JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

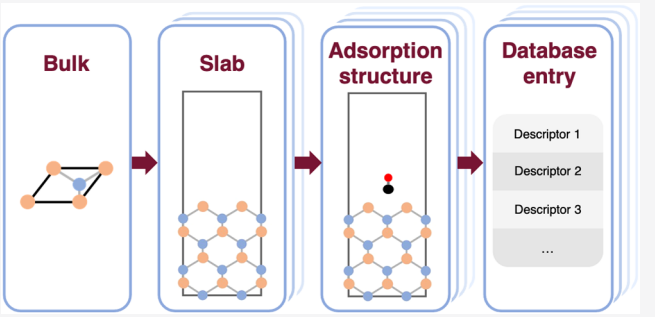
Automated Adsorption Workflow for Semiconductor Surfaces and the Application to Zinc Telluride

Oxana Andriuc,[▽] Martin Siron,[▽] Joseph H. Montoya, Matthew Horton, and Kristin A. Persson*

Cite This: https://doi.org/10.1021/acs.jcim.1c00340 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Surface adsorption is a crucial step in numerous processes, including heterogeneous catalysis, where the adsorption of key species is often used as a descriptor of efficiency. We present here an automated adsorption workflow for semiconductors which employs density functional theory calculations to generate adsorption data in a high-throughput manner. Starting from a bulk structure, the workflow performs an exhaustive surface search, followed by an adsorption structure construction step, which generates a minimal energy landscape to determine the optimal adsorbate−surface distance. An extensive set of energy-based, charge-based, geometric, and electronic descriptors tailored toward catalysis research are computed and saved to a personal user database. The application of the workflow to zinc telluride, a promising CO_2 reduction photocatalyst, is presented as a case study to illustrate the capabilities of this method and its potential as a material discovery tool.



INTRODUCTION

Adsorption plays an important role in a broad range of materials-based technologies, including heterogeneous catalysts,[1] energy storage,[2,3] and solid-state synthesis.[4,5] From a computational perspective, evaluation of adsorption energetics is a complex process, due to the breadth of possible adsorption structure configurations requiring comprehensive and generalized surface site-finding algorithms, and the complexity and elevated computational cost of first-principles calculations on such systems.

One application of adsorption that is of particular interest, given the need for novel, efficient renewable energy systems, is the photocatalytic conversion of carbon dioxide to fuels, which represents the driving force behind the work presented in this paper. Developing efficient and cost-effective solar-driven CO_2 reduction reaction (CO_2RR) devices has the potential to both alleviate greenhouse gas emissions and offer a sustainable liquid fuel production method. In recent years, many computational studies on CO_2RR photocatalysts have been carried out, offering insights into already existing, as well as potential new photocatalytic systems.[6] Nevertheless, the search for a robust CO_2RR photocatalytic system remains an active area of material discovery, motivating the need for a computational search approach.

While a number of high-throughput searches for photoelectrodes have been previously published,[7,8] existing literature frequently does not address the adsorption of relevant reactants and intermediates on the surface, despite it being a crucial step in the photocatalytic process. Existing adsorption workflows have been mainly designed for metallic systems,[9,10] or metal-organic frameworks,[11] and tend to focus on a limited set of descriptors.

The current work aims to provide a new and improved automated adsorption workflow, which can be applied to a variety of systems and chemistries, including semiconductors, which includes polar surfaces with more complicated electronic structure. With minimal user supervision or input, the workflow performs density functional theory (DFT) calculations on the bulk, all surfaces, and all adsorption structures with a given set of adsorbate species. This is followed by a comprehensive analysis step, which saves a set of computed adsorption properties to a personal user database that can be easily queried, and ultimately facilitates the screening of materials through a descriptor-based method. The generated data includes a series of quantities relevant to photocatalysis without impinging on the applicability of the workflow to other surface processes.

METHODS

Data and Software Availability. The adsorption workflow presented here is developed as part of the atomate open-source software package,[12] making use of the Fire-

Received: March 22, 2021

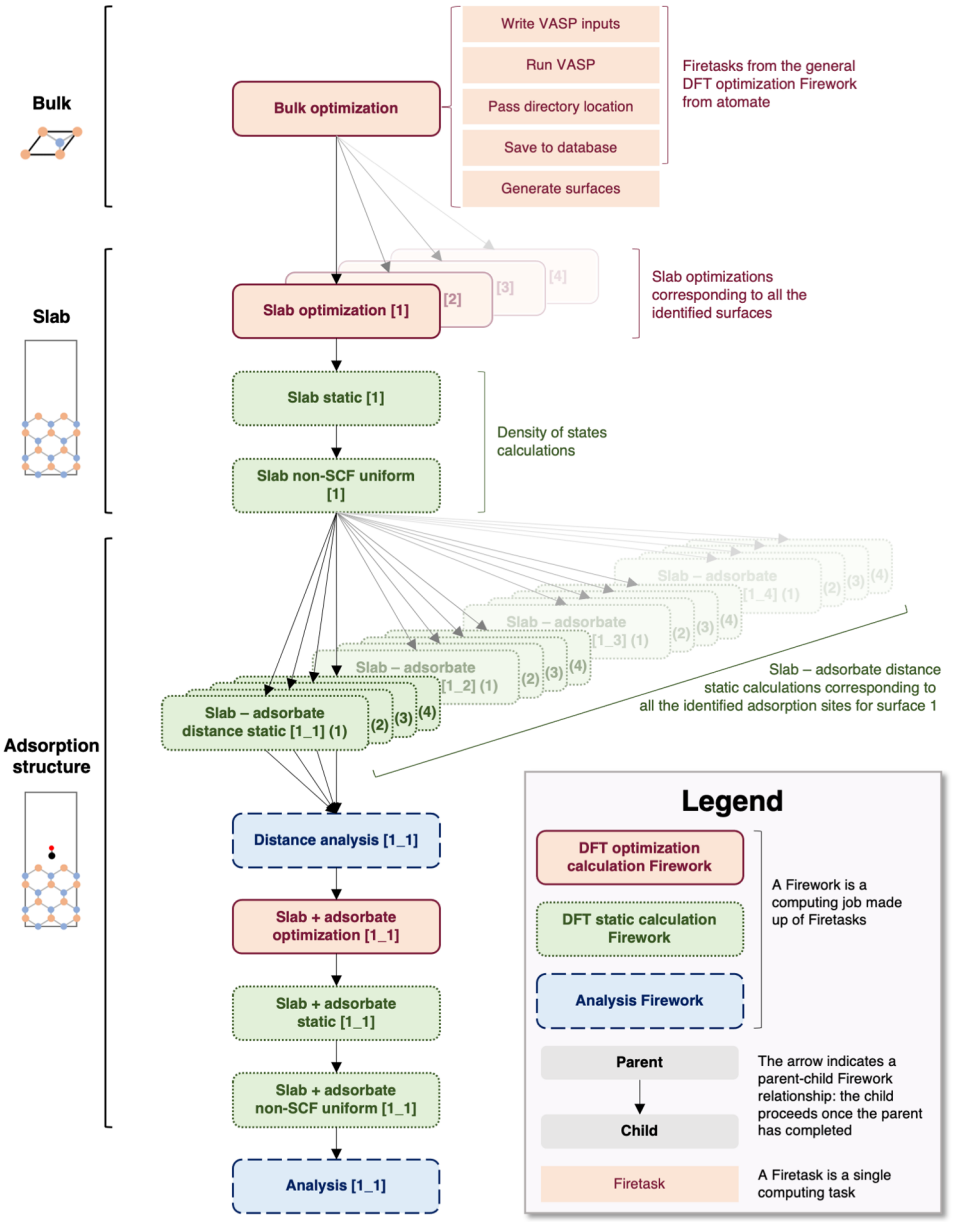


Figure 1. General workflow structure. The workflow consists of individual computing jobs (Fireworks), each of which is a sequence of computing tasks (Firetasks; example Firetasks shown for the bulk optimization Firework). At the slab and slab + adsorbate optimization levels, the flowchart shown here follows only one of the multiple possible branches corresponding to all of the surfaces, and adsorption structures, respectively.

Works,[13] pymatgen,[14] and custodian[14,15] software packages and is freely available at https://github.com/oxana-a/atomate/tree/ads_wf. The Fireworks workflow package facilitates the organization of various computing jobs (called Fireworks; an example being a density functional theory relaxation calculation) into a sequence of individual computing tasks (called Firetasks; an example being creating input files as part of a geometry optimization Firework). Density functional theory (DFT) calculations are performed with the Vienna Ab initio Simulation Package (VASP).[16,17] Details about how to run the workflow can be found in the Supporting Information.

Overview. The general structure of the workflow is shown in Figure 1. The first Firework comprises the crystal bulk structure optimization, which includes several Firetasks that are specific to DFT geometry optimizations (writing VASP input files, running VASP with custodian[14,15] to manage common DFT errors automatically, passing information about the location of the calculations to the next step, and saving the DFT outputs to an external user database) and a step that identifies all of the possible surfaces and adds the corresponding slab calculations to the workflow. Once the slab calculations are completed, the adsorption sites are identified for each surface, and the corresponding adsorption structure Fireworks are added to the workflow. The last step is the analysis Firework for each adsorption structure, in which the data generated throughout that specific branch of the workflow is organized, new descriptors are computed, and everything is saved to a personal user database.

Surface Generation. The surface identification is performed on the output structure of the bulk DFT geometry optimization using pymatgen functionality[14,18] up to a

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

maximum Miller index that can be specified by the user and is 1 by default. With this parameter value, the surface generation algorithm will return all unique surfaces corresponding to Miller planes such as {100}, {110}, and {111}, among others. Additionally, surfaces with no center of inversion will be mirrored and included. Each surface is modeled as a slab of a small finite number of atomic layers. The dimensionality of the bulk is determined[19] as part of the surface generation step, and for two-dimensional (2D) materials, only the terminations that correspond to the van der Waals layers in the material are considered.

For each of the identified slabs, a DFT geometry optimization step is performed. Slab atoms are labeled as surface, subsurface, or bottom surface. This categorization is based on a surface height parameter, which is 0.9 Å by default and determines the thickness at the top and bottom end of the slab within which atoms are labeled as surface and bottom surface, respectively, with all remaining atoms being subsurface. By default, the position of atoms that are more than 2 Å below the surface is kept fixed during the optimization using selective dynamics. Due to the computational limitations on the number of atoms that can be tractably modeled as periodic slabs, the use of selective dynamics ensures the bulklike character of atom layers below the surface and prevents any unphysical contributions from the relaxation of the exposed bottom surface to the total electronic energy. By default, the minimum length and width of slabs are set to 10 Å, the minimum height of the slab is set to 12 Å, and the minimum vacuum is set to 20 Å. These can all be tuned at the workflow level but were chosen to ensure that slab geometry optimization calculations are operating with enough vacuum to prevent slab−slab interactions, with enough thickness to ensure the bottom layers are bulklike, and with enough length and width to prevent adsorbate−adsorbate interactions in the upcoming steps of the workflow. Following the DFT geometry optimization step, the density of states (DOS) data for the slab is generated.

In surface science, interfaces are often dynamic. Surface atoms can rearrange, forming structures very different from bulk. Additionally, surfaces can also interact with the environment, e.g., a solvent medium. The default settings of the workflow, where the surface is generated from a bulk structure, can only capture local atomic relaxations, as described by density functional theory. However, the workflow can also be started from a slab structure, bypassing the bulk calculation and surface generation. A user who wishes to model a surface which has reconstructed and/or exhibits a different stoichiometry can start from there. As our method is designed to be a high-throughput exploration of adsorption properties, we expect that further studies are needed to model surfaces under more realistic environmental conditions, especially in fields such as photocatalysis.

Adsorption Structure Construction. To generate adsorption structures, the first step is identifying all of the possible adsorption sites. This is achieved using a Delaunay triangulation method which finds all 1-, 2-, and 3-fold ensembles on a surface, as described in previous work.[9] An adsorbate−surface pair structure is generated for each adsorption site on each surface. Next, an optimal adsorbate−surface distance is found prior to the adsorbate−surface DFT optimization. This step maximizes the success rate for the upcoming optimization step by accounting for the sensitivity of the optimization convergence on the input geometry in adsorption structures of nonmetallic systems and saves computational time by indicating whether an adsorption structure is promising enough to proceed to the costly optimization step. The energy landscape as a function of adsorbate−surface distance may be shallow and exhibit one or more local minima, making it difficult for commonly used DFT minimizers to locate the desired minimum. Static calculations are therefore performed at various adsorbate−surface distances, which are, by default, 0.5, 1.0, 1.5, and 2.0 Å. The total energies obtained from these calculations are then analyzed in the next step to either determine an optimal input structure for the geometry optimization or discard that specific adsorption structure and terminate the corresponding workflow branch if no structure successfully converges at the various distances tested. As the distance increases, if none of the final energies are below 0 eV, the lowest energy−distance pair is chosen. However, if the landscape is such that the 0 eV line is crossed twice, the average of the last two distance−energy pair with a negative total energy is chosen. If the energy landscape only crosses the 0 eV boundary once, the two most negative energy−distance pairs are averaged. The user also has the option to fit distance−energy pairs to a second-order polynomial and choose the distance corresponding to the minimum, or to simply use the lowest distance−energy pair. Each adsorption structure that is retained after the distance optimization step undergoes a DFT geometry optimization calculation in which atomic layers that are 2 Å below the surface are not allowed to move for the same reasons as in the case of the slab relaxations. The adsorbate−surface optimization is followed by a similar static and uniform non-self-consistent calculation to generate DOS data.

Data Analysis and Aggregation. The final step for each successful adsorption structure is an analysis step, which organizes all of the data from the previous steps, computes new data, and saves it as a JavaScript Object Notation (JSON) document to a new entry (per adsorption structure) in an external user database (Figure 2). The wide range of data computed and saved as part of the workflow allows for an in-depth descriptor-based analysis and further screening facilitated by the accessible querying of the user database. The computed properties include descriptors relevant for photoelectrocatalysis, which can be roughly split into four categories: energy-based, geometric, electronic, and charge-based properties.

Energy-based descriptors include the cleavage energy, which is calculated as

$$E_{\text{cleavage}} = \frac{E_{\text{slab}} - \alpha E_{\text{bulk}}}{2A} \quad (1)$$

where E_{slab} and E_{bulk} are the final DFT geometry optimization energies of the slab and the bulk, respectively, α is a scale factor that accounts for differences in the number of atoms in the slab unit cell and the bulk unit cell $\left(\alpha = \frac{N_{\text{slab}}}{N_{\text{bulk}}}\right)$, and A is the exposed surface area of either termination of the slab unit cell. The cleavage energy can be used as an indication of surface stability, however, noting that it only corresponds to the surface energy in the case of slabs that have equivalent surface and bottom surface terminations.

Particularly relevant to photoelectrocatalysis is the adsorption energy, which, for simple adsorbates, has been previously related to adsorption energies of more complex

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Figure 2. Schematic of database entry for each successful adsorption structure. Saved data includes properties for the corresponding bulk and empty slab.

adsorbates (through scaling relations),[20] as well as to transition state energies (through Brønsted−Evans−Polanyi relations).[21] These correlations can offer insight into the efficiency of the reaction under study. The adsorption energy is defined as

$$E_{\text{adsorption}} = E_{\text{slab+adsorbate}} - E_{\text{slab}} - E_{\text{adsorbate}} \quad (2)$$

where $E_{\text{slab+adsorbate}}$ and E_{slab} are the final DFT geometry optimization energies of the adsorption structure and the empty slab, respectively, and $E_{\text{adsorbate}}$ is the adsorbate ground-state reference energy. This reference value is supplied by the user through a dedicated parameter.

Geometric descriptors include translation vectors, coordination numbers, adsorption site type (both in the input and output structures), the distance between the adsorbate and the surface, and the adsorbate bond lengths and angles (if applicable). These site-specific descriptors give insight into the bonding changes due to the surface−adsorbate interaction: for example, a significant translation of surface atoms upon adsorption, a lengthening bond within the adsorbate, or a change in the adsorbate bond angle could all be an indication of strong chemisorption.

Electronic descriptors focus on the DOS calculations performed after each slab and surface−adsorbate pair relaxation calculation. From the slab DOS, surface-specific descriptors are calculated, including the surface work function, the conduction band minimum (CBM) and valence band maximum (VBM) orbital and elemental make-up, and the band center of the s, p, and d bands. The work function and its change with the introduction of an adsorbate give an indication of the charge transfer into the catalyst.[22] The orbital and elemental make-up can give hints about the degree of separation of excited charge carriers upon generation.[23] The work function is calculated using the work function analyzer from pymatgen[24] by averaging the local potential across the *c* axis and shifting it with respect to the vacuum energy level. The band centers are popular surface-specific descriptors used for photocatalysts in other material systems.[25,26] From the surface−adsorbate pair DOS calculation, a new work function analysis is performed, which allows for the calculation of the site-specific work function shift upon adsorption. The DOS overlap between surface species and adsorbate species is also calculated. Depending on the character of the surface states overlapping, stronger overlap could indicate the more repulsiveness or attractiveness of the adsorbate−surface interaction.[27]

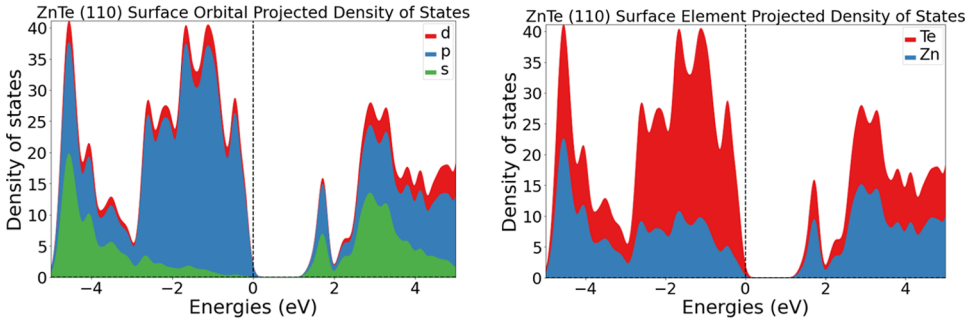
Charge-specific descriptors quantify the amount of charge transferred from the surface to the adsorbate at each site and are obtained from DDEC6[28] and Bader[29] charge partitioning analyses. Complemented by the work function shift upon adsorption, these quantities offer further insight into the extent of bonding present at the adsorption site.

Semiconductor Specificity. Modeling adsorption on semiconductor surfaces using DFT in a high-throughput manner poses challenges due to the less homogeneous electron density at the surface compared to metallic systems which can lead to numerical instability, potential dipoles due to polar terminations, and generally more complex surfaces which exhibit stronger site dependence of adsorption properties. To address the lack of already existing high-throughput adsorption methodologies that are suitable for semiconductor systems, our workflow employs a threefold approach:

1. Slabs and adsorption structures are generated from the outputs of bulk and slab relaxations, respectively, and bottom atomic layers are not allowed to move during their relaxations. This accounts for the propensity of semiconductor surfaces to relax more than metallic surfaces by forcing them to keep their bulklike positions and thus mitigates any energy contributions from the unphysical relaxation of the bottom surface, an artifact of computational limitations on the number of atomic layers that can be tractably modeled.
2. Static calculations are performed to determine the optimal adsorbate−surface distance prior to the adsorption structure relaxation. This is a relatively inexpensive way to fine-tune the input adsorption structure to maximize its relaxation convergence success, thus mitigating any issues where surface interactions are weak.
3. Descriptors specifically relevant to semiconductors are computed during the adsorption workflow, such as the CBM and VBM electronic structure, and the s-, p-, and d-band centers. Both the p-band center and d-band center have been previously used in CO_2 reduction studies as descriptors of interest.[25,26] In transition metals, the adsorption energy, in the d-band model, can be separated into the sum of two contributions: one from the broad *s*-band, and the narrow d-band typical of these materials.[26] A similar relationship has been ported to semiconductors with the p-band center.[25]

DFT Details. The DFT calculations performed throughout the workflow are carried out with the Vienna ab initio

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX



(a) Orbital (s-, p-, d-) pDOS (b) Elemental (Zn, Te) pDOS

Figure 3. Projected density of states (pDOS) at band edges for the ZnTe (110) surface, Gaussian-smoothed (σ = 0.1). Calculations show predominantly p-character at the VBM and predominantly s- and p-characters at the CBM. There is a high Te contribution to the VBM, whereas both Te and Zn states contribute to the CBM. This is in similar to other metal-chalcogenide semiconductors.

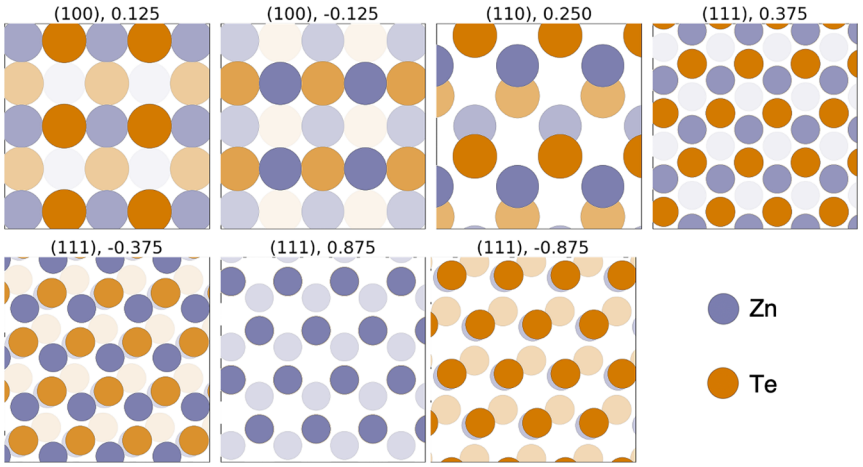


Figure 4. ZnTe surfaces generated as part of the workflow with a maximum Miller index of 1. Each surface is identified by a Miller plane, and the fractional shift is applied in the c direction to get to the specific termination.

simulation package,[16,17] using projector-augmented wave (PAW) pseudopotentials.[30,31] The default exchange−correlation functional used is RPBE, which is a revision of the Perdew−Burke−Ernzerhof (PBE)[32] functional and has been shown to give improved adsorption energies compared to its predecessor.[33] Dispersion is accounted for by Grimme's DFT-D3 correction method.[34] Slab calculations are performed using dipole corrections to the potential and forces, applied in the direction of the lattice vector along which the slab is oriented.[35] The default DFT parameters have been selected to provide reasonable accuracy while keeping the high-throughput workflow computationally tractable. All VASP calculations are performed using the custodian job management package, which identifies and corrects common VASP errors automatically.[14,15]

Other Parameters. The extensive data analysis is performed by the workflow without any user supervision or input. The only input parameters required for workflow to run are the bulk crystal structure, adsorbate structures, and adsorbate reference energies. The ground-state adsorbate structures and reference energies can be computed using the already existing dedicated atomate molecule workflow.[12] Additionally, the user can also control parameters such as VASP parameters, supercell size, vacuum size, adsorbate orientation, and maximum Miller index to consider when generating surfaces. The default values for parameters such as the slab length, width and height, and vacuum size have been chosen based on the zinc-blende-structured zinc telluride (ZnTe) benchmarking and convergence tests. Nevertheless, the use of the FireWorks framework allows for easy control over the workflow through both automated and manual changes to individual steps.

CASE STUDY: ZINC TELLURIDE

There are few semiconducting materials for which the experimental evidence of CO_2 reduction is clearly established and characterized.[36,37] Its aqueous stability under reducing conditions[38] and the existing evidence for the zinc-blende structure of zinc telluride being a promising component of CO_2RR photoelectrocatalytic systems, both experimental[39−43] and computational,[8] make ZnTe a suitable candidate for further adsorption studies. We present here the application of our workflow to zinc telluride (referenced in the Materials Project Database as mp-2176) as a case study and note that a detailed analysis of its performance as a CO_2RR photocatalyst is beyond the scope of this paper. The adsorbates under consideration are CO and H, as their adsorption energies on metallic catalytic surfaces have been well established to be strong indicators of CO_2RR and hydrogen evolution reaction (HER) capabilities, respectively,[44,45] thus warranting an investigation into the possibility of similar trends applying to semiconductors).

The computed conduction and valence band edges (Figure 3) are typical of other Zn-chalcogenide semiconductors: a

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

predominantly telluride, anion-p character for the valence band and a conduction band of both zinc and tellurium and s- and p-type character. Likewise, the valance and conductor band is rather well dispersed.[46] Additionally, for ZnTe, we calculated a cubic crystal system with a lattice constant of 6.13 Å, on par with published crystallographic data.[47]

Seven unique surfaces up to a Miller index of 1 are generated from the relaxed bulk ZnTe structure: two {100} surfaces with different terminations, one {110} surface, and four {111} surfaces with different terminations (Figure 4). The computed cleavage energies are similar in magnitude to previously reported surface energies[48] and suggest that the (110) surface is the most stable one, having the lowest cleavage energy of the seven slabs: 0.03 eV/Å^2 (0.52 J/m^2) compared to the next lowest energetic surface (111) with a cleavage energy of 0.06 eV/Å^2 (0.91 J/m^2). This is similar to previously published experimental literature in which the [110] and [111] are the typical growth directions of zinc-blende ZnTe.[48−50]

Moreover, the (110) surface in our study exhibits a surface reconstruction similar to what was previously observed experimentally, where surface tellurium anions remain close to their original positions, moving slightly outward by 0.2 Å, while zinc ions relaxed inward much more noticeably by 0.5 Å. This unique surface relaxation behavior exhibited by ZnTe compared to other zinc-blende semiconductors is attributed to its elevated ionicity.[51] Additionally, work function analysis shows the surface work function to match existing literature for ZnTe in the 5−6 eV range.[52]

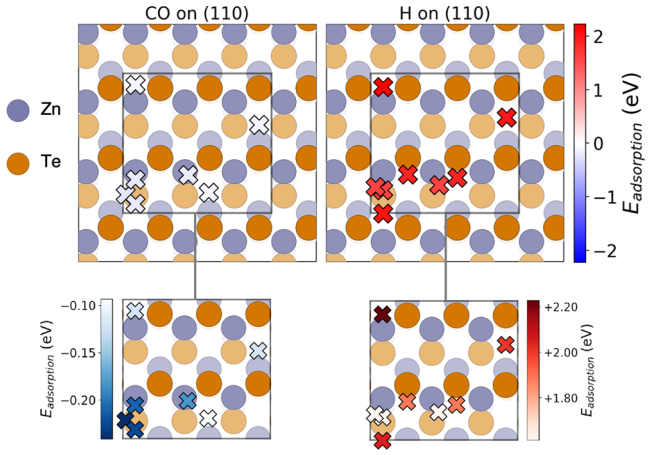


Figure 5. Adsorption energy plots for ZnTe (110): CO (left); H (right). Adsorption sites are marked by X and color coded by the adsorption energy value at that site: the data points on the top plots are on the same scale, whereas on the bottom, they are shown on narrower adsorption energy scales. One data point was excluded from the CO adsorption plot because the adsorbate dissociated at that site during the geometry optimization.

A total of 44 unique adsorption sites have been identified for the seven relaxed slabs, resulting in 44 CO adsorption structures and 44 H adsorption structures. Out of these, 23 CO structures and 28 H structures successfully converged in the geometry optimization step. The highest convergence rates were obtained for the most stable surfaces. We note that since unfavorable adsorption sites are predicted to lead to unstable input structures for geometry optimizations and hence a low probability of convergence to a minimum, an overall limited convergence rate for the DFT optimizations performed as part of the workflow is to be expected and does not impede its purpose. In total, over 700 Fireworks were launched as part of the ZnTe workflow with CO and H as adsorbates. Using a single 64-core node for bulk, slab, distance optimizations, and density of states' calculations, and two 64-core nodes for adsorption structure optimizations, the workflow run finished within a few weeks. Typically, bulk relaxations are completed in less than 2 h, slab relaxations in 15−25 h, static calculations in less than 4 h each on average, and slab−adsorbate calculations within 48 h. Density of states' calculations and static calculations finished in less than 3 h on average.

The use of the surface−adsorbate distance optimization algorithm was instrumental in the success rate of adsorption structure geometry optimizations, compared to the alternative (i.e., using the default 2 Å distance between the surface and the adsorbate for all input structures). For 6 out of the 23 CO adsorption structures and 4 out of the 28 H adsorption structures that were successfully relaxed to an equilibrium geometry, the energy landscape generated by the distance optimization step revealed a local maximum between the default 2 Å distance and the chosen optimal distance.

On the (110) ZnTe surface, eight unique sites were found by the site-finding Firetask. The adsorption energies for both CO and H on these eight sites are depicted in Figure 5. The two plots reveal weak CO adsorption and unfavorable interactions with H at all adsorption sites. No site exhibited significant charge transfer to the CO molecule, based on the Bader analysis. An example of an adsorption site's charge density differential between the adsorbate−surface pair and the (110) surface and adsorbate is shown in Figure 6. In this example, the bond between the adsorbate species has been depleted of electron density while C and the O have gained electron density outside of the bond. The lack of charge transfer to the CO absorbate combined with the weak CO-

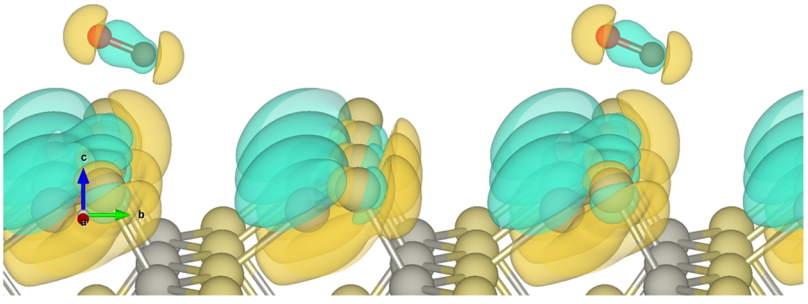


Figure 6. Difference of charge density between adsorbate−surface pair, adsorbate and surface for a CO molecule on a (110) site on ZnTe.

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

surface binding and unfavorable H-surface interactions on the (110) surface is commensurate with the previously experimentally observed CO and H being the predominant CO_2 reduction products on ZnTe.[41,43]

CONCLUSIONS

To address the lack of high-throughput adsorption methods for semiconductor surfaces, we have developed, as part of this work, an automated adsorption workflow suitable for such systems. Semiconductor specificity is achieved through the rigorous design of input structures for DFT optimizations, restrictions applied to DFT calculations, and the computation of descriptors relevant to semiconductors. While this work is primarily motivated by the need to generate comprehensive adsorption data sets for the purpose of photocatalytic material discovery, the workflow is suitable for any applications involving adsorption processes.

We applied the workflow to zinc telluride due to its potential as a CO_2RR photoelectrocatalytic material and presented the results here as a case study. Seven surfaces were identified, with a total of 44 unique adsorption sites, leading to 88 adsorption structures for the two considered adsorbates (CO, H). Overall, more than 700 distinct computations spanning bulk, slab, and adsorption structure DFT relaxation calculations and analysis tasks were completed as part of the workflow, with 51 adsorption structures successfully undergoing a geometry optimization. Computed properties such as equilibrium geometries, cleavage energies, adsorption energies, and charge transfer values were found to be in accordance with or support previously published experimental data, showing the reliability and robustness of the developed autonomous workflow. While it is beyond the scope of this paper to delve into an analysis of the generated ZnTe adsorption data to assess its CO_2RR capabilities, the presented case study establishes our adsorption workflow as a robust method for a previously overlooked class of materials, encompassing minimal user supervision, exhaustive surface and adsorption structure generation, and extensive data analysis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00340.

Workflow parameters' selection, VASP details for ZnTe calculations, and instructions on how to access and run the workflow (PDF)

AUTHOR INFORMATION

Corresponding Author

Kristin A. Persson − *Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States; Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0003-2495-5509; Email: kapersson@lbl.gov

Authors

Oxana Andriuc − *Department of Chemistry, University of California, Berkeley, California 94720, United States; Liquid Sunlight Alliance and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0002-4011-3339

Martin Siron − *Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States; Liquid Sunlight Alliance and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Toyota Research Institute, Los Altos, California 94022, United States;* orcid.org/0000-0002-4562-7814

Joseph H. Montoya − *Toyota Research Institute, Los Altos, California 94022, United States;* orcid.org/0000-0001-5760-2860

Matthew Horton − *Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States; Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0001-7777-8871

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00340

Author Contributions

$^{\nabla}$O.A. and M.S. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the Liquid Sunlight Alliance, which is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Fuels from Sunlight Hub under Award Number DE-SC0021266 and by the Joint Center for Artificial Photosynthesis, a DOE Energy Innovation Hub, supported through the Office of Science of the U.S. Department of Energy under Award Number DE-SC0004993. This work used data and software infrastructure from the Materials Project, which is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under Contract DE-AC02-05CH11231: Materials Project program KC23MP. Computations in this paper were performed using resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 and the Lawrencium computational cluster resource provided by the IT Division at the Lawrence Berkeley National Laboratory (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). M.S. acknowledges support from the National Science Foundation Graduate Research Fellowship under grant no. DGE 1752814.

REFERENCES

(1) Emmett, P. H. Adsorption and Catalysis. *J. Phys. Chem. A* **1959**, *63*, 449−456.

(2) Li, Y.; Wu, D.; Zhou, Z.; Cabrera, C. R.; Chen, Z. Enhanced Li Adsorption and Diffusion on MoS_2 Zigzag Nanoribbons by Edge Effects: A Computational Study. *J. Phys. Chem. Lett.* **2012**, *3*, 2221−2227.

(3) Suntivich, J.; Gasteiger, H. A.; Yabuuchi, N.; Nakanishi, H.; Goodenough, J. B.; Shao-Horn, Y. Design Principles for Oxygen-

reduction Activity on Perovskite Oxide Catalysts for Fuel Cells and Metal-air Batteries. *Nat. Chem.* **2011**, *3*, 546−550.
(4) Liu, L.; Corma, A. Metal Catalysts for Heterogeneous Catalysis: From Single Atoms to Nanoclusters and Nanoparticles. *Chem. Rev.* **2018**, *118*, 4981−5079.
(5) Ruffino, F.; Giannazzo, F. A Review on Metal Nanoparticles Nucleation and Growth on/in Graphene. *Crystals* **2017**, *7*, No. 219.
(6) Xu, S.; Carter, E. A. Theoretical Insights into Heterogeneous (Photo)electrochemical CO_2 Reduction. *Chem. Rev.* **2019**, *119*, 6631−6669.
(7) Pan, J.; Yan, Q. Data-driven Material Discovery for Photocatalysis: A Short Review. *J. Semicond.* **2018**, *39*, No. 071001.
(8) Singh, A. K.; Montoya, J. H.; Gregoire, J. M.; Persson, K. A. Robust and Synthesizable Photocatalysts for CO_2 Reduction: A Data-driven Materials Discovery. *Nat. Commun.* **2019**, *10*, No. 443.
(9) Montoya, J. H.; Persson, K. A. A High-throughput Framework for Determining Adsorption Energies on Solid Surfaces. *npj Comput. Mater.* **2017**, *3*, No. 14.
(10) Tran, K.; Palizhati, A.; Back, S.; Ulissi, Z. W. Dynamic Workflows for Routine Materials Discovery in Surface Science. *J. Chem. Inf. Model.* **2018**, *58*, 2392−2400.
(11) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Identifying Promising Metal-organic Frameworks for Heterogeneous Catalysis via High-throughput Periodic Density Functional Theory. *J. Comput. Chem.* **2019**, *40*, 1305−1318.
(12) Mathew, K.; et al. Atomate: A High-level Interface to Generate, Execute, and Analyze Computational Materials Science Workflows. *Comput. Mater. Sci.* **2017**, *139*, 140−152.
(13) Jain, A.; Ong, S. P.; Chen, W.; Medasani, B.; Qu, X.; Kocher, M.; Brafman, M.; Petretto, G.; Rignanese, G.-M.; Hautier, G.; Gunter, D.; Persson, K. A. Fireworks: A Dynamic Workflow System Designed for High-throughput Applications: Fireworks: A Dynamic Workflow System Designed for High-Throughput Applications. *Concurrency Computat.: Pract. Exper.* **2015**, *27*, 5037−5059.
(14) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A Robust, Open-source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314−319.
(15) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; et al. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, 011002.
(16) Kresse, G.; Furthmüller, J. Efficiency of Ab-initio Total Energy Calculations for Metals and Semiconductors Using a Plane-wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15−50.
(17) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-energy Calculations Using a Plane-wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169−11186.
(18) Sun, W.; Ceder, G. Efficient Creation and Convergence of Surface Slabs. *Surf. Sci.* **2013**, *617*, 53−59.
(19) Gorai, P.; Toberer, E. S.; Stevanović, V. Computational Identification of Promising Thermoelectric Materials among Known Quasi-2D Binary Compounds. *J. Mater. Chem. A* **2016**, *4*, 11110−11116.
(20) Nøkov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the Computational Design of Solid Catalysts. *Nat. Chem.* **2009**, *1*, 37−46.
(21) Cheng, J.; Hu, P.; Ellis, P.; French, S.; Kelly, G.; Lok, C. M. Brønsted- Evans- Polanyi Relation of Multistep Reactions and Volcano Curve in Heterogeneous Catalysis. *J. Phys. Chem. C* **2008**, *112*, 1308−1311.
(22) Burshtein, R. C.; Shurmovskaya, N. A. The Effect of Electronegative Gases on the Work Function of a Metal. *Surf. Sci.* **1964**, *2*, 210−216.
(23) Bai, Y.; Zhang, Q.; Xu, N.; Deng, K.; Kan, E. Efficient Carrier Separation and Band Structure Tuning of Two-Dimensional C_2N/GaTe van der Waals Heterostructure. *J. Phys. Chem. C* **2018**, *122*, 15892−15902.
(24) Tran, R.; Xu, Z.; Radhakrishnan, B.; Winston, D.; Sun, W.; Persson, K. A.; Ong, S. P. Surface Energies of Elemental Crystals. *Sci. Data* **2016**, *3*, No. 160080.
(25) Jacobs, R.; Hwang, J.; Shao-Horn, Y.; Morgan, D. Assessing Correlations of Perovskite Catalytic Performance with Electronic Structure Descriptors. *Chem. Mater.* **2019**, *31*, 785−797.
(26) Vojvodic, A.; Nørskov, J.; Abild-Pedersen, F. Electronic Structure Effects in Transition Metal Surface Chemistry. *Top. Catal.* **2014**, *57*, 25−32.
(27) Gauthier, J. A.; Fields, M.; Bajdich, M.; Chen, L. D.; Sandberg, R. B.; Chan, K.; Nørskov, J. K. Facile Electron Transfer to CO_2 during Adsorption at the Metal—Solution Interface. *J. Phys. Chem. C* **2019**, *123*, 29278−29283.
(28) Manz, T. A.; Limas, N. G. Introducing DDEC6 Atomic Population Analysis: Part 1. Charge Partitioning Theory and Methodology. *RSC Adv.* **2016**, *6*, 47771−47801.
(29) Henkelman, G.; Arnaldsson, A.; Jónsson, H. A Fast and Robust Algorithm for Bader Decomposition of Charge Density. *Comput. Mater. Sci.* **2006**, *36*, 354−360.
(30) Blöchl, P. E. Projector Augmented-wave Method. *Phys. Rev. B* **1994**, *50*, 17953−17979.
(31) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-wave Method. *Phys. Rev. B* **1999**, *59*, 1758−1775.
(32) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(33) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved Adsorption Energetics Within Density-functional Theory Using Revised Perdew-Burke-Ernzerhof Functionals. *Phys. Rev. B* **1999**, *59*, 7413−7421.
(34) 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*, No. 154104.
(35) Neugebauer, J.; Scheffler, M. Adsorbate-substrate and Adsorbate-adsorbate Interactions of Na and K Adlayers on Al(111). *Phys. Rev. B* **1992**, *46*, 16067−16080.
(36) Habisreutinger, S. N.; Schmidt-Mende, L.; Stolarczyk, J. K. Photocatalytic Reduction of CO_2 on TiO_2 and Other Semiconductors. *Angew. Chem., Int. Ed.* **2013**, *52*, 7372−7408.
(37) Lingampalli, S. R.; Ayyub, M. M.; Rao, C. N. R. Recent Progress in the Photocatalytic Reduction of Carbon Dioxide. *ACS Omega* **2017**, *2*, 2740−2748.
(38) Brookins, D. G. *Eh-pH Diagrams for Geochemistry*; Springer Science & Business Media, 2012.
(39) Jang, J.-W.; Cho, S.; Magesh, G.; Jang, Y. J.; Kim, J. Y.; Kim, W. Y.; Seo, J. K.; Kim, S.; Lee, K.-H.; Lee, J. S. Aqueous-Solution Route to Zinc Telluride Films for Application to CO_2 Reduction. *Angew. Chem., Int. Ed.* **2014**, *53*, 5852−5857.
(40) Ehsan, M. F.; Ashiq, M. N.; Bi, F.; Bi, Y.; Palanisamy, S.; He, T. Preparation and Characterization of $SrTiO_3$-ZnTe Nanocomposites for the Visible-light Photoconversion of Carbon Dioxide to Methane. *RSC Adv.* **2014**, *4*, 48411−48418.
(41) Ehsan, M. F.; He, T. In Situ Synthesis of ZnO/ZnTe Common Cation Heterostructure and Its Visible-light Photocatalytic Reduction of CO_2 into CH_4. *Appl. Catal. B* **2015**, *166−167*, 345−352.
(42) Jang, Y. J.; Lee, J.; Lee, J.; Lee, J. S. Solar Hydrogen Production from Zinc Telluride Photocathode Modified with Carbon and Molybdenum Sulfide. *ACS Appl. Mater. Interfaces* **2016**, *8*, 7748−7755.
(43) Jang, Y. J.; Jeong, I.; Lee, J.; Lee, J.; Ko, M. J.; Lee, J. S. Unbiased Sunlight-Driven Artificial Photosynthesis of Carbon Monoxide from CO_2 Using a ZnTe-Based Photocathode and a Perovskite Solar Cell in Tandem. *ACS Nano* **2016**, *10*, 6980−6987.
(44) Peterson, A. A.; Nørskov, J. K. Activity Descriptors for CO_2 Electroreduction to Methane on Transition-Metal Catalysts. *J. Phys. Chem. Lett.* **2012**, *3*, 251−258.

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(45) Parsons, R. The Rate of Electrolytic Hydrogen Evolution and the Heat of Adsorption of Hydrogen. *Trans. Faraday Soc.* **1958**, *54*, 1053.
(46) Woods-Robinson, R.; Han, Y.; Zhang, H.; Ablekim, T.; Khan, I.; Persson, K. A.; Zakutayev, A. Wide Band Gap Chalcogenide Semiconductors. *Chem. Rev.* **2020**, *120*, 4007−4055.
(47) Rabadanov, M. K. Anharmonic and "Disorder" Models of ZnTe Structure: High-temperature Single-crystal X-ray Diffraction Studies. *Kristallografiâ* **1995**, *40*, 505−510.
(48) Sun, Y.; Zhao, Q.; Gao, J.; Ye, Y.; Wang, W.; Zhu, R.; Xu, J.; Chen, L.; Yang, J.; Dai, L.; Liao, Z.-m.; Yu, D. In Situ Growth, Structure Characterization, and Enhanced Photocatalysis of High-quality, Single-crystalline ZnTe/ZnO Branched Nanoheterostructures. *Nanoscale* **2011**, *3*, 4418−4426.
(49) Lee, S. H.; Kim, Y. J.; Park, J. Shape Evolution of ZnTe Nanocrystals: Nanoflowers, Nanodots, and Nanorods. *Chem. Mater.* **2007**, *19*, 4670−4675.
(50) Shaygan, M.; Gemming, T.; Bezugly, V.; Cuniberti, G.; Lee, J.-S.; Meyyappan, M. In Situ Observation of Melting Behavior of ZnTe Nanowires. *J. Phys. Chem. C* **2014**, *118*, 15061−15067.
(51) Meyer, R. J.; Duke, C. B.; Paton, A.; So, E.; Yeh, J. L.; Kahn, A.; Mark, P. Surface Structure of ZnTe (110) as Determined from Dynamical Analysis of Low-energy-electron Diffraction Intensities. *Phys. Rev. B* **1980**, *22*, 2875−2886.
(52) Swank, R. K. Surface Properties of II-VI Compounds. *Phys. Rev.* **1967**, *153*, 844−849.

https://doi.org/10.1021/acs.jcim.1c00340
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX