JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau

Article

Screening and Understanding Li Adsorption on Two-Dimensional Metallic Materials by Learning Physics and Physics-Simplified Learning

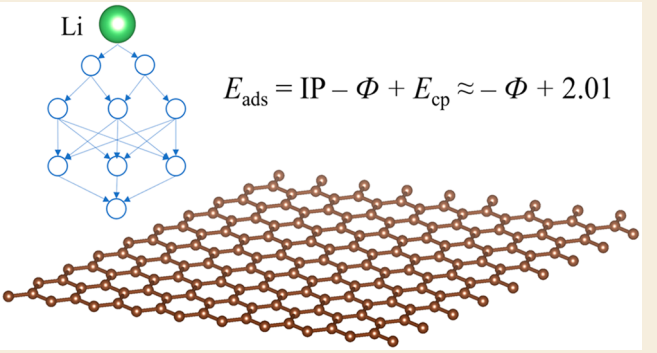
Sheng Gong,[#] Shuo Wang,[#] Taishan Zhu,[#] Xi Chen, Zhenze Yang, Markus J. Buehler, Yang Shao-Horn, and Jeffrey C. Grossman*

Cite This: *JACS Au* 2021, 1, 1904–1914

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Understanding and broad screening Li interaction energetics with surfaces are key to the development of materials for a wide range of applications including Li-based electrochemical capacitors, Li sensors, Li separation membranes, and Li-ion batteries. In this work, we build a high-throughput screening scheme to screen Li adsorption energetics on 2D metallic materials. First, density functional theory and graph convolution networks are utilized to calculate the minimum Li adsorption energies for some 2D metallic materials. The data is then used to find a dependence of the minimum Li adsorption energies on the sum of ionization potential, work function of the 2D metal, and coupling energy between Li^+ and substrate, and the dependence is used to screen all 2D metallic materials. Physics-simplified learning by splitting the property into different contributions and learning or calculating each component is shown to have higher accuracy and transferability for machine learning of complex materials properties.



KEYWORDS: *Li adsorption, 2D materials, work function, high-throughput screening, physics-simplified learning*

INTRODUCTION

Lithium adsorption is a fundamental electrochemical process in various applications such as Li-based electrochemical capacitors,[1] Li sensors,[2] Li separation membranes,[3] and Li-ion batteries (LIBs).[4,5] Two-dimensional (2D) materials have generated excitement for supercapacitors,[6] sensors,[7] separation membranes,[8] and LIBs[9] because of their distinct structural characteristics such as enlarged interlayer spacing and high aspect ratio, which can help alleviate structural instabilities, accelerate faster ion diffusion, and enhance Li adsorption.[10−12] Moreover, recently 2D materials have been used as components in composite materials such as in the form of mixed,[13−15] wrapped,[13,14,16] or encapsulated[13,14,17,18] composites[14,19] and layer-by-layer heterostructures,[13,19−21] showing broad applicability of 2D materials. Despite this promise, the interaction energetics between Li and 2D materials is not well understood due to the lack of both experimental and computational data,[22,23] which is a property important for not only devices built upon 2D materials but also applications with 2D materials-based composites. For example, for heterostructure electrodes, it is known that the Li insertion energetics between two different layers is largely dependent on the Li adsorption energetics on each single layer,[24−28] while for mixed, wrapped, and encapsulated electrodes, Li interaction energetics with the coating layer is used to evaluate the electrochemical stability of the coating layer.[17,29] Given the importance of the interaction between Li and 2D materials,, it would be useful to build a Li-2D materials interaction database for screening 2D materials with desired Li adsorption energetics and understanding the nature of Li adsorption.

Currently, there are several 2D materials databases with atomic and electronic structures from ab initio methods, C2DB,[30] Materials Cloud,[31] Jarvis,[32] and 2D Materials Encyclopedia,[33] and Jain et al.[34] summarized the databases into a comprehensive database with 7736 2D materials. Computation of Li interaction energetics with 2D materials requires ab initio methods such as density functional theory (DFT) to predict with sufficient accuracy across a large range of materials. Yet, it is challenging to screen over a large number of these materials with DFT since for each 2D material there are multiple possible adsorption sites, and different from the

Received: June 9, 2021
Published: October 6, 2021

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904–1914

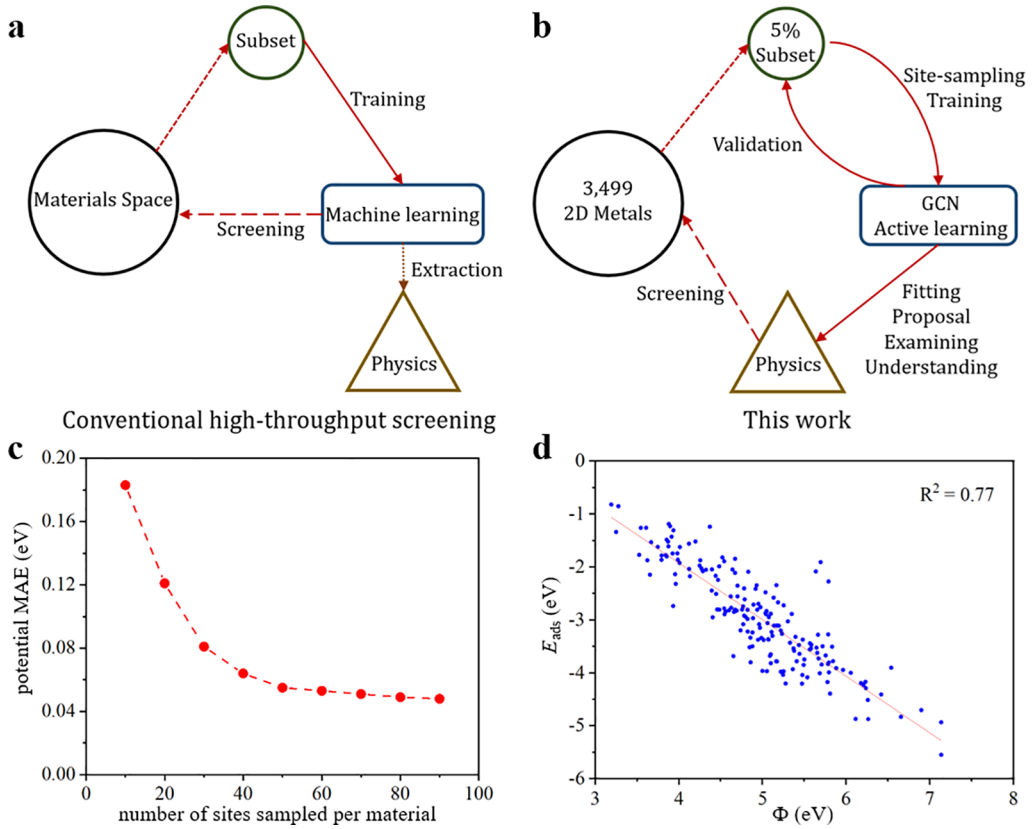


Figure 1. Illustrations of the high-throughput screening process. (a, b) Schematics of conventional machine learning-based high-throughput[49] screening workflow and the approach adopted in this work, respectively. (c) Mean average error (MAE, in eV) of the potentials for test sets versus number of adsorption sites sampled per material. (d) Minimum Li adsorption energy (E_{ads}, in eV) versus work function (Φ, in eV).

case of H adsorption where only "high-symmetry-sites" are screened,[35,36] Li is known to be adsorbed on sites far away from high-symmetry points,[17,37] which leads to a massive set of possible systems.

With the rapid development of machine learning (ML) techniques in the field of materials science, ML models are now used frequently in high-throughput screening,[38−41] where they are typically first trained on a small subset of the given materials space then directly applied to screen the larger materials space of interest. In the process, physical insights can be extracted as shown schematically in Figure 1a. Since Li adsorption is site dependent, here the most intuitive way to screen the minimum Li adsorption energy is to build machine learning potentials for all the materials and find the minimum adsorption energy from the potential-energy surface. However, it is still challenging to construct a universal potential that can generalize well for thousands of materials,[42,43] as will be shown later. Learning the minimum adsorption energy directly from the structure of the substrate is also not feasible, as it is too computationally demanding to generate the massive data of minimum Li adsorption energies, which limits the capacity of machine learning models,[44] and the low transferability of current machine learning models also makes the prediction on unseen data unreliable.[45−47] Therefore, it is important to explore approaches that go beyond conventional machine learning-based high-throughput screening that rely on purely data-driven machine learning models. Here, we incorporate physical insights related to Li binding energetics into the screening framework to simplify the learning problem.

For Li insertion into bulk materials, one approach is to estimate the energetics as a charge-transfer process between different energy levels[48−50] while leaving out other interactions, especially electrostatic due to their high complexity. Li adsorption on surfaces, including 2D materials, is considered to be a chemisorption process[51,52] in which the Li 2s orbital hybridizes with surface valence levels and the energy change depends on the degree of filling of the antibonding orbital. This understanding explains experimental observations well[53] and provides good qualitative trends. However, it is more challenging to estimate energetics quantitatively, especially when screening across different materials. In a recent work by Liu et al.,[54] Li adsorption is modeled via a two-step process: a Li atom first ionizes to a Li^+ and an electron and then the Li^+ couples with the charged 2D materials electrostatically. The system can be then approximated as an image-charge coupling where Li^+ is a +1 point charge and the 2D material is represented as a continuous conducting plate, giving the adsorption energy estimate as

$$E_{ads} = IP - \frac{14.38}{(2h)} \quad (1)$$

where E_{ads} is the minimum Li adsorption energy in eV, IP is the ionization potential of Li (5.39 eV), and h is the adsorption height in Å. Although this estimation explains trends between different alkali ions well,[54] the electronic structure of the substrate material is not accounted, which limits use across different materials.

Because it is hard to screen the minimum Li adsorption energy by either DFT calculations or conventional machine

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

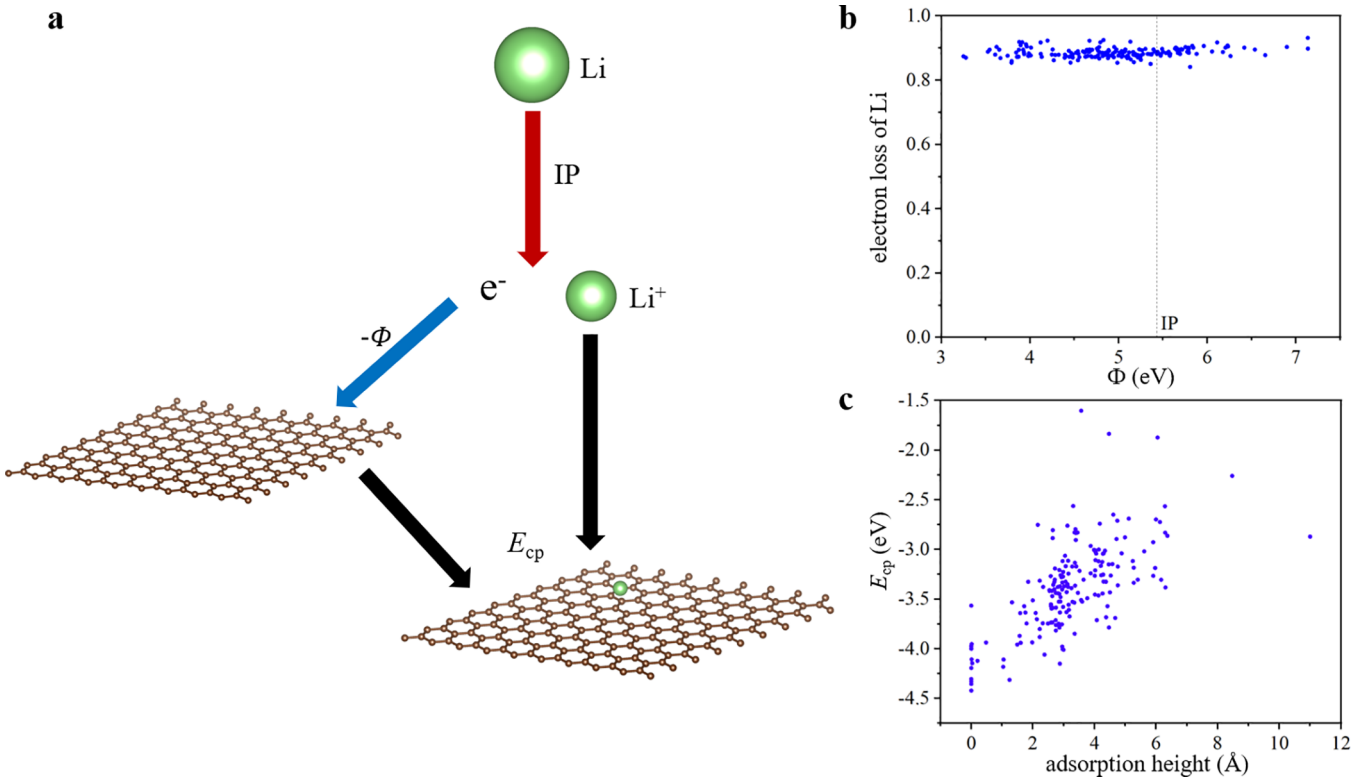


Figure 2. Proposed Li adsorption mechanism. (a) Illustration of the proposed three-step adsorption mechanism. Here, Φ is the work function, IP is the ionization potential of Li, and E_{cp} is coupling energy between Li^+ and the negatively charged substrate. (b) Electron loss of Li versus work function (Φ, in eV). The vertical line in the middle denotes the position of Li ionization potential. (c) Coupling energy (E_{cp}, in eV) versus adsorption height (from Li^+ to the central plane of the 2D material, in Å).

learning models, and because we know that the minimum Li adsorption energy is related to the process of ionization, charge transfer, and coupling, in this work we incorporate the process of learning physics into the screening circle (shown schematically in Figure 1b). The aim of this approach is to learn the underlying mechanism governing adsorption energy and summarize an empirical formula for higher efficiency, accuracy, and transferability. First, we use a subset with 5% of the 2D materials that are metallic (we restrict the present study to metallic materials since high electrical conductivity is desired for most electrochemical applications and they share different Li binding behaviors compared with semiconductors as will be discussed later). This subset of materials is then used to train graph convolutional network (GCN)[43,55−57] potentials to learn the adsorption energies with actively sampled adsorption sites. With the well-trained potential we then calculate the minimum Li adsorption energy for each material in the subset, and we use the data of the subset to fit and propose the physics of Li adsorption and then attempt to further examine and understand the physics based on a combination of human knowledge and machine learning. We find that the Li adsorption energy is the sum of three terms: the ionization potential of Li, the fermi level (work function) of the 2D metal, and coupling energy between Li^+ and the negatively charged surface. Finally, we use these correlations to screen the remaining 2D metals and find that fluorides and chromium oxides could be promising high-voltage materials, and we use the found physics as the design principle to enhance the Li adsorption ability of graphene. Through the comparison between learning the coupling energy and learning adsorption energy by random forest, we show that physics-simplified learning by decoupling a complex property into simpler contributions and learning or calculating each component separately has higher accuracy and transferability than unsimplified learning that directly learns the complex property.

The remainder of the paper is organized as follows. In the Minimum Li Adsorption Energies from GCN and DFT section, we describe how the minimum Li adsorption energy of the 180 materials subset is determined, in the Linear Relation between Li Adsorption Energy and Work Function section we introduce the linear relation between Li adsorption energy and work function, in the Prediction of the Coupling Energy section we discuss the components of coupling energy, machine learning of coupling energy, and the idea of physics-simplified learning, in the Transferability and Screening High-Voltage Materials section we test the transferability of the linear relation and physics-simplified learning and find materials with strong Li adsorption ability, in the Explaining Li Adsorption Behavior Based on the Found Li Adsorption Mechanism section we use the linear relation to explain some Li adsorption energetics, and in the Discussion section we further discuss the contributions and future applications of this work.

RESULTS

Minimum Li Adsorption Energies from GCN and DFT

In this work, we take advantage of graph convolutional networks (GCN) as the machine learning architecture for learning Li site energies at different adsorption sites. Recently, GCN has been shown to encode atomic and geometric information with high transferability,[47,56,58] and has been

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

utilized as a model form of interatomic potentials.[43,55] In order to efficiently learn Li adsorption energies at different sites and reduce the possibility of sampling bias, we actively sample sites from each material with site energies calculated by DFT, and then train GCN on all the calculated energies. For testing prediction performance, after each cycle of "sampling-DFT calculations-training" we sample a new set of site energies based on uncertainty, and these energies are added to the training data for the next generation GCN potential until the test error converges. The convergence is illustrated in Figure 1c, from which one can see that the error of the predicted site energies converges at around 0.05 eV, which is close to chemical accuracy (1 kcal/mol ≈ 0.04 eV/atom[59]) and much lower than the MAE of the naïve model (0.73 eV), where all predicted site energies are equal to the mean site energies from DFT. More details about the learning process are provided in Methods. As a comparison, a GCN potential trained on all 3499 2D metals, with the same number of total site energies as the training set (around 15000), gives an error of ∼ 0.5 eV (evaluated on the same test set as the last data point in Figure 1c), or 10 times higher than the error for the subset, which illustrates the difficulty of building a universal machine learning potential that generalizes well for the large materials space. Here we use the trained GCN potential to determine the minimum adsorption site for each material in the subset, and DFT optimization is followed for calculation of accurate minimum adsorption energies.

Linear Relation between Li Adsorption Energy and Work Function

From the aforementioned understanding of charge transfer between different energy levels, we expect that the minimum Li adsorption energy should have a strong correlation with the position of lowest unoccupied band or work function for zero-gap 2D materials. As shown in Figure 1d, the minimum Li adsorption energy has a strong linear relation (R^2 = 0.77) with work function

$$E_{ads} = -1.07\Phi + 2.35 \quad (2)$$

where Φ is the work function of the 2D metal in eV. In order to understand this linear relation, we propose a three-step adsorption mechanism based on the understanding described previously of charge-transfer and ionization-coupling mechanisms, as shown in Figure 2a. When approaching a 2D metal, a Li atom first ionizes to a Li^+ and an electron, which costs the energy equal to the ionization potential of Li, then the electron transfers to the 2D metal and releases energy equal to the work function, and finally the Li^+ couples with the negatively charged 2D metal with energy change of the coupling energy. According to this description, the minimum Li adsorption energy should have three terms

$$E_{ads} = IP - \Phi + E_{cp} \quad (3)$$

where E_{cp} is the coupling energy between Li^+ and the negatively charged substrate in eV.

Although eq 3 looks similar to eq 2, there is one "counterintuitive" behavior in our proposed mechanism from the understanding of charge transfer. In Figure 2a, we assume that the electron always transfers from the Li atom to the 2D material, and Figure 2b shows that, from Bader charge analysis,[60] in all cases Li indeed transfers ∼0.9 electron to the 2D metals. This behavior cannot be described by a picture wherein charge transfer results from the difference in energy levels, since for 2D metals with work function lower than the ionization potential of Li, their Fermi levels are higher than the Li 2s level, and electrons would transfer from the surface to the Li atom if electrons always flow from high energy levels to low ones. On the basis of the observed electron transfer direction in Figure 2b, charge transfer cannot simply occur from one isolated level to another, and according to chemisorption theory (Gurney model[61]), the DFT-computed electron transfer from Li to the surface can be explained by broadening and shift of the Li 2s orbital as it approaches the substrate.[51,61] Although not explicitly expressed in the proposed formula, these effects are included in our mechanism to support the proposed behavior.

Different from the proposed formula (eq 3), the slope of the line fitted from the subset (eq 2) is not −1, but rather −1.07. This deviation is likely due to bias of the data, because although the bonding between Li and 2D materials is not purely ionic,[52] there is almost no correlation (R^2 = 0.04) between work function and charge transfer from Bader charge analysis (Figure 2b), and therefore, there is no relation between the ratio of ionic/covalent bonding and work function. When we set the slope to be −1 and fit the intercept from the subset, we have a new fitted line:

$$E_{ads} = -\Phi + 2.01 \quad (4)$$

Comparing the numerical accuracy between eq 2 and eq 4, we find that the mean average error (MAE) and R^2 score between eq 2 and GCN/DFT-based data are 0.339 eV and 0.765, respectively, while those for eq 4 are 0.343 eV and 0.761, respectively, showing that switching from eq 2 to eq 4 does not have high impact on numerical accuracy. Based on this analysis and comparison, we choose eq 4 as the determined linear relation for the following parts.

While the qualitative trends between work function and adsorption energy have been established previously,eq 4 is of particular interest because of the "−1" slope, since on the one hand, previous studies of adsorption behavior based on the chemisorption theory have reported varying slopes between adsorption energy and electronic descriptors for different adsorbates,[62−64] and as described in the Discussion, Stavric et al.[22] reported a slope of ∼−0.92 by studying Li adsorption on 15 2D semiconductors, and on the other hand, we demonstrate that the linear relation with a "−1" slope is still valid for materials with low work function, which is beyond the understanding of charge transfer theory.

Prediction of the Coupling Energy

In addition to the linear relation between work function and adsorption energy, the coupling energy term in eq 3 is also of interest. Previously, Liu et al.[54] approximated the coupling effect as an electrostatic image-charge interaction where the coupling energy is inversely proportional to the adsorption height. From Figure 2c, one can see that the inverse trend holds well, which shows the nature of coupling energy is mainly electrostatic. More generally, the E_{cp} term should include the following contributions

$$E_{cp} = E_{electro}^{Li-subs} + E_{van\ der\ Waals}^{Li-subs} + \Delta E_{charge\ transfer}^{Li-subs} + E_{relaxation}^{subs} + E_{electro}^{Li-Li} \quad (5)$$

where $E_{electro}^{Li-subs}$ is the electrostatic interaction between Li^+ and negatively charged substrate and is approximated as an image-charge coupling by Liu et al.[54] as eq 1, $E_{van\ der\ Waals}^{Li-subs}$ is the van

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

Table 1. Comparison of Mean Average Error (MAE) of Predictions from Different Models from Leave-One-Out Cross-Validation[a].

models	Φ = RF(features)	E_{ads} = RF(features)	E_{cp} = RF(features) E_{ads} = IP − Φ + RF(features)	$E_{cp} = \bar{E}_{cp}$ $E_{ads} = -\Phi + 2.01$
MAE (eV)	0.268	0.345	0.216	0.343

[a]Here, RF(features) means that the quantity is from a random forest model with selected features, $\bar{E}_{cp}$ denotes the mean value of E_{cp}, IP is the ionization potential of Li, and the Φ terms in the right two cases are work functions from DFT. For the right two cases, the two models in each case share the same MAE according to eq 7.

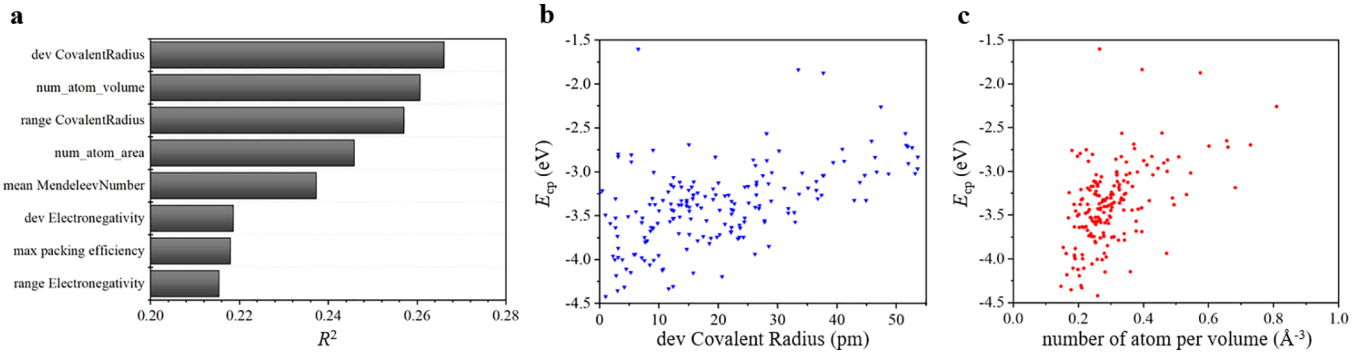


Figure 3. Correlations between features and coupling energy. (a) R^2 scores between compositional and structural features and E_{cp} (the R^2 score is obtained by building a linear regression between each feature and E_{cp}, and here, the R^2 score is the square of Pearson correlation coefficient[66]). Here, only features with R^2 larger than 0.2 are plotted. (b) E_{cp} versus standard deviation of covalent radius of elements in the 2D materials (in pm). (c) E_{cp} versus number of atoms per volume (area × thickness, in Å^{-3}).

der Waals interaction between Li and substrate, $\Delta E^{Li-subs}_{charge\ transfer}$ includes the energies associated with charge transfer that are not captured by the work function of 2D materials, $\Delta E^{subs}_{relaxation}$ is the energy change corresponds to the structural change of substrate, and $E^{Li-Li}_{electro}$ is the electrostatic repulsion between Li^+.

In this work, $E^{Li-subs}_{electro}$, $E^{Li-subs}_{van\ der\ Waals}$, and $\Delta E^{Li-subs}_{charge\ transfer}$ are included in the DFT calculation and machine learning of E_{cp}. Since the electrostatic interaction between Li^+ and negatively charged substrate is attractive, both of the first two energy contributions should follow the trend that, lower adsorption height, stronger adsorption (more negative adsorption energy),[54] while it is not clear whether $\Delta E^{Li-subs}_{charge\ transfer}$ correlates with adsorption height. In Figure 2c, we show that, lower adsorption height, smaller (more negative) E_{ads}, which suggests that the correlation between E_{ads} and adsorption height is similar to that of $E^{Li-subs}_{electro}$ and $E^{Li-subs}_{van\ der\ Waals}$. More discussions of $E^{Li-Li}_{electro}$ and $E^{subs}_{relaxation}$ are provided in the Supporting Information.

In order to improve our understanding of coupling and predict E_{cp} without calculating adsorption height by DFT (in this section, "E_{cp}" represents the coupling energy associated with the minimum Li adsorption energy, which is not site-dependent), we build a random forest model to learn E_{cp} from compositional and structural features as

$$E_{cp} = \text{RF(features)} \tag{6}$$

where RF represents the random forest model and "features" denote the compositional and structural features listed in Methods. As a result, we have a new equation for the minimum Li adsorption energy:

$$E_{ads} = \text{IP} - \Phi + E_{cp} = \text{IP} - \Phi + \text{RF(features)} \tag{7}$$

As shown in Table 1, using a random forest model lowers the mean average error (MAE) of the linear relation with fixed intercept (equation 4) by one-third, where the E_{ads} is solely dependent on Φ and as a result E_{cp} is seen as a constant value. Based on the data in Table 1, we note that eq 7 has higher accuracy than directly learning E_{ads} by random forest since the problem of predicting E_{ads} is decoupled into two separate problems (predicting Φ and predicting E_{cp}), where DFT is used to accurately and efficiently calculate the work function Φ from a single-point DFT calculation for each material. If E_{ads} is learned directly, we essentially learn E_{cp} and Φ at the same time (as shown in the feature importance in Figure S4), which makes the learning problem harder. This difference explains why the MAE for learning E_{ads} is larger than that for simultaneously learning E_{cp} and Φ, as shown in Table 1. We propose that the idea of "physics-simplified learning", which is splitting a complicated property into different contributions, and learning or calculating each component separately, could result in more accurate prediction performance, and more discussion about the idea is provided in the following sections.

In order to further understand E_{cp}, we show the R^2 scores between compositional and structural features and E_{cp} in Figure 3a (R^2 score obtained by building a linear regression between each feature and E_{cp}), from which the deviation and range of elemental properties and density-related features are those with the strongest correlation with E_{cp}. The latter is more straightforward to understand than the former. For density, the less densely packed atoms in the 2D materials the stronger the correlation (Figure 3c) since Li^+ is more likely to be adsorbed into the pores with lower adsorption height due to steric effects. Variances of elemental properties are known to be important for thermal and mechanical properties;[41] however, they are much less well understood in terms of correlation with Li adsorption energetics. As shown in Figure 3b, the more similar elements in 2D materials the stronger coupling. One explanation for this trend is that with more similar elements the bonds in the 2D materials become less polarized and more covalent because of more similar atoms at ends of bonds and because more covalently bound materials tend to contain

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

poorly packed structures with corner-shared or isolated polyhedral,[48] it is more likely that those materials have pores for Li^+ to be adsorbed in, as shown in Figure S5. A SHAP analysis[65] (impact of feature on model output) is provided in Figure S6 as a supplement for further understanding the relation between features and E_{cp}.

Transferability and Screening High-Voltage Materials

After building up the linear dependence of the work function and random forest model for coupling, we use the models to screen all the 2D metallic materials in the database by Jain et al.[34] to test transferability and search for novel battery materials. Since the energy density of electrochemical capacitors and Li-ion batteries is determined by voltage and capacity, and the voltage originates from the chemical potential difference between Li atoms in the anode and cathode, which corresponds to the Li adsorption energy on 2D materials, as an example, we select materials with the lowest minimum Li adsorption energies from our models as potential high-voltage materials and then use DFT to manually calculate their minimum Li adsorption energies, with results listed in Table 2.

Table 2. Prediction of the Minimum Li Adsorption Energy (in eV) on 2D Metallic Materials from Different Models[a]

	$E_{ads} = -\Phi + 2.01$	E_{ads} = IP − Φ + RF(features)	E_{ads} = RF(features)	DFT
Bi_2F_2	−7.29	−7.21	−4.64	−6.82
N_6F_{30} (2431)	−7.14	−7.19	−4.97	−7.48
N_6F_{30} (2783)	−6.87	−7.13	−3.93	−7.35
Cr_9O_{27}	−7.24	−7.13	−3.93	−7.10
$Cr_4P_8O_{32}$	−7.25	−7.12	−4.53	−8.04

[a]Materials with the top five lowest minimum Li adsorption energies from eq 7 are included, and the numbers after the two N_6F_{30} are their IDs in the database from Jain et al.,[34] respectively.

We find that some fluorides and chromium oxides are promising high-voltage materials with the minimum Li adsorption energies lower than −7 eV, which exceeds the previous records for 2D materials of ∼−5 eV.[22] As shown in Table 2, the predictions of ultrahigh binding strength are confirmed by the DFT calculations, which highlights the ability of our found empirical formula and physically simplified learning to extrapolate, since the lowest adsorption energy in the training set is around −6 eV. As a comparison, the unsimplified learning problem, E_{ads} = RF(features), fails to identify the found high-voltage materials, even though it has similar interpolation performance compared with equation 4, showing the poor transferability of the unsimplified learning problem.

Explaining Li Adsorption Behavior Based on the Found Li Adsorption Mechanism

In addition to high-throughput screening and machine learning, the found Li adsorption mechanism can be used as a tool for analyzing and modifying Li adsorption behavior. For example, the found linear relation between work function and Li adsorption energy can be used to enhance the Li adsorption ability of graphene, which is known to have superior mechanical strength and electrical conductivity[67] although weak Li adsorption.[68] Here we try to raise the work function of graphene by two types of p-type doping, B-doping and F-functionalization. For each modification, we construct three defect concentrations, and Li is adsorbed on the same site for all six defect graphene structures, as shown in Figure S7. In Figure 4a, the relation between the work function change and Li adsorption energy change is plotted, from which one can observe the trend that, for each type of modification, higher work function leads to lower Li adsorption energy (stronger adsorption), showing that raising the work function can serve to enhance Li adsorption on graphene. However, structural modifications also affect the coupling between Li^+ and the substrate, and here we find that for F-functionalized graphene, $|\Delta E_{ads}|$ is significantly smaller than $|\Delta\Phi|$, suggesting weaker coupling ($\Delta E_{cp} > 0$) according to eq 3. This is because for F-functionalized graphene each F atom on top of the graphene plane attracts ∼0.7 electrons away from the plane according to Bader charge analysis,[60] making the substrate more positively charged and thus weakening the electrostatic coupling.

Another phenomenon that can be explained by the found mechanism is the enhancement of Li adsorption on strained graphene.[69] Previously, it has been found that when stretching, Li adsorption on graphene and graphyne is enhanced,[69,70] and

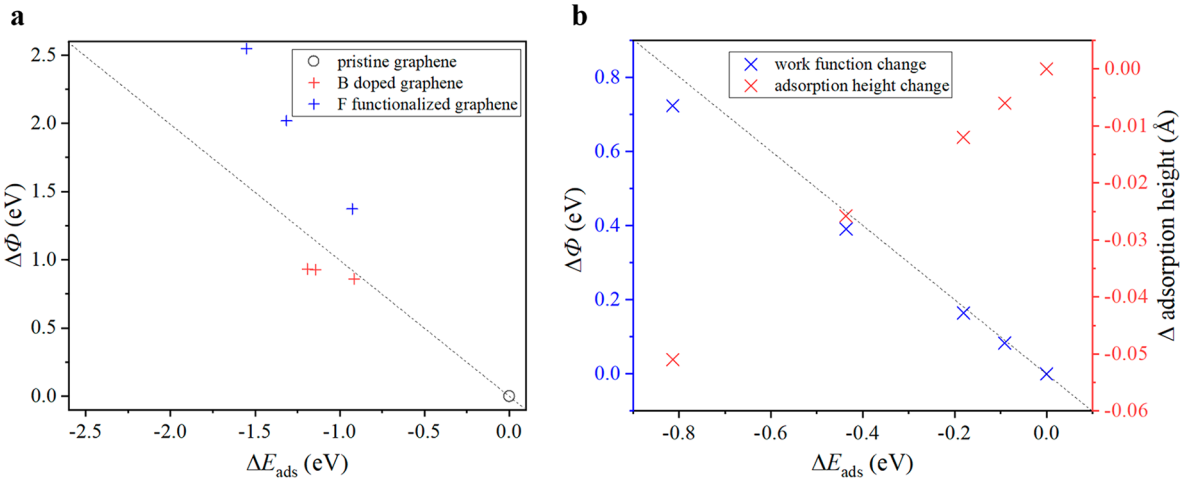


Figure 4. (a) Change of Li adsorption energy versus change of work function for B-doped graphene and F-functionalized graphene. (b) Change of Li adsorption energy versus change of work function and change of adsorption height for graphene with strain. The strain level from right to left is 0%, 1%, 2%, 5%, and 10%, respectively.

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

whether electrostatic attraction or charge transfer effect leads to the enhancement is under debate.[70] Here, as an example, we calculate the Li adsorption energy on graphene with in-plane tensile strain by DFT, and as shown in Figure 4b, we find that under strain, the changes of Li adsorption energy align well with the changes of work function (the change of work function under strain can be explained by tight-binding theory[71]). Although the difference between changes of Li adsorption energy and work function becomes larger with larger strain and adsorption height change, one can still conclude that, up to 10% strain, the main contribution of change of adsorption energy is work function change or, in other words, charge transfer effect as opposed to electrostatic interaction.

DISCUSSION

After showing the results for Li adsorption on 2D metallic materials, we briefly discuss the reasons why semiconductors are not included in this work. Semiconductors will possess different Li adsorption energetics compared with metals, which could be harder to capture. For semiconductors, if the Li 2s electron still transfers to the highest unoccupied band, the work function (Φ) term in eq 3 should be replaced by electron affinity (EA) as

$$E_{ads} = \mathrm{IP} - \mathrm{EA} + E_{cp} \quad (8)$$

However, there are a number of concerns about using eq 8 to screen minimum Li adsorption energy for 2D semiconductors: (1) EA is related to the position of the conduction band minimum (CBM), which is notoriously challenging to accurately predict using inexpensive (i.e., PBE) functionals;[72] (2) midgap states might appear when adsorbed with Li,[22] shifting EA from CBM into midgap states , in turn making it no longer reasonable to screen the minimum Li adsorption energy without explicitly introducing the Li atom into the system; (3) even if the effect of midgap states is considered, Stavric et al.[22] reported a line with slope around −0.92 by studying 15 2D semiconductors, and it is unclear whether the non "−1" slope is a result of data bias or other undiscovered adsorption mechanisms; and (4) In addition to charge transfer, for coupling, 2D semiconductors cannot be viewed as "conducting plates", further deviating from the image-charge coupling[54] compared with 2D metals.

The main contributions of this work and potential applications are summarized as follows:

(1) For high-throughput screening, as shown in Figure 1, we add the step of "learning physics" into the high-throughput screening framework, which is useful for problems where the use of a machine learning model is limited due to lack of generalizability and transferability while there exists previous knowledge regarding the target property of screening. For problems where the empirical formula is not as straightforward as the linear relation in this work, the technique of symbolic regression could be used to find the empirical expression.[73] Examples are screening thermal conductivity by building an empirical formula based on bulk modulus[74] and screening molecular adsorption by building empirical formulas based on electronic properties such as the d-band center.[62]

(2) For machine learning of materials properties, we illustrate a way to improve accuracy and transferability by physics-simplified learning or splitting a hard learning problem into simpler ones, as shown by the performances of learning E_{cp} and E_{ads} in Tables 1 and 2. This approach can be applied to learn complex properties which are known to be dependent on different sources of contributions. One example is learning interatomic potentials by learning different interactions, such as the potential used here. According to eq 3, E_{ads} can be partitioned into three contributions, and it is possible that learning contributions separately could result in higher learning accuracy. Since IP and Φ are intrinsic properties of Li and materials, here as a brief example, we train a GCN potential on E_{cp} on the same training and test set as the last data point in Figure 1c and find that the MAE is 0.038 eV. Compared with the MAE of 0.046 eV for learning E_{ads} in Figure 1c, we find that learning E_{cp} results in better prediction accuracy, as the target problem is simpler without the contribution of charge transfer. Note that here "E_{cp}" is a site-dependent property and different from that in the section of "Prediction of the coupling energy" above. This gives an example of the benefit of splitting energies from different interactions and learning each component separately, or physics-simplified interatomic potential learning, although further studies would elucidate the potential broader benefits of such an approach. In general, physics-simplified learning can be summarized as

$$\text{From } C = f(A, B);\ \text{minimizing loss} = \mathcal{L}(C_{true} - C_{pred})$$

$$\text{To minimizing loss} = \mathcal{L}(A_{true} - A_{pred})$$

and

$$\text{minimizing loss} = \mathcal{L}(B_{true} - B_{pred})$$

where f is a known functional form and A and B are known physical quantities. In the example of E_{ads} in this work, since $E_{ads} = \mathrm{IP} - \Phi + E_{cp}$, IP is a constant, and Φ can be calculated by DFT, we convert the problem of minimizing the difference between true and predicted E_{ads} into minimizing that between true and predicted E_{cp}. In eq 4, we use $E_{ads} = -\Phi + 2.01$ to screen the adsorption energy for all metallic materials, which is equivalent to using Φ as the "surrogate variable" to screen E_{ads}, and in eq 7, we use $E_{ads} = \mathrm{IP} - \Phi + \mathrm{RF(features)}$ to improve the accuracy and transferability, which is similar to the idea of "residual learning",[75−77] or learning the difference between high-fidelity data (E_{ads}) and low-fidelity data (Φ). For more general cases, if both A and B are known physical properties and require machine learning prediction, then the idea of "physics-simplified learning" is more similar to the concept of "physics-inspired machine learning architecture"[78] in the sense that although the base surrogate model is not changed (neural networks or random forests) multiple base models are employed in parallel and the outputs of the surrogate models are integrated by the known physical formula f. This approach is different from deriving formulation from machine learning, or symbolic regression,[73,74,79,80] where A, B, and C are all known physical properties while f is unknown, and different from learning physical representations[81] where C and f are known or human-designed while A and B are unknown.

(3) For materials selection, we build a database of Li adsorption energies, which can serve as a reference for selecting materials with ideal Li adsorption behavior. In addition to the high-voltage materials as shown above, materials with low and moderate Li adsorption energies are also important for various applications. For example, anode and coating materials require moderate Li adsorption energies for high energy density and stability[17,67,82] and Li adsorption

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

behavior is important for materials with ultralow work functions in electronic devices.[83]

(4) For understanding the Li bonding nature, although we do not expect that the proposed three-step adsorption mechanism is what actually happens in a real time sequence, the effective partitioning of adsorption energy can serve as an analysis tool and design principle, such as the analysis of enhancement of Li adsorption by p-doping and introducing strain into the substrate as discussed above.

SUMMARY

In summary, we use a high-throughput screening framework that combines first-principle calculations, graph convolutional networks, and the process of learning physics to screen Li adsorption on 2D metallic materials. We find that the adsorption energy can be partitioned into three terms, ionization potential, work function of 2D metals, and coupling energy. We further learn the coupling energy by random forest models, and by comparing the prediction accuracy between learning adsorption and coupling energy by random forest, we propose the idea of physics-simplified learning, or decoupling a complex property into simpler properties and learning those simpler properties. We hope that our high-throughput screening framework will be enlightening for other high-throughput screening tasks difficult for conventional first-principle and machine learning methods, the idea of physics-simplified learning will help to learn complex materials properties with higher accuracy and transferability, the built database of Li adsorption will serve as a reference for selecting materials with ideal Li adsorption energy, and the found Li adsorption mechanism will be used for analyzing Li adsorption energetics.

METHODS

GCN Potential

In this work, we use GCN potentials to learn the Li adsorption energies on different sites on different 2D materials. A key advantage of GCN is that it has high compositional transferability[56,58] because the atomic embedding is based on atomic properties.[58] Compared with other forms of deep neural potentials, such as Coulomb Matrix[84] and Radial Distribution Function,[85] which only take atomic positions without elemental information on these positions as input, GCN is a more suitable choice as the dimensions of feature vectors in GCN are independent of the number of elements in the system, which is important here as most of elements in the periodic table exist in our system, as shown in Figure S1. Here, we choose Crystal Graph Convolutional Neural Networks (CGCNN)[56] as the architecture for building graph convolution network potentials for Li adsorption because it is shown to be highly transferable across different compositions and geometries.[47,58] Note that iCGCNN[86] and other deep learning frameworks with both compositional and geometrical transferability might also be suitable as the form of potential. We use the same CGCNN architecture for learning formation energy of materials used by Xie et al.,[56] as it is shown to be a suitable form for learning energy.[56]

Here, we build GCN potentials for a subset of 180 2D metallic materials. More analysis about the subset, including the appearance frequency of elements and crystal system, is provided in the Supporting Information. For each material, we determine the candidate adsorption sites by splitting the 3D real space into grid points with the distance of 0.2 Å between each point in each dimension, and we only consider points that are within the distance of 2 to 5 Å to the constituent atoms of the 2D sheet. The setting of distance thresholds is based on atomic radii,[87] $r_{Li} + r_H = 2.2$ Å (H is the smallest atom in the periodic table without inert gas) and $r_{Li} + r_{Cs}$ = 4.65 Å (Cs is the largest atom), and we round the two values to 2 and 5 Å to provide more tolerance. For each grid point, we also check whether there are other symmetrically equivalent points in the space, and we only take one of those equivalent points into the candidate set.

For each material, 10 adsorption sites are randomly sampled to calculate the Li site energies, and then five GCN potentials are trained on the training set with 10 × 180 data points. Then, for the next batches of sampling, 50 adsorption sites are randomly sampled for each material, and DFT is used to calculate adsorption energies for the 10 sites with highest uncertainty (standard deviation) from the predictions from the 5 GCN potentials trained at the last batch. Note that to save computational resources we decide not to evaluate uncertainty of all candidate sites (more than 10^6) in each active learning cycle and instead evaluate 50*180 sites. After these DFT calculations, the newly calculated 10 × 180 energies are used as a test set to evaluate the prediction performance of the GCN potentials from the prior batch, and then the newly calculated energies are added to the training set (if we have sampled n batches, then we now have n × 10 × 180 training data, and 20% of these data are further randomly split out as the cross-validation set), and five new GCN potentials are trained based on the newly enlarged training set. The iterative process is stopped once the test errors converge to 0.05 eV, as shown in Figure 1c. As described above, there are 4.4 × 10^6 candidate adsorption sites for the 180 materials subset, and as shown in Figure 1c, the total number of sites queried by active sampling is 1.6 × 10^5; therefore, the efficiency of the sequential procedure is around 0.4%.

Descriptor-Based Models

For all of the descriptor-based models, we use the code from Scikit Learn.[88] For the descriptors used for learning E_{cp}, we add three features (thickness, number of atom per area, number of atom per volume) to the ones generated from the Matminer[89] platform. The modules used to generate features in this work are as follows: Element Property (magpie), Oxidation States, Electron Affinity, Band Center, Cohesive Energy, Miedema, TMetal Fraction, Valence Orbital, Yang Solid Solution, Global Symmetry Features, Structural Complexity, Chemical Ordering, Global Instability Index, Maximum Packing Efficiency, Minimum Relative Distances, Structural Heterogeneity, Average Bond Length, Average Bond Angle, Bond Orientational Parameter, and Coordination Number. We compare three models, linear regression, support vector regression and random forest. The results of the leave-one-out cross-validation are shown in Figure S3, from which one can see that random forest model is better in prediction accuracy than linear regression and support vector regression. Therefore, analysis about descriptors-based model is built on random forest in this work.

DFT Calculations

All first-principle calculations, including structural optimization, site energy, work function calculation, and charge transfer, are carried out using density functional theory (DFT) by employing the Vienna *Ab initio* Simulation Package (VASP).[90] The projector augmented wave (PAW) method[91] is used with a kinetic energy cutoff of 500 eV. The first Brillouin zone is sampled by the Monkhorst−Pack scheme[92] with a grid density of KSPACING = 0.4 $Å^{-1}$. The exchange-correlation interactions are treated using Perdew−Burke−Ernzerh functional (PBE) within the generalized gradient approximation (GGA).[93] DFT-D3 scheme[94] is used to estimate the van der Waals interactions, as it is thought to be more accurate than other semiempirical DFT-D methods.[94] Convergence criteria are set to be 10^{-4} eV for the total energy and $<10^{-2}$ eV/Å for atomic forces, respectively. To avoid Li−Li interactions between periodic images, supercells with lattice parameters larger than 7 Å[95] are used for all Li-contained calculations. For the work function Φ, we calculate it for each 2D material by first getting the fermi level and vacuum level of 2D materials from VASP output files vasprun.xml and LOCPOT, respectively, then subtracting fermi level from the vacuum level.

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.1c00260.

Distribution of materials in the subset, distribution of sampled Li adsorption sites, comparison between descriptor-based machine learning models, feature importance of learning E_{ads}, E_{cp}, and Φ by random forests, standard deviation of covalent radius versus adsorption height, SHAP analysis of learning E_{cp} by random forests, structures of B-doped graphene and F-functionalized graphene, and discussions about $E^{Li-Li}_{electro}$ and $E^{subs}_{relaxation}$ (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jeffrey C. Grossman − *Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0003-1281-2359; Email: jcg@mit.edu

Authors

Sheng Gong − *Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0002-7457-7959

Shuo Wang − *Department of Materials Science and Engineering, University of Maryland, College Park, Maryland 20742, United States*

Taishan Zhu − *Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States*

Xi Chen − *Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States*

Zhenze Yang − *Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States*

Markus J. Buehler − *Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0002-4173-9659

Yang Shao-Horn − *Department of Materials Science and Engineering and Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;* orcid.org/0000-0001-8714-2121

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.1c00260

Author Contributions

#S.G., S.W., and T.Z. contributed equally. S.G., T.Z. performed the machine learning analysis, S.G., S.W. performed the DFT calculations, S.G., S.W., T.Z., J.C.G., and Y.S.H. conceived the idea, J.C.G. supervised this work, and all authors contributed to the analysis of data and writing and editing of the paper.

Notes

The authors declare no competing financial interest.

The training set, high-throughput screening result, GCN potential and structures of the found high-voltage materials with adsorbed Li are provided at the linkhttps://github.com/shenggong1996/Screening-Li-adsorption-on-2D-metals. GCN potentials are constructed based on CGCNN at the link[56] https://github.com/txie-93/cgcnn. Random forest, linear regression and support vector regression are based on sklearn[88] at the link https://scikit-learn.org/stable/. Features are generated by the Matminer[89] at the link https://github.com/hackingmaterials/matminer. Other codes used in this work is available at the link https://github.com/shenggong1996/Screening-Li-adsorption-on-2D-metals.

■ ACKNOWLEDGMENTS

This work was supported by Toyota Research Institute. Computational support was provided by the DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and the Extreme Science and Engineering Discovery Environment, supported by National Science Foundation Grant No. ACI-1053575. Z.Y. and M.J.B. acknowledge support from the ARO (W911NF1920098) and AFOSR-MURI (FA9550-15-1-0514).

■ REFERENCES

(1) Ding, J.; Hu, W.; Paek, E.; Mitlin, D. Review of Hybrid Ion Capacitors: From Aqueous to Lithium to Sodium. *Chem. Rev.* **2018**, *118* (14), 6457−6498.

(2) Kamenica, M.; Kothur, R. R.; Willows, A.; Patel, B. A.; Cragg, P. J. Lithium Ion Sensors. *Sensors* **2017**, *17* (10), 2430.

(3) Razmjou, A.; Asadnia, M.; Hosseini, E.; Habibnejad Korayem, A.; Chen, V. Design principles of ion selective nanostructured membranes for the extraction of lithium ions. *Nat. Commun.* **2019**, *10* (1), 5793.

(4) Xu, N.; Li, L.; He, Y.; Tong, Y.; Lu, Y. Understanding the molecular mechanism of lithium deposition for practical high-energy lithium-metal batteries. *J. Mater. Chem. A* **2020**, *8* (13), 6229−6237.

(5) Zhao, D.; Li, S. J. F. i. C. Regulating the Performance of Lithium-Ion Battery Focus on the Electrode-Electrolyte Interface. *Front. Chem.* **2020**, *8*, 821.

(6) Forouzandeh, P.; Pillai, S. C. Two-dimensional (2D) electrode materials for supercapacitors. *Mater. Today: Proc.* **2021**, *41*, 498−505.

(7) Bolotsky, A.; Butler, D.; Dong, C.; Gerace, K.; Glavin, N. R.; Muratore, C.; Robinson, J. A.; Ebrahimi, A. Two-Dimensional Materials in Biosensing and Healthcare: From In Vitro Diagnostics to Optogenetics and Beyond. *ACS Nano* **2019**, *13* (9), 9781−9810.

(8) Cheng, L.; Liu, G.; Zhao, J.; Jin, W. Two-Dimensional-Material Membranes: Manipulating the Transport Pathway for Molecular Separation. *Acc. Mater. Res.* **2021**, *2* (2), 114−128.

(9) Rojaee, R.; Shahbazian-Yassar, R. Two-Dimensional Materials to Address the Lithium Battery Challenges. *ACS Nano* **2020**, *14* (3), 2628−2658.

(10) Tian, H.; Seh, Z. W.; Yan, K.; Fu, Z.; Tang, P.; Lu, Y.; Zhang, R.; Legut, D.; Cui, Y.; Zhang, Q. Theoretical Investigation of 2D Layered Materials as Protective Films for Lithium and Sodium Metal Anodes. *Adv. Energy Mater.* **2017**, *7* (13), 1602528.

(11) Banerjee, S.; Periyasamy, G.; Pati, S. K. Possible application of 2D-boron sheets as anode material in lithium ion battery: A DFT and AIMD study. *J. Mater. Chem. A* **2014**, *2* (11), 3856.

(12) Yang, Y.; Liu, X.; Zhu, Z.; Zhong, Y.; Bando, Y.; Golberg, D.; Yao, J.; Wang, X. The Role of Geometric Sites in 2D Materials for Energy Storage. *Joule* **2018**, *2* (6), 1075−1094.

(13) Raccichini, R.; Varzi, A.; Passerini, S.; Scrosati, B. The role of graphene for electrochemical energy storage. *Nat. Mater.* **2015**, *14* (3), 271−9.

(14) Cai, X.; Lai, L.; Shen, Z.; Lin, J. Graphene and graphene-based composites as Li-ion battery electrode materials and their application in full cells. *J. Mater. Chem. A* **2017**, *5* (30), 15423−15446.

(15) Kucinskis, G.; Bajars, G.; Kleperis, J. Graphene in lithium ion battery cathode materials: A review. *J. Power Sources* **2013**, *240*, 66−79.
(16) Yang, S.; Feng, X.; Ivanovici, S.; Mullen, K. Fabrication of graphene-encapsulated oxide nanoparticles: towards high-performance anode materials for lithium storage. *Angew. Chem., Int. Ed.* **2010**, *49* (45), 8408−11.
(17) Gong, S.; Wang, S.; Liu, J.; Guo, Y.; Wang, Q. Graphdiyne as an ideal monolayer coating material for lithium-ion battery cathodes with ultralow areal density and ultrafast Li penetration. *J. Mater. Chem. A* **2018**, *6* (26), 12630−12636.
(18) Gullapalli, H.; Kalaga, K.; Vinod, S.; Rodrigues, M.-T. F.; George, A.; Ajayan, P. M. 2D material integrated macroporous electrodes for Li-ion batteries. *RSC Adv.* **2017**, *7* (52), 32737−32742.
(19) Pomerantseva, E.; Gogotsi, Y. Two-dimensional heterostructures for energy storage. *Nat. Energy* **2017**, *2* (7), 17089.
(20) Xiong, P.; Zhang, F.; Zhang, X.; Wang, S.; Liu, H.; Sun, B.; Zhang, J.; Sun, Y.; Ma, R.; Bando, Y.; Zhou, C.; Liu, Z.; Sasaki, T.; Wang, G. Strain engineering of two-dimensional multilayered heterostructures for beyond-lithium-based rechargeable batteries. *Nat. Commun.* **2020**, *11* (1), 3297.
(21) Zhen, M.; Wang, J.; Guo, S.-q.; Shen, B. Vertically aligned nanosheets with MoS2/N-doped-carbon interfaces enhance lithium-ion storage. *Appl. Surf. Sci.* **2019**, *487*, 285−294.
(22) Stavrić, S.; Popović, Z. S.; Šljivančanin, Ž. Understanding trends in lithium binding at two-dimensional materials. *Phys. Rev. Mater.* **2018**, *2* (11), 114007.
(23) Ye, C.; Chao, D.; Shan, J.; Li, H.; Davey, K.; Qiao, S.-Z. Unveiling the Advances of 2D Materials for Li/Na-S Batteries Experimentally and Theoretically. *Matter* **2020**, *2* (2), 323−344.
(24) Lin, H.; Jin, R.; Zhu, S.; Huang, Y. C3N/blue phosphorene heterostructure as a high rate-capacity and stable anode material for lithium ion batteries: Insight from first principles calculations. *Appl. Surf. Sci.* **2020**, *505*, 144518.
(25) Demiroglu, I.; Peeters, F. M.; Gulseren, O.; Cakir, D.; Sevik, C. Alkali Metal Intercalation in MXene/Graphene Heterostructures: A New Platform for Ion Battery Applications. *J. Phys. Chem. Lett.* **2019**, *10* (4), 727−734.
(26) Zhu, Y.; Peng, W.; Li, Y.; Zhang, G.; Zhang, F.; Fan, X. Multiple roles of a heterointerface in two-dimensional van der Waals heterostructures: insights into energy-related applications. *J. Mater. Chem. A* **2019**, *7* (41), 23577−23603.
(27) Ding, Y.; Xiao, B.; Li, J.; Deng, Q.; Xu, Y.; Wang, H.; Rao, D. Improved Transport Properties and Novel Li Diffusion Dynamics in van der Waals C2N/Graphene Heterostructure as Anode Materials for Lithium-Ion Batteries: A First-Principles Investigation. *J. Phys. Chem. C* **2019**, *123* (6), 3353−3367.
(28) Lee, H. W.; Jung, H.; Yeo, B. C.; Kim, D.; Han, S. S. Atomistic Sodiation Mechanism of a Phosphorene/Graphene Heterostructure for Sodium-Ion Batteries Determined by First-Principles Calculations. *J. Phys. Chem. C* **2018**, *122* (36), 20653−20660.
(29) Aykol, M.; Kim, S.; Hegde, V. I.; Snydacker, D.; Lu, Z.; Hao, S.; Kirklin, S.; Morgan, D.; Wolverton, C. High-throughput computational design of cathode coatings for Li-ion batteries. *Nat. Commun.* **2016**, *7*, 13779.
(30) Haastrup, S.; Strange, M.; Pandey, M.; Deilmann, T.; Schmidt, P. S.; Hinsche, N. F.; Gjerding, M. N.; Torelli, D.; Larsen, P. M.; Riis-Jensen, A. C.; Gath, J.; Jacobsen, K. W.; Jørgen Mortensen, J.; Olsen, T.; Thygesen, K. S. The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Mater.* **2018**, *5* (4), 042002.
(31) Mounet, N.; Gibertini, M.; Schwaller, P.; Campi, D.; Merkys, A.; Marrazzo, A.; Sohier, T.; Castelli, I. E.; Cepellotti, A.; Pizzi, G.; Marzari, N. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* **2018**, *13* (3), 246−252.
(32) Choudhary, K.; Kalish, I.; Beams, R.; Tavazza, F. High-throughput Identification and Characterization of Two-dimensional Materials using Density functional theory. *Sci. Rep.* **2017**, *7* (1), 5179.
(33) Zhou, J.; Shen, L.; Costa, M. D.; Persson, K. A.; Ong, S. P.; Huck, P.; Lu, Y.; Ma, X.; Chen, Y.; Tang, H.; Feng, Y. P. 2DMatPedia, an open computational database of two-dimensional materials from top-down and bottom-up approaches. *Sci. Data* **2019**, *6* (1), 86.
(34) Jain, A.; Wang, Z.; Nørskov, J. K. Stable Two-Dimensional Materials for Oxygen Reduction and Oxygen Evolution Reactions. *ACS Energy Lett.* **2019**, *4* (6), 1410−1411.
(35) Yang, T.; Zhou, J.; Song, T. T.; Shen, L.; Feng, Y. P.; Yang, M. High-Throughput Identification of Exfoliable Two-Dimensional Materials with Active Basal Planes for Hydrogen Evolution. *ACS Energy Lett.* **2020**, *5* (7), 2313−2321.
(36) Karmodak, N.; Andreussi, O. Catalytic Activity and Stability of Two-Dimensional Materials for the Hydrogen Evolution Reaction. *ACS Energy Lett.* **2020**, *5* (3), 885−891.
(37) Hashmi, A.; Farooq, M. U.; Khan, I.; Son, J.; Hong, J. Ultra-high capacity hydrogen storage in a Li decorated two-dimensional C2N layer. *J. Mater. Chem. A* **2017**, *5* (6), 2821−2828.
(38) Ahmad, Z.; Xie, T.; Maheshwari, C.; Grossman, J. C.; Viswanathan, V. Machine Learning Enabled Computational Screening of Inorganic Solid Electrolytes for Suppression of Dendrite Formation in Lithium Metal Anodes. *ACS Cent. Sci.* **2018**, *4* (8), 996−1006.
(39) Sendek, A. D.; Cubuk, E. D.; Antoniuk, E. R.; Cheon, G.; Cui, Y.; Reed, E. J. Machine Learning-Assisted Discovery of Solid Li-Ion Conducting Materials. *Chem. Mater.* **2019**, *31* (2), 342−352.
(40) Cubuk, E. D.; Sendek, A. D.; Reed, E. J. Screening billions of candidates for solid lithium-ion conductors: A transfer learning approach for small data. *J. Chem. Phys.* **2019**, *150* (21), 214701.
(41) Zhu, T.; He, R.; Gong, S.; Xie, T.; Gorai, P.; Nielsch, K.; Grossman, J. C. Charting Lattice Thermal Conductivity of Inorganic Crystals. *Energy Environ. Sci.* **2021**, *14* (14), 3559.
(42) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **2016**, *145* (17), 170901.
(43) Mueller, T.; Hernandez, A.; Wang, C. Machine learning for interatomic potential models. *J. Chem. Phys.* **2020**, *152* (5), 050902.
(44) Zhang, Y.; Ling, C. A strategy to apply machine learning to small datasets in materials science. *npj Comput. Mater.* **2018**, *4* (1), 1.
(45) Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R. Solving the electronic structure problem with machine learning. *npj Comput. Mater.* **2019**, *5* (1), 1.
(46) Schmidt, J.; Marques, M. R. G.; Botti, S.; Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **2019**, *5* (1), 1.
(47) Gong, S.; Xie, T.; Zhu, T.; Wang, S.; Fadel, E. R.; Li, Y.; Grossman, J. C. Predicting charge density distribution of materials using a local-environment-based graph convolutional network. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2019**, *100* (18), 184103.
(48) Melot, B. C.; Tarascon, J. M. Design and preparation of materials for advanced electrochemical storage. *Acc. Chem. Res.* **2013**, *46* (5), 1226−38.
(49) Cherkashinin, G.; Hausbrand, R.; Jaegermann, W. Performance of Li-Ion Batteries: Contribution of Electronic Factors to the Battery Voltage. *J. Electrochem. Soc.* **2019**, *166* (3), A5308−A5312.
(50) Wang, Z.; Wang, D.; Zou, Z.; Song, T.; Ni, D.; Li, Z.; Shao, X.; Yin, W.; Wang, Y.; Luo, W. Efficient potential-tuning strategy through p-type doping for designing cathodes with ultrahigh energy density. *Natl. Sci. Rev.* **2020**, *7* (11), 1768−1775.
(51) Persson, B. N.; Ishida, H. Depolarization and metallization in alkali-metal overlayers. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1990**, *42* (5), 3171−3174.
(52) Kaack, M.; Fick, D. Binding of Li atoms adsorbed on metal surfaces. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *51* (24), 17902−17909.
(53) Horn, K.; Hohlfeld, A.; Somers, J.; Lindner, T.; Hollins, P.; Bradshaw, A. M. Identification of the s-derived valence-electron level in photoemission from alkali-metal adlayers on aluminum. *Phys. Rev. Lett.* **1988**, *61* (21), 2488−2491.
(54) Liu, Y.; Merinov, B. V.; Goddard, W. A. 3rd, Origin of low sodium capacity in graphite and generally weak substrate binding of

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914

Na and Mg among alkali and alkaline earth metals. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (14), 3735−9.
(55) Wang, W.; Yang, T.; Harris, W. H.; Gomez-Bombarelli, R. Active learning and neural network potentials accelerate molecular screening of ether-based solvate ionic liquids. *Chem. Commun.* **2020**, *56* (63), 8920−8923.
(56) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120* (14), 145301.
(57) Schutt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Muller, K. R. SchNet - A deep learning architecture for molecules and materials. *J. Chem. Phys.* **2018**, *148* (24), 241722.
(58) Xie, T.; Grossman, J. C. Hierarchical visualization of materials space with graph convolutional neural networks. *J. Chem. Phys.* **2018**, *149* (17), 174111.
(59) Sauer, J. Ab Initio Calculations for Molecule-Surface Interactions with Chemical Accuracy. *Acc. Chem. Res.* **2019**, *52* (12), 3502−3510.
(60) Henkelman, G.; Arnaldsson, A.; Jónsson, H. A fast and robust algorithm for Bader decomposition of charge density. *Comput. Mater. Sci.* **2006**, *36* (3), 354−360.
(61) Gurney, R. W. Theory of Electrical Double Layers in Adsorbed Films. *Phys. Rev.* **1935**, *47* (6), 479−482.
(62) Gao, W.; Chen, Y.; Li, B.; Liu, S. P.; Liu, X.; Jiang, Q. Determining the adsorption energies of small molecules with the intrinsic properties of adsorbates and substrates. *Nat. Commun.* **2020**, *11* (1), 1196.
(63) Garcia-Muelas, R.; Lopez, N. Statistical learning goes beyond the d-band model providing the thermochemistry of adsorbates on transition metals. *Nat. Commun.* **2019**, *10* (1), 4687.
(64) Chen, W.; Schneider, W. F.; Wolverton, C. Trends in Atomic Adsorption on Pt3M(111) Transition Metal Bimetallic Surface Overlayers. *J. Phys. Chem. C* **2014**, *118* (16), 8342−8349.
(65) Lundberg, S. M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J. M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S. I. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat Mach Intell* **2020**, *2* (1), 56−67.
(66) Anthony, L. J. R. R. The Cambridge Dictionary of Statistics; Cambridge University Press, 2003.
(67) Gong, S.; Wang, Q. Boron-Doped Graphene as a Promising Anode Material for Potassium-Ion Batteries with a Large Capacity, High Rate Performance, and Good Cycling Stability. *J. Phys. Chem. C* **2017**, *121* (44), 24418−24424.
(68) Shaidu, Y.; Küçükbenli, E.; de Gironcoli, S. Lithium Adsorption on Graphene at Finite Temperature. *J. Phys. Chem. C* **2018**, *122* (36), 20800−20808.
(69) Hao, F.; Chen, X. First-principles study of lithium adsorption and diffusion on graphene: the effects of strain. *Mater. Res. Express* **2015**, *2* (10), 105016.
(70) Zhang, Q.; Tang, C.; Zhu, W.; Cheng, C. Strain-Enhanced Li Storage and Diffusion on the Graphyne as the Anode Material in the Li-Ion Battery. *J. Phys. Chem. C* **2018**, *122* (40), 22838−22848.
(71) Reich, S.; Maultzsch, J.; Thomsen, C.; Ordejón, P. Tight-binding description of graphene. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2002**, *66* (3), 035412.
(72) Crowley, J. M.; Tahir-Kheli, J.; Goddard, W. A. 3rd, Resolution of the Band Gap Prediction Problem for Materials Design. *J. Phys. Chem. Lett.* **2016**, *7* (7), 1198−203.
(73) de Franca, F. O.; Aldeia, G. S. I. Interaction-Transformation Evolutionary Algorithm for Symbolic Regression. *Evol. Comput.* **2021**, 1−25.
(74) Loftis, C.; Yuan, K.; Zhao, Y.; Hu, M.; Hu, J. Lattice Thermal Conductivity Prediction Using Symbolic Regression and Machine Learning. *J. Phys. Chem. A* **2021**, *125* (1), 435−450.
(75) Patra, A.; Batra, R.; Chandrasekaran, A.; Kim, C.; Huan, T. D.; Ramprasad, R. A multi-fidelity information-fusion approach to machine learn and predict polymer bandgap. *Comput. Mater. Sci.* **2020**, *172*, 109286.
(76) Pilania, G.; Gubernatis, J. E.; Lookman, T. Multi-fidelity machine learning models for accurate bandgap predictions of solids. *Comput. Mater. Sci.* **2017**, *129*, 156−163.
(77) Egorova, O.; Hafizi, R.; Woods, D. C.; Day, G. M. Multifidelity Statistical Machine Learning for Molecular Crystal Structure Prediction. *J. Phys. Chem. A* **2020**, *124* (39), 8065−8078.
(78) Rai, R.; Sahu, C. K. J. I. A. Driven by data or derived through physics? a review of hybrid physics guided machine learning techniques with cyber-physical system (cps) focus. *IEEE Access* **2020**, *8*, 71050−71073.
(79) Weng, B.; Song, Z.; Zhu, R.; Yan, Q.; Sun, Q.; Grice, C. G.; Yan, Y.; Yin, W. J. Simple descriptor derived from symbolic regression accelerating the discovery of new perovskite catalysts. *Nat. Commun.* **2020**, *11* (1), 3513.
(80) Segler, M. H. S.; Preuss, M.; Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **2018**, *555* (7698), 604−610.
(81) Iten, R.; Metger, T.; Wilming, H.; Del Rio, L.; Renner, R. J. P. r. l. Discovering physical concepts with neural networks. *Phys. Rev. Lett.* **2020**, *124* (1), 010508.
(82) Er, D.; Li, J.; Naguib, M.; Gogotsi, Y.; Shenoy, V. B. Ti(3)C(2) MXene as a high capacity electrode material for metal (Li, Na, K, Ca) ion batteries. *ACS Appl. Mater. Interfaces* **2014**, *6* (14), 11173−9.
(83) Cui, Z.; Wang, X.; Li, E.; Ding, Y.; Sun, C.; Sun, M. Alkali-metal-adsorbed g-GaN monolayer: ultralow work functions and optical properties. *Nanoscale Res. Lett.* **2018**, *13* (1), 207.
(84) Faber, F.; Lindmaa, A.; von Lilienfeld, O. A.; Armiento, R. Crystal structure representations for machine learning models of formation energies. *Int. J. Quantum Chem.* **2015**, *115* (16), 1094−1101.
(85) Schütt, K. T.; Glawe, H.; Brockherde, F.; Sanna, A.; Müller, K. R.; Gross, E. K. U. How to represent crystal structures for machine learning: Towards fast prediction of electronic properties. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *89* (20), 205118.
(86) Park, C. W.; Wolverton, C. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Phys. Rev. Mater.* **2020**, *4* (6), 063801.
(87) Clementi, E.; Raimondi, D. L.; Reinhardt, W. P. Atomic Screening Constants from SCF Functions. II. Atoms with 37 to 86 Electrons. *J. Chem. Phys.* **1967**, *47* (4), 1300−1307.
(88) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(89) Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N. E. R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; Chard, K.; Asta, M.; Persson, K. A.; Snyder, G. J.; Foster, I.; Jain, A. Matminer: An open source toolkit for materials data mining. *Comput. Mater. Sci.* **2018**, *152*, 60−69.
(90) Kresse, G.; Furthmuller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54* (16), 11169−11186.
(91) Blochl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50* (24), 17953−17979.
(92) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Physical Review B* **1976**, *13* (12), 5188−5192.
(93) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865.
(94) 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* (15), 154104.
(95) Khantha, M.; Cordero, N. A.; Alonso, J. A.; Cawkwell, M.; Girifalco, L. A. Interaction and concerted diffusion of lithium in a (5,5) carbon nanotube. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *78* (11), 115430.

https://doi.org/10.1021/jacsau.1c00260
JACS Au 2021, 1, 1904−1914