Supplementary Material to: Structural search for stable Mg-Ca alloys accelerated with a neural network interatomic model

Wilfredo Ibarra-Hernández, $^{1,\,2}$ Samad Hajinazar, 3 Guillermo Avendaño-Franco, 1

Alejandro Bautista-Hernández,² Aleksey N. Kolmogorov,³ and Aldo H. Romero^{1, 2}

¹Department of Physics and Astronomy, West Virginia University, Morgantown, WV-26505-6315, USA

²Facultad de Ingeniería-BUAP, Apartado Postal J-39, Puebla, Pue. 72570, México

³Department of Physics, Applied Physics and Astronomy,

Binghamton University, State University of New York,

PO Box 6000, Binghamton, New York 13902-6000, USA

(Dated: October 4, 2018)

THE PYCHEMIA-MAISE INTERFACE

PyChemia and MAISE interact together for the purpose of the structural search. The roles of PyChemia are basically three: (i) generate structures for evaluation, (ii) coordinate the execution of MAISE and (iii) use global search algorithms that produce new structures based on the geometries and energies returned by MAISE. The inventory of structures that has been evaluated on successive generations is also a role taken by PyChemia[1] by using a central MongoDB database[2]. While the role of MAISE is to use the trained neural network that then is used to locally optimize the geometry of a given structure as well as determine the related energy (enthalpy), atomic forces and the stress tensor of the cell. The search is only performed in the primitive cells are considered for the next iteration.

For the creation of new structures, we have used two different methods, the first consists of a random structure generated by selecting different random lattice parameters, random angles and random atomic positions for a given composition. Cell parameters are chosen such that no acute angles are allowed. The second method consists in creating a smaller structural database by selecting a crystal structure from OQMD database[3, 4] which, by the time of this work, contains more than 297 thousand of structures for the stoichiometries of interest. For such purpose, before the structural search is launched the smaller database is built with the following criterion: (i) the considered structure from the database has the same number of chemical elements that the object of study (two in this case) and (ii) the total number of atoms is an integer multiple of the given stoichiometry. With this database we guarantee that we also consider very symmetric structures within the search, if they correspond to low energy configuration is dictated by the method used to evaluate forces and stresses. At each iteration, the population is analyzed and the lower energy structures are passed to the next iteration. High energy structures are replaced by new ones with 70% probability from the database and the remained 30% are randomly generated.

STATISTICS ON FOUND STRUCTURES

In Fig. S4, we show the percentage of structures for a given crystal family obtained by the two methods. For the MHM-DFT case, we have considered a total 847 structures where almost 57% belong to the triclinic crystal system. The fact that there is a large number of low symmetry structures can be due to the tendency of the material to disorder easily (see our discussion on SQS phases), which seems to be captured by this search method. Another reason for such number of low symmetry structures could be due to the used geometrical tolerance factors, which were selected quite strict to define the crystal symmetry (the tolerance in deviations from the specific crystal configuration was 0.001). As the search try to look for new phases, Mg-Ca prefers to have only few high symmetric ones and the search method can only create structures which are more energetic and less symmetric to keep the balance between exploration and structure diversity. We have changed the deviation tolerance to determine the space group to 0.01, and the only change that we observe was that some structures from triclinic phase have monoclinic phase. In general, high symmetry structure with the lowest formation energy has concentration of x=1/3 (Mg₂Ca) which is the right ground state concentration found experimentally. However, we cannot say that we have found the ground state since the C14 Laves phase, which contains only 6 atoms per cell, this result is consistent for both methodologies.

On the other hand, for the FF-NN methodology we have found up to 3292 different structures. From them, 49% of those are triclinic, a percentage that is close to the one found with the MHM-DFT methodology (58%).

Additionally, high symmetry structures (space group higher than 142) represent only 1.4% of the total structures under this methodology, a value also similar to one obtained from the MHM-DFT method.

SUPPORT TABLE

Ι	Table. S1 Mode frequencies with	Raman and IR response		2
---	---------------------------------	-----------------------	--	----------

SUPPORT FIGURES

Ι	Fig. S1 Neural network test results for equation of state (EOS)	3
II	Fig. S2 Neural network test results for defect formation energies	3
III	Fig. S3 Phonon dispersion curves obtained with DFT and NN	3
IV	Fig. S4 Space group histogram	4
V	Fig. S5 DFT convergence test results for Mg ₂ Ca Laves phases	4
VI	Fig. S6 SQS crystal structure	5
VII	Fig. S7 Partial density of electronic states	5
VIII	Fig. S8 Projected phonon density of states	5

TABLE S1: Frequencies in THz with either Raman or infrared (IR) response for the (meta)stable phases reported in this work. For the mS18-Mg₇Ca₂ we were unable to identify the specific frequencies that will have either Raman or IR response. Nevertheless, we can determine by group theory that the modes A_g and B_g will have Raman response while A_u and B_u modes will exhibit IR response.

	$\mathbf{oS36-Mg}_7\mathbf{Ca}_2$	$cF16-Mg_3Ca$	B2-MgCa	$hP8-MgCa_3$
	1.76, 1.82, 2.96, 3.19, 3.36			
	3.47, 3.62, 3.63, 3.67, 3.98	4.74		2.37, 2.68, 3.54, 4.30, 5.40
Raman	4.23, 4.27, 4.42, 4.46, 4.46			
response	4.69, 5.00, 5.11, 5.29, 5.57			
	5.99, 6.03, 6.56, 6.83, 6.86			
	7.08, 8.78			
	2.26, 3.44, 3.48, 3.57, 3.97	4.98, 6.76	5.93	4.27, 4.55, 4.74
IR	4.07, 4.17, 4.17, 4.23, 4.67			
response	4.78, 4.83, 4.99, 5.37, 5.92			
	6.17, 6.82, 6.85, 8.78			



FIG. S1: (a-c) Energy versus shortest interatomic distance for select Mg, Ca, and Mg-Ca structures evaluated with the developed NN model (red lines) and the DFT (black circles). The data sets were generated by rescaling the structures' volume and included in the NN training to ensure the correct description of configurations with short distances. As in our previous study[5], NN models trained without such data showed similarly good accuracy. (d) Energy versus c/a ratio in the volume-constrained hcp-Mg and hcp-Ca structures. The data illustrates the NN performance for structures subjected to anisotropic deformations.



FIG. S2: Vacancy and substitution formation energies calculated at the DFT and NN levels. The settings are discussed in our previous study.[5] From left to right, the data points correspond to the bcc, fcc, hcp, and sc structures.



FIG. S3: Phonon dispersion curves for the ground state structures of (a) magnesium, (b) C14-Mg₂Ca and (c) Calcium, obtained at DFT (black) and NN (red) level.



FIG. S4: Histogram of the percent of structures corresponding to each crystal system. For the case of MHM-VASP with crystal systems trigonal, hexagonal and cubic the percentage is 0, 0.24 and 0.35, respectively.



FIG. S5: k-point convergence tests for the Mg₂Ca Laves phases. In order to show the energy sensitivity to the choice of the unit cell shape, C15 was modeled with the conventional (24 atoms), primitive (6 atoms), and hexagonal (36 atoms) unit cells representing the same structure. The hexagonal unit cells for C14 (12 atoms), C36 (24 atoms), and C15 (36 atoms) had Γ -centered meshes specified in the figure for C14 with scaled down *c*-axis grid by a factor 2 for C36 and a factor of 3 for C15. The non-hexagonal unit cells were sampled with the Monkhorst-Pack sampling [6] with comparable densities; for example, the finest grids for the conventional and primitive unit cells of C15 were $12 \times 12 \times 12$ and $16 \times 16 \times 16$, respectively. The tetrahedron integration scheme was used in all the cases.



FIG. S6: SQS structure with 207 atoms in the cell. The motif used for this structure was $oS36-Mg_7Ca_2$. Orange and blue balls represent Mg and Ca atoms, respectively.



 $\label{eq:FIG.S7:Partial electronic density of states (a) oS36-Mg_7Ca_2, (b) mS18-Mg_7Ca_2, (c) cF16-Mg_3Ca, (d) B2-MgCa and (e) hP8-MgCa_3.$



FIG. S8: Projected phonon density of states for (a) oS36-Mg₇Ca₂, (b) mS18-Mg₇Ca₂, (c) cF16-Mg₃Ca, (d) B2-MgCa and (e) hP8-MgCa₃. The forces to extract the dynamical matrix were calculated with NN-model. Dashed lines from hP8-MgCa₃ corresponds to values at zero pressure.

- G. Avendao-Franco and A. H. Romero, Journal of Chemical Theory and Computation 12, 3416 (2016), pMID: 27232694, http://dx.doi.org/10.1021/acs.jctc.5b01157, URL http://dx.doi.org/10.1021/acs.jctc.5b01157.
- [2] Mongo web page, https://www.mongodb.com/es, accessed: 2017-08-13.
- [3] J. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, JOM 65, 1501 (2013), ISSN 1047-4838.
- [4] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Ruhl, and C. Wolverton, npj Computational Materials 65, 15010 (2015).
- [5] S. Hajinazar, J. Shao, and A. N. Kolmogorov, Phys. Rev. B 95, 014114 (2017), URL https://link.aps.org/doi/10.1103/ PhysRevB.95.014114.
- [6] H. J. Monkhorst and J. D. Pack, Phys. Rev. B 13, 5188 (1976), URL http://link.aps.org/doi/10.1103/PhysRevB.13.
 5188.