RESEARCH ARTICLE

Predicting solubility of ion pairs in aqueous inorganic chemistry

Tasnim Rahman,[b]Ψ Enric Petrus,[a] Ψ Mireia Segado,[a] Nicolas P. Martin, [b] Lauren Palys, [b] Mark A. Rambaran,[c] C. Andre Ohlin, [c] Carles Bo*[a,d] May Nyman*[b]

[a] Enric Petrus, Dr. Mireia Segado, Dr. Professor Carles Bo
Institute of Chemical Research of Catalonia (ICIQ), Barcelona Institute of Science & Technology (BIST)
Av. Països Catalans, 16
43007 Tarragona (Spain)
email: cbo@iciq.cat

[b] Tasnim Rahman, Dr. Nicolas P. Martin, Lauren N. Palys, Dr. Professor May Nyman
Department of Chemistry
Oregon State University
Corvallis, OR 97331 USA
email: may.nyman@oregonstate.edu

[c] Mark A. Rambaran and Dr. Professor C. Andre Ohlin
Department of Chemistry
Faculty of Science and Technology
Umeå University
901 87 Umeå, Sweden.

[d] Enric Petrus, Dr. Professor Carles Bo
Departament de Química Física i Inorgànica, Universitat Rovira i Virgili (URV)
Marcel·lí Domingo s/n
43007 Tarragona (Spain)

Ψ These two authors contributed equally as co first-authors

Supporting information for this article is given via a link at the end of the document

Abstract: Polyoxometalates (POMs), ranging in size from several to 100's of ångströms, represent building blocks of inorganic materials. Elucidating their complex solubility behavior with alkali countercations can inform and model natural and synthetic processes in aqueous media. In the study of POMs ($[Nb_{24}O_{72}H_9]^{15-}$, Nb_{24}) we discovered an unusual solubility trend (termed anomalous solubility) of alkali-POMs, in which Nb_{24} is most soluble with the smallest (Li^+) and largests (Rb/Cs^+) alkalis, and least soluble with Na/K^+. Via computation, we define a descriptor (σ-profile) and use an artificial neural network (ANN) to accurately predict all three described alkali-anion solubility trends: amphoteric, normal ($Li^+ > Na^+ > K^+ > Rb^+ > Cs^+$), and anomalous ($Cs^+ > Rb^+ > K^+ > Na^+ > Li^+$). Testing predicted amphoteric solubility along with X-ray scattering and single-crystal X-ray diffraction studies affirmed the accuracy of the desciptor, provided solution-phase snapshots of alkali-POM interactions, yielded a new POM formulated $[Ti_6Nb_{14}O_{54}]^{14-}$, and provided guidlines to exploit alkali-POM interactions for new POMs discovery.

Introduction

Interactions of cations or anions with aqueous macromolecules or interfaces orchestrate many natural and synthetic processes. Natural processes influenced by ion association include DNA helix-formation and stabilization,[1] protein folding,[2] and fate and transport of contaminants in the environment.[3] Aqueous synthesis of inorganic materials (i.e. zeolites, oxides, phosphates, etc.[4]) involves dissolution, self-assembly, precipitation and crystallization over the course of the reaction; and all these processes are dictated by interactions of ions with the growing nuclei. In turn, aqueous cation-anion pairing directly affects aforementioned interfacial interactions and crystal growth, since the first step to dissolution is separating ions, and the first step to precipitation or crystallization is forming ion-pairs in solution.

Lattice energy, e.g. calculated by the Born-Haber cycle, is sometimes a predictor of solubility for simple binary compounds, where low solubility equates with high lattice energy. Such models less effectively predict the energy of lattices containing complex ions; i.e. oxoanions with multiple modes of interaction with each other and with the counter-cations. While solid-state structures provide a starting point for modeling ion-interactions in solution, rigorous experimental characterization is challenging. These interactions are dynamic, and usually observed as an averaged 'snapshot' by experimental measurements including X-ray scattering, calorimetry, and vibrational spectroscopy, coupled with simulations.[5] An additional complicating factor of understanding ion-pairs and solubility of solids in water is the interactions between ions and water. This concept embodies the ability of ions to polarize water molecules and influence hydration spheres that prevent ion interactions; as well as the acid-base behavior of oxoanions and multivalent metal cations. Deprotonation of water molecules bound to metal cations, or protonation of oxoanions leads to pH changes and different modes of ion interaction, i.e. by hydrogen-bonding. Together these phenomena; hydration spheres, hydrogen-bonding, ion-pairing, lattice energy, and acid-base properties, engender very complex behavior, and there is no algorithm that accurately predicts solubility of any cation-anion pair in water, or as addressed here, *solubility trends* of anions with the alkali series.

The Hofmeister series[6] was defined empirically to describe the effect of simple ions on protein solubility, and this is sometimes adopted in discussion of aqueous-phase behavior of other macromolecules.[7] This series was developed to describe the ability of ions to 'salt-in' (dissolve) and 'salt-out' (precipitate) proteins. While poorly understood, the order of ion-protein solubility can be correlated with hydration spheres of the ions.[8] A similar classification of ions is kosmotropic and chaotropic.[9] Kosmotropic ions generally have high charge density and dictate assembly of hydration spheres. Chaotropic ions are large with low charge, and weak interaction with water. Polyoxometalates (POMs, anionic metal-oxo clusters of the early d^0-tranistion metals) are deemed superchaotropes, and they exhibit hydrophobicity, evidenced by adsorption onto hydrophobic surfaces from water.[10] Elaborated computational models for simple ions (halides and alkali cations) reproduce well both the normal and the reversed Hofmeister trends when ions interact with model hydrophobic or hydrophilic surfaces. (NEW REF 11: N. Schwierz, D. Horinek, R. R. Netz *Langmuir* **2013**, *29*, 2602−2614). Yet these models that attempt to predict ion-interactions with surfaces and macromolecules must consider anion-cation pair interactions coupled with solubility, because one cannot exist in solution without the other

POMs are excellent tools to understand and predict solubility behavior of ions, specifically in the realm of aqueous synthesis of inorganic materials. They can serve as models for soluble metal-oxo complexes that are intermediates to metal oxides in synthesis. They are large and contain heavy metals, conducive to X-ray scattering experiments that reveal interactions between ions in solution, as well as any change in speciation. Although they are prior-described as superchaotropes; the immense compositional and structural diversity of this family of polyanions yields behavior that spans from chaotropic to kosmotropic. For example, we identified opposing solubility trends of POMs with alkali countercations.[11] Whereas W, V, Mo and uranyl POM solubility with alkalis mostly trend Li>Na>K>Rb>Cs; Nb and Ta POM solubility trends are generally opposite. We have termed the former trend 'normal solubility' because it is what we expect based on alkali-cation hydration spheres. For example, Li^+ has a large hydration sphere, preventing formation of ion pairs, while Cs^+ readily forms ion pairs, initiating precipitation. The latter trend we have termed 'anomalous solubility'. Anomalous solubility also describes simple oxoanion-salts of PO_4^{3-} and CO_3^{2-}; Cs-salts are highly soluble for both, while Li^+-salts are insoluble. Recognition of these opposing solubility trends leads to many questions including; 1) How do cation-anion-water interactions differ for ion-pairs with opposite trends? 2) Can we identify a single descriptor to computationally predict solubility trends?

Here we have (a) Experimentally identified a third solubility trend, amphoteric solubility for the Nb_{24}-POM[12] $(Nb_{24}O_{72}H_9)^{15-}$ (Na/K^+-salts are least soluble); (b) Found that computed □-profiles of POMs and simple oxoanions are molecular descriptors that enabled developing a machine learning model to accurately predict normal, anomalous, and amphoteric solubility trends; (c) Tested the predictive model experimentally, confirming amphoteric solubility for $[TiNb_9O_{28}]^{7-}$ and $[Ti_2Nb_8O_{28}]^{8-}$ (respectively $TiNb_9$ and Ti_2Nb_8),[13] whose □-profiles are similar to that of Nb_{24}. Solution phase small-angle X-ray scattering (SAXS) of Nb_{24}, $TiNb_9$ and Ti_2Nb_8 plus isolation of a new titanoniobate POM, $[Ti_6Nb_{14}O_{54}]^{14-}$ (Ti_3Nb_7-dimer) suggest POMs with amphoteric solubility, predicted by their corresponding □-profiles, are dynamic in solution and provide opportunity for new POM discovery.

Results and Discussion

Discovering amphoteric solubility trend. Recently we showed that Nb_{24}[12b] forms upon addition of an alkali chloride salt to a solution of tetramethylammonium decaniobate $[TMA_6(Nb_{10}O_{28})\cdot xH_2O]$, Nb_{10}. SAXS identifies the Nb_{24} units as dimer or tetramer forms, where alkali-cations promote linking in solution. On the other hand, alkali carbonate (instead of chloride) yields solutions of predominantly monomeric Nb_{24} units (**fig. 1a**),[14] appropriate for comparing solubility behavior. Due to the insolubility of lithium carbonate, we instead used $LiNO_3$, and the solution was buffered to pH-match that of the Nb_{10}-alkali carbonate solutions, which also yielded the Nb_{24}-monomer (see SI for details). The maximum solubility of Nb_{24} as a function of alkali added is summarized in **Table S1** (and **fig. 3** discussed later). Note; solubility of A^+-**Nb_{24}** exhibits the same trend with alkali nitrate at two different alkali:POM ratios, and both alkali carbonate and alkali nitrate (**Table S1**), highlighting the robustness of this trend. However, SAXS studies were performed on the alkali carbonate series. SAXS in **fig. 1a** illustrates solution scattering at maximum concentration, intensity normalized for comparison. The solubility trends Li>Na<K<Rb<Cs, with solubility minimized at Na, but generally very high concentrations (2.5 molar Nb with Li, 3.5 molar Nb with Cs) are observed.

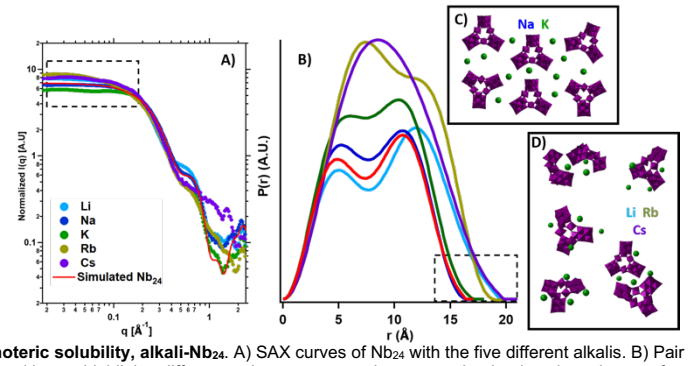


Figure 1. Solution speciation for amphoteric solubility, alkali-Nb_{24}. A) SAX curves of Nb_{24} with the five different alkalis. B) Pair distance distribution function (PDDF) of scattering curves. For A, and B, black dotted boxes highlights differences between scattering curves that leads to hypotheses of amphoteric solubility origin. C) Schematic of proposed Na-Nb_{24} and K-Nb_{24} solution arrangement, leading to the observed SAXS curves and lowest solubility. D) Schematic of proposed Li-Nb_{24}, Rb-Nb_{24} and Cs-Nb_{24} solution arrangements, leading to the observed SAXS curves and higher solubility. Differences between scattering curves at q>1 $Å^{-1}$ is related to solvent and alkali-scattering. Excessively high noise for the Cs-POM solution is due to X-ray absorption of this nucleus.

The SAXS data, in addition to confirming the Nb_{24} speciation in the presence of each alkali, reveal some subtle differences that can help understand the origin of the amphoteric solubility trend. The curves are intensity-normalized to match the Guinier region (the 'elbow' at q~0.08-0.4 $Å^{-1}$), which gives an approximation of the cluster size. This comparison shows good agreement with all alkalis, and agreement with the simulated data, indicating the dominant cluster speciation is the same. By aligning the Guinier region, differences in the low-q region becomes apparent, highlighted in the dashed box in **fig. 1a**. The Nb_{24}-solutions that exhibit the lowest solubility across the alkali-series (Na and K) have a negative deviation from the simulated curve, indicating ordering in solution based on anion-anion repulsion, incompletely shielded by countercations (illustrated in **fig. 1c**). With alkalis Li, Rb, and Cs, Nb_{24}-scattering exhibits a positive deviation relative to simulated scattering, as well as slight shifts to lower-q in the Guinier region. This indicates stronger interactions with alkalis, leading to some aggregate formation (i.e. of two or more POMs), and this is illustrated in **fig. 1d**. The pair-distance distribution function analyses (PDDF, probability distribution map of scattering vectors through dissolved species and aggregates, **fig. 1b**) provide a real-space representation of the scattering particles and size information (radius of gyration, R_g obtained from the fit **Table S4, figure S2**). The maximum linear extent (where probability, (P(r)), goes to 0, representing largest diameter) is highlighted in the dashed black box. The maximum linear extent for the simulated Nb_{24} and experimental Na-Nb_{24} and K-Nb_{24} is ~17 Å. The PDDFs of Li-Nb_{24}, Rb-Nb_{24} and Cs-Nb_{24} have a maximum linear extent approaching 20 Å, due to some clustering of Nb_{24} and association of the alkalis (**fig. 1d**). Other PDDF differences less important to solubility are described in the SI.

*How are these differences illustrated in **figs 1c** and **1d** related to solubility and why amphoteric solubility*? Ordering of species in solution with regular spacing of cations and anions can represent a first step toward crystallization or precipitation, thereby leading to poorer solubility for Na-Nb_{24} and K-Nb_{24} (**fig. 1c**). For A^+ =Li, Rb and Cs, A^+-Nb_{24} interactions partially neutralize the cluster charge,

minimizing cation-anion attractions between aggregates in solution. In addition, there is polydispersity, prohibiting ordering into a solid lattice. To understand *why* there is amphoteric solubility, we need to consider the hydration spheres of both cations and anions. The alkalis have well-defined hydration spheres and associated hydration energies; both increasing from Cs^+ to Li^+, directly related to increasing charge-density of the alkali-cation. For any oxoanion, it is far more complex, as protonation, basicity of individual oxo-ligands, and shape influence hydration spheres, but we can form hypotheses based on their interactions with predictable alkalis. For contact ion-pairing of Nb_{24} with Li, Rb and Cs to occur (as suggested by SAXS), the alkalis must be polarizing enough to strip the POMs' hydration sphere (Li^+); or have a weak hydration sphere (Rb/Cs^+), easily polarized and 'stripped' by the POM. Alkalis of intermediate polarizing strength (Na and K^+) have either hydration shells too strongly bound, or are not polarizing enough to strip the solvation shells of the POMs to promote interactions, inhibiting the partial neutralization of the POM that may promote higher solubility observed for $Li/Rb/Cs^+$.

Computational approach: In the quest for predicting the solubility of ionic species, after many years of efforts,[15] computational methods achieved considerable success for neutral molecules.[16] Dealing with the solubility of salts is much more cumbersome since the combinations of cation/anion/solvent all affect the properties of such solutions. However, the solubility trend we are discussing herein is an anion-dependent property only. Therefore, we sought a computational model that could correlate the properties of the electron density distribution of an anion with its solubility behaviour. We considered a set of representative anions (45 species), which include 1) simple anions (i.e. Cl^-), 2) mono-oxoanions (i.e. $[MO_x]^{y-}$), 3) common polyoxometalates of different shapes and charges,[17] 4) metal halides hexamers,[18] and 5) uranyl peroxide nanocapsules,[11c] a selection illustrated in **figure 2a**. By including such a variety of structures and compositions in the dataset, we expect to model anions spanning more than one order of magnitude in size (i.e. approx. 1 to 20 Å) and charge density with parallel range.

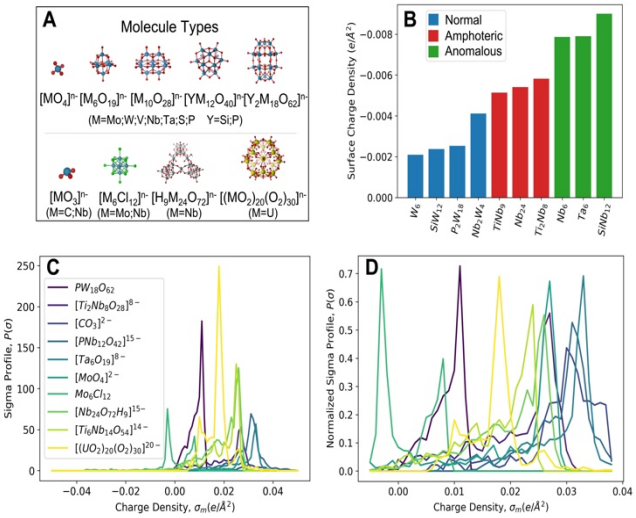


Figure 2. Sigma profiles of test set of oxoanions . **(A)** Type of molecules present in this work: simple anions, POMs, metal halides and derived structures. **(B)** Surface Charge Density (SCD) of 10 selected POMs ($[W_6O_{19}]^{2-}$, $[SiW_{12}O_{42}]^{6-}$, $[P_2W_{18}O_{62}]^{6-}$, $[Nb_2W_4O_{19}]^{4-}$, $[TiNb_9O_{28}]^{7-}$, $[H_9Nb_{24}O_{72}]$, $[Ti_2Nb_8O_{28}]^{8-}$, $[Nb_6O_{19}]^{8-}$, $[Ta_6O_{19}]^{8-}$ and $[SiNb_{12}O_{42}]^{20-}$); colour circles indicate the experimental solubility behaviour. **(C)** Raw □-profiles for nine topologically and compositionally distinct anions ($[P_2W_{18}O_{62}]^{6-}$, $[Ti_2Nb_8O_{28}]^{8-}$, $[CO_3]^{2-}$, $[PNb_{12}O_{42}]^{19-}$, $[Ta_6O_{19}]^{8-}$, $[MoO_4]^{2-}$, Mo_6Cl_{12}, $[H_9Nb_{24}O_{72}]^{15-}$, $[Ti_6Nb_{14}O_{54}]^{14-}$ and $[(UO_2)_{28}(O_2)_{30}]^{28-}$). **(D)** Zoomed in plot of the normalized □-profiles for the same molecule species than in **C**. Note that the □-profiles and the SCDs show opposite sign since the definition of the molecular surface charges distribution.

Next, we sought a molecular descriptor as the basis for a predictive computational model. It has been extensively reported that the charge/size ratio is paramount for determining the stability of metal oxides at distinct pH.[19] For example, higher-charge anions are generally stable at high pH, as is the case of the niobium Lindqvist, $[Nb_6O_{19}]^{8-}$ (Nb_6) which dominates at pH greater than 11.[20] An easy-to-compute descriptor for such charge/size ratio is the surface charge density (SCD), which we defined as the average value of the total molecular charge per surface area (see Supporting Information). **Figure 2b** shows the SCD for ten oxoanions with reported solubility data based on prior[21] and current work (discussed later). SCD correctly assesses that Group V POMs have higher surface charge densities than Group VI POMs. Moreover, it seems that the surface charge density is qualitatively related to solubility. The highest surface charge densities correspond to the compounds with an anomalous solubility trend (depicted with green circles). Intermediate values are consistent with an amphoteric behaviour (red circles) and the lowest surface charge densities match with a normal solubility (blue circles). Although **figure 2b** is satisfying and it seems to represent well the solubility trends for that small set, the borders between solubility groups are rather ambiguous. For example, the SCD values of $[TiNb_9O_{28}]^{7-}$ and $[Nb_2W_4O_{19}]^{4-}$ are very close, yet as we will show later, they correspond to different groups. The single SCD value does not account for obvious differences in the size and charge distribution of those two anions. Thus, such a 0-dimensional descriptor is not suitable for our purposes.

To improve on the latter results, we relied on a more elaborate charge/size ratio descriptor: the σ-profile, introduced by Klamt.[22] This 1-dimensional descriptor is defined as the histogram that quantifies the distribution of molecular surface segments that have a specific charge density named sigma, σ (see Supporting Information). **Figure 2c** shows the raw σ-profiles of ten compounds, zoomed in **figure 2d**. Although we defined a wider charge density domain, most of the peaks are centred in a narrow region, between -0.01 and

+0.03 e/Å^2. Moreover, peak intensities vary significantly depending on the anion size. For instance, the uranyl capsule, $[(UO_2)_{20}(O_2)_{30}]^{20-}$, shows the highest intensity whereas other compounds, such as $[CO_3]^{2-}$, present intensities ten times smaller. Sigma-profiles effectively capture those rich details on the size and charge distribution of molecular systems. However, to utilize them in a computational model, that extreme variance requires adjustment. To do so, we scaled the σ-profiles to obtain a unit norm (see figure S8).

Experimentally testing sigma profiles. Because Nb_{24} is not systematically crystallized without adding non-alkali cations (i.e. copper amines,[12a, 12c] , uranyl[23]), we targeted (1) experimental evaluation of a POM with predicted amphoteric solubility trend, and (2) crystallization of POM-alkali series to assess alkali-POM interactions. In particular, we targeted investigating the relationship between solution-phase interactions inferred by SAXS, and solid-state interactions, ultimately seeking a deeper understanding of amphoteric solubility behavior. Additionally, these experiments served to confirm both the generality of the phenomenon of amphoteric solubility, and to test the predictive nature of sigma profiles on solubility behavior. TMA salts of decametalates $[TiNb_9O_{28}]^{7-}$ and $[Ti_2Nb_8O_{28}]^{8-}$ ($TiNb_9$ and Ti_2Nb_8, respectively) were prepared using prior-described protocol,[13, 24] but adapted to exploit microwave synthesis for bulk and compositional purity.[25] Respectively seven and eight equivalents of alkali chloride were added to each to determine maximum solubility of ion-pairs (see SI for details). For consistency of these studies, we chose the process of adding the alkalis to the TMA-anion solutions, as was done with the Nb_{24} studies (**Table S2**). Indeed, both the Ti_2Nb_8 and $TiNb_9$ series demonstrated amphoteric solubility, as predicted by the σ-profiles. Solubility data for the titanoniobates, Nb_{24}, and representative normal and anomalous solubility POMs are depicted in **figure 3**. We show the data as both POM concentration and metal concentration, illustrating the very high solubility of the Nb_{24}, which we are currently exploiting in deposition of niobate thin films. Interestingly, all three of the amphoteric-solubility clusters have minimal solubility with Na^+ and not K^+, but we suspect other POM clusters not yet studied will exhibit minimal solubility with K^+.

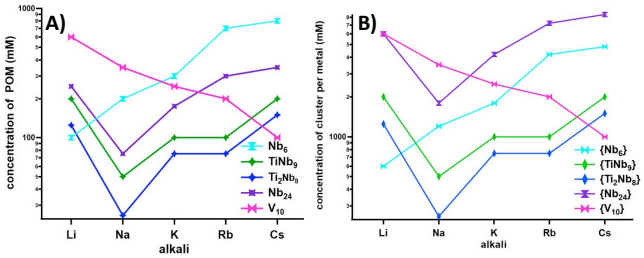


Figure 3. Log-concentration plot of experimentally-determined solubility of POMs, demonstrating all three solubility trends. A) represented as concentration of POM, B) represented as concentration of POM-forming metals. (see also **Table S2**).

Following the solubility studies, all A^+-$TiNb_9$ salts crystallized with seven charge-balancing alkalis per cluster except Li^+-$TiNb_9$, whose lattice contains six Li^+ and one TMA+ per cluster (**figure S3,** crystallographic information summarized in **Table S5**). We also obtained crystals from the A^+-Ti_2Nb_8 series, which were similar in alkali-POM interactions. For brevity we only present the A^+-$TiNb_9$ structures, briefly summarizing the lattice alkali-POM interactions. Li-$TiNb_9$ and Na-$TiNb_9$ both feature A^+ ligated and linked to each other via water molecules with no direct bonding to the POMs except via H-bonding between water molecules to cluster oxos (H_2O----O-POM distances ~2.7 Å). $Li(H_2O)_4$-tetrahedra are corner-linked into linear trimers. Six of seven octahedrally-coordinated $Na(H_2O)_6$ are edge-linked in infinite chains, and the seventh is isolated at the corners and center of the unit cell. For K-$TiNb_9$, there is a three-dimensional K-H_2O network of $K(H_2O)_{7\text{-}8}$ polyhedra that are edge- and face-sharing with each other (**figure S4**). This network contains tunnels and voids in which the $TiNb_9$-clusters reside, where the most basic terminal oxo-sites[24] bond directly to K^+. Rb-$TiNb_9$ can be described as a one-dimensional Rb-H_2O network, with the $TiNb_9$-oxos (μ_3-oxos and terminal, η-oxo) bonded to the Rb^+-cations. Finally, the Cs^+-cations bond to all the oxos of $TiNb_9$ (μ_3, μ_2, and η) and bridge the clusters. In both Rb-$TiNb_9$ and Cs-$TiNb_9$, there is alkali disorder and partial occupancies of the alkalis. The trends in alkali-water-POM bonding with increasing alkali-size can be summarized: (1) increased direct-bonding to POMs with increased alkali size, and (2) decreased selectivity of POM-oxo-alkali bonding (based on basicity of individual oxos[24]). The trends are similar to those observed prior for Nb_{10}-frameworks,[26] but without the strong directionality exhibited by Nb_{10}.

These lattice arrangements are representative of properties of the alkalis with respect to ion-pair formation[17] and hydration spheres, but *how does it relate to amphoteric solubility behavior*? Considering the solubility trend from K-$TiNb_9$ ~ Rb-$TiNb_9$ < Cs-$TiNb_9$; increase in A^+-O_{POM} μ_3-bonding (A=Cs) gives rise to POM-neutralization in solution, decreasing POM-POM interactions that leads to precipitation. Comparing Li-$TiNb_9$ and Na-$TiNb_9$; the latter has infinite alkali-water chains, which may be related to its poorer solubility, whereas Li-water chains are only trimers. The infinite Na-water chains were also observed prior for Na-Ti_2Nb_8.[27] An additional comparison to make is the structuring of Li-water to that observed for a POM with distinct anomalous solubility, $[Nb_6O_{19}]^{8-}$. In previously described $Li_8[Nb_6O_{19}]\cdot nH_2O$ structures,[28] Li-water formed adamantane-like clusters that were also directly bonded to the POM. The distinct anomalous solubility of the highly-charged Nb_6-POM is attributed to its ability to penetrate the Li hydration sphere, leading to direct association. The high charge-density Nb_6 and high-charge density Li^+ bridge extensively, leading to the poorest solubility. By comparison, Li^+ and $TiNb_9$ (lower charge-density) combined are not strong network formers, leading to high solubility. The delicate balance between electrostatic forces and particle sizes as predicted theoretically [11] is observed here again.

SAXS of A^+-$TiNb_9$ and A^+-Ti_2Nb_8 plus one additional structure yielded a surprising result, comparable to the tendency of Nb_{10} to convert to Nb_{24} upon simple addition of alkalis. Single-crystal X-ray analysis of a crystalline product of the Ti_2Nb_8 plus KCl solution revealed some degree of disassembly and dimerization of Ti_2Nb_8, yielding $[Ti_6Nb_{14}O_{54}]^{14-}$ (Ti_3Nb_7-dimer **figure 4A**). The Ti_3Nb_7-dimer,

fully formulated $K_{14}[Ti_6Nb_{14}O_{54}]\cdot 18H_2O$, crystallized in the orthorhombic *Immm* space group with a unit cell of 4432.3(2) $Å^3$ (**Table S6**). The composition suggests it is not a simple dimerization. In addition to the two Ti-sites of the cluster (blue in **figure 4A**), the linking polyhedra (yellow) are disordered, each hosting ½ Ti and ½ Nb (see SI for details). This is a clear indication that there is some disassembly of the decametalates in solution, perhaps limited to only these sites that are capped with the most basic and labile η-oxos.[20, 24] The topology of the Ti_3Nb_7-dimer is very similar to prior-reported Nb_{10}-dimer[29] obtained from organic solution, and it is only the second such dimerized decametalate POM. The major difference between the Ti_3Nb_7-dimer and the Nb_{10}-dimer is the M-O-M dihedral angle between the two M_{10}-units: it is bent in the Nb_{10}-dimer (158.7°), and straight in the Ti_3Nb_7-dimer (179.6°, **figure 4a and S6**). The Ti/Nb-O-Ti/Nb bond lengths that bridge the two decamers for the Ti_3Nb_7-dimer are 1.825(2) Å, shorter than those of the Nb_{10}-dimer (189-1.91 Å), reflecting partial substitution of Ti for Nb.

In the scattering data of Ti_2Nb_8 with added alkalis (**figure 4B**), there is a shift in the Guinier region to lower-q that correlates perfectly with increasing alkali size, indicating growth of a larger scattering specie or increasing ion-pairing. Also shown in **figure 4B** is simulating scattering of Ti_2Nb_8 and the Ti_3Nb_7-dimer. Comparison of the Guinier region (the elbow between 0.1-0.4 $Å^{-1}$) with experimental scattering suggests increased formation of this dimer with increasing alkali size could account for this change. Indeed, a two-phase fit (**figure S5**) modeling the size of the Ti_2Nb_8 and Ti_3Nb_7-dimer gave robust data fits that preserved the expected size of the clusters, and suggest increased dimerization, promoted by interaction with the larger alkalis (**Table S7**). Qualitatively, the scattering curves for $TiNb_9$ reflect a similar trend, but attempts to fit this scattering data using the same approach failed. This could have many explanations ranging from formation of other larger species that have not yet been identified by crystallization, or formation of polydisperse aggregates promoted by ion-pairing. With limited corollary information, we did not present further quantitative analysis of these data.

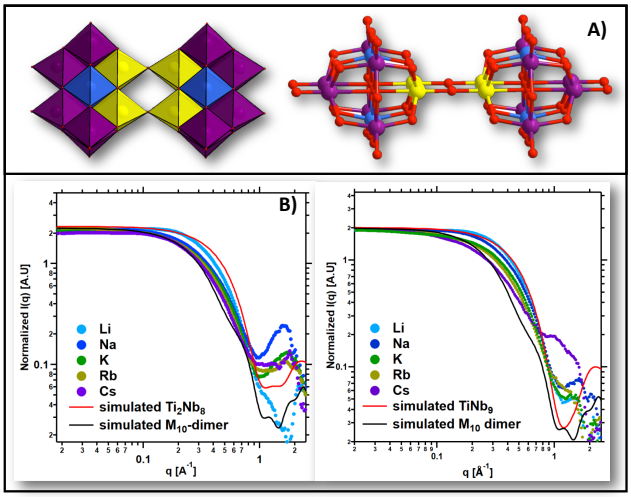


Figure 4. Solution behavior of $TiNb_9$ and Ti_2Nb_8. A) polyhedral representation (left) of Ti_3Nb_7-dimer, $[Ti_6Nb_{14}O_{54}]^{14-}$, and ball-and-stick representation (right) that highlights the ~180° Ti/Nb-O-Ti/Nb dihedral angle between the two Ti_3Nb_7-units. Blue is fully-occupied Ti, yellow is M-sites that are half Nb and half Ti, purple is Nb and red is O. **B)** SAXS of TMA-Ti_2Nb_8 with added alkalis. **C)** SAXS of TMA-$TiNb_9$ with added alkalis. For both SAXS spectra, simulated M_{10} and M_{10}-dimer scattering is also shown for comparison to the experimental scattering.

The predictive computational model we present herein is based on a Self-Organized Map (SOM), also known as Kohonen Neural Network (KNN)**. Figure 5A** shows a schematic representation of the workflow followed in this part. First, we have split the raw σ-profile data in two independent groups: training (34) and test (10) sets (see Supporting Information). We have defined our neural network as a two-dimensional grid of *N* x *N* neurons where *N* is equal to the size of the training set. Next, we initialize the KNN with the first two principal components and introduce the normalized σ-profiles of the training set to the neural network. The training of the KNN is carried out using unsupervised learning, which implies that we do not specify the solubility behaviour of the compounds. We chose this approach due to the lack of solubility data values available for the anomalous and amphoteric groups. The progress of the training is plotted in **figure 5B**, showing the quantization error rapidly converges to zero. In fact, the training phase was concluded after 3000 iterations, when the error was below $9\cdot10^{-5}$. Throughout the training, the neuron layer undergoes a self-organization process to arrange data into clusters. **Figure 5A** shows a hypothetical example where the neuron layer has formed four well-distinguished groups. The next step is to provide classification ability to the KNN, by adding a label to each cluster. To assign a property to a cluster, all the compounds within that cluster must present the same behaviour. In case there were discrepancies within a cluster, it would indicate that either the KNN or the descriptor are not adequate (details in SI). Finally, we would pass the test set to the trained KNN and it would automatically predict the solubility behaviour. Note that once the neural network has been trained, the neurons and thus the shape of the KNN remains constant.

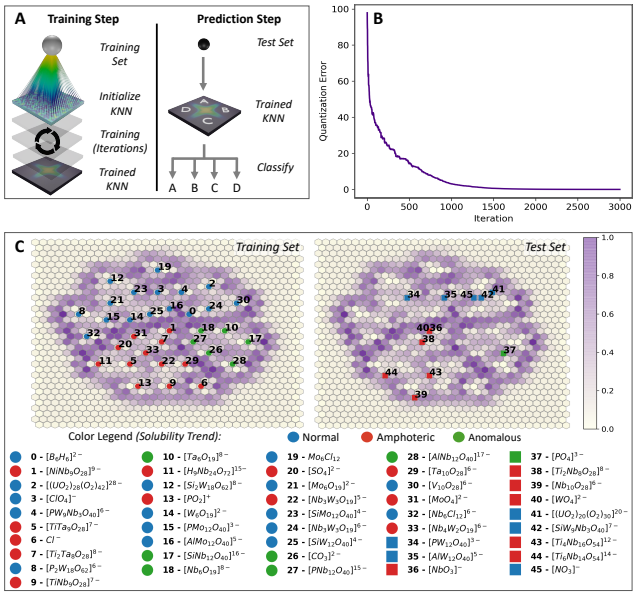


Figure 5. A) Schematic protocol of the KNN divided in two steps: training (Figure 5C left) and prediction (Figure 5C right). **B)** Number of iterations for the training phase as function of the quantization error. **C)** Neighbourhood weight distance plots on 34 x 34 hexagonal sheets for the training (left) and test (right) sets, marked with circles and squares respectively. Experimental solubility behaviour of simple anions are summarized in **Tables S1- S3**.

Figures 5C collects two *neighborhood weight* distance plots, corresponding to the training process and test sets respectively. Neighborhood plots compute the distance between input samples using the function learnt by the neural network. Within the SOM framework, distance is inversely related to similarity. Thus, the closer a group of data is (marked with light colors), the more similar their σ-profile shapes will be. Regarding the neighborhood plot for the training set, we can observe that the 34 σ-profiles (marked with circles) have been automatically clustered in three groups, marked with blue, red and green colours.

The biggest cluster, depicted in blue, is assigned to compounds that display a normal solubility trend. This cluster includes s $[(UO_2)_{20}(O_2)_{42}]^{20-}$, $[ClO_4]^-$, $[P_2W_{18}O_{62}]^{5-}$, $[W_6O_{19}]^{2-}$, B_6H_6 and $[PMo_{12}O_{40}]^{3-}$. Note that the rest of the compounds in this group are slight modifications of the six previous molecules. Furthermore, molybdenum and tungsten metals clearly dominate the oxoanions of this group. The second cluster, marked in red, corresponds to oxoanions with an amphoteric behaviour. In this group we find $[H_9Nb_{24}O_{72}]^{15-}$ and $[TiNb_9O_{28}]^{7-}$, both being experimentally assigned as amphoteric in this work. All the decametalate structures, $[M_{10}O_{28}]$ with the exception of V_{10} (experimentally determined to be normal solubility, **fig. 3**) are found in this group. Notably, we have not been able to experimentally-determine the solubility trend for alkai-Nb_{10} in water, given its propensity to convert entirely to Nb_{24}. Because Ti_2Nb_8 also shows evidence of alteration upon interaction with alkalis, we included the Ti_2Nb_8-dimer in the final test set, and it too is predicted to have amphoteric solubility, based on this model. The third and last cluster, marked in green, is attributed to compounds with an anomalous solubility. It includes the Lindqvist $[Nb_6O_{19}]^{8-}$, which is widely known to be more soluble in presence of Cs than Li or Na cations. In this group, $[CO_3]^{2-}$, $[Ta_6O_{19}]^{8-}$ and $[SiNb_{12}O_{40}]^{16-}$ all (qualitatively) display anomalous solubility. The trained KNN not only clusters and classifies compounds of the training set, but it also can be employed to predict the solubility of any oxoanion from its σ-profile. That is the case for the neighborhood plot of the test set shown in **figure 5C** in which twelve σ-profiles (marked with squares) were clustered and classified according to their solubility trend. Note that the shape of the 34x34 neuron layer remains unmodified once the training phase has been completed that includes both simple oxoanions and larger POMs and other clusters.

Conclusion

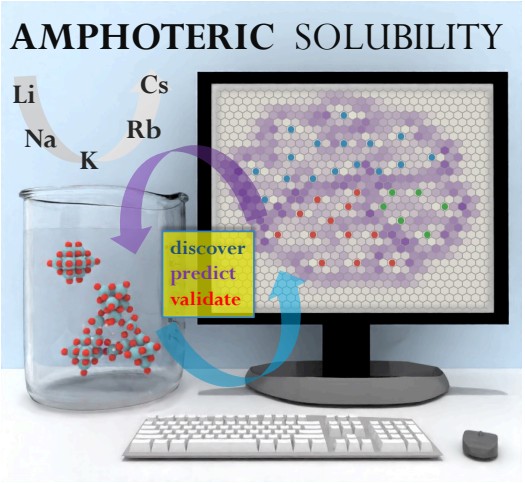
Dissolution, crystallization and precipitation of inorganic ion-pairs is extremely important in natural processes and synthesis, yet few predictive models for solubility exist that go beyond the inverse relationship of lattice energies to solubility, a relationship that is not always true. Here we have identified a third general solubility trend of alkali-POM pairs that is intermediate between normal solubility (trending Li-anion solubility highest, Cs-anion solubility lowest) and anomalous solubility (trending Cs-anion solubility highest, Li-anion solubility lowest). Anions with amphoteric solubility are least soluble with intermediate alkalis (Na/K) and was experimentally first described for Nb_{24}-POMs. We have developed the normalized σ-profile as a computational descriptor to predict all three solubility trends for anions ranging in size and complexity from simple halides to oxoanions that are several nanometers in diameter and contain >100 atoms. Accuracy of predictive σ-profiles was validated by experimentally demonstrating amphoteric solubility for two additional POMs, $TiNb_9$ and Ti_2Nb_8. Both solution and solid-state studies of alkali-Ti_2Nb_8 ion interactions in water showed its strong tendency to dimerize, and yielded a new POM $[Ti_6Nb_{14}O_{54}]^{14-}$ that is also computationally predicted to have anomalous solubility. With this result, we now have two examples of POMs with predicted/experimentally determined amphoteric solubility trend with alkalis (Nb_{10} and Ti_2Nb_8) that undergo solution-phase transformation to a different POM that also exhibits amphoteric solubility (Nb_{24} and Ti_3Nb_7-dimer); simply promoted by interaction with alkali-countercations. This result affirms that in aqueous metal oxide chemistry, while pH is a well-known and accepted driver of solubility and speciation, the often-neglected supporting counterions and electrolytes are active and important participants in many reactions.

Acknowledgements
This study was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Material Sciences and Engineering, under award DE SC0010802 (OSU and ICIQ). ICIQ acknowledges the Spanish Ministerio de Ciencia e Innovación (PID2020-112806RB-I00) and the Severo Ochoa Excellence Accreditation 2020 –2023 (CEX2019-000925-S, MCI/AEI), the ICIQ Foundation and the CERCA program of the Generalitat de Catalunya. We thank CBMM, Araxá, Minas Gerais, Brazil for the generous gift of hydrous niobia. We acknowledge the Murdock Charitable Trust (grant SR-2017297) for acquisition of the single-crystal X-ray diffractometer. CAO thanks the Swedish Research Council (grant 2018-07039) and the Kempe foundation (grant JCK-2029.1) for funding.

Keywords: polyoxometalate • ion-pairing • SAXS • solubility • machine learning

[1] S. A. Woodson, *Curr Opin Chem Biol* **2005**, *9*, 104-109.
[2] a) H. Maity, A. N. Muttathukattil, G. Reddy, *J Phys Chem Lett* **2018**, *9*, 5063-5070; b) H. I. Okur, J. Hladilkova, K. B. Rembert, Y. Cho, J. Heyda, J. Dzubiella, P. S. Cremer, P. Jungwirth, *J Phys Chem B* **2017**, *121*, 1997-2014.
[3] a) G. A. Waychunas, C. S. Kim, J. F. Banfield, *Journal of Nanoparticle Research* **2005**, *7*, 409-433; b) P. M. Fox, J. A. Davis, J. M. Zachara, *Geochimica et Cosmochimica Acta* **2006**, *70*, 1379-1387.
[4] a) D. Gebauer, A. Völkel, H. Cölfen, *Science* **2008**, *322*, 1819-1822; b) J. J. De Yoreo, P. U. Gilbert, N. A. Sommerdijk, R. L. Penn, S. Whitelam, D. Joester, H. Zhang, J. D. Rimer, A. Navrotsky, J. F. Banfield, *Science* **2015**, *349*, aaa6760; c) J. D. Rimer, D. D. Roth, D. G. Vlachos, R. F. Lobo, *Langmuir* **2007**, *23*, 2784-2791.
[5] a) N. F. van der Vegt, K. Haldrup, S. Roke, J. Zheng, M. Lund, H. J. Bakker, *Chemical Reviews* **2016,** *116*, 7626–7641; b) A. V. Verde, M. Santer, R. Lipowsky, *Physical Chemistry Chemical Physics* **2016**, *18*, 1918-1930.
[6] a) A. Salis, B. W. Ninham, *Chem Soc Rev* **2014**, *43*, 7358-7377; b) P. Jungwirth, P. S. Cremer, *Nat Chem* **2014**, *6*, 261-263.
[7] a) Y. Chu, J. H. Chen, F. Haso, Y. Y. Gao, J. E. S. Szymanowski, P. C. Burns, T. B. Liu, *Chem-Eur J* **2018**, *24*, 5479-5483; b) J. H. Chen, K. Qian, K. X. Xiao, J. C. Luo, H. Li, T. Ma, U. Kortz, M. Tsige, T. B. Liu, *Langmuir* **2020**, *36*, 10519-10527.
[8] M. Andreev, J. J. de Pablo, A. Chremos, J. F. Douglas, *J Phys Chem B* **2018**, *122*, 4029-4034.
[9] K. I. Assaf, W. M. Nau, *Angew Chem Int Ed Engl* **2018**, *57*, 13968-13981.
[10] a) T. Buchecker, P. Schmid, I. Grillo, S. Prevost, M. Drechsler, O. Diat, A. Pfitzner, P. Bauduin, *J Am Chem Soc* **2019**, *141*, 6890-6899; b) C. Drummond, L. Pérez-Fuentes, D. Bastos-González, *J. Phys. Chem. C* **2019**, *123*, 28744-28752; c) A. Sole-Daura, J. M. Poblet, J. J. Carbo, *Chem-Eur J* **2020**, *26*, 5799-5809.
[11] a) M. Nyman, *Dalton Trans* **2011**, *40*, 8049-8058; b) Y. Hou, M. Nyman, M. A. Rodriguez, *Angew Chem Int Edit* **2011**, *50*, 12514-12517; c) M. Nyman, P. C. Burns, *Chem Soc Rev* **2012**, *41*, 7354-7367.
[12] a) R. P. Bontchev, M. Nyman, *Angew Chem Int Edit* **2006**, *45*, 6670-6672; b) D. Sures, M. Segado, C. Bo, M. Nyman, *J Am Chem Soc* **2018**, *140*, 10803-10813; c) J. Y. Niu, P. T. Ma, H. Y. Niu, J. Li, J. W. Zhao, Y. Song, J. P. Wang, *Chem-Eur J* **2007**, *13*, 8739-8748.
[13] C. A. Ohlin, E. M. Villa, J. C. Fettinger, W. H. Casey, *Dalton Trans* **2009**, 2677-2678.
[14] M. Amiri, N. P. Martin, C. L. Feng, J. K. Lovio, M. Nyman, *Angew Chem Int Edit* **2021**, *60*, 12461-12466.
[15] a) O. Engkvist, P. Wrede, *J Chem Inf Comp Sci* **2002**, *42*, 1247-1249; b) E. Mullins, Y. A. Liu, A. Ghaderi, S. D. Fast, *Ind Eng Chem Res* **2008**, *47*, 1707-1725; c) M. C. Sorkun, J. M. V. A. Koelman, S. Er, *iScience* **2021**, *24*, 101961.
[16] S. Boobier, D. R. J. Hose, A. J. Blacker, B. N. Nguyen, *Nat Commun* **2020**, *11,* 5753.
[17] A. Misra, K. Kozma, C. Streb, M. Nyman, *Angew Chem Int Edit* **2020**, *59*, 596-612.
[18] N. G. Naumov, K. A. Brylev, Y. V. Mironov, S. Cordier, V. E. Fedorov, *J Struct Chemistry* **2014**, *55*, 1371-1389.
[19] X. Lopez, J. A. Fernandez, J. M. Poblet, *Dalton Trans* **2006**, *9*, 1162-1167.
[20] E. M. Villa, C. A. Ohlin, E. Balogh, T. M. Anderson, M. D. Nyman, W. H. Casey, *Angew Chem Int Edit* **2008**, *47*, 4844-4846.
[21] D. J. Sures, P. I. Molina, P. Miro, L. N. Zakharov, M. Nyman, *New J Chem* **2016**, *40*, 928-936.
[22] A. Klamt, *J Phys Chem* **1995**, *99*, 2224-2235.
[23] N. P. Martin, E. Petrus, M. Segado, A. Arteaga, L. N. Zakharov, C. Bo, M. Nyman, *Chem-Eur J* **2019**, *25*, 10580-10584.
[24] E. M. Villa, C. A. Ohlin, W. H. Casey, *Journal of the American Chemical Society* **2010**, *132*, 5264-5272.
[25] a) M. A. Rambaran, M. Pascual-Borrs, C. A. Ohlin, *European Journal of Inorganic Chemistry* **2019**, *2019*, 3913-3918; b) M. A. Rambaran, A. Gorzsas, M. Holmboe, C. A. Ohlin, *Dalton Trans* **2021**, *50*, 16030-16038.
[26] N. P. Martin, M. Nyman, *Angew Chem Int Edit* **2021**, *60*, 954-960.
[27] M. Nyman, L. J. Criscenti, F. Bonhomme, M. A. Rodriguez, R. T. Cygan, *J Solid State Chem* **2003**, *176*, 111-119.
[28] a) Y. Hou, D. B. Fast, R. E. Ruther, J. M. Amador, L. B. Fullmer, S. R. Decker, L. N. Zakharov, M. R. Dolgos, M. Nyman, *J Solid State Chem* **2015**, *221*, 418-425; b) T. M. Anderson, S. G. Thoma, F. Bonhomme, M. A. Rodriguez, H. Park, J. B. Parise, T. M. Alam, J. P. Larentzos, M. Nyman, *Cryst Growth Des* **2007**, *7*, 719-723.
[29] M. Maekawa, Y. Ozawa, A. Yagasaki, *Inorganic Chemistry* **2006**, *45*, 9608-9609.

Entry for the Table of Contents



Discover. Predict. Validate. Using alkali-polyoxometalate salts, we discover a previously-undescribed solubility trend, termed anomalous solubility. Mapping charge-density on inorganic polyatomic anions along with an artificial neural network accurately describes and predicts solubility trends of alkali-anion pairs. Solubility experiments to validate computation leads to new POMs discovery.

@NymanMay @Carles_Bo @OregonState @ICIQchem