Computational design of nanoclusters by property-based genetic algorithms: Tuning the electronic properties of (TiO_{2})_{n} clusters

Saswata Bhattacharya, Benjamin H. Sonin, Christopher J. Jumonville, Luca M. Ghiringhelli, and Noa Marom

Phys. Rev. B **91**, 241115 — Published 26 June 2015

DOI: 10.1103/PhysRevB.91.241115

Computational Design of Nanoclusters by Property-Based Genetic Algorithms: Tuning the Electronic Properties of $(TiO_2)_n$ Clusters

Saswata Bhattacharya[1], Benjamin H. Sonin[2], Christopher J. Jumonville[2], Luca M. Ghiringhelli,[*,1] and Noa Marom,[*,2]

[1]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195, Berlin, Germany*
[2]*Physics and Engineering Physics, Tulane University, New Orleans, Louisiana 70118, USA*

In order to design clusters with desired properties, we have implemented a suite of genetic algorithms tailored to optimize for low total energy, high vertical electron affinity (VEA), and low vertical ionization potential (VIP). Applied to $(TiO_2)_n$ clusters, the property-based optimization reveals the underlying structure-property relations and the structural features that may serve as active sites for catalysis. High VEA and low VIP are correlated with the presence of several dangling-O atoms and their proximity, respectively. We show that the electronic properties of $(TiO_2)_n$ up to n=20 correlate more strongly with the presence of these structural features than with size.

The rapidly growing field of nanocatalysis [1-4] harnesses the tunability of nano-structured materials to increase the activity and selectivity of catalysts, while lowering their cost by reducing the amount of rare elements. The properties of nanocatalysts depend strongly on the size, composition, and local structure of their constituents. Nanocatalysts often comprise metal or metal-oxide clusters dispersed on a high surface area support, owing to the enhanced reactivity of clusters compared to bulk materials [5-13]. In addition, clusters may serve as models to reveal the structure and function of catalytically active sites on surfaces [14-16].

The role of first principles simulations has been expanding from elucidating the active sites and reaction mechanisms involved in catalytic processes to computational screening and design of potential catalysts [17-21]. Efforts to computationally design catalysts have largely focused on solid alloys and surfaces. At the same time, the vast majority of computational studies of clusters have focused on searching for their global minimum structures [10, 22-25]. However, the most stable structures of atomic clusters are not necessarily optimal for catalysis. Rather, the presence of active sites has been shown to be a key factor in catalysis by clusters [26-31].

Our quest to computationally design atomic clusters for applications in nanocatalysis embarks from the hypothesis that clusters possessing a high vertical electron affinity (VEA) or a low vertical ionization potential (VIP) may be more promising candidate catalysts because they would be more chemically reactive (as they would accept or donate an electron more readily). We further assume that these electronic properties are correlated with the presence of potentially active sites [32]. Based on these premises, we have implemented a suite of three massively parallel cascade genetic algorithms (GAs). The first is the energy-based GA (EGA), described in detail in [33, 34]. The second is tailored to search for clusters with a high VEA (VEA-GA) and the third is tailored to search for a low VIP (VIP-GA). Here, we apply these to the case of $(TiO_2)_n$ clusters with n=2-10,15,20.

Analysis of the optimal structures found by the property-based GAs reveals the underlying structure-property relations. We show that the electronic properties of $(TiO_2)_n$ clusters depend more strongly on structure than on size. In particular, we find a clear correlation between certain structural motifs and the magnitude of the VEA, VIP, and fundamental gap. Contrary to common belief, we do not find a clear size trend. We attribute this to the fact that particular structural features related to high VEA and low VIP are less favorable energetically and therefore less likely to be found in larger clusters.

GAs have been used extensively for finding the energy global minimum (GM) of crystalline solids [35-41] and clusters [22-25, 42]. The strategy of a GA is to perform global optimization by simulating an evolutionary process [22-24, 43]. First, local optimizations are performed for an initial population of randomly generated structures. The scalar descriptor (or combination of descriptors) being optimized is mapped onto a fitness function. The structures with the highest fitness are assigned a higher probability for mating. In the mating step a crossover operator creates a child structure by combining fragments (or structural "genes") of two parent structures. The child structure is locally optimized and added to a common pool if determined to be different than the existing structures. The cycle of local optimization, fitness evaluation, and generation of new structures is repeated until convergence is achieved. Advantageously, the fitness function may be based on any property of interest, not necessarily the total or free energy [22, 44-48]. We rely on this to tailor GAs to explore the configuration space of desired electronic properties.

Within our cascade GA approach [33, 34] successive steps employ an increasingly accurate level of theory. Structural information is passed between steps of the cascade and some structures are filtered out. This

considerably increases the efficiency of the GA and reduces the time to solution.

Initially, the reactive force field ReaxFF [49-51] is used for an exhaustive GA pre-screening of possible structures. Structures found within an energy window of 4 eV from the energy GM are then transferred to a density-functional theory (DFT) based GA. DFT calculations are performed with the all-electron numeric atom-centered orbitals code FHI-aims [52, 53]. The cascade flow proceeds as follows: In the first step local optimizations are performed with the Perdew-Burke-Ernzerhof (PBE) [54, 55] functional, using *lower-level* settings [56]. Structures that are already in the pool or outside an energy window of 2 eV above the running GM are rejected [57]. The remaining structures are passed to the second step of the cascade, where their energies are calculated with the PBE-based hybrid functional (PBE0) [58] to evaluate their fitness [59]. For the electronic property based GAs, the VEA/VIP are evaluated based on the total energy difference between the neutral species and the anion/cation. After the GA cycle converges the 50 fittest isomers (for each cluster size) are post-processed at the *higher-level* settings [55]. The structures are first re-optimized using PBE. Then, PBE0 is used for the final single point total energy and VEA/VIP evaluations. We have verified that among isomers the *hierarchy* of the PBE0-level electronic properties (total energy, VEA, VIP) is largely conserved upon switching from *lower-level* to *higher-level* settings. Since the hierarchy of the optimized quantity is more important than the absolute values, this enables a considerable reduction of the computational effort by running the potential energy surface scan with *lower level* settings. A complete account of the GA implementation and validation is provided in the SI, including details on how crossover, mutation, and similarity checks are performed and how convergence is determined [60].

For the first application of the property-based GAs we have chosen the case of $(TiO_2)_n$ clusters with up to 20 stoichiometric units [61]. The three GAs ran independently starting from the initial pool generated by the force field based GA until convergence was achieved. In Figure 1 their performance is compared in terms of the ability to find structures that optimize different target properties. Histograms of the number of structures found for each property value by the respective GA are shown for $(TiO_2)_n$ clusters at selected sizes. The data shown here were obtained using the PBE0 functional with the *lower-level* settings, reflecting the level of theory used to evaluate the fitness function in the final step of the cascade scheme.

For small clusters (n=2,3) the three algorithms find the same isomers because the configuration space is small and there are only a few structures in the search window. Starting from n=4, it is evident that the three GAs explore different regions of the configuration space. The centers of the VEA-GA histograms in panel 1a are clearly shifted to higher VEAs, compared to the EGA histograms and there are more structures in the tail regions with particularly high VEA. Similarly, the VIP-GA histograms in panel 1b are shifted to lower VIPs. From the energy histograms in panel 1c, it is evident that the VEA-GA and VIP-GA systematically explore higher energy regions of the configuration space than the EGA. Hence, they efficiently find more structures with the desired electronic properties than the EGA [62].

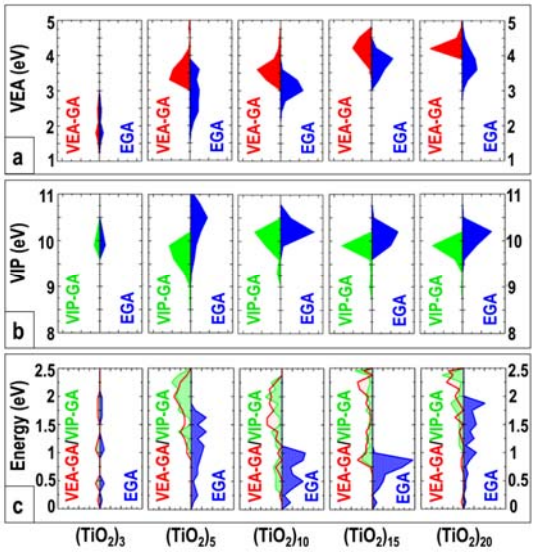

Figure 1. Comparison of the three GAs in terms of their ability to find structures with the target properties of (a) high VEA, (b) low VIP, and (c) low energy. Histograms represent the number of isomers found by the respective GA for $(TiO_2)_n$ clusters with n=3,5,10,15,20.

Analysis of the structures found by the property-based GAs reveals the structural features correlated with a high VEA and a low VIP. Figure 2a shows the Ti-O pair correlation functions (PCFs) of the ten best $(TiO_2)_5$ clusters found by the VEA-GA and the EGA. A peak at 1.65 Å is clearly more prominent in the VEA-GA PCF, compared to the EGA PCF. This is observed for all cluster sizes (see SI). This peak corresponds to the bond length of dangling-O atoms, shown in Figure 2b. Figure 2c shows the average and maximum VEA (upper panel) and number of dangling-O atoms (lower panel) for the same sample of isomers as for the PCFs. The data shown from here on (including in Figures 3, 4) correspond to the *higher-level* settings of the final post-processing stage.

For the smallest clusters with n=2,3 the graphs overlap because both algorithms find the same isomers. Starting from n=4, the graphs diverge. The high VEA isomers clearly have a larger number of dangling-O atoms than the low energy isomers. Therefore, a correlation can be drawn between the number of dangling-O atoms and a high VEA. Many of the high-VEA isomers with n≥4 have 3-4 dangling-O atoms. The number of dangling-O atoms decreases for high-VEA clusters with n≥10 because this structural motif becomes increasingly unfavorable energetically with size. Many of the high VEA clusters also have the tri-coordinated Ti site, reported in [32], as shown in the SI. Thus, the VEA-GA has revealed a new structure-property relation.

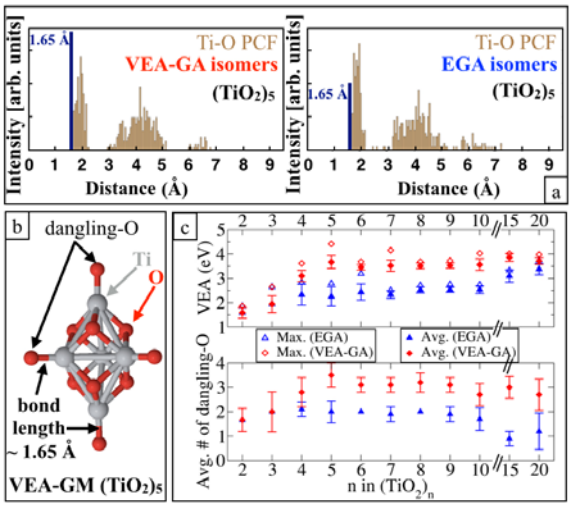

Figure 2. a) Cumulative Ti-O PCFs of the ten best $(TiO_2)_5$ isomers found by the VEA-GA (top) vs. the EGA (bottom); b) Visualization of the $(TiO_2)_5$ isomer with the highest VEA, showing its four dangling O atoms; c) The average and maximum VEA (top) and number of dangling O atoms (bottom) of the 10 isomers with the highest VEA vs. the 10 isomers with the lowest energy for all cluster sizes.

Visual inspection of the structures found by the VIP-GA reveals that they tend to have two dandling-O atoms in proximity to each other. The corresponding descriptor is the *bond connectivity* (the number of Ti-O bonds along the shortest path between two Ti atoms attached to dangling-O atoms, see Fig. 3a). Figure 3b shows the average and minimum VIP (upper panel) and the bond connectivity (lower panel) for the 10 isomers with the lowest VIP (found by the VIP-GA) vs. the 10 isomers with the lowest energy (found by the EGA) for all cluster sizes. Starting from n=4, the bond connectivity of the low-VIP isomers is consistently lower than that of the low-energy isomers. For n=4-9, the lowest VIP isomers have a bond connectivity of 2. For larger clusters (n≥10), the bond connectivity is higher. The reason is that energetically unfavorable under-coordinated O atoms are less likely to occur in larger clusters and if they do occur they are less likely to be in proximity to each other. This observation is supported by the fact that the EGA does not produce any isomers with more than one dangling-O for n=15. The VIP-GA does find structures with a low VIP and a bond connectivity of 2 for n≥10, however they fall outside of the 2 eV energy window and are therefore not shown in Figure 3.

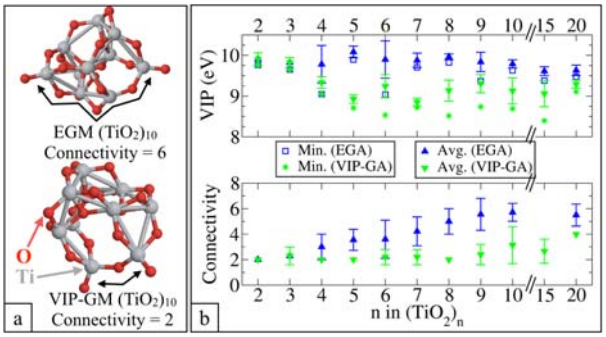

Figure 3. a) Bond connectivity of the energy global minimum (EGM) vs. the VIP-GM for $(TiO_2)_{10}$; b) The average and minimum VIP (top) and bond connectivity between dangling-O atoms (bottom) of the 10 isomers with the lowest VEA vs. the 10 isomers with the lowest energy for all cluster sizes.

Owing to quantum confinement effects, one would expect size trends in the electronic properties of clusters, whereby the VEA increases, the VIP decreases, and the gap narrows with size. The expected size trends are not readily apparent in Figures 2c and 3b. The VEA increases fast for the smallest isomers and then stabilizes and becomes almost constant. The VIP decreases for the smallest clusters and then fluctuates. We further investigate whether there is a size trend in the fundamental gap (VIP□VEA) of the structures found in the three searches.

Figure 4a shows the VIP vs. VEA for the best 10 structures found in each search for all cluster sizes. The loci of constant fundamental gap are indicated by diagonal lines. The results of the EGA, VEA-GA, and VIP-GA are clustered in different regions of the graph. The clusters found by the VEA-GA and VIP-GA tend to have gaps of 5.5-6.5 eV while the clusters found by the EGA tend to have gaps of 7-8 eV. Considering that a narrow gap requires a combination of a high VEA and a low VIP, it is not surprising that the VEA-GA and VIP-GA generally find structures with smaller gaps than the EGA [63].

What is more unexpected is that the clusters with the smallest gaps are not necessarily the largest ones. The three clusters with gaps below 5 eV have n=6,15, and 10. Between 5 to 5.5 eV we find clusters with n=5,7,8,9,10,15, and 20. Figure 4b shows the average and minimum gap of the 10 best structures found in

each search for all cluster sizes. A clear trend of narrowing gap with increasing size is seen only for the smallest clusters. For the larger clusters the EGA isomers show a weak trend of gap narrowing with size. The isomers found by the VEA-GA and VIP-GA exhibit significant fluctuations of the minimum and average gap and no clear size trend is visible. Our findings are consistent with those of Shevlin *et al.* for certain families of nitride clusters [64].

The absence of the expected size trends may be explained by the structural features associated with a high VEA, a low VIP, and a narrow gap. These features become less energetically favorable with increasing size because they involve a number of dangling-O atoms, some of which are in proximity to each other. Therefore, they are less likely to appear in larger clusters. We have thus demonstrated that the electronic properties of TiO_2 clusters with up to 20 stoichiometric units correlate more strongly with the presence of specific structural features than with the cluster size. For larger clusters we expect to eventually reach a size regime where isomers possessing multiple dangling-O atoms completely disappear from the search window and the electronic properties correlate more strongly with size. Indeed, such size trends have been reported for the IP, EA, and gap of bulk-like rutile TiO_2 nanocrystals without any dangling-O sites [65].

In summary, we have implemented a suite of three cascade genetic algorithms tailored to optimize cluster structures for low total energy (EGA), high vertical electron affinity (VEA-GA), and low vertical ionization potential (VIP-GA). Analysis of the structures found by the VEA-GA and the VIP-GA vs. the EGA reveals the following structure-property relations: A high VEA is correlated with a number of dangling-O atoms (typically 3-4) and a low VIP is correlated with low bond connectivity (typically 2) between two dangling-O atoms. These structure-property relations explain the absence of the expected size trends. Smaller clusters may have a higher VEA, a lower VIP, and a narrower gap than larger clusters because the structural features associated with these properties become less favorable with increasing size.

We further suggest that the presence of dangling-O atoms on TiO_2 clusters or surfaces may be associated with enhanced catalytic activity and that these O atoms may serve as the active sites. Our findings hint at a new approach to the computational design of cluster-based nanocatalysts, employing property-based GAs. These broadly applicable algorithms may be modified to search for any desired electronic property (that can be mapped onto a fitness function) and may be extended to clusters of varying composition as well as cluster/support systems. The process of optimization for a target property, associated with enhanced reactivity, reveals the underlying structure-property relations and the structural features that may serve as active sites for catalysis. This may provide valuable physical insight and design rules for better nanocatalysts.

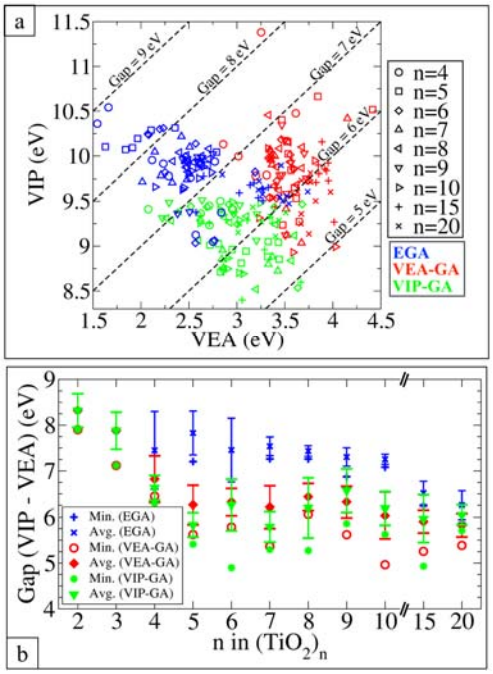

Figure 4. a) VIP vs. VEA for the best 10 structures found in each search for all cluster sizes. The loci of constant fundamental gap are indicated by diagonal lines. b) The average and minimum gap of the 10 best structures found in each search for all cluster sizes.

We thank Kristen Fichthorn and Muralikrishna Raju from Penn State for sharing the reaxFF parameters they used for TiO_2 surfaces. We thank Matthias Scheffler and Sergey Levchenko from FHI-Berlin for helpful discussions and a critical reading of the manuscript. Work at Tulane University was supported by the Louisiana Alliance for Simulation-Guided Materials Applications (LA-SiGMA), funded by the National Science Foundation (NSF) award number #EPS-1003897. Computer time was provided by the Argonne Leadership Computing Facility (ALCF) at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357 and by the Rechenzentrum Garching (RZG) of the Max-Planck Gesellschaft

[1] U. Heiz, and U. Landman, *Nanocatalysis* (Springer-Verlag, Berlin Heidelberg, 2007).

[2] F. Zaera, The Journal of Physical Chemistry Letters **1**, 621 (2010).

[3] F. Zaera, Chemical Society Reviews **42**, 2746 (2013).
[4] Y. Li, and G. A. Somorjai, Nano Lett. **10**, 2289 (2010).
[5] L. B. Vilhelmsen, and B. Hammer, Phys. Rev. Lett. **108**, 126101 (2012).
[6] A. A. Herzing, C. J. Kiely, A. F. Carley, P. Landon, and G. J. Hutchings, Science **321**, 1331 (2008).
[7] J. D. Stiehl, T. S. Kim, S. M. McClure, and C. B. Mullins, J. Am. Chem. Soc. **126**, 13574 (2004).
[8] T. S. Kim, J. D. Stiehl, C. T. Reeves, R. J. Meyer, and C. B. Mullins, J. Am. Chem. Soc. **125**, 2018 (2003).
[9] J. Oliver-Meseguer, J. R. Cabrero-Antonino, I. Domínguez, A. Leyva-Pérez, and A. Corma, Science **338**, 1452 (2012).
[10] R. V. Chepulskii, and S. Curtarolo, ACS Nano **5**, 247 (2010).
[11] R. Kretschmer, Z.-C. Wang, M. Schlangen, and H. Schwarz, Angewandte Chemie International Edition **52**, 9513 (2013).
[12] M. n. Calatayud, L. Maldonado, and C. Minot, The Journal of Physical Chemistry C **112**, 16087 (2008).
[13] Z.-C. Wang, S. Yin, and E. R. Bernstein, The Journal of Physical Chemistry Letters **3**, 2415 (2012).
[14] D. K. Böhme, and H. Schwarz, Angewandte Chemie International Edition **44**, 2336 (2005).
[15] H. Schwarz, Angewandte Chemie International Edition **50**, 10096 (2011).
[16] A. W. Castleman, and K. H. Bowen, The Journal of Physical Chemistry **100**, 12911 (1996).
[17] J. K. Norskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen, Nat Chem **1**, 37 (2009).
[18] P. M. Holmblad *et al.*, Catal Lett **40**, 131 (1996).
[19] J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff, and J. K. Norskov, Nat Mater **5**, 909 (2006).
[20] J. Greeley, and M. Mavrikakis, Nat Mater **3**, 810 (2004).
[21] J. Greeley, and J. K. Nørskov, Surf Sci **601**, 1590 (2007).
[22] R. L. Johnston, Dalton Trans., 4193 (2003).
[23] S. Heiles, and R. L. Johnston, Int. J. Quantum Chem. **113**, 2091 (2013).
[24] M. Sierka, Progress in Surface Science **85**, 398 (2010).
[25] C. R. A. Catlow *et al.*, Phys. Chem. Chem. Phys. **12**, 786 (2010).
[26] P. J. Roach, W. H. Woodward, A. W. Castleman, A. C. Reber, and S. N. Khanna, Science **323**, 492 (2009).
[27] F. Shimojo, S. Ohmura, R. K. Kalia, A. Nakano, and P. Vashishta, Phys. Rev. Lett. **104**, 126102 (2010).
[28] M. B. Abreu, C. Powell, A. C. Reber, and S. N. Khanna, J. Am. Chem. Soc. **134**, 20507 (2012).
[29] Z. Luo, J. C. Smith, W. H. Woodward, and A. W. Castleman, The Journal of Physical Chemistry Letters **3**, 3818 (2012).
[30] W. H. Woodward, A. C. Reber, J. C. Smith, S. N. Khanna, and A. W. Castleman, The Journal of Physical Chemistry C **117**, 7445 (2013).
[31] O. A. Syzgantseva, P. Gonzalez-Navarrete, M. Calatayud, S. Bromley, and C. Minot, The Journal of Physical Chemistry C **115**, 15890 (2011).
[32] N. Marom, M. Kim, and J. R. Chelikowsky, Phys. Rev. Lett. **108**, 106801 (2012).
[33] S. Bhattacharya, S. V. Levchenko, L. M. Ghiringhelli, and M. Scheffler, Phys. Rev. Lett. **111**, 135501 (2013).
[34] S. Bhattacharya, S. V. Levchenko, L. M. Ghiringhelli, and M. Scheffler, New J. Phys. **16**, 123016 (2014).
[35] A. R. Oganov, and C. W. Glass, The Journal of Chemical Physics **124**, 244704 (2006).
[36] C. W. Glass, A. R. Oganov, and N. Hansen, Computer Physics Communications **175**, 713 (2006).
[37] S. Wu *et al.*, Phys. Rev. B **83**, 184102 (2011).
[38] S. M. Woodley, P. D. Battle, J. D. Gale, and C. R. A. Catlow, Phys. Chem. Chem. Phys. **1**, 2535 (1999).
[39] G. Trimarchi, and A. Zunger, Journal of Physics: Condensed Matter **20**, 295212 (2008).
[40] D. C. Lonie, and E. Zurek, Computer Physics Communications **182**, 372 (2011).
[41] G. H. Jóhannesson *et al.*, Phys. Rev. Lett. **88**, 255506 (2002).
[42] D. M. Deaven, and K. M. Ho, Phys. Rev. Lett. **75**, 288 (1995).
[43] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
[44] N. M. O'Boyle, C. M. Campbell, and G. R. Hutchison, The Journal of Physical Chemistry C **115**, 16200 (2011).
[45] A. Jain, I. Castelli, G. Hautier, D. Bailey, and K. Jacobsen, J Mater Sci **48**, 6519 (2013).
[46] M. d'Avezac, J.-W. Luo, T. Chanier, and A. Zunger, Phys. Rev. Lett. **108**, 027401 (2012).
[47] L. Zhang, J.-W. Luo, A. Saraiva, B. Koiller, and A. Zunger, Nat Commun **4**, 2396 (2013).
[48] A. L. S. Chua, N. A. Benedek, L. Chen, M. W. Finnis, and A. P. Sutton, Nat Mater **9**, 418 (2010).
[49] M. Raju, S.-Y. Kim, A. C. T. van Duin, and K. A. Fichthorn, The Journal of Physical Chemistry C **117**, 10558 (2013).

[50] M. Raju, A. C. T. van Duin, and K. A. Fichthorn, Nano Lett. **14**, 1836 (2014).
[51] S. Monti, A. C. T. van Duin, S.-Y. Kim, and V. Barone, The Journal of Physical Chemistry C **116**, 5141 (2012).
[52] V. Blum *et al.*, Computer Physics Communications **180**, 2175 (2009).
[53] V. Havu, V. Blum, P. Havu, and M. Scheffler, Journal of Computational Physics **228**, 8367 (2009).
[54] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[55] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).
[56] The *lower-level* settings correspond to "light" integration grids and compact "tier 1" basis sets. The *higher-level* settings correspond to "tight" integration grids and larger "tier 2" basis sets. The numerical settings and basis sets of FHI-aims are described in detail in Ref. 51.
[57] Based on [32], isomers with energy as high as 1.25 eV above the GM may have been observed in experiments, in which TiO_2 clusters were formed by laser vaporisation. Restricting the search window to 2 eV would not eliminate any isomer that can conceivably form in such experiments.
[58] C. Adamo, and V. Barone, J Chem Phys **110**, 6158 (1999).
[59] The 2 eV energy window originally set at the PBE level expands to 2.5 eV upon switching to PBE0 because PBE0 tends to increase the energy differences between isomers. This is reflected in Figure 1c.
[60] All input and output files for the clusters reported here may be downloaded from http://nomad-repository.eu.
[61] Since the property-based GAs are based on the GA implementation described in [33-34] they have the inherent capability of treating non-stoichiometric structures on the same footing.
[62] The EGA found all the low energy isomers reported in [12, 31, 32] for all cluster sizes. For n=5 our choice of PBE0 as exchange-correlation functional and *higher-level* settings changes the ordering of the two most stable isomers, which are very close in energy, with respect to [32]. For n=10, we have found lower energy structures than those previously reported.
[63] The VEA-GA selects for structures with a large number of dangling-O atoms, regardless of their connectivity, whereas the VIP-GA selects for structures with a low connectivity regardless of the number of dangling-O atoms. We note that more structures with a narrow gap and more definitive structure-property relations may be found by tailoring a GA to search specifically for this property.
[64] S. A. Shevlin *et al.*, Phys. Chem. Chem. Phys. **10**, 1944 (2008).
[65] L. Hung, K. Baishya, and S. Öğüt, Phys. Rev. B **90**, 165424 (2014).