Check for updates



Catalysis Hot Paper

 How to cite:
 Angew. Chem. Int. Ed. 2021, 60, 5132–5135

 International Edition:
 doi.org/10.1002/anie.202013256

 German Edition:
 doi.org/10.1002/ange.202013256

To Every Rule There is an Exception: A Rational Extension of Loewenstein's Rule

Magnus Fant, Mattias Ångqvist, Anders Hellman, and Paul Erhart*

Abstract: Loewenstein's rule, which states that Al–O–Al motifs are energetically unstable, is fundamental to the understanding and design of zeolites. Here, using a combination of electronic structure calculations and lattice models, we show under which circumstances this rule becomes invalid and how it can be rationally extended using the chabasite framework for demonstration.

Zeolites are aluminosilicate minerals, used in many different industrial applications, including detergents, adsorbents/desiccants, and catalysts.^[1-4] They can occur in a staggering number of frameworks with distinct pore architectures and sizes, which are obtained by different arrangements of the underlying tetrahedral SiO₄ building blocks.^[5] For functionalization, Si⁺⁴ in the structure are substituted with Al⁺³, where the Al/Si ratio varies from zeolite to zeolite. The substitution introduces a net negative charge into the framework that needs to be balanced by counterions such as H⁺, Na⁺ or Cu⁺. Depending on the character of these counterions, different chemistry is introduced, let it be Brønsted acid chemistry (H⁺) or redox chemistry (Cu⁺/Cu⁺²).

The possibility for zeolites to have different pore sizes makes them ideal for separating various chemicals,^[6,7] e.g., separation of CO₂ in natural gas and hydrogen purification, and to enforce shape selectivity in catalytic transformations,^[8] e.g., differentiating between linear and branched hydro-carbons. Furthermore, the acidity of the Brønsted sites, whose strength can be tuned by, e.g., isomorphic substitution, plays a crucial role in many hydrocarbon reactions.^[9,10] Simultaneously, the redox abilities of metal cations in ion-exchanged zeolites are essential in many oxidation-reduction reactions, e.g., the selective catalytic reduction of NO_x.^[11, 12]

The catalytic performance of the zeolite is, to a large extent, controlled by the distribution of Al^{+3} sites, hence

[*]	M. Fant, Dr. M. Ångqvist, Prof. P. Erhart
	Department of Physics, Chalmers University of Technology
	Gothenburg (Sweden)
	E-mail: erhart@chalmers.se
	Prof. A. Hellman
	Department of Physics and Competence Centre for Catalysis,
	Chalmers University of Technology
	Gothenburg (Sweden)
	Supporting information and the ORCID identification number(s)
	the author(s) of this article can be found under:

https://doi.org/10.1002/anie.202013256.

understanding and controlling this distribution is a crucial part of developing more predictive synthesis-structure-function relationships.^[13–16] The Al distribution is often rationalized using Loewenstein's rule,^[17,18] which states that Al–O–Al motifs are unstable. This rule is so widely applied and so firmly established that violations warrant special status.^[19–21]

This situation motivates the present study, in which we undertake a critical examination of Loewenstein's rule using the prototypical SSZ-13 chabasite structure as a model system. While we show that it works as expected when applied in its original context, more importantly, we identify the conditions under which it falls short.

The number of distinct chemical configurations increases exponentially as a zeolite structure is loaded with Al⁺³ and counterions. This combinatorial explosion does not only exhaust any enumeration approach^[22] but highlights the importance of configurational entropy. To account for this aspect computationally, we constructed so-called alloy cluster expansions (CEs),^[23,24] which provide computationally efficient yet very accurate lattice models for the energy of materials as a function of the chemical distribution. Such models have already been successfully applied to other group 13/14-group based cage structures.^[25]

We studied charge compensation by H⁺, Na⁺, K⁺, and Rb⁺ as well as free charge carriers. While the latter compensation mechanism is not available in reality due to the large band gap, it serves as an extreme limit that provides useful insight, as shown below. Initially we considered seven different Wyckoff sites for each counterion. Based on density functional theory (DFT) calculations^[26-29] for each site in the dilute limit, we reduced the set of possible sites to four in the case of H⁺ (each associated with one oxygen site) and three in the case of Na^+ , K^+ , and Rb^+ (located along the channels and in the pore of the chabasite structure; see Supplementary Information for details^[25, 30-34]). Our CE models were then trained to the energies from DFT calculations for approximately 100 to 200 configurations, in which both cell metrics and ionic positions were allowed to relax. The final CEs reproduce these reference data within an average root mean square error over the validation set between 1.7 meV/atom (Rb) and 6.6 meV/atom (free carriers). These models were subsequently sampled by Monte Carlo (MC) simulations in the variance constrained semi-grand canonical (VCSGC) ensemble^[33] to obtain the fraction of Al-Al nearest neighbours (NNs) (i.e. Al-O-Al motifs) as a function of Al content. We intentionally sampled a very wide composition range up to approximately 55% in order to illustrate the fundamental factors that drive the composition dependence.

The results are strongly dependent on the type of counterion (Figure 1). When charge compensation is ach-

for

^{© 2021} The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

GDCh



Figure 1. Variation of the fraction of Al–Al nearest-neighbor pairs as a function of Al content in a chabasite framework (SSZ-13) using either ions (H⁺, Na⁺, K⁺, Rb⁺) or free carriers for charge compensation.

ieved by Na⁺, K⁺ or Rb⁺, the fraction of Al–Al NNs is almost zero up to a concentration of 33%. At high temperatures, a finite, but still small, number of Al–Al NNs is present since they create disorder, which yields an entropic contribution.

While the behavior observed for the alkali ions is largely compatible with Loewenstein's rule, hydrogen presents a very different case (Figure 1). The fraction of Al-Al NNs rises sharply at small Al content to about 40%, after which it continues to increases gradually. This behavior can be traced to the effectively *attractive* interaction between Al³⁺ if charge compensated by H⁺, which was already noted in Ref. [19] and which will be rationalized below in terms of a competition between electrostatics and strain. Notably this attraction leads to the formation of Al "clusters" (Figure 3). The temperature dependence of the number of Al-Al NNs is more pronounced than in the case of alkali counterions, which is consistent with a larger entropic contribution thanks to a larger number of available sites (4 sites with multiplicity 18 vs. 3 sites with multiplicities between 3 and 9) and smaller energy differences between these sites.

Finally, in the case of compensation by free carriers, the fraction of Al–Al NNs is *exactly* zero regardless of temperature up to 50% Al, at which point it is geometrically impossible to avoid Al–Al NNs.

One thus observes three different types of behavior depending on whether compensation is achieved by H^+ , alkali ions or free carriers. Strictly speaking, Loewenstein's rule is thus only obeyed in the case of free carrier compensation. This begs the question of what factors are at play in the other two cases. To resolve this question, it is instructive to consider the electronic structure of the different counterions in the dilute limit (Figure 2).

First, we consider the electronic structure of *isolated* impurity ions in SSZ-13. The exchange of a single Si⁺⁴ with



Figure 2. Electronic energy levels associated with different defects in SSZ-13 (top) and their real space representations shown as isosurfaces (bottom). Occupied and empty levels are shown in blue and orange, respectively. Visualized using ovito.^[35]

Al⁺³ (an acceptor in semi-conductor terminology) leads to the emergence of several levels near the valence band maximum (VBM). These states are semi-local in character and arise from the hybridization of O-2p orbitals (similar to the orbital shown in Figure 2 (ii)). In the case of H⁺ and Na⁺ (donors in semi-conductor terminology), on the other hand, one observes the appearance of deep localized levels near the conduction band minimum (CBM). (K⁺ and Rb⁺ exhibit characteristics similar to Na⁺ and are not discussed here in detail.)

In the lowest energy configuration, H^+ is situated next to an oxygen site; the lowest unoccupied level is localized on the H^+ site and exhibits p-character with an orientation along the H^-O bond axis, corresponding to the lowest unoccupied molecular orbital of OH^- (similar to Figure 2 (i)). Na⁺, on the other hand, prefers 6c sites, which reside along the axes of the channels of the SSZ-13 structure. The lowest unoccupied levels exhibit semi-local character as the corresponding charge density is distributed over several sites surrounding the Na⁺ site (similar to Figure 2 (ii)).

Next, we consider the electronic structure of the compensated systems, in which, generally speaking, the electron from the cation is transferred to unsaturated O–Al bonds, formally turning all framework oxygens into O^{-2} . In the case of H^+ – Al^{+3} , the attraction between H^+ and the saturated O^{-2} renders the O-sites close to Al^{+3} energetically preferred. This configuration leads to two levels in the band gap (see H^+ – Al^{+3} in Figure 2): an empty level near the CBM, which exhibits p-like character as in the case of the individual H^+ (Figure 2 (i)), and an occupied level near the VBM, which is comprised of p-orbitals localized at the four oxygen neighbors of the Al^{+3} site (Figure 2 (ii)).

One observes several localized levels in the band gap also in the case of Na⁺–Al⁺³ (Figure 2). While the levels near the VBM exhibit similar characteristics as in the case of H^+ –O⁻² (Figure 2(iii)), the levels in the vicinity of the CBM are much less localized as the charge density is not only located at the Na⁺ site but distributed over sites up to three neighbor shells away (Figure 2(iii)). This behavior can be attributed to the much larger size of Na⁺, which forces it to occupy sites along the channel, an effect that is even more pronounced for K⁺ and Rb⁺.

Further insight is provided by analyzing the charge density redistribution upon compensation. To this end, we consider the difference between the charge densities of the compensated system, the system with only Al^{+3} ions, and the free counterions in their atomic state. This analysis shows that H^+ - O^{-2} pairing leads to the formation of small dipoles but crucially—no isolated monopoles (Figure 3 a), which considerably reduces the repulsive interaction between like-charged species; in other words the leading electrostatic interaction term falls of as $1/R^3$ as opposed to 1/R. The insertion of Al gives rise to a notable structural relaxation as the average Al– Si nearest-neighbor distance is 3.21 Å to be compared with an



Figure 3. Low energy configurations in perspective (top row) and viewed along the *c*-axis (bottom row) of (a) H^+-AI^{+3} , (b) Na^+-AI^{+3} , (c) $2 \times H^+-AI^{+3}$, and (d) $4 \times H^+-AI^{+3}$. Yellow and blue isosurfaces show, respectively, negative and positive charge redistribution upon compensation with the different counterions.

average Si–Si distance of 3.12 Å in the ideal SSZ-13 structure. The strain field associated with each Al^{+3} site gives rise to an effective attraction, which in the absence of strong electrostatic repulsion leads to the clustering of Al^{+3} (Figure 3 c,d), a behavior that can also be observed for (effectively) charge-neutral defects in other materials.^[36]

In sharp contrast, in the case of Na⁺, one observes a localization of negative charge on the Na⁺ site with the compensating positive charge distributed over the surrounding O^{2-} sites (Figure 3b). As a result, the strain-mediated attraction is overruled by electrostatic repulsion (falling of as 1/R), preventing Al⁺³ clustering. As the system strives to maximize the separation between equally charged Na⁺ species, the attractive O-Na interaction also forces a separation of Al⁺³ species, since these are indirectly associated with saturated O⁻². The most extreme form of this behavior is obtained if compensation is achieved via free carriers (Figure 1). In this case, one effectively obtains a system of point charges (distributed over O⁻² sites) and a homogeneous background charge (corresponding to free carriers), for which the electrostatic energy is minimized by maximizing the Al-Al spacing.

The above analysis suggests that the underlying factors are largely electrostatic in nature and thus relatively insensitive to the framework structure. In fact, calculations of Al clusters in other frameworks constructed in analogy to the ones found for the chabasite framework yield very similar binding energies and show compensation with free carriers and H⁺ to, respectively, prevent and favor Al clustering (see Supporting Information for details).

To summarize, the thermodynamic distribution of Al⁺³ in the prototypical SSZ-13 zeolite is sensitive to the type of counterion used for charge compensation. While the Al-Al interaction is effectively attractive when compensating with H^+ , it is repulsive in the case of Na⁺, K⁺ or Rb⁺. This difference can be understood by considering counterion size, level of charge localization, and local structural rearrangements. In particular, H⁺ counterions enable full saturation of O^{-2} and very localized charge compensation, which effectively reduces the strong electrostatic repulsion between monopoles. This allows the strain-mediated attraction between Al⁺³ sites to take over, leading to clustering, and hence violation of the Loewenstein rule. By contrast, the larger size of $Na^{\scriptscriptstyle +},\ K^{\scriptscriptstyle +}$ or $Rb^{\scriptscriptstyle +}$ prevents the formation of localized bonds with O^{-2} sites and the compensation charge is much more delocalized. As a result, repulsive electrostatic interactions govern the Al⁺³ distribution.

While H-SSZ-13 is important for applications, H^+ species are usually not present during synthesis. The Al distribution that is commonly encountered in these materials will thus not correspond to an equilibrium state but is instead preserved kinetically due to the very large barriers for Al redistribution. The present insight raises the question of whether synthesis routes and annealing procedures can be devised that exploit the mechanisms described above to control the Al distribution more consciously. Here, lattice geometry and the locations of counterion sites play important roles. While the channels and pores in SSZ-13 are very small, it could be possible to realize local charge compensation more efficiently in the presence of larger channels and pores also with cations larger than $\mathrm{H}^+.$

Finally, we note that the present approach adds another computational method to the toolbox available for understanding and designing zeolite structure and chemistry.^[14,16,37] Future work in this direction should address for example the distribution of divalent and trivalent species,^[1,38,39] molecular counterions as well as other frameworks.

Acknowledgements

Funding from the Knut and Alice Wallenberg Foundation (2014.0226) as well as the Swedish Research Council (2018-06482, 2020-04935) are gratefully acknowledged. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at NSC, C3SE and PDC partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

Conflict of interest

The authors declare no conflict of interest.

Keywords: Al distribution · catalysis · Loewenstein rule · zeolites

- [1] A. M. Beale, F. Gao, I. Lezcano-Gonzalez, C. H. F. Peden, J. Szanyi, Chem. Soc. Rev. 2015, 44, 7371.
- [2] Y. Li, L. Li, J. Yu, Chem 2017, 3, 928.
- [3] B. Yilmaz, U. Müller, in Zeolites in Sustainable Chemistry: Synthesis Characterization and Catalytic Applications, Vol. 473 (Eds.: F.-S. Xiao, X. Meng), Springer, Berlin, Heidelberg.
- [4] Zeolites in Catalysis, Catalysis Series (Eds.: J. Čejka, R. E. Morris, P. Nachtigall), The Royal Society of Chemistry, Cambridge, 2017.
- [5] G. D. Gatta, P. Lotti, in *Modified Clay and Zeolite Nanocompo*site Materials, Micro and Nano Technologies 1-25 (Eds.: M. Mercurio, B. Sarkar, A. Langella), Elsevier, Amsterdam, 2019.
- [6] J. Caro, M. Noack, P. Kölsch, R. Schäfer, Microporous Mesoporous Mater. 2000, 38, 3.
- [7] J. Caro, M. Noack, *Microporous Mesoporous Mater.* 2008, 115, 215.
- [8] U. Olsbye, S. Svelle, M. Bjorgen, P. Beato, T. V. W. Janssens, F. Joensen, S. Bordiga, K. P. Lillerud, *Angew. Chem. Int. Ed.* **2012**, *51*, 5810; *Angew. Chem.* **2012**, *124*, 5910.
- [9] A. Bhan, E. Iglesia, Acc. Chem. Res. 2008, 41, 559.
- [10] A. A. Arvidsson, P. N. Plessow, F. Studt, A. Hellman, J. Phys. Chem. C 2020, 124, 14658.
- [11] C. Paolucci, I. Khurana, A. A. Parekh, S. Li, A. J. Shih, H. Li, J. R. Di Iorio, J. D. Albarracin-Caballero, A. Yezerets, J. T. Miller, W. N. Delgass, F. H. Ribeiro, W. F. Schneider, R. Gounder, *Science* **2017**, *357*, 898.

- [12] L. Chen, H. Falsig, T. V. Janssens, H. Grönbeck, J. Catal. 2018, 358, 179.
- [13] K. Muraoka, W. Chaikittisilp, T. Okubo, J. Am. Chem. Soc. 2016, 138, 6184.
- [14] M. Jeffroy, C. Nieto-Draghi, A. Boutin, *Chem. Mater.* 2017, 29, 513.
- [15] B. C. Knott, C. T. Nimlos, D. J. Robichaud, M. R. Nimlos, S. Kim, R. Gounder, ACS Catal. 2018, 8, 770.
- [16] J. R. Di Iorio, S. Li, C. B. Jones, C. T. Nimlos, Y. Wang, E. Kunkes, V. Vattipalli, S. Prasad, A. Moini, W. F. Schneider, R. Gounder, J. Am. Chem. Soc. 2020, 142, 4807.
- [17] W. Loewenstein, Am. Mineral. 1954, 39, 92.
- [18] R. G. Bell, R. A. Jackson, C. R. A. Catlow, *Zeolites* **1992**, *12*, 870.
- [19] R. E. Fletcher, S. Ling, B. Slater, *Chem. Sci.* **2017**, *8*, 7483.
- [20] V. V. Klepov, C. A. Juillerat, E. V. Alekseev, H.-C. zur Loye, *Inorg. Chem.* 2019, 58, 724.
- [21] C. J. Heard, L. Grajciar, P. Nachtigall, Chem. Sci. 2019, 10, 5705.
- [22] G. L. W. Hart, R. W. Forcade, *Phys. Rev. B* 2009, 80, 014120.
- [23] J. M. Sanchez, F. Ducastelle, D. Gratias, Phys. A 1984, 128, 334.
- [24] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, P. Erhart, Adv. Simul. Theory 2019, 2, 1900015.
- [25] M. Ångqvist, P. Erhart, Chem. Mater. 2017, 29, 7554.
- [26] J. P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 1996, 77, 3865.
- [27] C. Adamo, V. Barone, J. Chem. Phys. 1999, 110, 6158.
- [28] G. Kresse, J. Furthmüller, Comput. Mater. Sci. 1996, 6, 15.
- [29] G. Kresse, J. Furthmüller, Phys. Rev. B 1996, 54, 11169.
- [30] C. Baerlocher, L. McCusker, *Database of zeolite structures*, http://www.iza-structure.org/databases/.
- [31] G. Makov, M. C. Payne, Phys. Rev. B 1995, 51, 4014.
- [32] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, J. Machine Learning Res. 2011, 12, 2825.
- [33] B. Sadigh, P. Erhart, Phys. Rev. B 2012, 86, 134204.
- [34] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, J. Phys. Condens. Matter 2017, 29, 273002.
- [35] A. Stukowski, Modell. Simul. Mater. Sci. Eng. 2010, 18, 015012.
- [36] P. Erhart, J. Appl. Phys. 2012, 111, 113502.
- [37] V. Van Speybroeck, K. Hemelsoet, L. Joos, M. Waroquier, R. G. Bell, C. R. A. Catlow, *Chem. Soc. Rev.* 2015, 44, 7044.
- [38] D. W. Fickel, R. F. Lobo, J. Phys. Chem. C 2010, 114, 1633.
- [39] L. Chen, H. Falsig, T. V. W. Janssens, J. Jansson, M. Skoglundh, H. Grönbeck, *Catal. Sci. Technol.* 2018, 8, 2131.

Manuscript received: September 30, 2020 Revised manuscript received: December 3, 2020 Version of record online: January 22, 2021



Contents

Supporting Notes	1
1. Total energy calculations	1
2. Extension to other frameworks	2
3. Compensation by homogeneous background charge	2
4. Electronic structure calculations	3
5. Alloy cluster expansions	3
6. Monte Carlo simulations	3
7. Data and code availability	4
Supporting References	4

Supporting Notes

Supporting Note 1: Total energy calculations.

Electronic structure calculations were carried out within density functional theory in the projector augmented wave (PAW) formalism as implemented in the Vienna ab-initio simulation package. Standard PAW setups were employed for Si (4 valence electrons), O (6), H (1), and Na (1); in the case of K (7) and Rb (9) the 3p and 4s/4p semicore states, respectively, were included as well. The PBE functional was employed to describe exchange and correlation [1]. The plane-wave energy cutoff was set to 520 eV throughout. Both atomic positions and cell metric were relaxed until the maximum force fell below 30 meV/angstrom and stresses did not exceed 0.5 GPa. The Brillouin zone was sampled using a k-point mesh with a spacing of at most 0.35 Å⁻¹, equivalent to a 2 × 2 × 2 Monkhorst-Pack mesh for the primitive cell.

The initial evaluation of different crystallographic sites for the CHA (chabasite) framework [2], was carried out based on the primitive (36-atom) unit cell (Supporting Figure 1), employing a homogeneous background charge to describe the *isolated* charged species (Al⁺³, H⁺, Na⁺, K⁺, Rb⁺). In the case of H⁺, the sites associated with O⁻² ions (18f, 18g, two types of 18h sites; see Supporting Table 3) to be energetically strongly favored over sites in the channels or pores (3b, 6c, 9e). The latter are at least 1.8 eV higher in energy and are thus energetically prohibitive, whence they have been excluded from the cluster expansion (CE) (see Supporting Note 5). The opposite is observed for the alkaline species (Na⁺, K⁺, Rb⁺), for which the sites that directly neighbor an O⁻² site (18f, 18g, two types of 18h sites) are even unstable and hence have not been included in the CE.

Reference structures for the construction of alloy CEs for CHA were generated by enumeration [3, 4] for configurations with up to two Al atoms and by randomization over the entire composition range. Calculations for the construction of CEs were primarily carried out using the primitive (36-atom) unit cell. While the Al-rich end of the composition range is not of interest in itself here, we considered the entire concentration range in order to improve the stability and reliability of the CEs derived from these data. In the case of compensation by H^+ , Na^+ , K^+ , and Rb^+ the counterion-to-Al ratio was fixed to one to ensure charge neutrality.

To model compensation by free carriers configurations with only Al (no counterions) were created and charge compensation was achieved by a homogeneous background charge. Since the calculations become unreliable for large background charges, the Al/Si ratio in these cases was limited to be below 3/9 (also see Supporting Note 3).



Supporting Figure 1: Left: Location of 3b, 6c, and 9e Wyckoff sites in the SSZ-13 structure. Right: Energy differences between different Wyckoff sites for Na⁺, K⁺, and Rb⁺.

Supporting Note 2: Extension to other frameworks.

To demonstrate that clustering of Al in the presence of H counterions is not specific for the CHA structure, additional calculations were carried out for ACO, MFI, MOR and SOD frameworks, which were retrieved from the Database of Zeolite Structures [2]. For these calculations we increased the primitive unit cell to reach 192 (ACO), 108 (CHA), 288 (MFI), 144 (MOR) and 288 (MOR) sites. We then inserted 2 or 4 Al atoms in the structure, either placed far apart from each other or in the form of a cluster, following the motifs observed for CHA (see Fig. 3 of the main paper). These structures were subsequently relaxed using either a homogeneous background charge or by hydrogen ions, placed at the oxygen sites neighboring the Al sites. Relaxation and evaluation of the total energies of these structures were carried out using the same computational parameters as for the calculations described above. We then computed the binding energies as the total energy difference between clustered and dissociated configurations with negative binding energies indicating that association is favorable. The results demonstrate that the behavior observed in CHA is qualitatively and possibly even partly quantitatively transferable to other zeolite frameworks (Supporting Table 1).

Supporting Table 1: Binding energies in eV of Al clusters obtained via PBE calculations for different framework types.

Compensation by	Free carriers		H^+	
Number of Al	2	4	2	4
ACO	0.58	1.44	-0.16	-0.36
CHA	0.51	1.08	-0.19	-0.36
MFI	0.49	1.31	-0.19	-0.49
MOR	0.46	1.10	-0.13	-0.23
SOD	0.56	1.43	0.01	-0.08

Supporting Note 3: Compensation by homogeneous background charge.

Compensation with a homogeneous background charge is widely employed for treating different charge states of defects in solids. In these cases one is usually interested in the so-called dilute limit, i.e. very small defect concentrations, and the combination of a homogeneous background charge with a localized defect charge state can lead to image charge interactions (along with other finite-size effects) [5]. Since the latter can make a sizable contribution to the total energy, a host of correction schemes has been devised to remove this contribution. The latter is especially large in materials with small dielectric constants and thus little screening (such as zeolites) as well as for large defect charge states (since the correction scales to leading order with the square of the excess charge).

In the present case, we are, however, decidedly not interested in the dilute but the concentrated limit and hence these image interactions are not an artifact but rather an important part of the energetics. We therefore did not apply any corrections and used the as-calculated energies for the construction of the CE. We still limited the Al/Si ratio to be below 3/9 in order to capture potential chemical interactions for smaller Al concentrations. If the full concentration range was sampled (which could also lead to convergence problems in the density functional theory (DFT) self-consistency cycle), one should merely recover the fully ionic picture dominated by electrostatics, leading to the Loewenstein rule in its original form.

Finally, we note that in the case of the explicit counterions, charge compensation is achieved through two (relatively) localized charges and the interaction decays rather quickly. In contrast, in the case of charge-compensation through delocalized charges the interaction is more long-ranged. This is reflected in a longer-ranged CE and a higher cross-validation (CV)-root-mean-square error (RMSE) score.

Supporting Note 4: Electronic structure calculations.

The electronic levels of different species in the CHA framework were calculated using $2 \times 2 \times 2$ (288-atom) supercells to minimize defect-defect interactions and to obtain clean defect signatures. As further validation we carried out calculations using the PBE0 exchange-correlation functional [6]. While the latter, as expected, yields a larger band gap, the positions of the levels relative to the band edges varies by 0.1 eV or less.

Supporting Table 2: Level positions in CHA obtained from PBE and PBE0 calculations using 288-atom supercells.

Configuration	Level		relative to
	PBE	PBE0	
H^+	-0.80	-0.73	CBM
Na^+	-0.59	-0.63	CBM
K^+	-0.51	-0.56	CBM
Rb^+	-0.50	-0.55	CBM
H–Al	0.16	0.18	VBM
	-0.20	-0.21	VBM

Supporting Note 5: Alloy cluster expansions.

Alloy CEs [7] were constructed by automatic relevance determination regression using the ICET package [4] (version 1.3), which uses SCIKIT-LEARN for optimization tasks [8]. Alloy CEs can be written in the general form

$$E = E_0 + \sum_{\alpha} m_{\alpha} J_{\alpha} \bar{\Pi}_{\alpha}(\boldsymbol{\sigma}), \tag{1}$$

where E denotes the energy. The summation runs over all symmetry inequivalent clusters α with multiplicity m_{α} and effective cluster interaction (ECI) J_{α} . The cluster correlations α are computed as symmetrized averages of products over the pseudospin vector $\boldsymbol{\sigma}$. The latter represent the occupation of lattice sites by, e.g., Al, H, Na etc. To describe Al as well as counterion distributions, both respective sublattices were included in the construction of the CE. The Wyckoff sites included in each case are shown in Supporting Table 3 and Supporting Figure 1. We note that this approach has already been successfully applied to describe chemical ordering in clathrates, another class of inclusion compounds based on group 13 and 14 elements [9].

A systematic convergence study with respect to clusters to include in the summation was carried out, considering pair terms up to 6.9 Å and triplet terms up to 5.0 Å. The performance of these CE models was evaluated by the cross-validated root mean square error (CV-RMSE), which was computed by k-fold cross-validation (see Ref. 4 for details). The final CEs that were used in the subsequent Monte Carlo (MC) simulations included only pair terms up to a range of 5.5 Å as the inclusion of more terms did not improve the CV-RMSE. CV-RMSE values and other pertinent information regarding the CEs used for MC simulations are compiled in Supporting Table 3.

Supporting Note 6: Monte Carlo simulations.

MC simulations were carried using both the canonical and the variance constrained semi-grand canonical (VCSGC) ensemble [10] as implemented in the MCHAMMER module of ICET. The acceptance probability in the VCSGC is given by (here we adopt the notation in Ref. 11, where further details can be found)

$$\mathcal{P} = \min\left\{1, \exp\left[-\beta\Delta E - \kappa\Delta N_B(\phi + \Delta N_B/N + 2N_B/N)\right]\right\},\tag{2}$$

where ΔE is the energy change during a trial move, ΔN_B is the change in the number of *B* species, *N* is the number of available sites and $\beta = 1/k_BT$ is the inverse temperature. The VCSGC ensemble is controlled via ϕ and κ where κ determines the strength of the constraint on the variance of the concentration while ϕ determines the average concentration. Inspection of the expression above, shows that for $\kappa = 0$ one recovers the canonical ensemble. A full derivation and an example application of the VCSGC ensemble can be found in Refs. 10 and 11, respectively. Since the VCSGC ensemble led to better statistics and higher acceptance probabilities than the canonical ensemble [10], here we only report results from the former.

Simulations were conducted using $2 \times 2 \times 2$ supercells (96 Al/Si sites, 192 O/H sites, 48 Na/K/Rb sites) and run for up to 10⁶ MC trial moves. To preserve the counterion-to-Al ratio, a trial move consisted of simultaneously adding (or removing) both an Al and a counterion. Since the binding energy between Al and H is on the order of 1 eV, trial moves were furthermore constrained such that Al–H pairs remained intact. The concentration range was sampled by scanning the ϕ parameter of the VCSGC ensemble from -2.25 to 0.05 while keeping the κ parameter at 100. The number of Al–Al nearest neighbors (NNs) was determined by counting the number of oxygen sites with two Al neighbors. The number of Al–Al NNs was averaged over the MC trajectory to obtain the data shown in Figure 1 of the paper.

Compensation	Number of reference structures	CV-RMSE (meV/atom)	Singlets	Pairs	Wyckoff sites
Н	875	2.7	5	58	18f, 18g, $2 \times 18h$
Na	228	3.0	4	16	3b, 6c, 9e
Κ	233	1.7	4	14	3b, 6c, 9e
Rb	69	5.4	4	14	3b, 6c, 9e
free carriers	248	6.6	1	7	_

Supporting Table 3: Alloy cluster expansions used for MC sampling.

Supporting Note 7: Data and code availability.

The results of the DFT calculations have been compiled in a set of ASE databases [12]. They are available as a Zenodo dataset, which also contains scripts for the construction and sampling of CEs as well as selected results of these simulations [13].

Supporting References

- J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Physical Review Letters 77, 3865 (1996). doi:10.1103/PhysRevLett.77.3865.
- [2] C. Baerlocher and L. McCusker, *Database of Zeolite Structures*, http://www.iza-structure.org/databases/.
- [3] G. L. W. Hart and R. W. Forcade, Generating derivative structures from multilattices: Algorithm and application to hcp alloys, Physical Review B 80, 014120 (2009). doi:10.1103/PhysRevB.80.014120.
- [4] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, and P. Erhart, *icet A Python library for constructing and sampling alloy cluster expansions*, Advanced Simulation and Theory 2, 1900015 (2019). doi:10.1002/adts.201900015.
- [5] G. Makov and M. C. Payne, Periodic boundary conditions in ab initio calculations, Physical Review B 51, 4014 (1995). doi:10.1103/PhysRevB.51.4014.
- [6] C. Adamo and V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, Journal of Chemical Physics 116, 6158 (1999). doi:10.1063/1.478522.

- [7] J. M. Sanchez, F. Ducastelle, and D. Gratias, Generalized cluster description of multicomponent systems, Physica A: Statistical Mechanics and its Applications 128, 334 (1984). doi:10.1016/0378-4371(84)90096-7.
- [8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *Scikit-Learn: Machine Learning in Python*, Journal of Machine Learning Research 12, 2825 (2011). http://www.jmlr.org/papers/ v12/pedregosa11a.html.
- M. Ångqvist and P. Erhart, Understanding chemical ordering in intermetallic clathrates from atomic scale simulations, Chemistry of Materials 29, 7554 (2017). doi:10.1021/acs.chemmater.7b02686.
- [10] B. Sadigh and P. Erhart, Calculations of excess free energies of precipitates via direct thermodynamic integration across phase boundaries, Physical Review B 86, 134204 (2012). doi:10.1103/PhysRevB.86.134204.
- [11] M. Ångqvist, J. M. Rahm, L. Gharaee, and P. Erhart, Structurally driven asymmetric miscibility in the phase diagram of W-Ti, Physical Review Materials 3, 073605 (2019). doi:10.1103/PhysRevMaterials.3.073605.
- [12] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Du\lak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, *The atomic simulation environment* — *a Python library for working with atoms*, Journal of Physics: Condensed Matter **29**, 273002 (2017). doi:10.1088/1361-648X/aa680e.
- [13] Data and code that was used to generate the results presented here is available via zenodo at doi:10.5281/zenodo.4287250.