JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. 2018, 140, 9844−9853
pubs.acs.org/JACS

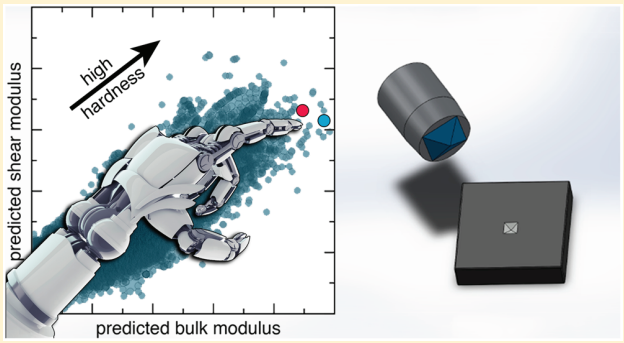
Machine Learning Directed Search for Ultraincompressible, Superhard Materials

Aria Mansouri Tehrani,[†,⊥] Anton O. Oliynyk,[†,⊥] Marcus Parry,[‡] Zeshan Rizvi,[†] Samantha Couper,[§] Feng Lin,[§] Lowell Miyagi,[§] Taylor D. Sparks,[‡] and Jakoah Brgoch*[,†]

[†]Department of Chemistry, University of Houston, Houston, Texas 77204, United States
[‡]Department of Materials Science and Engineering, University of Utah, Salt Lake City, Utah 84112, United States
[§]Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: In the pursuit of materials with exceptional mechanical properties, a machine-learning model is developed to direct the synthetic efforts toward compounds with high hardness by predicting the elastic moduli as a proxy. This approach screens 118 287 compounds compiled in crystal structure databases for the materials with the highest bulk and shear moduli determined by support vector machine regression. Following these models, a ternary rhenium tungsten carbide and a quaternary molybdenum tungsten borocarbide are selected and synthesized at ambient pressure. High-pressure diamond anvil cell measurements corroborate the machine-learning prediction of the bulk modulus with less than 10% error, as well as confirm the ultraincompressible nature of both compounds. Subsequent Vickers microhardness measurements reveal that each compound also has an extremely high hardness exceeding the superhard threshold of 40 GPa at low loads (0.49 N). These results show the effectiveness of materials development through state-of-the-art machine-learning techniques by identifying functional inorganic materials.



INTRODUCTION

Accelerating the discovery of next-generation, functional inorganic materials is currently centered on predicting a material's properties *a priori* using quantum chemical calculations.[1−5] By determining a compound's electronic and vibrational structure, it is possible to estimate a myriad of physical properties.[6] This has inspired large-scale, high-throughput first-principles density functional theory (DFT) calculations through frameworks such as The Materials Project,[7] AFLOW,[8] OQMD,[9] MPDS, and JARVIS-DFT.[10] These research programs have not only facilitated the generation of vast DFT-level data sets but also made the results readily accessible. Despite being a great resource, there are still limitations to using DFT to predict some properties. For example, band gap cannot be accurately reproduced by DFT with commonly used exchange−correlation functionals such as Perdew, Becke, and Ernzerhof (PBE), and implementation of hybrid functionals and Green's function (GW) methods is not yet feasible within a high-throughput framework.[11−13] Moreover, stress−strain or phonon calculations are computationally expensive, inhibiting large-scale calculations of elastic constants. Finally, DFT cannot easily account for atomic disorder such as site sharing, which is common in inorganic solids.

A potential pathway to shift the paradigm of computer-directed materials discovery is to employ data-driven approaches such as machine learning (ML). The implementation of ML in materials chemistry, however, remains in its infancy, with only a limited number of successful applications reported. For example, Support Vector Machine (SVM) models have been used to classify the crystal structures of transition metal phosphides[14] as well as predict equiatomic binary compounds leading to the discovery of a novel phase.[15] Likewise, a hybrid high-throughput first-principles computation and ML technique based on a probabilistic model was developed to identify stable novel compositions and their crystal structures.[16] Physical properties such as band gap,[17,18] elastic constants,[17,19] and Debye temperature[17] have also been predicted using an array of ML techniques. ML has even been employed for designing efficient organic light-emitting diodes by rapidly screening and fabricating efficient emitter materials.[20] Similarly, utilizing a machine-learning approach through balancing model prediction and uncertainty has led researchers to discover a lead-free $BaTiO_3$-based piezoelectric.[21] The development of new thermoelectric materials is another example that illustrates the use of machine learning

Received: March 9, 2018
Published: July 16, 2018

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

in materials science.[22,23] Each of these research achievements highlights the potential of ML in materials development as an emergent method.

One area where ML may also be useful is in the search for materials with exceptional mechanical properties, such as high incompressibility or extreme hardness. Traditionally, the search for new superhard materials, which by convention have a Vickers hardness (H_V) > 40 GPa,[24] has relied on trial-and-error methods or simple design rules.[25] These researches have largely concentrated on discovering materials that form strong covalent bonds using light main group elements as in diamond, *c*-BN, B_6O, and *c*-BC_2N or identifying compounds that combine light elements and transition metals with high valence electron density like in ReB_2 and WB_4.[26−28] While the first class benefits from relying on abundant, cheap elements, their synthesis requires extreme pressures and temperatures, making their preparation cumbersome. The second type of compounds employs conventional metallurgical techniques but often includes expensive and scarce transition metals.[25] To aid the discovery of new superhard materials in either class, attempts have been made to transcend these simple design rules and screen materials by computationally predicting hardness. However, DFT cannot directly calculate hardness because it is governed by surface plastic deformation upon indentation and, therefore, spans multiple lengths scales ranging from local atomic structure to bulk microstructure. Nevertheless, researchers have developed empirical models of hardness within DFT-based genetic algorithms that have suggested the existence of numerous high-hardness materials. For example, compounds such as ReB_4,[29] B_4C_4,[30] and (*t*,*m*,*o*)-Si_3N_4[31] are all predicted as potential superhard materials. Further, *t*-SiCN[32] is predicted to be a thermodynamically stable compound with a hardness of 41 GPa, while a novel CN phase with sp^3-hybridization is predicted to have a hardness of 62 GPa, although its synthesis requires an applied pressure of 10.9 GPa, limiting large-scale preparation. Unfortunately, these genetic algorithm-based predictions are computationally time-intensive, the versatility of the hardness models are mainly limited to strictly covalent compounds, and they are generally centered on discovering high-pressure phases. Of greater concern is that very few of these predicted materials have ever been experimentally verified.

Given the obvious difficulty in using computation to explicitly predict hardness, an alternative route is to utilize a proxy, which is a distinct, quantifiable material property that scales with hardness. Elastic moduli such as the bulk modulus (B) and shear modulus (G) are both reported to correlate with hardness[33,34] and could therefore be potential proxies. More importantly, B and G are both intrinsic material properties that can be reliably estimated through first-principles calculations. In fact, DFT provides great accuracy in predicting elastic constants. This has allowed high-throughput programs such as the Materials Project to implement density functional perturbation theory (DFPT) to calculate elastic constants.[35] Although this approach has, to date, calculated the elastic constants for more than 6000 ordered, crystalline materials, this is <10% of all reported inorganic, solid-state compounds.

Here, we develop a method based on machine learning to vastly expand the number of materials with their elastic moduli predicted. This approach employs a combination of compositional and structural descriptors to build a ML model using the Materials Project data as a training set. Our approach is capable of predicting B and G for ∼120 000 binary, ternary, and quaternary inorganic solids that are compiled in Pearson's Crystal Database[36] regardless of their unit cell size, atomic disorder, or high electron correlation. We then use both the ML-predicted B and G values as a proxy to identify compositions with potential high hardness, leading us to synthesize $ReWC_2$ and $Mo_{0.9}W_{1.1}BC$. According to our model, these compounds have the highest expected B and G for a ternary and quaternary phase, respectively. More importantly, these compositions can be prepared at ambient pressure using arc melting, making them ideal prospective ultraincompressible, high-hardness materials. This unprecedented ML-directed search for new superhard materials strays from the traditional design rules and shows the potential to significantly advance the rate of new functional materials development using machine learning.

EXPERIMENTAL METHODS

Data Extraction and Machine-Learning Training Model. To create the training set, 3248 elastic moduli (B and G) were extracted from the Materials Project database.[37] The Materials Project implements a high-throughput framework based on DFPT[7] using the Vienna *ab Initio* Simulation Package (VASP)[38] with the PBE functional[39] used for exchange and correlation, while the DFT+U (Hubbard parameter) scheme is employed for highly correlated systems.[40] The elastic constants are determined using DFPT coupled with the Voigt−Reuss−Hill approximation to convert the elastic tensors into the elastic moduli.[41]

The compositions extracted from the Materials Project were first cross-referenced against an inorganic crystal structure database, Pearson's Crystal Database (PCD), to ensure only phases accessible at ambient pressure and temperature were used in the machine-learning model. Additionally, the data extracted from Materials Project were manually checked for unreasonable values, e.g., negative moduli. Only binary, ternary, and quaternary compositions were considered, and compounds containing group 18 (noble gases) were excluded, as were any phases with the elements hydrogen, Tc, and Z > 83 (except for U and Th). Any crystal structures reported as thin films, foils, or suspensions were not included. These criteria reduced the final training data set to 2572 elastic moduli. To distinguish between polymorphs present in the training set, each compound was labeled as "composition, space group number". For example, PtV adopts both an orthorhombic (space group number 51) crystal structure and a tetragonal (space group number 123) crystal structure, which we label accordingly as PtV,51 and PtV,123.

The machine-learning model was next created based on a Support Vector Machine Regression (SVR)[42,43] algorithm using PLS_Toolbox Version 8.2.1 (Eigenvector Research Inc., Wenatchee, WA, USA) within the MATLAB 2016a (The Mathworks, Natick, MA, USA) environment. The SVR employed a radial basis function (RBF) as the kernel function and was trained with a 10-fold cross-validation scheme. The cross-validation method helps to avoid overfitting and artificially high statistics. The descriptors used in this study included 34 distinct compositional variables describing the elemental properties such as position on the periodic table, electronic structure, and physical properties as well as their associated math expressions (difference, average, largest value, and smallest value). Additionally, 14 structure descriptors related to variables including crystal system, space group, and unit cell volume, among others, were incorporated in the ML model. The full list of descriptors and mathematical expressions is available in the Supporting Information (Figure S1). In total, 150 descriptors were used to build the ML model constituting a 2572 × 150 matrix, which was normalized, mean-centered, and rescaled to the unit variance. A genetic algorithm-based variable selection was then performed with partial least-squares regression choice in 25 Latent Value space with a population size of 256, while the maximum number of generations was set at 100. A 20-fold random split cross-validation was used for the variable selection. The variable importance on projection of the full descriptor set as well

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

as the subset of descriptors selected by the genetic algorithm is discussed below and included in the Supporting Information Figure S1. All calculations were performed using an Intel Corei5-4690K CPU on a Windows PC.

After finalizing the machine learning training model, 118 287 compounds available in the PCD that follow our general composition criteria, discussed above, were assembled in a database. Their elastic moduli, bulk and shear, were predicted using ML, while Young's modulus and Poisson's ratio were calculated from these values. The data for 258 compounds with the highest B and G are provided in Table S1 of the Supporting Information.

Synthesis and Structure Determination. Samples of $ReWC_x$ ($0.8 \leq x \leq 2$) and $Mo_{0.9}W_{1.1}BC$ were prepared by arc melting the elements using a water-cooled copper hearth under an argon flow atmosphere. The starting materials, Re (Alfa Aesar, 99.997%), W (Alfa Aesar, 99.5%), C (Sigma-Aldrich, ≥99.99%), Mo (Alfa Aesar, 99.95%), and crystalline B (Alfa Aesar, 99.5%), were weighed out in the desired stoichiometry ratios (total mass ≈ 0.4 g), pressed into pellets, and then arc melted using a current of 80 A until homogeneous melting occurred. The products were melted at least twice including flipping the ingot to ensure mixing of the elements. The weight loss after arc melting was <5%, attributed mostly to volatilization of graphite during the arc melting process. To confirm the products were fully dense, density was measured using a Mettler Toledo density kit (MS-DNY-43) and compared to the X-ray-determined density.

The ingots were then split in two, with one-half ground into a fine powder using a Diamonite mortar and pestle and analyzed for phase purity using powder X-ray diffraction on a PanAnalytcal X'Pert powder diffractometer equipped with Cu Kα radiation (λ = 1.54059 Å). The crystal structures of the products were confirmed by whole diffraction pattern fitting using the LeBail method. The backscattered electron microscopy images were collected for $ReWC_{0.8}$, ReWC, and $Mo_{0.9}W_{1.1}BC$ using a JEOL JSM 6400 scanning electron microscope (SEM). A single crystal was also selected from the $ReWC_{0.8}$ and $Mo_{0.9}W_{1.1}BC$ arc-melted samples and analyzed using a Bruker PLATFORM single-crystal X-ray diffractometer equipped with a SMART APEX II CCD area detector using graphite-monochromated Mo Kα radiation (λ = 0.710 73 Å) at room temperature using ω scans at eight different ϕ angles with a 0.30° frame width and an exposure time of 12 s per frame. The initial structural model, utilized previously, reported and standardized atomic positions from $ReWC_2$ and $Mo_{0.9}W_{1.1}BC$ structures, while the structure refinement was performed using the SHELXL (version 6.12)[44] software package. Single-crystal refinement data are available in Tables S2, S3, and S4 of the Supporting Information. All crystal structure drawings were produced using the program VESTA.[45]

Mechanical Property Measurements. A portion of the $Mo_{0.9}W_{1.1}BC$ and $ReWC_{0.8}$ samples analyzed using laboratory X-ray powder diffraction were examined under compression to measure the equation of state (EoS) to compare the experimental and machine-learning-predicted bulk modulus. The powdered sample was mixed with a small amount of platinum powder as a pressure calibrant.[46] High-pressure conditions were generated using a symmetric diamond anvil cell (DAC) with 200 μm diameter flat cutlet diamond anvils. A stainless steel gasket (initial thickness of 250 μm) was preindented to a thickness of 40 μm, and a 100 μm diameter sample chamber was laser milled into the center of the indentation. The sample/Pt powder mixtures were gas loaded into the sample chamber with a neon pressure medium.

Samples were compressed to ∼50 GPa in 1−3 GPa increments at ambient temperature. Angle-dispersive synchrotron X-ray diffraction spectra were collected in axial geometry at beamline 12.2.2 of the Advanced Light Source (ALS) of Lawrence Berkeley National Laboratory using a mar3450 image plate and monochromatic X-rays (λ = 0.4983 Å). Sample to detector distance (330.5 mm), detector tilt, and detector rotation were calibrated with a cerium dioxide standard. Diffraction data were analyzed for unit cell parameters using the Le Bail method as implemented in the software package MAUD (Materials Analysis Using Diffraction).[47] Lattice strain was not refined, as the measurements were near hydrostatic. Relative cell volume vs pressure curves were fit to a Birch−Murnaghan EoS using EosFit to determine bulk modulus (B_0) and the first pressure derivative of the bulk modulus (B_0').[48]

The other half of the arc-melted ingots were used for Vickers microindentation measurements. The samples were mounted in epoxy resin and then polished using SiC (600 to 1200 grit) followed by a 1 μm diamond slurry. The Vickers hardness was measured using microindentation (Leitz instrument) at loads of 0.49, 0.98, 1.96, 2.94, and 4.90 N, calibrating the instrument for hardness testing with external standards (Wilson test block). The indentations were then analyzed using a VHX-600 digital microscope at 1000× magnification and the VHX digital microscope image analysis software. A minimum of 10 indentations with a dwell time of 10 s, following ASTM Designation E384-17,[49] were made and analyzed at each load to ensure reliable statistics.

RESULTS AND DISCUSSION

Machine-Learning Bulk and Shear Modulus. Machine learning of elastic moduli begins with the construction of a robust training set. In this case, the Materials Project DFT calculated B and G are ideal to train the SVR machine-learning model.[42,43] Mathematically, the training data are in the form $\{(x_1, y_1), ..., (x_n, y_n)\} \subset \mathbb{R}^d \times \mathbb{R}$, where n = 2572 and d = 150 corresponding to the number of compounds and the number of descriptors in the feature set, respectively. The objective of SVR is to find a function, $f(x)$, that contains all the training data within the bound of $\pm\epsilon$, which is 0.1 in this model. The linear SVR is then implemented following $f(x) = \langle w, x\rangle + b$, where w and b are the variables that can be optimized and $w \in \mathbb{R}^d$, $b \in \mathbb{R}$, and $\langle w, x\rangle$ is the dot product in the input (descriptor) space. To acquire flatness, which is necessary for a generalized solution, $\|w\|$ should be minimized and slack parameters (ξ) are introduced to improve the optimization and allow error in the data. For the ϵ-insensitive SVR model employed here, the loss function is defined as zero for $|\xi| \leq \epsilon$ and $|\xi| - \epsilon$ for $|\xi| \geq \epsilon$. A constant, C, is introduced to control the compromise between the flatness and the influence of ξ. In these ML models, our cost (C) and gamma (γ) have optimized values of 10 and 0.01, respectively, where γ is the kernel parameter governing the influence of each support vector. The optimization problem is solved by constructing the Lagrange function $f(x) = \sum_{i=1}^{2572}(\alpha_i - \alpha_i^*)\langle x_i, x\rangle + b$, where α_i and α_i^* are Lagrange multipliers and w is described as a linear combination of x_i. However, this formalism is essentially a nonlinear SVR that must be transformed into a feature space where linear SVR can be applied. This is straightforwardly achieved using a kernel function, $\Phi(x, x') = \langle \phi(x), \phi(x')\rangle$, which in this model is an RBF kernel function.

Following this SVR method, the B and G moduli were trained independently using 2572 moduli. As shown in Figure 1a, remarkable agreement is obtained between the DFT-calculated bulk modulus (B_{DFT}) and the ML-predicted values (B_{SVR}) with the cross-validated root-mean-square error (RMSECV) and coefficient of determination (r^2) being 17.2 GPa and 0.94, respectively. Examining the linear fit shows a slight overestimation of the B_{SVR} compared to the B_{DFT} at low values (<75 GPa) and a slight underestimation at high values (>250 GPa) (Figure 1a). Decomposing the error in Figure 1b reveals the high accuracy of our ML model with more than 86% of the compounds' B_{SVR} differing by less than ±20% from the DFT-calculated values (typically, DFT values correlate with experimental values within ±15% accuracy).[50] The best agreement is obtained for intermetallics such as HfVGe,139

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

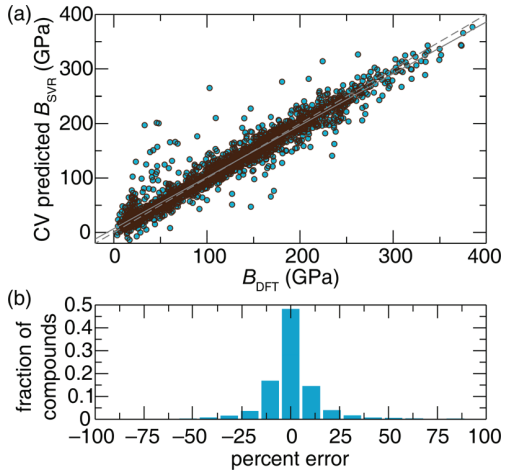


Figure 1. (a) Cross-validated values of B using the SVR model against the training set (B_{DFT}). The dashed line is the ideal 1:1 ratio, whereas the solid line is the fit to the predicted values. The deviation between the dashed line and the solid line shows a slight overestimation at low values in contrast to the slight underestimation at high values. (b) Fraction of compounds predicted within the relative percent error for the cross-validated B model.

and Zr_2Al,140 with negligible error. In fact, 308 compounds are predicted with a cross-validated error of less than 1%. A few compounds are exceedingly inaccurate, such as HgO,62, a polymorph of SiO_2,33, and BiIO,129, which have B_{SVR} values that differ by 451%, 411%, and 180%, respectively, relative to B_{DFT}. Overall, better agreement is achieved for compounds with metallic/covalent bond character, and the error increases for highly ionic compounds. In fact, the compounds generally missed by our model tend to have a very low bulk modulus and largely appear to be generally oxides and halides. Calculation of the statistical error while exclusively considering oxides (highlighted in Figure S2a of the Supporting Information) indicates the B_{SVR} relative error is only ~15% compared to B_{DFT}, only slightly higher than that of the full data set. Overall, the bulk modulus SVR model shows excellent statistics and predictive power.

Examination of the variables selected by the genetic algorithm, listed in Figure S1 of the Supporting Information, provides insight into the crystal−chemical properties that control B_{SVR}. Overall, only 52% of the variables survived after feature selection and are essential to predict the bulk modulus. In general, position on the periodic table accounts for the largest number of variables included in the final model, even though the variable importance projection (VIP) score from a partial least-squares regression model suggests these variables do not carry significant weight. Among the size variables, the covalent and crystal radius strongly affected the model, while the electron count shows two variables for the metallic valence (largest and smallest values of constituent elements) as well as all four variables for the number of p and d electrons are important for prediction of B_{SVR}. The only physical properties that significantly influence the bulk modulus are density, cohesive energy, and polarizability. The cohesive energy, in particular, stands out among the compositional variables with all four mathematical expressions dominating the VIP score. Regarding the structural variables, unit cell density, volume per atom, Gilman electron density (Gilman scale[34] postulates an order for counting valence electrons in transition metals to yield the best correlation with bulk modulus), and outer shell electron density dictate B_{SVR}. The compressibility of a material is clearly governed by values relating to the compound's various measures of bond strength and its density. Although many of the conclusions drawn from this analysis are unsurprising, they provide a quantitative foundation for classical perceptions regarding compressibility in materials.

Shear modulus is notoriously more difficult than bulk modulus to calculate and to measure experimentally. The ML model developed here exhibits similar difficulty in accurately determining the shear modulus. Figure 2a illustrates the cross-validated G_{SVR}, which has an RMSECV of 16.5 GPa and a cross-validated r^2 of 0.84. The plotted data clearly show that G_{SVR} is underestimated above 200 GPa, which is potentially due to the lack of data with an extremely high shear modulus accentuated by the fact that only two data points, stemming from boron nitride, are present above 300 GPa. This imbalance in the training data leads to imbalance in the predictions as well a bias toward lower predicted values. Adding more data with high shear modulus could potentially curtail this issue.

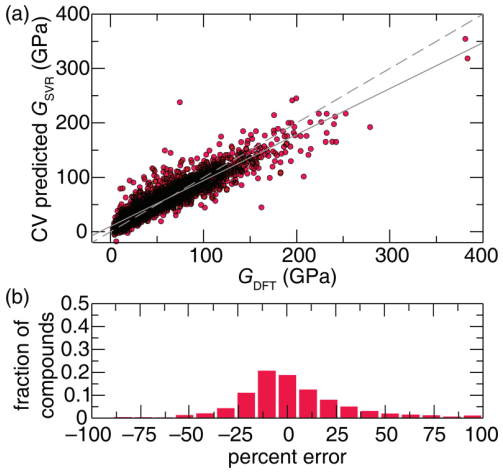


Figure 2. (a) Cross-validated values of G using the SVR model against the training set (G_{DFT}). The dashed line is the ideal 1:1 ratio, whereas the solid line is fit to the predicted values. The deviation between the dashed line and the solid line shows a slight overestimation at low values in contrast to the significant underestimation at high values. (b) Fraction of compounds predicted within the relative percent error for the cross-validated G model.

Nevertheless, similar to B_{SVR}, the G_{SVR} model provides accurate predictions for many intermetallics. For example, MgSbPt,216 and Y_3Al,221 are predicted exactly. Moreover, 87 compounds are predicted with a relative error of less than 1%. From Figure 2b it is discernible that around 61% of compounds are within the 20% relative error range for the G model. Contrary to B, the errors here are not dominated by oxides (Supporting Information Figure S2b). Instead, compounds with rather diverse chemistries are predicted erroneously such as Ti_3Ir,223, TI_2PdCl_4,123, HfS_3,11, and NaVSe,166 with the cross-validated relative errors of 511%, 386%, 230%, and 206%, respectively.

The genetic algorithm suggests that only 44% of all variables are necessary to accurately predict the shear modulus. Similar to the bulk modulus, G_{SVR} does not have a particular set of compositional variables that regulate the shear modulus. The atomic, covalent, and crystal radii all slightly influence the

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

model according to their VIP scores, whereas pseudopotential radius is not considered whatsoever. Neither the electronegativity scales nor their derivatives are particularly highlighted by the genetic algorithm. The metallic valence and the number of valence electrons all show significant importance to the model, while there is less emphasis on the number of p or d electrons. Analogous to B_{SVR}, the physical properties are not prominent in the G_{SVR} model with the exception of the largest heat of fusion as well as the sum and difference of the cohesive energy. The importance of structural variables for G_{SVR} is similar to those for the bulk modulus; only the valence electron density variable was additionally selected for shear modulus prediction. The volume per atom, unit cell density, and Gilman electron density remain the most important structural variables owing to their close relationship to a material's response under stress.

The statistics of our models are comparable with the recent research using 5-fold cross-validated prediction of bulk and shear using a gradient boosting decision tree technique. In this model, a unique set of 2494 universal property-labeled fragment descriptors was employed, leading to a cross-validated, root mean squared error (RMSECV) value of 14.3 GPa and an r^2 value of 0.97 for the bulk modulus and an RMSECV of 18.4 GPa and an r^2 of 0.88 for the shear modulus.[17] Another study uses gradient boosting machine local polynomial regression for modeling B and G that only implements 17 descriptors; however, some of the descriptors in this approach include a material's properties, such as band gap and formation energy per atom, that require a prior DFT calculation.[19] Nevertheless, most of the ML approaches developed thus far, irrespective of ML method or choice of descriptors, are able to predict elastic moduli with impressive accuracy in a fraction of the time that it currently takes *ab initio* calculations.

Screening Crystal Structure Databases for Ultraincompressible, High-Hardness Materials. The ultimate goal of ML is to employ a training set that is capable of predicting the elastic moduli for entire crystallographic databases, which can then act as a proxy to direct the search for materials with desired mechanical properties. Here, our model is employed to predict the B and G of 118 287 inorganic compounds contained within the PCD, including 15 770 binaries, 56 266 ternaries, and 46 251 quaternaries. Importantly, of these compounds, 27 698 (23%) contain rare-earth elements that cannot be readily calculated using conventional DFT and 19 384 (16%) have either disordered positions or site mixing, which is also generally inaccessible with DFT. Since our machine-learning approach is based on composition and crystal structure, the effect of disorder and mixing is indirectly accounted for by descriptors such as "electron density" and "volume per atom" of the crystal structure, although specific local coordination environments are not captured.

The predicted bulk and shear moduli values are illustrated in Figure 3, plotted as G_{SVR} vs B_{SVR}. The darker regions occur where there are overlapping data and indicate that most of the B_{SVR} values are concentrated between 1 and 180 GPa and G_{SVR} values are concentrated between 1 and 90 GPa. It is discernible that the plotted distribution mostly follows a rough linear trend, meaning that it is unlikely for a material with a high bulk modulus to possess a low shear modulus and *vice versa*. The cluster of compounds at $B_{SVR} \approx 350$ GPa and $G_{SVR} \approx 350$ GPa, which are conspicuously isolated from the rest of the materials calculated, are BN, underscoring their industrial dominance as

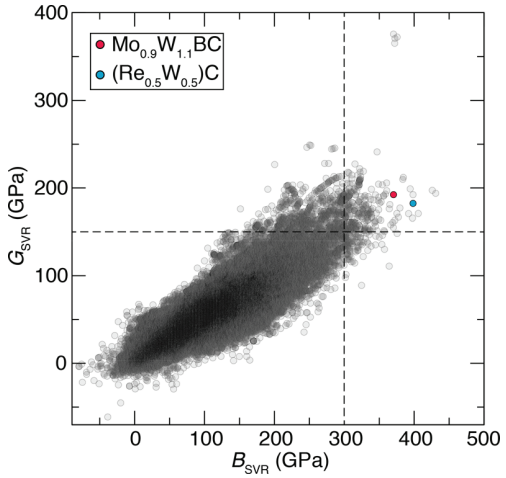


Figure 3. Plotting the SVR-predicted G and B moduli of 118 288 inorganic compounds derived from PCD. The vertical dashed line separates the compounds with potential high B, while the horizontal dashed line separates the compounds with a potential high G modulus. The compounds encompassed in the right top corner are potentially hard or superhard. The red and blue circles correspond to $Mo_{0.9}W_{1.1}BC$ and $Re_{0.5}W_{0.5}C$, respectively, which are selected for subsequent synthesis and characterization.

high-hardness materials. It can also be noted that a small portion (~0.5%) of compounds are predicted with negative values, mostly pertaining to regression errors in the SVR model. Many of these materials predicted to have negative G_{SVR} and/or B_{SVR} contain group 17 elements such as Cl_3Ga,2, I_3Nb,59, Cl_4,121, BrHgI,36, and $BrCCl_3$,225, where the chemistry as discussed above is not entirely captured by our model.

Screening for materials with superior mechanical properties, hardness in particular, is now practical using this two-dimensional plot. First, only compounds with $B_{SVR} \geq 300$ GPa, shown by the horizontal dashed line in Figure 3, are considered as potential ultraincompressible materials. This high cutoff instantly reduces the number of candidate materials from 118 287 to only 359 phases. However, this criterion is not sufficient while screening for hard materials considering that the bulk modulus has only a minor correlation with hardness. Although it is well known that all superhard materials are ultraincompressible (high bulk modulus), the reverse is clearly not valid; for example, elemental Os has an extremely high bulk modulus (~400 GPa)[51] but a low hardness (~4 GPa)[52] due to its metallic character, leading to negligible resistance to the generation and movement of dislocations. Shear modulus, G, demonstrates a more substantial relationship with hardness. A high shear modulus is essential to hardness because it is correlated with the generation of dislocation loops (e.g., Frank–Read sources) and a Peierls–Nabarro stress barrier, which regulate the movement of dislocations in a unit cell. Since dislocation glide is the main contributor to plastic deformation in ductile materials, shear modulus is associated with hardness. Thus, we only considered compositions with a $G_{SVR} \geq 150$ GPa, as highlighted by the vertical dashed line in Figure 3. This leaves only 258 potential ultraincompressible, hard compounds in our screening database (Table S1 in the Supporting Information). Examining the chemistry of these materials indicates that a majority of these compounds are carbides, borides, and nitrides, as expected for hard materials.

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844–9853

In fact, most of the noteworthy binary, high-hardness materials occur in this region, including WB_4,194 (B_{SVR} = 315, G_{SVR} = 227), WC,187 (B_{SVR} = 384, G_{SVR} = 212), TaC,225 (B_{SVR} = 323, G_{SVR} = 210), and ReB_2,194 (B_{SVR} = 330, G_{SVR} = 191). These binaries are already known hard/superhard materials and are predicted as potential materials with exceptional mechanical properties, e.g., superhardness, confirming the effectiveness of our approach. It is worth noting that our model and screening method is based on DFT-predicted elastic moduli and does not account for extrinsic hardening factors such as grain boundaries or secondary-phase hardening.

With the success of locating a majority of known high-hardness materials in the top right corner of Figure 2, our screening method was then used to identify unexplored phases with potential high hardness. Compound selection was achieved by sorting the compounds ($B_{SVR} \geq 300$ GPa; $G_{SVR} \geq 150$). Since most of the binary compounds predicted by ML as potential superhard materials were either already known materials or required high-pressure synthesis, we focused instead on identifying the best ternary and quaternary candidates. Among these compounds, with respect to the bulk modulus, $Re_{0.5}W_{0.5}C$,225 ranks first in ternaries (5th overall) and $Mo_{0.9}W_{1.1}BC$,63 ranks first in quaternaries (19th overall). Interestingly, $ReWC_2$ and $Mo_{0.9}W_{1.1}BC$ have been previously investigated as superconductors with transition temperatures of 3.8 K[53] and 4.2 K,[54] respectively. Moreover, both compounds can be synthesized at ambient pressure. Thus, we directed our search for ultraincompressible, high-hardness materials toward these two compositions.

Synthesis and Characterization of Two High-Hardness Materials. Starting with stoichiometric $ReWC_2$ ($Re_{0.5}W_{0.5}C$, as reported in PCD), the arc-melted samples showed unreacted graphite peaks in the powder X-ray diffraction. Therefore, subsequent synthesis was carried out with reduced carbon content following $ReWC_x$ ($0.8 \leq x \leq 2.0$). The unreacted graphite peak disappeared and a pure phase sample was achieved when $0.8 \leq x \leq 1.0$, as shown in Supporting Information Figure S3. Carbon deficiency in this range is not surprising considering carbon nonstoichiometry is rife in the literature of refractory transition metal carbides; for example, MC_x (M = Ti, Zr, Hf, V, Nb, and Ta) have been reported as carbon deficient, with the carbon content falling in the range $0.5 < x < 0.97$.[55] Since $ReWC_{0.8}$ had the least porous surface based on the SEM micrographs (Figure S4a and b in the Supporting Information), this composition was selected for further analysis. The refined powder X-ray diffraction pattern of $ReWC_{0.8}$ is shown in Figure 4a and indicates the product is phase pure, although the diffraction peaks appear surprisingly broad. Careful analysis of the diffraction pattern, especially at high angles, showed the presence of two peaks with slightly different d spacing. Refining the crystal structure using the Le Bail method required the incorporation of two phases with slightly different unit cell volumes (V = 68.427(2) Å and V = 68.985(2) Å). Interestingly, the SEM micrographs (Figure S4 in the Supporting Information) collected with backscattered electrons indicate a homogeneous composition across the sample. However, the difference in unit cell volumes likely arises from a phase width with the product having regions of slightly different compositions not captured by backscattered electron microscopy. Subsequent hardness measurements performed on both $ReWC_{0.8}$ and $ReWC_{1.0}$ samples show only a slight difference in hardness (Figure S5 in the Supporting Information) and show that slight variations in composition do not greatly influence the bulk mechanical properties. To further confirm the synthesized ingot is dense, the density was measured to be 17.4(4) g/cm^3, which in comparison to the density calculated based on X-ray diffraction refinement shows the sample is 93(2)% dense.

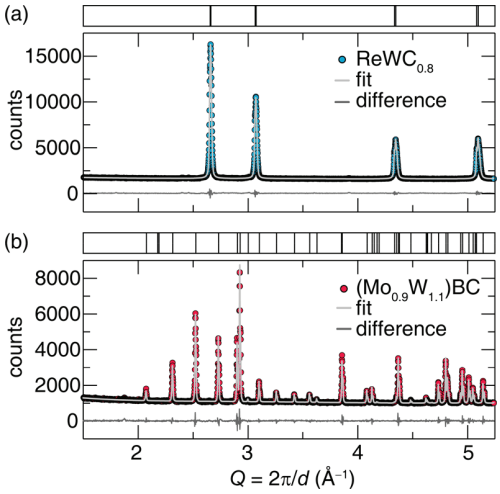


Figure 4. Le Bail refinement plot of (a) $ReWC_x$ and (b) $Mo_{0.9}W_{1.1}BC$ powder X-ray diffraction data. The experimental data are blue ($ReWC_x$) and red ($Mo_{0.9}W_{1.1}BC$); the Le Bail fit is in light gray, and the difference curve is in dark gray.

The $ReWC_{0.8}$ crystal structure as depicted in Figure 5a shows the compound adopts a NaCl-type ($Fm\bar{3}m$) crystal structure, with transition metals sites (Wyckoff 4a) equally occupied by Re and W, while the carbon site (Wyckoff 4b) is deficient. The carbon−transition metal bond distance is 2.0498(1) Å, which is shorter than the carbon−rhenium bond length (2.157(4) Å) in the high-pressure Re_2C phase.[56] The carbon−transition metal bond distance in $ReWC_{0.8}$ is also significantly shorter than the carbon−tungsten bond length (2.1989(11) Å) in WC.[57] These differences in bonding are

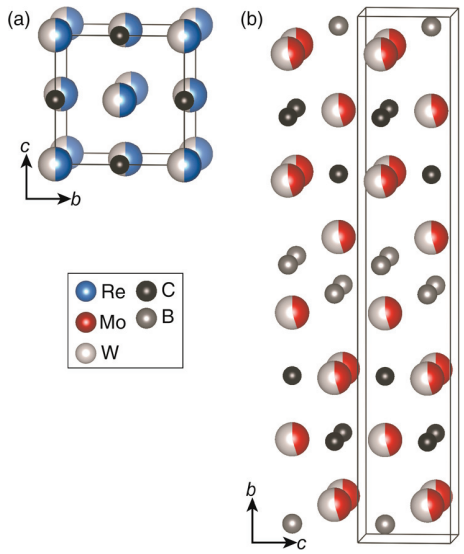


Figure 5. (a) Crystal structure of $ReWC_x$ with x = 2 and (b) crystal structure of $Mo_{0.9}W_{1.1}BC$. Re is blue, Mo is red, W is light gray, B is dark gray, and carbon is black.

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

likely the origin of the superior predicted intrinsic mechanical properties in $ReWC_{0.8}$ compared to Re_2C or WC.

The quaternary target compound, $Mo_{0.9}W_{1.1}BC$, was obtained as a single-phase product directly from arc melting, as illustrated by the powder X-ray diffraction pattern (Figure 4b). Density measurements indicate that the synthesized ingot is 95(3)% dense. The compound is isostructural to the parent phase Mo_2BC crystallizing in orthorhombic space group *Cmcm* (space group no. 63) with refined lattice parameters of a = 3.07593(3) Å, b = 17.3142(1) Å, and c = 3.04184(3) Å. The highly anisotropic unit cell is pseudotetragonal, with a long b-axis arising from an alternating stacking of boron-rich and carbon-rich planes. The crystal structure, illustrated in Figure 5b, contains infinite zigzag chains of boron atoms along the [001] direction with a B−B bond length of ∼1.83(1) Å, which is similar to the B−B bond length in ReB_2. The crystal structure also contains two crystallographically independent transition metal positions. One position (Wyckoff 4*c*) centers the basal plane of a square-based pyramidal geometry and is coordinated to five carbon atoms. The second transition metal (Wyckoff 4*c*) position centers a trigonal prismatic, square-face monocapped prism formed by six boron atoms and one carbon atom. These transition metal sites are a statistical mixing of both transition metals.

To confirm the SVR bulk modulus predictions of $ReWC_{0.8}$ and $Mo_{0.9}W_{1.1}BC$, high-pressure DAC diffraction experiments were performed to determine B_0. MAUD was used to refine the diffraction patterns, and the (111), (200), and (220) platinum diffraction peaks were used as an internal standard for pressure calibration. The pressure-dependent diffraction patterns for $ReWC_{0.8}$ (Figure 6a) and $Mo_{0.9}W_{1.1}BC$ (Figure 6b) both show compression with increasing pressure based on the shift of the peaks to large Q-space values, which is also supported by plotting the refined relative volume (V/V_0) as a function of pressure. The lattice parameters and associated error for the carbide, borocarbide, and platinum are provided for each pressure in the Supporting Information Tables S5 and S6. Interestingly, above 30 GPa, $ReWC_{0.8}$ exhibits a kink in the compression curve; the origin of this deviation is not entirely clear from these data. However, because of the unexpected high-pressure (>30 GPa) behavior of $ReWC_{0.8}$, the pressure data were fit up to P = 28.6(3) GPa. A second-order B-M fit was implemented due to the limited pressure range for determining B_0'. According to the fit, the bulk modulus of $ReWC_{0.8}$ is B_0 = 380(8) GPa, deviating by less than 5% from the ML-predicted value of 398 GPa. Additionally, the refined standard state unit cell volume (V_0) is 68.43(1) Å^3, in agreement with the originally refined unit cell volume of 68.427(2) Å^3.

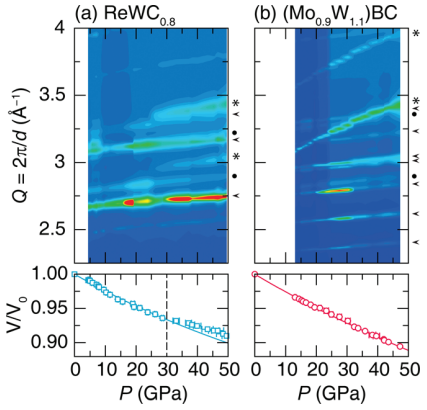


Figure 6. Diffraction patterns are plotted as a function of applied pressure for (a) $ReWC_{0.8}$ and (b) $Mo_{0.9}W_{1.1}BC$ (top panel). The prominent carbide and borocarbide peaks are marked by (<),while platinum peaks are indicated by (•), and peaks related to the gasket material or pressure medium are shown by (∗). B-M equations of state are fit to the relative volume vs pressure curves (bottom panel). $ReWC_{0.8}$ was compressed from 4.4(1) GPa to 49(1) GPa, but due to unexpected high-pressure behavior was only fit to 28.6(3) GPa.

$Mo_{0.9}W_{1.1}BC$ could be reliably measured up to 47.3 GPa (Figure 6b) without any anomalous changes in the unit cell volume as a function of temperature. Fitting the refined volume with a third-order B-M EoS across the entire pressure range yielded a bulk modulus of B_0 = 373(4) GPa. The agreement between this experimentally measured value and the ML prediction of B_{SVR} = 370 GPa is excellent, deviating by less than 0.25%. The refined standard state unit cell volume, V_0 = 162.000(3) Å^3, also aligns with the initially refined volume of 162.000(3) Å^3. The first derivative of the bulk modulus for $Mo_{0.9}W_{1.1}BC$ was determined to be B_0' = 2.3(2); however, because this value is lower than reported for most other compounds, a second-order fit was also employed returning V_0 = 162.001(6) Å^3 and B_0 = 342(2) GPa, which is still a <10% deviation from the ML-predicted B_{SVR} value. The results obtained from these DAC experiments indicate not only that both $ReWC_{0.8}$ and $Mo_{0.9}W_{1.1}BC$ are ultraincompressible materials with a bulk modulus greater than WB_4 but that our machine-learning prediction is excellent at predicting the bulk modulus of these two inorganic materials.

The Vickers hardness of the samples was determined by performing microindentation at five loads. As shown in Figure 7a and b, the load-dependent hardness of $ReWC_{0.8}$ (similar hardness values were obtained for ReWC with the results provided in Figure S5 of the Supporting Information) and $Mo_{0.9}W_{1.1}BC$ suggests that at low load (0.49 N) both compounds exceed the superhard limit of 40 GPa. The hardness ranges from H_V = 40(3) GPa at 0.49 N load to H_V = 22.5(7) GPa at 4.9 N load for $ReWC_{0.8}$. The Vickers hardness is slightly higher for $Mo_{0.9}W_{1.1}BC$, with hardness values spanning from H_V = 42(2) GPa at 0.49 N load to H_V = 24.9(7) GPa at 4.9 N load. This H_V surpasses the reported hardness of the parent phase, Mo_2BC, of only 1823 kg/mm^2 (17.8 GPa) by Vickers microindentation.[54] More recently, nanoindentation measurements of Mo_2BC, used as a thin film for hard coatings, report hardness values of 28−29 GPa.[58,59] These measurements show that the incorporation of tungsten most likely enhances hardness, likely due to an increase in valence electron density. However, it should be noted that comparing nanoindentation with microindentation measurements is not straightforward. Further, these results at high load are in relative agreement with previous microhardness measurements on $Mo_{0.9}W_{1.1}BC$ of 2190 kg/mm^2 (21.5 GPa);[54] although the indentation load is not reported, it is likely under high load (>4.9 N). It is of interest for future studies to explore the anisotropic hardness response of $Mo_{0.9}W_{1.1}BC$ in correlation to its crystal structure; here the reported hardness is regardless of this structure anisotropy. In fact, the higher standard deviation in the hardness values (Figure 7) compared to $ReWC_{0.8}$ could stem from this anisotropy. More importantly, these hardness values are comparable (slightly lower) to the recent reports of superhard materials such as WB_4, which range from H_V = 43.3 GPa at

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

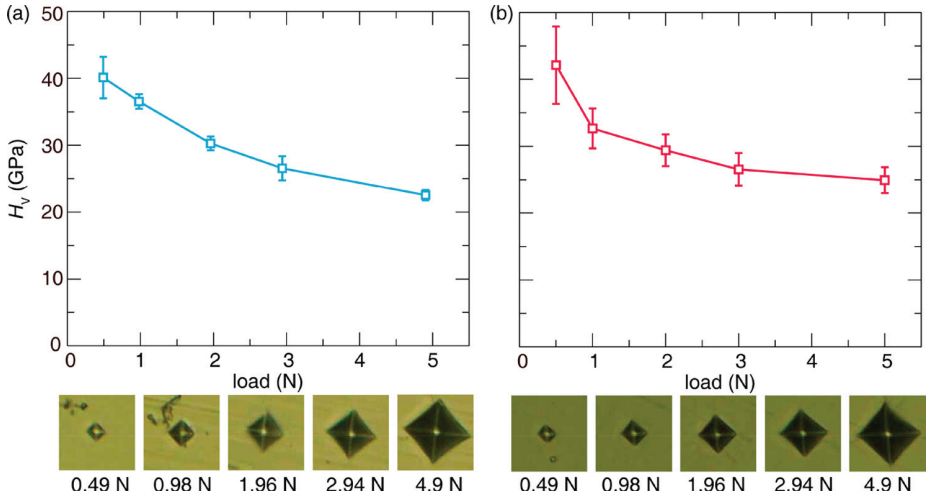


Figure 7. (a) $ReWC_{0.8}$ and (b) $Mo_{0.9}W_{1.1}BC$ Vickers microindentation (H_V) measurements as a function of applied load with error bars corresponding to errors from 10 independent measurements. Representative indentation marks at each load are represented at shown below each plot.

0.49 N to H_V = 28.1 GPa at 4.9 N.[28] $Mo_{0.9}W_{1.1}BC$ is also of significant interest considering this phase has a more earth abundant and a cheaper composition compared to WB_4 due to the partial substitution of Mo for W and C for B.[25] Noting that WB_4 is already an economic choice, $Mo_{0.9}W_{1.1}BC$ has great promise to be a cost-effective ultraincompressible, high-hardness material used in a wide range of applications.

Although these materials both fall in the superhard regime at low load, there is a ~40% decrease in the hardness at high load. In fact, this load-dependent mechanical response is observed in most metallic hard/superhard materials reported.[27,28,60] This observation is often attributed to the indentation size effect. There are also multiple alternative explanations for this response such as the presence of friction at the interface, which is proportional to the surface, whereas the plastic deformation resistance is proportional to the volume of the indentation. Conversely, it is possible that the small indentation volume probed at low load does not contain any dislocations; therefore, the hardness approaches the theoretical limit of the perfect crystal.[61,62] Nevertheless, the macroscopic (experimental measurement) and microscopic behavior relationship is still ambiguous,[63] and categorizing compounds such as ReB_2, WB_4, and now $ReWC_{0.8}$ and $Mo_{0.9}W_{1.1}BC$ as superhard materials has been controversial.[27] For example, the ability of ReB_2 to scratch diamond supports the superhard classification, which is supported by its high hardness at low indentation loads. However, at high loads the hardness of ReB_2 decreases significantly to 30 GPa.[27]

Generally, the hardness of a material is considered at the asymptotic hardness region before major cracks are observed. It is highlighted that low indentation loads are not suitable for the assessment of ultrahard materials since plastic deformation does not prevail and the hardness values are not viable.[27] Yet, even reporting hardness values at the asymptotic region could be intricate and problematic. For example, recent Vickers hardness measurements of δ-MoN single crystals are only reported at 0.245 and 0.49 N indentation loads (33 and 29 GPa), while the asymptotic hardness is assessed to be ~30 GPa. Even at such a low load, the indentations demonstrate formation and propagation of significant cracks, which is most likely due to the brittle nature of δ-MoN.[60] Considering most hard materials are predominantly brittle, aggravating this issue, we believe that the asymptotic hardness value should always be reported as a standard. However, the full load-dependence hardness behavior alongside the details of indentation such as the formation of cracks is more informative, leading to a better understanding of each material's specific indentation response.

Here, no major cracks were observed for most indentations for either the carbide or the borocarbide phase under high load (Figure 7), which is an indication of relative ductile behavior of these materials. Pugh's ratio (G_{SVR}/B_{SVR}) is an indication of ductile/brittle behavior of a material. For $ReWC_{0.8}$ and $Mo_{0.9}W_{1.1}BC$, Pugh's ratio is determined to be 0.46 and 0.52, respectively, which fall in the ductile regime (<0.57). However, since the ML-predicted shear modulus is likely underestimated, these values are probably closer to 0.57. Comparing the Pugh's ratio of the carbide and borocarbides discussed here to Pugh's ratio of ReB_2 (0.58), WB_4 (0.72), and *c*-BN (1.01) determined from our ML model establishes that our compounds are significantly more ductile. In fact, the ductility of our phases is likely the origin of the slightly lower hardness of $ReWC_{0.8}$ and $Mo_{0.9}W_{1.1}BC$ at the asymptotic region corresponding to the delocalization of electrons and the moderately high mobility of dislocations. Modifying the composition to make the carbide and borocarbide slightly less ductile is one potential avenue to improve the hardness of these phases.

CONCLUSION

Combining quantum mechanical calculations and advanced machine-learning techniques provides a unique framework to predict mechanical properties of inorganic materials. This method is employed to screen more than 180 000 compounds collected in the PCD regardless of unit cell size, chemical composition, or atomic disorder. Synthesis and characterization of top candidates of ternary and quaternary phases revealed discovery of an ultraincompressible, high-hardness metal ($ReWC_{0.8}$) as well as strengthen the case for $Mo_{0.9}W_{1.1}BC$ as an inexpensive, earth-abundant ultraincompressible hard metal. Even though the hardness of these phases does not yet surpass light main group high-pressure phases such as diamond and *c*-BN, transition-metal-based hard

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

materials offer a wide range of intriguing chemistries with impressive mechanical properties. Therefore, the search for superhard materials should not be limited to B/C/N compounds. Finally, based on the success of this research and the advances in record keeping of data as well as big data analysis tools, we foresee a paradigm shift of traditional synthesis of functional inorganic materials to be driven by informatics methods.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b02717.

Feature selection and variable importance on projection; list of predicted bulk, shear, and Young moduli and Poisson's ratio from PCD database; crystallographic data; cross-validated training set highlighting oxides; X-ray diffraction for $ReWC_x$; SEM images; lattice parameters and volumes refined from diamond anvil X-ray diffraction; Vickers hardness measurement of $ReWC_{0.8}$ and ReWC (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*jbrgoch@uh.edu

ORCID

Aria Mansouri Tehrani: 0000-0003-1968-0379
Anton O. Oliynyk: 0000-0003-0732-7340
Taylor D. Sparks: 0000-0001-8020-7711
Jakoah Brgoch: 0000-0002-1406-1352

Author Contributions

⊥A. Mansouri Tehrani and A. O. Oliynyk contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge Prof. Ken White in the Department of Mechanical Engineering at the University of Houston for allowing us to use the microindenter and associated resources. Financial support was provided by the National Science Foundation through No. NSF-CMMI 15-62142 (J.B. and A.M.), NSF-CMMI 15-62226 (T.D.S. and M.P.), EAR-1344579 and EAR-1654687 (L.M.). J.B. also thanks the donors of the American Chemical Society Petroleum Research Fund (55625-DNI10) and Seed Funding for Advanced Computing (SeFAC) at the University of Houston for supporting this research. Z.R. thanks the HCC-REEMS REU program (DMR-1460564) for supporting his summer of research. A.O.O. gratefully acknowledges the Eby Nell McElrath Postdoctoral Fellowship at University of Houston for financial support. L.M. also acknowledges support from the Carnegie DOE Alliance Center (CDAC). The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. COMPRES, the Consortium for Materials Properties Research in Earth Sciences under NSF Cooperative Agreement EAR-160685 also supported this project through funding beamline and gas loading equipment. We would like to thank Andrew Doran and Martin Kunz for beamline technical support and for assistance with gas loading.

■ REFERENCES

(1) Curtarolo, S.; Hart, G. L. W.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. *Nat. Mater.* **2013**, *12* (3), 191−201.
(2) Nishijima, M.; Ootani, T.; Kamimura, Y.; Sueki, T.; Esaki, S.; Murai, S.; Fujita, K.; Tanaka, K.; Ohira, K.; Koyama, Y.; Tanaka, I. *Nat. Commun.* **2014**, DOI: 10.1038/ncomms5553.
(3) Gautier, R.; Zhang, X.; Hu, L.; Yu, L.; Lin, Y.; Sunde, T. O. L.; Chon, D.; Poeppelmeier, K. R.; Zunger, A. *Nat. Chem.* **2015**, *7* (4), 308−316.
(4) Jain, A.; Shin, Y.; Persson, K. A. *Nat. Rev. Mater.* **2016**, *1* (1), 15004.
(5) Sanvito, S.; Oses, C.; Xue, J.; Tiwari, A.; Zic, M.; Archer, T.; Tozman, P.; Venkatesan, M.; Coey, M.; Curtarolo, S. *Sci. Adv.* **2017**, *3* (4), e1602241.
(6) Togo, A.; Tanaka, I. *Scr. Mater.* **2015**, *108*, 1−5.
(7) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. *APL Mater.* **2013**, *1* (1), 011002.
(8) Curtarolo, S.; Setyawan, W.; Hart, G. L. W.; Jahnatek, M.; Chepulskii, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. *Comput. Mater. Sci.* **2012**, *58*, 218−226.
(9) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. *npj Comput. Mater.* **2015**, *1* (1), 15010.
(10) JARVIS-DFT | NIST Materials Genome Initiative https://mgi.nist.gov/jarvis-dft (accessed Feb 12, 2018).
(11) Perdew, J. P. *Int. J. Quantum Chem.* **1985**, *28* (S19), 497−523.
(12) Heyd, J.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *121* (3), 1187−1192.
(13) Hedin, L. *Phys. Rev.* **1965**, *139* (3A), A796−A823.
(14) Oliynyk, A. O.; Adutwum, L. A.; Rudyk, B. W.; Pisavadia, H.; Lotfi, S.; Hlukhyy, V.; Harynuk, J. J.; Mar, A.; Brgoch, J. *J. Am. Chem. Soc.* **2017**, *139* (49), 17870−17881.
(15) Oliynyk, A. O.; Adutwum, L. A.; Harynuk, J. J.; Mar, A. *Chem. Mater.* **2016**, *28* (18), 6672−6681.
(16) Hautier, G.; Fischer, C. C.; Jain, A.; Mueller, T.; Ceder, G. *Chem. Mater.* **2010**, *22* (12), 3762−3767.
(17) Isayev, O.; Oses, C.; Toher, C.; Gossett, E.; Curtarolo, S.; Tropsha, A. *Nat. Commun.* **2017**, *8*, 15679.
(18) Xie, T.; Grossman, J. C. *Phys. Rev. Lett.* **2018**, *120* (14), 145301.
(19) De Jong, M.; Chen, W.; Notestine, R.; Persson, K.; Ceder, G.; Jain, A.; Asta, M.; Gamst, A. *Sci. Rep.* **2016**, *6*, 34256.
(20) Gómez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Duvenaud, D.; Maclaurin, D.; Blood-Forsythe, M. A.; Chae, H. S.; Einzinger, M.; Ha, D. G.; Wu, T.; Markopoulos, G.; Jeon, S.; Kang, H.; Miyazaki, H.; Numata, M.; Kim, S.; Huang, W.; Hong, S. I.; Baldo, M.; Adams, R. P.; Aspuru-Guzik, A. *Nat. Mater.* **2016**, *15* (10), 1120−1127.
(21) Du, X. P.; Lo, V. C.; Wang, Y. X. *J. Comput. Chem.* **2012**, *33* (1), 18−24.
(22) Oliynyk, A. O.; Sparks, T. D.; Gaultois, M. W.; Ghadbeigi, L.; Mar, A. *Inorg. Chem.* **2016**, *55* (13), 6625−6633.
(23) Gaultois, M. W.; Oliynyk, A. O.; Mar, A.; Sparks, T. D.; Mulholland, G. J.; Meredig, B. *APL Mater.* **2016**, *4* (5), 053213.
(24) Kaner, R. B.; Gilman, J. J.; Tolbert, S. H. *Science* **2005**, *308* (5726), 1268−1269.
(25) Mansouri Tehrani, A.; Ghadbeigi, L.; Brgoch, J.; Sparks, T. D. *Integr. Mater. Manuf. Innov.* **2017**, *6* (1), 1−8.
(26) Akopov, G.; Yeung, M. T.; Kaner, R. B. *Adv. Mater.* **2017**, *29* (21), 1604506.
(27) Chung, H. Y.; Weinberger, M. B.; Levine, J. B.; Kavner, A.; Yang, J. M.; Tolbert, S. H.; Kaner, R. B. *Science* **2007**, *316* (5823), 436−439.

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853

(28) Mohammadi, R.; Lech, A. T.; Xie, M.; Weaver, B. E.; Yeung, M. T.; Tolbert, S. H.; Kaner, R. B. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (27), 10958−10962.
(29) Wang, B.; Wang, D. Y.; Wang, Y. X. *J. Alloys Compd.* **2013**, *573*, 20−26.
(30) Zhang, Y.; Wang, L. *J. Phys. Soc. Jpn.* **2013**, *82* (7), 073702.
(31) Cui, L.; Hu, M.; Wang, Q.; Xu, B.; Yu, D.; Liu, Z.; He, J. *J. Solid State Chem.* **2015**, *228*, 20−26.
(32) Cui, L.; Wang, Q.; Xu, B.; Yu, D.; Liu, Z.; Tian, Y.; He, J. *J. Phys. Chem. C* **2013**, *117* (42), 21943−21948.
(33) Liu, A. Y.; Cohen, M. L. *Science* **1989**, *245* (4920), 841−842.
(34) Gilman, J. J. *Electronic Basis of the Strength of Materials*; Cambridge University Press, 2003.
(35) de Jong, M.; Chen, W.; Angsten, T.; Jain, A.; Notestine, R.; Gamst, A.; Sluiter, M.; Krishna Ande, C.; van der Zwaag, S.; Plata, J. J.; Toher, C.; Curtarolo, S.; Ceder, G.; Persson, K. A.; Asta, M. *Sci. Data* **2015**, *2*, 150009.
(36) Villars, P.; Cenzual, K. *Pearson's Crystal Data: Crystal Structure Database for Inorganic Compounds*; 2014.
(37) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. *APL Mater.* **2013**, *1* (1), 011002.
(38) Kresse, G.; Furthmüller, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54* (16), 11169−11186.
(39) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77* (18), 3865−3868.
(40) Anisimov, V. I.; Zaanen, J.; Andersen, O. K. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *44* (3), 943−954.
(41) Hill, R. *Proc. Phys. Soc., London, Sect. A* **1952**, *65* (5), 349−354.
(42) Drucker, H.; Burges, C. J. C.; Kaufman, L.; Smola, A.; Vapnik, V. *Adv. Neural Inf. Process. Dystems* **1997**, *1*, 155−161.
(43) Chang, C.-C.; Lin, C.-J. *ACM Trans. Intell. Syst. Technol.* **2011**, *2* (3), 1−27.
(44) Sheldrick, G. M. *SHELXTL*, Version 6.12; Bruker AXS Inc.: Madison, WI, 2001.
(45) Momma, K.; Izumi, F. *J. Appl. Crystallogr.* **2011**, *44* (6), 1272−1276.
(46) Fei, Y.; Ricolleau, A.; Frank, M.; Mibe, K.; Shen, G.; Prakapenka, V. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (22), 9182−9186.
(47) Wenk, H. R.; Lutterotti, L.; Kaercher, P.; Kanitpanyacharoen, W.; Miyagi, L.; Vasin, R. *Powder Diffr.* **2014**, *29* (3), 220−232.
(48) Angel, R. J.; Gonzalez-Platas, J.; Alvaro, M. *Z. Kristallogr. - Cryst. Mater.* **2014**, *229* (5), 405−419.
(49) ASTM International. E384-17 Standard Test Method for Microindentation Hardness of Materials; 2017, pp 1−40.
(50) De Jong, M.; Chen, W.; Angsten, T.; Jain, A.; Notestine, R.; Gamst, A.; Sluiter, M.; Ande, C. K.; Van Der Zwaag, S.; Plata, J. J.; Toher, C.; Curtarolo, S.; Ceder, G.; Persson, K. A.; Asta, M. *Sci. Data* **2015**, *2*, 150009.
(51) Cynn, H.; Klepeis, J. E.; Yoo, C. S.; Young, D. A. *Phys. Rev. Lett.* **2002**, *88* (13), 4.
(52) Charlesby, A. *Radiat. Phys. Chem.* **1997**, *49* (2), 297.
(53) Lawson, A. C. *J. Less-Common Met.* **1971**, *23* (1), 103−106.
(54) Lejay, P.; Chevalier, B.; Etourneau, J.; Hagenmuller, P.; Peshev, P. *Synth. Met.* **1981**, *4* (2), 139−145.
(55) Williams, W. S. *Mater. Sci. Eng., A* **1988**, *105−106*, 1−10.
(56) Zhao, Z.; Cui, L.; Wang, L. M.; Xu, B.; Liu, Z.; Yu, D.; He, J.; Zhou, X. F.; Wang, H. T.; Tian, Y. *Cryst. Growth Des.* **2010**, *10* (12), 5024−5026.
(57) Page, K.; Li, J.; Savinelli, R.; Szumila, H. N.; Zhang, J.; Stalick, J. K.; Proffen, T.; Scott, S. L.; Seshadri, R. *Solid State Sci.* **2008**, *10* (11), 1499−1510.
(58) Emmerlich, J.; Music, D.; Braun, M.; Fayek, P.; Munnik, F.; Schneider, J. M. *J. Phys. D: Appl. Phys.* **2009**, *42* (18), 185406.
(59) Anders, A. *J. Appl. Phys.* **2017**, *121* (17), 75305−171101.
(60) Wang, S.; Antonio, D.; Yu, X.; Zhang, J.; Cornelius, A. L.; He, D.; Zhao, Y. *Sci. Rep.* **2015**, DOI: 10.1038/srep13733.
(61) Manika, I.; Maniks, J. *Acta Mater.* **2006**, *54* (8), 2049−2056.
(62) Bull, S. J. *Z. Metallkd.* **2003**, *94* (7), 787−792.
(63) Brazhkin, V.; Dubrovinskaia, N.; Nicol, M.; Novikov, N.; Riedel, R.; Solozhenko, V.; Zhao, Y. *Nat. Mater.* **2004**, *3* (9), 576−577.

DOI: 10.1021/jacs.8b02717
J. Am. Chem. Soc. 2018, 140, 9844−9853