Stability and metastability of clusters in a reactive atmosphere: theoretical evidence for unexpected stoichiometries of Mg_MO_x

Saswata Bhattacharya, Sergey V. Levchenko, Luca M. Ghiringhelli, and Matthias Scheffler
Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany

Supplemental Material

I. Thermodynamic stability of stoichiometric vs non-stoichiometric Mg_MO_x clusters

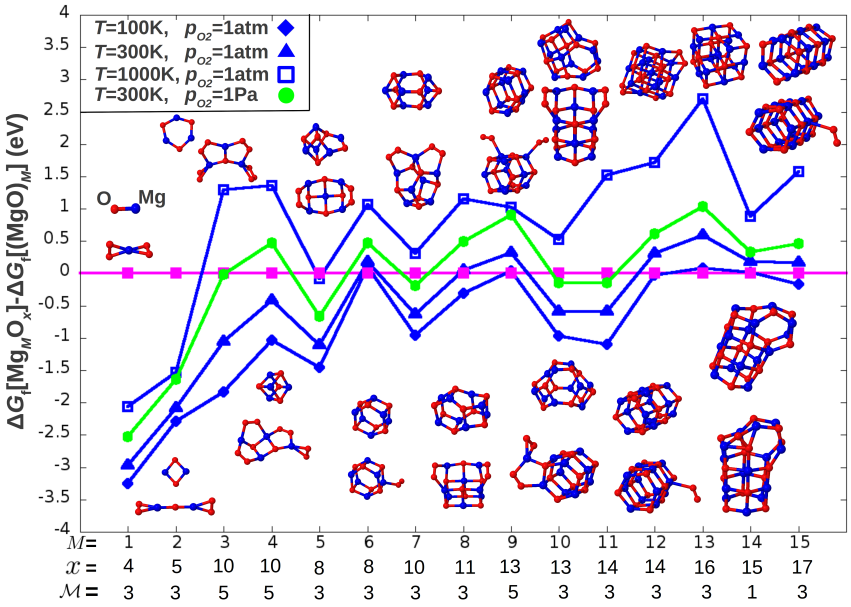


FIG. 1: Free energy of formation of thermodynamically most stable non-stoichiometric (Mg_MO_x with $M \neq x$) relative to stoichiometric ($M = x$) clusters at several (T, p_{O_2}) conditions. The geometries were optimized with PBE+vdW, and the electronic energy was calculated using PBE0+vdW. The label of the horizontal axis shoes, below the amount M of Mg atoms, the amount x of O atoms for the thermodynamically most stable non-stoichiometric cluster at the corresponding M, at $p_{O_2} = 1$ atm and $T = 300$ K. For the same thermodynamic condition, the third line reports the spin multiplicity $\mathcal{M}$ of the lowest free-energy non-stoichiometric Mg_MO_x (the stoichiometric clusters are all singlets, i.e., $\mathcal{M} = 1$).

II. Details on the implemented genetic algorithm (GA) scheme

The benchmark and full detail on validation is found in Ref. [1]

Schematically, our cascade GA (cGA) algorithm proceeds as follows (all terms in italic are explained afterwards):

(1) Selection of a composition of the clusters and formation of an initial pool of random structures, locally optimized by a classical force field (FF).

(2) Evaluation of the *fitness* function for all structures (using FF binding energy).

(3) GA global optimization using the classical FF. This consists in the iteration of steps (i)–(v):

  (i) *Selection* of two structures (in GA jargon, *parents*).

  (ii) Assemblage of a trial structure (*child*) through *crossover* and *mutation*.

  (iii) Local optimization (force minimization) of child structure using the classical FF.

  (iv) Evaluation of the *fitness* function. Comparison of the optimized child with existent structures; reject if *similar*, jump to (i).

  (v) Check whether convergence has been reached. If so, stop FF-GA and go the next step, DFT-GA.

(4) Formation of a new pool of structures using best fit structures from FF-GA, locally optimized at the DFT level (PBE+vdW, *low-level settings*).

(5) Calculation of fitness function for all structures (using energy at the PBE0+vdW level).

(6) GA scheme using DFT. In practice iteration of steps (a)–(i):

  (a) *Selection* of two structures.

  (b) Assemblage of a child structure through *crossover* and *mutation*.

  (c) Local optimization of the child structure with PBE+vdW, *low-level settings*.

  (d) Comparison of the optimized child with existent structures. *Early rejection* if *similar*; jump to (a).

  (e) Further local optimization of the child with PBE+vdW, *high-level settings*.

  (f) Comparison of the further optimized child with existent structures. *Rejection* if *similar*; jump to (a).

  (g) Harmonic analysis of the optimized child; if unstable, perturb along the unstable mode and go back to (c).

  (h) Evaluation of *fitness* function based on PBE0+vdW total energy.

  (i) Check whether convergence has been reached. If so, stop.

In the following, we analyze one by one the key words introduced in the detailed scheme above.

Fitness function

Each cluster i in the population is assigned a normalized fitness value, ρ_i, based on its total energy (binding energy for the FF):

$$\rho_i = \frac{\epsilon_i}{\sum_i \epsilon_i} \tag{1}$$

and ϵ_i is the relative energy of the i^{th} cluster as defined below:

$$\epsilon_i = \frac{E_{\max} - E_i}{E_{\max} - E_{\min}} \tag{2}$$

Where E_i is the total energy of the i^{th} cluster of the population and $E_{\min}, E_{\max}$ correspond to the dynamically updated lowest and highest total energies in the population, respectively.

With this definition, low (more negative) energy clusters have high fitness and high (less negative) energy clusters have low fitness.

Selection rule

We use a "roulette-wheel" selection criterion [2] with selection probability proportional to the value of the normalized fitness function. The idea is that the lower the total (or binding) energy (i.e., large negative value) of a certain configuration, the larger the probability to be chosen from the population. A cluster is picked at random and is selected for mating if its normalized fitness value (ρ_i) is greater than Rand$[0,1]$, a randomly generated number between 0 and 1 (i.e., if $\rho_i >$ Rand$[0,1]$); where ρ_i is the normalized fitness function defined in section-.

The above " best-fit" selection scheme can take significantly long time to reach another basin in the PES. In such situations, adding a little diversity by selecting one "bad" (high-energy) structure in the population is found to help in moving out to the next basin. Therefore, we define a complementary fitness function $\tilde{\rho}_j = (1 - \rho_j)$ and we select one structure with high ρ_i and another with high $\tilde{\rho}_j$. This choice, when the the mixing ratio among different selection rules is optimized, greatly helps the convergence of the GA scheme and we also show how we optimized the mixing ratio among different selection rules.

Crossover

The crossover operator takes care of combining the two parent clusters selected as explained above. It is implemented as a modified version of the cut-and-splice crossover operator of Deaven and Ho.[3] In our implementation of the cut-and-splice operation, first a random rotation is given (keeping the center of geometry of the cluster at the origin of the coordinate system) to both the parent clusters. Both clusters are then cut horizontally parallel to the xy-plane ($z = 0$). Atoms with positive z-value are selected from one cluster and atoms with negative z-value are selected form the other cluster. These complementary fragments are spliced together. Importantly, this cut-and-splice operation does not ensure the preservation of the chosen cluster size (i.e., the total number of atoms) and the specific composition. We have adapted here three different kind of crossovers to maintain size and composition.

(i) Crossover-1: Strictly speaking, this is a combined crossover and mutation (see below) step. After cut-and-splice we always maintain the same ordering of atoms that is given in the parent clusters. As an example, let us consider a small cluster like Mg_2O_2. In the parent cluster the ordering of atomic coordinate is given as Mg(1), O(2), Mg(3), O(4). When the cut-and-splice operation is applied, we get, for instance, a child with atomic coordinates from cluster one as Mg(1), Mg(3) (i.e., above the xy-plane) and that of from cluster two as Mg(3'), O(4') (i.e., below the xy-plane). Therefore, the entire atomic coordinates of the child are [Mg(1), Mg(3)], [Mg(3'), O(4')]. If this is the case, we replace the species of Mg(3) with O (without changing its coordinate) to impose the correct composition to the child. Thus the new ordering of atoms is: Mg(1), O(2), Mg(3), O(4) (i.e., the same composition as the parents). Therefore, it is possible that after the cut-and-splice operation a Mg atom of the parent cluster is replaced by an O atom in the child and vice versa.

Also the total spin of the clusters is left free to evolve together with the spatial coordinates of the atoms. In this way we sample on equal footing the configurational space of atomic coordinates and the spin. The crossover of the spin coordinates is performed in the following way: when we create a new child by grabbing the atomic coordinates from the parents as explained above, we also make note of the atom-projected spin moments (via Hirshfeld partitioning of the electron density) for each atom. Such spin moments are given as initial moments of the individual atoms of the child. During the optimization process, these atom-projected moments are left free to change.

(ii) Crossover-2: In this procedure, after cut-and-splice we check whether the stoichiometry of the parents is maintained in the child. If it is maintained, we accept the child, otherwise we reject it and we iterate until until the child has the required stoichiometry.[4, 5] The advantage of this procedure is that it helps to maintain winning features of the parent molecule but most of the time it takes many iterations to obtain a valid child, even for a moderately sized cluster. In case one or more pairs of atom are too close, we adopt the same remedy as for crossover-1. The spin coordinates are taken care of the same as in the crossover-1 case.

(iii) Crossover-3: After re-orientation of the selected parent clusters we take all the metal (Mg) atoms from one parent molecule and all the oxygen atoms from another parent molecule. This crossover helps introducing diversity in the genetic pool, but the rate of rejection during the assemblage of the child can be rather high due to the high likelihood that two atoms are too close.

Mutation

After crossover, which generates a child, mutation is introduced. Different mutation operators can be defined. We have adopted a) a translation between the two halves of the clusters (this is performed if atoms coming from the two different parents find themselves too close upon splicing of the two halves) and b) exchange of the atom species without perturbing their coordinates.

Similarity of structures

In order to decide whether a newly found structure was already seen previously during the GA scan, after the local optimization we a) compare the energy of the new structure with that of all the others seen before and b) use a criterion based on the distances between all the atoms' pairs. In practice, we construct a coarse grained radial distribution function (rdf) of the clusters, consisting of 14 bins conveniently spaced. Each bin contains the (normalized) number of atom-pairs whose distance is between the distances that define the boundaries of the bin. For each cluster we have then a 14-dimensional rdf-array and the euclidean distance (i.e., the square root of the sum of the squared difference between corresponding elements in the two arrays) between the arrays arranged for two clusters is evaluated. If this distance (note that it is a pure number) is greater than a convenient threshold (we used 0.01), then the structures are considered as different.

Local optimization and early rejection scheme

Although the geometry and the energy of the structures is not fully converged with PBE+vdW @ *low-level settings*, we have realized that there is a one-to-one mapping between the structures found at this level and those fully converged. In other words, if two structures are *similar* according to PBE+vdW @ *low-level settings*, they are also at the PBE0+vdW @ *high-level settings* (see below). Furthermore, if a structure at the PBE0+vdW @ *high-level settings* is within ~ 0.5 eV from the running GM, with PBE+vdW @ *low-level settings* the structure is not further than 0.2 eV from the same running GM (with energy evaluated with PBE+vdW @ *low-level settings*). This implies that, with our conservative choice of not optimizing with the *high-level settings* structures that with *low-level settings* result positive to the *similarity* test or are more than 1.5 eV higher in energy than the current GM, we are not risking to reject structures that would eventually result in the GM or close to it.

In the PBE+vdW, *high-level settings* optimization, atomic forces were converged to less than 10^{-6} eV/Å. The grid settings where set to "tight" and the basis set was tier-2. In cascade, for the structure optimized in this way (i.e., without further optimization), we evaluated the PBE0+vdW energy with the tier-2 basis set. This energy is later used for the calculation of fitness of that particular cluster. The difference in binding energy between an isomer optimized with PBE0+vdW forces (tight / tier 1 / forces converged to 10^{-5} eV/Å) and the same optimized with PBE+vdW (tight / tier 2 / forces converged to 10^{-5} eV/Å), when the energy of both geometries is evaluated via PBE0+vdW (tight / tier 2), is small, i.e. at most 0.04 eV among all cases we checked. The computational cost of the PBE0+vdW further optimization would be thus not worthy (we estimated a gain of up to a factor 2 of overall computational time just by skipping the latter optimization)

Parallelization

The operation of selecting from the genetic pool two structures for the mating and the subsequent local optimization of the child, is an operation that can be performed at any moment also when a local optimization of a child is already running. The algorithm is thus suitable for a very efficient parallelization.

On top of FHI-aims parallelization (i.e. local optimization is run in parallel on an optimized number of CPUs) we add a second level of parallelization, i.e., we run at the same time several local optimizations, independently. The only communication among such replicas is the selection of the parents that is performed from a common genetic pool. The latter is also updated by each replica at the end of each local optimization. The local optimizations run independently, i.e., each replica can start a new mating + local optimization cycle right after one is concluded; hence, there is no idling time between cycles. Thus, we have n local optimizations running in parallel, each requiring p cores , that fill the $n \times p$ cores required for the algorithm. The scaling behavior is about $O(p^{1.5})$ with the number of cores for the local optimization part [6] . The number p is indeed tuned in order to be sure that the speed-up is

still $O(p^{1.5})$ for the specific system. The scaling with respect to the n replicas is linear, because the replicas are for the most of the time independent and only at the beginning and at the end of each local optimization, information is shared among the replicas.The first level of parallelization is performed within the FHI-aims code, by means of the MPI environment. The second level is script based: The total $n \times p$ number of cores is divided into n groups, n subdirectories are created and in each of them a cycle of local optimization job runs, each using p cores.

[1] S. Bhattacharya, S. V. Levchenko, L. M. Ghiringhelli, M. Scheffler, to be submitted (2013).
[2] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA, 1989.
[3] D. M. Deaven and K. M. Ho, Phys. Rev. Lett., **75**, 288 (1995).
[4] R. L. Johnston, Dalton Trans. **22** 4193 (2003).
[5] M. Sierka, Progress in Surface Science **85**, 398 (2010).
[6] V. Havu, V. Blum, P. Havu, and M. Scheffler, J. Comp. Phys. **228**, 8367 (2009).

III. Performance of reaxFF

System	ReaxFF	PBE+vdW	PBE0+vdW	HSE06+vdW	Experiment
MgO bulk modulus [GPa]	167	149 [1]	169 [1]	169 [1]	165 [2]
MgO lattice constant [Å]	4.24	4.26 [1]	4.21 [1]	4.21 [1]	4.21 [2]
O_2 binding energy [eV]	5.38	6.23	5.37	5.31	5.22 [3]
O_2 bond distance [Å]	1.24	1.22	1.20	1.20	1.21 [4]
O_2 stretching freqency [cm^{-1}]	1694	1555	1605	1663	1580 [4]

TABLE I: Comparison of reaxFF with xc functionals on some of the properties used for training Mg and O parameters in reaxFF. The calculations for PBE+vdW, PBE0+vdW, and HSE06+vdW are done with tight-setting and tier-2 basis set.

The reaxFF parameters for Mg and O were fit [5] to a quantum-mechanics training set including MgO−bulk and O_2−molecule properties. We checked that this parametrization yields for bulk MgO lattice constant within 1 % and bulk modulus within 10 % from the experimental values; similarly, for the MgO and O_2 dimer, it gives bond length and vibrational frequency within 1 % and 10% from the respective reference values. On this set of properties, reaxFF performs remarkably well, comparable to PBE+vdW, PBE0+vdW, and HSE06+vdW, but it fails very clearly for small clusters. In particular, ReaxFF yields a qualitatively good pre-scanning for stoichiometric clusters, while for non-stoichiometric ones it is far from desirable accuracy. This is not unexpected, as not even PBE is able to account for the charge redistributions associated with non-stoichiometry clusters.

[1] J. Heyda and G. E. Scuseria, J. Chem. Phys., **121**, 1187 (2004). J. Paier, M. Marsman, K. Hummer, and G. Kresse, J. Chem. Phys. **124**, 154709 (2006). Erratum: J. Chem. Phys. **125**, 249901 (2006).
[2] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, Phys. Rev. B **69**, 075102 (2004); Erratum: **78**, 239907(E) (2008). K. Marklund and S. A. Mahmoud, Physica Scripta. **3**, 75 (1971).
[3] M. W. Chase Jr., NIST-JANAF Themochemical Tables, Fourth Edition, J. Phys. Chem. Ref. Data, Monograph 9, 1 (1998) (web: `http://webbook.nist.gov/cgi/cbook.cgi?ID=C7782447&Units=SI&Mask=2`). Zero-point energy correction taken from Huber, K. P. and Herzberg, G., in *Molecular Spectra and Molecular Structure: Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979), Vol. 4.
[4] K.P. Huber, G. Herzberg, Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules, Van Nostrand Reinhold Co. (1979).
[5] R. Zhu, F. Janetzko, Y. Zhang, A. C. T. van Duin, W. A. Goddard and D. R. Salahub, Theor. Chem. Account, **120** 479 (2008).

IV. O_2-adsorption energy on MgO_x, with functionals corrected by the experimental value of O_2 binding energy

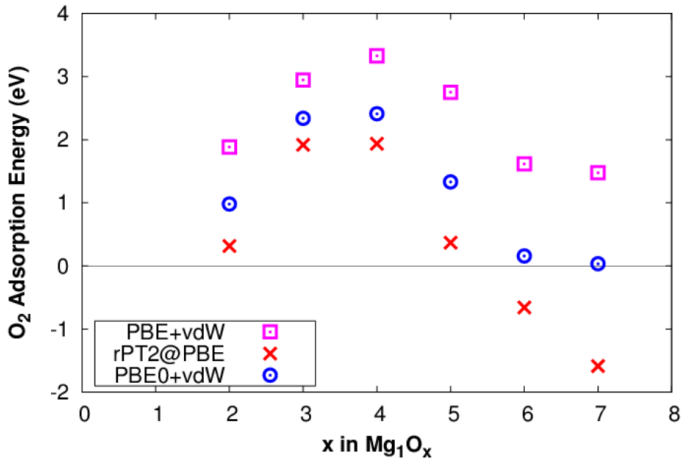


FIG. 2: O_2-adsorption energy of MgO_x cluster using different xc functionals, calculated with O_2 molecule total energy corrected for the experimental value of O_2 binding energy (-5.21 eV [1]).

[1] M. W. Chase Jr., NIST-JANAF Themochemical Tables, Fourth Edition, J. Phys. Chem. Ref. Data, Monograph 9, 1 (1998) (web: `http://webbook.nist.gov/cgi/cbook.cgi?ID=C7782447&Units=SI&Mask=2`). Zero-point energy correction taken from Huber, K. P. and Herzberg, G., in *Molecular Spectra and Molecular Structure: Constants of Diatomic Molecules* (Van Nostrand Reinhold, NewYork, 1979), Vol. 4.

V. O_2-adsorption energy on Mg_2O_x and Mg_3O_x clusters

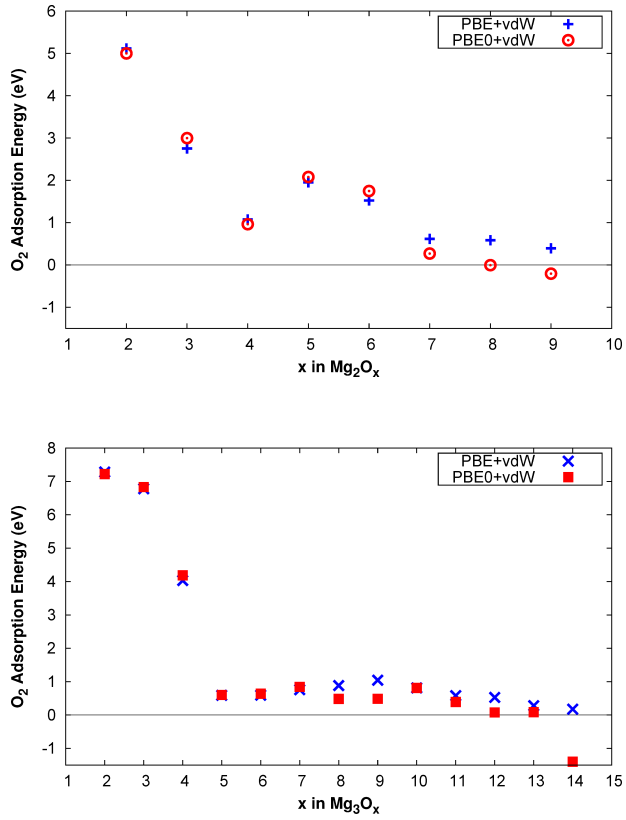
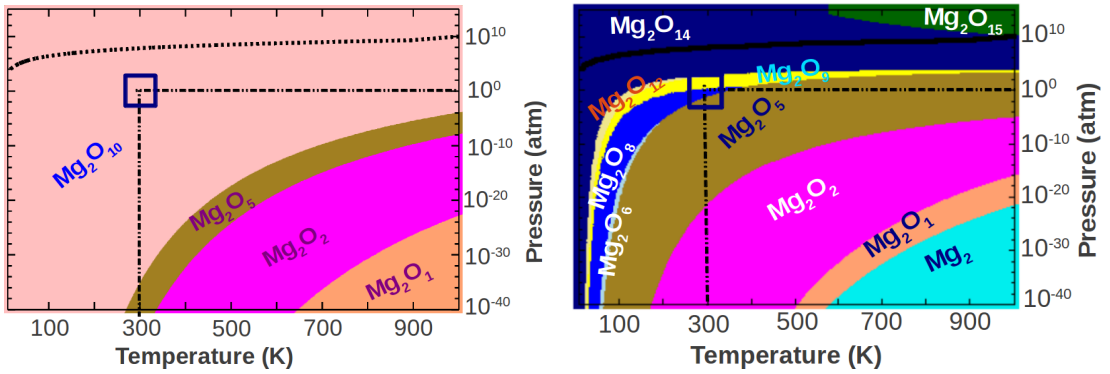


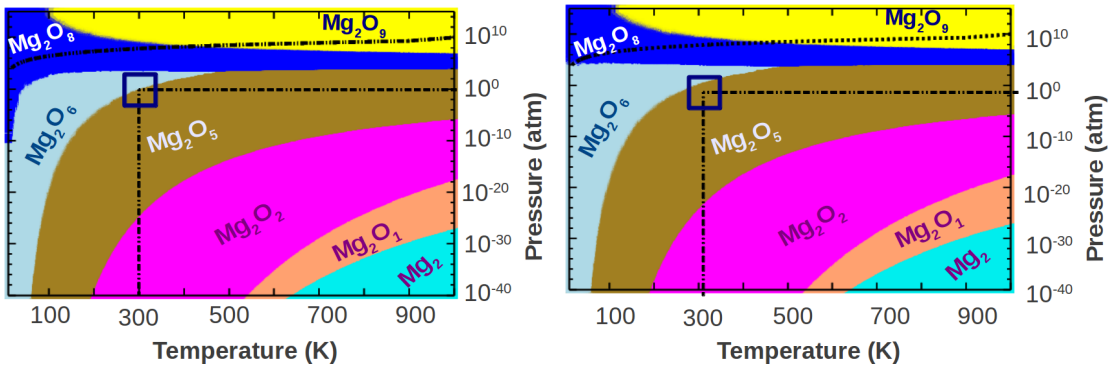
FIG. 3: O_2-adsorption energy of Mg_2O_x (top) and Mg_3O_x (bottom) cluster using PBE+vdW and PBE0+vdW functionals.

VI. Mg_2O_x phase diagrams with various functionals



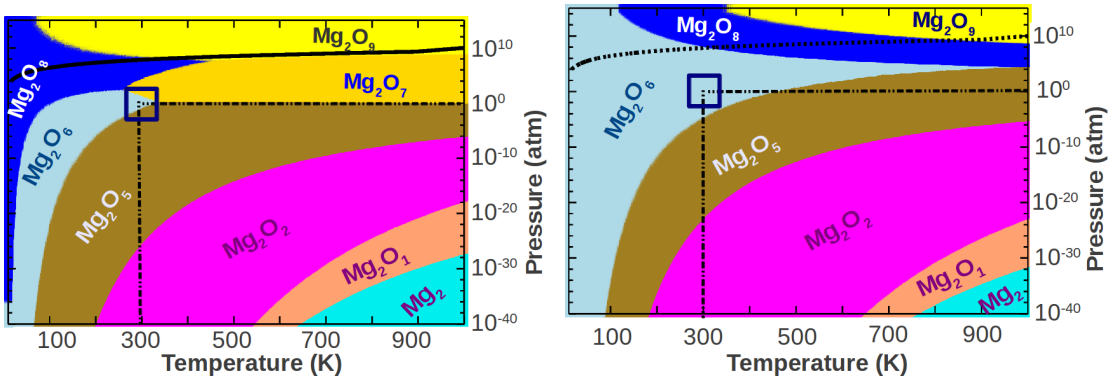
(a) ReaxFF

(b) PBE+vdW



(c) PBE0+vdW

(d) PBE0 (no vdW)

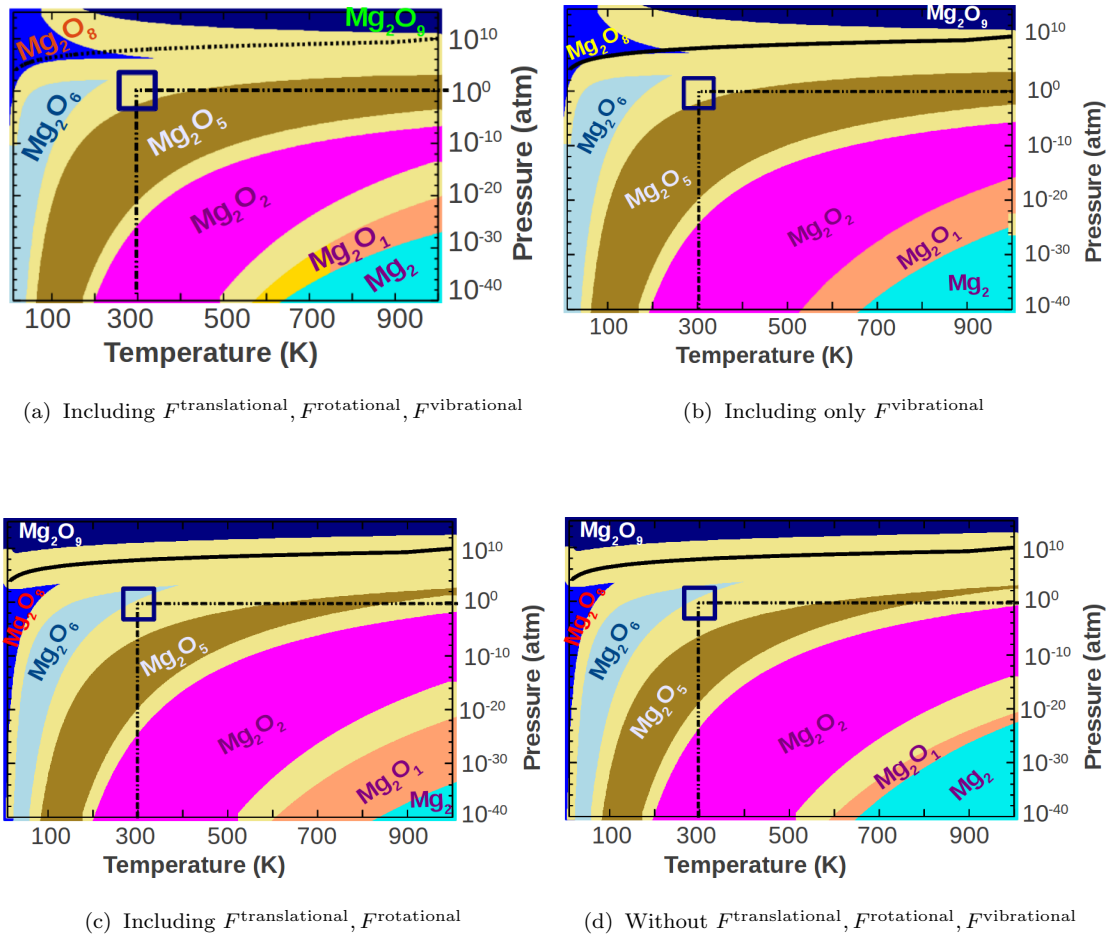


(e) PBE0+vdW, exp. O_2 binding energy

(f) rPT2@PBE

FIG. 4: (T, p_{O_2}) phase diagrams of Mg_2O_x at different levels of theory. The square encompasses the region around normal conditions ($T = 300$ K, $p_{O_2} = 1$ atm) and the dashed-dotted lines are guides for the eyes for identifying the point at normal conditions on the diagram.

VII. Effect of translational, rotational, vibrational contributions to the free energy on Mg_2O_x phase diagram



(a) Including $F^{\text{translational}}, F^{\text{rotational}}, F^{\text{vibrational}}$

(b) Including only $F^{\text{vibrational}}$

(c) Including $F^{\text{translational}}, F^{\text{rotational}}$

(d) Without $F^{\text{translational}}, F^{\text{rotational}}, F^{\text{vibrational}}$

FIG. 5: (T, p_{O_2}) phase diagrams of Mg_2O_x with different contributions in the free energy of the clusters. The geometries are optimized with PBE+vdW and the electronic energy are calculated using PBE0+vdW. The sand-colored unlabeled regions are regions where different compositions (at least the adjacent ones) coexist (free energy of the coexisting species within 3 k_BT, see text).

VIII. Effect of anharmonic contributions to configurational free energy

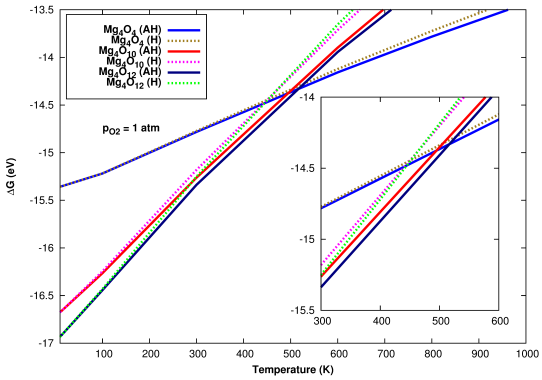
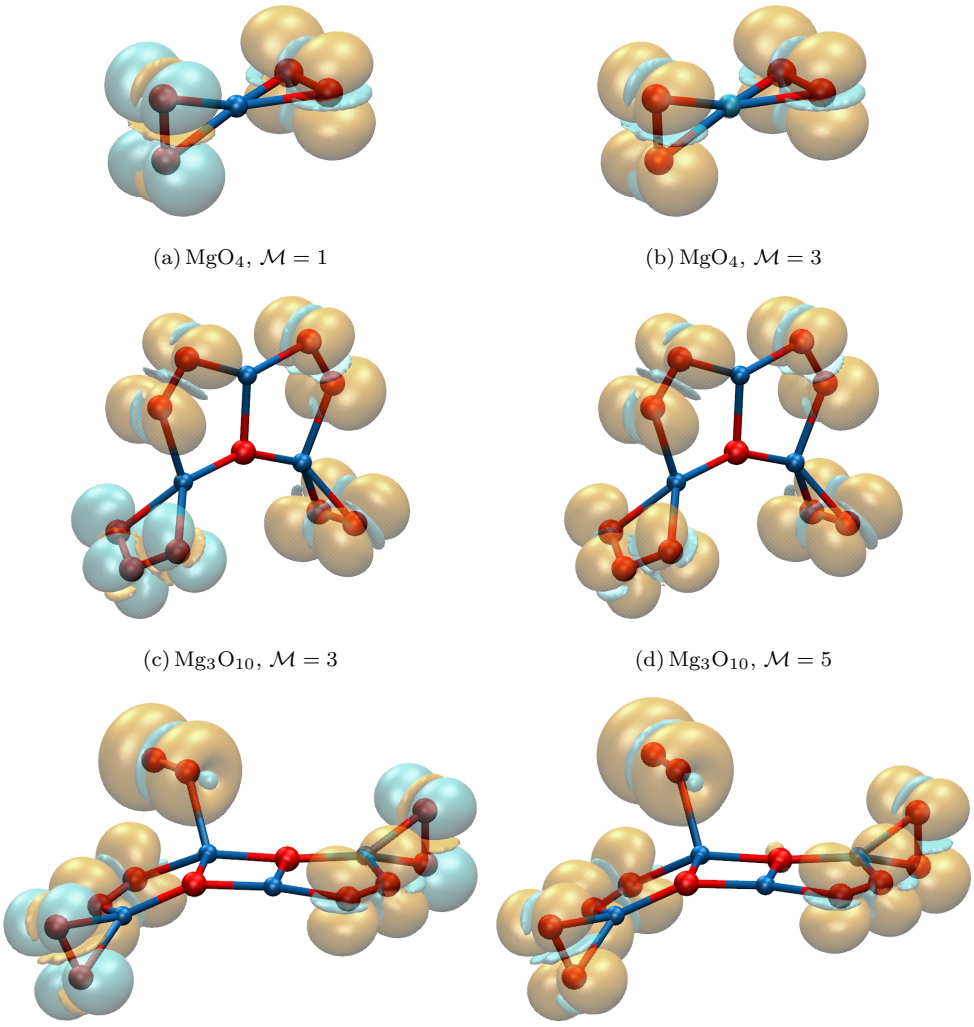


FIG. 6: Anharmonic (AH) and harmonic (H) contribution to the Free energy at $p_{O_2} = 1$ atm of Mg_4O_4, Mg_4O_{10}, Mg_4O_{12}.

IX. Examples of spin densities on non-stochiometric Mg_MO_x



(a) MgO_4, $\mathcal{M} = 1$ (b) MgO_4, $\mathcal{M} = 3$

(c) Mg_3O_{10}, $\mathcal{M} = 3$ (d) Mg_3O_{10}, $\mathcal{M} = 5$

(e) Mg_4O_{12}, $\mathcal{M} = 3$ (f) Mg_4O_{12}, $\mathcal{M} = 7$

FIG. 7: Spin density (difference between the majority-spin and the minority-spin electronic densities) for three representative clusters: MgO_4 ($\mathcal{M} = 1$ and $\mathcal{M} = 3$), Mg_3O_{10} ($\mathcal{M} = 3, 5$), and Mg_4O_{12} ($\mathcal{M} = 3, 7$). Red (blue) spheres represent O (Mg) nuclei. Orange (blue) lobes are spin-density isosurfaces at 0.01 (-0.01) $e^-/\text{Å}^3$ level. The two spin states of MgO_4 ($\mathcal{M} = 1$ and $\mathcal{M} = 3$) are nearly degenerate in energy (at the PBE0+vdW level). Mg_3O_{10} ($\mathcal{M} = 3$, isoenergetic with $\mathcal{M} = 1$) contains an O_3 moiety, which hosts one unpaired electron in the same fashion as the O_2 moieties. Having Mg_3O_{10} four O_2/O_3 moieties, its largest low-energy (0.02 eV higher than the singlet ground state) multiplicity is $\mathcal{M} = 5$. With its five O_2 moieties, Mg_4O_{12} shows at 0.05 eV from the singlet ground state an isomer with spin multiplicity $\mathcal{M} = 7$. The shown $\mathcal{M} = 3$ state is 0.03 eV from the singlet ground state. The off-plane O_2 moiety hosts two unpaired electrons.