Supporting Information

Nucleation Mechanisms and Speciation of Metal Oxide Clusters

Enric Petrus, † Mireia Segado, † Carles Bo*, †, ‡

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

1. Computational Details ........................................................................................................... 2
2. Code Overview ................................................................................................................... 3
3. Chemical Reactions ............................................................................................................ 5
   3.1. Reactions in Figure 1 .................................................................................................... 5
   3.2 Reactions in Figure 6 .................................................................................................... 7
4. Speciation Model Definition .............................................................................................. 8
5. Binomial Coefficient and Computational Cost .................................................................... 9
6. Formation Constant Definition ........................................................................................ 10
7. Parametrized Formation Constants ............................................................................... 11
8. DFT Functional Dependence ......................................................................................... 12
9. Ionic Strength Dependence ............................................................................................ 13
10. Ad-hoc correction ............................................................................................................ 14
Bibliography ......................................................................................................................... 18
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 PBE0,⁴ BLYP,⁵ B3LYP-D3⁷ and BP86.³,⁶ 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 radiis 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 repository¹² and can be accessed via:

https://iochem-bd.iciq.es/browse-review-collection/100/24719/8650a4b8f150d4055850764.
2. Code Overview

Scheme S1. Summary of the tasks implemented in the POMSimulator. Orange box indicates tasks performed by the python code developed by the authors. Dashed boxes collect relevant outputs (reaction maps and speciation models) whereas bold boxes correspond to important tasks.

Scheme S1 collects a detailed sequences of steps which summarize the protocol followed in this work:

1) A dataset of 30 polyoxomolybdates (POMos) cartesian files was prepared (see top Figure 1).

2) Using the ADF software package we performed geometry optimizations, analytical frequencies and Bader topological analysis.

3) Atom types, chemical bonds depicted by Bader’s Theory and free energies were parsed from the 30 ADF text output files.
4) Each POMos was converted to a molecular graphs using NetworkX python library.\textsuperscript{11} Atoms, bonds and free energies were converted to nodes, edges and attributes respectively.

5) We mapped the isomorphism property for all the 30 molecular graphs.

6) Isomorphic relationships were converted to chemical reactions according to the missing stoichiometry of the two isomorphically-related graphs. 72 chemical reactions were obtained: 20 acid-base, 23 condensations and 29 additions.

7) We set up $\sim 10^8$ speciation models (more information in Section 5).

8) Two approximations were applied (i) speciation models with equations representing all the nuclearities (ii) all the condensation and addition reactions had to consider the monomer, $[H_mMoO_4]^{2-m}$, as a reactant. Thus, the computational cost was reduced to 1620 speciation models.

9) Models were solved using the parameters already defined in the principal manuscript.

10) Representation of the octamolybdate $[H_nMo_8O_{26}]^{n-4}$ formation mechanism as a reaction map of the best speciation model.
3. Chemical Reactions

3.1. Reactions in Figure 1

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

\[
\begin{align*}
\text{MoO}_4^{2-} + H_5O_2^+ & \rightarrow \text{HMoO}_4^- + 2H_2O \quad \Delta G = -29.4 \\
\text{HMoO}_4^- + H_5O_2^+ & \rightarrow \text{H}_2\text{MoO}_4 + 2H_2O \quad \Delta G = -9.83 \\
\text{HMoO}_4^- + \text{HMoO}_4^- & \rightarrow \text{Mo}_2O_7^{2-} + H_2O \quad \Delta G = -1.8 \\
\text{HMoO}_4^- + \text{H}_2\text{MoO}_4 & \rightarrow \text{HMoO}_7^- + H_2O \quad \Delta G = -4.92 \\
\text{Mo}_2O_7^{2-} + H_5O_2^+ & \rightarrow \text{H}_2\text{MoO}_7^- + 2H_2O \quad \Delta G = -12.95 \\
\text{H}_2\text{MoO}_4 + \text{H}_2\text{MoO}_4 & \rightarrow \text{H}_2\text{MoO}_7^- + H_2O \quad \Delta G = -0.9 \\
\text{H}_2\text{MoO}_7^- + H_5O_2^+ & \rightarrow \text{H}_2\text{MoO}_7^- + 2H_2O \quad \Delta G = 5.81 \\
\text{HMoO}_4^- + \text{H}_2\text{MoO}_7^- & \rightarrow \text{Mo}_3O_{10}^{2-} + H_2O \quad \Delta G = -2.51 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_2O_7^{2-} & \rightarrow \text{Mo}_3O_{10}^{2-} + H_2O \quad \Delta G = -5.63 \\
\text{HMoO}_4^- + \text{H}_2\text{MoO}_7^- & \rightarrow \text{HMo}_3O_{10}^- + H_2O \quad \Delta G = -4.87 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_2O_7^- & \rightarrow \text{HMO}_3O_{10}^{2-} + H_2O \quad \Delta G = -0.85 \\
\text{Mo}_3O_{10}^{2-} + H_5O_2^+ & \rightarrow \text{HMO}_3O_{10}^- + 2H_2O \quad \Delta G = -8.17 \\
\text{H}_2\text{MoO}_4 + \text{H}_2\text{MO}_2O_7 & \rightarrow \text{H}_2\text{MO}_3O_{10}^- + H_2O \quad \Delta G = -0.46 \\
\text{HMO}_3O_{10}^- + H_5O_2^+ & \rightarrow \text{H}_2\text{MO}_3O_{10} + 2H_2O \quad \Delta G = -5.42 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_3O_{10}^{2-} & \rightarrow \text{Mo}_4O_{13}^{2-} + H_2O \quad \Delta G = -9.85 \\
\text{Mo}_3O_{10}^{2-} + \text{H}_2\text{MoO}_4 & \rightarrow \text{Mo}_4O_{13}^{2-} + H_2O \quad \Delta G = -14.58 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_2O_7^- & \rightarrow \text{Mo}_4O_{13}^{2-} + H_2O \quad \Delta G = -7.44 \\
\text{H}_2\text{MO}_4 + \text{HMO}_3O_{10}^- & \rightarrow \text{HMO}_4O_{13}^- + H_2O \quad \Delta G = -1.17 \\
\text{HMO}_4O_{13}^- + \text{H}_2\text{MoO}_4 & \rightarrow \text{HMO}_4O_{13}^- + 2H_2O \quad \Delta G = -1.12 \\
\text{Mo}_4O_{13}^{2-} + H_5O_2^+ & \rightarrow \text{HMO}_4O_{13}^- + 2H_2O \quad \Delta G = 0.51 \\
\text{H}_2\text{MoO}_4 + \text{H}_2\text{MO}_3O_{10} & \rightarrow \text{H}_2\text{MO}_4O_{13} + H_2O \quad \Delta G = 8.7 \\
\text{H}_2\text{MO}_2O_7 + \text{H}_2\text{MO}_3O_{10} & \rightarrow \text{H}_2\text{MO}_4O_{13} + H_2O \quad \Delta G = 9.14 \\
\text{HMO}_4O_{13}^- + H_5O_2^+ & \rightarrow \text{H}_2\text{MoO}_4O_{13} + 2H_2O \quad \Delta G = 4.45 \\
\text{H}_2\text{MO}_4 + \text{Mo}_4O_{13}^{2-} & \rightarrow \text{Mo}_5O_{16}^{2-} + H_2O \quad \Delta G = 18.74 \\
\text{H}_2\text{MoO}_4 + \text{HMO}_3O_{10}^- & \rightarrow \text{Mo}_5O_{16}^{2-} + H_2O \quad \Delta G = 12.15 \\
\text{H}_2\text{MO}_2O_7 + \text{Mo}_5O_{16}^{2-} & \rightarrow \text{Mo}_6O_{16}^{2-} + H_2O \quad \Delta G = 9.79 \\
\text{H}_2\text{MoO}_4 + \text{HMO}_3O_{10}^- & \rightarrow \text{H}_2\text{MO}_5O_{16}^- + H_2O \quad \Delta G = 4.41 \\
\text{Mo}_5O_{16}^{2-} + H_5O_2^+ & \rightarrow \text{HMO}_5O_{16}^- + 2H_2O \quad \Delta G = -13.55 \\
\text{H}_2\text{MO}_2O_7 + \text{H}_2\text{MO}_3O_{10} & \rightarrow \text{H}_2\text{MO}_5O_{16} + H_2O \quad \Delta G = 35.65 \\
\text{HMO}_5O_{16}^- + H_5O_2^+ & \rightarrow \text{H}_2\text{MO}_5O_{16} + 2H_2O \quad \Delta G = 25.82 \\
\text{Mo}_4O_2^{2-} + \text{Mo}_4O_{13}^{2-} & \rightarrow \text{Mo}_5O_{17}^{4-} \quad \Delta G = 11.95 \\
\text{Mo}_2O_7^{2-} + \text{Mo}_3O_{10}^{2-} & \rightarrow \text{Mo}_5O_{17}^{4-} \quad \Delta G = 23.47 \\
\text{HMO}_4O_2^- + \text{Mo}_4O_{13}^{2-} & \rightarrow \text{HMO}_5O_{17}^{3-} \quad \Delta G = 13.39 \\
\text{Mo}_2O_7^{2-} + \text{HMO}_3O_{10}^- & \rightarrow \text{HMO}_5O_{17}^{3-} \quad \Delta G = 3.68 \\
\text{HMO}_2O_7^- + \text{HMO}_3O_{10}^- & \rightarrow \text{HMO}_5O_{17}^{3-} \quad \Delta G = 8.46 \\
\text{Mo}_5O_{17}^{4-} + H_5O_2^+ & \rightarrow \text{HMO}_5O_{17}^{3-} + 2H_2O \quad \Delta G = -27.96 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_4O_{13}^{2-} & \rightarrow \text{H}_2\text{Mo}_5O_{17}^{2-} \quad \Delta G = 13.99 
\end{align*}
\]
\[
\begin{align*}
\text{Mo}_2\text{O}_7^{2-} + \text{H}_2\text{Mo}_5\text{O}_{10} & \rightarrow \text{H}_2\text{Mo}_5\text{O}_{17}^{2-} \quad \Delta G = -0.13 \\
\text{HMO}_2\text{O}_7^- + \text{HMO}_3\text{O}_{10}^- & \rightarrow \text{H}_2\text{Mo}_5\text{O}_{17}^{2-} \quad \Delta G = 7.4 \\
\text{H}_2\text{MoO}_2\text{O}_7^- + \text{Mo}_3\text{O}_{10}^{2-} & \rightarrow \text{H}_2\text{Mo}_5\text{O}_{17}^{2-} \quad \Delta G = 5.04 \\
\text{HMO}_5\text{O}_{17}^{3-} + \text{H}_2\text{O} & \rightarrow \text{H}_2\text{Mo}_5\text{O}_{17}^{2-} + 2\text{H}_2\text{O} \quad \Delta G = -9.23 \\
\text{MoO}_4^{2-} + \text{Mo}_5\text{O}_{10}^{2-} & \rightarrow \text{Mo}_6\text{O}_{20}^{4-} \quad \Delta G = -15.63 \\
\text{H}_2\text{MoO}_4^- + \text{Mo}_5\text{O}_{17}^{4-} & \rightarrow \text{Mo}_6\text{O}_{20}^{4-} + \text{H}_2\text{O} \quad \Delta G = -8.84 \\
\text{Mo}_2\text{O}_{10}^{2-} + \text{Mo}_3\text{O}_{10}^{2-} & \rightarrow \text{Mo}_6\text{O}_{20}^{4-} \quad \Delta G = 20.26 \\
\text{HMO}_4^- + \text{Mo}_5\text{O}_{16}^{2-} & \rightarrow \text{HMO}_8\text{O}_{23}^{3-} \quad \Delta G = -0.05 \\
\text{Mo}_5\text{O}_{10}^{2-} + \text{HMO}_3\text{O}_{10}^- & \rightarrow \text{HMO}_6\text{O}_{26}^{3-} \quad \Delta G = 14.61 \\
\text{Mo}_6\text{O}_{20}^{4-} + \text{H}_2\text{O} & \rightarrow \text{HMO}_6\text{O}_{20}^{3-} + 2\text{H}_2\text{O} \quad \Delta G = -13.82 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_5\text{O}_{16}^{2-} & \rightarrow \text{H}_2\text{Mo}_8\text{O}_{26}^{2-} \quad \Delta G = 0.34 \\
\text{Mo}_5\text{O}_{10}^{2-} + \text{H}_2\text{Mo}_3\text{O}_{10}^- & \rightarrow \text{H}_2\text{Mo}_6\text{O}_{26}^{2-} \quad \Delta G = 10.59 \\
\text{HMO}_3\text{O}_{10}^- + \text{HMO}_3\text{O}_{10}^- & \rightarrow \text{H}_2\text{Mo}_6\text{O}_{26}^{2-} \quad \Delta G = 13.34 \\
\text{HMO}_6\text{O}_{26}^{3-} + \text{H}_2\text{O} & \rightarrow \text{H}_2\text{Mo}_6\text{O}_{26}^{2-} + 2\text{H}_2\text{O} \quad \Delta G = -9.44 \\
\text{H}_2\text{MoO}_4 + \text{MoO}_6\text{O}_{20}^{4-} & \rightarrow \text{H}_2\text{Mo}_{12}\text{O}_{34}^{4-} + \text{H}_2\text{O} \quad \Delta G = -6.45 \\
\text{MoO}_4^{2-} + \text{Mo}_4\text{O}_{13}^{2-} & \rightarrow \text{Mo}_7\text{O}_{23}^{4-} \quad \Delta G = 23.66 \\
\text{HMO}_3\text{O}_{10}^- + \text{HMO}_4\text{O}_{13}^- & \rightarrow \text{HMO}_7\text{O}_{23}^{3-} \quad \Delta G = 7.99 \\
\text{HMO}_3\text{O}_{10}^- + \text{MoO}_4\text{O}_{13}^{2-} & \rightarrow \text{HMO}_7\text{O}_{23}^{3-} \quad \Delta G = 16.67 \\
\text{MoO}_6\text{O}_{26}^{4-} + \text{H}_2\text{O} & \rightarrow \text{HMO}_7\text{O}_{23}^{3-} + 2\text{H}_2\text{O} \quad \Delta G = -15.16 \\
\text{HMO}_3\text{O}_{10}^- + \text{HMO}_4\text{O}_{13}^- & \rightarrow \text{H}_2\text{Mo}_7\text{O}_{23}^{2-} \quad \Delta G = 20.7 \\
\text{H}_2\text{Mo}_3\text{O}_{10}^+ + \text{MoO}_4\text{O}_{13}^- & \rightarrow \text{H}_2\text{Mo}_7\text{O}_{23}^{2-} \quad \Delta G = 26.63 \\
\text{HMO}_7\text{O}_{23}^{3-} + \text{H}_2\text{O} & \rightarrow \text{H}_2\text{Mo}_7\text{O}_{23}^{2-} + 2\text{H}_2\text{O} \quad \Delta G = 4.54 \\
\text{MoO}_4^{2-} + \text{Mo}_6\text{O}_{20}^{4-} & \rightarrow \text{Mo}_7\text{O}_{24}^{6-} \quad \Delta G = 25.36 \\
\text{MoO}_2\text{O}_7^- + \text{Mo}_5\text{O}_{17}^{4-} & \rightarrow \text{Mo}_7\text{O}_{24}^{6-} \quad \Delta G = 37.89 \\
\text{HMO}_4^- + \text{Mo}_6\text{O}_{20}^{4-} & \rightarrow \text{HMO}_7\text{O}_{24}^{6-} \quad \Delta G = 26.51 \\
\text{MoO}_2\text{O}_7^- + \text{HMO}_5\text{O}_{17}^{3-} & \rightarrow \text{HMO}_7\text{O}_{24}^{5-} \quad \Delta G = 37.6 \\
\text{MoO}_2\text{O}_7^- + \text{H}_2\text{O} & \rightarrow \text{HMO}_7\text{O}_{24}^{5-} + 2\text{H}_2\text{O} \quad \Delta G = -28.25 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_6\text{O}_{20}^{4-} & \rightarrow 2\text{H}_2\text{Mo}_7\text{O}_{24}^{4-} \quad \Delta G = 18.52 \\
\text{HMO}_7\text{O}_{24}^{5-} + \text{H}_2\text{O} & \rightarrow \text{H}_2\text{Mo}_7\text{O}_{24}^{4-} + 2\text{H}_2\text{O} \quad \Delta G = -17.82 \\
\text{H}_2\text{MoO}_4 + \text{Mo}_7\text{O}_{23}^{4-} & \rightarrow \text{Mo}_8\text{O}_{26}^{4-} + \text{H}_2\text{O} \quad \Delta G = 11.14 \\
\text{HMO}_3\text{O}_{12}^2^- + \text{Mo}_4\text{O}_{13}^{2-} & \rightarrow \text{Mo}_8\text{O}_{26}^{4-} \quad \Delta G = 22.37 \\
\text{MoO}_2\text{O}_7^- + \text{H}_2\text{O} & \rightarrow \text{HMO}_8\text{O}_{26}^{3-} + 2\text{H}_2\text{O} \quad \Delta G = -9.31 \\
\text{Mo}_3\text{O}_{13}^2^- + \text{H}_2\text{MoO}_4\text{O}_{13}^- & \rightarrow \text{H}_2\text{Mo}_8\text{O}_{26}^{2-} \quad \Delta G = 8.2 \\
\text{HMO}_4\text{O}_{13}^- + \text{HMO}_4\text{O}_{13}^- & \rightarrow \text{H}_2\text{Mo}_8\text{O}_{26}^{2-} \quad \Delta G = 12.14
\end{align*}
\]
3.2 Reactions in Figure 6

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.

\[
\begin{align*}
\text{MoO}_{4}^{2−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMOO}_{4}^{−} + 2\text{H}_{2}\text{O} \quad ΔG=−29.4 \\
\text{HMOO}_{4}^{−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{4}^{2−} + 2\text{H}_{2}\text{O} \quad ΔG=−18.83 \\
\text{HMOO}_{4}^{−} + \text{H}_{2}\text{MoO}_{4} & \rightarrow \text{HMO}_{2}\text{O}_{7}^{−} + \text{H}_{2}\text{O} \quad ΔG=4.08 \\
\text{MoO}_{2}\text{O}_{7}^{2−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{2}\text{O}_{7}^{−} + 2\text{H}_{2}\text{O} \quad ΔG=−12.95 \\
\text{HMO}_{2}\text{O}_{7}^{−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{2}\text{O}_{7}^{−} + 2\text{H}_{2}\text{O} \quad ΔG=−5.81 \\
\text{H}_{2}\text{MoO}_{4} + \text{HMO}_{2}\text{O}_{7}^{−} & \rightarrow \text{HMO}_{3}\text{O}_{10}^{−} + \text{H}_{2}\text{O} \quad ΔG=8.15 \\
\text{MoO}_{3}\text{O}_{10}^{2−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{3}\text{O}_{10}^{−} + 2\text{H}_{2}\text{O} \quad ΔG=−8.17 \\
\text{MoO}_{3}\text{O}_{10}^{−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{3}\text{O}_{10} + 2\text{H}_{2}\text{O} \quad ΔG=−5.42 \\
\text{H}_{2}\text{MoO}_{4} + \text{HMO}_{3}\text{O}_{10}^{−} & \rightarrow \text{HMO}_{4}\text{O}_{13}^{−} + \text{H}_{2}\text{O} \quad ΔG=7.83 \\
\text{MoO}_{4}\text{O}_{13}^{2−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{4}\text{O}_{13}^{−} + 2\text{H}_{2}\text{O} \quad ΔG=0.51 \\
\text{HMO}_{4}\text{O}_{13}^{−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{4}\text{O}_{13} + 2\text{H}_{2}\text{O} \quad ΔG=4.45 \\
\text{H}_{2}\text{MoO}_{4} + \text{MoO}_{4}\text{O}_{13}^{2−} & \rightarrow \text{MoO}_{5}\text{O}_{16}^{2−} + \text{H}_{2}\text{O} \quad ΔG=27.74 \\
\text{MoO}_{5}\text{O}_{16}^{2−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{5}\text{O}_{16}^{−} + 2\text{H}_{2}\text{O} \quad ΔG=−13.55 \\
\text{HMO}_{5}\text{O}_{16}^{−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{5}\text{O}_{16} + 2\text{H}_{2}\text{O} \quad ΔG=25.82 \\
\text{HMO}_{5}\text{O}_{16}^{−} + \text{MoO}_{4}\text{O}_{13}^{2−} & \rightarrow \text{HMO}_{5}\text{O}_{17}^{3−} \quad ΔG=13.39 \\
\text{MoO}_{5}\text{O}_{17}^{4−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{5}\text{O}_{17}^{3−} + 2\text{H}_{2}\text{O} \quad ΔG=−27.96 \\
\text{HMO}_{5}\text{O}_{17}^{3−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{5}\text{O}_{17}^{2−} + 2\text{H}_{2}\text{O} \quad ΔG=−9.23 \\
\text{HMO}_{5}\text{O}_{17}^{3−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{5}\text{O}_{17}^{2−} + 2\text{H}_{2}\text{O} \quad ΔG=−9.23 \\
\text{MoO}_{6}\text{O}_{20}^{4−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{6}\text{O}_{20}^{3−} + 2\text{H}_{2}\text{O} \quad ΔG=−13.82 \\
\text{MoO}_{6}\text{O}_{20}^{3−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{6}\text{O}_{20}^{2−} + 2\text{H}_{2}\text{O} \quad ΔG=−9.44 \\
\text{H}_{2}\text{MoO}_{4} + \text{MoO}_{6}\text{O}_{20}^{3−} & \rightarrow \text{Mo}_{7}\text{O}_{23}^{4−} + \text{H}_{2}\text{O} \quad ΔG=2.55 \\
\text{Mo}_{7}\text{O}_{23}^{4−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{7}\text{O}_{23}^{3−} + 2\text{H}_{2}\text{O} \quad ΔG=−15.16 \\
\text{MoO}_{7}\text{O}_{23}^{3−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{7}\text{O}_{23}^{2−} + 2\text{H}_{2}\text{O} \quad ΔG=4.54 \\
\text{MoO}_{8}\text{O}_{24}^{5−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{8}\text{O}_{24}^{4−} + 2\text{H}_{2}\text{O} \quad ΔG=−28.25 \\
\text{H}_{2}\text{MoO}_{4} + \text{MoO}_{8}\text{O}_{24}^{4−} & \rightarrow \text{H}_{2}\text{MoO}_{7}\text{O}_{24}^{4−} \quad ΔG=27.52 \\
\text{HMO}_{7}\text{O}_{24}^{5−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{7}\text{O}_{24}^{4−} + 2\text{H}_{2}\text{O} \quad ΔG=−17.82 \\
\text{H}_{2}\text{MoO}_{4} + \text{Mo}_{7}\text{O}_{23}^{4−} & \rightarrow \text{MoO}_{8}\text{O}_{26}^{4−} + \text{H}_{2}\text{O} \quad ΔG=−2.14 \\
\text{MoO}_{8}\text{O}_{26}^{3−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{HMO}_{8}\text{O}_{26}^{3−} + 2\text{H}_{2}\text{O} \quad ΔG=−9.31 \\
\text{HMO}_{8}\text{O}_{26}^{3−} + \text{H}_{5}\text{O}_{2}^{+} & \rightarrow \text{H}_{2}\text{MoO}_{8}\text{O}_{26}^{2−} + 2\text{H}_{2}\text{O} \quad ΔG=0.1
\end{align*}
\]
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.

\[ \prod [\text{Products}] - e^{-\Delta G_{\text{free}}/RT} \cdot \prod [\text{Reactants}] = 0 \]  \hspace{1cm} \text{Eq. S1}

\[ \sum_i [\text{Compound}]_i - [0.005 \text{M}] = 0 \]  \hspace{1cm} \text{Eq. S2}

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

\[ \text{Activity} = [\text{Compound}]_i \cdot \gamma_i = [\text{Compound}]_i \cdot 10^{-0.5 \cdot \Delta T \cdot (\sqrt{\frac{\gamma_i}{1+\gamma_i}} - 0.30)} \]  \hspace{1cm} \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).

\[ f = \binom{\text{Reactions}}{9} \]

**Figure S1.** Number of speciation models as a function of the total number of chemical reactions.

Figure S1 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-linear equations scale up to $10^{10}$ which is too computational demanding.
6. Formation Constant Definition

Formation reactions are formulated as follows:

\[ \lambda \text{MoO}_4^{2-} + (2 \cdot \mu) \text{H}^+ \rightarrow \text{Mo}_\lambda\text{O}_{4\cdot\lambda-\mu}^{6\cdot\lambda-2\cdot4\cdot\lambda-\mu} + \mu \text{H}_2\text{O} \]

Hence, the corresponding formation constant is:

\[ pK_f = -\log \left( \frac{[\text{H}_2\text{O}]^{\mu}[\text{Mo}_\lambda\text{O}_{4\cdot\lambda-\mu}]^{6\cdot\lambda-2\cdot4\cdot\lambda-\mu}}{[\text{MoO}_4]^\lambda[\text{H}^+]^{2\mu}} \right) \quad \text{Eq. S4} \]
7. Parametrized Formation Constants

Figure S2. Box plot for the parametrized formation constants (pK_{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

Figure S3. Best linear regressions between pK$_{\text{Exp}}$ vs pK$_{\text{DFT}}$ depending on the DFT functional employed. Color scheme: blue (B3LYP-D3), orange (BLYP), green (BP86), red (PBE) and purple (PBE0).
9. Ionic Strength Dependence

Figure 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.\(^\text{12-15}\) 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:

\[
\begin{align*}
\text{MoO}_2(\text{OH})_2 + 2\text{H}_2\text{O} & \rightarrow \text{Mo(OH)}_6 & \Delta G_1 &= +49.63 \text{ kcal}\cdot\text{mol}^{-1} \\
\text{MoO}_2(\text{OH})_2 + 2\text{H}_2\text{O} & \rightarrow \text{MoO}_2(\text{OH})_2(\text{OH}_2)_2 & \Delta G_2 &= +22.11 \text{ kcal}\cdot\text{mol}^{-1} \\
\text{MoO}_2(\text{OH})_2 + 2\text{H}_2\text{O} & \rightarrow \text{MoO}_3(\text{OH}_2)_2 & \Delta G_3 &= +18.07 \text{ kcal}\cdot\text{mol}^{-1}
\end{align*}
\]

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 ΔG₁, ΔG₂ and ΔG₃ would decrease to 45.66, 18.14 and 14.10 kcal·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.\(^\text{16-18}\) Alternatively, we performed a static DFT calculations but with explicit water clusters instead (free energies collected in Table S1) as Figure S5 shows. The corresponding reactions energies were 63.47, 12.80 and 105.36 kcal·mol⁻¹ for ΔG₁, ΔG₂ and...
Therefore, adding explicit water molecules did not lead to any enhancement, which is in agreement with the recent work by Zhang, N et al.\textsuperscript{12}

\[
\Delta G_3 \text{ respectively.}
\]

Figure S5. Water clusters for molybdic acid isomers. (a) H\textsubscript{2}MoO\textsubscript{4} with 30 explicit H\textsubscript{2}O, (b) Mo(OH)\textsubscript{6}, (c) MoO\textsubscript{2}(OH)\textsubscript{2}(OH\textsubscript{2})\textsubscript{2} and (d) MoO\textsubscript{3}(OH\textsubscript{2})\textsubscript{3} with 32 explicit H\textsubscript{2}O respectively.

Therefore, we relied on the Car-Parrinello molecular dynamic (CPMD) results reported by Vilà-Nadal et al.\textsuperscript{13} where they observed that the six-coordinated acid was 6 kcal·mol\textsuperscript{-1} more stable than the tetrahedral. We applied an \textit{ad-hoc} correction to the free energy of the molybdic acid (δG=9 kcal·mol\textsuperscript{-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 \textit{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 \textit{ad-hoc} constant between 0 and 20 kcal·mol\textsuperscript{-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 S6 which corroborates the fact that static DFT underestimates the free energy of the molybdic acid systematically.

![Graph](image)

**Figure S6.** 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·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                     |
| $MoO_2(OH)_{2}(OH_2)_2$·$32H_2O$ | -11438.86                  |
| $Mo(OH)_6$·$30H_2O$  | -11375.39                     |
| Compound                                      | Value       |
|----------------------------------------------|-------------|
| MoO$_3$·(H$_2$O)$_3$·30H$_2$O                | -11426.06   |
| MoO$_2$(OH)$_2$(OH)$_2$·30H$_2$O              | -11333.50   |
Bibliography

(1) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* 2001, 22 (9), 931–967.

(2) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* 1986, 33 (12), 8822–8824.

(3) Perdew, J. P. Erratum: Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* 1986, 34 (10), 7406–7406.

(4) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* 1999, 110 (13), 6158–6170.

(5) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* 1988, 37 (2), 785–789.

(6) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *J. Chem. Phys.* 1988, 38, 3098–3100.

(7) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* 1993, 98, 5648–5652.

(8) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* 2010, 132 (15).

(9) Van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two-component Hamiltonians. *J. Comput. Chem.* 1993, 99 (6), 4597–4610.

(10) Van Lenthe, E.; Baerends, E. J. Optimized Slater-type Basis Sets for the Elements 1–118. *J. Comput. Chem* 2003, 24 (9), 1142–1156.

(11) Klamt, A. J. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* 1995, 99 (7), 2224–2235.

(12) Oyerinde, O. F.; Weeks, C. L.; Anbar, A. D.; Spiro, T. G. Solution Structure of Molybdic Acid from Raman Spectroscopy and DFT Analysis. *Inorganica Chim. Acta* 2008, 361 (4), 1000–1007.

(13) Vilà-Nadal, L.; Wilson, E. F.; Miras, H. N.; Rodríguez-Fortea, A.; Cronin, L.; Poblet, J. M. Combined Theoretical and Mass Spectrometry Study of the Formation-Fragmentation of Small Polyoxomolybdates. *Inorg. Chem.* 2011, 50 (16), 7811–7819.

(14) Liu, X.; Cheng, J.; Sprik, M.; Lu, X. Solution Structures and Acidity Constants of Molybdic Acid. *J. Phys. Chem. Lett.* 2013, 4 (17), 2926–2930.

(15) Zhang, N.; Königsberger, E.; Duan, S.; Lin, K.; Yi, H.; Zeng, D.; Zhao, Z.; Hefter, G. Nature of Monomeric Molybdenum(VI) Cations in Acid Solutions Using Theoretical Calculations and Raman Spectroscopy. *J. Phys. Chem. B* 2019, 123 (15), 3304–3311.

(16) López, X.; Fernández, J. A.; Romo, S.; Paul, J. F.; Kazansky, L.; Poblet, J. M. Are the Solvent Effects Critical in the Modeling of Polyoxoanions? *J. Comput. Chem.*
(17) Romo, S.; Fernández, J. A.; Maestre, J. M.; Keita, B.; Nadj, L.; De Graaf, C.; Poblet, J. M. Density Functional Theory and Ab Initio Study of Electronic and Electrochemistry Properties of the Tetranuclear Sandwich Complex [Fe IIII(II)2(H2O)2(PW9O 34)2]6-. *Inorg. Chem.* **2007**, *46* (10), 4022–4027.

(18) Rodríguez-Forte, A.; Vilà-Nadal, L.; Poblet, J. M. Hydration of Hydrogentungstate Anions at Different PH Conditions: A Car-Parrinello Molecular Dynamics Study. *Inorg. Chem.* **2008**, *47* (17), 7745–7750.