Supporting Information - Assessing Arsenic Species in Foods Using Regularized Linear Regression of the Arsenic K-edge X-ray Absorption Near Edge Structure

Evan P. Jahrman, Lee L. Yu, William P. Krekelberg, David A. Sheen, Thomas C. Allison, and John L. Molloy

Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

This document presents micro-X-ray fluorescence (μXRF) measurements consisting of roughly 10,000 spots collected over a 2.2 mm x 7.5 mm region, see Fig. S1. Measurements were performed with an Orbis μXRF analyzer (EDAX Ametek Inc.) equipped with a silicon drift detector possessing an active area of 50 mm^2. A detector live time of 10 seconds per spot resulted in measurement times of several hours per sample. The X-ray tube source contained a Rh anode and was operated with an accelerating potential of 30 kV and a filament current of 1 mA. The incident X-rays were focused to a nominally 30 μm spot size. Results from the polyethylene glycol (PEG) bound pellet are shown relative to a boron nitride (BN) bound pellet. In Fig. S2, the As K-edge XANES spectra are shown for each pentavalent organic compound. Fig. S3 shows cacodylic acid (DMA) in an aqueous solution. For comparison, solid-phase DMA was also measured, and its spectrum is shown in the figure to emphasize the spectral effects induced by interactions with the solvent. The results of the peak fitting procedure are provided in Fig. S4. This figure also shows each component used in the peak fitting algorithm, scaled to its relative contribution to the fitted spectrum. Fig. S5 compares the As K-edge XANES spectrum measured for BCR-627 to three pentavalent organic arsenic compounds with nearby white line positions. Fig. S6 shows the As K-edge XANES spectrum of BCR-627 and a spectrum representing its fit result when including As_Sug in the model's training set. Fig. S7 provides a comparison between the As K-edge XANES spectrum measured for the eluent extracted from SRM 3232 and a spectrum representing its fit result. Finally, Table S1 provides the As mass fractions reported in the certificates for BCR-627, CRM 7405-b, and SRM 3232. Table S2 presents the results of the bootstrap analysis for BCR-627 with As_Sug included in the model's training set. Table S3 lists the arsenic speciation distribution for each extraction by-product. The results are categorized by species class with the values and uncertainties determined by the bootstrap method provided. Likewise, Table S2 lists the arsenic speciation distribution of the eluent from the extraction of SRM 3232, but with the results categorized by species.

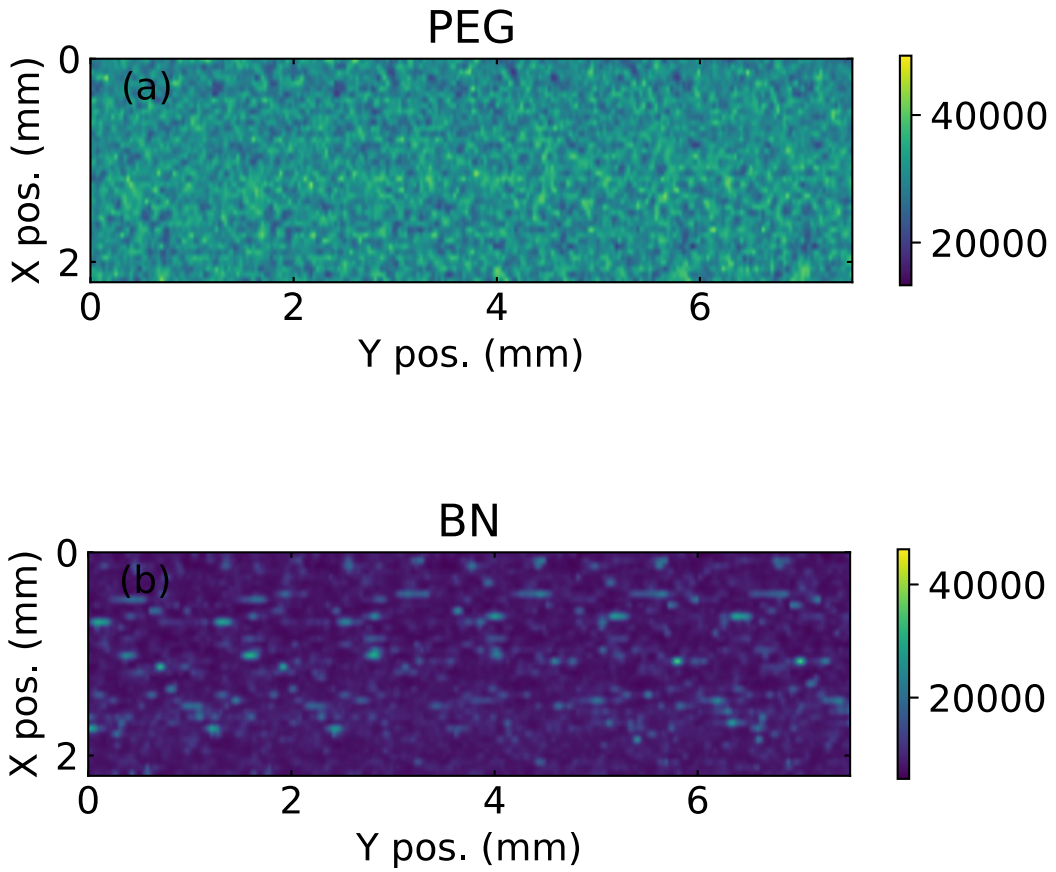

Figure S1: The μXRF maps were used to construct heat maps for DMA pelletized in one of two binders: a) polyethylene glycol (PEG) or b) boron nitride (BN). The intensity corresponds to counts from the As Kα fluorescence lines. Note that both materials contained roughly the same amount of arsenic by mass.

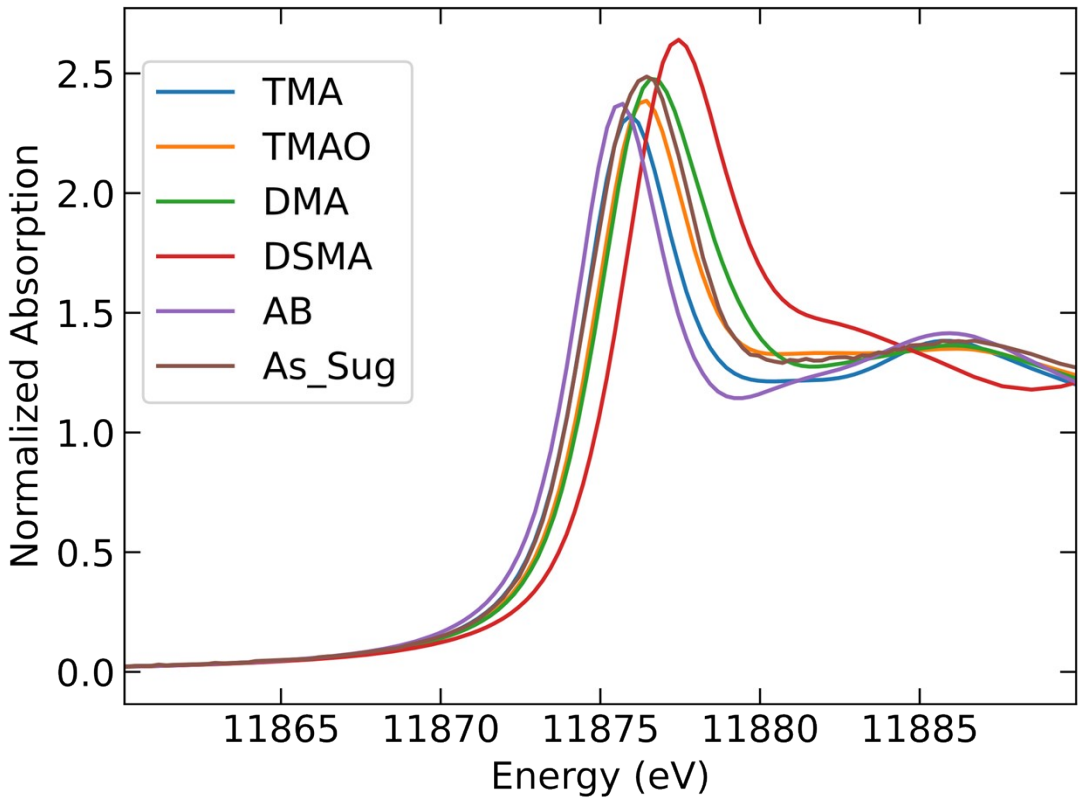

Figure S2: The As K-edge XANES spectra are shown for the series of pentavalent organic arsenic compounds. Specifically, the compounds represented are tetramethylarsonium iodide (TMA), trimethylarsine oxide (TMAO), dimethylarsinic acid (DMA), disodium methyl arsonate (DSMA), arsenobetaine (AB), and arsenosugar from the eluent of SRM 3232 (As_Sug).

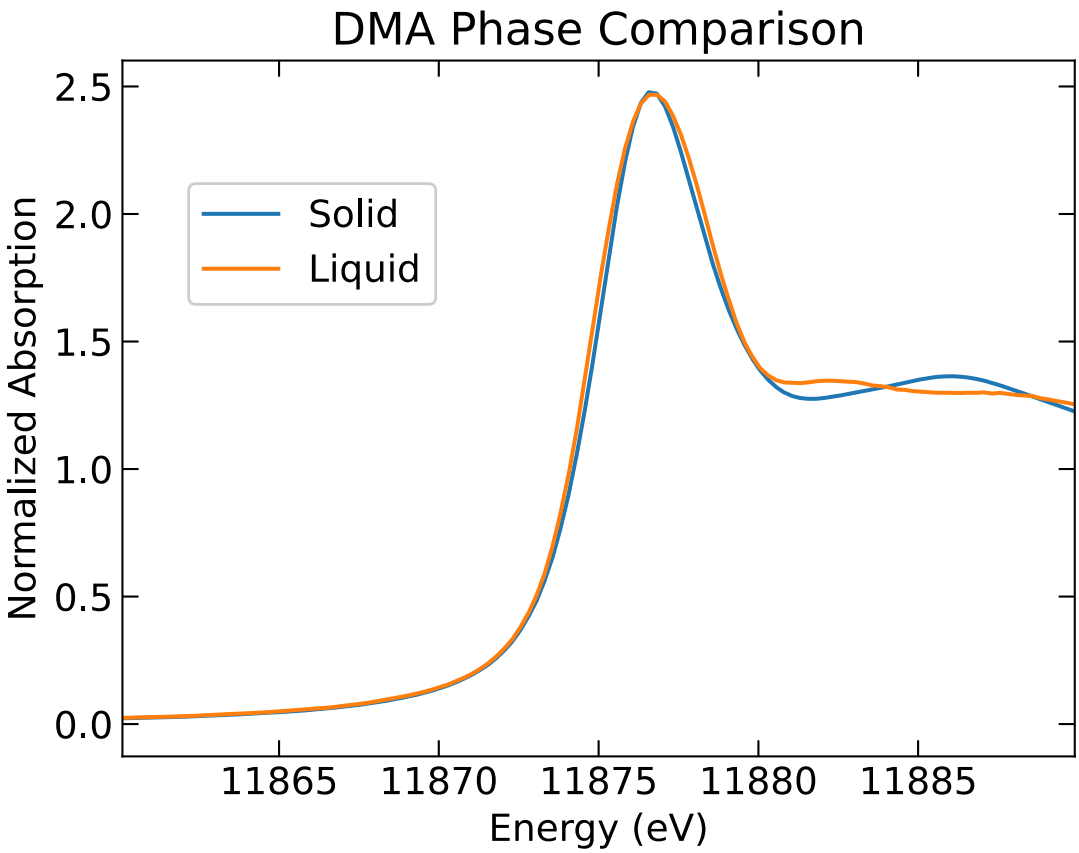

Figure S3: The As K-edge XANES spectra are shown for DMA in the solid-phase and in solution.

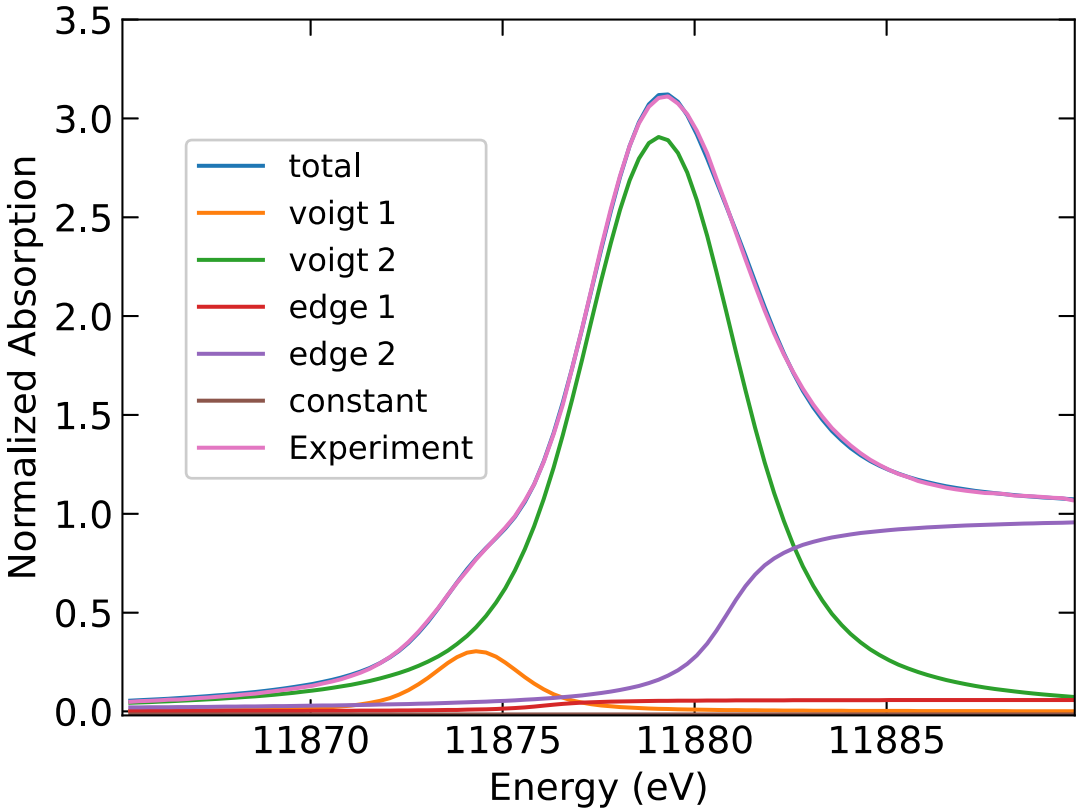

Figure S4: The As K-edge XANES spectrum of the acidic glycerol solution is shown along with the fitted result and its components.

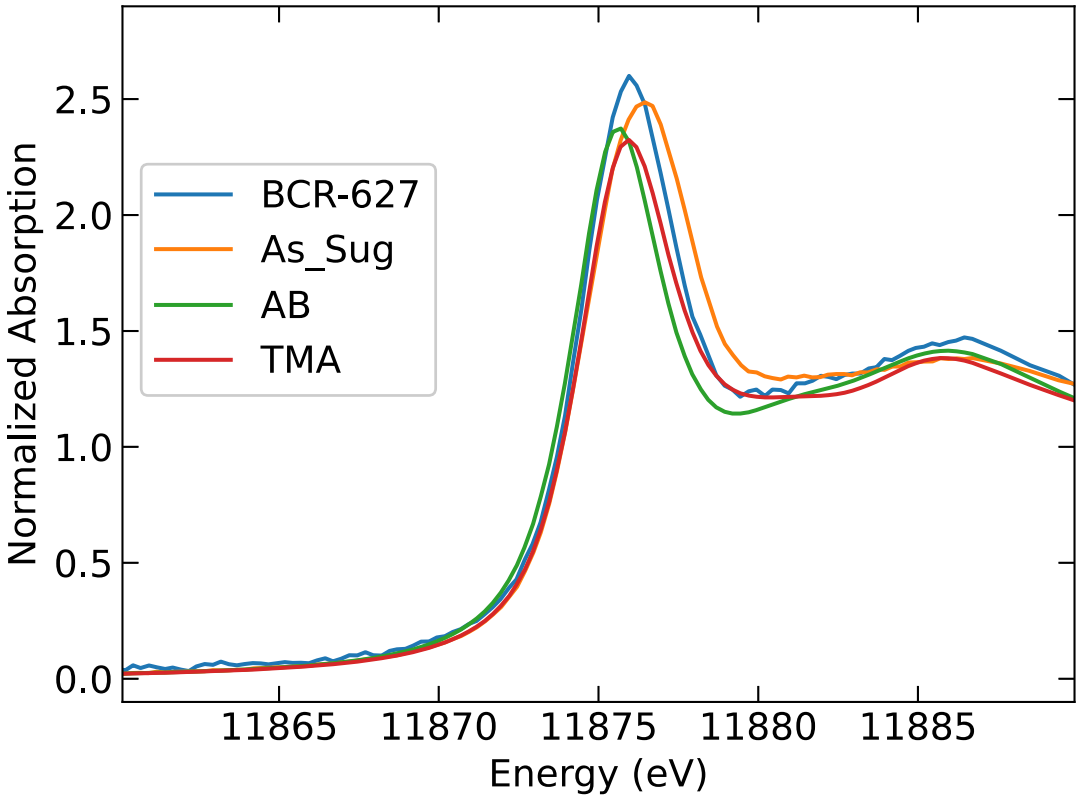

Figure S5: The As K-edge XANES spectrum measured for BCR-627 shown alongside three pentavalent organic arsenic compounds exhibiting white line features nearby.

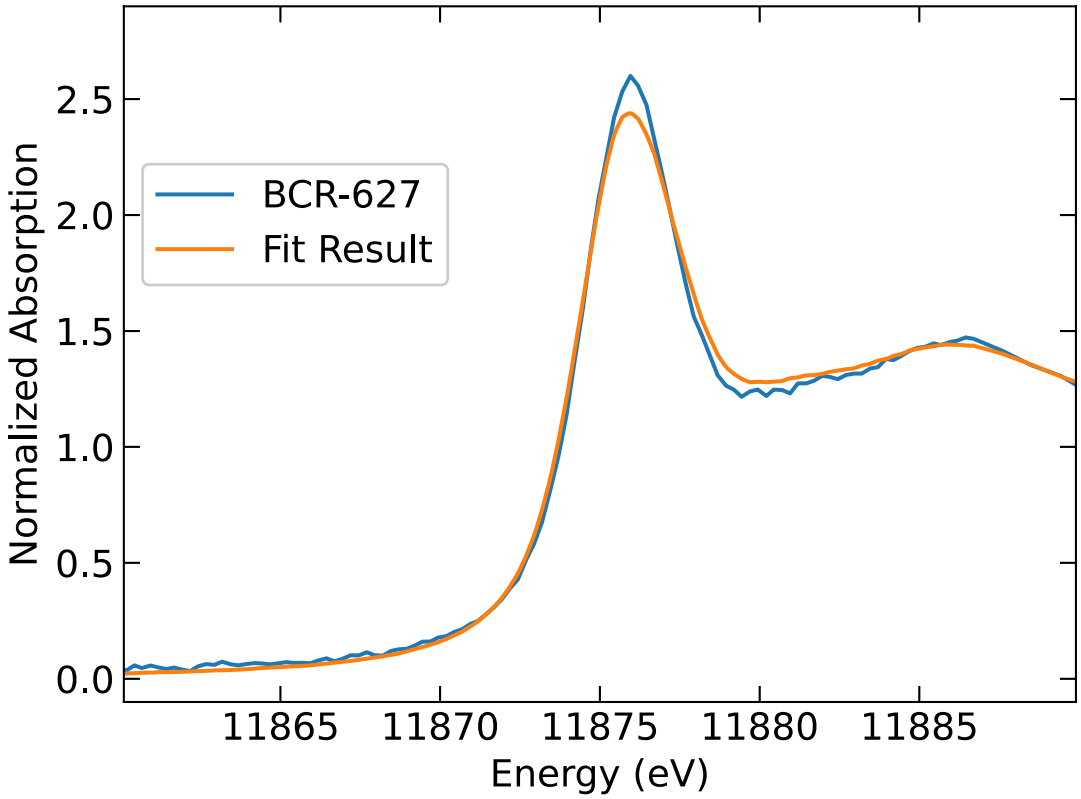

Figure S6: The As K-edge XANES spectrum of BCR-627 shown relative to a synthetic spectrum constructed from the corresponding fit result which included As_Sug in its training set.

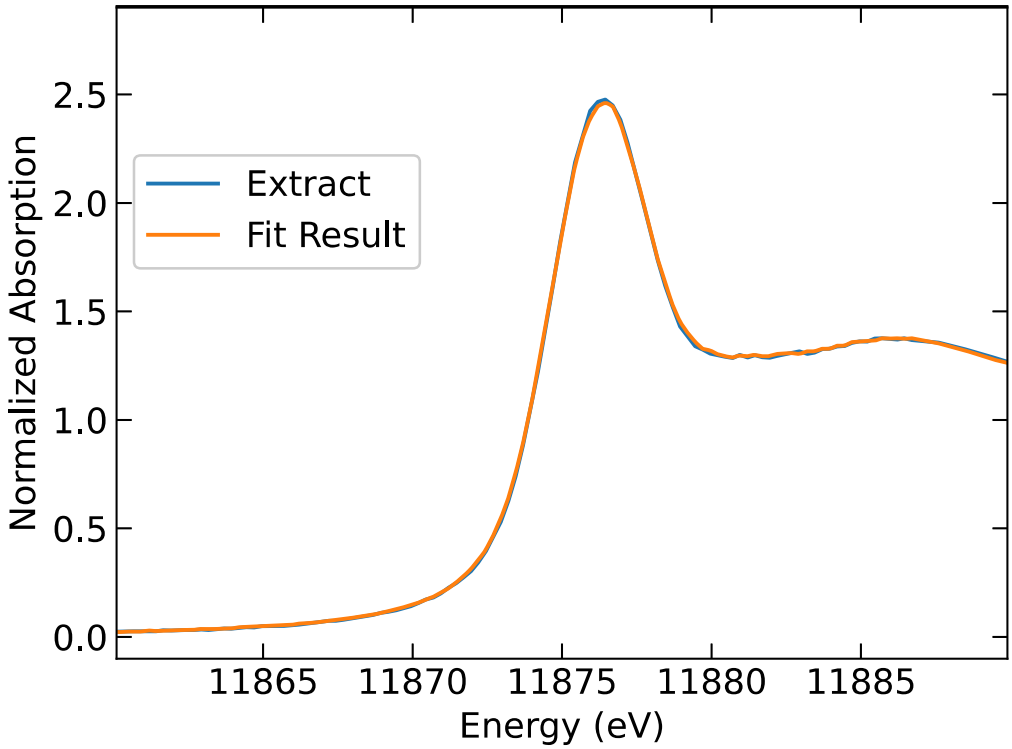

Figure S7: The As K-edge XANES spectrum measured for the eluent extracted from SRM 3232 is shown relative to a synthetic spectrum constructed from the corresponding fit result.

Table S1: Arsenic mass fractions reported in the certificates for BCR-627, CRM 7405-b, and SRM 3232. In all units, the denominator refers to kg of sample mass and the numerator refers to the amount of arsenic in the species or in the sample for the 'Total Arsenic' row. The 'n.r.' entries specify that the relevant species was not reported in that material. For BCR-627, uncertainties are taken as the half-width of the 95 % confidence interval of the mean. For CRM 7405-b, uncertainties are the half-width of the expanded uncertainty interval calculated using a coverage factor of 2. For SRM 3232, the expanded uncertainties are reported with a coverage factor of 2.00, except for Arsenosugar-392 which was reported with a coverage factor of 2.03. All values are certified, except where specified.

	BCR-627	CRM 7405-b	SRM 3232
	Mass Fraction (units below)	Mass Fraction (mg/kg)	Mass Fraction (mg/kg)
Total Arsenic	4.8 ± 0.3 (mg/kg)	49.5 ± 1.0	38.3 ± 1.3
Arsenic acid [As(V)]	n.r.	24.4 ± 0.7	n.r.
Arsenosugar-408	n.r.	1.41 ± 0.04	n.r.
Arsenosugar-328	n.r.	0.44 ± 0.02	1.20 ± 0.14[b]
Arsenosugar-482	n.r.	0.2[a]	5.59 ± 0.51[b]
Arsenosugar-392	n.r.	0.16[a]	14.06 ± 0.72[b]
DMA	2.0 ± 0.3 (µmol/kg)	0.24[a]	0.479 ± 0.077[b]
Arsenobetaine	52 ± 3 (µmol/kg)	n.r.	n.r.

[a]Values were provided as technical information.
[b]Values were provided as reference values.

Table S2: The results of the bootstrap analysis are presented for BCR-627 and include As_Sug. This result estimates the fraction of arsenic atoms in each of the given species. A confidence interval (CI) is provided for each estimate and corresponds to the 5th and 95th percentile of each population constructed in the bootstrap analysis. Also provided are the CoA values. These are presented both as the species fractions in the extract and also compared to the species fractions that could be present in the raw material. The CoA uncertainties are based on the half-width of the 95 % confidence interval around the mean of a series of independent measurements. Note that many of the species listed were not reported (n.r.) in the certificate.

	BCR-627					
	As K-edge XANES		CoA (Based on Extract)		Known Species Fraction in Bulk	
Species	Result (%)	Confidence Interval (%)	Result (%)	Confidence Interval (%)	Result (%)	Lower Bound (%)
As2S3	0.0	-	n.r.	-	n.r.	-
As2O3	0.1	[0.0, 0.7]	n.r.	-	n.r.	-
AB	48.8	[26.0, 65.9]	96.3	[88.6, 100.0]	81.2	74.3
As_Sug	51.1	[33.7, 74.0]	n.r.	-	n.r.	-
TMA	0.0	-	n.r.	-	n.r.	-
TMAO	0.0	-	n.r.	-	n.r.	-
DMA	0.0	-	3.7	[3.1, 4.3]	3.1	2.6
DSMA	0.0	-	n.r.	-	n.r.	-
As2O5	0.0	-	n.r.	-	n.r.	-
pH=1.55	0.0	-	n.r.	-	n.r.	-
pH=4.69	0.0	-	n.r.	-	n.r.	-
pH=9.39	0.0	-	n.r.	-	n.r.	-
pH=12.75	0.0	-	n.r.	-	n.r.	-
Basic Glycerol	0.0	-	n.r.	-	n.r.	-
Acidic Glycerol	0.0	-	n.r.	-	n.r.	-

Table S3: The results of the bootstrap analysis are presented for each of the extraction by-products. This result estimates the fraction of arsenic atoms in each of the given species classes. A confidence interval (CI) is provided for each estimate and corresponds to the 5th and 95th percentile of each population constructed in the bootstrap analysis.

	Residue 1		Residue 2		Extract	
Class	Result (%)	Confidence Interval (%)	Result (%)	Confidence Interval (%)	Result (%)	Confidence Interval (%)
Sulfides (III)	0.0	[0.0, 0.1]	0.0	[0.0, 0.1]	0.0	[0.0, 0.2]
Inorganics (III)	0.1	[0.0, 1.8]	0.2	[0.0, 2.1]	0.9	[0.0, 4.6]
Organics (V)	99.9	[98.2, 100.0]	99.8	[97.7, 100.0]	98.9	[95.4, 100.0]
Inorganics (V)	0.0	-	0.0	-	0.2	[0.0, 1.1]
Glycerol (V)	0.0	[0.0, 0.2]	0.0	[0.0, 0.4]	0.0	[0.0, 0.1]

Table S4: The results of the bootstrap analysis are presented for the eluent extracted from SRM 3232. This result estimates the fraction of arsenic atoms in each of the given species. A confidence interval (CI) is provided for each estimate and corresponds to the 5th and 95th percentile of each population constructed in the bootstrap analysis.

	Extract	
Species	Result (%)	Confidence Interval (%)
As_2S_3	0.0	[0.0, 0.2]
As_2O_3	0.9	[0.0, 4.6]
AB	2.1	[0.0, 12.4]
As_Sug	91.6	[66.5, 99.7]
TMA	0.0	-
TMAO	0.0	-
DMA	4.6	[0.0, 31.2]
DSMA	0.6	[0.0, 3.2]
As_2O_5	0.1	[0.0, 1.0]
pH=1.55	0.0	[0.0, 0.3]
pH=4.69	0.0	[0.0, 0.3]
pH=9.39	0.0	-
pH=12.75	0.0	-
Basic Glycerol	0.0	-
Acidic Glycerol	0.0	[0.0, 0.1]