



**HAL**  
open science

# Structurally Constrained Evolutionary Algorithm for the Discovery and Design of Metastable Phases

Busheng Wang, Katerina P. Hilleke, Samad Hajinazar, Gilles Frapper, Eva Zurek

► **To cite this version:**

Busheng Wang, Katerina P. Hilleke, Samad Hajinazar, Gilles Frapper, Eva Zurek. Structurally Constrained Evolutionary Algorithm for the Discovery and Design of Metastable Phases. 2023. hal-04123022

**HAL Id: hal-04123022**

**<https://hal.science/hal-04123022>**

Preprint submitted on 9 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Structurally Constrained Evolutionary Algorithm for the Discovery and Design of Metastable Phases**

Busheng Wang,<sup>†</sup> Katerina P. Hilleke,<sup>†</sup> Samad Hajinazar,<sup>†</sup> Gilles Frapper,<sup>‡</sup>  
and Eva Zurek<sup>\*,†</sup>

<sup>†</sup>*Department of Chemistry, State University of New York at Buffalo, Buffalo, NY 14260-3000, USA*

<sup>‡</sup>*Applied Quantum Chemistry Group, E4 Team, IC2MP UMR 7285, Université de Poitiers-CNRS,  
86073, Poitiers, France*

E-mail: ezurek@buffalo.edu

## Abstract

Metastable materials are abundant in nature and technology, showcasing remarkable properties that inspire innovative materials design. However, traditional crystal structure prediction methods, which rely solely on energetic factors to determine a structure's fitness, are not suitable for predicting the vast number of potentially synthesizable phases that represent a local minimum corresponding to a state in thermodynamic equilibrium. Here, we present a new approach for the prediction of metastable phases with specific structural features, and interface this method with the XTALOPT evolutionary algorithm. Our method relies on structural features that include the local crystalline order (*e.g.*, the coordination number or chemical environment), and symmetry (*e.g.*, Bravais lattice and space group) to filter the parent pool of an evolutionary crystal structure search. The effectiveness of this approach is benchmarked on three known metastable systems: XeN<sub>8</sub>, with a two-dimensional polymeric nitrogen sublattice, brookite TiO<sub>2</sub>, and a high pressure BaH<sub>4</sub> phase that was recently characterized. Additionally, a newly predicted metastable melamine salt, *P*-1 WC<sub>3</sub>N<sub>6</sub>, was found to possess an energy that is lower than two phases proposed in a recent computational study. The method presented here could help in identifying the structures of compounds that have already been synthesized, and developing new synthesis targets with desired properties.

## Introduction

Thermodynamically metastable phases, which are kinetically trapped and do not correspond to the global minimum in the free energy landscape, are commonly found in nature and can be synthesized in the laboratory and mass produced industrially.<sup>1</sup> According to the Materials Project database (accessed in June 2023), 53.13% of the experimentally verified phases are metastable.<sup>2</sup> These compounds can exhibit a range of unique properties, which can be attributed to a larger number of possible atomic arrangements and chemical bonding configurations they can adopt as compared to their ground states. Just a few examples of the broad spectrum of applications that metastable phases may be used for include superconductors,<sup>3,4</sup> photocatalysts,<sup>5</sup> photovoltaics,<sup>6,7</sup> ion conductors,<sup>8</sup> and steels.<sup>9</sup> These examples illustrate that metastable phases may present novel opportunities for materials innovation and design.

In the last decade computational crystal structure prediction (CSP) has become a powerful tool facilitating materials discovery.<sup>10–12</sup> A number of methods have been developed and applied to predict the structures of materials without input from any experimental data including random search,<sup>13</sup> simulated annealing,<sup>14</sup> metadynamics,<sup>15</sup> minima hopping,<sup>16</sup> basin hopping,<sup>17–19</sup> particle swarm optimization,<sup>20</sup> evolutionary algorithms,<sup>21–25</sup> as well as template-based elemental substitution.<sup>26</sup> Traditionally, these methods have been devoted towards finding the global (free) energy minimum (ground state), and predicting the crystal structure for a desired metastable material – at best, a secondary target – has remained a formidable challenge. The reason for this is that uncovering the structures of low-lying metastable phases that could potentially be synthesized is complicated by the vast number of local minima that are not typically accessed using traditional CSP methods.

In recent years, significant progress has been made in advancing metastable phase prediction through data-assisted CSP searches. Various computed physical and chemical properties have been employed as objectives to be optimized (either in lieu of, or in addition to the enthalpy/energy) including but not limited to hardness,<sup>27–29</sup> density,<sup>30</sup> band gap,<sup>31,32</sup> the degree of interstitial electron localization,<sup>33</sup> phase transitions,<sup>34</sup> as well as structural information (*e.g.*, predefined molec-

ular structures<sup>3,35–38</sup>). Experimental data (*e.g.*, comparison of simulated and experimental X-ray diffraction (XRD) patterns<sup>39–41</sup>) has also been employed within such CSP searches.

Yet, comparison of computed observables with experiment may, at times, be problematic. This is especially true when the structural characteristics of a synthesized compound cannot be fully measured and only partial geometric details can be determined. Examples of data that are helpful, but not sufficient, to characterize a particular crystal include local order or coordination environments that can be probed using infrared and Raman spectroscopy,<sup>42,43</sup> or the symmetry of a particular lattice or sub-lattice obtained via XRD.<sup>44,45</sup> In such cases, the absence of complete structural information can impede precise compound characterization. For example, determining the crystal structure of hydrides is a particularly challenging task,<sup>44–46</sup> as hydrogen provides weak X-ray scattering in contrast to heavier metallic atoms, given the proportionality of the scattering power to the atomic number.<sup>47,48</sup> Additionally, when it is known that specific structural features are beneficial for a particular property, these can be leveraged in the design of hypothetical materials. Some examples include thermodynamically metastable oxalate-based  $\text{Li}_x(\text{CO}_2)_y$  compounds that can be employed as organic materials for lithium-ion batteries.<sup>49</sup> The impact of such theoretical insights cannot be understated, as they offer valuable avenues for expanding the design possibilities of important materials in a forward-thinking manner. Therefore, the development of a versatile global search method for identifying crystal structures of metastable phases that incorporates partial structural features derived from theoretical models and experimental data is highly desirable.

In this study, we present a new approach for determining the crystal structure of a metastable phase with specific structural features based on the XTALOPT<sup>23,50</sup> evolutionary algorithm (EA). Our method automatically filters out structures that do not adhere to predefined criteria including but not limited to the local crystalline order (*e.g.*, coordination number and chemical environment) and symmetry (*e.g.*, Bravais lattice and space group) before ranking according to total energy or enthalpy. The accuracy and effectiveness of this method was evaluated by benchmarking its performance against three systems:  $\text{XeN}_8$ <sup>51</sup> containing 2D polymeric nitrogen motifs at 50 GPa, brookite  $\text{TiO}_2$  at ambient pressure, and the  $\text{BaH}_4$  phase synthesized at 50 GPa that possesses an

$I4/mmm$  Ba sub-lattice.<sup>44</sup> Additionally, we have predicted a new metastable melamine salt,  $P-1$   $WC_3N_6$ , which is found to be lower in energy compared to the compounds reported recently by Chen, Wang and Dronskowski.<sup>52</sup> Our results demonstrate that a constrained search is more efficient in identifying metastable phases with user-defined structural features compared to a traditional CSP search.

## Methods

*Structure prediction.* Our CSP searches were carried out using the in-house developed XTALOPT<sup>23,50</sup> EA (release 12),<sup>53</sup> which was designed to find stable and metastable structures of crystals given only their composition. PYMATGEN,<sup>54</sup> an open-source Python library for materials analysis, was utilized to determine the degree of local crystalline order (*e.g.*, coordination number, bond lengths and chemical environments) and symmetry (*e.g.*, space group and Bravais lattice) as global variables for each structure that could be used in a constrained search for metastable phases, by restricting the breeding pool to crystals that possessed the desired structural feature. The *pymatgen.symmetry.analyzer* module in the PYMATGEN library includes the *SpacegroupAnalyzer* class that provides two methods, namely, *get\_space\_group\_number* and *get\_crystal\_system*, which serve to extract the international space group number and the crystal system (Bravais lattice), respectively, for a given structure. In this study, the two aforementioned methods were implemented with a symmetry tolerance of 0.02 for both  $BaH_2$  and  $TiO_2$ . Additionally, the *pymatgen.core.structure* module contains the *get\_neighbors* method, which facilitates the retrieval of the coordination number of an atom within the structure. A cutoff value of 1.5 Å was used for the N-N and C-N bond distance (see the Supporting Information, or SI, for more detailed information on how this value was chosen). The energies or enthalpies associated with the phases that did not adhere to the structural constraints were assigned an exceptionally large value (*e.g.*, 99 eV per formula unit or FU). Because the breeding pool was chosen to include only the 50 most stable structures, this artificially large value prevented the structures that did not satisfy the constraints

from populating the breeding pool. The CSP searches for  $\text{XeN}_8$  (1-2),  $\text{TiO}_2$  (2, 4, 8),  $\text{BaH}_4$  (2, 4), and  $\text{WC}_3\text{N}_6$  (1-2) were conducted with the number of FU in the primitive cell given in the brackets. PyXtal<sup>55</sup> was used to generate molecular crystals as seeds for  $\text{WC}_3\text{N}_6$ , which contains melamine,  $(\text{C}_3\text{N}_6)^{6-}$ , motifs and tungsten atoms. Detailed information regarding the scripts utilized for applying the constraints are provided in the SI.

*First-principles calculations.* The geometry optimizations and electronic structure calculations were performed using the Vienna *Ab Initio* Simulation Package (VASP),<sup>56,57</sup> coupled with the Perdew-Burke-Ernzerhof (PBE) gradient-corrected exchange and correlation functional,<sup>58</sup> which was used for all of the CSP searches. To test the effect of dispersion interactions on the relative energies or enthalpies the D3 (BJ)<sup>59,60</sup> and optB88<sup>61</sup> approaches were used in the case of  $\text{WC}_3\text{N}_6$  and  $\text{XeN}_8$ . The electron-ion interaction was described by the projector-augmented-wave method,<sup>62</sup> in which the Xe  $4d^{10}5s^25p^6$ , N  $2s^22p^3$ , Ti  $3d^34s^1$ , O  $2s^22p^4$ , Ba  $5s^25p^86s^2$ , H  $1s^1$ , W  $5d^56s^1$ , and C  $2s^22p^2$  configurations were treated as valence. In the structure searches, the geometries were optimized with the “accurate” precision setting in VASP using a three-step process. First, the atomic positions were allowed to relax within a set unit cell, which was followed by a volume-only relaxation. Lastly, a full relaxation was conducted. A plane-wave basis set with an energy cutoff of 520 eV, and a  $\Gamma$ -centered Monkhorst-Pack  $k$ -mesh<sup>63</sup> where the number of divisions along each reciprocal-lattice vector was chosen such that its product with the real lattice constant was 30 Å, was employed. These values were increased to 600 eV and 50 Å for precise reoptimizations and electronic structure calculations. For  $\text{XeN}_8$  and  $\text{BaH}_4$  under pressure, energy cutoffs of 1000 and 700 eV were utilized, respectively. The phonon spectra of the new phase of  $\text{WC}_3\text{N}_6$  was computed using the frozen phonon method implemented in the PHONOPY<sup>64</sup> code.

# Results and Discussion

## Constrained Evolutionary Search – Workflow

Our methodology for setting up the constrained CSP searches, which are described in this manuscript, is built upon the open-source EA, XTALOPT,<sup>50</sup> developed in our group. In its most basic form, XTALOPT constructs possible unit cells and atomic coordinates of structures that sample the full  $3N+3$ -dimensional potential energy surface (PES), and are subsequently relaxed to the nearest stationary point via an external program. Most often, a periodic density functional theory (DFT) package is employed, though codes that make use of interatomic potentials have also been interfaced with XTALOPT. The goal of traditional XTALOPT searches is to find the location in the PES that corresponds to a minimum in the energy or enthalpy, termed the global minimum. Without including effects such as temperature, configurational entropy or reaction barriers, this would correspond to the structure expected to be dominant in a system at chemical equilibrium. In addition to the global minimum, many local minima separated by energy barriers are scattered throughout the PES, which correspond to alternative potential configurations of atoms in unit cells, albeit ones that are higher in energy than the global minimum. Along the way, XTALOPT may also identify structures at these local minima, although only the global minimum is explicitly targeted. The local minima represent metastable structures, which does not preclude them from experimental observation nor from persisting once formed. Diamond is metastable at ambient conditions, representing a local minimum in the PES of elemental carbon, whereas graphite is the global minimum. Regardless of its metastable status, the diamond allotrope of carbon still – famously – persists “forever” at ambient temperature/pressure conditions.

In the same way that evolution in a biological population adapts to its environment over time, with favorable traits being passed down from generation to generation, the XTALOPT EA will selectively choose favorable (low-energy) configurations for propagation to future generations of structures. The “fitness” of each of these configurations is evaluated, typically according to their thermodynamic stability represented by enthalpy or energy. This procedure must be modified to

target local rather than global minima. Therefore, in the current work we introduce additional criteria, by constraining the crystals that make up the parent pool to those that contain user defined structural features. In such a way the selection, mutation, and recombination operators will work on the population of atomic configurations *that match the structural criteria* to generate the lowest-energy structures that possess the desired motifs. This approach effectively enhances the accuracy and efficiency in predicting metastable phases that contain user-defined structural features.

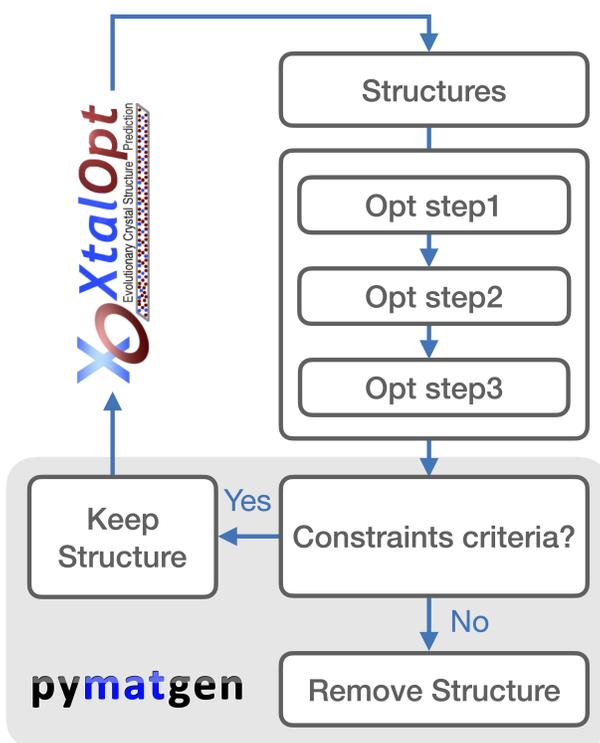


Figure 1: Schematic workflow for conducting constrained crystal structure prediction searches through the utilization of the XTALOPT EA coupled with the PYMATGEN library. Structures are generated as in a traditional XTALOPT search (open area), then passed to an external program for local relaxation (white box). The relaxed structures are analyzed via PYMATGEN (grey box), and the parent pool upon which XTALOPT can apply evolutionary operators is filtered. The constraints, which may include specified coordination numbers and bond lengths, space group symmetry or Bravais lattice, are applied after the geometry optimizations are complete, and these constraints remove unsuitable structures from the breeding pool.

Figure 1 shows a schematic workflow for conducting constrained searches through the utilization of XTALOPT coupled with the PYMATGEN library. As in a traditional XTALOPT search the initial set of structures may either be seeded by the user, or generated randomly, for example via

the RANDSPG algorithm<sup>65</sup> that can be employed to create random symmetric crystals. Next, the geometries of these initial structures may be optimized loosely using any of the total energy calculation methods available in the XTALOPT code. In the present work, the local relaxations were performed by first-principles DFT in a three-stage optimization step as described in the *Methods* section above. This reduces the noise of the potential energy surfaces and favors the generation of chemically sensible structures to accelerate the speed at which favorable geometries are identified. Post-optimization, duplicates are removed from the breeding pool using the XTALCOMP algorithm<sup>66</sup> to maintain structural diversity. What is new in this work is that the crystals in the breeding pool are further subject to structural constraints that are applied via the PYMATGEN library,<sup>54</sup> as described above. Crystals that do not meet the specified constraints are removed from the breeding pool. Next, as in a traditional search, the evolutionary operators are applied on the optimized structures (parents) to generate the child structures.

In what follows, we provide examples of the utility of this method on a number of systems chosen to illustrate how appropriately defining the desired coordination numbers, chemical environments, as well as the space group and Bravais lattice type can be used to filter the parent pool. The constrained technique described in Figure 1 has been integrated with the XTALOPT software, and we demonstrate its utility for XeN<sub>8</sub>, and brookite TiO<sub>2</sub>, which are metastable by 102 and 13 meV/atom, respectively, at 50 GPa and ambient pressure. Furthermore, we also used the method to predict the structure of a synthesized high-pressure phase of BaH<sub>4</sub>, whose Ba sublattice was characterized with the aid of XRD.<sup>44</sup> Finally, utilizing the methodology proposed herein a new metastable H-free melanine phase, W(C<sub>3</sub>N<sub>6</sub>), is predicted. This polymorph is calculated to be lower in energy than the previously reported potential multifunctional materials based upon the (C<sub>3</sub>N<sub>6</sub>)<sup>6-</sup> unit.<sup>52</sup>

## **Constraining Coordination Numbers for Prediction of 2D Polymeric Nitrogen**

The prediction and discovery of novel polymeric nitrogen or nitrogen-rich compounds is of great interest in contemporary times because of their potential as environmentally-friendly high-energy

density materials (HEDMs). Such compounds should ideally possess N-N single or N=N double bonds, which may release a vast amount of energy upon decomposition to the inert and thermodynamically stable N≡N triple bond found in N<sub>2</sub>. Several polymeric nitrogen configurations have been previously reported such as chain,<sup>67</sup> layered,<sup>68–71</sup> cubic-gauche allotrope (cg-N),<sup>72</sup> cage-like,<sup>73</sup> N<sub>5</sub>, N<sub>6</sub>, and N<sub>8</sub> molecular crystals.<sup>74–77</sup> Note that four of these computationally predicted phases have been experimentally characterized, *i.e.*, bulk cg-N,<sup>72</sup> layered polymeric nitrogen (LP-N),<sup>68</sup> a layered black-phosphorus-type nitrogen allotrope (BP-N),<sup>69</sup> and hexagonal layered polymeric nitrogen (HLP-N).<sup>70</sup> Contrary to LP-N, BP-N, and HLP-N, which are composed of 3-coordinated nitrogen atoms and possess layers that are at least two atoms thick, a novel polymeric nitrogen with layers a single layer thick has recently been predicted, with intercalated alkali metal (K<sup>78</sup>) or noble gas (Xe<sup>51</sup>) atoms. This new family of structures characterized by two-dimensional (2D) polymeric nitrogen possesses crown ether-like nanopores, which contain both 2- and 3-coordinated N atoms.

To illustrate the evolution of a CSP search carried out with the traditional XTALOPT algorithm versus one performed using constraints, it is useful to plot the DFT energies or enthalpies of the optimized crystals versus their structure index, which is directly related to the time at which the structure was generated during the course of the run. Figures 2(a) and (b) illustrate the progress of the unconstrained and constrained searches, respectively, carried out on the XeN<sub>8</sub> stoichiometry at 50 GPa. Before describing the exact constraints that were employed, let us comment on how these two plots differ. In the constrained search, the discarded structures (denoted using a thick line at 99 eV/FU) corresponded to 57.4% of all optimized phases, and were deemed ineffective structures that have been excluded from consideration. From the remaining crystals, 99.1% of the low-lying structures were located within the range of -25.0 to -29.0 eV/FU, as shown in Figure 2(b). Meanwhile, it is evident that the majority of these effective structures can be broadly categorized into three distinct enthalpy regions: -25.0 to -25.5 eV/FU, -27.0 to -27.5 eV/FU, and -28.3 to -29.0 eV/FU. The structures falling within these specific enthalpy ranges contain the desired 2D polymeric nitrogen, with the primary distinction between them being the torsion angles of the

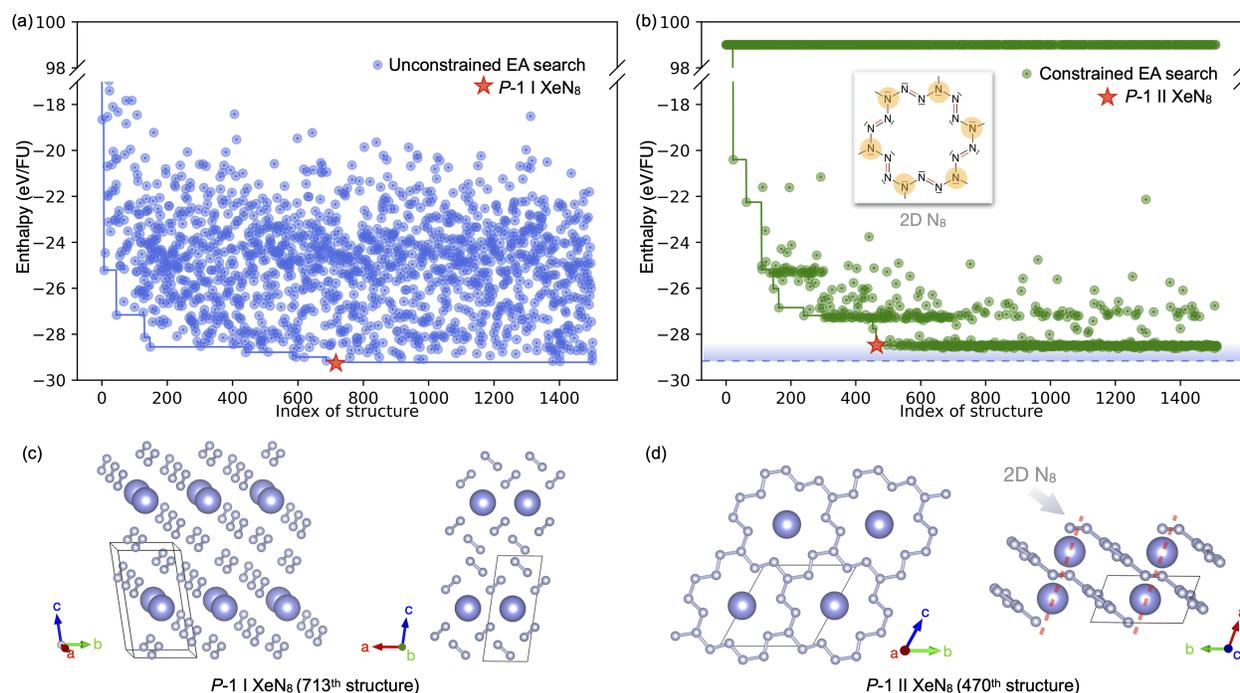


Figure 2: (a-b) The evolution of the enthalpy versus the structure index for the XeN<sub>8</sub> stoichiometry at 50 GPa performed with a CSP search without (a) and with (b) constraints is shown by the blue and green dots, respectively. Approximately 1500 crystals were optimized in both runs. The most stable structures found within these searches are indicated by the red stars. (c-d) The crystal structures of  $P-1 I$  (the global minimum) and metastable  $P-1 II$  XeN<sub>8</sub>. The Lewis structure of the 18-crown-6 N<sub>18</sub> unit in the 2D-N<sub>8</sub> layer of  $P-1 II$  XeN<sub>8</sub> is shown in (b). The solid thin blue and green lines in (a) and (b) denote the lowest enthalpy structure found during the course of the search. In (b) the blue dashed line represents the enthalpy of the ground state  $P-1 I$  XeN<sub>8</sub> phase. The enthalpies of the discarded phases in the constrained search are represented with a constant value of 99 eV/FU. Nitrogen atoms are small and xenon atoms are big balls.

2-coordinate nitrogen atoms. For instance, the group of structures with enthalpies from -25.0 to -25.5 eV/FU exhibit a more crumpled N layer in comparison to those found between -28.3 to -29.0 eV/FU, which possess a relatively flat 2D N layer. The unconstrained EA search results in a larger diversity of optimized structures, which is reflected by their greater range in enthalpies, as depicted in Figure 2(a). Specifically, 96.0% of the low-lying structures are located within the enthalpy range of -18.0 to -30.0 eV/FU. The blue or green jagged lines in the two searches highlight the enthalpy of the most stable phase found, and the red stars the first instance that the target phases were identified.

The ground state  $\text{XeN}_8$  phase, whose unit cell is illustrated in Figure 2(c), features exclusively  $\text{N}_2$  dumbbells that surround the Xe atoms, and it possesses  $P-1$  symmetry. It was the 713<sup>th</sup> structure to be generated during the course of the unconstrained EA search (Figure 2(a)), and was found four times within the run. Because  $\text{N}_2$  molecules comprise this polymorph, it cannot be considered as a candidate for a HEDM. An alternative  $\text{XeN}_8$  crystal that contains 2D polymeric nitrogen motifs could be based upon intercalating the noble gas atom into the pores of 18-crown-6  $\text{N}_{18}$  units, whose Lewis structure is presented in Figure 2(b, inset).<sup>51</sup> It is noteworthy that the 2D polynitrogen layer based upon the 18-crown-6  $\text{N}_{18}$  motifs exhibits a specific ratio of nitrogen atoms with different coordination numbers: 1/3 of the nitrogen atoms are three-coordinate (highlighted by yellow dots in Figure 2(b), while 2/3 are two-coordinate. Each ring of the aromatic 2D polynitrogen layer contains ten  $\pi$  electrons per  $\text{N}_{18}$  unit. These electrons are distributed such that each  $\text{N}^t$  (tri-coordinate) atom has one  $\pi$  lone pair shared by three rings ( $6 \times 1/3 \pi$  pairs), and each  $\text{N}^d$  (di-coordinate) atom has one  $\pi \text{N}^d=\text{N}^d$  pair shared by two rings ( $6 \times 1/2 \pi$  pairs). This molecular orbital configuration follows the Hückel rule and results in a  $4n + 2$  electron count, which confers chemical stability to the 2D polynitrogen layer.

What type of constraints could be used to find a metastable crystal lattice based upon this  $\text{N}_{18}$  motif? As alluded to above, one possibility could be to constrain the coordination numbers of the nitrogen atoms. Requiring 1/3 of the N atoms to be 3-coordinate and the remainder to be 2-coordinate in the parent pool of a CSP search performed at 50 GPa resulted in the discovery

of the  $P-1$  symmetry phase illustrated in Figure 2(d), which is consistent with previous predictions.<sup>51</sup> As desired, this phase contains slightly puckered 2D polynitrogen layers in the  $bc$  plane that are separated by Xe atoms along the  $a$  lattice vector, and the layers are stacked directly on top of each other. At 50 GPa this desired  $P-1$  II phase is computed to be 0.918 eV/FU higher in enthalpy than the  $P-1$  I phase, but by 100 GPa it becomes the ground state  $\text{XeN}_8$  polymorph. Despite the fact that  $P-1$  II  $\text{XeN}_8$  remains dynamically stable, and could be kinetically trapped upon decompression to ambient pressure, it becomes increasingly thermodynamically disfavored at lower pressures, where structures featuring  $\text{N}_2$  molecular units, like  $P-1$  I, dominate. Indeed,  $P-1$  II was not predicted in the traditional search illustrating that without imposing the aforementioned constraints, this metastable phase would never be found at 50 GPa. This example highlights the effectiveness of our constrained search method in isolating a specific metastable phase, here, based on coordination number, among a large number of possible configurations. Though this EA search was performed using the PBE functional, enthalpy comparisons based upon functionals that include vdW dispersion yielded the same trends (see Table S3 in SI).

## Finding Polymorphs through Bravais Lattice Constraints

In addition to considering the local crystalline order (represented by coordination number for the test case of  $\text{XeN}_8$ ), unit cell symmetry could also be leveraged as a criterion for predicting the crystal structure of a desired metastable compound. Here, we take  $\text{TiO}_2$  as an example, aiming to predict the polymorphs that possess an orthorhombic Bravais lattice. Multiple phases of  $\text{TiO}_2$  are known with the three main polymorphs being anatase (FU = 4), rutile (FU = 2), and brookite (FU = 8),<sup>79,80</sup> whose crystal structures are shown in Figure 3(a). Using the PBE functional at 0 GPa anatase turns out to be the ground state,<sup>79</sup> though various interatomic potentials have predicted that either rutile or brookite are preferred.<sup>81</sup> Within the PBE functional, which we use here, brookite and rutile turned out to be metastable with their energies being 42 and 93 meV/FU higher than that of anatase, respectively. As such, an unconstrained XTALOPT search would be expected to yield anatase  $\text{TiO}_2$  as the lowest energy phase.

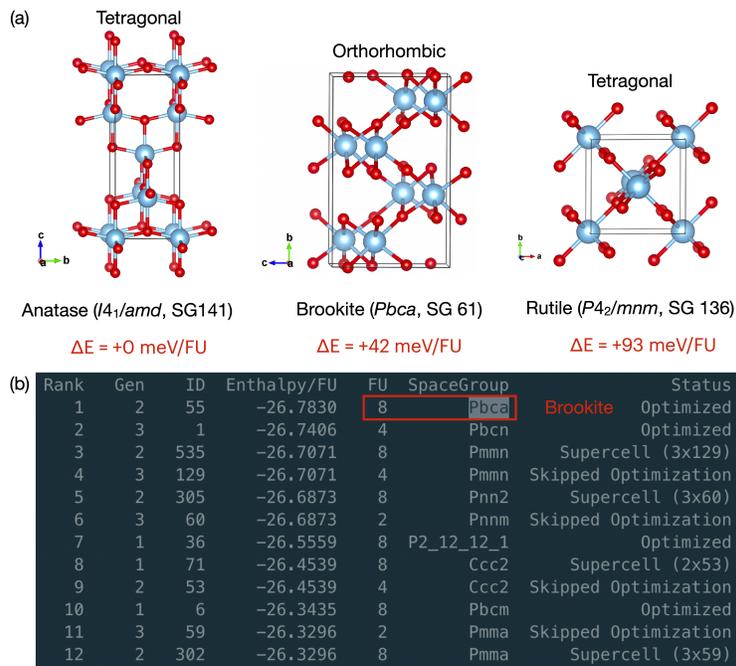


Figure 3: (a) The crystal structures of the three main TiO<sub>2</sub> phases: anatase ( $I4_1/amd$  space group), brookite ( $Pbca$  space group), rutile ( $P4_2/mnm$  space group). Their PBE-calculated relative enthalpies are provided, with anatase being the most stable. (b) The output file (results.txt) of an XTALOPT EA search constrained so that the structures in the breeding pool possessed an orthorhombic Bravais lattice. The brookite phase of TiO<sub>2</sub> is calculated to be the most stable one that meets the desired criteria. Oxygen atoms are red and titanium atoms are blue balls.

In all three of these polymorphs the titanium atoms are six coordinated either in a trigonal prismatic (brookite), or a perfect octahedral (rutile) or distorted octahedral (anatase) configuration. The oxygen atoms, on the other hand, exhibit a 3-coordinate arrangement in all forms of rutile, anatase, and brookite. Therefore, if our goal was to uncover the brookite phase using a CSP search, constraining the coordination numbers on the titanium and oxygen atoms might not yield the desired target. Could the crystal lattices of these polymorphs be used instead? While both anatase (space group  $I4_1/amd$ ) and rutile (space group  $P4_2/mnm$ ) adopt a tetragonal lattice, brookite (space group  $Pbca$ ) possesses an orthorhombic Bravais lattice. PYMATGEN's *SpacegroupAnalyzer.get\_crystal\_system* library allows us to obtain the Bravais lattice for a crystal structure. Consequently, in an attempt to specifically locate the orthorhombic brookite phase, we carried out a constrained evolutionary search in which exclusively those structures with an orthorhombic Bravais lattice were retained in the pool, while all other structures were eliminated.

During the course of an XTALOPT run, a running list of all of the structures analysed, ranked from lowest to highest in enthalpy or energy, is maintained. The contents of this output file, named 'results.txt', for the constrained search for orthorhombic  $TiO_2$  is presented in Figure 3(b). In this search, wherein around 1300 structures were generated and optimized, brookite appeared as the 345<sup>th</sup> crystal and was highlighted as the highest ranked polymorph in the results.txt file. An analogous unconstrained search ranked anatase, the 187<sup>th</sup> crystal found, as the most stable polymorph with the highest rank, whereas brookite and rutile were found as the 1141<sup>th</sup> and 559<sup>th</sup> structures with rank #16 and #25, respectively, see Section S4 in SI. Though both the constrained and unconstrained searches identified brookite, the former was much more efficient finding it as the 345<sup>th</sup> as compared to the 1141<sup>th</sup> structure. This indicates the inadequacy of the conventional evolutionary search process when compared to the constrained approach for identifying a specific metastable phase.

## Sublattice Symmetry for Aiding Experimental Structural Determination

Though XRD is an invaluable tool for the experimental characterization of a crystalline material, sometimes it cannot be used alone to determine a compound's structure. This is particularly true when elements with similar scattering factors comprise the phase (*e.g.*, it is difficult to distinguish between boron and carbon via XRD), heavy elements with very strong scattering factors mask signals arising from weaker scatters, or when light elements – hydrogen being the most notorious – are present. Another example is high pressure research where XRD patterns obtained in either static<sup>44,45</sup> or dynamic compression<sup>82,83</sup> experiments may only provide incomplete data due to the experimental set-up.

In such situations, an EA search constrained by the available XRD data may be extremely useful in uncovering the structure of a synthesized phase. In fact, in the field of high pressure research unconstrained CSP searches have already proven to be invaluable in predicting targets for synthesis and identifying the materials that have been made.<sup>11</sup> A case in point is a recent report that described the synthesis of BaH<sub>4</sub> using high-pressure diamond anvil cell experiments.<sup>44</sup> Synchrotron measurements were performed to obtain the XRD pattern of the synthesized polyhydride, with the experimental XRD pattern depicted in Figure 4(a). Analysis of the data led the researchers to conclude that it resulted from a co-existence of two phases. One of these was a *P6<sub>3</sub>/mmc* BaH<sub>2</sub> compound, which was previously known,<sup>84</sup> and the second was attributed to an unknown BaH<sub>4</sub> phase. Given the weak scattering power of hydrogen atoms,<sup>47,48</sup> the symmetry that was associated with the second unknown phase, *I4/mmm*, could only provide information about the structure of the heavy barium sublattice (see inset). Though the hydrogen atom positions could not be extracted directly from the experimental data, their assignment could be made based upon DFT calculations that relied on CSP combined with analogy to other known structures. The ground state found via CSP performed using the *ab initio* random structure searching (AIRSS)<sup>13</sup> method was a *Cmcm* symmetry crystal<sup>44</sup> that did not quite reach the Ba-H convex hull. However, the experiment was in better agreement with a metastable phase whose enthalpy lay  $\sim 300$  meV/FU above the ground state, which was isotypic with many of the tetrahydrides of an electropositive metal element that

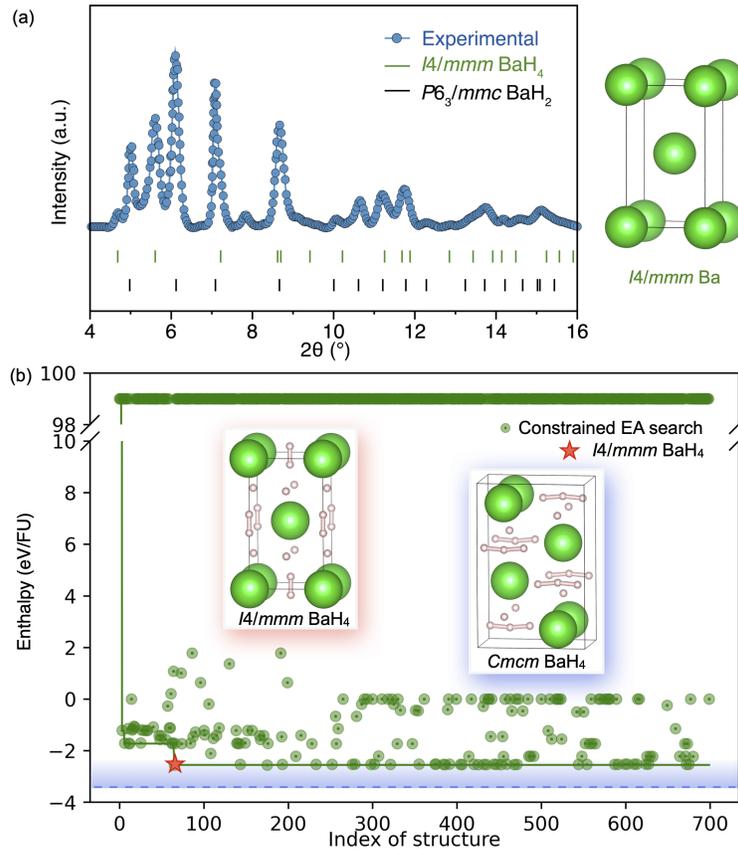


Figure 4: (a) The XRD pattern of a phase mixture of barium polyhydrides at 45 GPa from Ref.<sup>44</sup> The tick marks indicate Bragg peaks from both  $I4/mmm$   $BaH_4$  and  $P6_3/mmc$   $BaH_2$ . The right panel displays the unit cell of an  $I4/mmm$   $Ba$  configuration. (b) The evolution of the enthalpy versus the structure index for  $BaH_4$  at 50 GPa in an XTALOPT search constrained so that the breeding pool only contains structures with an  $I4/mmm$  symmetry  $Ba$  sublattice. The 65<sup>th</sup> structure corresponded to the most stable phase satisfying this constraint, as indicated by the red star. The blue dashed line represents the calculated enthalpy of the ground state ( $Cmcm$   $BaH_4$ ), which was found in an unconstrained search. The crystal structures of the  $I4/mmm$  and  $Cmcm$   $BaH_4$  phases are inserted in (b). Barium atoms are green and hydrogen atoms white balls.

have been predicted or synthesized under pressure.<sup>85</sup> These two structures are compared in the inset in Figure 4(b). To determine if a more stable phase could explain the experimental data, it is imperative to perform a search that constrains the breeding pool to  $I4/mmm$  Ba sublattices.

To illustrate the utility of a constrained XTALOPT search in such a situation, the *SpacegroupAnalyzer.get\_space\_group\_number* method in the PYMATGEN library was employed to determine the space group of the Ba sublattice of each optimized structure. Figure 4(b) illustrates the evolution of the calculated enthalpy as a function of the structure index. The phase with the lowest enthalpy identified in this search, with an  $I4/mmm$  space group, first appeared as the 65<sup>th</sup> structure, indicated with a red star in Figure 4(b). It was identical to the  $I4/mmm$  BaH<sub>4</sub> proposed in Ref.,<sup>44</sup> and isotypic with many of the previously predicted or synthesized metal tetrahydrides,<sup>85</sup> confirming the structural assignment made in Ref.<sup>44</sup>

Our DFT calculations found that the enthalpy of  $I4/mmm$  BaH<sub>4</sub> was 277 meV/FU higher than that of the  $Cmcm$  phase, which appeared as the 375<sup>th</sup> structure in an unconstrained search, confirming it was the ground state. In the unconstrained EA run  $I4/mmm$  BaH<sub>4</sub> was ranked as the 354<sup>th</sup> lowest enthalpy candidate in the ‘results.txt’ file, and it is therefore unlikely that it would have been chosen for further analysis from an unconstrained search without the availability of experimental data. This highlights the inefficiency of the traditional evolutionary search method, as compared to the constrained one, for identifying a specific metastable phase. The large enthalpy difference between the  $I4/mmm$  and  $Cmcm$  tetrahydrides substantiates the hypothesis that the former could be an intermediate phase resulting from hydrogen migration through the Ba lattice.<sup>44</sup>

Furthermore, it is worth mentioning that the methodology described in this section was used by us in an attempt to uncover the structure of an unknown-stoichiometry N-doped lutetium hydride,<sup>86</sup> which was suggested to comprise the putative room-temperature superconductor recently reported under kbar pressures.<sup>45</sup> The measured XRD pattern could only provide evidence about the Lu sublattice, which was proposed to be face-centered-cubic ( $fcc$ ) and did not reveal the positions of the hydrogen or nitrogen atoms. In our work several metastable Lu-N-H structures featuring  $fcc$  Lu were successfully identified using constrained XTALOPT searches, although their estimated  $T_c$ s

were far below room temperature.<sup>86</sup>

## Molecular Units by Constraining the Chemical Environment

In a search for new photocatalysts for water splitting, Chen and coworkers used CSP methods to uncover two energetically favorable polymorphs of  $\text{WC}_3\text{N}_6$ , a salt comprised of  $(\text{C}_3\text{N}_6)^{6-}$ , a fully deprotonated form of melamine ( $\text{C}_3\text{H}_6\text{N}_6$ ).<sup>52</sup> The two polymorphs, which adopt space groups  $P1$  and  $P3$ , were predicted using the particle swarm optimization algorithm.<sup>20</sup> Within them, the tungsten atom is found in either a 5-coordinated ( $P1$ ) or 6-coordinated ( $P3$ ) environment, as depicted in Figure 5(a,b). Given that the melaminate anion,  $(\text{C}_3\text{N}_6)^{6-}$ , has a higher energy than three carbodiimide anions,  $(\text{CN}_2)^{2-}$ , we wondered if a constrained XTALOPT search would be successful in locating novel polymorphs of  $\text{WC}_3\text{N}_6$ ?

The constraints that were the most appropriate for our target could be derived from the chemical environments and coordination numbers of C and N within the melaminate anion. Specifically, each C atom is coordinated to three N atoms, and whereas half of the N atoms are expected to be coordinated to a single C atom, the remaining half will be coordinated to two C atoms, see the melaminate in Figure 5(a). The cutoff value for the C-N bond distance of 1.5 Å was chosen based upon a comparison with the expected single and double bond lengths between carbon and nitrogen, as discussed more fully in the SI. The current version of XTALOPT is unable to create the first generation of structures by making use of molecular units, unless they conform to the well-known valence shell electron pair repulsion (VSEPR) geometries. Since the melaminate anion cannot be represented by a VSEPR geometry, it is highly unlikely that the structures generated via this method or using the RANDSPG algorithm would possess suitable chemical environments. Therefore, in addition to utilizing the aforementioned constraints to filter the breeding pool, the initial generation was populated with 100 random structures generated by RANDSPG, the previously predicted  $P1$  and  $P3$  configurations,<sup>52</sup> and with 98 molecular crystals of  $\text{WC}_3\text{N}_6$  generated by the PYXTAL code.<sup>55</sup> The initial generation was created in this way for both the constrained and traditional searches.

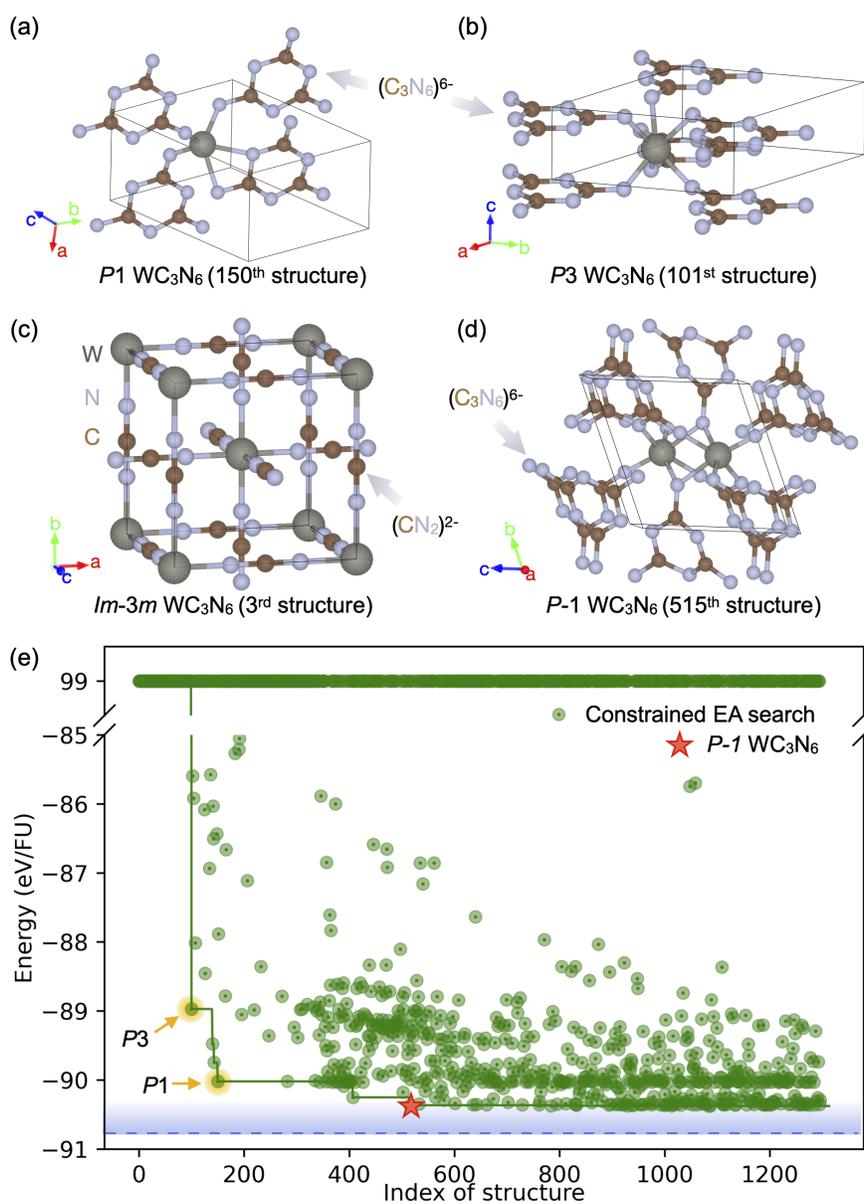


Figure 5: (a-d) The crystal structures of multiple phases of  $WC_3N_6$ . (a, b) The previously proposed  $P1$  and  $P3$  polymorphs with  $(C_3N_6)^{6-}$  units.<sup>52</sup> (c) The  $Im-3m$  phase, which is the ground state, found in a traditional search. This compound contains  $W^{6+}$  octahedrally coordinated by  $[N=C=N]^{2-}$  units. (d) The most stable  $P-1 WC_3N_6$  salt-like phase containing  $(C_3N_6)^{6-}$  motifs found in a constrained search. (e) The evolution of the energy versus structure index in an ambient pressure constrained CSP search for  $WC_3N_6$ . The 515<sup>th</sup> structure with space group  $P-1$  is indicated by the red star, while the 150<sup>th</sup> and 101<sup>st</sup> structures (seeds) are indicated by the orange arrows. The blue dashed line represents the calculated energy of the ground state ( $Im-3m WC_3N_6$ ) from an unconstrained search. The energies of the discarded phases in the constrained search are represented as a constant value of 99 eV/FU.

The most stable phase found in the unconstrained search was a carbodiimide-like salt in the  $Im-3m$  ( $Z = 2$ ) space group featuring W(VI) and 3  $(CN_2)^{2-}$  moieties (Figure 5(c)), which was found in the initial random generation as the third structure. Though we did not locate any other polymorphs in the literature with the  $WC_3N_6$  stoichiometry featuring carbodiimide anions, a  $W_2(CN_2)_3$  phase with these units, which is isostructural with an experimentally known  $R-3c$   $Cr_2(CN_2)_3$ <sup>87</sup> compound has been reported in the Materials Project database.<sup>2</sup>

In the constrained search, however, the most stable structure located corresponded to a  $P-1$  ( $Z = 2$ )  $WC_3N_6$  metastable phase with two W atoms in a distorted octahedral coordination, as shown in Figure 5(d), with an average W-N distance of 2.04 Å. This distance is extremely close to the expected W-N bond length of 2.06 Å for a  $W^{6+}$  ion coordinated to  $N^{3-}$  in a sixfold coordination state.<sup>52</sup> Phonon calculations confirmed that this structure was a local minimum. The previously reported  $P3$  and  $P1$  structures were found as the 101<sup>st</sup> and 150<sup>th</sup> phases in the constrained search (since they were employed as seeds), as illustrated in the energy versus structure index plot in Figure 5(d), and the newly predicted polymorph was found as the 515<sup>th</sup> structure. The energy of the new  $P-1$  phase was calculated to be 53 meV/atom higher than the global minimum,  $Im-3m$   $W(CN_2)_3$ , but 29 meV/atom (28 kJ mol<sup>-1</sup>) lower than the previously reported metastable  $P1$  phase at the D3 (BJ) level of theory. The relative stability of these phases has also been confirmed by the optB88 and PBE theories as displayed in Table 1. It is worth mentioning that the energy difference between  $P3$  and  $P1$   $WC_3N_6$  is calculated to be 100 meV/atom (97 kJ mol<sup>-1</sup>), consistent with the value of 96 kJ mol<sup>-1</sup> reported in Ref.<sup>52</sup>

**Table 1: The relative energies of various  $WC_3N_6$  phases as calculated at the D3 (BJ), optB88, and PBE levels of theory.**

Phases	$\Delta E$ (meV/atom)		
	D3 (BJ)	optB88	PBE
$Im-3m$	-53	-42	-45
$P-1^a$	0	0	0
$P1^b$	29	56	11
$P3^b$	129	145	120

<sup>a</sup> This work.

<sup>b</sup> Ref<sup>52</sup>.

## Conclusions

A method to predict metastable phases with specific structural features has been developed and implemented through a constrained scheme that utilizes the XTALOPT software and the PYMATGEN library. This approach involves automating global crystal prediction algorithms so the structures kept in their breeding pools meet user defined criteria pertaining to either local crystalline order or symmetry. The accuracy and effectiveness of this method was evaluated by benchmarking its performance against three metastable systems: XeN<sub>8</sub> with 2D polymeric nitrogen motifs at 50 GPa, brookite TiO<sub>2</sub> at ambient pressure, and a synthesized BaH<sub>4</sub> phase that exhibits an *I4/mmm* Ba sub-lattice, as deduced via experimental X-ray diffraction (XRD) measurements at 50 GPa. Additionally, we have predicted a new metastable melaminite salt, *P*-1 WC<sub>3</sub>N<sub>6</sub>, whose energy was found to be lower than a recently reported phase.<sup>52</sup>

Our study demonstrates that a constrained search is more efficient in identifying metastable phases with predefined structural features as compared to a traditional crystal structure search. The constrained method presented herein offers a substantial degree of flexibility, as it permits users to develop constraint criteria according to their specific requirements, and with the help of shared scripts provided in the SI. In situations where experimental data is available, infrared or Raman data may suggest coordination number constraints, while XRD data might indicate constraints related to symmetry, like the Bravais lattice or space group of the crystal, or even of its sub-lattice. Additionally, the local crystalline order and chemical environment, which are associated with specific material properties, could serve as the constraint criteria for the discovery of advanced materials with desired functionalities. We believe this method has the potential to aid in the identification of experimental compounds and the design of metastable phases important for basic science and for a broad range of potential applications.

## Associated Content

The Supporting Information can be accessed at no cost through <https://pubs.acs.org/doi/xxx>.

This resource provides comprehensive information on the constrained search methodology employed in conjunction with XTALOPT and PYMATGEN, including the associated scripts. Furthermore, the calculated structural parameters for the phases discussed are presented therein. The relative enthalpy of the Xe<sub>6</sub> phases was studied by utilizing the vdW correction methods. Additionally, the dynamic stability of the *P*-1 WC<sub>3</sub>N<sub>6</sub> phase was investigated.

## Author Information

### Corresponding Author

Eva Zurek - Department of Chemistry, State University of New York at Buffalo, Buffalo, NY 14260-3000, USA; ORCID: 0000-0003-0738-867X

Email: [ezurek@buffalo.edu](mailto:ezurek@buffalo.edu)

### Authors

Busheng Wang - Department of Chemistry, State University of New York at Buffalo, Buffalo, NY 14260-3000, USA; ORCID: 00000002-7743-9471

Katerina P. Hilleke - Department of Chemistry, State University of New York at Buffalo, Buffalo, NY 14260-3000, USA; ORCID: 0000-0003-4322-8403

Samad Hajinazar - Department of Chemistry, State University of New York at Buffalo, Buffalo, NY 14260-3000, USA; ORCID: 0000-0002-7255-5932

Gilles Frapper - Applied Quantum Chemistry Group, E4 Team, IC2MP UMR 7285, Université de Poitiers-CNRS, 86073 Poitiers, France; ORCID: 0000-0001-5177-6691

## **Author Contributions**

E.Z., G.F. and B.W., conceived the research. B.W. carried out the calculations and E.Z. supervised the study. K.H. and S.H. contributed to the analysis and made improvements to the XTALOPT code. All authors participated in discussing the results. The manuscript was written in collaboration with all authors. All authors have given approval to the final version of the manuscript.

## **Notes**

The authors declare no competing financial interest.

## **Acknowledgements:**

We acknowledge the U.S. National Science Foundation (DMR-2136038) for financial support. K.H. acknowledges the Chicago/DOE Alliance Center under Cooperative Agreement Grant No. DE-NA0003975. Calculations were performed at the Center for Computational Research at SUNY Buffalo (USA).<sup>88</sup> Gilles Frapper acknowledges the High-Performance Computing Center of Joliot-Curie Rome/TGCC GENCI (France) under Project No. A0140807539.

## References

- (1) Sun, W.; Dacek, S. T.; Ong, S. P.; Hautier, G.; Jain, A.; Richards, W. D.; Gamst, A. C.; Persson, K. A.; Ceder, G. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2016**, *2*, e1600225.
- (2) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G. et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1*, 011002.
- (3) Budden, M.; Gebert, T.; Buzzi, M.; Jotzu, G.; Wang, E.; Matsuyama, T.; Meier, G.; Laplace, Y.; Pontiroli, D.; Riccò, M. et al. Evidence for metastable photo-induced superconductivity in  $K_3C_{60}$ . *Nat. Phys.* **2021**, *17*, 611–618.
- (4) Yoshida, M.; Kudo, K.; Nohara, M.; Iwasa, Y. Metastable superconductivity in two-dimensional  $IrTe_2$  crystals. *Nano lett.* **2018**, *18*, 3113–3117.
- (5) Sclafani, A.; Herrmann, J. Comparison of the photoelectronic and photocatalytic activities of various anatase and rutile forms of titania in pure liquid organic phases and in aqueous solutions. *J. Phys. Chem.* **1996**, *100*, 13655–13661.
- (6) Nagabhushana, G.; Shivaramaiah, R.; Navrotsky, A. Direct calorimetric verification of thermodynamic instability of lead halide hybrid perovskites. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 7717–7721.
- (7) Kuech, T. F.; Babcock, S. E.; Mawst, L. Growth far from equilibrium: Examples from III-V semiconductors. *Appl. Phys. Rev.* **2016**, *3*.
- (8) Sanna, S.; Esposito, V.; Andreasen, J. W.; Hjelm, J.; Zhang, W.; Kasama, T.; Simonsen, S. B.; Christensen, M.; Linderoth, S.; Pryds, N. Enhancement of the chemical stability in confined  $\delta$ - $Bi_2O_3$ . *Nat. Mater.* **2015**, *14*, 500–504.

- (9) Li, Z.; Pradeep, K. G.; Deng, Y.; Raabe, D.; Tasan, C. C. Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off. *Nature* **2016**, *534*, 227–230.
- (10) Oganov, A. R. *Modern methods of crystal structure prediction*; John Wiley & Sons, 2011.
- (11) Zurek, E.; Grochala, W. Predicting crystal structures and properties of matter under extreme conditions via quantum mechanics: the pressure is on. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2917–2934.
- (12) Oganov, A. R.; Pickard, C. J.; Zhu, Q.; Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **2019**, *4*, 331–348.
- (13) Pickard, C. J.; Needs, R. Structures at high pressure from random searching. *Phys. Status Solidi B* **2009**, *246*, 536–540.
- (14) Kirkpatrick, S.; Gelatt Jr, C. D.; Vecchi, M. P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680.
- (15) Martoňák, R.; Laio, A.; Parrinello, M. Predicting crystal structures: the Parrinello-Rahman method revisited. *Phys. Rev. Lett.* **2003**, *90*, 075503.
- (16) Goedecker, S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *J. Chem. Phys.* **2004**, *120*, 9911–9917.
- (17) Wales, D. J.; Doye, J. P. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **1997**, *101*, 5111–5116.
- (18) Burnham, C. J.; English, N. J. Crystal structure prediction via basin-hopping global optimization employing tiny periodic simulation cells, with application to water–ice. *J. Chem. Theory Comput.* **2019**, *15*, 3889–3900.

- (19) Yang, S.; Day, G. M. Exploration and optimization in crystal structure prediction: Combining basin hopping with quasi-random sampling. *J. Chem. Theory Comput.* **2021**, *17*, 1988–1999.
- (20) Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B* **2010**, *82*, 094116.
- (21) Oganov, A. R.; Glass, C. W. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *J. Chem. Phys.* **2006**, *124*, 244704.
- (22) Trimarchi, G.; Zunger, A. Global space-group optimization problem: Finding the stablest crystal structure without constraints. *Phys. Rev. B* **2007**, *75*, 104113.
- (23) Lonie, D. C.; Zurek, E. XtalOpt: An open-source evolutionary algorithm for crystal structure prediction. *Comput. Phys. Commun.* **2011**, *182*, 372–387.
- (24) Tipton, W. W.; Hennig, R. G. A grand canonical genetic algorithm for the prediction of multi-component phase diagrams and testing of empirical potentials. *J. Phys. Condens. Matter.* **2013**, *25*, 495401.
- (25) Hajinazar, S.; Thorn, A.; Sandoval, E. D.; Kharabadze, S.; Kolmogorov, A. N. MAISE: Construction of neural network interatomic models and evolutionary structure optimization. *Comput. Phys. Commun.* **2021**, *259*, 107679.
- (26) Wei, L.; Fu, N.; Siriwardane, E. M.; Yang, W.; Omeo, S. S.; Dong, R.; Xin, R.; Hu, J. TCSP: a template-based crystal structure prediction algorithm for materials discovery. *Inorg. Chem.* **2022**, *61*, 8431–8439.
- (27) Lyakhov, A. O.; Oganov, A. R. Evolutionary search for superhard materials: Methodology and applications to forms of carbon and TiO<sub>2</sub>. *Phys. Rev. B* **2011**, *84*, 092103.
- (28) Zhang, X.; Wang, Y.; Lv, J.; Zhu, C.; Li, Q.; Zhang, M.; Li, Q.; Ma, Y. First-principles structural design of superhard materials. *J. Chem. Phys.* **2013**, *138*, 114101.

- (29) Avery, P.; Wang, X.; Oses, C.; Gossett, E.; Proserpio, D. M.; Toher, C.; Curtarolo, S.; Zurek, E. Predicting superhard materials via a machine learning informed evolutionary structure search. *npj Comput. Mater.* **2019**, *5*, 1–11.
- (30) Zhu, Q.; Oganov, A. R.; Salvadó, M. A.; Pertierra, P.; Lyakhov, A. O. Denser than diamond: Ab initio search for superdense carbon allotropes. *Phys. Rev. B* **2011**, *83*, 193410.
- (31) Xiang, H.; Huang, B.; Kan, E.; Wei, S.-H.; Gong, X. Towards direct-gap silicon phases by the inverse band structure design approach. *Phys. Rev. Lett.* **2013**, *110*, 118702.
- (32) Zeng, Q.; Oganov, A. R.; Lyakhov, A. O.; Xie, C.; Zhang, X.; Zhang, J.; Zhu, Q.; Wei, B.; Grigorenko, I.; Zhang, L. et al. Evolutionary search for new high-k dielectric materials: methodology and applications to hafnia-based oxides. *Acta Crystallogr. C: Struct. Chem.* **2014**, *70*, 76–84.
- (33) Zhang, Y.; Wang, H.; Wang, Y.; Zhang, L.; Ma, Y. Computer-assisted inverse design of inorganic electrides. *Phys. Rev. X* **2017**, *7*, 011017.
- (34) Zhu, Q.; Oganov, A. R.; Lyakhov, A. O.; Yu, X. Generalized evolutionary metadynamics for sampling the energy landscapes and its applications. *Phys. Rev. B* **2015**, *92*, 024106.
- (35) Zhu, Q.; Oganov, A. R.; Glass, C. W.; Stokes, H. T. Constrained evolutionary algorithm for structure prediction of molecular crystals: methodology and applications. *Acta Crystallogr. B: Struct. Sci.* **2012**, *68*, 215–226.
- (36) Curtis, F.; Li, X.; Rose, T.; Vazquez-Mayagoitia, A.; Bhattacharya, S.; Ghiringhelli, L. M.; Marom, N. GAtor: A first-principles genetic algorithm for molecular crystal structure prediction. *J. Chem. Theory Comput.* **2018**, *14*, 2246–2264.
- (37) Case, D. H.; Campbell, J. E.; Bygrave, P. J.; Day, G. M. Convergence properties of crystal structure prediction by quasi-random sampling. *J. Chem. Theory Comput.* **2016**, *12*, 910–924.

- (38) Kilgour, M.; Rogal, J.; Tuckerman, M. Geometric deep learning for molecular crystal structure prediction. *J. Chem. Theory Comput.* **2023**,
- (39) Gao, P.; Tong, Q.; Lv, J.; Wang, Y.; Ma, Y. X-ray diffraction data-assisted structure searches. *Comput. Phys. Commun.* **2017**, *213*, 40–45.
- (40) Ward, L.; Michel, K.; Wolverton, C. Automated crystal structure solution from powder diffraction data: Validation of the first-principles-assisted structure solution method. *Phys. Rev. Mater.* **2017**, *1*, 063802.
- (41) Ling, H.; Montoya, J.; Hung, L.; Aykol, M. Solving inorganic crystal structures from X-ray powder diffraction using a generative first-principles framework. *Comput. Mater. Sci.* **2022**, *214*, 111687.
- (42) Gard, P.; Sourisseau, C.; Ouvrard, G.; Brec, R. Infrared study of lithium intercalated phases in the  $\text{Li}_x\text{FeS}_2$  system ( $0 \leq x \leq 2$ ). Characterization of a new iron disulfide. *Solid State Ion.* **1986**, *20*, 231–238.
- (43) Wang, B.; Braems, I.; Sasaki, S.; Guégan, F.; Cario, L.; Jovic, S.; Frapper, G. Prediction of a new layered polymorph of  $\text{FeS}_2$  with  $\text{Fe}^{3+}\text{S}^{2-}(\text{S}_2^{2-})_{1/2}$  structure. *J. Phys. Chem. Lett.* **2020**, *11*, 8861–8866.
- (44) Peña-Alvarez, M.; Binns, J.; Marqués, M.; Kuzovnikov, M. A.; Dalladay-Simpson, P.; Pickard, C. J.; Ackland, G. J.; Gregoryanz, E.; Howie, R. T. Chemically assisted precompression of hydrogen molecules in alkaline-earth tetrahydrides. *J. Phys. Chem. Lett.* **2022**, *13*, 8447–8454.
- (45) Dasenbrock-Gammon, N.; Snider, E.; McBride, R.; Pasan, H.; Durkee, D.; Khalvash-Sutter, N.; Munasinghe, S.; Dissanayake, S. E.; Lawler, K. V.; Salamat, A. et al. Evidence of near-ambient superconductivity in a N-doped lutetium hydride. *Nature* **2023**, *615*, 244–250.

- (46) Einaga, M.; Sakata, M.; Ishikawa, T.; Shimizu, K.; Eremets, M. I.; Drozdov, A. P.; Troyan, I. A.; Hirao, N.; Ohishi, Y. Crystal structure of the superconducting phase of sulfur hydride. *Nat. Phys.* **2016**, *12*, 835–838.
- (47) Schmidtman, M.; Coster, P.; Henry, P. F.; Ting, V. P.; Weller, M. T.; Wilson, C. C. Determining hydrogen positions in crystal engineered organic molecular complexes by joint neutron powder and single crystal X-ray diffraction. *CrystEngComm* **2014**, *16*, 1232–1236.
- (48) Cheng, H.; Lu, C.; Liu, J.; Yan, Y.; Han, X.; Jin, H.; Wang, Y.; Liu, Y.; Wu, C. Synchrotron radiation X-ray powder diffraction techniques applied in hydrogen storage materials-A review. *Prog. Nat. Sci.* **2017**, *27*, 66–73.
- (49) Huang, B.; Frapper, G. Pressure-Induced Polymerization of CO<sub>2</sub> in Lithium–Carbon Dioxide Phases. *J. Am. Chem. Soc.* **2018**, *140*, 413–422.
- (50) Falls, Z.; Avery, P.; Wang, X.; Hilleke, K. P.; Zurek, E. The XtalOpt evolutionary algorithm for crystal structure prediction. *J. Phys. Chem. C* **2020**, *125*, 1601–1620.
- (51) Wang, B.; Guégan, F.; Frapper, G. Putting xenon and nitrogen under pressure: towards new layered and two-dimensional nitrogen allotropes with crown ether-like nanopores. *J. Mater. Chem. C* **2022**, *10*, 10374–10381.
- (52) Chen, D.; Wang, Y.; Dronskowski, R. Computational design and theoretical properties of WC<sub>3</sub>N<sub>6</sub>, an H-free melamine and potential multifunctional material. *J. Am. Chem. Soc.* **2023**, *145*, 6986–6993.
- (53) Avery, P.; Toher, C.; Curtarolo, S.; Zurek, E. XtalOpt Version r12: An open-source evolutionary algorithm for crystal structure prediction. *Comput. Phys. Commun.* **2019**, *237*, 274–275.
- (54) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Guntentor, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **2013**, *68*, 314–319.

- (55) Fredericks, S.; Parrish, K.; Sayre, D.; Zhu, Q. PyXtal: A Python library for crystal structure generation and symmetry analysis. *Comput. Phys. Commun.* **2021**, *261*, 107810.
- (56) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169.
- (57) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.
- (58) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (59) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (60) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (61) Klimeš, J.; Bowler, D. R.; Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys. Condens. Matter.* **2009**, *22*, 022201.
- (62) Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **1994**, *50*, 17953.
- (63) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13*, 5188.
- (64) Togo, A.; Oba, F.; Tanaka, I. First-principles calculations of the ferroelastic transition between rutile-type and CaCl<sub>2</sub>-type SiO<sub>2</sub> at high pressures. *Phys. Rev. B* **2008**, *78*, 134106.
- (65) Avery, P.; Zurek, E. RandSpg: An open-source program for generating atomistic crystal structures with specific spacegroups. *Comput. Phys. Commun.* **2017**, *213*, 208–216.

- (66) Lonie, D. C.; Zurek, E. Identifying duplicate crystal structures: XtalComp, an open-source solution. *Comput. Phys. Commun.* **2012**, *183*, 690–697.
- (67) Mattson, W. D.; Sanchez-Portal, D.; Chiesa, S.; Martin, R. M. Prediction of new phases of nitrogen at high pressure from first-principles simulations. *Phys. Rev. Lett.* **2004**, *93*, 125501.
- (68) Tomasino, D.; Kim, M.; Smith, J.; Yoo, C.-S. Pressure-induced symmetry-lowering transition in dense nitrogen to layered polymeric nitrogen (LP-N) with colossal Raman intensity. *Phys. Rev. Lett.* **2014**, *113*, 205502.
- (69) Laniel, D.; Geneste, G.; Weck, G.; Mezouar, M.; Loubeyre, P. Hexagonal layered polymeric nitrogen phase synthesized near 250 GPa. *Phys. Rev. Lett.* **2019**, *122*, 066001.
- (70) Laniel, D.; Winkler, B.; Fedotenko, T.; Pakhomova, A.; Chariton, S.; Milman, V.; Prakapenka, V.; Dubrovinsky, L.; Dubrovinskaia, N. High-pressure polymeric nitrogen allotrope with the black phosphorus structure. *Phys. Rev. Lett.* **2020**, *124*, 216001.
- (71) Wang, X.; Tian, F.; Wang, L.; Cui, T.; Liu, B.; Zou, G. Structural stability of polymeric nitrogen: A first-principles investigation. *J. Chem. Phys.* **2010**, *132*, 024502.
- (72) Eremets, M. I.; Gavriluk, A. G.; Trojan, I. A.; Dzivenko, D. A.; Boehler, R. Single-bonded cubic form of nitrogen. *Nat. Mater.* **2004**, *3*, 558–563.
- (73) Sun, J.; Martinez-Canales, M.; Klug, D. D.; Pickard, C. J.; Needs, R. J. Stable all-nitrogen metallic salt at terapascal pressures. *Phys. Rev. Lett.* **2013**, *111*, 175502.
- (74) Huang, B.; Wang, B.; Wu, S.; Guégan, F.; Hu, W.; Frapper, G. Predicted polymeric and layered covalent networks in transition metal pentazolate  $M(\text{cyclo-N}_5)_x$  phases at ambient and high pressure: up to 20 nitrogen atoms per metal. *Chem. Mater.* **2021**, *33*, 5298–5307.
- (75) Wang, B.; Larhlimi, R.; Valencia, H.; Guégan, F.; Frapper, G. Prediction of novel tin nitride  $\text{Sn}_x\text{N}_y$  phases under pressure. *J. Phys. Chem. C* **2020**, *124*, 8080–8093.

- (76) Greschner, M. J.; Zhang, M.; Majumdar, A.; Liu, H.; Peng, F.; Tse, J. S.; Yao, Y. A new allotrope of nitrogen as high-energy density material. *J. Phys. Chem. A* **2016**, *120*, 2920–2925.
- (77) Hirshberg, B.; Gerber, R. B.; Krylov, A. I. Calculations predict a stable molecular crystal of N<sub>8</sub>. *Nat. Chem.* **2014**, *6*, 52–56.
- (78) Steele, B. A.; Oleynik, I. I. Novel potassium polynitrides at high pressures. *J. Phys. Chem. A* **2017**, *121*, 8955–8961.
- (79) Trail, J.; Monserrat, B.; Ríos, P. L.; Maezono, R.; Needs, R. J. Quantum Monte Carlo study of the energetics of the rutile, anatase, brookite, and columbite TiO<sub>2</sub> polymorphs. *Phys. Rev. B* **2017**, *95*, 121108.
- (80) Guo, Q.; Zhou, C.; Ma, Z.; Yang, X. Fundamentals of TiO<sub>2</sub> photocatalysis: concepts, mechanisms, and challenges. *Adv. Mater.* **2019**, *31*, 1901997.
- (81) Woodley, S.; Catlow, C. Structure prediction of titania phases: Implementation of Darwinian versus Lamarckian concepts in an evolutionary algorithm. *Comput. Mater. Sci.* **2009**, *45*, 84–95.
- (82) Polsin, D. N.; Lazicki, A.; Gong, X.; Burns, S. J.; Coppari, F.; Hansen, L. E.; Henderson, B. J.; Huff, M. F.; McMahon, M. I.; Millot, M. et al. Structural complexity in ramp-compressed sodium to 480 GPa. *Nat. Commun.* **2022**, *13*, 2534.
- (83) Wang, B.; Hilleke, K. P.; Wang, X.; Polsin, D. N.; Zurek, E. Topological electride phase of sodium at high pressures and temperatures. *Phys. Rev. B* **2023**, *107*, 184101.
- (84) Pena-Alvarez, M.; Binns, J.; Martinez-Canales, M.; Monserrat, B.; Ackland, G. J.; Dalladay-Simpson, P.; Howie, R. T.; Pickard, C. J.; Gregoryanz, E. Synthesis of Weaire–Phelan barium polyhydride. *J. Phys. Chem. Lett.* **2021**, *12*, 4910–4916.

- (85) Bi, T.; Zurek, E. Electronic structure and superconductivity of compressed metal tetrahydrides. *Chem. Eur. J.* **2021**, *27*, 14848–14870.
- (86) Hilleke, K. P.; Wang, X.; Luo, D.; Geng, N.; Wang, B.; Zurek, E. Structure, stability and superconductivity of N-doped lutetium hydrides at kbar pressures. *arXiv preprint arXiv:2303.15622* **2023**,
- (87) Tang, X.; Xiang, H.; Liu, X.; Speldrich, M.; Dronskowski, R. A ferromagnetic carbodiimide:  $\text{Cr}_2(\text{NCN})_3$ . *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 4738–4742.
- (88) *Center for Computational Research, University at Buffalo.*  
<http://hdl.handle.net/10477/79221>.

For Table of Contents Only

