Prediction of Novel High Pressure H_2O-NaCl and Carbon Oxide Compounds with Symmetry-Driven Structure Search Algorithm

Rustin Domingos,[1] Kareemullah M. Shaik,[2] and Burkhard Militzer[2,3]

[1]*Department of Physics, University of California, 94720 Berkeley*
[2]*Department of Earth and Planetary Science, University of California, 94720 Berkeley*
[3]*Department of Astronomy, University of California, 94720 Berkeley*

Crystal structure prediction with theoretical methods is particularly challenging when unit cells with many atoms need to be considered. Here we employ a symmetry-driven structure search (SYDSS) method and combine it with density functional theory (DFT) to predict novel crystal structures at high pressure. We sample randomly from all 1,506 Wyckoff positions of the 230 space groups to generate a set of initial structures. During the subsequent structural relaxation with DFT, existing symmetries are preserved, but the symmetries and the space group may change as atoms move to more symmetric positions. By construction, our algorithm generates symmetric structures with high probability without excluding any configurations. This improves the search efficiency, especially for large cells with 20 atoms or more. We apply our SYDSS algorithm to identify stoichiometric $(H_2O)_n$-$(NaCl)_m$ and C_nO_m compounds at high pressure. We predict a novel H_2O-NaCl structure with *Pnma* symmetry to form at 3.4 Mbar, which is within the range of diamond anvil experiments. In addition, we predict a novel C_2O structure at 19.8 Mbar and C_4O structure at 44.0 Mbar with *Pbca* and *C2/m* symmetry respectively.

PACS numbers:

INTRODUCTION

Crystal structure prediction with theoretical methods is a challenging subject even for the simplest materials [1, 2]. For complex materials with a large number of atoms in the unit cells, N, finding the ground state crystal structure is particularly difficult. The dimensionality of the search space grows as $3N + 3$, while the number of local minima increases exponentially with the dimensionality, thus the effort required to find the global Gibbs free energy minimum increases exponentially with N. This problem is classified as NP-hard, and for large systems, searching all configurations is unfeasible. However, significant progress has been made with evolutionary algorithms [3–5], random search techniques [6, 7], and others methods including simulated annealing [8], minima hopping [9], and metadynamics [10]. In principle these methods do not require experimental input, however, efficiency may be improved if it is incorporated. For example, powder diffraction data has been used to restrict the search space to within a known space group [11]. In addition, knowledge of energetically preferred structural elements like molecules and functional groups may guide the structure generation process. One could, for example, place entire H_2O molecules or NaCl pairs instead of individual atoms. However, one must be aware that these geometries may not be preserved at megabar pressures and thus the implementation of such constraints may eliminate the most favorable structures. While no search method offers a rigorous path to finding the most stable structure, they have all lead to many novel low-enthalpy candidate structures and enriched our understanding of materials at high-pressure. Most importantly, a number of theoretical predictions have later been confirmed experimentally [2, 12–22]. Knowledge of the globally stable structure allows one to calculate physical properties under extreme conditions, where experimental results are not yet obtainable.

For large unit cells, crystal structure prediction has remained a challenge. When we applied the *ab initio* random structure search technique [6], with no symmetry constraints, to look for novel $FeSiO_3$ structures [23], we found 74% of the randomly generated 5-atom cells relaxed into symmetric structures, while the remaining ones had no (or P_1) symmetry. During the relaxation of 10-atom cells, the fraction of symmetric, non-P_1 structures decreased to 43%. For 15-, 20-, 25- and 40-atom cells, the fraction of non-P_1 structures dropped 0.6%, 0.9%, 0.01%, and 0.02%, respectively. This means more than 100 20-atom structures needed to be relaxed in order to generate one symmetric structure that had a chance of being the global enthalpy minimum. This argument adopts the common assumption that the most stable structure has at least one symmetry operation. The wealth of experimental data shows that most compounds crystallize into symmetric structures at low temperature. This tendency is expressed by Pauling's *rule of parsimony* [24] and supported by the energetics of symmetry calculations [25].

Symmetry and Structure Prediction

It has been recognized that symmetries are key to studying large clusters [26–28] and crystal structures [4, 7, 29, 30]. In Ref. [7], this point is addressed by choosing N_{op} specific symmetry operations. A subset of the atom positions are chosen randomly and the remaining images are generated according to symmetry. For the example of a mirror plane, the positions for half of the atoms are chosen randomly while the other half are placed on their mirror images. During the subsequent structural relaxation, a more symmetric structure may emerge. With the added symmetry operation, this structure can be derived from a supergroup of the original group with N_{op} op-

erations. The evolutionary algorithm in Refs. [4, 29] relies on the particle swarm optimization method to move from one generation to then next. The set of 230 space groups have been used to generate the initial set of structures by selecting Wyckoff positions within a given space group that are consistent with the chosen composition. Once a structure is generated from a particular space group, the algorithm introduces a penalty to prevent the generation of another structure from the same space group. When this method was applied to the structural optimization of TiO_2 with classical potentials [29] it was shown in that the implementation of symmetry constraints improved the efficiency of the search algorithm. With constraints, more low energy structures were generated and approximately half as many generations were needed to find the optimal structure. In Ref. [30] symmetry constraints are implemented by placing atoms on the most general Wyckoff position with the option of merging nearby atoms onto more symmetric Wyckoff positions, while allowing for symmetry breaking in subsequent generations. It was shown in [30] that initializing the evolutionary algorithm with symmetric structures improved efficiency while using classical potentials to determine ground state structures of $MgAl_2O_4$. In addition, implementation of symmetry constraints allowed the determination of the ground state of $Mg_{24}Al_{16}Si_{24}O_{96}$, a 160 atom unit cell structure that was not found previously without symmetry constraints. In our approach, we directly sample from the 230 space groups and all associated 1,506 Wyckoff positions. This allows us to include all Wyckoff positions consistently and select them with a high probability without excluding any structure in principle. All space groups and Wyckoff positions are treated with equal probability until we eliminate structures in which atoms are very close. The first water-salt structures that we generated with this method were reported in Ref. [31]. Independent of this work, a similar approach was developed in Ref. [32]. First a set of space groups is selected. For every space group, a list of all possible combinations of Wyckoff positions is assembled that are consistent with the given composition. For large systems, this list may become exceptionally large. For this reason, the size of this list was reduced by putting similar Wyckoff positions into groups. This made the algorithm more efficient but also changed the probability of how often certain combinations of Wyckoff positions are selected. Conversely, in our algorithm, we sample Wyckoff positions without generating such a list and have thus no need to restrict its size. In Ref. [32], 10 space groups were chosen when the initial generation of TiO_2 structures were derived for the subsequent evolutionary algorithm. Using classical potentials, it was shown that symmetry constraints increase the probability of finding low energy structures, but also the probability of generating high energy structures, resulting in an increased average energy overall.

High-pressure water-salt and carbon oxides

Our goal is to design an efficient method to predict the crystal structure of real materials at arbitrary pressures, without requiring experimental input. Here we developed a symmetry-driven structure search (SYDSS) technique to identify novel crystalline compounds at high pressure. We applied our SYDSS algorithm to search for $(H_2O)_n$-$(NaCl)_m$ and C_nO_m compounds at megabar pressures. While salt dissolves in water up to a maximum concentration, to our knowledge, no stoichiometric H_2O-NaCl compound has been found in nature, generated with laboratory experiments, or predicted theoretically. However, at high pressure, the properties of materials change and compounds that, while immiscible at ambient conditions, may form stoichiometric compounds [33]. In Ref. [34], a novel $LiCl{\cdot}6H_2O$ structure was shown to form at 2 GPa and it was suggested that other salt-ice compounds may exists at higher pressure. The discovery of novel high pressure compounds may improve our understanding of the interior structure and dynamics of ice giant planets [35]. If we assume, as an example, NaCl were available in sufficient quantities, a separate H_2O-NaCl layer [36] would form below the ice layer because of its higher density. The density contrast of the two layers would also introduce a convective barrier into the interior and potentially prolong the cooling process of an ice giant planet.

The properties of carbon and oxygen are of high interest in planetary science because together with hydrogen and nitrogen, they form the planetary ices H_2O, CH_4, and NH_3 that make up the bulk of the interiors of ice giant planets [35]. Depending on the formation conditions and the composition of the building materials, a variety of planets and different interior structures are expected to form [36–38]. Terrestrial planets like Venus have thick and hot atmospheres that are rich in CO_2. In the atmospheres of more massive exoplanets, we can expect to find carbon-oxygen compounds that are exposed to yet higher pressure. However, the properties of such compounds are not yet well characterized at extreme conditions. With density functional molecular dynamics simulations, Boates et al. [39] predicted CO_2 to exhibit a liquid-liquid phase transition at 0.5 Mbar. Leonhardi and Militzer [40] predicted a similar phase transition for CO to occur between 0.1 and 0.2 Mbar. In the simulations, CO was also observed to change phase from a molecular to a polymeric fluid. At yet higher pressures, CO was found to spontaneously freeze into an amorphous solid. Even though amorphous CO_2 structures have been generated with high-pressure laboratory experiments [41], one may expect that the amorphous CO structures seen in the simulations do not correspond to the thermodynamic ground state and that there is exists at least one ordered solid CO structure with a lower free energy. Here we thus use our SYDSS method to look for novel crystalline carbon-oxygen structures with a carbon-to-oxygen ratio of 1:1 and variety of other compositions.

In the original formulation, the *ab initio* random structure

search technique did not take advantage of crystal symmetries and worked in the space group P_1 [6]. Symmetric structures emerge, however, when atoms move onto Wyckoff positions during the relaxation. If one wants to start with, and maintain a certain set of crystal symmetries during the entire search process, a special handling of the Wyckoff positions is unavoidable as the following simple example of a mirror plane illustrates.

If an atom is placed exactly on the mirror plane then there exists only one instance of it, otherwise there are two. Switching continuously from one case to the other is difficult within the context of *ab initio* simulations because when one atom moves closer to the mirror plane, the distance between the atoms becomes small, repulsive forces become large and *ab initio* calculations with pseudopotentials typically do not converge. One could, of course, remove one of the atoms if the distance between the pair becomes too small but then the remaining structure would no longer be symmetric. This means, for structural relaxation algorithms that preserve the symmetry of the mirror plane, one needs to decide at the very beginning whether the atom is on or off the mirror plane. In both cases, the atom can still move in the subsequent relaxation and occupy a more symmetric position.

While for a single mirror plane, only two cases need to be considered, for a typical space group there exists a series of Wyckoff positions that all need to be treated separately. Thus we decided to treat then 230 space groups and associated 1,506 Wyckoff positions [42, 43] in a consistent fashion. This means even screw axis symmetries are included and atoms, that are far away from each other, have a higher chance of being placed on symmetry positions. It is our goal to construct an algorithm that does not exclude any structure but drastically increases the probability that symmetric structures are generated successfully. To prevent convergence issues in *ab initio* calculations, we exclude, however, structures where atoms are unphysically close. For this project, we conservatively chose the following minimum distances between different Na, Cl, H, C and O species: r_{NaNa} = 1.4, r_{NaCl} = 1.2, r_{ClCl} = 1.4, r_{NaH} = 0.8 r_{ClH} = 0.8, r_{HH} = 0.7, r_{NaO} = 1.2, r_{ClO} = 1.2, r_{OH}= 0.8, r_{CO}= 1.1, r_{CC}= 1.2, and r_{OO} = 1.2 Å.

METHOD

Our SYDSS algorithm repeatedly steps through all 230 space groups until a user-defined number of structures, N_S, have been generated successfully. For a chosen space group, it selects lattice parameters and angles at random, applies any constraints of the space group, and scales the unit cell so that its volume matches a chosen target volume. Then it builds a list of all atoms to be placed in the cell (chemical composition times the number of formula units, N_{FU}). As long as this list is not yet exhausted, our code loops over all Wyckoff positions of the selected space group. Then it loops over all atom types that have a sufficient number of atoms remaining to fill all instances of the selected Wyckoff position. It chooses random values for all free parameters of this position, generates the coordinates, and checks whether the atoms satisfy all minimum distance criteria [44]. If they do, our algorithm continues to place the remaining atoms. If it fails to meet the distance criteria at any point in this process, it discards the current configuration and continues with the next space group until all N_S initial structures have been generated.

In our current implementation, the SYDSS algorithm has only a minimal set of adjustable parameters: the atoms in the cell, the set of minimum distances, and the target volume. The initial unit cell angles are chosen between 40 and 140 degrees [6] because the primitive cells of most structures can be represented in this way. This range could be broadened or one could sample from a smooth prior distribution that includes all angles. These choices, in particular the distance criteria, imply that not all space groups occur in the list of generated structures with equal probability. Figure 1 shows the probability distribution of space groups in the set of initial structures of H_2O-NaCl structures. Many space groups occur with very low or even zero probability because an insufficient number of atoms remain to fill all instances of a chosen Wyckoff position, or the inability to do so and satisfy all minimum distance criteria. In particular, cubic systems (space groups 195-230) occur rarely among our generated H_2O-NaCl structures but they occur frequently when we apply our algorithm to monatomic metals. Because of its lack of symmetry constraints, space group P_1 is still among the space groups that are generated most often but its total weight is now closer to 10% compared to 100% in [6]. If needed, additional biases could be introduced into the current implementation of our SYDSS algorithm in order to reduce the P_1 probability further.

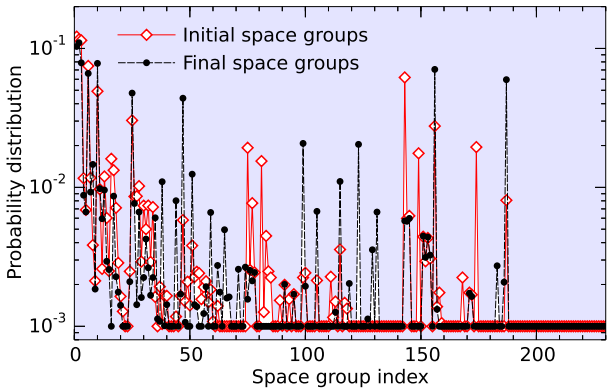

FIG. 1: Distribution of initial and final space groups among the 57,347 H_2O-NaCl structures that we generated with our SYDSS algorithm. Both distributions are not identical because the space group may change during the structural relaxation with DFT forces. Some space groups have a low or zero occurrence probability. We set their probabilities to 10^{-3} to include them in this graph.

Starting with these initial structures, structural relaxation at constant pressure was carried out in the framework of density functional theory, using the Perdew-Burke-Enzerhof functional [45] and the projector augmented wave method [46] as

implemented in the Vienna *ab initio* simulation package [47]. A basis-set cutoff energy of 980 eV used for the plane-wave expansion of the wave functions. For the first round of relaxation, we used k-point grids of 4×4×4 for cells with less than 15 atoms and 2×2×2 for cells with greater than 15 atoms. The best of these structures were then re-relaxed using higher density grids (6×6×6 to 12×12×12) to ensure accurate enthalpies. This method allowed us to search the structures more efficiently by removing unlikely candidates early.

During relaxations, the symmetry of the initial space group was preserved. This still allowed structures to attain higher symmetries of a supergroup during the relaxation if the atoms move to more symmetric positions while maintaining the symmetry operations of the original space group. This means the space groups of the initial and the relaxed structures may differ. Such transitions are illustrated in Fig. 2. Many transitions occurred between space groups in the same crystal system but we also noticed transitions from monoclinic (space groups 3-15) to orthorhombic (16-74), from orthorhombic to tetragonal (75-142), and from trigonal (143-167) to hexagonal (168-194) systems. In a few instances, a smaller primitive unit cell emerged during the relaxation. Such a transition may plot below the diagonal in Fig. 2 because the smaller unit cell may have a lower space group number.

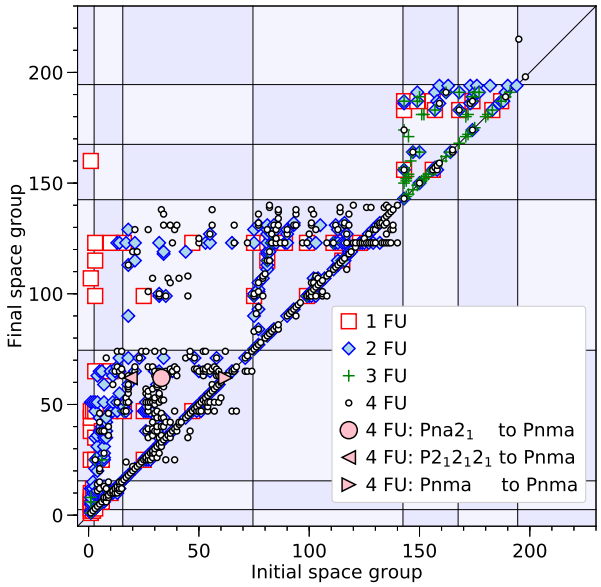

FIG. 2: In this transition diagram, final versus initial space groups are plotted for the relaxation of 1 to 4 formula unit structures. Transitions to higher space groups are much more frequent because the structural relaxation often increases the symmetry. The large circle indicates one possible pathway to our *Pnma* structure. The lines separate the 7 crystal systems.

RESULTS

Water-Salt Structure Search: $(H_2O)_n$-$(NaCl)_m$

We generated and relaxed over 55,000 structures with compositions $(H_2O)_n$-$(NaCl)_n$ having between 1 and 4 formula units at pressures between 1 and 10 Mbar. We also explored additional water-salt mixing ratios by relaxing over 11,000 $(H_2O)_n$-$(NaCl)_m$ structures with n:m=4:1, 3:1, 2:1, 3:2, 1:2, and 1:3. However, we were not able to find any thermodynamically stable structures with $n \neq m$ and we will thus focus the following discussion on structures with equal water-salt ratios where our structure search was more successful.

Enthalpies of the computed $(H_2O)_n$-$(NaCl)_n$ structures were then compared with the enthalpy of the H_2O and NaCl endmembers. NaCl endmember enthalpies were calculated using the B2 structure which is stable in this pressure range [48]. H_2O endmember enthalpies were calculated using the high pressure ice phases predicted at each pressure. At 1-2 Mbar, enthalpies were calculated using the ice X structure [49]. For pressures of 3-7 Mbar, enthalpies were calculated using the *Pbcm* structure [50]. At 8 Mbar, the *Pbca* structure [51] was used. At 9-10 Mbar, the $P3_121$ structure [52] was assumed. To further test our SYDSS method, we also applied it to pure water ice. We relaxed 2,000 H_2O structures at 9 Mbar and reproduced the $P3_121$ structure from Ref. [52].

After comparison with endmember data, three H_2O-NaCl structures were found to have enthalpies lower than that of the combined endmembers, suggesting a novel H_2O-NaCl structure would form at high pressure. The enthalpy comparisons of the three best structures with that of the endmembers is given in Fig. 3. The $P\bar{1}$ structure was found by relaxing a structure with one formula unit of H_2O-NaCl, the $P2_1$ structure was found from two formula unit structures, and the *Pnma* structure was found from four formula unit structures. This enthalpy data predicts a novel *Pnma* symmetric H_2O-NaCl structure forming at 3.4 Mbar, which is within the pressure range of diamond anvil cell experiments [53].

Out of 7,909 four-formula-unit H_2O-NaCl structures that were successfully relaxed, 185 (∼2%) relaxed into the *Pnma* space group. Of these 185 *Pnma* structures, 74 relaxed from structures with $P2_12_12_1$ symmetry, 35 from *Cc* symmetric structures, and 76 initially started from *Pnma* symmetric structures. The fact that this structure was never generated from a nonsymmetric initial structure (space group P_1) and the rate of occurrence illustrates the advantages of implementing symmetry constraints in our algorithm.

The parameters of our novel orthorhombic H_2O-NaCl structure are given in Tab. I and two pictures are shown in Fig. 4. We verified that this structure is dynamically stable by performing phonons calculations with the Phonopy code [54] using 1x1x2, 1x2x1, and 2x1x1 supercells. The structure can be explained best by analyzing the layering parallel to the a-b planes. Layers with Cl^- ions alternate with layers of Na^+ and O^{2-} ions. The Cl^- ion always occupy the same position

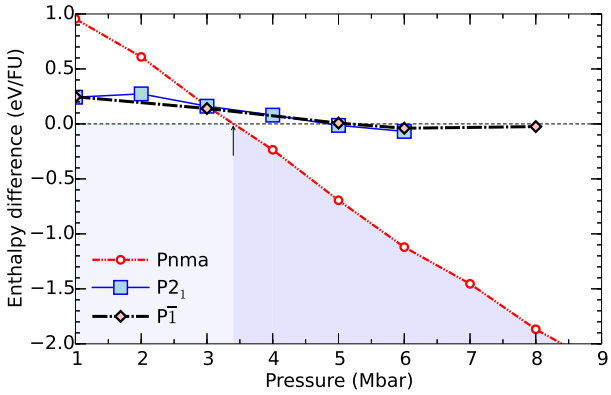

FIG. 3: The difference in enthalpy, $H_{\mathrm{H_2ONaCl}} - H_{\mathrm{H_2O}} - H_{\mathrm{NaCl}}$, per formula unit as function of pressure. The arrow marks the pressure of 3.4 Mbar where the H_2O-NaCl structure with *Pnma* symmetry is predicted to form from H_2O and NaCl. The $P2_1$ and $P\bar{1}$ structures were also shown to have lower enthalpies than the endmembers at 5 Mbar but the *Pnma* structure is energetically favored.

TABLE I: Parameters of the orthorhombic H_2O-NaCl structure with *Pnma* symmetry at 4 Mbar. The lattice parameters are a=3.942, b=3.849, and c=5.187 Å.

Atom	Wyckoff	x	y	z
Na	b	0	0	1/2
Cl	c	−0.476	1/4	−0.370
H	d	0.355	−0.437	−0.310
O	c	0.278	1/4	0.267

in every layer. The Cl^--Cl^- distances are thus smaller than the separation between other ion pairs of the same type. From layer to layer, the Na^+ and O^{2-} ions alternate between two positions, leading to unit cell with 20 atoms. The typical geometry of a H_2O molecule is well preserved. The H_2O dipole moments lie in the a-b planes and are arranged in clockwise or anticlockwise direction around each column of Cl^- ion. Overall the charges are reasonably well balanced in this structure.

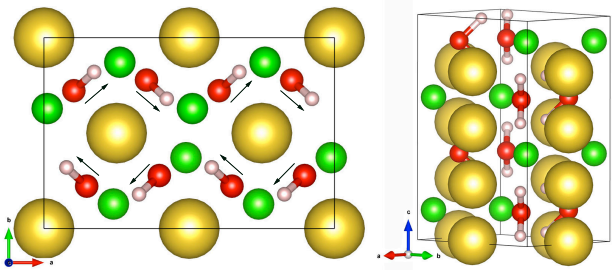

FIG. 4: Novel orthorhombic NaCl-H_2O crystal structure with *Pnma* symmetry. With decreasing size, the spheres denote the positions of Cl, Na, O and H atoms. The structure has 20 atoms per unit cell but has been doubled in c direction in the right image.

C

We generated and relaxed over 700,000 C_n-O_m structures with up to 52 atoms per unit cells with ratios from n:m ranging from 1:7 to 6:1 at pressures between 1 and 50 Mbar. Enthalpies of the resulting structures were compared to the C and O endmembers, which were calculated using the stable carbon phases of diamond for 1-10 Mbar, BC8 for 15-25 Mbar, and SC1 for 30-50 Mbar [55]. For oxygen endmembers ζ-C2/m oxygen for 1-15 Mbar [56, 57] and the *Cmcm* oxygen structure above 20 Mbar [58] were used.

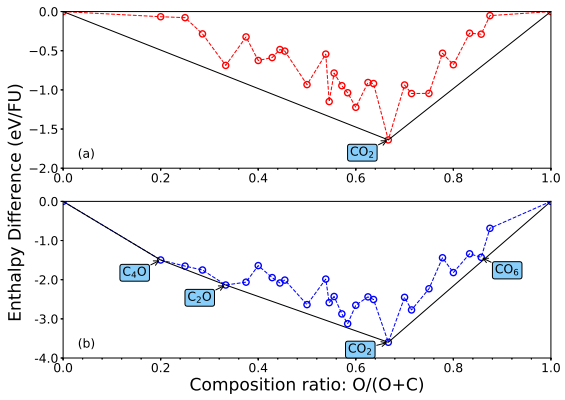

FIG. 5: Enthalpy difference per formula unit as a function of composition is plotted. (a) Representative enthalpy calculations for varying compositions at 7 Mbar. (b) Enthalpy calculations at 50 Mbar suggest C_4O and C_2O structures would be favorable in systems that contain more carbon than CO_2.

In Fig. 5, we plot the enthalpy difference per atom between various C_nO_m compounds that of the endmember phases, $\Delta H = H_{\mathrm{C_nO_m}} - (nH_{\mathrm{C}} + mH_{\mathrm{O}})$. At every pressure under consideration, the most stable CO_2 structures, that we obtained, reproduced previous results [59]. The convex hull at 7 Mbar in Fig. 5a shows that no stable structures are expected to exist besides CO_2 and the two endmembers. However, convex hull diagram at 50 Mbar in Fig. 5b revealed the existence of two new stable carbon-rich structures with C_4O and C_2O compositions. Interestingly, no stable CO structures were found over the entire pressure range. All CO structures, that we generated, were found to have a higher enthalpy than a combination of carbon and CO_2. Also none of our oxygen-rich compounds were found to be stable. Only one structure with a C:O = 1:6 composition came close to matching the combined enthalpies of pure oxygen and CO_2 but was not found to be stable in the pressure range up to 50 Mbar.

The C_4O structure is monoclinic and has *C2/m* symmetry. It was found from relaxing structures with 2 formula units (10 atoms). The image in Fig. 6 reveals a layered structure where thin oxygen planes alternate with thick carbon layers. The oxygen atoms form a 2D hexagonal lattice in planes spanned by the crystal lattice vectors b and c. The carbon atoms are arranged on four, tightly stacked hexagonal layers in between.

The C_2O structure can also be viewed as a layered structure but the bonding is more complex and three dimensional. The structure is orthorhombic and has *Pbca* symmetry. In Fig. 7, the unit cell with 8 formula units (24 atoms) has been double in b direction to illustrate the layers and 3D bonding. The

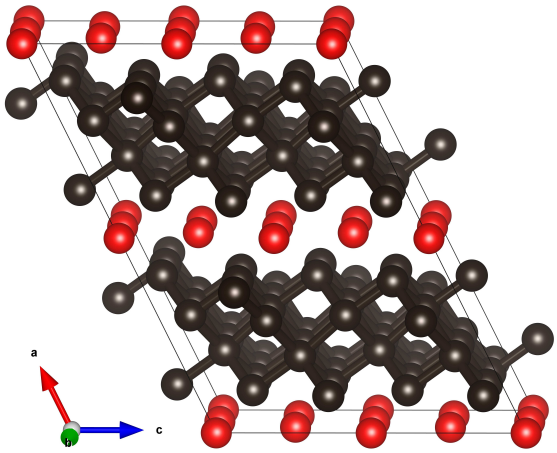

FIG. 6: Moniclinic C_4O crystal structure with *C2/m* symmetry at 45 Mbar. The unit cell with 10 atoms as been doubled along every lattice vector to better illustrate the C and O layers in the structure. The C and O atoms are shown in dark and light color, respectively.

shortest bonds occur between the C and O atoms in the layers but C-O bond distances vary considerably between 1.11 and 1.30 Å at 25 Mbar. The C-O layers are connected by C-C bonds that are all between 1.17 and 1.18 Å long. Again, we verified the C_2O and C_4O structures structure were dynamically stable by performing phonons calculations with the Phonopy code [54] using 2x2x2 supercells.

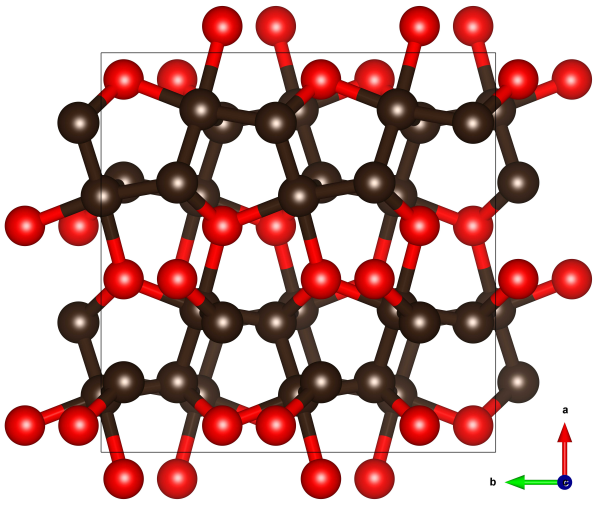

FIG. 7: Orthorhombic C_2O crystal structure with *Pbca* symmetry at 25 Mbar. The unit cell with 24 atoms as been doubled in *b* direction. The C and O atoms are shown in dark and light color, respectively.

We performed enthalpy calculations of these new C_4O and C_2O structures in order to determine the pressure at which they are favored over a decomposition into pure carbon and CO_2. In Fig. 8, we choose to plot the resulting enthalpy difference with respect to a mixture of pure carbon and C_2O because this allows us to illustrate the C_4O and the C_2O formulation pressures in a single diagram. We predict the C_2O structure to form at 19.8 Mbar while the C_4O structure becomes stable at 44.0 Mbar. The parameters of both structure given in tables II and III. The formation pressures of both structures are considerably larger than those that are typically reached with diamond anvil cell experiments. This is not unexpected because the diamond anvils would otherwise have reacted with the samples in any experiment that contained sufficient amounts of free oxygen. However, such pressures are accessible with dynamic compression techniques that use ramp waves to compress the sample at lower temperature than with standard shock wave experiments [60].

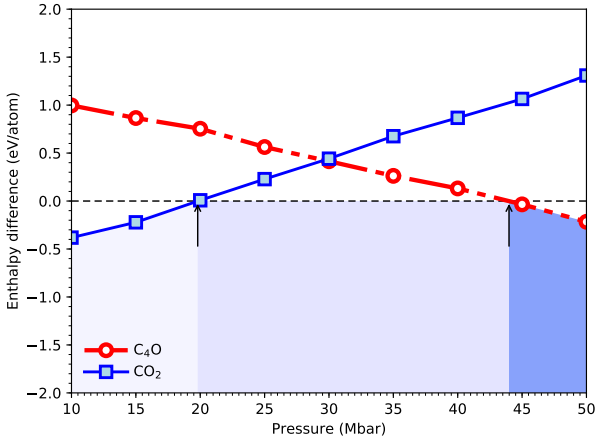

FIG. 8: The difference in enthalpy, $H_{C_nO_m} - \left[(n-2m) \times H_C + m \times H_{C_2O}\right]$, per atom as function of pressure. The first arrow marks the pressure of 19.8 Mbar where the C_2O structure with *Pbca* symmetry is predicted to form. The second arrow marks the pressure 44.0 Mbar where the C_4O structure with *C2/m* symmetry is predicted to form.

TABLE II: Parameters of the orthorhombic C_2O structure with *Pbca* symmetry at 25 Mbar. The lattice parameters are a=5.845, b=2.890, and c=2.899 Å.

Atom	Wyckoff	x	y	z
C	c	0.848	0.497	0.850
C	c	0.325	0.115	0.531
O	c	0.567	0.386	0.757

TABLE III: Parameters of the monoclinic C_4O structure with *C2/m* symmetry at 45 Mbar. The lattice parameters are a=2.960, b=1.434, c=3.916 Å and β = 106.73°.

Atom	Wyckoff	x	y	z
C	i	0.347	0.000	0.220
C	i	0.893	0.000	0.605
O	b	0.000	0.500	0.000

CONCLUSION

Our SYDSS algorithm provides a systematic and consistent way to generate symmetric candidate structures for relaxation with DFT forces with the goal of predicting novel crystal structures at high pressure. While no structure is excluded in principle, symmetric structures are generated with high probability. This significantly improves the efficiency of our structure search algorithm for large unit cells with 20 atoms or more, if one adopts the common view that ground state crystal structures are symmetric.

We applied our SYDSS technique to search for novel stoichiomtric H_2O-NaCl compounds at high pressure because we assumed large unit cells would be needed to accommodate atoms from both materials in an optimal way. Indeed, our best structure has a comparatively large primitive unit cell of 20 atoms. However, as with any random search method, there is no guarantee there does not exist yet another H_2O-NaCl structure with lower enthalpy unless our prediction is confirmed with experiments. The predicted formation pressure of 3.4 Mbar is well within the reach of diamond anvil cell experiments [53]. If indeed a yet more stable H_2O-NaCl compounds exists, x-ray diffraction measurements should reveal such a structure.

When we applied our SYDSS method to search for novel carbon-oxygen compounds at megabar pressures, we identified two novel carbon-rich but no oxygen-rich structures. At 19.8 Mbar, we predict an orthorhombic C_2O structure to form from dense carbon and CO_2. At 44.0 Mbar, a novel monoclinic C_4O structure is expect to become thermodynamically stable. Both transition pressures are beyond the reach of static high pressure experiments but can in principle be generated with dynamic compression techniques.

Also, we cannot completely rule out the existence of unknown low-enthalpy structures of the H_2O, NaCl, carbon, and oxygen endmembers even though one has looked for such structures carefully with DFT methods carefully already. If a novel H_2O, NaCl, carbon, or oxygen structure existed, the formation pressures of the predicted novel compounds would be shifted to higher values than we have predict here. However, in a diamond cell or ramp compression experiment one would see such novel endmember structures. In either case, new compounds or endmember structures are exprected to produced when H_2O-NaCl and C-O mixtures are exposed to pressures of 3.4 and 19.8 Mbar, respectively.

The authors acknowledge support from the U.S. National Science Foundation (grant 1412646), the U.S. Department of Energy (grant DE-SC0010517), and University of California's lab fee program as well as encouraging discussions with R. Caracas and other participants of the program "Dynamics and Evolution of Earth-like Planets" at the Kavli Institute for Theoretical Physics. In part, this work used the National Energy Research Scientific Computing Center and the Extreme Science and Engineering Discovery Environment. K. Driver provided comments on this manuscript.

[1] J. Maddox. *Nature*, 335:201, 1988.
[2] R. J. Needs and C. J. Pickard. *APL Mater.*, 4:053210, 2016.
[3] A. R. Oganov and Colin W. Glass. *J. Chem. Phys.*, 124:244704, 2006.
[4] Y. Wang, J. Lv, L. Zhu, and Y. Ma. *Phys. Rev. B*, 82:094116, 2010.
[5] D. C. Lonie and E. Zurek. *Comp. Phys. Comm.*, 182:372, 2011.
[6] C. J. Pickard and R. J. Needs. *Phys. Rev. Lett.*, 97:045504, 2006.
[7] C. J. Pickard and R. J. Needs. *J. Phys. Condens. Matter*, 23:053201, 2011.
[8] S. M. Woodley and R. Catlow. *Nature Mat.*, 7:937, 2008.
[9] S. Goedecker. *J. Chem. Phys*, 120:9911–9917, 2004.
[10] A. Laio and M. Parrinello. *Proc Natl Acad Sci USA*, 99:12562, 2002.
[11] B. Meredig and C. Wolverton. *Nature Mat.*, 12:123, 2013.
[12] A. R. Oganov, C. W. Glass, and S. Ono. *Earth and Planetary Science Letters*, 241:95 – 103, 2006.
[13] S. Ono, T. Kikegawa, and Y. Ohishi. *Am. Mineral*, 92:1246, 2007.
[14] Y. Ma, M. Eremets, R. Oganov A, Y. Xie, I. Trojan, S. Medvedev, A. O. Lyakhov, M. Valle, and V. Prakapenka. *Nature*, 458:182, 2008.
[15] M. Marqués, M. I. McMahon, E. Gregoryanz, M. Hanfland, C. L. Guillaume, C. J. Pickard, G. J. Ackland, and R. J. Nelmes. *Phys. Rev. Lett.*, 106:095502, 2011.
[16] Y. Ma, M. Eremets, A. R. Oganov, Y. Xie, I. Trojan, S. Medvedev, A. O. Lyakhov, M. Valle, and V. Prakapenka. *Nature*, 458:182, 2009.
[17] J. Lv, Y. Wang, L. Zhu, and Y. Ma. *Phys. Rev. Lett.*, 106:015503, 2011.
[18] B. Monserrat, R. J. Needs, E. Gregoryanz, and C. J. Pickard. *Phys. Rev. B*, 94:134101, 2016.
[19] C. J. Pickard, A. Salamat, M. J. Bojdys, R. J. Needs, and P. F. McMillan. *Phys. Rev. B*, 94:094104, 2016.
[20] A. Dewaele, N. Worth, C. J. Pickard, R. J. Needs, S. Pascarelli, O. Mathon, M. Mezouar, and T. Irifune. *Nature Chem.*, 8:784, 2016.
[21] V. V. Struzhkin, D. Y. Kim, E. Stavrou, T. Muramatsu, H. K. Mao, C. J. Pickard, R. J. Needs, V. B. Prakapenka, and A. F. Goncharov. *Nature Comm.*, 7:12267, 2016.
[22] A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, et al. *Acta Crystallographica Section B*, 72:439, 2016.
[23] L. Zhang, Y. Meng, W. Yang, L. Wang, W. L. Mao, Q.-S. Zeng, J. S. Jeong, A. J. Wagner, K. A. Mkhoyan, W. Liu, R. Xu, and H. k. Mao. *Science*, 344:877, 2014.
[24] L. Pauling *J. Am. Chem. Soc.*, 51:1010, 1929.
[25] Wales D J. *Chem. Phys. Lett.*, 285:330, 1998.
[26] S. E. Wheeler, P. v. R. Schleyer, and H. F. Schaefer. *J. Chem. Phys.*, 126:104104, 2007.
[27] J. Lv, Y. Wang, L. Zhu, and Y. Ma. *J. Chem. Phys.*, 137:084104, 2012.
[28] M. T. Oakley, R. L. Johnston, and D. J. Wales. *Phys. Chem. Chem. Phys.*, 15:3965, 2013.
[29] Y. Wang, J. Lv, L. Zhu, and Y. Ma. *Comp. Phys. Comm.*, 183:10, 2012.
[30] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, and Q. Zhu. *Comp. Phys. Comm.*, 184:1172, 2013.
[31] R. Domingos, K. M. Shaik and B. Militzer. *AGU Fall Meeting 2016*, Abstract MR31A-2669, 2016.

[32] P. Avery, and E. Zurek *Comp. Phys. Comm.*, 213:208, 2017.

[33] J. Feng, R. G. Hennig, N. W. Ashcroft, and R. Hoffmann. *Nature*, 451:445, 2008.

[34] S. Klotz, L. E. Bove, T. Strässle, T. C. Hansen, and A. M. Saitta. *Nature Mat.*, 8:405, 2009.

[35] H. F. Wilson, M. L. Wong, and B. Militzer. *Phys. Rev. Lett.*, 110:151102, 2013.

[36] H.F. Wilson and B. Militzer. *Astrophys. J.*, 793:34, 2014.

[37] J. C. Bond, D. P. O'Brien, and D. S. Lauretta. *Astrophys. J.*, 715:1050, 2010

[38] N. Madhusudhan, K. K. M. Lee, and O. Mousis. *Astrophys. J. Lett.*, 759, 2012

[39] B. Boates, S. Hamel, E. Schwegler, and S. A. Bonev. *J. Chem. Phys.*, 134(6):064504, 2011

[40] T. Leonhardi, and B. Militzer. *J. High Energy Density Physics*, 22:41, 2017

[41] M. Santoro, F. A. Gorelli, R. Bini, G. Ruocco, S. Scandolo, and W. A. Crichton. *Nature*, 441:857-860, 2006

[42] T. Hahn (ed.),*International Tables for Crystallography*, Vol. A: Space-group Symmetry, 5th edn. (Springer, 2006).

[43] Bilbao crystallographic server, http://www.cryst.ehu.es.

[44] Eqs. (4-7) in B. Militzer, *J. High Energy Density Physics* 21:8, 2016.

[45] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 77:3865, 1996.

[46] G. Kresse and J. Furthmüller. *Phys. Rev. B*, 54:11169, 1996.

[47] G. Kresse and D. Joubert. *Phys. Rev. B*, 59:1758, 1999.

[48] W. A. Bassett, T. Takahashi, H.-k. Mao, and J. S. Weaver. *J. Appl. Phys.*, 39:319–325, 1968.

[49] A. Polian and M. Grimsditch. *Phys. Rev. Lett.*, 52:1312–1314, 1984.

[50] M. Benoit, M. Bernasconi, P. Focher, and M. Parrinello. *Phys. Rev. Lett.*, 76:2934–2936, 1996.

[51] B. Militzer and H. F. Wilson. *Phys. Rev. Lett.*, 105:195701, 2010.

[52] C. J. Pickard, M. Martinez-Canales, and R. J. Needs. *Phys. Rev. Lett.*, 110:245701, 2013.

[53] S. S. Lobanov, V. B. Prakapenka, C. Prescher, Z. Konopkova, H.-P. Liermann, K. L. Crispin, C. Zhang, and A. F. Goncharov. *Journal of Applied Physics*, 118:035905, 2015.

[54] A. Togo and I. Tanaka. *Scr. Mater.*, 108:1–5, 2015

[55] L. X. Benedict, K. P. Driver, S. Hamel, B. Militzer, T. Qi, A. A. Correa, A. Saul and E. Schwegler. *Phys. Rev. B*, 89:224109, 2014.

[56] Y. Akahama, H. Kawamura, D. Häusermann, M. Hanfland, O. Shimomura. *Phys. Rev. Lett.*, 74:4690, 1995.

[57] Y. Ma, A. R. Oganov, and C. W. Glass. *Phys. Rev. B*, 76:064101, 2007.

[58] J. Sun, M. Martinez-Canales, D. D. Klug, C. J. Pickard, and R. J. Needs. *Phys. Rev. Lett.*, 108:045503, 2012.

[59] C. Lu, M. Miao, and Y. Ma. *Journal of the American Chemical Society*, 135:14167, 2013.

[60] R. F. Smith, J. H. Eggert, R. Jeanloz, T. S. Duffy, D. G. Braun, J. R. Patterson, R. E. Rudd, J. Biener, A. E. Lazicki, A. V. Hamza, J. Wang, T. Braun, L. X. Benedict, P. M. Celliers and G. W. Collins. *Nature*, 511:330-333, 2014.