Electronic Supplementary Information

How natural materials remove heavy metals from water: mechanistic insights from molecular dynamics simulations

Fabio Pietrucci^a, Mauro Boero^b, Wanda Andreoni^{c,d}

^aSorbonne Université, Muséum National d'Histoire Naturelle, CNRS, UMR 7590, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC), 4 Pl Jussieu, F-75005 Paris, (France) fabio.pietrucci@sorbonne-universite.fr

^b Université de Strasbourg, Institut de Physique et Chimie des Matériaux de Strasbourg (IPCMS), CNRS, UMR 7504, 23 rue du Loess, F-67034 Strasbourg (France) mauro.boero@ipcms.unistra.fr

^cEcole Polytechnique Fédérale de Lausanne (EPFL), Institut de Physique, CH-1015 Lausanne (Switzerland) <u>wanda.andreoni@epfl.ch</u>

^dIstituto Italiano di Tecnologia (IIT), Via Morego 30, I-16163 Genova (Italy)

Text S1. METHODS

1.1 Classical MD

1.1.1 Model: Notation <u>http://www.genome.jp/kegg/catalog/codes2.html</u>). We employed the Glycam-web server <u>http://glycam.org/tools/molecular-dynamics/oligosaccharide-builder/build-glycan</u> to construct an initial 3D structure of the polymer. As reported in the main text, The sequence is

(Man20GlcA10Man20)Gal(Man20GlcA10Man20)Gal(Man20GlcA10Man20)Gal(Man20GlcA 10Man20) where Man = D-Mannose, Gal = D-Galactose, GlcA = D-Glucuronic acid. (see Figure S1). This structure was then solvated with water and Pb ions were added, resulting in 4186 atoms in the organic + 20 Pb²⁺ + 31782 waters. In particular, the organic system contains the following oxygen-based functional groups: 40 carboxylate anions (COO⁻), 570 hydroxy (C-OH) and 406 ether groups (-COC-). Total atoms = 99552, total charge = 0, cubic box of side 100.476 A.

Note that the pK_a values of carboxylates are ~4 to be compared to pK_a=14 of water at room temperature. For this reason, in our models these groups are deprotonated. (https://clas.sa.ucsb.edu/staff/Resource%20Folder/Chem109ABC/Acid,%20Base%20Strength/T

able%20of%20Acids%20w%20Kas%20and%20pKas.pdf)



Figure S1 - Three-dimensional structure of the organic polymer, focusing on the links between different monomers: a) D-glucuronic acid and D-mannose, b) D-mannose and D-galactose. The complete sequence is given in the text.

1.1.2 MD simulations

All calculations used the Gromacs code $(5.1.43)^1$. Files initially in Amber format were converted to Gromacs using the tool acpype. As force-fields we use GLYCAM_06² for the organic part (version j-1 compatible with amber ff14SB), TIP3P for water³ and a simple Lennard-Jones potential for the lead ions⁴.

After an initial relaxation of the as-built solvated model, simulations in the NPT ensemble were performed at T=300K using stochastic velocity rescaling with time constant 1 ps^5 and at P=1 bar using an isotropic Berendsen barostat with time constant 1 ps^6 . We adopted a timestep of 1 fs.

Electrostatic interactions are calculated with the particle-mesh Ewald method (Fourier spacing 0.12). The cutoff for non-bonded interactions was set at 9 Å.

1.2 Ab Initio MD

Periodic boundary conditions are applied with a cubic box of side 21A. Simulations are performed in the NVT ensemble, at 300K using the Nosé-Hoover thermostat⁷. We adopted a timestep of 0.072 fs. Norm-conserving angular-momentum dependent pseudopotentials are used to represent the core-valence interaction and plane waves as basis set for the valence electron wavefunctions. Goedecker-Teter-Hutter-Krack^{8,9} pseudopotentials with PBE exchange-correlation functionals¹⁰ are from <u>https://www.cp2k.org/static/potentials/</u>. In particular, the Pb pseudopotentials correspond to the Pb¹⁴⁺ configuration. The expansion in plane waves is up to a 120Ry cutoff for the valence wavefunctions and 480Ry for the electron density. All calculations are performed with the CPMD code.¹¹

1.3 Metadynamics

The duration of classical metadynamics simulations was between 450 and 850 ps, that of ab initio metadynamics was of 25 and 27 ps for the forward simulations, and less than 10 ps for the backward simulations

All simulations were performed employing the open source Plumed version 1.3 software¹² using a modified version that can be freely downloaded (together with examples of use) at <u>https://sites.google.com/site/fabiopietrucci/Home/download</u>.

1.3.1 Path CVs

All metadynamics simulations employed path CVs^{13} based on the Euclidean distance between coordination patterns as a metric¹⁴

$$s = \frac{\sum_{k=1}^{n} k \ e^{-\lambda D[x(t), x_k]}}{\sum_{k=1}^{n} e^{-\lambda D[x(t), x_k]}} \qquad z = -\frac{1}{\lambda} \log \left[\sum_{k=1}^{n} e^{-\lambda D[x(t), x_k]} \right]$$
$$D[x(t), x_k] = \sum_{i\beta} \left[C_{i\beta}(t) - C_{i\beta}^k \right]^2 \qquad C_{i\beta}(t) = \sum_{j \in \beta} \frac{1 - \left[\frac{r_{ij}(t)}{d_{\alpha\beta}} \right]^n}{1 - \left[\frac{r_{ij}(t)}{d_{\alpha\beta}} \right]^m}$$

In the above formulas, the progress along the path (*s*) and the distance from the path (*z*) are expressed as a function of the current configuration of the system, x(t), and of *n* reference configurations x_k , that are meant to represent an initial guess of the reaction mechanism. To avoid making any specific guess, in our simulations we adopted only 2 reference states (*n*=2), i.e., the reactants (*k*=1, Pb bound to the organic) and the products (*k*=2, Pb dissociated from the organic and solvated by water). The distance *D* is computed from the coordination numbers $C_{i\beta}$ of atom *i*

with species β , with parameters *n*=6, *m*=9, *d*_{PbO} = 3.3 A for classical metadynamics simulations, and *m*=12 for ab initio metadynamics simulations. The parameter λ is chosen so that $\lambda D_{12} = 2.3$ as a good compromise between a path CV landscape with very narrow regions for reactants and products (λ large) and one with very broad, overlapping reactants and products, leading to a bad resolution of the barrier region (λ small). Reference coordination patterns are obtained from equilibrium trajectories of the reactants and of the products.

The following reference coordination patterns have been employed for classical metadynamics simulations:

| k=1 | O from organic | O from water |
|-----|-------------------|-----------------|
| Pb* | 21 | 26 |

| k=2 | O from organic | O from water |
|-----|-------------------|-----------------|
| Pb* | 0 | 50 |

The following reference coordination patterns have been employed for the first ab initio metadynamics simulation (the last four columns refer to oxygen atoms belonging to carboxylate groups that coordinate Pb in the reactant state):

| k=1 | O from | O from | O from | O from | O from | O from |
|-----|---------|--------|--------|------------------|--------|--------|
| | organic | water | COO1 | COO ₂ | COO₃ | COO4 |
| Pb* | 6.087 | 2.507 | 1.092 | 1.255 | 0.717 | 0.752 |

| k=2 | O from | O from | O from | O from | O from | O from |
|-----|---------|--------|--------|------------------|--------|--------|
| | organic | water | COO1 | COO ₂ | COO₃ | COO4 |
| Pb* | 0 | 7.4 | 0 | 0 | 0 | 0 |

The following reference coordination patterns have been employed for the second ab initio metadynamics simulation (the last five columns refer to oxygen atoms belonging to carboxylate groups that coordinate Pb in the reactant state):

| k=1 | O from organic | O from water | O from COO ₁ | O from COO ₂ | O from COO ₃ | O from COO4 | O from COO ₅ |
|-----|-------------------|-----------------|----------------------------|----------------------------|----------------------------|----------------|-------------------------|
| Pb* | 6.984 | 1.381 | 1.404 | 1.419 | 1.001 | 0.827 | 1.105 |

| k=2 | O from organic | O from water | O from COO1 | O from COO ₂ | O from COO₃ | O from COO4 | O from COO ₅ |
|-----|-------------------|-----------------|----------------|----------------------------|----------------|----------------|-------------------------|
| Pb* | 0 | 7.4 | 0 | 0 | 0 | 0 | 0 |

1.3.2 Bias potential

We performed metadynamics simulations on the classical models by depositing Gaussian hills of height 2.4 kJ/mol every 300 fs, with width 0.01 and 1 along the *s* and *z* directions, respectively. For the ab initio simulations, we deposited hills of height 4.2 kJ/mol every 36 fs, with width 0.03 and 0.8 along the *s* and *z* directions, respectively. As discussed in the main text, for the sake of comparison, we also performed classical metadynamics simulations using the ab initio hills parameters (see Fig. S6)

Text S2. Results and Discussion



Figure S2. Results of 3 unbiased simulations: Fraction of ions that are bound to at least one carboxylate, as a function of time.



Figure S3. Results of one of the unbiased MD simulations for each of the 20 Pb ions as a function of time. We denote by "1" a case when Pb is at a binding distance from at least one of the carboxylate oxygens; "0" when the contact is lost.



Figure S4 a-d. Free energy surface for the unbinding of Pb reconstructed from four independent classical metadynamics simulations in different environments (full unbinding). The shaded area (S>1.8) represents the region where the bias profile is not representative of the free energy, due to lack of reverse transition (binding). Below: Evolution of coordination number of Pb and O atoms of water molecules and different groups in the organic (see legend) as a function of the S path coordinate. Vertical bars represent standard deviations.



Figure S5: Example of a configuration with solvated Pb ions. a) Pair distribution function and b) integral (coordination number) for the distance between Pb and O atoms of water, when Pb is far from the organic, from classical unbiased MD simulation of Pb capture.



Figure S6 a-b Free energy surface for the unbinding of Pb reconstructed from two independent ab initio metadynamics simulations in different environments. The shaded area (S>1.7) represents the region where the bias profile is not representative of the free energy, due to lack of reverse transition (binding). Below: Evolution of the coordination numbers of Pb and O belonging to water and to different organic functional groups, as a function of the S path coordinate. The vertical bars represent standard deviations of coordination numbers.



Figure S7 a-b Free energy surface for the binding of Pb reconstructed from ab initio metadynamics simulation started from the end (unbound state, S>1.8) of the simulations in Fig. 4. In the shaded area - S<1.2 – the bias is not representative of the free-energy. Evolution of the coordination of Pb with oxygen atoms belonging to water and to different organic functional groups as function of the S path coordinate.



Figure S8 a-d Free energy surface for the unbinding of Pb reconstructed from four independent classical metadynamics simulations in different environments (full unbinding) with protocol adjusted to reproduce that of ab initio metadynamics. The shaded area (S>1.8) represents the region where the bias profile is not representative of the free energy, due to lack of reverse transition (binding). The barriers estimated from the bias potentials are a) 50 kcal/mol, b) 70 kcal/mol, c) 100 kcal/mol, d) 180 kcal/mol.



Figure S9: Vibrational density of states (arbitrary units) calculated from (unbiased) MD trajectories of the DFT model and projected onto the carboxylate groups. The spectral features correspond to the symmetric and antisymmetric stretching modes. $VDOS(\omega) = \int dt \ e^{i\omega t} < \mathbf{v}_{CO}(t) \ \mathbf{v}_{CO}(0) >$ where $\mathbf{v}_{CO}(t) = \mathbf{v}_{O}(t) - \mathbf{v}_{C}(t) = d[\mathbf{r}_{O}(t) - \mathbf{r}_{C}(t)]/dt$.

References

(1) <u>www.gromacs.org</u>

(2) K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley and R. J. Woods, *J. Comput. Chem.* 2008, **29**, 622-655.

(3) W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* 1983, **79**, 926–935.

- (4) A. S. de Araujo, M. T. Sonoda, O. E. Piro and E. E. Castellano, J. Phys. Chem. B 2007, 111, 2219-2224
- (5) G. Bussi, D. Donadio and M. Parrinello, J. Chem. Phys. 2007, 126, 014101-014107.

(6) H. J. C Berendsen, J. P. M. Postma., W. F. van Gunsteren, A. Di Nola and J. R. Haak, *J. Chem. Phys.* 1984, **81**, 3684-90.

(7) S. Nosé, J. Chem. Phys. 1984, 81, 511-519.

(8) S. Goedecker, M. Teter and J. Hutter J., Phys. Rev. B 1996, 54, 1703-1710

- (9) M. Krack, Theor. Chem. Acc. 2005, 114, 145-152
- (10) J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. 1996, 77, 3865-3868; ibid. 1998, 80, 891-894.

(11) CPMD, Copyright IBM Corp. 1990–2019, Copyright MPI für Festkörperforschung Stuttgart 1997–2001 www.cpmd.org

(12) M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Comput. Phys. Commun.* 2009, **180**, 1961–1972.

(13) D. Branduardi, F. L. Gervasio and M. Parrinello, J Chem Phys. 2007, 126, 054103-054112

(14) F. Pietrucci and A. M. Saitta, Proc. Natl Acad. Sci. USA 2015, 112, 15030 - 15035.