216-017-0659-1
27. Dunn, P. J. H., Malinovsky, D., and Goenaga-Infante, H. (2020). Calibration hierarchies for light element isotope delta reference materials. *Rapid Communications in Mass Spectrometry* 34(9e8711). DOI: 10.1002/rcm.8711

Appendix 1. Registration

Registration Form CCQM-K167/P211: Carbon isotope delta measurements of vanillin

National Metrology Institutes (NMIs) and appropriate Designated Institutes (DIs) in accordance with the CIPM MRA are invited to participate in the Key Comparison.

Participant's Name	
Describe your affiliation (NMI or DI)	
Name of the Institute	
Address	
Country	
E-mail of contact	
Telephone number of contact	
Participate in CCQM-K167?	Yes or No
Participate in CCQM-P211?	Yes or No

Shipping instructions

Please indicate any special instructions (for importation) and the full shipping address and telephone number of a contact.

--

Please send the completed form by e-mail before November 26, 2019 to:

Dr. Michelle Chartrand
E-mail: Michelle.Chartrand@nrc-cnrc.gc.ca

National Research Council Canada
1200 Montreal Rd., Building M-12
Ottawa, ON, K1A 0R6 CANADA
Phone: 1 613 991 4606

Appendix 2. Measurement Protocol Sent to Participants

CCQM-K167/P211: Carbon isotope delta measurements of vanillin

Background

The Key Comparison is coordinated by Michelle Chartrand and Juris Meija at NRC Canada.

Verification of the authenticity of food items is essential to ensure the quality and safety of food products. Carbon isotope delta, $\delta_{VPDB}(^{13}C)$, measurements are routinely used to determine the authenticity of food products by determining the source of the food product, and detecting adulteration of the food products via addition of unreported additives.

The first Key Comparison for $\delta_{VPDB}(^{13}C)$ measurements (CCQM-K140) was conducted in 2015-2016, and honey samples were measured. The October 2018 meeting of the IRWG in Ottawa, Canada, included a need for more key comparisons for light stable isotope measurements.

This proposed Key Comparison will support Calibration and Measurement Capability (CMC) claims for $\delta_{VPDB}(^{13}C)$ measurements, allowing for institutions to demonstrate and improve core capabilities in this area. The goal of this Key Comparison is to establish current best achievable uncertainties for $\delta_{VPDB}(^{13}C)$ measurements.

This Key Comparison will support measurement claims for bulk $\delta_{VPDB}(^{13}C)$ measurements where quantitative conversion of C to CO₂ is demonstrated, within the isotopic range of available reference materials [1]; $\delta_{VPDB}(^{13}C) = -47.32 \text{ ‰ to } +535.3 \text{ ‰}$.

This study was designed using vanillin, a high purity, easily combustible organic material, and it is expected that the $\delta_{VPDB}(^{13}C)$ measurement uncertainty achieved is the best demonstrable uncertainty for the method employed. More complex matrices such as bulk plant materials and materials with high protein content can be supported where quantitative conversion of C to CO₂ is demonstrated. For continuous flow isotope ratio mass spectrometry methods, where there is a high degree of similarity in the instrumental approach, this Key Comparison can also support bulk nitrogen isotope ($\delta_{AIR-N_2}(^{15}N)$) measurements where quantitative conversion of N to N₂ is demonstrated, within the isotopic range of available reference materials [1]; $\delta_{AIR-N_2}(^{15}N) = -30.41 \text{ ‰ to } +375.3 \text{ ‰}$.

Material

Pure vanillin (> 0.99 g/g) was purchased from a major chemical supplier in Canada. The sample was sieved using standard US 30, 40 and 50 mesh sieves, and the portion that fell through the 50 mesh sieve was collected. Aliquots of 0.20 to 0.25 g of the sieved vanillin were portioned into 2 mL glass vials and stored in a dry box at room temperature until distributed.

Each vial of vanillin provided in this study has been evaluated for its $\delta_{VPDB}(^{13}C)$ value at the NRC, and vial-to-vial homogeneity was determined to be fit for purpose.

Measurand

Carbon isotope delta relative to the Vienna Pee Dee Belemnite, $\delta_{VPDB}(^{13}\text{C})$ where VPDB is a virtual material defined by reference material NBS 19 such that $\delta_{VPDB}(^{13}\text{C}, \text{NBS 19}) = +1.95 \text{ ‰}$ exactly. The target value is $\delta_{VPDB}(^{13}\text{C}) = -20 \text{ ‰}$ to -40 ‰ .

Test Sample Receipt / Handling

Samples will be distributed by courier to the participants. Each laboratory participating in the Key Comparison will receive one sample vial of vanillin. Please inform the coordinator immediately if the test sample has been compromised in any way, or if it arrives in questionable condition. It is recommended to keep the vial tightly capped and stored in a cool dry place out of direct sunlight when not in use.

Choice of Method / Procedure

Participants in the Key Comparison are requested to use a suitable method for $\delta_{VPDB}(^{13}\text{C})$ measurements, and are welcome to select any appropriate $\delta_{VPDB}(^{13}\text{C})$ reference material to realize the VPDB scale, or perform absolute carbon isotope ratio measurement and report it on the VPDB scale. A list of certified reference materials that could be used to identify accepted references for delta value isotope ratio traceability statements is published and maintained by IUPAC [1]. The use of LSVEC is not acceptable.

Reporting

Key Comparison results are to be reported as $\delta(^{13}\text{C})$ values relative to the VPDB. We request you to report a single value, standard uncertainty, and expanded uncertainty including the coverage factor. A comprehensive uncertainty budget should be reported with the results. A detailed description of the instrument used, calibration methods including traceability to the VPDB scale, and carbon isotope delta analysis details is requested. The correction of ^{17}O should be done using the procedure recommended by IUPAC [2].

Additionally, we request that at least five (5) replicates NRC Vanillin be measured and reported. As participants may use common $\delta_{VPDB}(^{13}\text{C})$ reference materials, individual measurements are required to establish the correlation between laboratory results [3]. The Key Comparison Reference Value (KCRV) and the associated degrees of equivalence will be assigned from all participating laboratories using the general principles outlined in guidance document CCQM 13-22.

Please complete and submit the report to Michelle Chartrand via email before the deadline.

Time Schedule

Registration deadline: November 26, 2019

Ship materials: December 1, 2019

Deadline for receipt of data: March 31, 2020

Prepare/distribute draft report: April 15, 2020

Discussion of the results and draft report at IRWG meeting: April 18-24, 2020

Draft A final report: May 30, 2020
Draft B final report: September 15, 2020
Finalize report: After October 2020 IRWG meeting

Participation and Registration

National Metrology Institutes or an appropriate Designated Institute in accordance with the CIPM MRA, are invited to participate in the Key Comparison. If you decide to participate in the Key Comparison, please fill in the Registration Form and send to Michelle Chartrand via email.

Contact Information

Dr. Michelle Chartrand
National Research Council Canada
1200 Montreal Rd., Building M-12
Ottawa, ON K1A 0R6, CANADA
Phone: 1-613-991-4606
E-mail: Michelle.Chartrand@nrc-cnrc.gc.ca

References

1. Brand, W. A., Coplen, T. B., Vogl, J., Rosner, M., and Prohaska, T. (2014). Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). *Pure Appl. Chem.* 86: 425-467. DOI: 10.1515/pac-2013-1023
2. Brand, W. A., Assonov, S. S., and Coplen, T. B. (2010). Correction for the ^{17}O interference in $\delta(^{13}\text{C})$ measurements when analyzing CO_2 with stable isotope mass spectrometry (IUPAC Technical Report). *Pure Appl. Chem.* 82: 1719-1733. DOI: 10.1351/PAC-REP-09-01-05
3. Meija, J. and Chartrand, M.M.G. (2017). Uncertainty evaluation in normalization of isotope delta measurements results against international reference materials. *Anal. Bioanal. Chem.* 410(3): 1061-1069. DOI: 10.1007/s00216-017-0659-1

Appendix 3. Submission of Results

CCQM-K167/P211 Results Submission Form

Name of Reporting Scientist	
Institute	
Address	
E-mail	
Are you participating in CCQM-K167?	Yes or No
Are you participating in CCQM-P211?	Yes or No
Analysis Technique (e.g. EA-IRMS, CRDS, off-line)	
Instrumentation Used (Brand and Manufacturer)	

The Reported Result and its Uncertainty

Reported $\delta_{VPDB}(^{13}C)$ (‰)	Standard Uncertainty (‰)	Expanded Uncertainty (‰)	Coverage Factor, k

Details of the Analysis Method

Please describe the correction methods applied such as ^{17}O , blank, drift, etc., for each analysis	
Please describe your measurement sequence (e.g. frequency of calibrant measurements, etc.)	
What metrics do you employ to ensure optimal instrument performance?	

Reference Materials for Calibration

(for establishing the correlation between the Laboratory results)

	Name of Reference Material	Certified $\delta_{VPDB}(^{13}C)$ (‰)	Standard Uncertainty (‰)
RM #1			
RM #2			

insert extra lines as needed

Individual Replicate Measurements (for establishing the correlation between the Laboratory results and assessing the reported measurement uncertainty)

Determination	Sample mass weighed (mg)	Reported $\delta_{VPDB}(^{13}C)$ (‰)	Date of Measurement
1			
2			
3			
4			
5			

insert extra lines as needed

Please explain briefly how these individual measurements were combined to produce the Reported Result and its Uncertainty	
---	--

Uncertainty Budget for the Reported Result

Please provide details of your uncertainty budget for the final reported $\delta_{VPDB}(^{13}\text{C})$ value	
---	--

Additional Comments

Please add any additional comments if necessary	
---	--