Creation of an Fe_3P Scheibersite Density Functional Tight Binding Model for Astrobiological Simulations

Riccardo Dettori*,†,‡ and Nir Goldman*,†,¶

†*Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, United States*

‡*Department of Physics, University of Cagliari, Monserrato, CA, 09042 Italy*

¶*Department of Chemical Engineering, University of California, Davis, California 95616, United States*

E-mail: riccardo.dettori@dsf.unica.it; ngoldman@llnl.gov

Abstract

The mineral schreibersite, e.g., Fe_3P, is commonly found in iron-nickel meteorites and could have served as an abiotic phosphorus source for prebiotic chemistry. However, atomistic calculations of its degradation chemistry generally require quantum simulation approaches, which can be too computationally cumbersome to study sufficient time and length scales for this process. In this regard, we have created a computationally efficient semi-empirical quantum Density Functional Tight Binding (DFTB) model for iron and phosphorus-containing materials by adopting an existing semiautomated workflow that represents many-body interactions by linear combinations of Chebyshev polynomials. We have utilized a relatively small training set to optimize a DFTB model that is accurate for schreibersite physical and chemical properties, including its bulk properties, surface energies, and water bulk absorption. We then show

that our model shows strong transferability to several iron phosphide solids as well as multiple allotropes of iron metal. Our resulting DFTB parameterization will allow us to interrogate schreibersite aqueous decomposition at longer time and length scales than standard quantum approaches, allowing for the potential to investigate its role in prebiotic chemistry on early Earth.

Introduction

Phosphorous plays a central role in the formation of many life-building compounds, including nucleotides, phospholipids, adenosine triphosphate (ATP) for metabolic energy storage, and calcium phosphate (apatite) found in bone.[1–5] Extraterrestrial sources such as meteoritic impacts could have served as one possible source of prebiotic phosphorous on early Earth,[4] including those with the mineral schreibersite, $(Fe,Ni)_3$-P.[6,7] Schreibersite is a ferromagnetic material with tetrahedral symmetry (orthogonal lattice constants with $a = b > c$). It has been suggested to play a role in the formation of prebiotic organics,[4,6,8,9] where experimental studies have shown that it can react with a wide variety of organic compounds and can catalyze the formation of amino acids, sugars, and nucleotides.[1,9,10] However, early Earth likely experienced a wide range of thermodynamic conditions, including electrical storms[11] and extreme thermodynamic conditions due to impacts,[12–14] that are difficult to probe with experimental trial and error alone. Schreibersite degradation studies would thus benefit from atomistic simulations, which identify specific reaction mechanisms and kinetics, and can help narrow the relevant thermodynamic conditions for prebiotic synthesis.

To date, computational studies on schreibersite have largely focused on quantum simulations of bulk properties and water surface adsorption using Kohn-Sham Density Functional Theory (DFT).[15–17] DFT is known to yield accurate descriptions of condensed phase physical and chemical data, as well as the energetics of chemical bond breaking and forming under reactive conditions (e.g., Refs. 12,14). However, DFT simulations generally require heavy computational effort, which limits calculations to systems containing few hundreds of

atoms and to molecular dynamics (MD) timescales of tens of picoseconds. These scales are generally orders of magnitude smaller than those probed by experiments, where for example, chemistry at elevated pressure and temperature can nanoseconds or longer to equilibrate,[18,19] and extended structures can form that are nanometer-scale in diameter.[20,21] Reactive MD models can help alleviate this time and length scale gap optimizing empirical[22] or machine-learning functional forms[23–26] to DFT or other quantum mechanical data. These methods are generally fit to the properties of specific chemical reactants over a narrow range of thermodynamic conditions. As a result, they can yield high computational efficiency, but tend to yield completely unphysical results outside of their fitting regime.

The Density Functional Tight Binding method (DFTB) represents a promising alternative for simulations of reactive materials. DFTB is an approximate quantum simulation technique whereby the majority of the electronic interactions are represented by a minimal atomic basis set and are precomputed and mapped to a radial grid.[27–29] The DFTB total energy expression contains an empirical repulsive energy E_{rep} that can be fit to experimental or high-level computational data in order to extend the technique to a wide variety of materials and conditions.[30] There is considerable flexibility in the E_{Rep} functional form, including cubic splines,[31] Gaussian Process Regression,[32] many-body polynomial based representations,[33] and neural network potentials.[34] The DFTB methodology allows for several orders of magnitude increase in computational efficiency while retaining most of the accuracy of standard DFT-MD simulations. Consequently, DFTB simulations can achieve close to chemical equilibrium timescales for reactivity under extreme conditions ($\sim$ 1 ns) using relatively low computational effort. One of the main limitations of DFTB, though, is the availability of existing repulsive energies in literature, as one repository of interaction parameters[27,28,35–44] does not contain specific interactions between iron and phosphorus. In addition, the DFTB method itself has several hyperparameters, mainly the wavefunction and density confining radii (discussed below), that require challenging optimizations as well.

In this work, we overcome these issues for Fe_3P schreibersite by determining the Fe-Fe

and Fe-P DFTB interaction potentials through a semi-automated workflow that leverages the Chebyshev Interaction Model for Efficient Simulation (ChIMES).[45–47] ChIMES is a many-body reactive force field model for MD simulation based on linear combinations of Chebyshev polynomials. This linear parameterization allows for rapid model optimization, allowing us to cycle through large numbers of DFTB and ChIMES hyperparameters. DFTB/ChIMES models have been created for a wide variety of systems, including TiH_2,[48] energetic materials,[19] actinides,[49] and as a general purpose model for organic materials.[50] Here, we adopt the ChIMES framework to determine iron-iron and iron-phosphorus interaction for DFTB, specifically. We first briefly review the DFTB and ChIMES approaches and explain our workflow for optimization of both DFTB and CHIMES hyperparameters. We then downselect our model through validation against a wide range of DFT that was not included in our training set, such as the Fe_3P lattice constants, bulk moduli, surface facet energies, and bulk water absorption. We further examine our optimal DFTB/ChIMES by computing surface facet energies, water surface adsorption on different facets, and MD simulation at ambient and extreme conditions. Finally, we explore the transferability of our model through the prediction of properties for different materials, including FeP, Fe_2P, and three allotropes of iron. Our results indicate that our DFTB/ChIMES model yields comparable accuracy to DFT for properties related to prebiotic synthesis, including the intense pressures and temperatures that can be accessed during planetary impact.

Methods

DFT calculations and training set construction

All DFT simulations were performed using the Vienna Ab initio Simulation Package (VASP),[51–53] with projector augmented wave (PAW) pseudopotentials[54,55] and the Perdew, Burke, and Ernzerhof (PBE) exchange-correlation functional.[56] Partial occupancies of the electronic states were set with fourth-order Methfessel-Paxton smearing,[57] using a width of 0.2 eV. We

observed converged energies for the bulk system with a plane-wave energy cut-off of 500 eV, with a self-consistent field (SCF) convergence criteria of 10^{-6} eV and a $6 \times 6 \times 6$ **k**-point Monkhorst-Pack mesh sampling of Brillouin zone.[58] The force convergence tolerance for geometry optimizations was set to 0.01 eV/Å for each atom in each direction. Our calculations were performed with collinear spin polarization initialized in a ferromagnetic configuration, which was maintained during all optimizations. As mentioned is Ref. 15, the inclusion of van der Waals (vdW) correction to our calculations resulted in a constant offset of the water absorption energies, with minimal change to optimized configurations. Furthermore, the observed effect is strongly dependent upon the choice of the vdW functional.[59] Hence, for the model development in this study, we have chosen to adopt standard PBE functional without vdW, though we note that those effects can be included at a later date.

Our training set was built by performing DFT-MD simulations of a 32 atom Fe_3P schreibersite supercell, initially optimized to dimensions of 9.045Å ×9.045Å ×4.380Å, yielding a density of 7.36 g/cm^3. We then used the relaxed structure to initiate 20 ps long DFT-MD trajectories with a timestep of 1.0 fs at temperatures of $300, 1500$, and 2000 K and pressures of $0, 10, 20, 30$, and 40 GPa. From these simulations, we then aggregated the following 310 training configurations in total:

- 120 equally spaced configurations at $P = 0$ GPa and $T = 300, 1500$, and 2000 K.
- 120 equally spaced configurations at $P = 10, 20, 30$, and 40 GPa and $T = 1500$ K.
- 70 configurations obtained by randomly displacing by 0.1 Å the atoms of the optimized Fe_3P structure, similar to previous reactive MD model development.[26]

DFTB calculations

DFTB calculations were performed with the DFTB+ code[29] using orbital resolved self-consistent charge calculations (SCC)[60] with a convergence criteria of 2.72×10^{-5} eV (10^{-6} au). We adopted the same electron thermal smearing and **k**-point mesh of our DFT calculations.

Similarly, all DFTB calculations were performed with collinear spin polarization initialized in a ferromagnetic spin geometry. The DFTB formalism has been extensively discussed in literature,[27–29,61,62] with a brief outline provided here. The total energy expression with SCC is derived by assuming spherically symmetric charge densities and then expanding the Kohn-Sham total DFT total energy expression to either second[27] or third order[63] about small charge fluctuations. We note that in this work we rely on the second-order charge fluctuation model only.

The resulting total energy expression reads as:

$$E_{\mathrm{DFTB}} = E_{\mathrm{BS}} + E_{\mathrm{Coul}} + E_{\mathrm{rep}}. \tag{1}$$

Here, E_{BS} is the band structure energy, which is computed from the Kohn-Sham eigenstates determined from diagonalization of the approximate Hamiltonian with a minimal, Slater-type basis set. The term E_{Coul} corresponds to the charge-transfer energy from second-order charge fluctuations and is computed self-consistently. Finally, E_{rep} is the repulsive energy, which contains the ion-ion repulsions, as well as Hartree and exchange-correlation double counting terms. E_{rep} is generally represented as a short-ranged empirical function that includes first coordination shell or bonded interactions only. Its parameters are usually fit to reproduce DFT or experimental data, and it can describe pairwise[30,64] or many-body interactions.[19,33,49,50,65] The Hamiltonian and overlap matrices used in E_{BS} are determined from pre tabulated Slater-Koster (SK) tables that assume two-center interactions, only. However, prior to SK tabulation, both the electronic wave functions and electron density are subjected to separate confining or compression potentials, which has the effect of improving transferability of the calculations. For our work, we use the following simple power function for confinement:

$$V_{conf}(r) = \left(\frac{r}{R_\chi}\right)^2, \tag{2}$$

where $\chi = \Psi$ or n, R_Ψ corresponds to wavefunction compression, and R_n to density compression. Previous efforts have used larger exponents in the power function or alternate functional forms (e.g., Ref. 66, which can be investigated in future work.

In general, optimal choice of confining radii is system-dependent and cannot be determined a priori, frequently requiring a semi-exhaustive search.[33,67] In our study, we choose to focus on optimizing the Fe density and wavefunction confining radii, only, in order to make our search more tractable. Similarly, we have only created E_{Rep} models for Fe-Fe and Fe-P interactions. The remaining DFTB+ interaction potentials (including tabulated Hamiltonian and overlap matrices, and/or E_{Rep} functions) were taken from known parameter sets. All non-iron interactions, such as those involving P, O, and H were taken from the mio-1-1 parameter set.[27,63] The Fe-O and Fe-H repulsive energies were determined by first computing new tabulated quantum interactions using our choice of confining radii for Fe and those from mio-1-1 for O and H. This was followed by matching energetics for dimer calculations from the trans3d parameter set[44] to approximate the repulsive energies.

ChIMES Methodology and Many-body E_{Rep} determination

The ChIMES method has been discussed in detail elsewhere,[45–47,67] and hence will only be given a short introduction here. The starting point in ChIMES is a many-body expansion of the total energy of the system:

$$E_{\mathrm{DFT}} = \sum_{i_1}^{n_\mathrm{a}} {}^1E_{i_1} + \sum_{i_1>i_2}^{n_\mathrm{a}} {}^2E_{i_1 i_2} + \sum_{i_1>i_2>i_3}^{n_\mathrm{a}} {}^3E_{i_1 i_2 i_3} + \sum_{i_1>i_2>i_3>i_4}^{n_\mathrm{a}} {}^4E_{i_1 i_2 i_3 i_4} + \cdots + \sum_{i_1>i_2\ldots\ i_{n_\mathrm{B}-1}>i_{n_\mathrm{B}}}^{n_\mathrm{a}} {}^{n_\mathrm{B}}E_{i_1 i_2 \ldots i_{n_\mathrm{B}}}. \tag{3}$$

Here, the index n_a allows for a sum over the total number of atoms. The one-body contributions (${}^1E_{i_1}$) are element-specific constants, the two-body or dimer contributions (${}^2E_{i_1 i_2}$) correspond to pair-wise interactions, the three-body or trimer contributions (${}^3E_{i_1 i_2 i_3}$) to triplet energies, etc., until a user-defined maximum of n_B is reached.

We now express each of the greater than one-body terms as a sum of linear combinations

of Chebyshev polynomials of the first kind. In this case, $T_m(x)$ corresponds to a Chebyshev polynomial of order m, and $x = \cos\theta$, which limits the Chebyshev polynomial range to $[-1:1]$. For the 2B energy ${}^2E_{i_1 i_2}$, we use the following expression:

$$ {}^2E_{i_1 i_2} = f_{\mathrm{p}}\left(r_{i_1 i_2}\right) + f_{\mathrm{c}}^{e_{i_1} e_{i_2}}\left(r_{i_1 i_2}\right) \sum_{m=1}^{N_{2B}} C_m^{e_{i_1} e_{i_2}} T_m(s_{i_1 i_2}^{e_{i_1} e_{i_2}}) \tag{4} $$

Starting with the terms in the summation, $C_m^{e_{i_1} e_{i_2}}$ is the permutationally invariant coefficient for mth order polynomial (taken to maximum order of N_{2B}) for pairwise interactions between atoms i and j, which are represented by element types e_{i_1} and e_{i_2}. The term $s_{i_1 i_2}^{e_{i_1} e_{i_2}}$ is the transformed pairwise distance $(r_{i_1 i_2})$ such that the user-defined radial range $\left[r_{\min}^{e_{i_1} e_{i_2}} : r_{\max}^{e_{i_1} e_{i_2}}\right]$ is scaled to $[-1:1]$, using an element-pair specific scaling function. Similar to previous work,[45] we rely on the Morse-like exponential function, where $r_{i_1 i_2}$ is scaled by a user chosen constant $\lambda_{e_1 e_2}$, i.e., $s_{i_1 i_2}^{e_{i_1} e_{i_2}} \propto \exp\left(-r_{i_1 i_2}/\lambda_{e_1 e_2}\right)$. The value of $\lambda_{e_1 e_2}$ is usually taken to be the peak position of the first coordination shell from the radial distribution function for that element pair. We ensure smooth behavior and zero energy at the radial boundary $(r_{\max}^{e_{i_1} e_{i_2}})$ through the term $f_{\mathrm{c}}^{e_{i_1} e_{i_2}}(r_{i_1 i_2})$, which we take to be is a Tersoff smoothing function.[68] Lastly, we use a smoothly varying penalty function $f_{\mathrm{p}}(r_{i_1 i_2})$ in order to prevent sampling of distances below the minimum value sampled in our training set, thus preventing determination of Chebyshev polynomials outside of their natural range.

Greater than two-body terms can be determined through creation of many-body orthogonal polynomials. These can be obtained by defining a cluster of size n and taking the tensorial product of the Chebyshev polynomials derived from the constituent $\binom{n}{2}$ unique pairs. In this case, the three-body polynomials will be products of $\binom{3}{2} = 3$ two-body polynomials, i.e.:

$$ {}^3E_{i_1 i_2 i_3} = f_{\mathrm{c}}^{e_{i_1} e_{i_2}}\left(r_{i_1 i_2}\right) f_{\mathrm{c}}^{e_{i_1} e_{i_3}}\left(r_{i_1 i_3}\right) f_{\mathrm{c}}^{e_{i_2} e_{i_3}}\left(r_{i_2 i_3}\right) \sum_{m=0}^{N_{3B}} \sum_{p=0}^{N_{3B}} \sum_{q=0}^{N_{3B}{}'} C_{mpq}^{e_{i_1} e_{i_2} e_{i_3}} T_m\left(s_{i_1 i_2}^{e_{i_1} e_{i_2}}\right) T_p\left(s_{i_1 i_3}^{e_{i_1} e_{i_3}}\right) T_q\left(s_{i_2 i_3}^{e_{i_2} e_{i_3}}\right) \tag{5} $$

The triple sum over the three-body polynomials is taken over the hypercube up to a prede-

fined order N_{3B}, and the primed sum denotes that we include only terms for which two or more of the polynomial powers are greater than zero (to guarantee that three distinct triplets are evaluated). Four-body (4B) terms are similarly included in ChIMES optimizations,[26] where ${}^4E_{i_1 i_2 i_3 i_4}$ is now determined from the sum over the product of the $\binom{4}{2} = 6$ constituent pair-wise polynomials multiplied by a single permutationally invariant coefficient.

The training set for the E_{Rep} was then created by computing the DFTB quantum mechanical terms for each configuration (e.g., E_{BS} and E_{Coul}) and subtracting these quantities from the corresponding DFT data. This generally includes the cartesian forces, $\vec{F}$, stress tensor components, $\bar{\bar{\sigma}}$, and the total system energy. Optimal ChIMES parameters (the coefficients of linear combination) can then readily be determined through linear least squares fitting of the overdetermined matrix equation $\boldsymbol{wAC} = \boldsymbol{wB}_{\text{rep}}$, The matrix $\boldsymbol{A}$ corresponds to the derivatives of the ChIMES energy or force expression with respect to the fitting coefficients (i.e., the unweighted polynomial values, including the smoothing function). The column vectors $\boldsymbol{C}$ and $\boldsymbol{B}_{\text{rep}}$ correspond to the linear ChIMES coefficients for which we are solving and the numerical values for the training data, respectively. The symbol $\boldsymbol{w}$ corresponds to a diagonal matrix of weights to be applied to the elements of $\boldsymbol{B}_{\text{rep}}$ and rows of $\boldsymbol{A}$.

The Chebyshev polynomial coefficients can then be solved for by minimizing the objective function:

$$F_{\text{obj}} = \sqrt{\frac{1}{N_d} \times \left(\sum_{\tau=1}^{M} \sum_{i=1}^{n_a} \sum_{\alpha=1}^{3} w_{\text{F}_{\alpha_i}}^2 \left(\Delta \text{F}_{\alpha_i}\right)^2 + \sum_{\tau=1}^{M} \sum_{\alpha=1}^{3} w_{\sigma_{\alpha\alpha}}^2 \left(\Delta \sigma_{\alpha\alpha}\right)^2 + \sum_{\tau=1}^{M} w_{\text{E}_\tau}^2 \left(\Delta E_\tau\right)^2 \right)}. \tag{6}$$

Here, M is the total number of configurations in the training set, N_d is the total number of data entries ($3Mn_a$ force components plus $3M$ stress tensor components plus M energy components), the w values correspond to the weights, and $\Delta F_{\alpha_i} = F^{\tau}_{\text{ChIMES}_{\alpha_i}} - F^{\tau *}_{\text{Rep}_{\alpha_i}}$, $\Delta \sigma_{\alpha\beta} = \sigma^{\tau}_{\text{ChIMES}_{\alpha\beta}} - \sigma^{\tau *}_{\text{Rep}_{\alpha\beta}}$, and $\Delta E_i = E^{\tau}_{\text{ChIMES}} - E^{\tau *}_{\text{Rep}}$. In this case, $F^{\tau *}_{\text{Rep}_{\alpha_i}}$, $\sigma^{\tau *}_{\text{Rep}_{\alpha\beta}}$, and $E^{\tau *}_{\text{Rep}}$ correspond to the residuals determined by the difference between the DFT computed quantities and those computed from DFTB without the repulsive energy for ith atom in the τth configuration, and α/β correspond to the Cartesian direction. Previous work has indi-

cated that forces and the diagonal stress tensor components can be the most significant data for E_{Rep} determination.[19,30,67] Consequently, for Fe_3P DFTB/ChIMES model development, we have chosen to optimize to these quantities, only.

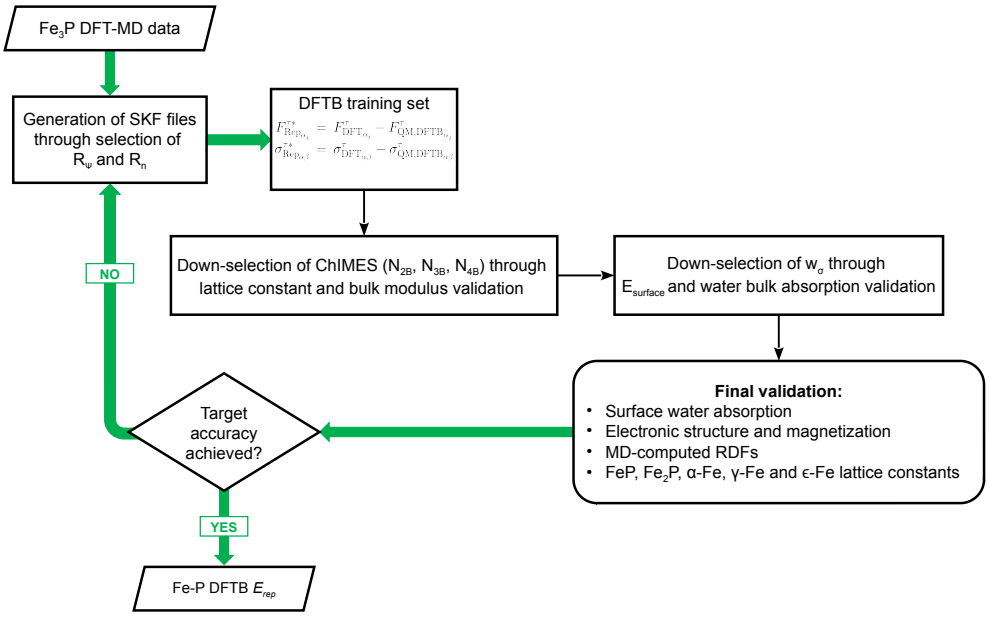


Figure 1: Workflow for DFTB/ChIMES optimization.

We then optimized the DFTB+ confining radii and ChIMES hyperparameters through a workflow similar to our previous efforts[67] (Figure 1). For ChIMES E_{Rep} parametrization, we set the minimum and cutoff radii to consider the first coordination shell of the configurations in our training set: $1.6 \leq r_{FeFe} \leq 3.5$Å and $1.5 \leq r_{\mathrm{FeP}} \leq 2.9$. We use Morse-like scaling constants of $\lambda_{\mathrm{FeFe}} = 2.6$ and $\lambda_{\mathrm{FeP}} = 2.25$, which correspond to the average nearest neighbor distances in our training set. We have limited our search to the following sets of $\{n_{2\mathrm{B}}, n_{3\mathrm{B}}, n_{4\mathrm{B}}\}$ ChIMES polynomials orders: $\{8, 4, 0\}$, $\{12, 8, 0\}$, $\{12, 8, 4\}$, and $\{16, 12, 0\}$. We have also applied different weights to stress tensor components, with values of $w_\sigma = 1, 5, 10, 25, 50$, in order to improve agreement with the validation data.

Results and discussion

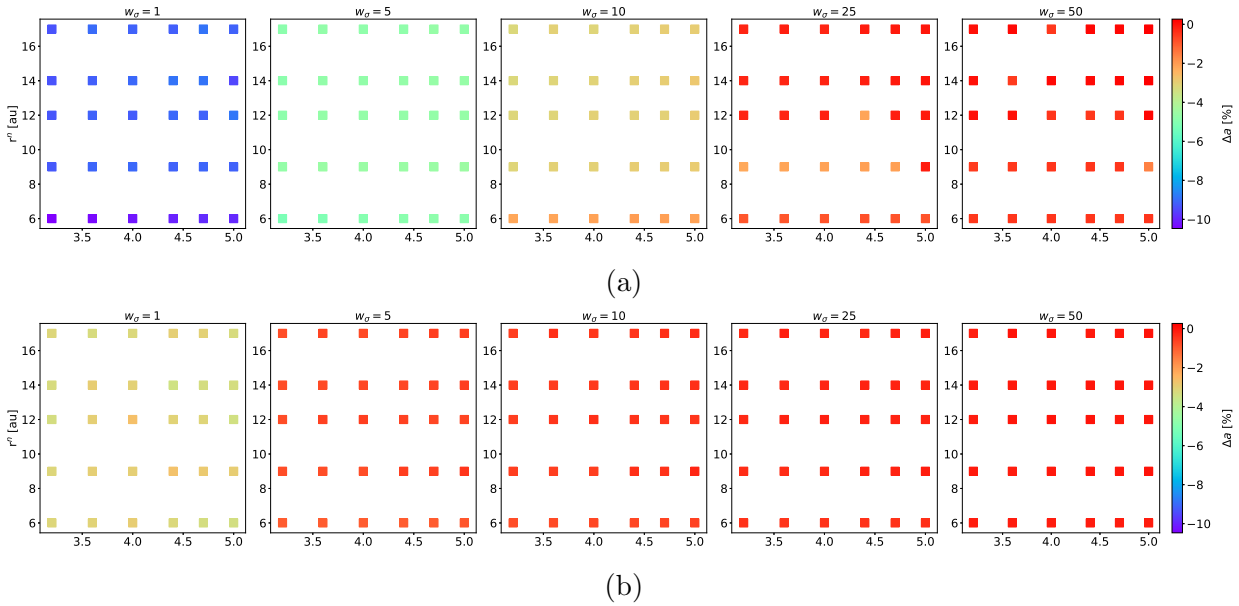
R_Ψ and R_n Optimization

We have computed a grid of DFTB Hamiltonian and overlap matrix tables that span values of $3.2 \leq R_\Psi \leq 5.0$ au and $6.0 \leq R_n \leq 17.0$ au, for a total of 30 different combinations of R_Ψ and R_n. These parameterizations were then passed through our workflow in order to create a unique ChIMES repulsive energy for each set of $\{R_\Psi, R_n\}$. We find that in general the root-mean-squared errors (RMSE) in the forces and stresses show some correlation with the compression radii. Smaller R_Ψ and R_n are associated with smaller force and stress errors. For example, the $\{8, 4, 0\}$ model set yielded Fe force errors on the order of 8.1 kcal/mol.Å, P force errors 5.4 kcal/mol.Å, and stress tensor errors of 4.5 GPa for $\{R_\Psi = 3.2\,\text{au}, R_n = 6.0\,\text{au}\}$, in contrast to values of 16.6 kcal/mol.Å, 13.4 kcal/mol.Å, and 4.14 GPa for $\{R_\Psi = 5.0\,\text{au}, R_n = 17.0\,\text{au}\}$ and $w_\sigma = 1$. The force RMSE exhibits relatively small changes with increasing stress tensor weight, though setting $w_\sigma = 50$ decreases the stress errors to < 1 GPa. Increasing the polynomial order of the ChIMES model has the overall effect of lowering the RMSE. For example, the $\{16, 12, 0\}$ model set yields Fe force errors of approximately 7.8 kcal/mol.Å, P forces of 5.1 kcal/mol.Å, and stress tensor errors of 2.7 GPa for $\{R_\Psi = 3.2\,\text{au}, R_n = 6.0\,\text{au}\}$ and $w_\sigma = 1$. (Please see Supporting Information for additional data.).

The resulting DFTB/ChIMES models were then validated against a number of properties previously computed results from DFT calculations[15] that were not a part of our training set. This includes the Fe_3P lattice constants, bulk modulus, surface energies for the (001), (100), and (110) crystal facets, and water bulk absorption and surface adsorption energies. All DFTB/ChIMES validation calculations were performed using the same k-point meshes as the DFT results, namely, bulk systems were run using a $6 \times 6 \times 6$ **k**-point Monkhorst-Pack grid, while a $1 \times 7 \times 7$ mesh was used for the surface calculations, where the smaller mesh value is related to the nonperiodic surface direction. We again used SCC[27] and charge convergence criteria of 2.72×10^{-5} eV (10^{-6} au) and a force convergence tolerance of 0.01

eV/Å for each atom in each direction.

We now report on the results from our validation calculations. In general, we find that the inclusion of 4B terms in the $\{12, 8, 4\}$ parameterization yields only very change small in accuracy relative to the $\{12, 8, 0\}$ set. Hence, for the sake of brevity, we limit our discussion to the $\{8, 4, 0\}$ and $\{16, 12, 0\}$ parameterizations, which are the lowest and highest limits of ChIMES complexity in our study (the complete set of results can be found in the Supporting Information).



(a)

(b)

Figure 2: Percentage deviation from DFT values of the schreibersite a lattice constant a for (a) {8,4,0}, and (b) {16,12,0} model sets, as a function of the stress tensor weights (reported above the plots). Results for the c lattice constant are included in the Supporting Information.

Lattice constants. Our results for the relative difference between the $\{8, 4, 0\}$ model and DFT for the schreibersite equilibrium lattice constants (Fig. 2 for the a constant, with results for c included in the Supporting Information for the sake of brevity) indicate that using equal weights in our optimizations for the forces and stress tensor components ($w_\mathrm{F} = w_\sigma = 1$) consistently yields an overly compact lattice, where the values of a and c are approximately 8% and 10% smaller. These higher density structures also optimize to antiferromagnetic spin

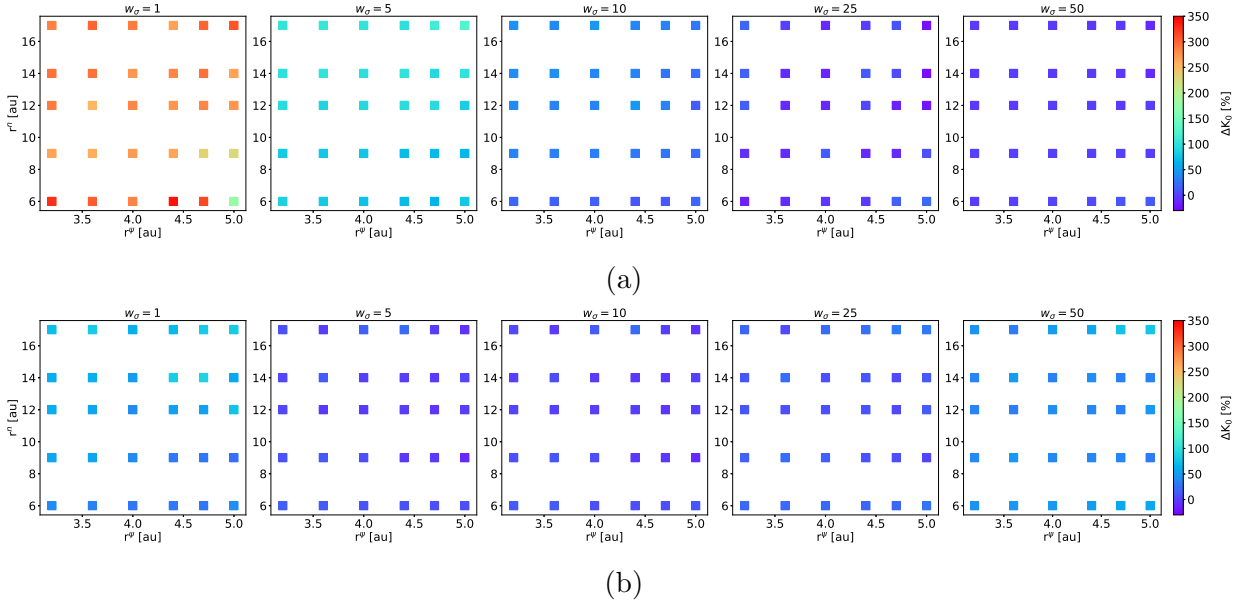
geometries, rather than the correct ferromagnetic result. In addition, we observe a relatively small dependence on values of R_Ψ and R_n. However, the introduction of weights of $w_\sigma > 1$ while keeping w_F at unity improves the results appreciably, where observe the errors improve to -5% for $w_\sigma = 5$ to $\sim 1\%$ for $w_\sigma = 25, 50$.

Increasing the polynomial order to $\{16, 12, 0\}$ uniformly improves the lattice constant results relative to DFT results, regardless of the value of w_σ. Here, the average error in a and c is $\sim 3.5\%$ for $w_\sigma = 1$, which decreases to $< 0.5\%$ for $w_\sigma = 25, 50$. We note that we observe only small differences between results for $\{12, 8, 0\}$ and $\{12, 8, 4\}$, where the lattice constants both agree within 1% for values of $w_\sigma = 10$ and 25 (see Supporting Information for detailed data).

Bulk modulus. We now investigate the influence the confining radii on the schreibersite bulk modulus in order to test the validity of our models for high-pressure material property prediction. In this case, we have computed the zero temperature compression curves by isotropically compressing the system pressure to 100 GPa with each DFTB/ChIMES model. We then fit our data with a Rose-Vinet equation of state equation of state, where the bulk modulus K_0 is one of the fitting parameters.[69] For reference, we previously obtained a DFT bulk modulus of $K_\mathrm{DFT} = 173.2$ GPa,[15] compared to two available experimental results of 172[70] and 157.9[71] GPa.

Fig. 3 reports the results in terms of relative difference ΔK_0 with respect to DFT. We find that all results for $w_\sigma = 1$ yield bulk moduli that are up to 400% too high for the $\{8, 4, 0\}$ set, and over 100% too high for the $\{16, 12, 0\}$ set. This is due to the excessively compact antiferromagnetic structure present in all of these DFTB/ChIMES models. We observe that applying values of w_σ greater than one yields improved results. The $\{8, 4, 0\}$ model set yields errors of approximately 100% for $w_\sigma = 5$, and errors approaching zero for $w_\sigma = 25$ and greater. Our results for $\{16, 12, 0\}$ initially converge more rapidly to low error results, where we observe a number of DFTB/ChIMES optimizations that yield errors of

close to zero for $w_\sigma = 5$ and 10. However, the relative values of ΔK_0 increase with higher values of w_σ. This is likely because the large stress tensor weights combined with the higher complexity of the model yield higher errors in the forces overall. In general, we find that the $\{16, 12, 0\}$ model set generally yields the highest level of accuracy for our validation studies. Hence, we choose to focus on this model set for the remainder of our confining radii study.



(a)

(b)

Figure 3: Percentage deviation from DFT bulk modulus of schreibersite for (a) the $\{8, 4, 0\}$ set and (b) $\{16, 12, 0\}$ set a function of stress tensor weights (reported above the plots).

Surface Energies. Next, we have investigated the surface formation energy for the (001),(100), and (110) low index surface facets. The surface formation energy is defined as

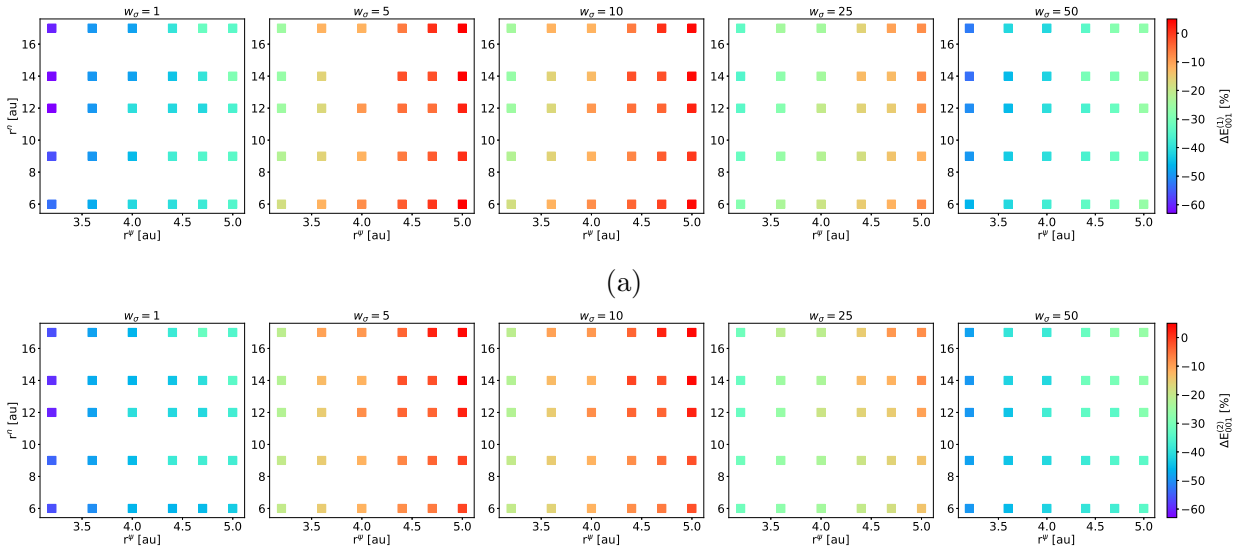
$$E_{\text{surface}} = \frac{E_{\text{slab}} - n\left(\frac{E_{\text{bulk}}}{n_{\text{bulk}}}\right)}{2A}. \tag{7}$$

Here, E_{slab} is the total electronic energy of the slab, $(E_{\text{bulk}}/n_{\text{bulk}})$ is the energy per atom of the bulk cell, n is the number of atoms contained in the slab, A is the surface area of a given facet, and the factor of two accounts for both surfaces present in a periodically replicated system. We use our previously computed surface structures from DFT as our initial configuration for DFTB/ChIMES validation test. In this case, each surface slab is

six atomic layers thick (corresponding to a total of 96 atoms) and with a vacuum region of 20 Å in each direction perpendicular to the slab surface in order to minimize nonphysical interactions between periodic images. Furthermore, due to the heterogeneity of the system, each facet was constructed with different surface terminations (i.e., the number of P or Fe atoms at the surface). By enforcing both symmetry (top and bottom layers are identical) and stoichiometry (i.e., the system maintains a 3:1 iron to phosphorus ratio), we were able to build two different surface terminations for each facet, where either a majority of iron (labeled 'type 1') or phosphorus atoms (labeled 'type 2') is present. The errors for the (001) and (100) surfaces were similar for all sets of confining radii. Hence, for our discussion, we choose to focus on both types of the (001) and (110) surfaces, only.

The complete set of formation energies is reported in the Supporting Information. Our results for $\{16, 12, 0\}$ show that the energies for both type 1 and type 2 (001) surfaces are highly sensitive to the value of w_σ (Fig. 4). When no weights are applied the stress tensor components, the surface energy can be up to 60% larger than DFT (first panel on the left of Fig. 4b), with the highest value of $E^{(1)}_{001} = 3.66$ J/m^2 for $R_\Psi = 3.2$ au and $R_n = 12$ au compression radii, compared to the DFT value of 2.24 J/m^2. We find that increasing the w_σ 5 or 10 improves the results significantly, where surface energies with $R_\Psi > 4.4$ au deviate from DFT for both type 1 and 2 surfaces by less than 5%, with values of $E^{(1)}_{001}$ between 2.15 and 2.33 J/m^2, each. Increasing w_σ to larger values yields significantly larger errors, where the $w_\sigma = 25$ set yields surface energies that are 20-30% higher for $R_\Psi \leq 4.4$ au, with values of $E^{(1)}_{001}$ between 2.67 and 3.02 J/m^2. This is again likely due to large errors in the forces for this set, which results in large errors in optimized geometries.

In contrast, the effect of different confining radii on the (110) surface energy was more difficult to characterize (Fig. 5), as our set of DFTB/ChIMES models for the $\{16, 12, 0\}$ set were more likely to be overfit. This resulted in a lack of convergence of a number of our surface optimizations, as indicated by the lack of data points for some combinations of $\{R_\Psi, R_n\}$. However, we still observe a similar trend for both type 1 and type 2 surfaces,



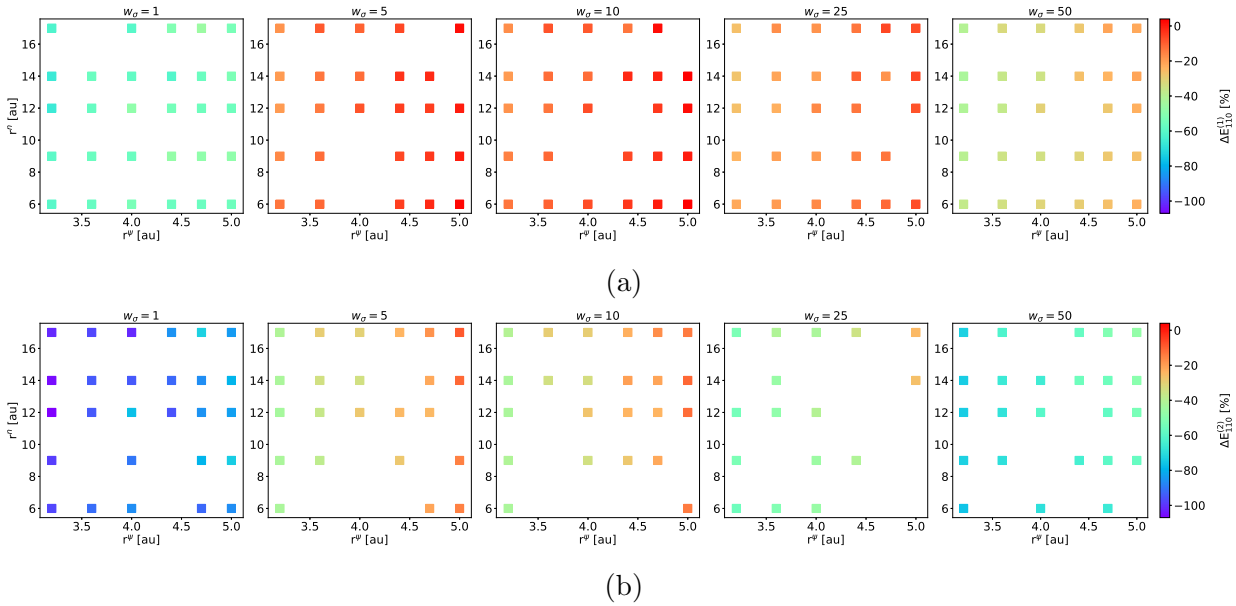
(a)

(b)

Figure 4: Percentage deviation from DFT values of schreibersite surface energy formation for (001) facets of termination type 1 (a) and type 2 (b) obtained for $\{16, 12, 0\}$.

where values of $R_\Psi > 4.0$ and $w_\sigma = 5$ or 10 yielded the smallest deviations from DFT. For example, the DFTB/ChIMES model obtained with $\{R_\Psi = 4.7$ au$, R_n = 9.0$ au$\}$ and $w_\sigma = 10$ yielded an error of $\sim 3\%$ for the type 1 surface and 22% of the type 2 surface, with values 2.61 and 1.88 J/m^2, compared to the DFT results of 2.53 and 1.54 J/m^2, respectively. It is likely the difficulties with convergence for this set would be improved with additional training data, which is the subject of future work.

Bulk interstitial water absorption. We now examine the interaction of schreibersite with water molecules in the bulk. We have used the PyMatGen code[72] to identify two distinct interstitial sites within the lattice (Fig. 6). Specifically, the first type of interstitial site (site 1) consists of a tetrahedral cage composed of iron atoms. The second type of interstitial site (site 2) is formed by a tetrahedron that incorporates a phosphorus atom at one of its vertices. The formation energies for these sites are calculated via the equation $E_{\text{def}} = E_{\text{Fe}_3\text{P}+\text{H}_2\text{O}} - (E_{\text{bulk}} + E_{\text{H}_2\text{O}})$, where $E_{\text{Fe}_3\text{P}+\text{H}_2\text{O}}$ represents the total energy of the relaxed defected structure, E_{bulk} is the total energy of the bulk schreibersite, and $E_{\text{H}_2\text{O}}$



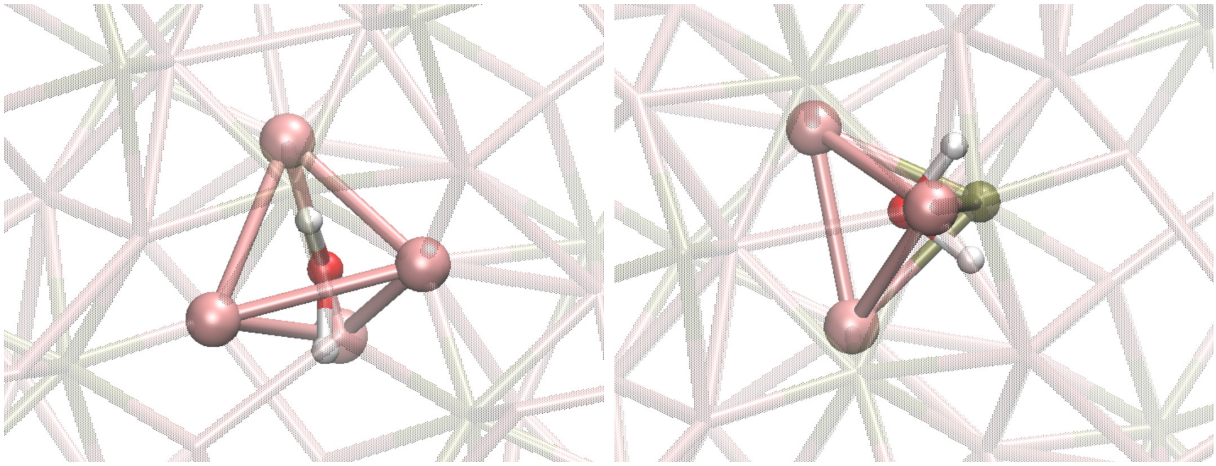
(a)

(b)

Figure 5: Percentage deviation from DFT values of schreibersite surface energy formation for (110) facets of termination type 1 (a) and type 2 (b) obtained for $\{16, 12, 0\}$ as a function of stress weights.

denotes the energy of an isolated water molecule in vacuum. In this case, both DFT and all DFTB/ChIMES calculations were initiated from the same initial configurations, which resulted in dissociation of the water molecule upon optimization. For reference, the values computed via DFT are $E_{\text{site1}} = -3.56$ eV and $E_{\text{site2}} = -3.49$ eV, respectively.

Our sweep tests of confining radii for the $\{16, 12, 0\}/w_\sigma = 5, 10$ model sets indicate that in general these dissociative absorption energies are more accurately computed when considering longer confining radii (Fig. 7). For example, errors relative to DFT are minimized for the $w_\sigma = 5$ set when $R_\Psi = 5.0$ au and $R_n = 17.0$ au, with values of -3.49 and -3.88 eV, respectively for site 1 and site 2. However, we observe that smaller confining radii such as $R_\Psi = 4.7$ au and $R_n = 9.0$ au yield relatively small errors as well, with values of -3.64 and -3.22 eV for site 1 and site 2, respectively, when $w_\sigma = 5$. It is worth noting that, although the $w_\sigma = 10$ in general also exhibits high accuracy, it tends to be overfit when $R_\Psi = 5.0$ (indicated by the missing points in Fig. 7). The increased error with higher stress tensor weighting is commensurate with the higher error in predicted forces, discussed above. The errors are also generally highest for the $\{R_\Psi = 3.2$ au, $R_n = 6.0$ au$\}$ models due to over



(a) (b)

Figure 6: (Bulk interstitial sites for water absorption in schreibersite for (a) site 1 with only Fe atoms, (b) site 2 with one P atom as a vertex of the tetrahedral cage.

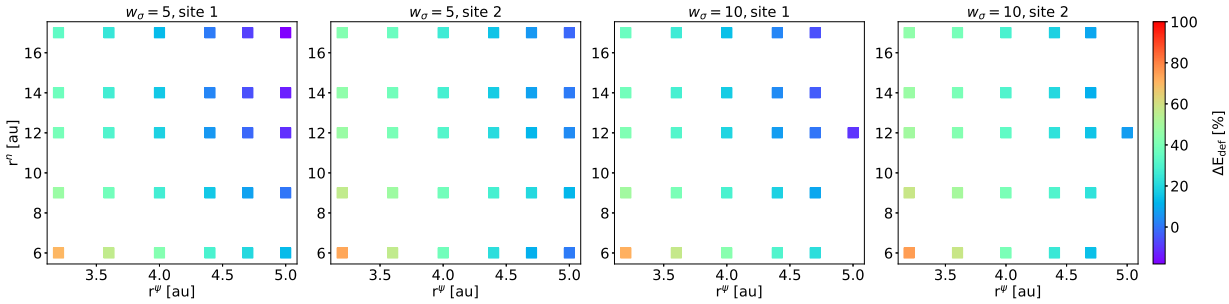


Figure 7: Percentage deviation from DFT values of schreibersite bulk water absorption for $\{16, 12, 0\}$ obtained for both interstitial sites and for $w = 5$ and $w = 10$.

compression of the wavefunction and density.

Optimal Hyperparameters. Our findings thus far provide a broad picture of how the ChIMES hyperparameters and DFTB configuration influence the precision and accuracy of our model. Given our results, we have chosen $\{R_{\Psi}^{\mathrm{Fe}} = 4.7\,\mathrm{au}, R_{n}^{\mathrm{Fe}} = 9.0\,\mathrm{au}\}$, $\{n_{2\mathrm{B}} = 16, n_{3\mathrm{B}} = 12, n_{4\mathrm{B}} = 0\}$ as our optimal set of DFTB/ChIMES hyperparameters, with $w_{\sigma} = 5$ for the weights assigned to the stress tensor components (final results shown in Table 1). This model yielded the highest overall degree of accuracy while avoiding overfitting for our validation calcuations. This model is able to predict the correct energetic ordering of absorption sites ($E_{\mathrm{site2}} > E_{\mathrm{site1}}$) with a somewhat large ΔE of 0.41 eV relative to our DFT calculations, though this value likely changes with the choice of functional and basis set.

Table 1

		DFTB/ ChIMES	DFT
Lattice constant [Å]	a	9.008	9.045
	c	4.366	4.380
K0 [GPa]		170.3	173.2
(001) [J/m^2]	type 1	2.32	2.24
	type 2	2.10	2.00
(100) [J/m^2]	type 1	2.20	2.05
	type 2	2.05	1.95
(110) [J/m^2]	type 1	2.61	2.53
	type 2	1.88	1.54
E_{def} [eV]	Site 1	-3.64	-3.78
	Site 2	-3.22	-3.49

Water Surface Absorption Energies

We now test our optimal DFTB/ChIMES model for water absorption on the most stable facet, i.e., the (110) type 2 (phosphorus terminated) surface. Similarly to interstitial defects, we defined the adsorption energies as $E_{\mathrm{ads}} = E_{\mathrm{slab+H_2O}} - (E_{\mathrm{slab}} + E_{\mathrm{H_2O}})$ where $E_{\mathrm{slab+H_2O}}$ is the energy of the Fe_3P slab with the adsorbed molecule in its final state, E_{slab} is the

energy of the pristine slab, and again E_{H_2O} is the energy of an isolated water molecule in vacuum. In our previous work,[15] we identified three different adsorption sites, i.e., 'ontop', 'bridge' and 'hollow' (Fig. 8). Each type of site can involve different combinations of Fe or P atoms, resulting in two different site types for each classification. We thus validate our model against results for ontop sites on an iron or on a phosphorous atom, bridge sites among two iron atoms or one iron and one phosphorous, and two different hollow sites, with or without phosphorus, depending on the packing of the surface atoms. In addition, the computed adsorption energies for each site depend on the initial orientation of the water molecule, i.e., whether one of the hydrogen atoms from the water molecule is pointing up ('u') into the vacuum, down into the slab ('d'), or the H_2O molecular plane is parallel to the schreibersite surface ('p').[15,17] Hence, each adsorption site is examined for its $u, p,$ and d dependence as well, for a total of 18 adsorption energies (Table 2).

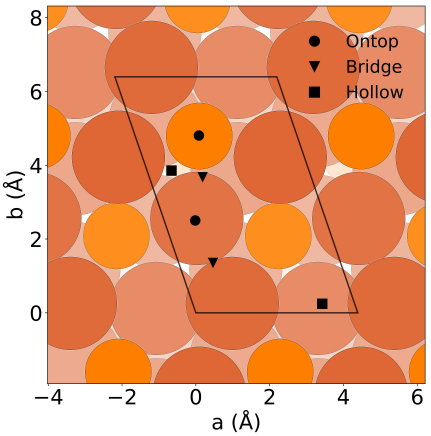


Figure 8: Different adsorption sites on (110) surface. In this image, the big brown circles represent iron atoms, while the smaller orange circles stand for the phosphorous atoms. Graphic taken from Ref. 15.

We observe that the strength of the adsorption energies involving iron atoms are generally overestimated by DFTB/ChIMES relative to DFT, yielding values that are ∼0.22-0.25 eV too negative for the Fe-Fe bridge site, ∼0.38-0.40 eV too negative for the Fe-Fe-P hollow site, and 0.43-0.57 eV too negative for the Fe ontop site. One exception is the Fe-Fe-Fe p

Table 2: Adsorption energies for the different sites shown in Fig. 8, and for the initial molecular configurations investigated here. We note that 'p' stands for planar, 'u' stands for upward and 'd' for downward. Values for $u, p,$ and d that are nearly isoenergetic indicate optimization to a similar final water molecule geometry.

			DFTB/ ChIMES	DFT
Bridge	Fe - Fe	(p)	-0.649	-0.429
		(u)	-0.694	-0.444
		(d)	-0.693	-0.445
	Fe - P	(p)	-0.152	-0.211
		(u)	-0.156	-0.168
		(d)	-0.156	-0.184
Hollow	Fe - Fe - P	(p)	-0.710	-0.331
		(u)	-0.710	-0.338
		(d)	-0.728	-0.334
	Fe - Fe - Fe	(p)	-0.188	-0.195
		(u)	-0.729	-0.484
		(d)	-0.725	-0.484
Ontop	Fe	(p)	-0.799	-0.227
		(u)	-0.720	-0.269
		(d)	-0.702	-0.268
	P	(p)	-0.145	-0.173
		(u)	-0.131	-0.162
		(d)	-0.125	-0.164

hollow site, where our model agrees with DFT within 0.07 eV.

In contrast, the adsorption energies from DFTB/ChIMES predominantly involving phosphorus atoms tend to be more accurate, where the Fe-P bridge site results differ from DFT by less than 0.05 eV and the P ontop site by less than 0.04 eV. These discrepancies in the Fe results are likely due in part to the fact that the Fe-O and Fe-H pairwise repulsive energies were determined by matching to the trans3d parameter set, which was optimized for bond length and dissociation energies, whereas chemical surface interactions with transition metals can require the inclusion of many-body interactions.[67]

Magentization as a Function of Pressure

Given our objective of developing a model suitable for high-pressure impact studies, we consider the variation of magnetization in response to external pressure. Here, we have computed the relative magnetization in scheibersite as it is compressed at zero Kelvin from zero to 100 GPa (Fig. 9). Results from DFTB/ChIMES and DFT are very similar throughout the entire pressure range, with a gradual decrease up to 30 GPa, where we observe a rapid drop that corresponds to a potential first-order phase transition.[70,71,73] This diminishing trend culminates in a non-magnetic state at $P > 50$ GPa in both DFT and DFTB/ChIMES, where the iron atoms have largely lost their magnetic moments, as previously noted by high-pressure experiments.[73]

Molecular Dynamics

In order to test the dynamics of our model under ambient and extreme conditions, we have performed molecular dynamics (MD) calculations using DFTB/ChIMES at pressures and temperatures from ambient up to 50 GPa and 1500 K. Both DFTB/ChIMES and DFT NVT simulations were initiated using an optimized 32-atom schreibersite bulk simulation cell, run for 5 ps and using a timestep of 1 fs. Here, we have used a $2 \times 2 \times 2$ **k**-point Monkhorst-Pack grid for both sets of simulations.

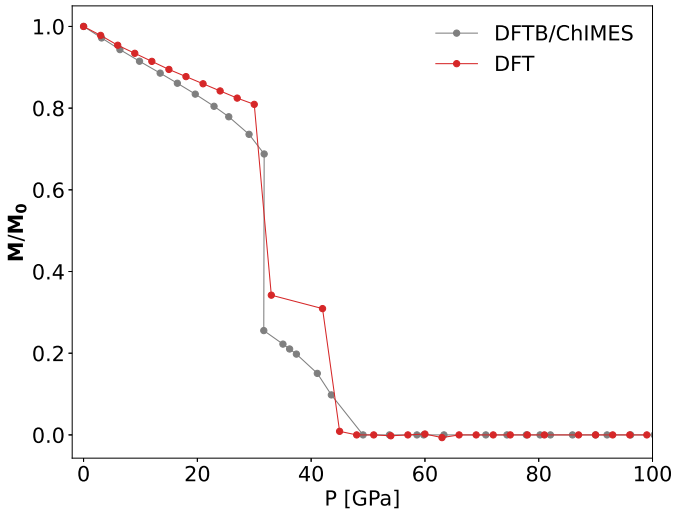
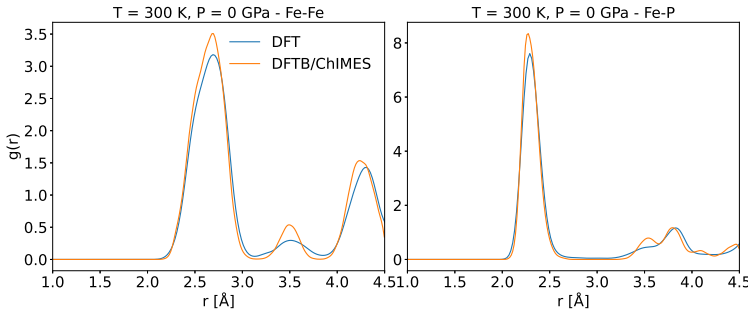
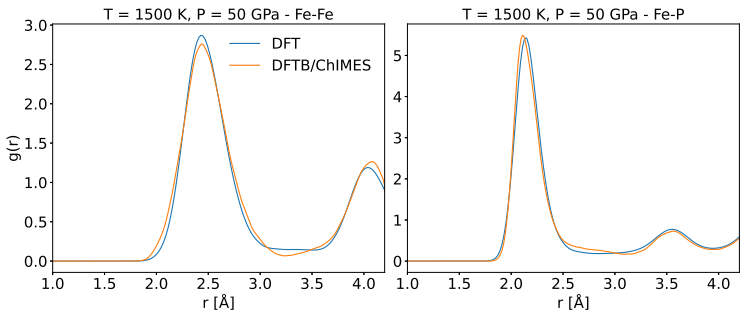


Figure 9: Normalized magnetization of Fe_3P as a function of external pressure.



(a)



(b)

Figure 10: Radial distribution functions obtained for Fe-Fe and Fe-P interatomic distances in schreibersite, as a function of temperature and pressure. Top row (a) shows results for $T = 300$K and $P = 0$GPa, while other rows display respectively $T = 1500$K and $P = 10, 30, 50$GPa.

Our results for the Fe-Fe and Fe-P RDFs indicate overall very good agreement between the DFTB/ChIMES and DFT datasets (Fig. 10), where we have excluded the P-P RDFs for the sake of brevity. Simulations under ambient conditions (Fig. 10a) indicate that DFTB/ChIMES can effectively capture the first coordination shells, where we observe a close match with DFT for the first peak in both the Fe-Fe and Fe-P RDFs. DFTB/ChIMES reveals a slightly less structured atomic distribution for the second and third coordination shells, though the differences are relatively small. At elevated conditions (10b), we observe strong agreement with DFT for all solvation shells captured within our simulations. Notably, configurations at $P = 50$ GPa were absent from our dataset, indicative of the ability of our model to represent state points beyond our training regime accurately.

Transferability of the model

Finally, we examine the transferability of our DFTB/ChIMES model through lattice optimizations of materials other than Fe_3P schreibersite. All optimizations were performed using the material unit cells and the DFTB+ convergence criteria discussed above. We first performed calculations on two other iron phosphide structures, FeP and Fe_2P. Iron monophosphide (FeP) belongs to a family of transition-metal monopnictides, crystallizing in the MnP structure type (orthorhombic, space group Pnma). Recently, FeP has garnered attention due to its distinctive magnetic and superconducting properties, including the formation of a helical magnetic structure.[74] Di-iron phosphide (Fe_2), has been investigated as an anode material in sodium-ion batteries.[75] Fe_2P crystallizes in a hexagonal crystal structure that can be described as layers of Fe atoms in a hexagonal close-packed arrangement with P atoms occupying half of the octahedral interstices.

Our optimization results for these two iron phosphides are summarized in Tab. 3. The FeP b and c lattice constants compare very well with DFT and experiment, with errors of $\sim$1% and 4%, respectively. The a lattice constant differs by roughly 10%, though DFTB/ChIMES is able to retain the orthorhombic primitive cell. The Fe_2P optimization yielded somewhat

Table 3: Comparison between DFTB/ChIMES, DFT, and experiments for the FeP and Fe_2P lattice constants.

	FeP			Fe_2P		
	DFTB/ ChIMES	DFT	expt[74]	DFTB/ ChIMES	DFT	expt[76]
a [Å]	3.340	3.034	3.089	5.828	5.802	5.876
b [Å]	5.356	5.133	5.174			
c [Å]	5.745	5.694	5.772	3.340	3.408	3.449

closer comparison with DFT and experiment overall, where both a and c lattice constants agree within ∼1-2%.

We have also performed transferability studies on three allotropic forms of iron metal, including α-Fe, γ-Fe, and ϵ-Fe. Ferrite, or α-iron, is the ground state and corresponds to a ferromagnetic body-centered cubic (bcc) structure that remains stable up to 1185 K. Beyond this temperature but at low pressure, α-Fe transitions the face-centered cubic (fcc) γ-Fe structure, also named austenite. Under pressures exceeding approximately 10-13 GPa and temperatures up to around 700 K, α-Fe adopts the hexagonal close-packed (hcp) ϵ-iron structure, also called hexaferrum.[77] The results of our calculations indicate strong agreement for the lattice constants of all three phases, with agreement of $\lesssim$ 1% α-Fe and γ-Fe. The agreement for ϵ-iron is also reasonable, where DFTB/ChIMES differs by ∼1% or the a lattice constant, and by ∼4% for the c lattice constant (Table 4). However, we observe some disagreement with DFT for the energetic ordering of the allotropes. In this case, DFT yields an energetic ordering of α-Fe followed by ϵ-Fe and γ-Fe, DFTB/ChIMES yields ϵ-Fe as the most stable phase. However, ϵ-Fe is nearly isoenergetic with α-Fe for both sets of calculations, and the energetic difference between α-Fe and γ-Fe is similar, with a value of 0.13 eV from DFTB/ChIMES and 0.09 eV from DFT. Regardless, our results indicate the relatively strong transferability of DFTB/ChIMES to Fe/P-containing materials, given that none of these materials were present in our training set.

Table 4: Comparison of the DFTB/ChIMES with DFT and experimental values of the different iron allotropes' lattice constants and energy per atom. In particular, the energetic ordering is given with respect to the most stable structures, which has a 0.0 eV value in the table.

	$\alpha-$Fe			$\gamma-$Fe			$\epsilon-$Fe		
	DFTB/ ChIMES	DFT	expt[78]	DFTB/ ChIMES	DFT	expt[79]	DFTB/ ChIMES	DFT	expt[80]
a [Å]	2.827	2.831	2.856	3.647	3.630	3.6519	2.464	2.484	2.461
c [Å]							3.715	3.898	3.952
ΔE [eV]	0.022	0.000		0.157	0.093		0.000	0.072	

Conclusions

In this work, we have determined an optimal DFTB/ChIMES model in order to conduct quantum simulations of chemical reactivity due to the meteoritic mineral schreibersite. We have used our workflow to optimize the DFTB semi-empirical Hamiltonian and overlap matrices, while simultaneously determining a many-body ChIMES repulsive energy. The advantage to our workflow stems from the rapidity with which the linear ChIMES parameters can be determined for each set of quantum interactions, as well as the relatively small data requirements for our repulsive energy optimizations. Our DFTB/ChIMES workflow can be performed with a semiautomated approach, allowing for a rapid downselection of DFTB hyperparameters for the required application.

Our DFTB/ChIMES model yields accurate results for the validation of our study, indicative of a robust model for simulations of Fe_3P degradation due to planetary impact. This included a wide diversity of data, such as the scheribersite lattice constant, bulk modulus, and low-index surface energies. We also compute relatively accurate results for bulk water absorption and water surface adsorption, though these results could be improved further by including these types of systems in our optimization workflow. Regardless, DFTB/ChIMES can determine a highly accurate representation of the magnetization as a function of increasing pressure, where results compare closely with DFT in terms of a schreibersite phase transition at approximately 30 GPa and the absence of magnetization above 50 GPa. Fur-

thermore, our model is shown to be robust for molecular dynamics simulations under extreme pressures and temperatures. Finally. we show that our model shows transferability to different materials, including iron monophosphide, di-iron phosphide, and three allotropes of iron metal. Future work can involve further refinement of our Fe/P DFTB model, by including training data from additional materials, phases, and thermodynamic state points. Ultimately, our model will help with simulations of prebiotic impact synthesis, where the role of aqueous schreibersite systems in phosphate production is difficult to discern, and there is subsequently a need for quantum simulations for predictions of chemical and physical properties to help elucidate future experiments.

Acknowledgement

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The authors gratefully thank the Exobiology Program Element NNH20ZDA001N-EXO for support (proposal #20-EXO20-0149).

Supporting Information Available

References

(1) Pasek, M. A.; Gull, M.; Herschy, B. Phosphorylation on the early earth. Chemical Geology **2017**, 475, 149 – 170.

(2) Gulick, A. Phosphorus as a factor in the origin of life. American Scientist **1955**, 43, 479–489.

(3) Pasek, M. A.; Kee, T. P. Origins of Life: The Primal Self-Organization; Springer Berlin Heidelberg, 2011; pp 57–84.

(4) Schwartz, A. W. Phosphorus in prebiotic chemistry. Philos. Trans. R. Soc. Lond. B Biol. Sci. **2006**, 361, 1743–1749.

(5) Pasek, M. A. Schreibersite on the early Earth: Scenarios for prebiotic phosphorylation. Geoscience Frontiers **2017**, 8, 329–335.

(6) Pasek, M. A.; Lauretta, D. A. Aqueous Corrosion of Phosphide Minerals From Iron Meteorites: A Highly Reactive Source of Prebiotic Phosphorus on the Surface of the Early Earth. Astrobiology **2005**, 5, 515–535.

(7) Pasek, M.; Kee, T.; Bryant, D.; Pavlov, A.; Lunine, J. Production of Potentially Prebiotic Condensed Phosphates by Phosphorus Redox Chemistry. Angewandte Chemie International Edition **2008**, 47, 7918–7920.

(8) Gull, M. Prebiotic Phosphorylation Reactions on the Early Earth. Challenges **2014**, 5, 193–212.

(9) Gull, M.; Mojica, M. A.; Fernández, F. M.; Gaul, D. A.; Orlando, T. M.; Liotta, C. L.; Pasek, M. A. Nucleoside phosphorylation by the mineral schreibersite. Scientific Reports **2015**, 4, 17198.

(10) Cruz, N. L. L.; Qasim, D.; Abbott-Lyon, H.; Pirim, C.; McKee, A. D.; Orlando, T.; Gull, M.; Lindsaya, D.; Pasek, M. A. The evolution of the surface of the mineral schreibersite in prebiotic chemistry. Phys. Chem. Chem. Phys. **2016**, 18, 20160–20167.

(11) Miller, S. L. A production of amino acids under possible primitive Earth conditions. Science **1953**, 117, 528.

(12) Manaa, M. R.; Reed, E. J.; Fried, L. E.; Goldman, N. Nitrogen-rich heterocycles as reactivity retardants in shocked insensitive explosives. J. Am. Chem. Soc. **2009**, 131, 5493–5487.

(13) Goldman, N.; Reed, E. J.; Fried, L. E.; Kuo, I.-F. W.; Maiti, A. Synthesis of glycine-containing complexes in impacts of comets on early Earth. Nat. Chem. **2010**, 2, 949–954.

(14) Goldman, N.; Tamblyn, I. Prebiotic chemistry within a simple impacting icy mixture. J. Phys. Chem. A **2013**, 117, 5124 – 5131.

(15) Dettori, R.; Goldman, N. First-Principles Surface Characterization and Water Adsorption of Fe3P Schreibersite. ACS Earth and Space Chemistry **2022**, 6, 512–520.

(16) Pantaleone, S.; Corno, M.; Rimola, A.; Balucani, N.; Ugliengo, P. Ab Initio Computational Study on Fe_2NiP Schreibersite: Bulk and Surface Characterization. ACS Earth and Space Chemistry **2021**, 5, 1741–1751.

(17) Pantaleone, S.; Corno, M.; Rimola, A.; Balucani, N.; Ugliengo, P. Water Interaction with Fe2NiP Schreibersite (110) Surface: a Quantum Mechanical Atomistic Perspective. The Journal of Physical Chemistry C **2022**, 126, 2243–2252.

(18) Cannella, C.; Goldman, N. Carbyne fiber synthesis within evaporating metallic liquid carbon. J. Phys. Chem. C **2015**, 119, 21605–21611.

(19) Lindsey, R. K.; Bastea, S.; Goldman, N.; Fried, L. E. Investigating 3,4-bis(3-nitrofurazan-4-yl)furoxan detonation with a rapidly tuned density functional tight binding model. J. Chem. Phys. **2021**, 154, 164115.

(20) Armstrong, M. R.; Lindsey, R. K.; Goldman, N.; Nielsen, M. H.; Stavrou, E.; Fried, L. E.; Zaug, J. M.; Bastea, S. Ultrafast shock synthesis of nanocarbon from a liquid precursor. Nature Communications **2020**, 11.

(21) Lindsey, R. K.; Goldman, N.; Fried, L. E.; Bastea, S. Chemistry-mediated Ostwald ripening in carbonrich C/O systems at extreme conditions. Nature Communications **2022**, 13, 1424.

(22) Nomora, K.; Kalia, R. K.; Nakano, A.; Vashista, P.; van Duin, A. C. T.; Goddard, W. A. Dynamic transition in the shock structure of an energetic material. Phys. Rev. Lett. **2007**, 99, 148303.

(23) Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. Phys. Rev. B **2019**, 99, 014104.

(24) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. Nat. Commun. **2019**, 10, 1–8.

(25) Kovács, D. P.; van der Oord, C.; Kucera, J.; Allen, A. E. A.; Cole, D. J.; Ortner, C.; Csányi, G. Linear Atomic Cluster Expansion Force Fields for Organic Molecules: Beyond RMSE. J. Chem. Theory. Comput. **2021**, 17, 7696–7711.

(26) Pham, C. H.; Lindsey, R. K.; Fried, L. E.; Goldman, N. Calculation of the detonation state of HN_3 with quantum accuracy. J. Chem. Phys. **2021**, 153, 224102.

(27) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. Phys. Rev. B **1998**, 58, 7260–7268.

(28) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). Journal of Chemical Theory and Computation **2011**, 7, 931–948.

(29) Hourahine, B. et al. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. The Journal of Chemical Physics **2020**, 152, 124101.

(30) Goldman, N.; Koziol, L.; Fried, L. E. Using force-matched potentials to improve the accuracy of density functional tight binding for reactive conditions. J. Chem. Theory Comput. **2015**, 11, 4530–4535.

(31) Kandy, A. K. A.; Wadbro, E.; Aradi, B.; Broqvist, P.; Kullgren, J. Curvature Constrained Splines for DFTB Repulsive Potential Parametrization. J. Chem. Theory Comput. **2021**, 1771-1781, 21.

(32) Wengert, S.; Csányi, G.; Reuter, K.; Margraf, J. T. Data-efficient machine learning for molecular crystal structure prediction. Chem. Sci. **2021**, 12, 4536.

(33) Goldman, N.; Fried, L. E.; Lindsey, R. K.; Pham, C. H.; Dettori, R. Enhancing the accuracy of density functional tight binding models through ChIMES many-body interaction potentials. J. Chem. Phys. **2023**, 158, 144112.

(34) Stöhr, M.; Medrano Sandonas, L.; Tkatchenko, A. Accurate Many-Body Repulsive Potentials for Density-Functional Tight Binding from Deep Tensor Neural Networks. J. Phys. Chem. Lett. **2020**, 11, 6835–6843.

(35) Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. Description of Phosphate Hydrolysis Reactions with the Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB) Theory. 1. Parameterization. Journal of Chemical Theory and Computation **2008**, 4, 2067–2084.

(36) Niehaus, T.; Elstner, M.; Frauenheim, T.; Suhai, S. Application of an approximate density-functional method to sulfur containing compounds. Journal of Molecular Structure: THEOCHEM **2001**, 541, 185–194.

(37) Guimarães, L.; Enyashin, A. N.; Frenzel, J.; Heine, T.; Duarte, H. A.; Seifert, G. Imogolite Nanotubes: Stability, Electronic, and Mechanical Properties. ACS Nano **2007**, 1, 362–368.

(38) Luschtinetz, R.; Oliveira, A. F.; Frenzel, J.; Joswig, J.-O.; Seifert, G.; Duarte, H. A. Adsorption of phosphonic and ethylphosphonic acid on aluminum oxide surfaces. Surface Science **2008**, 602, 1347–1359.

(39) Sieck, A.; Frauenheim, T.; Jackson, K. A. Shape transition of medium-sized neutral silicon clusters. physica status solidi (b) **2003**, 240, 537–548.

(40) Köhler, C.; Seifert, G.; Frauenheim, T. Density functional based calculations for Fen (n≤32). Chemical Physics **2005**, 309, 23–31.

(41) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. Journal of Chemical Theory and Computation **2012**, 9, 338–354.

(42) Kubillus, M.; Kubař, T.; Gaus, M.; Řezáč, J.; Elstner, M. Parameterization of the DFTB3 Method for Br, Ca, Cl, F, I, K, and Na in Organic and Biological Systems. Journal of Chemical Theory and Computation **2014**, 11, 332–342.

(43) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. Journal of Chemical Theory and Computation **2014**, 10, 1518–1537.

(44) Zheng, G.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K.; Lundberg, M.; Elstner, M.; Köhler, C.; Frauenheim, T. Parameter Calibration of Transition-Metal Elements for the Spin-Polarized Self-Consistent-Charge Density-Functional Tight-Binding (DFTB) Method: Sc, Ti, Fe, Co, and Ni. Journal of Chemical Theory and Computation **2007**, 3, 1349–1367.

(45) Lindsey, R. K.; Goldman, N.; Fried, L. E.; Bastea, S. Many-body reactive force field development for carbon condensation in C/O systems under extreme conditions. The Journal of Chemical Physics **2020**, 153, 054103.

(46) Lindsey, R. K.; Fried, L. E.; Goldman, N.; Bastea, S. Active learning for robust, high-complexity reactive atomistic simulations. The Journal of Chemical Physics **2020**, 153, 134117.

(47) Goldman, N.; Fried, L. E.; Lindsey, R. K.; Pham, C. H.; Dettori, R. Enhancing the accuracy of density functional tight binding models through ChIMES many-body interaction potentials. The Journal of Chemical Physics **2023**, 158, 144112.

(48) Goldman, N.; Kweon, K. E.; Sadigh, B.; Heo, T.-W.; Lindsey, R. K.; Pham, C. H.; Fried, L. E.; Aradi, B.; Holliday, K.; Jeffries, J. R.; Wood, B. C. Semi-Automated Creation of Density Functional Tight Binding Models through Leveraging Chebyshev Polynomial-Based Force Fields. J. Chem. Theory Comput. **2021**, 17, 4435–4448.

(49) Goldman, N.; Aradi, B.; Lindsey, R. K.; Fried, L. E. J. Chem. Theory. Comput. **2018**, 14, 2652–2660.

(50) Pham, C. H.; Lindsel, R. K.; Fried, L. E.; Goldman, N. High-Accuracy Semiempirical Quantum Models Based on a Minimal Training Set. J. Phys. Chem. Lett. **2022**, 13, 2934–2942.

(51) Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. Phys. Rev. B **1993**, 47, 558–561.

(52) Kresse, G.; Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. Phys. Rev. B **1994**, 49, 14251–14269.

(53) 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.

(54) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. Phys. Rev. B **1999**, 59, 1758–1775.

(55) Blöchl, P. E. Projector augmented-wave method. Phys. Rev. B **1994**, 50, 17953–17979.

(56) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. Phys. Rev. Lett. **1996**, 77, 3865–3868.

(57) Methfessel, M.; Paxton, A. T. High-precision sampling for Brillouin-zone integration in metals. Phys. Rev. B **1989**, 40, 3616–3621.

(58) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. Phys. Rev. B **1976**, 13, 5188–5192.

(59) Bedolla, P. O.; Feldbauer, G.; Wolloch, M.; Eder, S. J.; Dörr, N.; Mohn, P.; Redinger, J.; Vernes, A. Effects of van der Waals Interactions in the Adsorption of Isooctane and Ethanol on Fe(100) Surfaces. The Journal of Physical Chemistry C **2014**, 118, 17608–17615.

(60) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. Physical Review B **1998**, 58, 7260–7268.

(61) Aradi, B.; Hourahine, B.; Frauenheim, T. DFTB+, a sparse matrix-based implementation of the DFTB method. J. Phys. Chem. A **2007**, 111, 5678–5684.

(62) Koskinen, P.; Mäkinen, V. Density-functional tight-binding for beginners. Comp. Mater. Sci. **2009**, 47, 237–253.

(63) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). J. Chem. Theory Comput. **2011**, 7, 931–948.

(64) Vuong, V. Q.; Madridejos, J. M. L.; Aradi, B.; Sumpter, B. G.; Metha, G. F.; Irle, S. Density-functional tight-binding for phosphine-stabilized nanoscale gold clusters. Chemical Science **2020**, 11, 13113–13128.

(65) Dantanarayana, V.; Nematiaram, T.; Vong, D.; Anthony, J. E.; Troisi, A.; Cong, K. N.; Goldman, N.; Faller, R.; Moulé, A. J. Predictive Model of Charge Mobilities in Organic Semiconductor Small Molecules with Force-Matched Potentials. Journal of Chemical Theory and Computation **2020**, 16, 3494–3503.

(66) Chou, C.-P.; Nishimura, Y.; Fan, C.-C.; Mazur, G.; Irle, S.; Witek, H. A. Automatized parameterization of DFTB using particle swarm optimization. J. Chem. Theory Comput. **2016**, 12, 53–64.

(67) Goldman, N.; Kweon, K. E.; Sadigh, B.; Heo, T. W.; Lindsey, R. K.; Pham, C. H.; Fried, L. E.; Aradi, B.; Holliday, K.; Jeffries, J. R.; Wood, B. C. Semi-Automated Creation of Density Functional Tight Binding Models through Leveraging Chebyshev Polynomial-Based Force Fields. Journal of Chemical Theory and Computation **2021**, 17, 4435–4448.

(68) Tersoff, J. Empirical interatomic potential for carbon, with application to amorphous-carbon. Phys. Rev. Lett. **1988**, 61, 2879.

(69) Vinet, P.; Smith, J. R.; Ferrante, J.; Rose, J. H. Temperature effects on the universal equation of state of solids. Physical Review B **1987**, 35, 1945–1953.

(70) High-pressure behavior of Fe_3P and the role of phosphorus in planetary cores. Earth and Planetary Science Letters **2014**, 390, 296–303.

(71) Scott, H. P.; Huggins, S.; Frank, M. R.; Maglio, S. J.; Martin, C. D.; Meng, Y.; Santillán, J.; Williams, Q. Equation of state and high-pressure stability of Fe_3P-schreibersite: Implications for phosphorus storage in planetary cores. Geophysical Research Letters **2007**, 34.

(72) https://pymatgen.org.

(73) Scott, H.; Kiefer, B.; Martin, C. D.; Boateng, N.; Frank, M.; Meng, Y. P-V equation of state for Fe2P and pressure-induced phase transition in Fe3P. High Pressure Research **2008**, 28, 375–384.

(74) Chernyavskii, I. O. et al. Incommensurate magnet iron monophosphide FeP: Crystal growth and characterization. Physical Review Materials **2020**, 4.

(75) Huang, H.; Yu, C.; Yang, J.; Han, X.; Zhao, C.; Li, S.; Liu, Z.; Qiu, J. Ultrasmall diiron phosphide nanodots anchored on graphene sheets with enhanced electrocatalytic activity for hydrogen production via high-efficiency water splitting. Journal of Materials Chemistry A **2016**, 4, 16028–16035.

(76) Dera, P.; Lavina, B.; Borkowski, L. A.; Prakapenka, V. B.; Sutton, S. R.; Rivers, M. L.; Downs, R. T.; Boctor, N. Z.; Prewitt, C. T. High-pressure polymorphism of Fe_2P and its implications for meteorites and Earth's core. Geophysical Research Letters **2008**, 35.

(77) Mathon, O.; Baudelet, F.; Itié, J. P.; Polian, A.; d'Astuto, M.; Chervin, J. C.; Pascarelli, S. Dynamics of the Magnetic and Structural $\alpha - \epsilon$ Phase Transition in Iron. Physical Review Letters **2004**, 93.

(78) Introductory Chemistry for the Environmental Sciences; Cambridge University Press, 1996; pp 148–209.

(79) SEKI, I.; NAGATA, K. Lattice Constant of Iron and Austenite Including Its Supersaturation Phase of Carbon. ISIJ International **2005**, 45, 1789–1794.

(80) Takahashi, T.; Bassett, W. A.; Mao, H.-K. Isothermal compression of the alloys of iron up to 300 kilobars at room temperature: Iron-nickel alloys. Journal of Geophysical Research **1968**, 73, 4717–4725.

TOC Graphic