Supporting Information

Nucleation Mechanisms and Speciation of Metal Oxide Clusters

Enric Petrus,[a] Mireia Segado,[a] Carles Bo*[a,b]

[a] Institute of Chemical Research of Catalonia (ICIQ), The Barcelona Institute of Science and Technology (BIST), Av. Països Catalans, 16, 43007 Tarragona, Spain
[b] Departament de Química Física i Inorgánica, Universitat Rovira i Virgili, Marcel·lí Domingo s/n, 43007 Tarragona, Spain E-mail: cbo@iciq.cat

1. Computational Details

Molecular geometries for all species were fully optimized using a DFT method by employing ADF package (SCM ADF version 2019.1).[1] We have used several functionals: PBE,[2,3] PBE0,[4] BLYP,[5,6] B3LYP-D3[5,7,8] and BP86[2,3,6]. Unless otherwise stated, we have used PBE functional as reference. Relativistic corrections were taken into account by the scalar-relativistic zero-order regular approximation (ZORA)[9,10] at the TZP basis set level. Solvent effects were introduced using the continuous solvent model COSMO with Klamt radis for water.[11] Stationary points were characterized with analytic frequency calculations. Ground state free energies were computed using a temperature of 298 K and 1 atm of pressure. In addition, no standard state corrections have been considered so as not to introduce distorting effects which could affect the rescale of the formation constants. A dataset collection of the computational results is available in the ioChem-BD repository[12] and can be accessed via: http://dx.doi.org/10.19061/iochem-bd-1-178.

2. Bader Bond Critical Points as Edges

The chemical connectivity of each compound determines the edges of the subsequent molecular graphs. In general terms, the elucidation of the chemical bonds is straightforward, however, in case of considering non-covalent interactions, an alternative approach is needed. Actually, polyoxomolybdates assemble in acid media which emphasises the importance of taking into account hydrogen bonds that are known to stabilize these large negatively-charged structures.

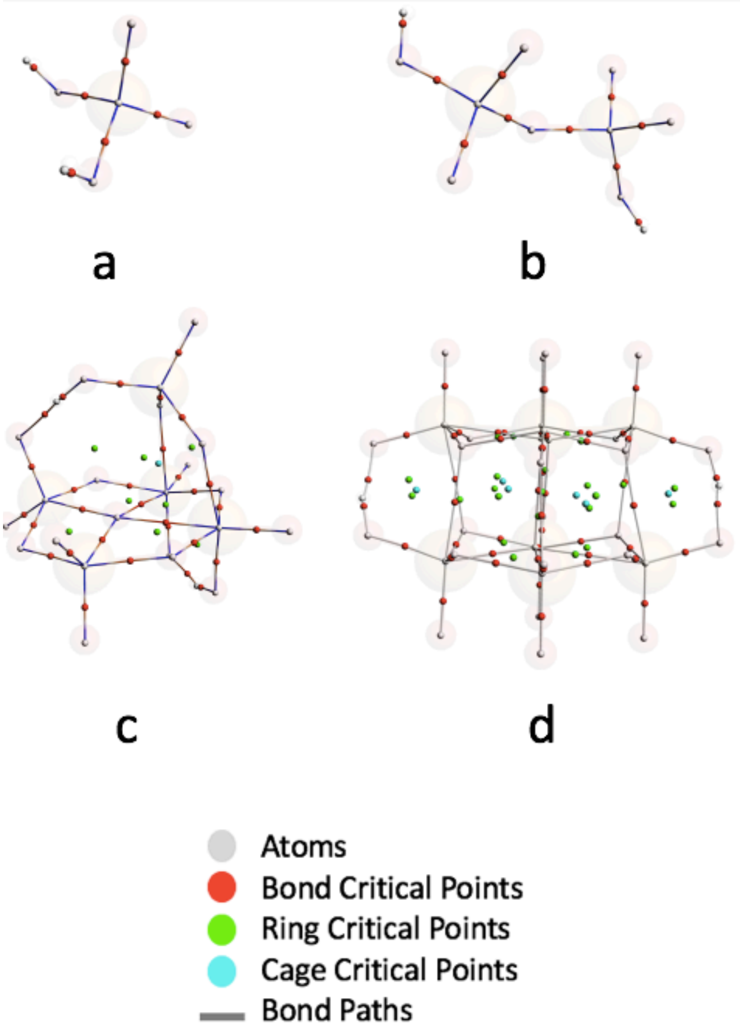


Fig S1 Bader topology for a) H_2MoO_4 b) $H_2Mo_2O_7$ c) $[H_2Mo_5O_{17}]^{2-}$ and d) $[H_2Mo_8O_{26}]^{2-}$. Grey, red, green and cyan circles correspond to the atoms, bonds, rings, and cage critical points, respectively. Grey line indicates the bond paths.

Figure S1 collects the Bader topology of four polyoxometalates considered in this work: H_2MoO_4, $H_2Mo_2O_7$, $[H_2Mo_5O_{17}]^{2-}$ and $[H_2Mo_8O_{26}]^{2-}$. From Figure S1a and S1b it can be observed that the chosen methodology correctly detects the metal-oxygen covalent bonds. Furthermore, Figure S1c and Figure S1d show that hydrogen bonds are described as an explicit connection. Therefore, the current approach provides an accurate description of the chemical topology which is very important so as to formulate reliable molecular graphs.

3. Chemical Reactions

3.1. Reactions in Figure 1

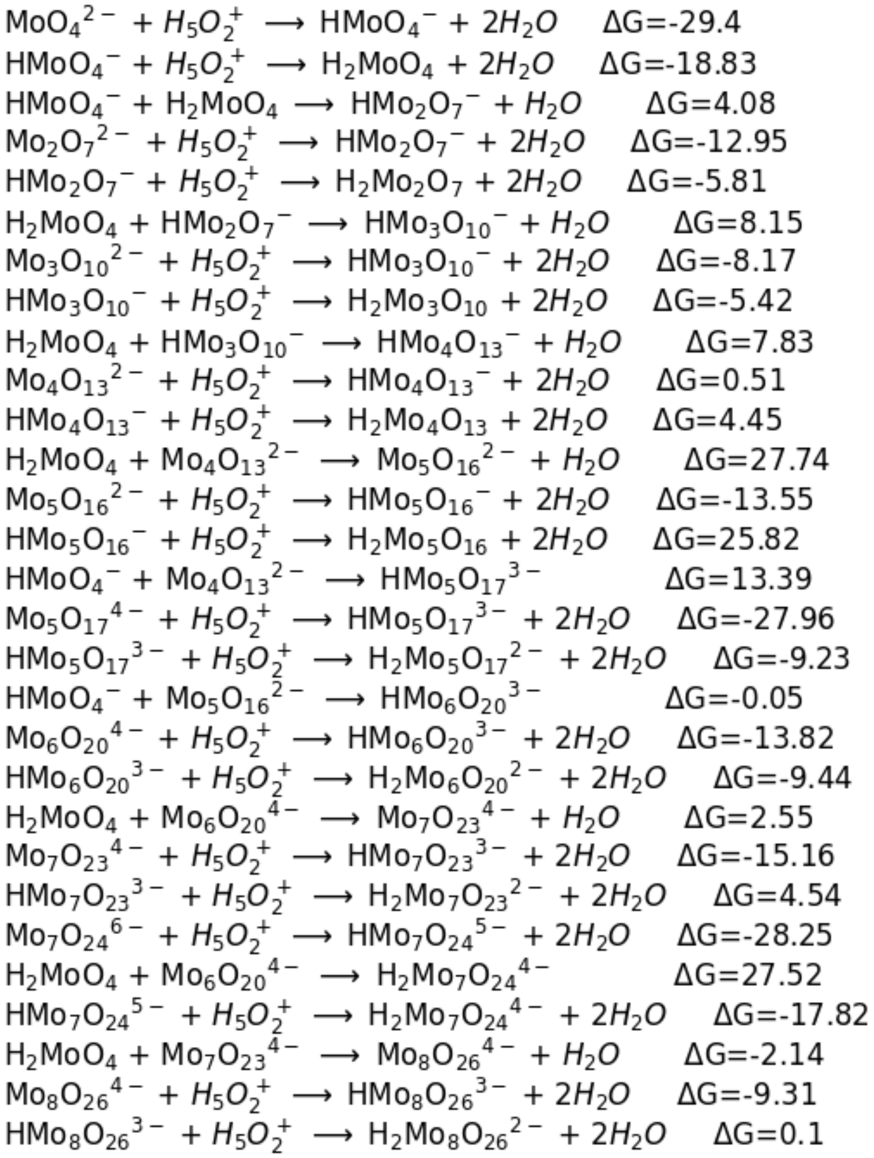
The 72 chemical reactions obtained from the transformation of the isomorphic matrix.

$MoO_4^{2-} + H_5O_2^+ \longrightarrow HMoO_4^- + 2H_2O \quad \Delta G=-29.4$
$HMoO_4^- + H_5O_2^+ \longrightarrow H_2MoO_4 + 2H_2O \quad \Delta G=-9.83$
$HMoO_4^- + HMoO_4^- \longrightarrow Mo_2O_7^{2-} + H_2O \quad \Delta G=-1.8$
$HMoO_4^- + H_2MoO_4 \longrightarrow HMo_2O_7^- + H_2O \quad \Delta G=-4.92$
$Mo_2O_7^{2-} + H_5O_2^+ \longrightarrow HMo_2O_7^- + 2H_2O \quad \Delta G=-12.95$
$H_2MoO_4 + H_2MoO_4 \longrightarrow H_2Mo_2O_7 + H_2O \quad \Delta G=-0.9$
$HMo_2O_7^- + H_5O_2^+ \longrightarrow H_2Mo_2O_7 + 2H_2O \quad \Delta G=-5.81$
$HMoO_4^- + HMo_2O_7^- \longrightarrow Mo_3O_{10}^{2-} + H_2O \quad \Delta G=-2.51$
$H_2MoO_4 + Mo_2O_7^{2-} \longrightarrow Mo_3O_{10}^{2-} + H_2O \quad \Delta G=-5.63$
$HMoO_4^- + H_2Mo_2O_7 \longrightarrow HMo_3O_{10}^- + H_2O \quad \Delta G=-4.87$
$H_2MoO_4 + HMo_2O_7^- \longrightarrow HMo_3O_{10}^- + H_2O \quad \Delta G=-0.85$
$Mo_3O_{10}^{2-} + H_5O_2^+ \longrightarrow HMo_3O_{10}^- + 2H_2O \quad \Delta G=-8.17$
$H_2MoO_4 + H_2Mo_2O_7 \longrightarrow H_2Mo_3O_{10} + H_2O \quad \Delta G=-0.46$
$HMo_3O_{10}^- + H_5O_2^+ \longrightarrow H_2Mo_3O_{10} + 2H_2O \quad \Delta G=-5.42$
$H_2MoO_4 + Mo_3O_{10}^{2-} \longrightarrow Mo_4O_{13}^{2-} + H_2O \quad \Delta G=-9.85$
$Mo_2O_7^{2-} + H_2Mo_2O_7 \longrightarrow Mo_4O_{13}^{2-} + H_2O \quad \Delta G=-14.58$
$HMo_2O_7^- + HMo_2O_7^- \longrightarrow Mo_4O_{13}^{2-} + H_2O \quad \Delta G=-7.44$
$H_2MoO_4 + HMo_3O_{10}^- \longrightarrow HMo_4O_{13}^- + H_2O \quad \Delta G=-1.17$
$HMo_2O_7^- + H_2Mo_2O_7 \longrightarrow HMo_4O_{13}^- + H_2O \quad \Delta G=-1.12$
$Mo_4O_{13}^{2-} + H_5O_2^+ \longrightarrow HMo_4O_{13}^- + 2H_2O \quad \Delta G=0.51$
$H_2MoO_4 + H_2Mo_3O_{10} \longrightarrow H_2Mo_4O_{13} + H_2O \quad \Delta G=8.7$
$H_2Mo_2O_7 + H_2Mo_2O_7 \longrightarrow H_2Mo_4O_{13} + H_2O \quad \Delta G=9.14$
$HMo_4O_{13}^- + H_5O_2^+ \longrightarrow H_2Mo_4O_{13} + 2H_2O \quad \Delta G=4.45$
$H_2MoO_4 + Mo_4O_{13}^{2-} \longrightarrow Mo_5O_{16}^{2-} + H_2O \quad \Delta G=18.74$
$HMo_2O_7^- + HMo_3O_{10}^- \longrightarrow Mo_5O_{16}^{2-} + H_2O \quad \Delta G=12.15$
$H_2Mo_2O_7 + Mo_3O_{10}^{2-} \longrightarrow Mo_5O_{16}^{2-} + H_2O \quad \Delta G=9.79$
$H_2Mo_2O_7 + HMo_3O_{10}^- \longrightarrow HMo_5O_{16}^- + H_2O \quad \Delta G=4.41$
$Mo_5O_{16}^{2-} + H_5O_2^+ \longrightarrow HMo_5O_{16}^- + 2H_2O \quad \Delta G=-13.55$
$H_2Mo_2O_7 + H_2Mo_3O_{10} \longrightarrow H_2Mo_5O_{16} + H_2O \quad \Delta G=35.65$
$HMo_5O_{16}^- + H_5O_2^+ \longrightarrow H_2Mo_5O_{16} + 2H_2O \quad \Delta G=25.82$
$MoO_4^{2-} + Mo_4O_{13}^{2-} \longrightarrow Mo_5O_{17}^{4-} \quad \Delta G=11.95$
$Mo_2O_7^{2-} + Mo_3O_{10}^{2-} \longrightarrow Mo_5O_{17}^{4-} \quad \Delta G=23.47$
$HMoO_4^- + Mo_4O_{13}^{2-} \longrightarrow HMo_5O_{17}^{3-} \quad \Delta G=13.39$
$Mo_2O_7^{2-} + HMo_3O_{10}^- \longrightarrow HMo_5O_{17}^{3-} \quad \Delta G=3.68$
$HMo_2O_7^- + Mo_3O_{10}^{2-} \longrightarrow HMo_5O_{17}^{3-} \quad \Delta G=8.46$
$Mo_5O_{17}^{4-} + H_5O_2^+ \longrightarrow HMo_5O_{17}^{3-} + 2H_2O \quad \Delta G=-27.96$
$H_2MoO_4 + Mo_4O_{13}^{2-} \longrightarrow H_2Mo_5O_{17}^{2-} \quad \Delta G=13.99$

$\mathrm{Mo_2O_7}^{2-} + \mathrm{H_2Mo_3O_{10}} \longrightarrow \mathrm{H_2Mo_5O_{17}}^{2-} \quad \Delta G=-0.13$
$\mathrm{HMo_2O_7}^{-} + \mathrm{HMo_3O_{10}}^{-} \longrightarrow \mathrm{H_2Mo_5O_{17}}^{2-} \quad \Delta G=7.4$
$\mathrm{H_2Mo_2O_7} + \mathrm{Mo_3O_{10}}^{2-} \longrightarrow \mathrm{H_2Mo_5O_{17}}^{2-} \quad \Delta G=5.04$
$\mathrm{HMo_5O_{17}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_5O_{17}}^{2-} + 2H_2O \quad \Delta G=-9.23$
$\mathrm{MoO_4}^{2-} + \mathrm{Mo_5O_{16}}^{2-} \longrightarrow \mathrm{Mo_6O_{20}}^{4-} \quad \Delta G=-15.63$
$\mathrm{H_2MoO_4} + \mathrm{Mo_5O_{17}}^{4-} \longrightarrow \mathrm{Mo_6O_{20}}^{4-} + H_2O \quad \Delta G=-8.84$
$\mathrm{Mo_3O_{10}}^{2-} + \mathrm{Mo_3O_{10}}^{2-} \longrightarrow \mathrm{Mo_6O_{20}}^{4-} \quad \Delta G=20.26$
$\mathrm{HMoO_4}^{-} + \mathrm{Mo_5O_{16}}^{2-} \longrightarrow \mathrm{HMo_6O_{20}}^{3-} \quad \Delta G=-0.05$
$\mathrm{Mo_3O_{10}}^{2-} + \mathrm{HMo_3O_{10}}^{-} \longrightarrow \mathrm{HMo_6O_{20}}^{3-} \quad \Delta G=14.61$
$\mathrm{Mo_6O_{20}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_6O_{20}}^{3-} + 2H_2O \quad \Delta G=-13.82$
$\mathrm{H_2MoO_4} + \mathrm{Mo_5O_{16}}^{2-} \longrightarrow \mathrm{H_2Mo_6O_{20}}^{2-} \quad \Delta G=0.34$
$\mathrm{Mo_3O_{10}}^{2-} + \mathrm{H_2Mo_3O_{10}} \longrightarrow \mathrm{H_2Mo_6O_{20}}^{2-} \quad \Delta G=10.59$
$\mathrm{HMo_3O_{10}}^{-} + \mathrm{HMo_3O_{10}}^{-} \longrightarrow \mathrm{H_2Mo_6O_{20}}^{2-} \quad \Delta G=13.34$
$\mathrm{HMo_6O_{20}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_6O_{20}}^{2-} + 2H_2O \quad \Delta G=-9.44$
$\mathrm{H_2MoO_4} + \mathrm{Mo_6O_{20}}^{4-} \longrightarrow \mathrm{Mo_7O_{23}}^{4-} + H_2O \quad \Delta G=-6.45$
$\mathrm{Mo_3O_{10}}^{2-} + \mathrm{Mo_4O_{13}}^{2-} \longrightarrow \mathrm{Mo_7O_{23}}^{4-} \quad \Delta G=23.66$
$\mathrm{Mo_3O_{10}}^{2-} + \mathrm{HMo_4O_{13}}^{-} \longrightarrow \mathrm{HMo_7O_{23}}^{3-} \quad \Delta G=7.99$
$\mathrm{HMo_3O_{10}}^{-} + \mathrm{Mo_4O_{13}}^{2-} \longrightarrow \mathrm{HMo_7O_{23}}^{3-} \quad \Delta G=16.67$
$\mathrm{Mo_7O_{23}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_7O_{23}}^{3-} + 2H_2O \quad \Delta G=-15.16$
$\mathrm{HMo_3O_{10}}^{-} + \mathrm{HMo_4O_{13}}^{-} \longrightarrow \mathrm{H_2Mo_7O_{23}}^{2-} \quad \Delta G=20.7$
$\mathrm{H_2Mo_3O_{10}} + \mathrm{Mo_4O_{13}}^{2-} \longrightarrow \mathrm{H_2Mo_7O_{23}}^{2-} \quad \Delta G=26.63$
$\mathrm{HMo_7O_{23}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_7O_{23}}^{2-} + 2H_2O \quad \Delta G=4.54$
$\mathrm{MoO_4}^{2-} + \mathrm{Mo_6O_{20}}^{4-} \longrightarrow \mathrm{Mo_7O_{24}}^{6-} \quad \Delta G=25.36$
$\mathrm{Mo_2O_7}^{2-} + \mathrm{Mo_5O_{17}}^{4-} \longrightarrow \mathrm{Mo_7O_{24}}^{6-} \quad \Delta G=37.89$
$\mathrm{HMoO_4}^{-} + \mathrm{Mo_6O_{20}}^{4-} \longrightarrow \mathrm{HMo_7O_{24}}^{5-} \quad \Delta G=26.51$
$\mathrm{Mo_2O_7}^{2-} + \mathrm{HMo_5O_{17}}^{3-} \longrightarrow \mathrm{HMo_7O_{24}}^{5-} \quad \Delta G=37.6$
$\mathrm{Mo_7O_{24}}^{6-} + H_5O_2^+ \longrightarrow \mathrm{HMo_7O_{24}}^{5-} + 2H_2O \quad \Delta G=-28.25$
$\mathrm{H_2MoO_4} + \mathrm{Mo_6O_{20}}^{4-} \longrightarrow \mathrm{H_2Mo_7O_{24}}^{4-} \quad \Delta G=18.52$
$\mathrm{HMo_7O_{24}}^{5-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_7O_{24}}^{4-} + 2H_2O \quad \Delta G=-17.82$
$\mathrm{H_2MoO_4} + \mathrm{Mo_7O_{23}}^{4-} \longrightarrow \mathrm{Mo_8O_{26}}^{4-} + H_2O \quad \Delta G=-11.14$
$\mathrm{Mo_4O_{13}}^{2-} + \mathrm{Mo_4O_{13}}^{2-} \longrightarrow \mathrm{Mo_8O_{26}}^{4-} \quad \Delta G=22.37$
$\mathrm{Mo_8O_{26}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_8O_{26}}^{3-} + 2H_2O \quad \Delta G=-9.31$
$\mathrm{Mo_4O_{13}}^{2-} + \mathrm{H_2Mo_4O_{13}} \longrightarrow \mathrm{H_2Mo_8O_{26}}^{2-} \quad \Delta G=8.2$
$\mathrm{HMo_4O_{13}}^{-} + \mathrm{HMo_4O_{13}}^{-} \longrightarrow \mathrm{H_2Mo_8O_{26}}^{2-} \quad \Delta G=12.14$

3.2 Reactions in Figure 5

The 29 Chemical reactions used to set up model 790. Note that the free energy of the molybdic acid has been stabilized by -9.0 kcal·mol^{-1}.



$$\mathrm{MoO_4}^{2-} + H_5O_2^+ \longrightarrow \mathrm{HMoO_4}^{-} + 2H_2O \quad \Delta G=-29.4$$
$$\mathrm{HMoO_4}^{-} + H_5O_2^+ \longrightarrow \mathrm{H_2MoO_4} + 2H_2O \quad \Delta G=-18.83$$
$$\mathrm{HMoO_4}^{-} + \mathrm{H_2MoO_4} \longrightarrow \mathrm{HMo_2O_7}^{-} + H_2O \quad \Delta G=4.08$$
$$\mathrm{Mo_2O_7}^{2-} + H_5O_2^+ \longrightarrow \mathrm{HMo_2O_7}^{-} + 2H_2O \quad \Delta G=-12.95$$
$$\mathrm{HMo_2O_7}^{-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_2O_7} + 2H_2O \quad \Delta G=-5.81$$
$$\mathrm{H_2MoO_4} + \mathrm{HMo_2O_7}^{-} \longrightarrow \mathrm{HMo_3O_{10}}^{-} + H_2O \quad \Delta G=8.15$$
$$\mathrm{Mo_3O_{10}}^{2-} + H_5O_2^+ \longrightarrow \mathrm{HMo_3O_{10}}^{-} + 2H_2O \quad \Delta G=-8.17$$
$$\mathrm{HMo_3O_{10}}^{-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_3O_{10}} + 2H_2O \quad \Delta G=-5.42$$
$$\mathrm{H_2MoO_4} + \mathrm{HMo_3O_{10}}^{-} \longrightarrow \mathrm{HMo_4O_{13}}^{-} + H_2O \quad \Delta G=7.83$$
$$\mathrm{Mo_4O_{13}}^{2-} + H_5O_2^+ \longrightarrow \mathrm{HMo_4O_{13}}^{-} + 2H_2O \quad \Delta G=0.51$$
$$\mathrm{HMo_4O_{13}}^{-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_4O_{13}} + 2H_2O \quad \Delta G=4.45$$
$$\mathrm{H_2MoO_4} + \mathrm{Mo_4O_{13}}^{2-} \longrightarrow \mathrm{Mo_5O_{16}}^{2-} + H_2O \quad \Delta G=27.74$$
$$\mathrm{Mo_5O_{16}}^{2-} + H_5O_2^+ \longrightarrow \mathrm{HMo_5O_{16}}^{-} + 2H_2O \quad \Delta G=-13.55$$
$$\mathrm{HMo_5O_{16}}^{-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_5O_{16}} + 2H_2O \quad \Delta G=25.82$$
$$\mathrm{HMoO_4}^{-} + \mathrm{Mo_4O_{13}}^{2-} \longrightarrow \mathrm{HMo_5O_{17}}^{3-} \quad \Delta G=13.39$$
$$\mathrm{Mo_5O_{17}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_5O_{17}}^{3-} + 2H_2O \quad \Delta G=-27.96$$
$$\mathrm{HMo_5O_{17}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_5O_{17}}^{2-} + 2H_2O \quad \Delta G=-9.23$$
$$\mathrm{HMoO_4}^{-} + \mathrm{Mo_5O_{16}}^{2-} \longrightarrow \mathrm{HMo_6O_{20}}^{3-} \quad \Delta G=-0.05$$
$$\mathrm{Mo_6O_{20}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_6O_{20}}^{3-} + 2H_2O \quad \Delta G=-13.82$$
$$\mathrm{HMo_6O_{20}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_6O_{20}}^{2-} + 2H_2O \quad \Delta G=-9.44$$
$$\mathrm{H_2MoO_4} + \mathrm{Mo_6O_{20}}^{4-} \longrightarrow \mathrm{Mo_7O_{23}}^{4-} + H_2O \quad \Delta G=2.55$$
$$\mathrm{Mo_7O_{23}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_7O_{23}}^{3-} + 2H_2O \quad \Delta G=-15.16$$
$$\mathrm{HMo_7O_{23}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_7O_{23}}^{2-} + 2H_2O \quad \Delta G=4.54$$
$$\mathrm{Mo_7O_{24}}^{6-} + H_5O_2^+ \longrightarrow \mathrm{HMo_7O_{24}}^{5-} + 2H_2O \quad \Delta G=-28.25$$
$$\mathrm{H_2MoO_4} + \mathrm{Mo_6O_{20}}^{4-} \longrightarrow \mathrm{H_2Mo_7O_{24}}^{4-} \quad \Delta G=27.52$$
$$\mathrm{HMo_7O_{24}}^{5-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_7O_{24}}^{4-} + 2H_2O \quad \Delta G=-17.82$$
$$\mathrm{H_2MoO_4} + \mathrm{Mo_7O_{23}}^{4-} \longrightarrow \mathrm{Mo_8O_{26}}^{4-} + H_2O \quad \Delta G=-2.14$$
$$\mathrm{Mo_8O_{26}}^{4-} + H_5O_2^+ \longrightarrow \mathrm{HMo_8O_{26}}^{3-} + 2H_2O \quad \Delta G=-9.31$$
$$\mathrm{HMo_8O_{26}}^{3-} + H_5O_2^+ \longrightarrow \mathrm{H_2Mo_8O_{26}}^{2-} + 2H_2O \quad \Delta G=0.1$$

4. Speciation Model Definition

The combination of 29 equations S1 and one equation S2 are needed to construct and solve each speciation model. Note that the total amount of molybdenum is fixed to 0.005M in the present study.

$$\prod[\mathrm{Products}] - \mathrm{e}^{-\Delta \mathrm{G_{DFT}}/\mathrm{RT}} \cdot \prod[\mathrm{Reactants}] = 0 \qquad \text{Eq. S1}$$

$$\sum_{\mathrm{i}}[\mathrm{Compound}]_{\mathrm{i}} - [0.005\mathrm{M}] = 0 \qquad \text{Eq. S2}$$

Followed by the adjust of the concentrations according to Davies equation:

$$\mathrm{Activity} = [\mathrm{Compound}]_{\mathrm{i}} \cdot \gamma_{\mathrm{i}} = [\mathrm{Compound}]_{\mathrm{i}} \cdot 10^{-0.5 \cdot \mathrm{z}^2 \cdot (\frac{\sqrt{\mathrm{I}}}{1+\sqrt{\mathrm{I}}} - 0.30\mathrm{I})} \qquad \text{Eq. S3}$$

5. Binomial Coefficient and Computational Cost

The binomial coefficient function provides an estimation of the number of speciation models that can be proposed. We are setting up models of 30 equations, but 20 refer to acid-base equilibria and 1 to mass balance, both considered to be invariant. However, the 9 equations left are variable and they refer to condensation and addition reactions. Thus the combinatorial function is defined by a numerator that corresponds to the total number of condensation and addition reactions (42 reactions) and a denominator which accounts for the number of variant equations (9 equations).

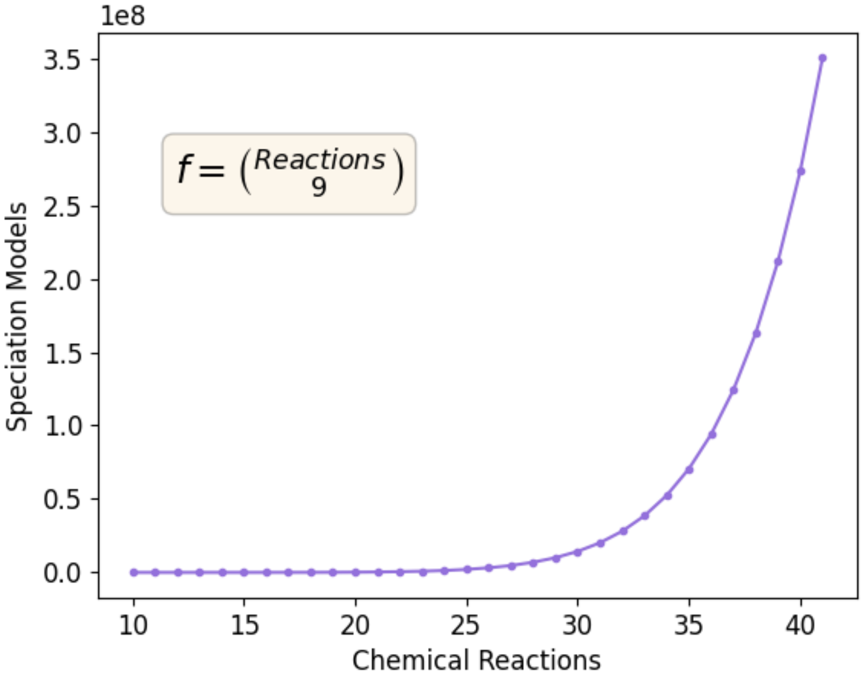


Fig S2 Number of speciation models as a function of the total number of chemical reactions.

Figure S2 shows the exponential trend of the binomial coefficient in a x domain from 10 to 42 chemical reactions. It can be observed that the function rapidly scales to the 10^8 magnitude. In addition, every speciation model has to be solved for a range of pH between 0 and 35 in steps of 0.2 units (i.e, 175 times). Therefore, the number of systems of non-lineal equations scale up to 10^{10} which is too computational demanding.

6. Formation Constant Definition

Formation reactions are formulated as follows:

$\lambda\ MoO_4^{2-} + (2\cdot\mu)\ H^+ \rightarrow Mo_\lambda O_{4\cdot\lambda-\mu}{}^{6\cdot\lambda-2\cdot4\cdot\lambda-\mu} + \mu\ H_2O$

Hence, the corresponding formation constant is:

$$\mathrm{pK_f} = -\log\left(\frac{[\mathrm{H_2O}]^{\mu}\left[\mathrm{Mo_\lambda O_{4\cdot\lambda-\mu}}\right]^{6\cdot\lambda-2\cdot4\cdot\lambda-\mu}}{[\mathrm{MoO_4}]^{\lambda}[\mathrm{H^{+1}}]^{2\mu}}\right) \qquad \text{Eq. S4}$$

7. Parametrized Formation Constants

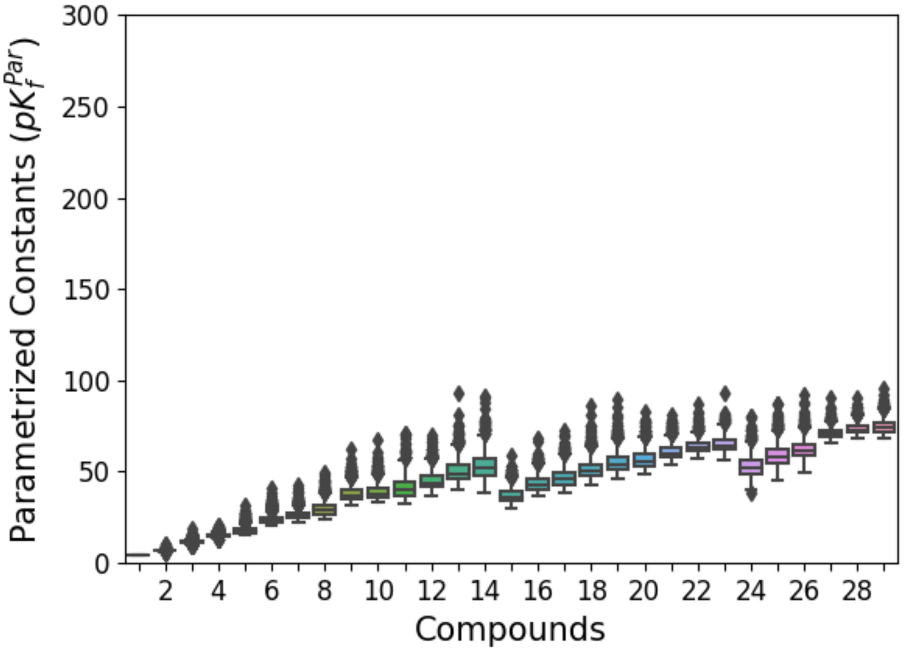


Fig S3 Box plot for the parametrized formation constants (pK_f^{Par}) of species 1 to 29 obtained by each speciation model and rescaled using the linear rescale of model 1225: y=0.33x-3.05.

8. DFT Functional Dependence

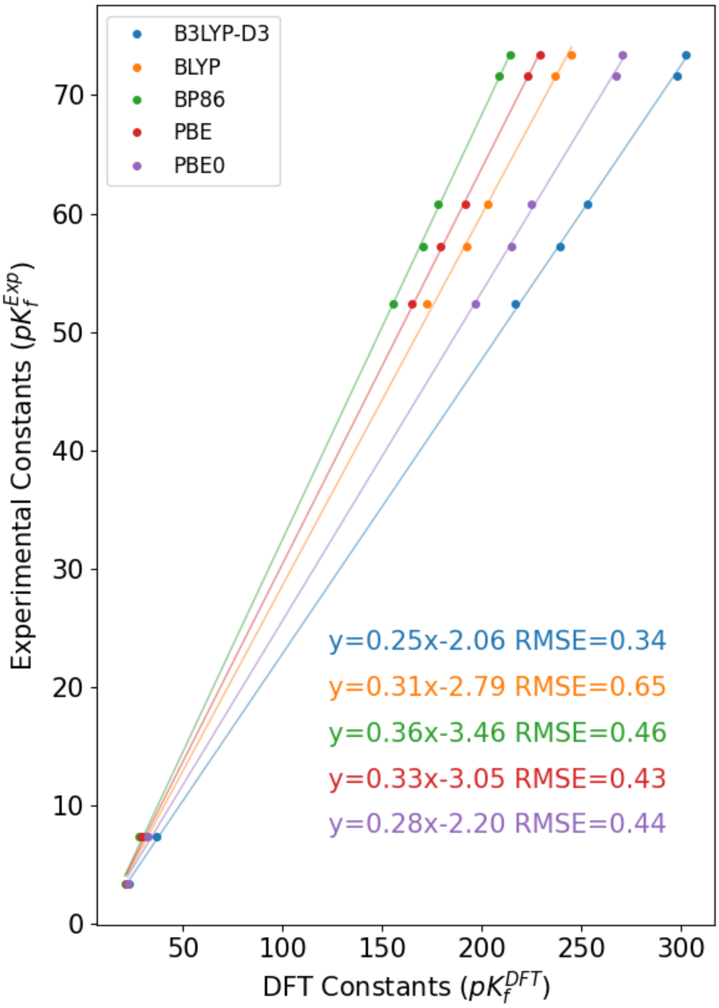


Fig. S4 Best linear regressions between pK_f^{Exp} vs pK_f^{DFT} depending on the DFT functional employed. Color scheme: blue (B3LYP-D3), orange (BLYP), green (BP86), red (PBE) and purple (PBE0).

9. Ionic Strength Dependence

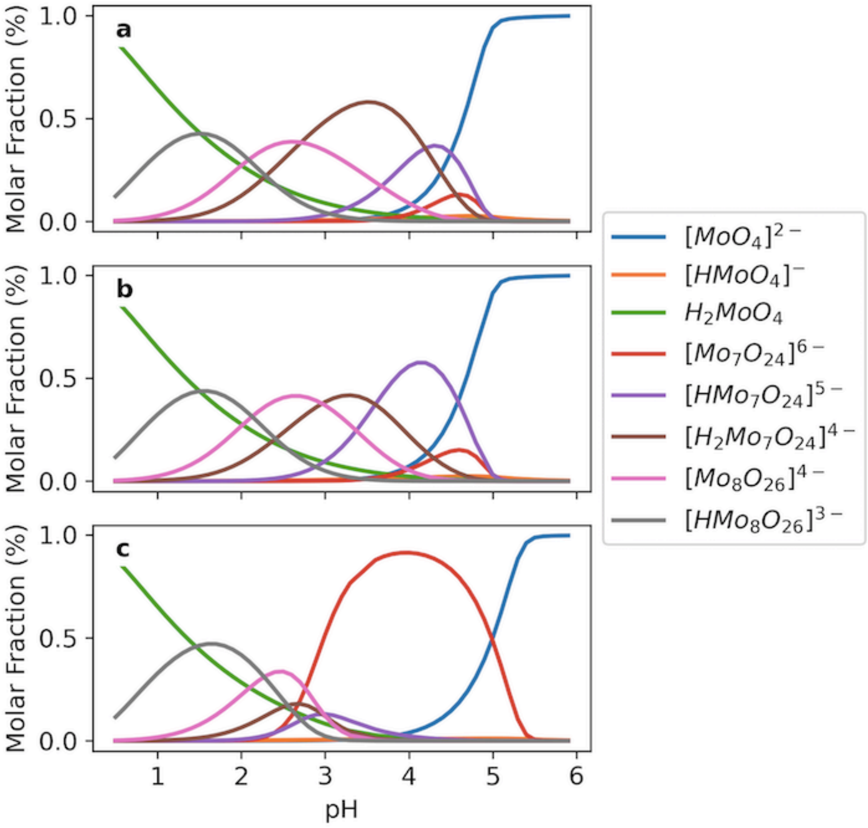


Fig. S5 Speciation diagrams for model 790 at different ionic strengths (a) 0 M, (b) 0.01 M, (c) 0.5 M.

10. *Ad-Hoc* Correction

We propose the addition of an *ad-hoc* correction to the free energy of the molybdic acid (δG) given its clearly underestimated concentration showed in Figure 3d. The problematic stemmed from the uncertainty of its exact structure in aqueous solution even though it has been a matter of discussion in the recent decades.[13–16] It is well-known that a decrease of pH causes the expansion of the molybdenum coordination sphere thus disrupting the initial tetrahedral disposition. The nucleophilic attacks of water molecules lead to the formation of six-coordinated structures such as: $Mo(OH)_6$, $MoO_2(OH)_2(OH_2)_2$ and $MoO_3(OH_2)_3$. Static DFT calculations were performed for all the 3 molybdic acid structures and the results are collected in Table S1. Assuming that the formation of $Mo(OH)_6$, $MoO_2(OH)_2(OH_2)_2$ and $MoO_3(OH_2)_3$ are the products of the reaction between two water molecules with $MoO_2(OH)_2$, the following reactions can be set up:

$MoO_2(OH)_2 + 2H_2O \rightarrow Mo(OH)_6$ $\Delta G_1 = +49.63\ kcal{\cdot}mol^{-1}$

$MoO_2(OH)_2 + 2H_2O \rightarrow MoO_3(OH_2)_3$ $\Delta G_2 = +22.11\ kcal{\cdot}mol^{-1}$

$MoO_2(OH)_2 + 2H_2O \rightarrow MoO_2(OH)_2(OH_2)_2$ $\Delta G_3 = +18.07\ kcal{\cdot}mol^{-1}$

None of the ΔG expressed above indicated that the formation of the six-coordinate species is spontaneous. At first sight, it can be attributed to the entropic penalty as three moles of reactant are transformed into one mole of product. If we settled the same reactions but using a cluster of two waters instead, the ΔG_1, ΔG_2 and ΔG_3 would decrease to 45.66, 18.14 and 14.10 $kcal{\cdot}mol^{-1}$, still remaining strongly endergonic. Rather, the source of error is related to the poor description of the implicit solvent models when accounting for explicit aqua ligands.[17–19] Alternatively, we performed a static DFT calculations but with explicit water clusters instead (free energies collected in Table S1) as Figure S6 shows. The corresponding reactions energies were 63.47, 12.80 and 105.36 $kcal{\cdot}mol^{-1}$

for ΔG_1, ΔG_2 and ΔG_3 respectively. Therefore, adding explicit water molecules did not lead to any enhancement, which is in agreement with the recent work by Zhang, N et al.[12]

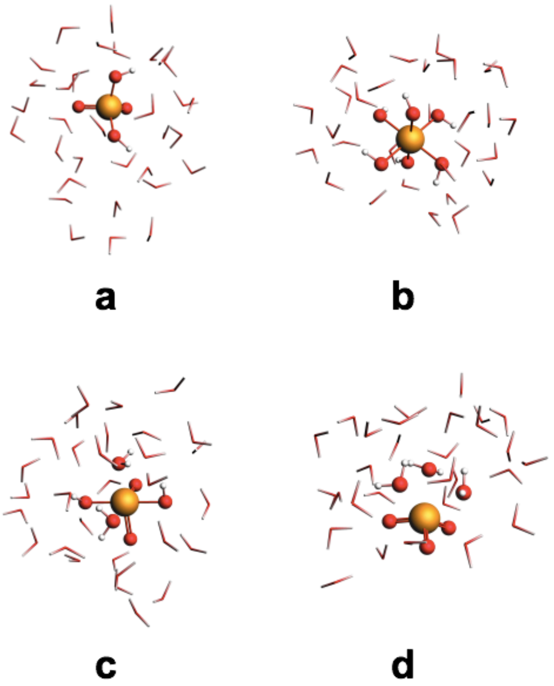


Fig. S6 Water clusters for molybdic acid isomers. (a) H_2MoO_4 with 30 explicit H_2O, (b) $Mo(OH)_6$, (c) $MoO_2(OH)_2(OH_2)_2$ and (d) $MoO_3(OH_2)_3$ with 32 explicit H_2O respectively.

Therefore, we relied on the Car-Parrinello molecular dynamic (CPMD) results reported by Vilà-Nadal et al.[14] where they observed that the six-coordinated acid was 6 kcal·mol^{-1} more stable than the tetrahedral. We applied an *ad-hoc* correction to the free energy of the molybdic acid ($\delta G = -9$ kcal·mol^{-1}) and we recalculated the 1620 speciation models. Consequently, the description of the speciation diagrams significantly improved as Figure 3e and Figure 4b show. As a matter of fact, the RMSE between the parametrized model and the experiments decreased 0.34 pK units. Nonetheless, the addition of *ad-hoc* constants must be dealt with especial attention since they complicate the model and might cause alternative problems. This is why, for the sake of completeness, we screened the *ad-hoc* constant between 0 and 20 kcal·mol^{-1} so as to observe its overall behavior. Figure S6 collects the lowest RMSE value of the best speciation model as function of the correction free energy. The bar marked in purple corresponds to the chosen constant in this work while the rest have been depicted in grey. It is worth noting the parabolic trend in Figure S7 which corroborates the fact that static DFT underestimates the free energy of the molybdic acid systematically.

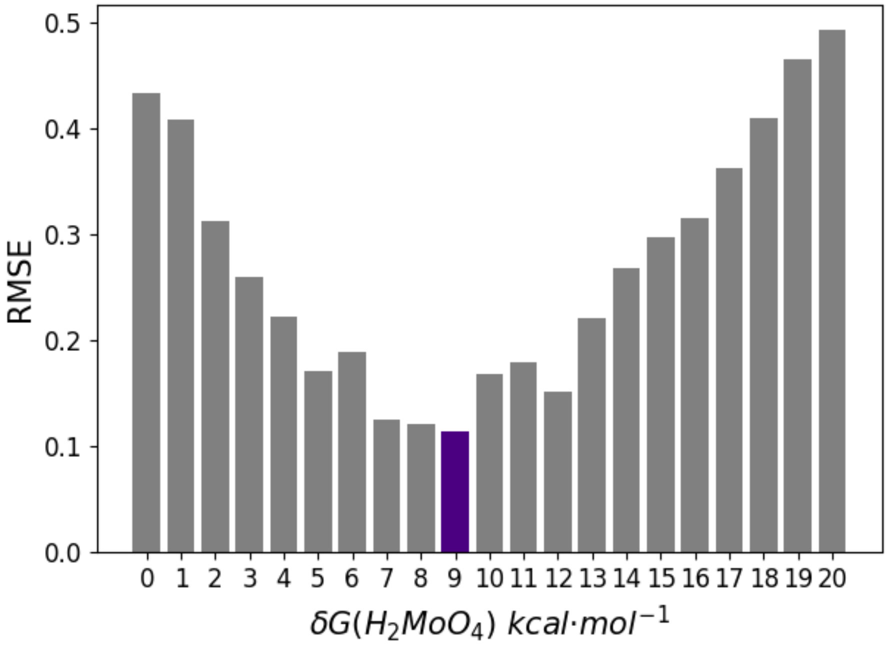


Fig. S7 Bar plot of the RMSE for the best speciation models as function of the ad-hoc correction, $\delta G(H_2MoO_4)$. Bar marked in magenta correspond constant applied to Figure 3e and Figure 4b.

Table S1. Gibbs free energies of molybdic acids computed at level PBE/TZP.

	Gibbs Energy ($kcal{\cdot}mol^{-1}$)
H_2O	-330.29
$2H_2O$	-656.61
$MoO_2(OH)_2(OH_2)_2$	-1059.56
$Mo(OH)_6$	-1670.51
$MoO_3(H_2O)_3$	-1698.03
$MoO_2(OH)_2(OH_2)_4$	-1702.07
$MoO_2(OH)_2(OH_2)_2{\cdot}32H_2O$	-11438.86
$Mo(OH)_6{\cdot}30H_2O$	-11375.39
$MoO_3{\cdot}(H_2O)_3{\cdot}30H_2O$	-11426.06

$MoO_2(OH)_2(OH_2)_4 \cdot 30H_2O$	-11333.50

Bibliography

1 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.

2 J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822–8824.

3 J. P. Perdew, *Phys. Rev. B*, 1986, **34**, 7406–7406.

4 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

5 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

6 A. D. Becke, *J. Chem. Phys.*, 1988, **38**, 3098–3100.

7 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

8 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, , DOI:10.1063/1.3382344.

9 E. Van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys*, 1993, **99**, 4597–4610.

10 E. Van Lenthe and E. J. Baerends, *J. Comput. Chem*, 2003, **24**, 1142–1156.

11 A. J. Klamt, *J. Phys. Chem.*, 1995, **99**, 2224–2235.

12 C. . Àlvarez, M.; M.; Graaf, C.; López, N.; Maseras, F.; Poblet, M, J.; Bo, *J. Chem. Inf. Model.*, 2015, **55**, 95–103.

13 O. F. Oyerinde, C. L. Weeks, A. D. Anbar and T. G. Spiro, *Inorganica Chim. Acta*, 2008, **361**, 1000–1007.

14 L. Vilà-Nadal, E. F. Wilson, H. N. Miras, A. Rodríguez-Fortea, L. Cronin and J. M. Poblet, *Inorg. Chem.*, 2011, **50**, 7811–7819.

15 X. Liu, J. Cheng, M. Sprik and X. Lu, *J. Phys. Chem. Lett.*, 2013, **4**, 2926–2930.

16 N. Zhang, E. Königsberger, S. Duan, K. Lin, H. Yi, D. Zeng, Z. Zhao and G. Hefter, *J. Phys. Chem. B*, 2019, **123**, 3304–3311.

17 X. López, J. A. Fernández, S. Romo, J. F. Paul, L. Kazansky and J. M. Poblet, *J. Comput. Chem.*, 2004, **25**, 1542–1549.

18 S. Romo, J. A. Fernández, J. M. Maestre, B. Keita, L. Nadjo, C. De Graaf and J. M. Poblet, *Inorg. Chem.*, 2007, **46**, 4022–4027.

19 A. Rodríguez-Fortea, L. Vilà-Nadal and J. M. Poblet, *Inorg. Chem.*, 2008, **47**, 7745–7750.