

## ARTICLE

Received 19 Dec 2016 | Accepted 8 May 2017 | Published 15 Jun 2017

DOI: 10.1038/ncomms15852

OPEN

# Group additivity-Pourbaix diagrams advocate thermodynamically stable nanoscale clusters in aqueous environments

Lindsay A. Wills<sup>1</sup>, Xiaohui Qu<sup>2</sup>, I-Ya Chang<sup>1</sup>, Thomas J.L. Mustard<sup>1</sup>, Douglas A. Keszler<sup>1</sup>, Kristin A. Persson<sup>2</sup> & Paul Ha-Yeon Cheong<sup>1</sup>

The characterization of water-based corrosion, geochemical, environmental and catalytic processes rely on the accurate depiction of stable phases in a water environment. The process is aided by Pourbaix diagrams, which map the equilibrium solid and solution phases under varying conditions of pH and electrochemical potential. Recently, metastable or possibly stable nanometric aqueous clusters have been proposed as intermediate species in non-classical nucleation processes. Herein, we describe a Group Additivity approach to obtain Pourbaix diagrams with full consideration of multimeric cluster speciation from computations. Comparisons with existing titration results from experiments yield excellent agreement. Applying this Group Additivity-Pourbaix approach to Group 13 elements, we arrive at a quantitative evaluation of cluster stability, as a function of pH and concentration, and present compelling support for not only metastable but also thermodynamically stable multimeric clusters in aqueous solutions.

<sup>1</sup>Department of Chemistry, Oregon State University, 153 Gilbert Hall, Corvallis, Oregon 97331-4003, USA. <sup>2</sup>Department of Materials Science and Engineering, UC Berkeley, Berkeley, California 94720, USA. Correspondence and requests for materials should be addressed to D.A.K. (email: douglas.keszler@oregonstate.edu) or to K.A.P. (email: kristinpersson@berkeley.edu) or to P.H.-Y.C. (email: paulc@science.oregonstate.edu).

Cluster formation in aqueous solutions have attracted renewed interest during the last decade<sup>1</sup>. Aqueous inorganic clusters are nanoscale molecular solvates composed of several metal ions linked via oxo and hydroxo ligands. They are crucial in a diverse range of chemical processes such as solution deposition of device thin films<sup>2</sup>, biomineralization<sup>3,4</sup> and the natural cycling of minerals through soil, water and biomass<sup>5</sup>. Aqueous clusters are also known to be causative agents of neurodegenerative disease and image enhancement additives<sup>6</sup>. Recently, stable or metastable aqueous cluster formation have been suggested to play an important role in an alternative nucleation process as compared to classical nucleation theory (CNT)<sup>3,7–9</sup>. In this scenario, nucleation towards a solid bulk state proceeds via aggregation of solvated clusters, and not via a one-at-a-time attachment of monomeric species to a critical nucleus as assumed in CNT<sup>10</sup>. The occurrence of nanoscale aqueous clusters, in systems such as CaCO<sub>3</sub> (refs 3,7,11,12), AlPO<sub>4</sub> (ref. 13), Ga (ref. 14), Fe (ref. 15) Fe oxyhydroxides (ref. 16), Hf/Zr (ref. 17) and Al (ref. 18) has been inferred from a variety of experimental visualization, speciation and spectroscopic observations. From the evidence, it is clear that these clusters (i) form through a reversible process<sup>3,19</sup> whose speciation is governed by equilibria of molecular species in solution<sup>3</sup> and (ii) are stable in solution for an extended time<sup>3,7,8</sup>. While the majority of published Pourbaix diagrams<sup>20–22</sup> exhibit exclusively monomeric aqueous ion stability domains, growing thermochemical data on clusters is being incorporated into speciation and Pourbaix analyses. As examples, we note that stable multimetric species, beyond dimers and trimers, are suggested for Mo, W and V<sup>22</sup> and for Bi<sup>20</sup> at 0.1 M and high potential. On the other hand, known cluster-forming systems such as Al, Fe, Ga, Hf and so on, exhibit monomeric aqueous ion stability in these reference works. Hence, there are still many questions surrounding nanometric cluster formation in solution, for example, as to the structure, the stability, the concentration ranges, counter-ion effects and solvation structure as well as the role of clusters in the nucleation process<sup>23</sup>. Theory can provide a complementary means to elucidate some of these questions. For example, quantum mechanical (QM) studies of the structure, electronic and spectroscopic properties, absorption and reaction sites for clusters are well established<sup>24,25</sup>. Recently, molecular dynamics simulations have been used to unravel solute-solvent and ion-pairing interactions as well as the dynamics of specific clusters<sup>24,26,27</sup>.

In this work, we present two distinct theoretical procedures, which together provide an efficient and theoretically rigorous route towards obtaining Pourbaix diagrams with full consideration of cluster stability and speciation from first-principles. First, we reveal a group additivity (GA) method that can predict the Gibbs free energies of aqueous oxo/hydroxo metal clusters in excellent agreement with QM and experiments at a fraction of the cost. GA was developed by Benson in the 1950s to rapidly predict the thermodynamics of organic compounds<sup>28,29</sup>. No such analogous method exists for pure aqueous inorganic compounds. We demonstrate the methodology on group 13 metal clusters which have drawn considerable attention recently due to their importance in catalytic microporous solids<sup>30</sup>, green aqueous precursors for large-area metal oxide thin films<sup>2</sup>, water treatment<sup>18</sup> as well as tumour diagnosis and possible treatment of malignancies<sup>31</sup>. Second, we use the resulting cluster free energies in the Pourbaix formalism of the Materials Project<sup>32</sup> developed by us<sup>33</sup>. The most important aspect of this formalism is that it uses a combination of computed solid states and clusters with experimentally obtained Gibbs free energies for the monomeric aqueous ions and water, which means that effects such as aqueous ion solvation are inherently accounted for. By combining the GA method with this formalism, we build

a foundation for robustly evaluating the thermodynamic stability of clusters against the competing solid phases and monomeric aqueous ions.

## Results

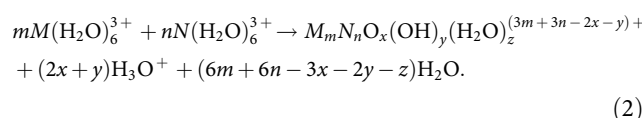
**Group additivity method for cluster energy prediction.** To demonstrate our approach and study the thermodynamic stability of clusters in aqueous solutions, we consider group 13 metal clusters; in particular, the Al and Ga systems. We emphasize that the monomeric aqueous species are single metal (*M*) aqueous ions, of the form  $[MO_xH_z]^{n+}$ , whereas the clusters are polymeric metal hydrolysis products  $[M_xO_yH_z]^{n+}$ . In the case of Al, common monomeric species such as  $[Al(H_2O)_6]^{3+}$ ,  $[AlOH(H_2O)_5]^{2+}$  and  $[Al(OH)_4]^{1-}$  as well as larger hydrolysis products can be identified by <sup>27</sup>Al-NMR as well as potentiometric studies<sup>34</sup>. Of the multimetric species, the dominant aluminium complex in aqueous systems is the large tridecamer  $[Al_{13}O_4(OH)_{24}(H_2O)_{12}]^{7+}$ , that is,  $\epsilon$ -Keggin ion, found across a broad pH range depending on aluminium ion concentration.

The aqueous behaviour of Ga has significant similarities to Al. Like Al, Ga forms a few monomeric species, such as  $[Ga(H_2O)_6]^{3+}$ ,  $[GaOH(H_2O)_5]^{2+}$  and  $[Ga(OH)_4]^{1-}$  under different concentration and pH conditions. Ga also forms larger hydrolysis products<sup>14</sup>, primarily in the form of the Ga<sub>13</sub> Keggin cluster, though other hydrolysis products have also been synthesized<sup>35</sup>.

To efficiently obtain the energy of a generic cluster, we present a GA method that predicts the thermodynamic stabilities of aqueous group 13 metal hydroxide and oxide clusters, regardless of size or structure. The method quantitatively reproduces first-principles computations in predicting hydrolysis reaction energies and stabilities of both homometallic and heterometallic group 13 clusters with mean absolute error of 3.0 kcal mol<sup>−1</sup>. These calculations take seconds, a fraction of the cost of computing the QM values. In this method, the Gibbs free energy of any given aqueous metal oxo/hydroxo cluster is expressed as the sum of the Gibbs free energies of the metal cations and the ligated oxo, hydroxo and aquo ligands. For instance, the absolute Gibbs free energy of Al dimer,  $[Al_2(OH)_2(H_2O)_8]$  is:

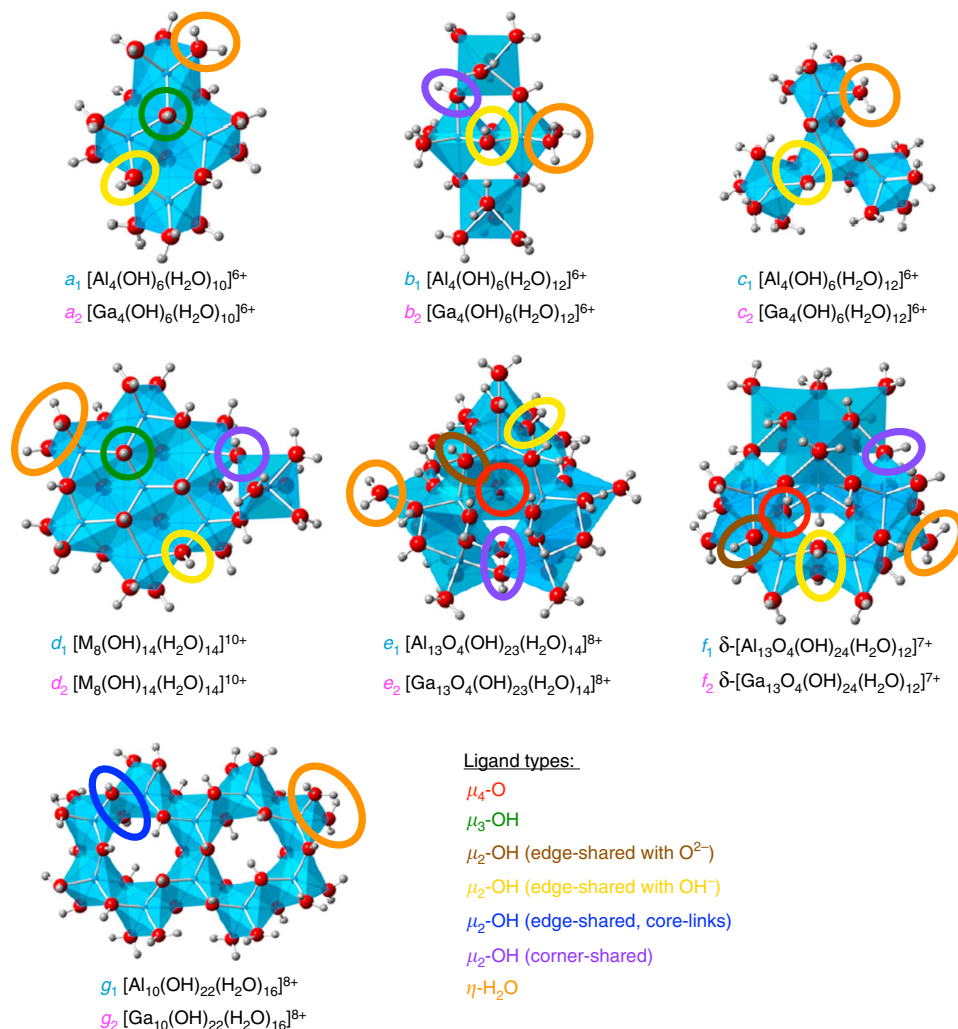
$$G_{Al_2} = 2G_{Al^{3+}} + 2G_{\mu_2-OH} + 8G_{\eta-H_2O}. \quad (1)$$

The general hydrolysis reaction can be written as the following equation, where  $\{M, N\}$  identify the group 13 metals and  $\{m, n\}$  denote the respective stoichiometries:

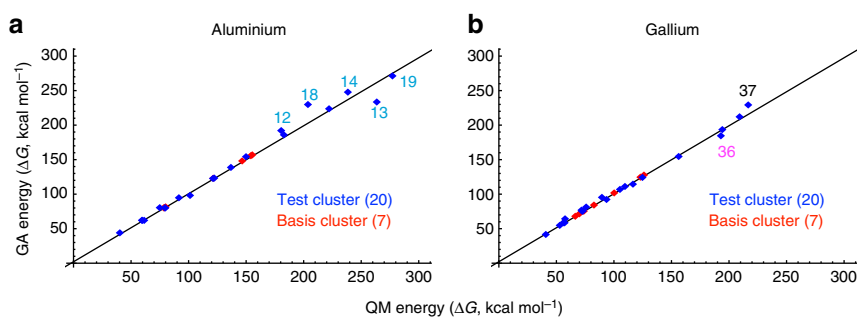


Through inspection of group 13 clusters structures that have either been discovered or conjectured to exist and their QM computed energies<sup>5,36</sup>, we identify seven unique ligands that form the basis for replicating any structure (Fig. 1). The Gibbs free energies of the seven basis cluster in Fig. 1 can be written as a set of coupled linear equations with the ligand energies as unknown variables. Using first-principles computed free energies<sup>37–39</sup> for a set of clusters, the Gibbs free energies of the seven common ligands are fit.

**Validation of group additivity method.** The regression plots comparing the GA predicted and QM computed hydrolysis energies are shown in Fig. 2. The linearity between the QM energy and GA predicted energy ( $R^2 = 0.99$ ) shows that the GA method accurately reproduces the QM computed hydrolysis reaction energies of Al, Ga and Al/Ga aqueous oxo/hydroxo clusters. The methodology, the ligand energies and the validation



**Figure 1 | The seven aqueous clusters used to compute ligand energies.** Seven metal aqueous oxo/hydroxo clusters were used to compute energies for the seven common ligands. The subscripts of  $\mu$  denote the number of metal cations bound to each oxo/hydroxo ligand. Detailed information about the ligands are found in SI.



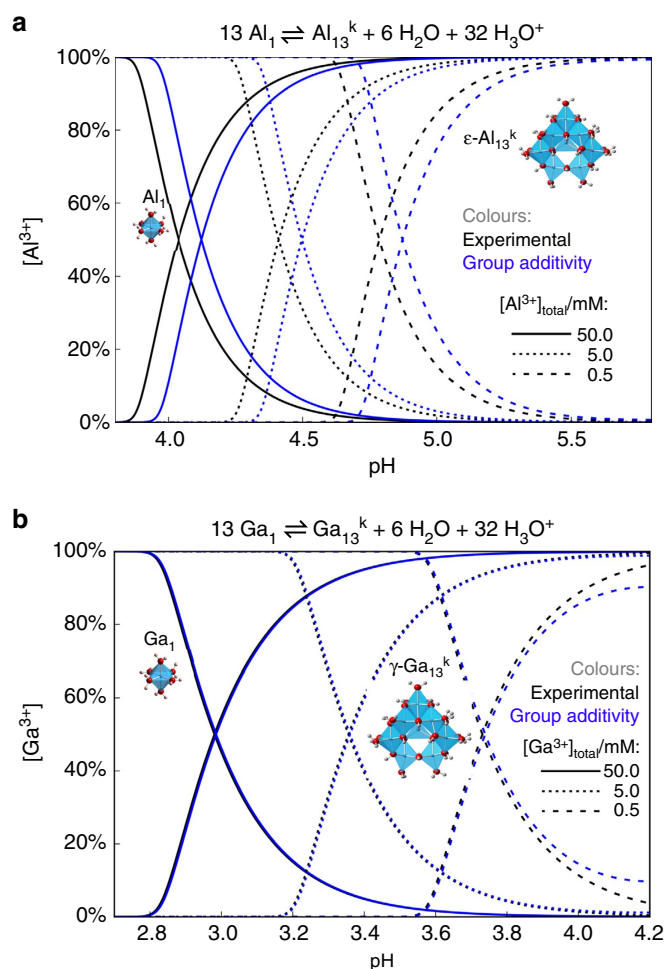
**Figure 2 | GA predicted versus QM computed hydrolysis energies for Al and Ga.** (a) Al and (b) Ga energies. The slope of the linear regression line is near unity (0.98), the intercept is 2.10 and the  $R^2$  value is 0.99. Clusters whose error are larger than  $5 \text{ kcal mol}^{-1}$  are labelled. Al clusters are in cyan, Ga in magenta and heterometallic clusters in black.

are described in more detail in the Supplementary Tables 1–3, Supplementary Figs 1,2 and Supplementary Notes 1,2.

The GA-Pourbaix formalism is able to quantitatively reproduce not only QM free energies, but also qualitatively the experimental free energies and speciation. To verify the accuracy, the equilibrium constant of the hydrolysis reaction leading to the Keggin aluminium or gallium tridecamer cluster was used to predict speciation. Specifically, the ratio of the monomer and Keggin concentrations was calculated with respect to pH and overall concentration of

metal ions. As shown in Fig. 3, the equilibria predicted from the GA method robustly reproduce the reported experimental speciation behaviour of both  $\gamma\text{-Ga}$  and  $\varepsilon\text{-Al}$  Keggin ions.

**Pourbaix diagram for nanoscale clusters.** To introduce the computed clusters to a rigorous thermodynamic framework which allows us to compare their stability with those of competing phases, for example, solids and monomeric aqueous



**Figure 3 | Speciation diagrams as a function of pH and concentrations.** Equilibrium constants are deduced from experiments and the group additivity method. **(a)** Aluminum clusters. **(b)** Gallium clusters.

ions, we adopt the Pourbaix formalism, as implemented in the Materials Project<sup>32</sup>. The formation energies for the different species (experimental aqueous ions and first-principles computed solids) are imported through a common reference state scheme allowing for a grand canonical convex hull of stable aqueous and solid species as a function of pH and hydrogen potential. The Pourbaix diagram is constructed by projecting the convex hull to the  $E$ -pH plane. This formalism has been benchmarked<sup>33</sup> on almost all non-radioactive elements in the periodic table and reproduces or surpasses (due to its access to large amounts of computed solid polymorphs which have not been measured for their thermodynamic properties), the available experimental Pourbaix diagrams. Since launch in 2013, the Materials Project Pourbaix App has been used to successfully predict and interpret the stability of multiple materials and materials classes such as  $\text{Mn}_2\text{V}_2\text{O}_7$  (ref. 40),  $\text{BiVO}_4$  (ref. 41),  $\text{MnNiO}_3$  (ref. 42), Cu vanadates<sup>41</sup>, 2D materials for photocatalyst applications<sup>43</sup> and  $\text{CaCO}_3$  (ref. 4), as a function of aqueous conditions. In this work we extend the framework by directly importing the cluster energies through equations similar to equation (1). The resulting values for  $\mu_{\text{cluster}}^{\text{ref}}$  are provided in Supplementary Table 4. We note that the Materials Project database contains not only the monomeric aqueous ions and stable solid state phases, but also many more metastable solids—against which the cluster Gibbs free energies are being evaluated, as a function of pH and potential<sup>32</sup>. To date, the Materials Project contains 106 unique

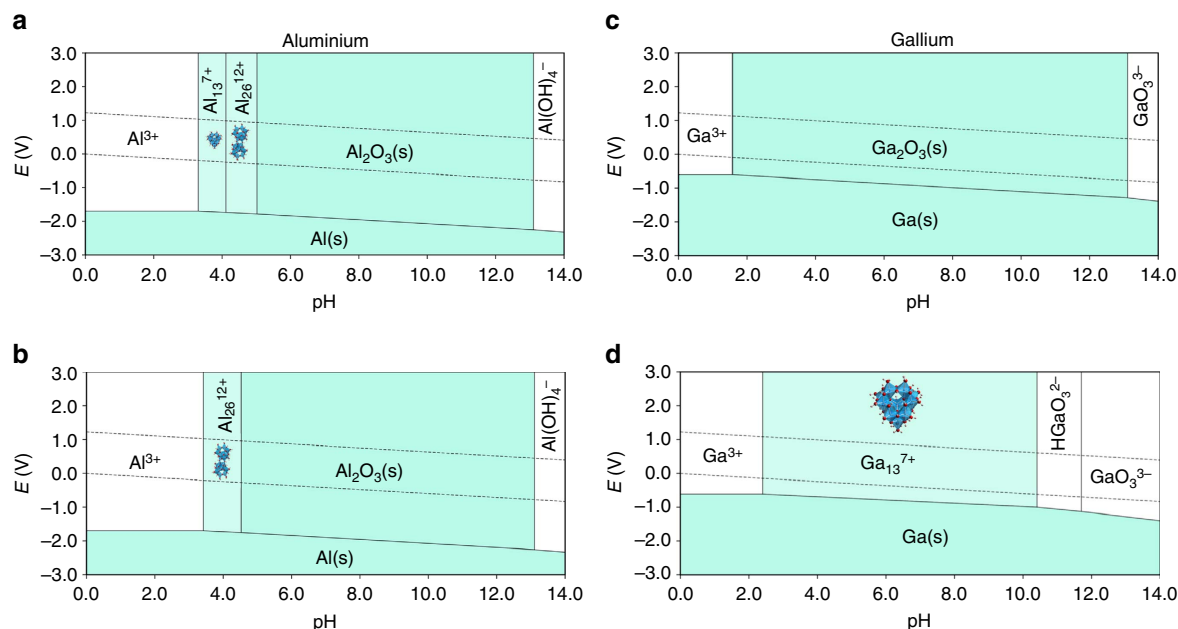
$\text{Al-O}_x\text{-H}_y$ , 41  $\text{Ga-O}_x\text{-H}_y$  solid compounds and four Al as well as seven Ga monomeric aqueous ions. Furthermore, we emphasize that the aqueous species analyses do not include any counter-ion or salt effects in the aqueous solution.

Figure 4a,b show the two resulting GA-Pourbaix diagrams for aluminium which illustrates the evolution of stable species as a function of pH, standard hydrogen potential—at two representative cluster concentrations. Similar to the well-known experimental aluminium Pourbaix diagram<sup>20,21,44</sup>, for  $\text{pH} < 3$ , the monomer  $\text{Al}^{3+}$  is stable in solution, while at high pH, the  $\text{Al(OH)}_4^-$  is preferred. Solid state  $\text{Al}_2\text{O}_3$  is expected to precipitate at neutral conditions and at reducing potentials lower than 2.0 V versus standard hydrogen potential, metallic aluminium represents the equilibrium state in aqueous solutions. Existing experimental Pourbaix diagrams are dominated by solid and monomeric species<sup>20,21</sup>; however, in Fig. 4 we also observe that multimeric aqueous cluster species are predicted to be thermodynamic ground states at slightly acidic conditions. At low cluster concentration ( $10^{-8}$  M) we find the  $\epsilon$ -Al Keggin (#15 in Supplementary Fig. 2) to be the most stable species between  $3.5 < \text{pH} < 4.1$ , while a ‘dimer-Keggin’ cluster ( $\text{Al}_{26}^{15+}$  in Fig. 4a) becomes stable at higher  $\text{pH} = 4.1\text{--}5.2$ . This dimer-Keggin resembles a dimer  $\epsilon$ -Keggin, but is actually a dimer of  $\delta$ -Keggins. The  $\delta$ -Keggin cluster differs from the  $\epsilon$ -Keggin by a  $60^\circ$  rotation of one of the four trimeric groups around the exterior of the  $\epsilon$ -Keggin cluster. At higher cluster concentrations ( $10^{-3}$  M) the dimer-Keggin outcompetes the  $\epsilon$ -Keggin, and stabilizes between  $3.3 < \text{pH} < 4.7$ . We note that larger oligomeric clusters are favoured at higher cluster concentrations, which suggests a route towards cluster aggregation and subsequent formation of a bulk Al oxide or hydroxide (for example,  $\text{AlOOH}$  or  $\text{Al}_2\text{O}_3$ )<sup>10</sup>. Indeed, the  $\delta$ -Keggin form is known to polymerize leading to the formation of larger aluminium polycations, whereof  $\text{Al}_{30}^k$ ,  $\text{Al}_{32}^k$  and possibly  $\text{Al}_{26}^k$ , as suggested here, are examples.

Interestingly, for gallium, we find that while the Keggin  $\text{Ga}_{13}^k$  shows the expected stability regime and speciation with respect to the monomeric  $\text{Ga}^{3+}$  ions (Fig. 3), in the full Pourbaix diagram it is outcompeted by other species, primarily the solid  $\text{Ga}_2\text{O}_3$  by  $\sim 166$  meV per Ga at a representative  $\text{pH} = 3.0$ , and 0.001 M cluster concentration. Hence, at regular concentrations, the gallium Pourbaix diagram—including the clusters—is exactly the same as the traditional one<sup>20,21,44</sup> without any included cluster energies (Fig. 4c). However, by removing the stable phases from the construction of the Pourbaix hull, we simulate scenarios where specific phases are unable to form. For example, if  $\text{Ga}_2\text{O}_3$  nucleation and growth is inhibited, the  $\text{Ga}_{13}^k$  cluster emerges as the most stable phase for  $2.4 < \text{pH} < 10.4$  (Fig. 4d). At lower pH, the monomeric aqueous ion  $\text{Ga}^{3+}$  is stable, while for  $\text{pH} > 10.4$ , first, the  $\text{HGa(OH)}_3^{2-}$  and then  $\text{GaO}_3^{3-}$  form at successively more basic conditions ( $\text{pH} > 11.8$ ).

## Discussion

The two complementary techniques presented in this work provide a robust and efficient framework for calculating the Gibbs free energy and evaluating the thermodynamic stability of clusters in controlled aqueous solutions. For Al, the resulting GA-Pourbaix diagrams advocate the existence of stable nanoscale clusters in aqueous solution which adds pertinent information to the current discussion on non-classical nucleation models<sup>3,7,8</sup>. Figure 4 shows an intuitive succession of the  $\epsilon$ -Keggin ion to dimerization of a similar structure ( $\delta$ -Keggin) as the concentrations of the clusters as well as the monomeric ions are increased. Furthermore, we note that the Al clusters are stabilized without any counter-ion or salt effects, which may be indirectly



**Figure 4 | Pourbaix diagrams with nanoscale aqueous clusters.** The diagrams incorporate solid phases from first-principles, aqueous ions from experiments and clusters from the GA method. In total, 20 Al clusters are evaluated to produce (a and b) and 16 Ga clusters are evaluated to produce (c and d). In (a) the respective concentrations are  $10^{-8}$  M for both monomeric and cluster aluminium species and in (b)  $10^{-2}$  M for monomeric species and  $10^{-3}$  M for the aluminium cluster species. In (c and d), the respective concentrations are  $10^{-2}$  M for monomeric species and  $10^{-3}$  M for the gallium cluster species. In (c) all phases are included whereas in (d) we simulate a situation where  $\text{Ga}_2\text{O}_3$  is kinetically inhibited to nucleate which shows the stability regime of the  $\text{Ga}_{13}$  Keggin. The concentrations for (a–d) refer to the metal ion and corresponds to the solubility of the stable species at the phase boundary. Hence the concentrations should be treated as a lower limit of solubility at the given potential and pH.

supported by the variety of different solutions in which Al Keggin clusters have been observed<sup>45–48</sup>.

For gallium we find that the  $\text{Ga}_{13}$  Keggin successfully competes with the  $\text{Ga}^{3+}$  monomeric aqueous ion, as the computed speciation diagram reproduces experimental trends (Fig. 3b). However, in the full GA-Pourbaix diagram, other species such as the aqueous ion  $\text{Ga}(\text{OH})_2^+$  and the solid  $\text{Ga}_2\text{O}_3$  phase are thermodynamically more stable for the pH range where the Ga clusters may emerge. Hence, the dilute GA-Pourbaix diagram—including cluster energies from 16 Ga clusters—is the same as the traditional experimental Pourbaix diagram<sup>20,21,44</sup>. However, simulating the stability of species under conditions where, for example, we assume  $\text{Ga}_2\text{O}_3$  nucleation to be kinetically hindered, we obtain the GA-Pourbaix diagram in Fig. 4d, where the  $\text{Ga}_{13}$  Keggin stability regime is evident.

The combined results of the aluminium and gallium GA-Pourbaix diagrams prompt us to the following comments. First of all, it is important to note that the results displayed in Fig. 4 assume non-interacting, dilute conditions and do not include any counter-ion effects, which are expected to selectively influence the stability of different clusters. Al Keggin clusters are clearly thermodynamically stable even under these conditions, while Ga clusters may be dependent on salt effects to outcompete the stable phases in the GA-Pourbaix diagram. For example, we stress that the energy difference between the  $\text{Ga}_{13}^k$  and  $\text{Ga}_2\text{O}_3$  as well as monomeric aqueous specie  $\text{Ga}(\text{OH})_2^+$  is 166 and 72 meV per Ga; respectively, which is within the energy of counter-ion interactions. In general, analysing the free energy difference between competing clusters as well as other possible species opens up the possibility to rationally influence cluster stability by tuning the solution composition. For example, the ‘flat’  $\text{Al}_{13}$  cluster is predicted to be thermodynamically unstable with respect to the  $\varepsilon$ -Keggin by 30 meV per Al. However, this cluster has been synthesized and observed under several different conditions<sup>5,49</sup>.

In addition, the intermediates that have been suggested to give rise to the ‘flat’  $\text{Al}_{13}$  cluster are not stable from these QM or GA calculations<sup>5</sup>. We speculate that there are many more metastable clusters that may be stabilized by strong salt association and ion-pairing in solution. Available experimental results also suggest that different cluster structures, that may or may not be related to the macroscopic bulk<sup>5,50</sup>, form under different aqueous conditions. In particular, nanoscale clusters exhibiting building units resembling metastable bulk materials may pre-condition the synthesis of target nanoparticles<sup>5,11</sup>, thin films<sup>2</sup> and bulk materials with select structure and functionality<sup>18</sup>.

Finally, we would also like to emphasize that no kinetic effects are included, which are known to be important in many of the observed nucleation processes, and for determining the specific condition-dependent mechanism by which the final crystal and its morphology is formed. Nevertheless, the presented results provide a rigorous thermodynamic framework for the analysis of stable species in solution against which comparison and interpretation of experimental results can be made.

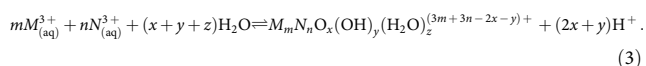
A general predictive computational framework to predict the vast array of aqueous metal clusters found in natural ecosystems, technological applications and biology is a long-standing challenge. In this work, we present the results of combining two procedures to improve our ability to predict and interpret cluster formation in aqueous solutions. First, we disclose a general robust GA method to calculate the free energy of inorganic clusters, which reproduces experimental equilibrium constants and speciation, rival the accuracy of quantum mechanically computed values, and—once developed for a particular chemical system—can be performed in seconds. The speciation diagrams comparing GA, quantum mechanics and experimental results show that the GA method is a general, accurate and expedient method to predict the stabilities of group 13 metal aqua-oxo/hydroxo clusters. Furthermore, we import the resulting cluster

thermodynamic data into the Materials Project Pourbaix framework, which provides a hybrid approach to classical Pourbaix diagrams—allowing for computed and experimental species to be evaluated within a single-reference state system. The result supports the emerging experimental evidences<sup>3,7–9,12,18</sup> of thermodynamically stable and mildly metastable solvated clusters, which may form as part of pre-nucleation pathways in non-CNT<sup>10</sup>. In the case of Al, we find that the  $\epsilon$ -Keggin and ‘dimer’  $\delta$ -Keggin outcompetes monomeric aqueous species at mildly acidic conditions and low cluster concentrations, whereas higher concentrations favour the dimer-Keggin alone. In the case of Ga, we find no cluster species to be more stable than the monomeric ions and the bulk  $\text{Ga}_2\text{O}_3$  solid under dilute, salt-free water conditions. However, the Keggin  $\text{Ga}_{13}$  cluster is found to be more stable than the monomeric  $\text{Ga}^{3+}$  under mildly acid conditions and between  $2.4 < \text{pH} < 10.4$  is the second most stable species, after bulk  $\text{Ga}_2\text{O}_3$ . We emphasize that the current framework does not include interactions with other species in solution, such as counter-ions, which are likely to provide an excellent tuning parameter to change the hierarchy of stable clusters and promote nucleation of metastable structures. We also stress that kinetic factors and the actual atomistic path towards nucleation, including structural rearrangement of cluster collectives, are beyond the current analysis. Nevertheless, this GA-Pourbaix methodology will contribute towards improved understanding on nucleation by providing a quantitative framework to aid analysis and ultimately guide rational synthesis of metastable clusters and structures in solution.

## Methods

**Quantum mechanical method for group additivity.** The ground-state structures and Gibbs free energies are computed by HF/6-31G(d,p) with the IEFPCM-UFF continuum solvation model for water. The electronic energy was refined using the B3LYP/6-311 + G(d) single point, and the solvation energy was recomputed using HF/6-311 + G(d) with CPCM-UAKS for water.

**Pourbaix diagram formalism.** To integrate the GA free energies into the Persson Pourbaix formalism we express the clusters formation with respect to the monomeric aqueous ions:



The corresponding free energy of cluster hydrolysis is then obtained by:

$$\Delta G_{\text{hydrolysis}}^{\circ} = \mu_{\text{M}_m\text{N}_n\text{O}_x(\text{OH})_y(\text{H}_2\text{O})_z^{(3m+3n-2x-y)+}}^{\text{ref}} + 0 - m \cdot \mu_{\text{M}_{\text{aq}}^{3+}}^{\text{ref}} - n \cdot \mu_{\text{N}_{\text{aq}}^{3+}}^{\text{ref}} - (x + y + z) \cdot \mu_{\text{H}_2\text{O}}^{\text{ref}} \quad (4)$$

where  $\mu_{\text{M}_m\text{N}_n\text{O}_x(\text{OH})_y(\text{H}_2\text{O})_z^{(3m+3n-2x-y)+}}^{\text{ref}}$ ,  $\mu_{\text{M}_{\text{aq}}^{3+}}^{\text{ref}}$ ,  $\mu_{\text{N}_{\text{aq}}^{3+}}^{\text{ref}}$  and  $\mu_{\text{H}_2\text{O}}^{\text{ref}}$  are the chemical potentials of the cluster, protons, metal ions and  $\text{H}_2\text{O}$ , respectively. All energies are provided at the standard state, and conform the consistent reference framework by Persson *et al.*<sup>29</sup>  $\mu_{\text{M}_{\text{aq}}^{3+}}^{\text{ref}}$ ,  $\mu_{\text{N}_{\text{aq}}^{3+}}^{\text{ref}}$  and  $\mu_{\text{H}_2\text{O}}^{\text{ref}}$  are known properties, while  $\Delta G_{\text{hydrolysis}}^{\circ}$  is calculated by the GA method presented here. Hence,  $\mu_{\text{M}_m\text{N}_n\text{O}_x(\text{OH})_y(\text{H}_2\text{O})_z^{(3m+3n-2x-y)+}}^{\text{ref}}$  is obtained from equation (4):

$$\mu_{\text{M}_m\text{N}_n\text{O}_x(\text{OH})_y(\text{H}_2\text{O})_z^{(3m+3n-2x-y)+}}^{\text{ref}} = \Delta G_{\text{hydrolysis}}^{\circ} + m \cdot \mu_{\text{M}_{\text{aq}}^{3+}}^{\text{ref}} + n \cdot \mu_{\text{N}_{\text{aq}}^{3+}}^{\text{ref}} + (x + y + z) \cdot \mu_{\text{H}_2\text{O}}^{\text{ref}} \quad (5)$$

**Code availability.** The code to plot Pourbaix diagram is available via <https://github.com/materialsproject>.

**Data availability.** The data that support the findings of this study are available within the paper and its supplementary information file.

## References

1. Hanrahan, G. in *Modelling of Pollutants in Complex Environmental Systems*, Vol. 2, Advanced Topics in Environmental Science (ILM Publications, 2010).

2. Meyers, S. T. *et al.* Solution-processed aluminum oxide phosphate thin-film dielectrics. *Chem. Mater.* **19**, 4023–4029 (2007).
3. Gebauer, D., Volkel, A. & Colfen, H. Stable prenucleation calcium carbonate clusters. *Science* **322**, 1819–1822 (2008).
4. Sun, W., Jayaraman, S., Chen, W., Persson, K. A. & Ceder, G. Nucleation of metastable aragonite  $\text{CaCO}_3$  in seawater. *Proc. Natl Acad. Sci. USA* **112**, 3199–3204 (2015).
5. Wang, W. *et al.* Electrolytic synthesis of aqueous aluminum nanoclusters and in situ characterization by femtosecond raman spectroscopy and computations. *Proc. Natl Acad. Sci. USA* **110**, 18397–18401 (2013).
6. Yokel, R. A. Brain uptake, retention, and efflux of aluminum and manganese. *Environ. Health Perspect.* **110**, 699–704 (2002).
7. Nielsen, M. H., Aloni, S. & De Yoreo, J. J. *In situ* TEM imaging Of  $\text{CaCO}_3$  nucleation reveals coexistence of direct and indirect pathways. *Science* **345**, 1158–1162 (2014).
8. Pouget, E. M. *et al.* The initial stages of template-controlled  $\text{CaCO}_3$  formation revealed by cryo-TEM. *Science* **323**, 1455–1458 (2009).
9. Navrotsky, A. Energetic clues to pathways to biomineralization: precursors, clusters, and nanoparticles. *Proc. Natl Acad. Sci. USA* **101**, 12096–12101 (2004).
10. Gebauer, D., Kellermeier, M., Gale, J. D., Bergström, L. & Cölfen, H. Pre-nucleation clusters as solute precursors in crystallisation. *Chem. Soc. Rev.* **43**, 2348 (2014).
11. Nielsen, M. H. *et al.* Investigating processes of nanocrystal formation and transformation via liquid cell TEM. *Microsc. Microanal.* **20**, 425–436 (2014).
12. Habraken, W. J. E. M. *et al.* Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nat. Commun.* **4**, 1507 (2013).
13. Grandjean, D., Beale, A. M., Petukhov, A. V. & Weckhuysen, B. M. Unraveling the crystallization mechanism of CoAPO-5 molecular sieves under hydrothermal conditions. *J. Am. Chem. Soc.* **127**, 14454–14465 (2005).
14. Bradley, S. M., Kydd, R. A. & Yamdagni, R. Detection of a new polymeric species formed through the hydrolysis of gallium(III) salt solutions. *J. Chem. Soc. Dalton Trans.* **1989**, 413 (1990).
15. Sadeghi, O., Zakharov, L. N. & Nyman, M. Aqueous formation and manipulation of the iron-oxo Keggin ion. *Science* **347**, 1359–1362 (2015).
16. Legg, B. A., Zhu, M., Comolli, L. R., Gilbert, B. & Banfield, J. F. Determination of the three-dimensional structure of ferrihydrite nanoparticle aggregates. *Langmuir* **30**, 9931–9940 (2014).
17. Chiavacci, L. A., Santilli, C. V., Pulcinelli, S. H., Bourgaux, C. & Briois, V. Role of the surface state and structural feature in the thermoreversible sol–gel transition of a zirconyl aqueous precursor modified by sulfuric acid. *Chem. Mater.* **16**, 3995–4004 (2004).
18. Stewart, T. A. *et al.* Enhanced water purification: a single atom makes a difference. *Environ. Sci. Technol.* **43**, 5416–5422 (2009).
19. Baumgartner, J. *et al.* Nucleation and growth of magnetite from solution. *Nat. Mater.* **12**, 310–314 (2013).
20. Brookins, D. G. *Eh-pH Diagrams for Geochemistry* (Springer-Verlag, 1988).
21. Takeno, N. *Atlas of Eh-pH Diagrams. Intercomparison of Thermodynamic Databases*. Geological Survey of Japan Open File Report No. 419. (National Institute of Advanced Industrial Science and Technology Research Center for Deep Geological Environments, 2005). <https://www.gsj.jp/researches/openfile/openfile2005/openfile0419.html>.
22. Schweitzer, G. K. & Pesterfield, L. L. *The Aqueous Chemistry of the Elements* (Oxford University Press, 2010).
23. Lee, J., Yang, J., Kwon, S. G. & Hyeon, T. Nonclassical nucleation and growth of inorganic nanoparticles. *Nat. Rev. Mater.* **1**, 16034 (2016).
24. Casey, W. H. & Rustad, J. R. Reaction dynamics, molecular clusters, and aqueous geochemistry. *Annu. Rev. Earth Planet. Sci.* **35**, 21–46 (2007).
25. Tiferet, E. *et al.* The energy landscape of uranyl-peroxide species. *Chem. Eur. J.* **20**, 3646–3651 (2014).
26. López, X., Nieto-Draghi, C., Bo, C., Avalos, J. B. & Poblet, J. M. Polyoxometalates in solution: molecular dynamics simulations on the  $\alpha$ -PW<sub>12</sub>O<sub>40</sub><sup>3-</sup> Keggin anion in aqueous media. *J. Phys. Chem. A* **109**, 1216–1222 (2005).
27. Cameron, J. M. *et al.* Investigating the transformations of polyoxoanions using mass spectrometry and molecular dynamics. *J. Am. Chem. Soc.* **138**, 8765–8773 (2016).
28. Benson, S. W. & Buss, J. H. Additivity rules for the estimation of molecular properties. thermodynamic properties. *J. Chem. Phys.* **29**, 546 (1958).
29. Benson, S. W. *et al.* Additivity rules for the estimation of thermochemical properties. *Chem. Rev.* **69**, 279–324 (1969).
30. Vicente, M. A., Belver, C., Sychev, M., Prihod'ko, R. & Gil, A. Relationship between the surface properties and the catalytic performance of Al-, Ga-, and AlGa-pillared saponites. *Ind. Eng. Chem. Res.* **48**, 406–414 (2009).
31. Jakupec, M. A. & Keppler, B. K. Gallium in cancer treatment. *Curr. Top. Med. Chem.* **4**, 1575–1583 (2004).

32. Jain, A. *et al.* Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 11002 (2013).
33. Persson, K. A., Waldwick, B., Lazic, P. & Ceder, G. Prediction Of solid-aqueous equilibria: scheme to combine first-principles calculations of solids with experimental aqueous states. *Phys. Rev. B* **85**, 235438 (2012).
34. Öhman, L.-O., Forsling, W., Songstad, J., Sjöblom, J. & Thylin, B. Equilibrium and structural studies of silicon(IV) and aluminium(III) in aqueous solution. 3. Potentiometric study of aluminium(III) hydrolysis and aluminium(III) hydroxo carbonates in 0.6 M NaCl. *Acta Chem. Scand.* **35a**, 795–802 (1981).
35. Mensinger, Z. L. *et al.* Synthesis of heterometallic group 13 nanoclusters and inks for oxide thin-film transistors. *Angew. Chemie Int. Ed.* **47**, 9484–9486 (2008).
36. Mensinger, Z. L., Wang, W., Keszler, D. A. & Johnson, D. W. Oligomeric group 13 hydroxide compounds—a rare but varied class of molecules. *Chem. Soc. Rev.* **41**, 1019–1030 (2012).
37. Frisch, M. J. *et al.* Gaussian 09 (Gaussian, Inc., Wallingford CT, 2009).
38. Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. *Ab initio* calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **98**, 11623–11627 (1994).
39. Tomasi, J., Mennucci, B. & Cancès, E. The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM *ab initio* level. *J. Mol. Struct. THEOCHEM* **464**, 211–226 (1999).
40. Yan, Q. *et al.*  $\text{Mn}_2\text{V}_2\text{O}_7$ : an earth abundant light absorber for solar water splitting. *Adv. Energy Mater.* **5**, 1401840 (2015).
41. Toma, F. M. *et al.* Mechanistic insights into chemical and photochemical transformations of bismuth vanadate photoanodes. *Nat. Commun.* **7**, 12012 (2016).
42. Yu, J. *et al.* First-principles study of electronic structure and photocatalytic properties of  $\text{MnNiO}_3$  as an alkaline oxygen-evolution photocatalyst. *Chem. Commun.* **51**, 2867–2870 (2015).
43. Singh, A. K., Mathew, K., Zhuang, H. L. & Hennig, R. G. Computational screening of 2D materials for photocatalysis. *J. Phys. Chem. Lett.* **6**, 1087–1098 (2015).
44. Pourbaix, M. *Atlas Of Electrochemical Equilibria In Aqueous Solutions* (National Association of Corrosion Engineers, 1974).
45. Johansson, G., Gullman, L.-O., Kjekshus, A. & Söderquist, R. On the crystal structure of some basic aluminium salts. *Acta Chem. Scand.* **14**, 771–773 (1960).
46. Fu, G., Nazar, L. F. & Bain, A. D. Aging processes of alumina sol-gels: characterization of new aluminum polyoxycations by aluminum-27 NMR spectroscopy. *Chem. Mater.* **3**, 602–610 (1991).
47. Armstrong, C. R., Casey, W. H. & Navrotsky, A. Energetics Of Al Keggin cluster compounds. *Proc. Natl Acad. Sci. USA* **108**, 14775–14779 (2011).
48. Drljaca, A., Hardie, M. J. & Raston, C. L. Selective isolation of Keggin ions using self-assembled superanion capsules. *J. Chem. Soc. Dalton Trans* 3639–3642 (1999).
49. Wang, W., Wentz, K. M., Hayes, S. E., Johnson, D. W. & Keszler, D. A. Synthesis of the hydroxide cluster  $[\text{Al}_{13}(\mu_3\text{-OH})_6(\mu\text{-OH})_{18}(\text{H}_2\text{O})_{24}]^{15+}$  from an aqueous solution. *Inorg. Chem.* **50**, 4683–4685 (2011).
50. Casey, W. H., Olmstead, M. M. & Phillips, B. L. A new aluminum hydroxide octamer,  $[\text{Al}_8(\text{OH})_{14}(\text{H}_2\text{O})_{18}](\text{SO}_4)_5 \cdot 16\text{H}_2\text{O}$ . *Inorg. Chem.* **44**, 4888–4890 (2005).

## Acknowledgements

P.H.-Y.C. is the Bert and Emelyn Christensen Professor of Chemistry at Oregon State University and gratefully acknowledges financial support from the Vicki & Patrick F. Stone Family and the National Science Foundation (NSF, CHE-1352663). L.A.W. acknowledges financial support from OSU Hedberg Fellowship. All authors acknowledge financial support and computing infrastructure in part provided by the NSF Phase-2 CCI, Center for Sustainable Materials Chemistry (NSF CHE-1102637).

## Author contributions

L.A.W. and P.H.-Y.C. developed the group additivity method. L.A.W. performed the quantum mechanical calculations. X.Q. and K.A.P. developed the formalism to integrate Pourbaix diagram with clusters. X.Q. implemented the code base and plotted the diagram. K.A.P., L.A.W., X.Q. and P.H.-Y.C. wrote the final manuscript. All authors participated in the final version of the article.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Wills, L. A. *et al.* Group additivity-Pourbaix diagrams advocate thermodynamically stable nanoscale clusters in aqueous environments. *Nat. Commun.* **8**, 15852 doi: 10.1038/ncomms15852 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017