THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCA Article

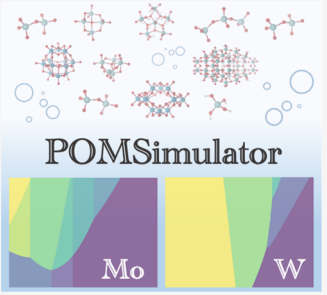
Unlocking Phase Diagrams for Molybdenum and Tungsten Nanoclusters and Prediction of their Formation Constants

Enric Petrus and Carles Bo*

Cite This: J. Phys. Chem. A 2021, 125, 5212−5219 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Understanding and controlling aqueous speciation of metal oxides are key for the discovery and development of novel materials, and challenge both experimental and computational approaches. Here we present a computational method, called POMSimulator, which is able to predict speciation phase diagrams (Conc. vs pH) for multispecies chemical equilibria in solution, and which we apply to molybdenum and tungsten isopolyoxoanions (IPAs). Starting from the MO_4 monomers, and considering dimers, trimers, and larger species, the chemical reaction networks involved in the formation of $[H_{32}Mo_{36}O_{128}]^{8-}$ and $[W_{12}O_{42}]^{12-}$ are sampled in an automatic manner. This information is used for setting up ~10^5 speciation models, and from there, we generate the speciation phase diagrams, which show an insightful picture of the behavior of IPAs in aqueous solution. Furthermore, we predict the values of 107 formation constants for a diversity of molybdenum and tungsten molecular oxides. Among these species, we could include several pentagonal-shaped species and very reactive tungsten intermediates as well. Last but not least, the calibration employed for correcting the density functional theory (DFT) Gibbs energies is remarkably similar for both metals, which suggests that a general rule might exist for correcting computed free energies for other metals.



INTRODUCTION

Self-assembly processes of discrete metal oxide nanoclusters are driven by a combination of many factors, the pH, the ionic force, and the presence of additional species being the main players. Although in this chemistry there are no well-established rules as in organic chemical synthesis,[1] stable anionic metal-oxo compounds have been characterized continuously for almost two centuries. One of the first examples of polyoxometalates (POMs) was molybdenum α-Keggin synthesized by Berzelius in 1826.[2] Since then, POMs have been synthesized with elements throughout the periodic table, thus we can find examples with lanthanides,[3] actinides,[4] and other transition metals.[5] This chemical variety has been accompanied by a broad range of topologies: molybdenum blue wheels,[6,7] keplerates,[8,9] and polyperoxouranates.[10,11] Metal-oxo compounds also present different applications in relevant fields such as biochemistry,[12−15] catalysis,[16−21] and nuclear reprocessing.[22,23]

Hitherto, polyoxometalate chemistry has largely succeeded in discovering new compounds and novel applications. However, it is becoming more and more apparent that a deeper understanding of metal-oxo speciation, and thus of the self-assembly process, is needed.[24] Currently, different techniques are used to study POMs in solution: electronic[25] and vibrational[26] spectroscopy, nuclear magnetic resonance,[27] mass-spectrometry experiments,[28] small-angle X-ray scattering,[29] and kinetic investigations.[30] Despite that, the experimental determination of unstable metal-oxo equilibrium constants is proving to be difficult. To the best of our knowledge, there is no information yet on the thermodynamic constants of pentagonal-based units $\{M_xM_6\}$, which are a recurrent motif for providing curvature and sphericity to metal-oxo nanoclusters.

Speciation of metal oxides in solution is determined by a myriad of chemical equilibria that coexist in the reaction medium. Therefore, the cornerstone for predicting the relative abundance of species is to elucidate the chemical reaction network (CRN)[31,32] along with its dependency on physical conditions: temperature, ionic strength, and concentration. Quantum mechanical calculations are consolidated as an effective and accurate method for predicting reaction mechanisms through the analysis of the potential energy surfaces (PES).[33] In fact, computing static energy profiles based on chemical intuition has been the workhorse of quantum chemistry for the past three decades.[34] For example, a group additivity method, based on ab initio calculations, can describe the aluminum cluster speciation and its Pourbaix diagrams.[35] Alternatively, metadynamics simulations have also achieved great success, for instance, in explaining the aggregation process of the Lindqvist structure in concordance with mass-spectrometry experiments.[36−38] However, both static and dynamic approaches strongly rely on human

Received: April 12, 2021
Revised: May 26, 2021
Published: June 4, 2021

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219

guidance in either locating the transition states (TS) or determining the collective variables. Because of this, the analysis of large-scale PES, purely based on human intuition and in a nonautomated fashion, is time-consuming and error-prone.[39]

The study of coupled reactions requires a meticulous determination of all chemical reactions involved. Indeed, how to build a robust and scalable reaction network is a subject of discussion in the literature.[40,41] First of all, it is important to choose a general-purpose molecular representation for the chemical species. Broadbelt et al. showed that molecular graphs are a straightforward option in which atoms are assigned as nodes and chemical bonds as edges.[42] Moreover, the parallelism between chemistry and graph theory can be further extended to chemical reaction networks.[43] Therefore, graphs provide a unified framework to both represent chemical compounds and ease the management of the CRNs. There are numerous examples of this methodology in very different chemical processes such as prebiotics,[44] hydroformylation,[45] water-shift,[46] and combustion[47] reactions.

We recently introduced the basics of a new methodology aimed at dealing with multispecies equilibria involved in the formation of metal-oxo clusters in aqueous solution based on first-principles and in an automated manner.[48] In the first implementation, we considered a small set of small molybdates (up to octamolybdate) and started by capturing raw results derived from density functional theory (DFT) calculations, which were transformed into molecular graphs based on the topological properties of the charge density. In the next step, our method performed heuristic searches to determine the CRNs connecting all species in the set. Finally, the reaction network was used to set up multiple speciation models, which were solved. We found that the DFT-derived formation constants were systematically overestimated, so we came up with a calibration procedure that enabled constructing the speciation diagram (% species vs pH, at a fixed metal concentration).

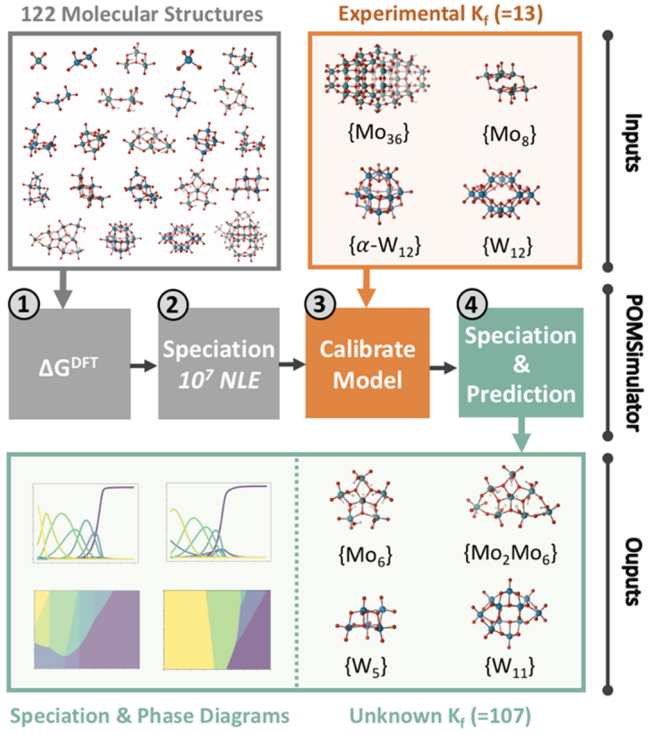
Herein, we present the first full application of our new method, called POMSimulator, toward the in silico prediction of speciation phase diagrams (Conc. vs pH) of molybdenum and tungsten medium-sized isopolyanion (IPA) nanoclusters. On the one hand, we have considered a large set of molybdates, including pentagonal $\{Mo_xMo_6\}$ and medium-sized $\{Mo_{36}\}$ nanoclusters. On the other hand, we have expanded this methodology to tungsten IPAs. A general overview of the workflow followed in this work is depicted in Scheme 1. We start by optimizing the geometries and compute the harmonic frequencies for all of the metal-oxo compounds. Then POMSimulator (1) creates the CRN and estimates the free reaction energy of each transformation, (2) sets up and solves all of the possible speciation models, (3) calibrates the DFT free reaction energies, ΔG^{DFT}, using experimental data, and (4) predicts speciation phase diagrams and not yet reported thermodynamic constants.

RESULTS AND DISCUSSION

We have started describing a class of anionic molecular oxides that contain a unique type of transition metals, also known as IPAs. On the one hand, the molecular set of polyoxomolybdates (POMos) is formed by a total of 72 compounds with the general formulas $[H_zMo_pO_{3p+1}]^{z-2}$, $[H_zMo_pO_{3p+2}]^{z-4}$, and $[H_zMo_pO_{3p+3}]^{z-6}$; pentagonal-based structures $\{Mo_xMo_6\}$; and the largest existing IPAs: $[H_{16+z}\ Mo_{18}O_{65}]^{z-6}$ and $[H_{32+z}Mo_{36}O_{128}]^{z-8}$, abbreviated as $\{Mo_{18}\}$ and $\{Mo_{36}\}$. On the other hand, the set of polyoxotungstates (POTs) is formed from 50 species with the general formulas: $[H_zW_pO_{3p+1}]^{z-2}$, $[H_zW_pO_{3p+2}]^{z-4}$, and $[H_zW_pO_{3p+3}]^{z-6}$. It also includes medium-sized structures, such as $[H_zW_{11}O_{40}]^{z-14}$, $[H_zW_{12}O_{42}]^{z-12}$, and α-$[H_zW_{12}O_{40}]^{z-8}$, abbreviated as $\{W_{11}\}$, $\{W_{12}\}$, and α-$\{W_{12}\}$. We have not included pentagonal structures in tungsten polyoxometalates because there are no isopolyoxoanions that include these motifs. In contrast, polyoxomolybdates show these pentagonal structures within the largest clusters: $\{Mo_{18}\}$ and $\{Mo_{36}\}$. In both groups, multiple protonation states are considered, and a detailed overview of the complete molecular set can be found in Figure S1. The molecular geometries have been optimized using a DFT methodology that is described in the Supporting Information. A dataset collection[49] is available in the ioChem-BD repository.[50]

In the first place, POMSimulator collects the results of the DFT calculations. It starts by capturing the optimized geometries of the metal oxides and transform them into molecular graphs. Atoms are translated to nodes and bond critical points to edges. Additionally, chemical properties such as the total charge and the Gibbs free energy are stored. Once all of the relevant information has been processed, POMSimulator is ready to create a CRN. To do so, the

Scheme 1. General Overview of the POMSimulator Workflow[a]



[a](1) Compute DFT Gibbs energies of metal oxides and determine free energies (195 reactions in total) of the reaction through heuristic searches. (2) Solve systems of nonlinear equations (NLE). Our parallelized code needs 48 h in a 28-core CPU to solve 1×10^7 systems of equations. (3) Calibrate DFT intrinsic error. Eight experimental formation constants (K_f) for molybdenum (e.g., $\{Mo_{36}\}$, $\{Mo_8\}$) and five for tungsten (e.g., $\{W_{12}\}$, $\{\alpha\text{-}W_{12}\}$) were used, respectively. (4) Generate speciation and phase diagrams using corrected DFT energies and predict unreported thermodynamic constants.

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219

isomorphism property is used (i.e., morphological-like graphs) to define the relationships between all of the species. If two graphs are isomorphic, we subtract the atoms between both compounds to determine which chemical reaction could interrelate them: acid−base, condensation, addition, or dimerization.

Following this methodology, we automatically sample large reaction networks with their associated reaction free energies (ΔG^{DFT}). Then, the CRNs are employed to set up the speciation models to predict the behavior of polyoxometalates in solution. However, the larger the size of the CRN, the larger the number of multispecies chemical equilibria (MSCE). Furthermore, the MSCE consists of systems of nonlinear equations, the resolution of which is slow and numerically unstable. Thus, the bottleneck of POMSimulator is the resolution of the speciation models derived from the CRNs. To find a good balance between thoroughness and computational efficiency, we have come up with two solutions. Firstly, we have parallelized the solving step to run POMSimulator in a multicore CPU. Secondly, POMSimulator applies heuristic searches to disregard meaningless reactions, which would otherwise increase the computational cost. Scheme S1 provides a complete overview of the protocol, as well as an exhaustive explanation of the implementation in detail.

Chemical reaction networks scale rapidly with the size of the molecular set. Because of this, CRNs tend to be dense and convoluted, as they contain an extensive number of transformations. One way of analyzing complex networks is to represent them as distributions. In our case, we have examined the thermochemistry of the CRNs employing violin plots. Figure 1A shows eight energy distribution plots, organized according to the type of reaction and IPA metal. Colors green and orange correspond to molybdenum and tungsten systems, respectively. Moreover, plots are also organized according to the reaction type: acid−base, condensation, addition, or dimerization. Note that acid−base equilibria are key reactions for driving the nucleation of metal oxides. For instance, molybdenum and tungsten oxides aggregate at low pH, whereas niobium and tantalum only form large clusters under alkaline conditions. The other three types of reactions handle the growth and disassembly transformations. The median values, depicted in Figure 1A, provide the first estimation of the thermochemistry within the CRNs. For example, the median rates in acid−base, condensation, and dimerization reactions are below 0 kcal·mol^{-1}, in contrast to addition reactions. Thus, acid−base, condensation, and dimerization reactions are mostly spontaneous whereas additions are mostly endergonic. Note that this trend is very similar for both metal-oxo families. Moreover, the distribution plots also provide information about the variances for each reaction type. POMos present a larger variance in condensation-type reactions, whereas POTs show an even greater variance for dimerization-type reactions. Overall, Figure 1A shows only a purely energy-based picture. Therefore, reactions that have a ΔG^{DFT} above 0 kcal·mol^{-1} might turn out to be favorable when pH or ionic strengths are considered. Tables S1 and S2 collect the complete list of chemical reactions, with their associated Gibbs reaction energies and reaction types.

Despite the chemical likeness of molybdenum and tungsten, they have few structures in common. The six alike compounds are the monomers: $[MO_4]^{2-}$, $[HMO_4]^{-}$, H_2MO_4 and the heptamers: $[M_7O_{24}]^{6-}$, $[HM_7O_{24}]^{5-}$, and $[H_2M_7O_{24}]^{4-}$. Figure 1B shows the DFT formation constants of these structures (in box plots). Analogously to Figure 1A, green and orange correspond to the molybdenum and tungsten systems, respectively. Additionally, there is an inset plot that shows the tridimensional geometry of the heptamer $[M_7O_{24}]^{6-}$. The complete set of 122 formation constants, computed with our software, can be found in Figure S2. It is worth noting that molybdates formation constants show a larger variance than tungstates due to the relative size of the molecular sets.

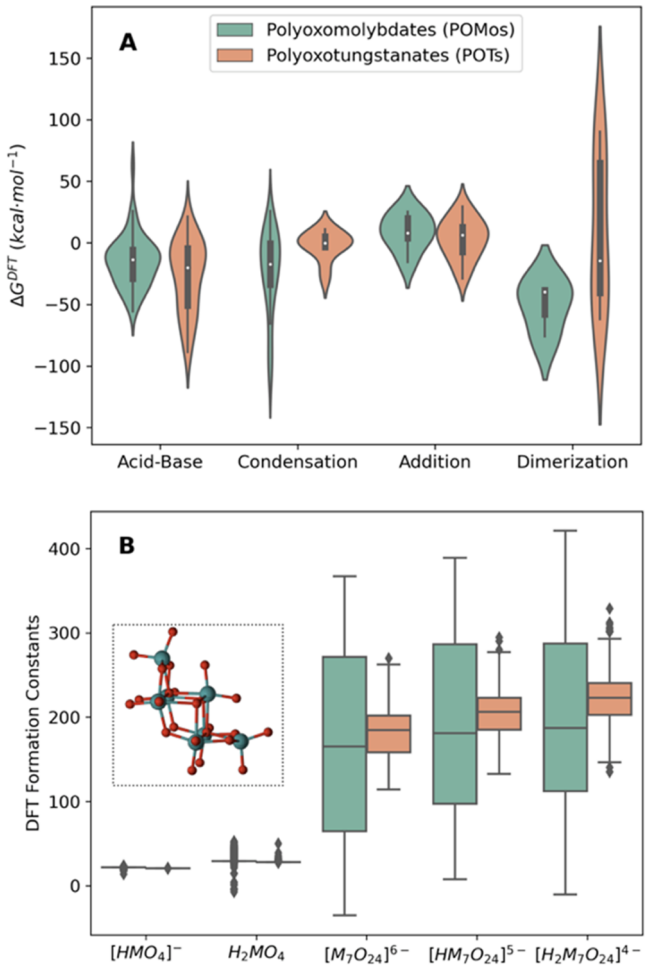


Figure 1. (A) Gibbs energy distribution plot (in kcal·mol^{-1}) for the main types of reactions present in the speciation models: acid−base, condensation, addition, and dimerization. (B) Box plots of the DFT formation constants for the monomer and heptamer. Samples for molybdenum and tungsten constants are of the order of 8×10^4 and 1×10^3, respectively. Inset plot of the $[M_7O_{24}]^{6-}$ tridimensional structure. Constants are referred to the formation reaction: $p[MO_4]^{2-} + qH^+ \rightleftharpoons [H_zM_pO_m]^{n-} + (q/2 + p - z/2)\ H_2O$ Polyoxomolybdates are marked in green and tungstates in orange.

A total of 116 644 speciation models were solved for molybdenum oxides, whereas 1536 models were solved for tungsten oxides. Furthermore, POMSimulator computes an average of 8×10^4 constants for each molybdenum cluster (see Table S3), whereas it computes an order of magnitude less for tungsten compounds (see Table S4). Interestingly, the median values in Figure 1B vary if we compare tungsten and molybdenum oxides. POMSimulator predicts that the median values for $[H_nW_7O_{24}]^{n-6}$ are greater than those for $[H_nMo_7O_{24}]^{n-6}$. The larger the formation constants, the more stable the metal-oxo compounds are. At this stage, our method successfully describes the fact that formation of tungsten oxides

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219

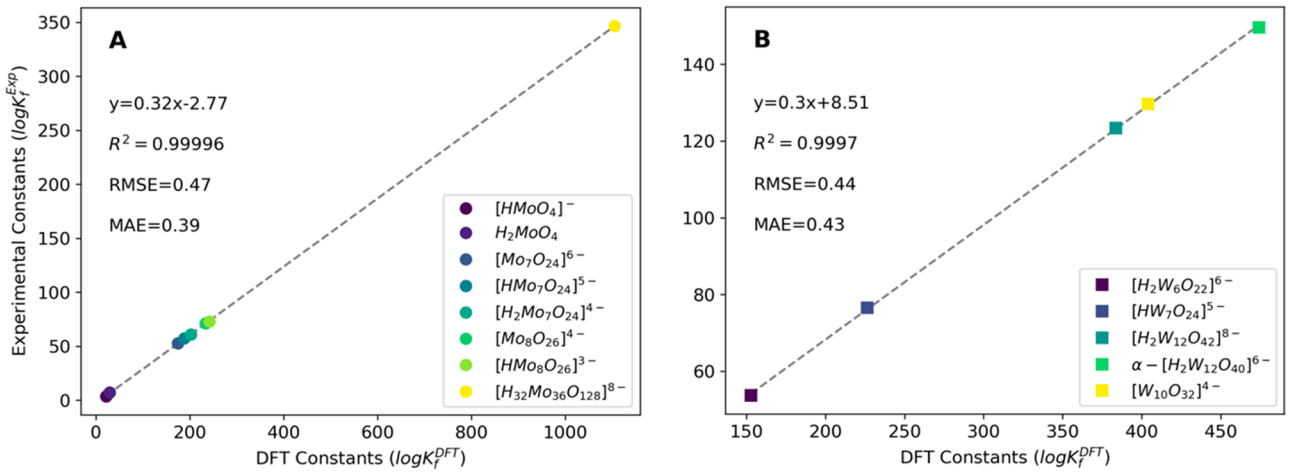


Figure 2. Linear regressions between log K_f^{Exp} and log K_f^{DFT}, with the lowest root mean squared error (RMSE) and mean average (MAE) values at 298.15 K and 1 atm for A polyoxomolybdates (circles) and B −polyoxotungstates (squares) at 1.0 and 0.25 M NaCl ionic strengths, respectively.

is more favorable than molybdenum oxides. This evidence is in very good agreement with experiments.[51] The chemical explanation of this phenomenon is attributed to the larger strength of the W−O bond compared to the Mo−O bond.[52] Even if this energy difference is relatively small, the accumulative effect in medium and large clusters gains importance.

There is a big disadvantage in quantum mechanical assessment of the Gibbs energies of small solvated ions: the proton species.[53] Although this error could be relatively small, the fact that the pH units are logarithmic amplifies the disturbing effect. The most common approaches to mitigate this error rely on setting up thermodynamic cycles and performing some sort of calibrations. For example, Grimme et al. studied 26 multiproton organic compounds to estimate the pK_a employing a linear free-energy relationship.[54] Besides, Machuqueiro et al. recently developed a method to predict pK_a values of titratable residues in biomolecules.[55] These great achievements contrast the more restricted developments in the area of inorganic chemistry.[56] We attribute this difference to the difficulties involved in treating distinctly highly anionic charged compounds in solution in the same manner. In this context, our group has recently discovered that POMSimulator offers an exceptionally accurate calibration for small molybdenum oxides.[48] We found that our method computed formation constants that relate just linearly to some experimental ones. Thereby, we could correct the intrinsic DFT error with a simple linear scaling while retaining high accuracy (R^2 = 0.9997 and root mean squared error (RMSE) = 0.43). Thanks to this successful calibration, POMSimulator computed speciation diagrams, which were in excellent agreement with the experimental ones. In light of these promising results, we have now applied our method to a larger molybdenum oxide set and a tungsten oxide set. So far, we have observed that POMSimulator correctly predicts the relative stabilities of molybdenum and tungsten oxides (see Figure 2). Even so, the most determining aspect for assessing the overall performance is the accuracy of the calibration step.

Figure 2 shows that POMSimulator predicts the equilibrium constants of two different metal oxide systems with extraordinary accuracy. The experimental formation constants employed were determined by Cruywagen et al. and Rozantsev et al. at ionic strengths of 1.0 and 0.25 M, respectively.[57,58] Figure 2A presents the linear correlation for molybdenum formation constants. The *x*-axis corresponds to the constants computed by our method, while the *y*-axis corresponds to the experimental constants reported by Cruywagen.[57] The molybdenum compounds used for the calibration are as follows: $[HMoO_4]^-$, H_2MoO_4, $[Mo_7O_{24}]^{6-}$, $[HMo_7O_{24}]^{5-}$, $[H_2Mo_7O_{24}]^{4-}$, $[Mo_8O_{26}]^{4-}$, $[HMo_8O_{26}]^{3-}$, and $[H_{32}Mo_{36}O_{128}]^{8-}$. The eight molybdenum clusters are represented as circles in the plot. Note that the linear regression in Figure 2A is the most accurate case out of the other 239 promising regressions depicted in Figure S3A. The selected regression presents the highest correlation coefficient, R^2 = 0.99996, and the lowest root mean squared error, RMSE = 0.47, and mean average error, MAE = 0.39. These excellent parameters prove that POMSimulator can predict equilibrium constants of medium-size molybdenum clusters. Figure 2B shows the linear correlation for tungsten formation constants. The experimental values are extracted from the work of Rozantsev and co-workers.[58]

The set of tungsten oxides used for the calibration are as follows: $[H_2W_6O_{22}]^{6-}$, $[HW_7O_{24}]^{5-}$, $[H_2W_{12}O_{42}]^{8-}$, a-$[H_2W1_2O_{40}]^{6-}$, and $[W_{10}O_{32}]^{4-}$. Figure 2B presents the best linear regression for the tungsten oxide system, even though there are ten additional regressions with an RMSE lower than 1.0 (Figure S3B). Tungsten oxide calibration offers outstanding accuracy with a correlation of R^2 = 0.9997 and prediction errors of RMSE = 0.44 and MAE = 0.43. Therefore, this calibration method not only applies to molybdenum oxides but also to tungsten oxides. In addition to this generalization, the prediction capabilities are remarkably high in both cases. Note that the values of the two slopes, which were obtained independently, are very similar (0.3 vs 0.32), but the value in the origin is not, indicating that this calibration method is metal-dependent.

POMSimulator aims at providing an exhaustive picture speciation in solution. We have already demonstrated that the calibration step is close to optimal, thus we can obtain accurate equilibrium constants. Therefore, the next step is the resolution of the multispecies chemical equilibria employing the corrected thermodynamic constants. In our previous work, we presented speciation diagrams, which consist of concentration distribution curves that vary according to the pH. Nonetheless, we have noticed that this approach only shows a

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219

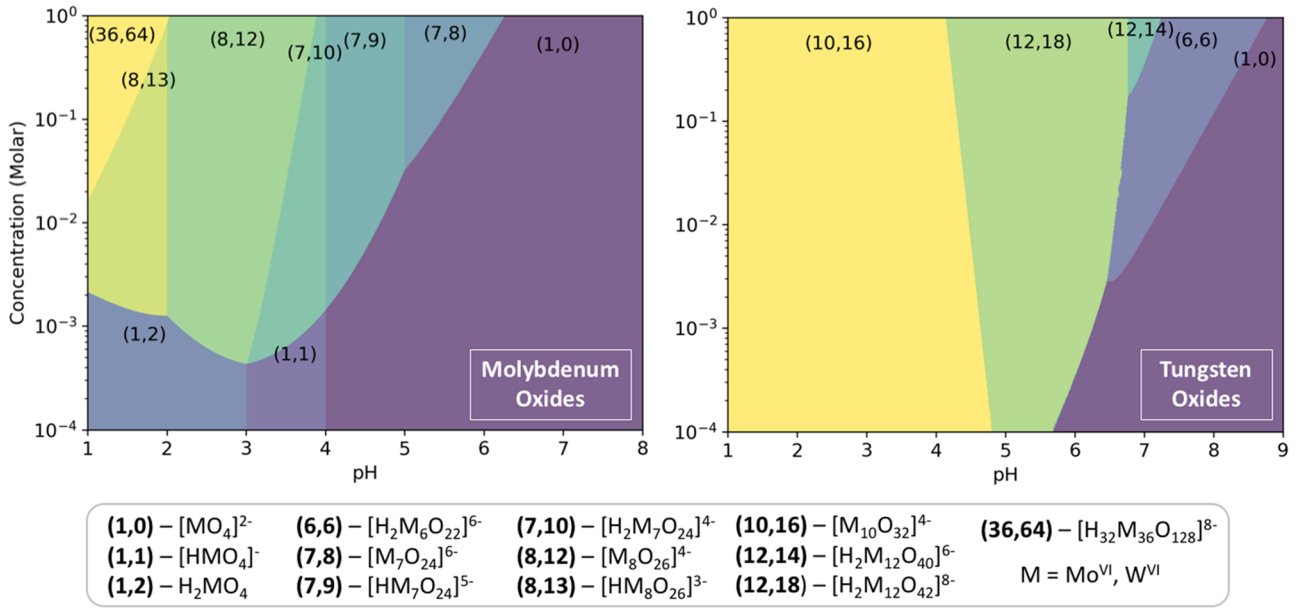


Figure 3. Speciation phase diagrams employing scaled ΔG^{DFT} values, at 298.15 K and 1 atm. Left: polyoxomolybdates at I = 1 M NaCl, with a grid of 5×10^9 solved models. Right: polyoxotungstates at I = 0.25 M NaCl, with a grid of 1×10^8 solved models. The labels in the legend indicate the general formula of polyoxometalates (M = Mo^{VI} and W^{VI}). Indices (p, q) refer to $p[MO_4]^{2-} + qH^+ \rightleftharpoons [H_zM_pO_m]^{n-} + (q/2 + p - z/2)\ H_2O$. Total metal molar concentration is fixed between 10^0 and 10^{-4} M. Concentration axes are on a logarithmic scale.

narrow window of the metal oxide speciation. As a matter of fact, the total metal concentration is a determining variable as well.

For instance, there is no aggregation at very diluted concentrations, whereas there are clusters that form only at high concentrations. Thus, it is necessary to scan both the pH and the total concentration to build speciation phase diagrams. For the very first time, we are presenting phase diagrams for polyoxometalate aqueous speciation based on quantum mechanical calculations. Figure 3 shows the speciation phase diagrams computed by POMSimulator employing a rather fine grid. The number of systems of equations solved in Figure 3A,B are 5×10^9 and 1×10^8, respectively. Note that not only the proton concentration (x-axis) is on a logarithmic scale but also the metal oxide concentration (y-axis). We have done so to better represent the details of the diagram at very low concentrations. Chemical species in Figure 3 are annotated with their stoichiometric formula, $[H_zM_pO_m]^{n-}$, and with the common notation in speciation studies, (p, q), which refers to stoichiometric coefficients of the formation reactions.

Figure 3 shows the speciation phase diagrams for molybdenum and tungsten oxides using the rescaled formation constants. At a pH higher than 7, the speciation is completely dominated by the monomer $[MoO_4]^{2-}$, regardless of the total concentration. Under more acidic conditions, molybdenum monoxide either undergoes polymerization or protonation reactions. This variability depends on the total concentration: at lower values ($<10^3$ M), it is less probable that two molecules will react due to the diffusion factor. Therefore, the only reactions that take place are the acid−base equilibria of the monomer to hydrogenmolybdate, $[HMoO_4]^-$, and molybdic acid, H_2MoO_4. However, at higher concentrations, the aggregation is favored and promotes the growth of larger clusters. In fact, the largest polyoxomolybdate, $\{Mo_{36}\}$, is only formed at concentrations larger than 5×10^2 M. Furthermore, the condensation zone (yellow-green area) includes other relevant oxides such as the octamolybdates $[Mo_8O_{26}]^{4-}$ and $[HMo_8O_{26}]^{-3}$, at a pH range between 1 and 3.5. Under more neutral conditions, between 4 and 6, the heptamer oxides $[H_2Mo_7O_{24}]^{4-}$, $[HMo_7O_{24}]^{5-}$, and $[Mo_7O_{24}]^{6-}$ predominate. It can be observed that POMos speciation is extraordinarily rich both in the number of species and topologies.

Likewise, the speciation phase diagram for polyoxotungstates, using the rescaled formation constants, is shown in Figure 3B. Tungsten monoxide is the most abundant compound under relatively alkaline conditions (pH > 9). In contrast to POMos, the concentrations for POTs does not have a sharp disturbing effect on the relative abundances. Despite that, in very diluted solutions, the polymerization process does not take place either. Decatungstate, $[W_{10}O_{32}]^{4-}$, is the predominant oxide at pH lower than 5. Between pH 5 and 6.5, the most abundant species are dodecatungstates $[H_2W_{12}O_{42}]^{8-}$ and to a lesser extend $[H_2W_{12}O_{40}]^{6-}$, which exist only at high concentrations ($>10^1$M). Before dissembling into the monomer, hexatungstate $[H_2W_6O_{22}]^{6-}$ predominates at neutral pH. Immense efforts have been made using experimental techniques to characterize metal oxide formation constants. Even so, there are many metal-oxo compounds that have proved to be difficult to isolate and characterize by the current state-of-the-art techniques. For example, pentagonal molybdenum oxides, $\{Mo_xMo_6\}$, are a recurrent motif in medium and large clusters. These pentagonal structures are assigned as important building blocks for adding curvature and sphericity to Keplerate architectures.[5] Despite that and as far as we know, no formation constants have been determined yet. Moreover, the formation of $[W_{12}O_{42}]^{8-}$, through the lacunary form $[W_{11}O_{40}]^{12-}$, is still under discussion.[24] In view of these knowledge gaps, our method offers a straightforward approach to estimate equilibrium constants of transient intermediates. Table 1 collects the predicted values for the formation constants for some molybdenum and tungsten IPAs. Note that Table 1 shows a single state of protonation for each cluster to ease the interpretation. Nevertheless, the complete set of 122 constants can be found in Tables S7 and S8. To estimate

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219

Table 1. DFT Computed Formation Constants for Polyoxomolybdates at 298.15 K, 1 atm and 1 M NaCl and Polyoxotungstates at 298.15 K, 1 atm, and 0.25 M NaCl

molybdenum oxides	formation Constant[a,b]	tungsten oxides	formation constant[a,c]
$[Mo_2O_7]^{2-}$	11.68	$[W_2O_7]^{2-}$	23.42
$[H_6Mo_2O_{10}]^{2-}$	2.66	$[W_3O_{10}]^{2-}$	38.85
$[Mo_3O_{10}]^{2-}$	22.14	$[W_3O_{11}]^{4-}$	26.27
$[H_6Mo_3O_{14}]^{4-}$	3.27	$[W_4O_{13}]^{2-}$	54.48
$[Mo_4O_{13}]^{2-}$	33.59	$[W_4O_{15}]^{6-}$	19.53
$[H_6Mo_4O_{18}]^{6-}$	12.86	$[W_5O_{16}]^{2-}$	58.09
$[Mo_5O_{16}]^{2-}$	38.02	$[W_5O_{17}]^{4-}$	51.06
$[Mo_5O_{17}]^{4-}$	35.19	$[W_5O_{19}]^{8-}$	22.01
$[H_6Mo_5O_{22}]^{8-}$	18.82	$[W_6O_{20}]^{4-}$	68.21
$[Mo_6O_{20}]^{4-}$	41.84	$[W_{11}O_{40}]^{14-}$	64.00
$[H_{11}Mo_6O_{27}]^{7-}$	32.37		
$[Mo_7O_{23}]^{4-}$	55.36		
$[H_{12}Mo_7O_{31}]^{8-}$	44.37		
$[H_{13}Mo_8O_{35}]^{9-}$	51.47		
$[H_{14}Mo_9O_{37}]^{6-}$	68.61		
$[H_{14}Mo_9O_{38}]^{8-}$	62.51		

[a]Constants refer to the formation reaction: $p[MO_4]^{2-} + qH^+ \rightleftharpoons [H_zM_pO_m]^{n-} + (q/2+ p - z/2)\ H_2O$. [b]Linear regression employed: $y = 0.32x - 2.77$ (RMSE = 0.47, MAE = 0.39) using Cruywagen et al. constants[57] for the calibration. [c]Linear regression employed: $y = 0.30x + 8.51$ (RMSE = 0.44, MAE = 0.43) using Rozantsev et al. constants[58] for the calibration.

these constants, the linear regressions, depicted in Figure 2, are used to calibrate the raw formation constants. Thanks to the excellent fitting, we can obtain accurate constants of unreported metal-oxo compounds. POMSimulator has determined 64 molybdenum equilibrium constants, but Table 1 collects only the 16 most relevant examples. This sample includes from small clusters (e.g., $[Mo_2O_7]^{2-}$, $[Mo_4O_{13}]^{2-}$) to pentagonal-like structures (e.g., $[H_6Mo_5O_{22}]^{8-}$, $[H_{14}Mo_9O_{38}]^{8-}$). Furthermore, our method has computed 47 tungsten formation constants, but Table 1 shows only the 10 most representative samples. This set is formed by small clusters (e.g., $[W_3O_{10}]^{2-}$, $[W_5O_{16}]^{2-}$) and larger and more anionic species (e.g., $[W_5O_{19}]^{8-}$, $[W_{11}O_{40}]^{14-}$). It is worth noting that even with this newly calculated data, all tungsten oxide formation constants are larger than the molybdenum oxide analogues, thus indicating higher stability in accordance with experimental facts.

CONCLUSIONS

The present work presents a robust methodology for dealing with complex multispecies chemical equilibria in solution in an automated manner, which allows predicting speciation phase diagrams and formation constants of transient species that are out of reach experimentally. This first complete application of the new method to the full set of molybdenum and tungsten IPAs demonstrates its potential for understanding and predicting metal oxide speciation in aqueous solution. Concentrations, pH, and ionic force play a key role in the nucleation and growth processes of larger clusters, and our method enables taking all of them into account.

Starting with a diverse set of molecular species at different protonation states, our method creates molecular graphs first using just quantum mechanical derived raw data (basically the electronic charge density and Gibbs free energies). From there, reaction networks are deduced and subsequently employed for setting up the speciation models, thus hundreds of thousands of nonlinear equations are solved relatively fast. The values of the equilibrium constants computed with our method show that polyoxotungstates are more stable than polyoxomolybdates, which is in very good agreement with experimental facts. So far, we have predicted the formation constants of 107 unreported compounds. For example, we have determined formation constants of the pentagonal structures: $\{Mo_6\}$, $\{MoMo_6\}$, $\{Mo_2Mo_6\}$, and $\{Mo_3Mo_6\}$ and reactive tungsten intermediates such as $\{W_5\}$ and $\{W_{11}\}$.

The unique capability to build speciation phase diagrams, a feature with which our code has been equipped, provides a powerful tool for guiding experiments. We expect that these kinds of plots will help optimize the experimental conditions toward the synthesis of the desired cluster. Molybdenum and tungsten IPAs, as those considered here, are well-known species and thus not the holy grail in polyoxometalate chemistry. Other cations and anions play a fundamental role in the growth of larger structures for which trial and error approaches still prevail. Although the present work proves that the multiequilibrium concept can be extended further, including kinetics equations in our simulation protocols is the next challenge.

Furthermore, and following common practice to rescale the DFT Gibbs free energies when computing acid/base equilibrium constants, we observed a strong linear dependence between computed and experimental formation constants, which allows us to calibrate the thermodynamic constants. The linear relationships exhibit exceptionally high correlations (R^2 = 0.9999) and low errors (RMSE = 0.45), both beyond the current benchmark. These results provide a solid ground to describe the systems included in this study.

Last but not least, we have noticed some insightful aspects of the calibration method. On the one hand, the actual linear regression of POMos is remarkably close so it extends what we reported previously. This fact is highly relevant since it adds robustness to the calibration scale. Moreover, tungsten oxides present a regression with the very same slope (m = 0.3) and different intercepts. This trend suggests that a universal calibration might exist that could correct the ΔG^{DFT} error for various systems. More experiments are currently underway to investigate this hypothesis further and expand the applicability of POMSimulator throughout the periodic table. Since molecular metal oxides are often viewed as inorganic polymers, we envisage that this kind of treatment could be applied to other condensation and addition reactions, like those in organic polymerization reactions.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.1c03292.

Computational details, overview of the Python code developed, full list of chemical reactions, and speciation models and formation constants plus additional data (PDF)

AUTHOR INFORMATION

Corresponding Author

Carles Bo − *Institute of Chemical Research of Catalonia (ICIQ), The Barcelona Institute of Science and Technology (BIST), 43007 Tarragona, Spain; Departament de Química*

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219

Física i Inorgánica, Universitat Rovira i Virgili, 43007 Tarragona, Spain; orcid.org/0000-0001-9581-2922; Email: cbo@iciq.cat

Author

Enric Petrus − *Institute of Chemical Research of Catalonia (ICIQ), The Barcelona Institute of Science and Technology (BIST), 43007 Tarragona, Spain;* orcid.org/0000-0002-8022-3153

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.1c03292

Author Contributions

E.P. created the Python codes used in this work, performed the calculations, and generated all drawings. C.B. designed and supervised the whole project. All authors contributed to the discussions and wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the ICIQ Foundation, the CERCA Program and AGAUR (grant 2017SGR00290) of the Generalitat de Catalunya, and the Spanish Ministerio de Ciencia, Innovacion y Universidades through project CTQ2017-88777-R for financial support.

■ REFERENCES

(1) Mikulak-Klucznik, B.; Gołębiowska, P.; Bayly, A. A.; Popik, O.; Klucznik, T.; Szymkuć, S.; Gajewska, E. P.; Dittwald, P.; Staszewska-Krajewska, O.; Beker, W.; et al. Computational Planning of the Synthesis of Complex Natural Products. *Nature* **2020**, *588*, 83−88.

(2) Berzelius, J. J. Beitrag Zur Näheren Kenntniss Des Molybdäns. *Ann. Phys.* **1826**, *82*, 369−392.

(3) Das, V.; Kaushik, R.; Hussain, F. Heterometallic 3d-4f Polyoxometalates: An Emerging Field with Structural Diversity to Multiple Applications. *Coord. Chem. Rev.* **2020**, *413*, No. 213271.

(4) Auvray, T.; Matson, E. M. Polyoxometalate-Based Complexes as Ligands for the Study of Actinide Chemistry. *Dalton Trans.* **2020**, *49*, 13917−13927.

(5) Vilà-Nadal, L.; Miras, H. N. Coordination Chemistry in Polyoxometalates and Metal Clusters. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Elsevier, 2020; pp 1−37.

(6) Müller, A.; Serain, C. Soluble Molybdenum Blues - "Des Pudels Kern.". *Acc. Chem. Res.* **2000**, *33*, 2−10.

(7) Garrido Ribó, E.; Bell, N. L.; Xuan, W.; Luo, J.; Long, D.-L.; Liu, T.; Cronin, L. Synthesis, Assembly, and Sizing of Neutral, Lanthanide Substituted Molybdenum Blue Wheels {Mo 90 Ln 10}. *J. Am. Chem. Soc.* **2020**, *142*, 17508−17514.

(8) Müller, A.; Krickemeyer, E.; Bögge, H.; Schmidtmann, M.; Peters, F. Organizational Forms of Matter: An Inorganic Super Fullerene and Keplerate Based on Molybdenum Oxide. *Angew. Chem., Int. Ed.* **1998**, *37*, 3360−3363.

(9) Lin, J.; Li, N.; Yang, S.; Jia, M.; Liu, J.; Li, X.-M.; An, L.; Tian, Q.; Dong, L.-Z.; Lan, Y.-Q. Self-Assembly of Giant Mo 240 Hollow Opening Dodecahedra. *J. Am. Chem. Soc.* **2020**, 13982−13988.

(10) Burns, P. C.; Kubatko, K. A.; Sigmon, G.; Fryer, B. J.; Gagnon, J. E.; Antonio, M. R.; Soderholm, L. Actinyl Peroxide Nanospheres. *Angew. Chem., Int. Ed.* **2005**, *44*, 2135−2139.

(11) Dembowski, M.; Colla, C. A.; Yu, P.; Qiu, J.; Szymanowski, J. E. S.; Casey, W. H.; Burns, P. C. The Propensity of Uranium-Peroxide Systems to Preserve Nanosized Assemblies. *Inorg. Chem.* **2017**, *56*, 9602−9608.

(12) Bijelic, A.; Aureliano, M.; Rompel, A. Polyoxometalates as Potential Next-Generation Metallodrugs in the Combat Against Cancer. *Angew. Chem., Int. Ed.* **2019**, *58*, 2980−2999.

(13) Vandebroek, L.; Noguchi, H.; Kamata, K.; Tame, J. R. H.; Van Meervelt, L.; Parac-Vogt, T. N.; Voet, A. R. D. Hybrid Assemblies of a Symmetric Designer Protein and Polyoxometalates with Matching Symmetry. *Chem. Commun.* **2020**, *56*, 11601−11604.

(14) Moons, J.; Azambuja, F.; Mihailovic, J.; Kozma, K.; Smiljanic, K.; Amiri, M.; Cirkovic Velickovic, T.; Nyman, M.; Parac-Vogt, T. N. Discrete Hf 18 Metal-oxo Cluster as a Heterogeneous Nanozyme for Site-Specific Proteolysis. *Angew. Chem.* **2020**, *132*, 9179−9186.

(15) Sciortino, G.; Aureliano, M.; Garribba, E. Rationalizing the Decavanadate(V) and Oxidovanadium(IV) Binding to G-Actin and the Competition with Decaniobate(V) and ATP. *Inorg. Chem.* **2021**, *60*, 334−344.

(16) Blasco-Ahicart, M.; Soriano-Lopez, J.; Carbo, J. J.; Poblet, J. M.; Galan-Mascaros, J. R. Polyoxometalate Electrocatalysts Based on Earthabundant Metals for Efficient Water Oxidation in Acidic Media. *Nat. Chem.* **2018**, *10*, 24−30.

(17) Maksimchuk, N. V.; Ivanchikova, I. D.; Maksimov, G. M.; Eltsov, I. V.; Evtushok, V. Yu.; Kholdeeva, O. A.; Lebbie, D.; Errington, R. J.; Solé-Daura, A.; Poblet, J. M.; et al. Why Does Nb(V) Show Higher Heterolytic Pathway Selectivity Than Ti(IV) in Epoxidation with H 2 O 2? Answers from Model Studies on Nb- and Ti-Substituted Lindqvist Tungstates. *ACS Catal.* **2019**, *9*, 6262−6275.

(18) Greiner, S.; Schwarz, B.; Ringenberg, M.; Dürr, M.; Ivanovic-Burmazovic, I.; Fichtner, M.; Anjass, M.; Streb, C. Redox-Inactive Ions Control the Redox-Activity of Molecular Vanadium Oxides. *Chem. Sci.* **2020**, *11*, 4450−4455.

(19) Gobbo, P.; Tian, L.; Kumar, B. V. V. S. P.; Turvey, S.; Cattelan, M.; Patil, A. J.; Carraro, M.; Bonchio, M.; Mann, S. Catalytic Processing in Ruthenium-Based Polyoxometalate Coacervate Protocells. *Nat. Commun.* **2020**, *11*, No. 41.

(20) Craig, M. J.; Barda-Chatain, R.; García-Melchor, M. Fundamental Insights and Rational Design of Low-Cost Polyoxometalates for the Oxygen Evolution Reaction. *J. Catal.* **2021**, *393*, 202−206.

(21) Zhong, J.; Pérez-Ramírez, J.; Yan, N. Biomass Valorisation over Polyoxometalate-Based Catalysts. *Green Chem.* **2021**, *23*, 18−36.

(22) Spano, T. L.; Simonetti, A.; Corcoran, L.; Smith, P. A.; Lewis, S. R.; Burns, P. C. Comparative Chemical and Structural Analyses of Two Uranium Dioxide Fuel Pellets. *J. Nucl. Mater.* **2019**, *518*, 149−161.

(23) Traustason, H.; Bell, N. L.; Caranto, K.; Auld, D. C.; Lockey, D. T.; Kokot, A.; Szymanowski, J. E. S.; Cronin, L.; Burns, P. C. Reactivity, Formation, and Solubility of Polyoxometalates Probed by Calorimetry. *J. Am. Chem. Soc.* **2020**, *142*, 20463−20469.

(24) Gumerova, N. I.; Rompel, A. Polyoxometalates in Solution: Speciation under Spotlight. *Chem. Soc. Rev.* **2020**, *49*, 7568−7601.

(25) Ruiz de La Oliva, A.; Sans, V.; Miras, H. N.; Long, D. L.; Cronin, L. Coding the Assembly of Polyoxotungstates with a Programmable Reaction System. *Inorg. Chem.* **2017**, *56*, 5089−5095.

(26) Marianski, M.; Seo, J.; Mucha, E.; Thomas, D. A.; Jung, S.; Schlögl, R.; Meijer, G.; Trunschke, A.; Von Helden, G. Structural Characterization of Molybdenum Oxide Nanoclusters Using Ion Mobility Spectrometry-Mass Spectrometry and Infrared Action Spectroscopy. *J. Phys. Chem. C* **2019**, *123*, 7845−7853.

(27) Centellas, M. S.; Piot, M.; Salles, R.; Proust, A.; Tortech, L.; Brouri, D.; Hupin, S.; Abécassis, B.; Landy, D.; Bo, C.; et al. Exploring the Self-Assembly of Dumbbell-Shaped Polyoxometalate Hybrids, from Molecular Building Units to Nanostructured Soft Materials. *Chem. Sci.* **2020**, *11*, 11072−11080.

(28) Hülsey, M. J.; Sun, G.; Sautet, P.; Yan, N. Observing Single-Atom Catalytic Sites During Reactions with Electrospray Ionization Mass Spectrometry. *Angew. Chem., Int. Ed.* **2021**, *60*, 4764−4773.

(29) Nyman, M. Small-Angle X-Ray Scattering to Determine Solution Speciation of Metal-Oxo Clusters. *Coord. Chem. Rev.* **2017**, *352*, 461−472.

(30) Miras, H. N.; Mathis, C.; Xuan, W.; Long, D.-L.; Pow, R.; Cronin, L. Spontaneous Formation of Autocatalytic Sets with Self-Replicating Inorganic Metal Oxide Clusters. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 10699−10705.
(31) Oster, G.; Perelson, A. Chemical Reaction Networks. *IEEE Trans. Circuits Syst.* **1974**, *21*, 709−721.
(32) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem. A* **1977**, *81*, 2340−2361.
(33) Funes-Ardoiz, I.; Schoenebeck, F. Established and Emerging Computational Tools to Study Homogeneous Catalysis—From Quantum Mechanics to Machine Learning. *Chem* **2020**, *6*, 1904−1913.
(34) Harvey, J. N.; Himo, F.; Maseras, F.; Perrin, L. Scope and Challenge of Computational Methods for Studying Mechanism and Reactivity in Homogeneous Catalysis. *ACS Catal.* **2019**, *9*, 6803−6813.
(35) Wills, L. A.; Qu, X.; Chang, I.-Y.; Mustard, T. J. L.; Keszler, D. A.; Persson, K. A.; Cheong, P. H.-Y. Group Additivity-Pourbaix Diagrams Advocate Thermodynamically Stable Nanoscale Clusters in Aqueous Environments. *Nat. Commun.* **2017**, *8*, No. 15852.
(36) Vilà−Nadal, L.; Rodríguez-Fortea, A.; Yan, L. K.; Wilson, E. F.; Cronin, L.; Poblet, J. M. Nucleation Mechanisms of Molecular Oxides: A Study of the Assembly-Dissassembly of [W6O19]2- by Theory and Mass Spectrometry. *Angew. Chem., Int. Ed.* **2009**, *48*, 5452−5456.
(37) 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*, 7811−7819.
(38) Cameron, J. M.; Vilà-Nadal, L.; Winter, R. S.; Iijima, F.; Murillo, J. C.; Rodríguez-Fortea, A.; Oshio, H.; Poblet, J. M.; Cronin, L. Investigating the Transformations of Polyoxoanions Using Mass Spectrometry and Molecular Dynamics. *J. Am. Chem. Soc.* **2016**, *138*, 8765−8773.
(39) Dewyer, A. L.; Argüelles, A. J.; Zimmerman, P. M. Methods for Exploring Reaction Space in Molecular Systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1354.
(40) Unsleber, J. P.; Reiher, M. The Exploration of Chemical Reaction Networks. *Annu. Rev. Phys. Chem.* **2020**, *71*, 121−142.
(41) Coley, C. W. Defining and Exploring Chemical Spaces. *Trends Chem.* **2021**, *3*, 133−145.
(42) Broadbelt, L. J.; Stark, S. M.; Klein, M. T. Computer Generated Pyrolysis Modeling: On-the-Fly Generation of Species, Reactions, and Rates. *Ind. Eng. Chem. Res.* **1994**, *33*, 790−799.
(43) Craciun, G.; Feinberg, M. Multiple Equilibria in Complex Chemical Reaction Networks: Extensions to Entrapped Species Models. *IEE Proc.: Syst. Biol.* **2006**, *153*, 179−186.
(44) Wołos, A.; Roszak, R.; Zadlo-Dobrowolska, A.; Beker, W.; Mikulak-Klucznik, B.; Spólnik, G.; Dygas, M.; Szymkuc, S.; Grzybowski, B. A. Synthetic Connectivity, Emergence, and Self-Regeneration in the Network of Prebiotic Chemistry. *Science* **2020**, *369*, No. eaaw1955.
(45) Kim, Y.; Kim, J. W.; Kim, Z.; Kim, W. Y. Efficient Prediction of Reaction Paths through Molecular Graph and Reaction Network Analysis. *Chem. Sci.* **2018**, *9*, 825−835.
(46) Ismail, I.; Stuttaford-Fowler, H. B. V. A.; Ochan Ashok, C.; Robertson, C.; Habershon, S. Automatic Proposal of Multistep Reaction Mechanisms Using a Graph-Driven Search. *J. Phys. Chem. A* **2019**, *123*, 3407−3417.
(47) Jinich, A.; Sanchez-Lengeling, B.; Ren, H.; Goldford, J. E.; Noor, E.; Sanders, J. N.; Segrè, D.; Aspuru-Guzik, A. A Thermodynamic Atlas of Carbon Redox Chemical Space. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 32910−32918.
(48) Petrus, E.; Segado, M.; Bo, C. Nucleation Mechanisms and Speciation of Metal Oxide Clusters. *Chem. Sci.* **2020**, *11*, 8448−8456.
(49) Álvarez-Moreno, M. M.; Graaf, C.; López, N.; Maseras, F.; Poblet, M. J.; Bo, C. Managing the Computational Chemistry Big Data Problem: The IoChem-BD Platform. *J. Chem. Inf. Model.* **2015**, *55*, 95−103.
(50) Petrus, E.; Bo, C. *ioChem Data Collection*. http://dx.doi.org/10.19061/iochem-bd-1-201.
(51) Pope, M. T.; Müller, A. Polyoxometalate Chemistry: An Old Field with New Dimensions in Several Disciplines. *Angew. Chem., Int. Ed.* **1991**, *30*, 34−48.
(52) Baker, L. C. W.; Glick, D. C. Present General Status of Understanding of Heteropoly Electrolytes and a Tracing of Some Major Highlights in the History of Their Elucidation. *Chem. Rev.* **1998**, *98*, 3−49.
(53) Seybold, P. G.; Shields, G. C. Computational Estimation of PKa Values. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2015**, *5*, 290−297.
(54) Pracht, P.; Wilcken, R.; Udvarhelyi, A.; Rodde, S.; Grimme, S. High Accuracy Quantum-Chemistry-Based Calculation and Blind Prediction of Macroscopic PKa Values in the Context of the SAMPL6 Challenge. *J. Comput.-Aided Mol. Des.* **2018**, *32*, 1139−1149.
(55) Reis, P. B. P. S.; Vila-Viçosa, D.; Rocchia, W.; Machuqueiro, M. PypKa: A Flexible Python Module for Poisson-Boltzmann-Based PKa Calculations. *J. Chem. Inf. Model.* **2020**, *60*, 4442−4448.
(56) Seybold, P. G. Quantum Chemical Estimation of the Acidities of Some Inorganic Oxoacids. *Mol. Phys.* **2015**, *113*, 232−236.
(57) Cruywagen, J. J. Protonation, Oligomerization, and Condensation Reactions of Vanadate(V), Molybdate(VI), and Tungstate-(VI). In *Advances in Inorganic Chemistry*; Elsevier, 1999; Vol. *49*, pp 127−182.
(58) Rozantsev, G. M.; Sazonova, O. I. Thermodynamic Parameters of Interconversions of Isopolyanions in Solutions of Tungsten(VI). *Russ. J. Coord. Chem.* **2005**, *31*, 552−558.

https://doi.org/10.1021/acs.jpca.1c03292
J. Phys. Chem. A 2021, 125, 5212−5219