

**Supplementary material**

**Synchrotron X-ray absorption spectroscopy analysis of arsenic chemical speciation in human nail clippings**

*Olena Ponomarenko,<sup>A,E</sup> Mihai R. Gherase,<sup>B</sup> Mark S. LeBlanc,<sup>B</sup> Chang-Yong Kim,<sup>C</sup> Elstan D. Desouza,<sup>D</sup> Michael J. Farquharson,<sup>D</sup> Fiona E. McNeill,<sup>D</sup> Susan Nehzati,<sup>A</sup> Ingrid J. Pickering,<sup>A</sup> Graham N. George<sup>A</sup> and David E. B. Fleming<sup>B,E</sup>*

<sup>A</sup>Molecular and Environmental Sciences Research Group, Department of Geological Sciences, University of Saskatchewan, SK, S7N 5E2 Canada.

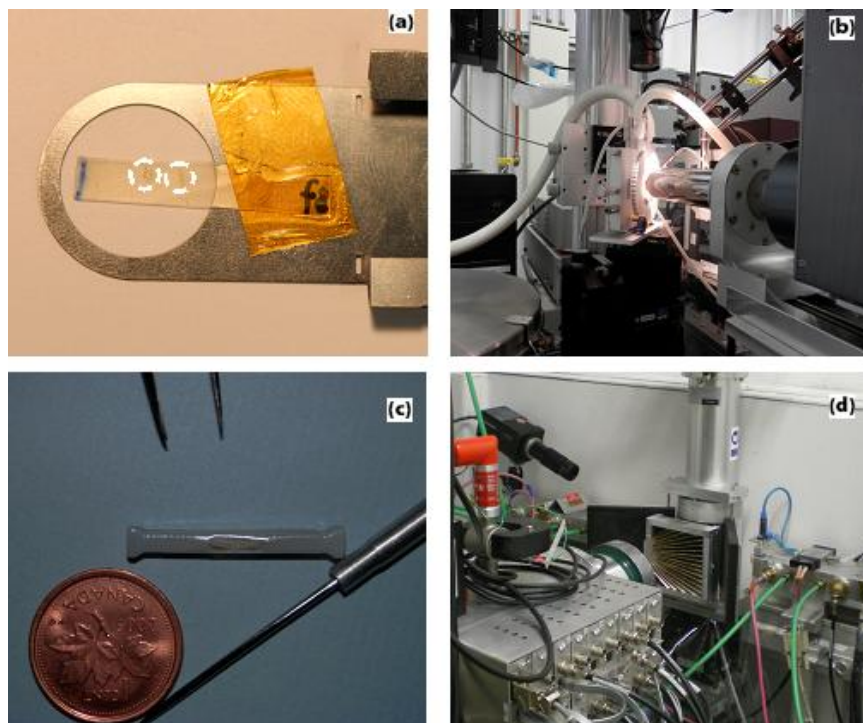
<sup>B</sup>Physics Department, Mount Allison University, Sackville, NB, E4L 1E6, Canada.

<sup>C</sup>Canadian Light Source, 44 Innovation Boulevard, Saskatoon, SK, S7N 2V3, Canada.

<sup>D</sup>Department of Medical Physics and Applied Radiation Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada.

<sup>E</sup>Corresponding authors. Email: olena.ponomarenko@usask.ca, dfleming@mta.ca

## Experimental setup



**Fig. S1.** (a) A cut of a polyester-resin with embedded fingernail clippings (encircled) mounted on the holder for microbeam XRF imaging and XAS. (b) The microbeam XRF imaging and  $\mu$ XAS set-up at the HXMA beamline. (c) Pieces of a fingernail clipping packaged in a ‘modified dogbone’ Lucite cell, prepared for bulk XAS. A Canadian 1¢ coin is shown to scale. (d) The fluorescence mode bulk XAS set-up at the SSRL beamline 7–3.

## Principal component analysis

PCA and TARGET routines in EXAFSPAK were used for the principal component analysis and target transform, respectively. PCA is based on single value decomposition (SVD) algorithm to generate a set of principal components (eigenvectors, or transformation basis) and weights (eigenvalues) representing the variation in the set of spectra.

The eigenvectors obtained from the PCA are mathematical abstractions with no simple relationship to the chemical components or species in the samples. The number of significant components with relatively large eigenvalues is an indication of the minimum number of chemical species needed to reconstruct a set of data.<sup>[1,2]</sup>

The Malinowski’s indicator function (IND) was used to estimate the minimum number of components required to fit the data. The correct value for the number of components is found by finding minimal IND values.<sup>[3]</sup> An additional test was to gradually decrease the number of

eigenvectors starting from those with the lowest eigenvalues and check the quality of the reconstructed spectra.<sup>[1,2]</sup>

To test whether the spectra of the selected reference standards were the likely components of the As spectra in nail samples, we performed target transforms that related the principal components to the spectra of the candidate physical reference standards. The target transformation algorithm tests whether the spectra of the selected reference standards are spanned by the transformation basis (principal components) obtained in SVD procedure, so that they can be used as possible components of the arsenic XAS spectra in the nail samples.<sup>[3]</sup> A superimposition of the reference spectrum ( $T$ ) and the resulting vector from PCA ( $T^*$ ) together with the variance ( $\Sigma(T-T^*)^2/N$ , with  $N$  data points) was used to determine which particular reference compound is a probable As species in nail samples.<sup>[2,3]</sup>

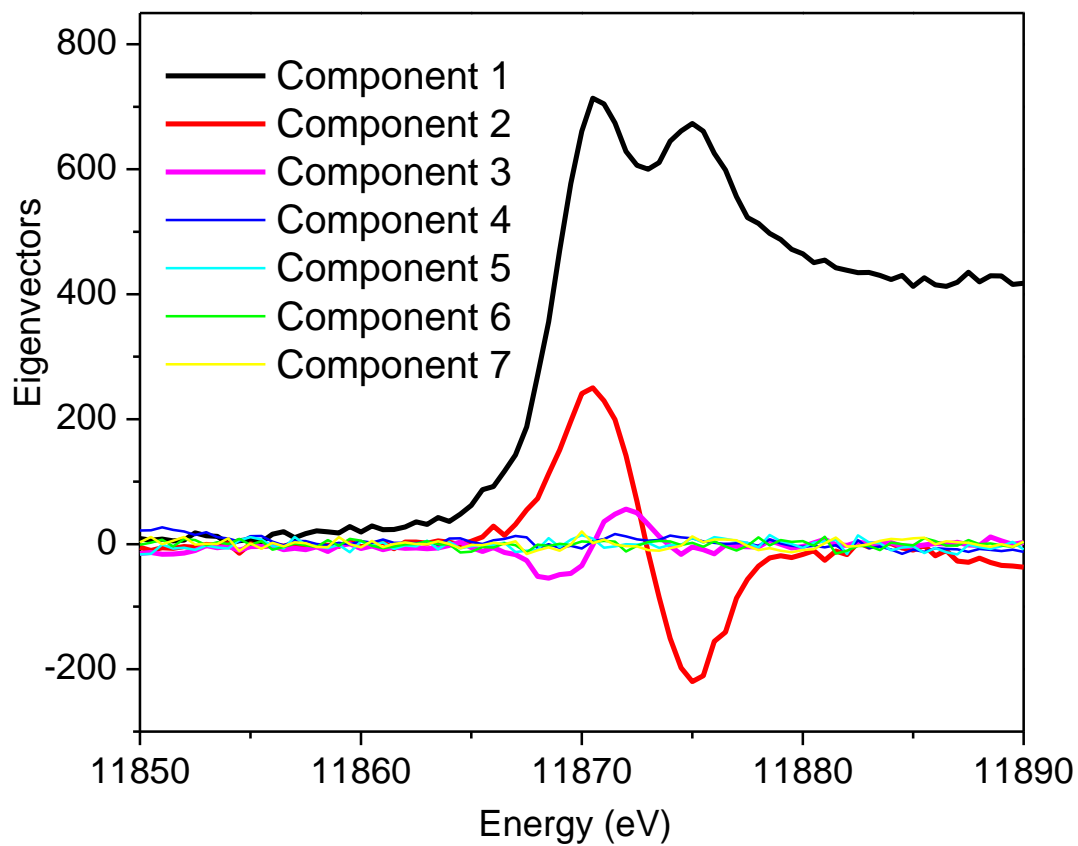
Initially we applied SVD decomposition (PCA subroutine) to the set of 7 spectra (1t-7-2, 1t-7-4, 2f-5-2, 2f-3-2, 2f-3-3, 2f-4-1, 6f-3) each consisting of 81 points (energy range 11850–11890 eV). The resulting eigenvalues and the values of IND functions are presented in Table S1. The principal components are shown in Fig. S2. The results indicate that three out of the seven components are needed to reproduce the entire set of the data. The results of target transform for the selected reference spectra are shown in Fig. S3. For example, the arsenobetaine and DMA<sup>V</sup> reference standards do not look promising based on both visual examination and high values of residuals (Table S2). On the other hand, As(GS)<sub>3</sub>, MMA<sup>III</sup>, DMA<sup>III</sup> arsenite, arsenate and TMS reasonably well represent chemical coordination in the nail As near-edge spectra. A similar analysis was applied to the reduced sets of data (i.e. without 6f-3, 2f-5-2, and/or 2f-4-1) to check if this affects the quality of the target transformation of the reference standard. At the next stage, the DATFIT subroutine (least-square fitting using linear combination of chemical standards) was used to analyse chemical speciation in the samples. For the testing purposes, we applied DATFIT subroutine using all reference standards listed in Table 1 (the main text), and found the results being consistent with the PCA analysis. After using DATFIT subroutine and selecting sets with smallest standard deviations for each component's fit, we were able to reduce this list to As(GS)<sub>3</sub>, AsO<sub>4</sub><sup>-</sup>, As(OH)<sub>3</sub>. Note, that all available information is used for the selection of best fitting set. For example, the As near edge XAS spectra of [(GS)<sub>2</sub>AsSe]<sup>-</sup> and As(GS)<sub>3</sub> reference compounds are very similar, but the first compound had to be rejected since it is very unstable in air atmosphere, and also because the results of our previous microbeam XRF multi-element imaging studies showed that selenium content in these samples is practically negligible.

**Table S1. Results of principal component analysis (PCA) of the set of nail clipping samples, represented by eigenvalues and IND function**

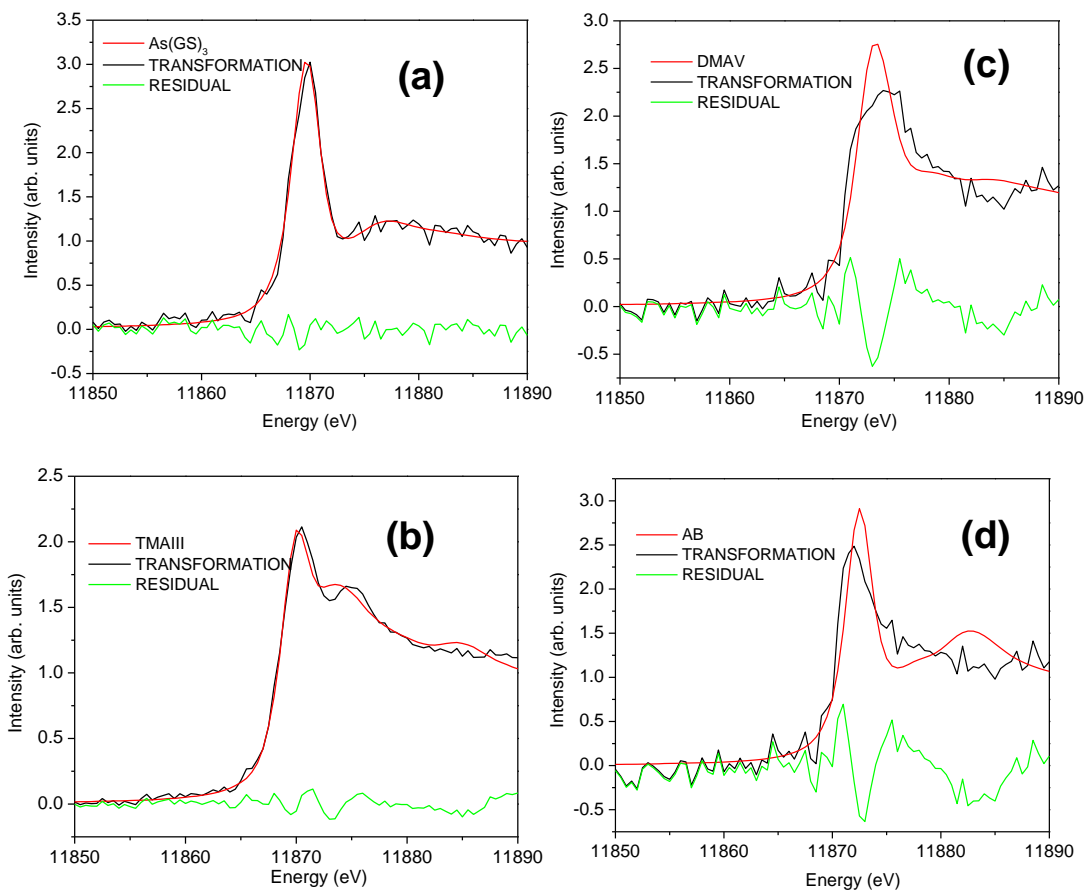
Component	Eigenvalue	IND
1	3491.28	0.973
2	742.615	0.421
3	162.073	0.472
4	86.4894	0.748
5	68.6892	1.559
6	51.6669	6.698
7	60.2804	9.999

**Table S2. Target transform results for reference standards**

Target reference standard	Residual $\times$ 1000
As(GS) <sub>3</sub>	0.611
[(GS) <sub>2</sub> AsSe] <sup>-</sup>	0.521
As(OH) <sub>3</sub>	0.447
AsO <sub>4</sub> <sup>-</sup>	0.236
MMA <sup>III</sup>	0.383
TMA <sup>III</sup>	0.224
MMA <sup>III</sup> -LIP	0.580
DMA <sup>III</sup>	0.294
MMA <sup>V</sup>	2.070
DMA <sup>V</sup>	3.556
AB	5.530
TMAO	4.095
TMAS	3.833
As(Cat) <sub>3</sub>	39.477



**Fig. S2.** Eigenvectors obtained using PCA for the complete set of nail As near edge XAS spectra. The analysis indicates that three components out of seven are sufficient to reproduce the entire dataset.



**Fig. S3.** Target transforms of reference standard spectra: (a) arsenic-*tris*-glutathione, (b) trimethylarsine, (c) dimethylarsenic acid, (d) arsenobetaine.

## References

- [1] G. N. George, I. J. Pickering, X-Ray absorption spectroscopy in biology and chemistry, in *Brilliant Light in Life and Material Sciences* (Eds V. Tsakanov, H. Wiedemann) 2007, pp. 97–119 (Springer: Dordrecht, the Netherlands).
- [2] J. Essilfie-Dughan, I. J. Pickering, M. J. Hendry, G. N. George, T. Kotzer, Molybdenum speciation in uranium mine tailings using X-ray absorption spectroscopy. *Environ. Sci. Technol.* 2011, 45, 455. doi:10.1021/es102954b
- [3] E. Malinowski, D. Howery, *Factor Analysis in Chemistry* 1980 (Wiley: New York).