

Elemental vacancy diffusion database from high-throughput first-principles calculations for FCC and HCP structures

Thomas Angsten, Tam Mayeshiba, Henry Wu, Dane Morgan

Department of Materials Science and Engineering, University of Wisconsin-Madison, 1509 University Avenue, Madison, WI 53706

E-mail: angsten@wisc.edu

Abstract. This work demonstrates how databases of diffusion-related properties can be developed from high-throughput *ab initio* calculations. The formation and migration energies for vacancies of all adequately stable pure elements in both the face-centered cubic (FCC) and hexagonal close packing (HCP) crystal structures were determined using *ab initio* calculations. For HCP migration, both the basal plane and z-direction nearest-neighbor vacancy hops were considered. Energy barriers were successfully calculated for forty-nine elements in the FCC structure and forty-four elements in the HCP structure. These data were plotted against various elemental properties in order to discover significant correlations. The calculated data show smooth and continuous trends when plotted against Mendeleev numbers. The vacancy formation energies were plotted against cohesive energies to produce linear trends with regressed slopes of 0.317 and 0.323 for the FCC and HCP structures respectively. This result shows the expected increase in vacancy formation energy with stronger bonding. The slope of approximately 0.3, being well below that of one predicted by a simple fixed bond strength model, is consistent with a reduction in the vacancy formation energy due to many-body effects and relaxation. Vacancy migration barriers are found to increase nearly linearly with increasing stiffness, consistent with the local expansion required to migrate an atom. A simple semi-empirical expression is created to predict the vacancy migration energy from the lattice constant and bulk modulus for FCC systems, yielding estimates with errors of approximately 30%.

1. Introduction

Many materials phenomena are affected by atomic diffusion properties, and knowledge of these properties is of great technological importance in applications ranging from structural materials to electrochemical devices. Extensive experimental efforts have been exerted over many decades in order to determine the activation energies for vacancy diffusion, Q , in pure elements and alloys. The vacancy diffusion mechanism is the dominant mechanism of self-diffusion in monoatomic close-packed crystals [1]. As attempt frequencies are relatively similar, within a simple Arrhenius model, the quantity Q largely determines self-diffusion properties for elements in close-packed structures. The activation enthalpy for vacancy diffusion can be found by summing together the enthalpy of vacancy formation, H_{vf} , and the vacancy migration enthalpy, H_{vm} . Accurate experimental determination of these two quantities can be quite challenging. The large range of vacancy formation enthalpies reported from the different experimental methods (e.g., quenching, positron annihilation, etc.) indicates that H_{vf} determination is non-trivial. One complication that is commonly encountered in experimentally determining H_{vf} is the isolation of monovacancy contributions from those of other defects, such as divacancies [2]. H_{vm} can be directly measured in quenching experiments. However, this often leads to less accurate results than simply subtracting the vacancy formation enthalpy from Q found in diffusion experiments [1, 3]. Thus, the accuracy of H_{vm} is dependent on methods used to find H_{vf} , making both quantities difficult to determine consistently and accurately. These difficulties partially provide the incentive to use theoretical alternatives to calculate the diffusion properties of materials.

One theoretical approach that has gained popularity in recent decades is the use of *ab initio* Density Functional Theory (DFT) based quantum codes in order to determine materials properties. DFT is an efficient technique used to solve the fundamental quantum-mechanical equations for complex systems of atoms. DFT has had good success in predicting accurate diffusion properties [4-6]. The evolution of fast and robust DFT codes has made it possible to automate computational techniques in which large amounts

of data relevant to a specific materials property can be obtained [7]. High-throughput *ab initio* calculations are being used to study an increasing range of problems, including surface energy determination [8], crystal structure prediction [9], Band structure calculation [10], catalyst design [11], and Li-ion battery cathode design [12]. A number of tools to aid in high-throughput computation are also becoming available, including AFLOW [10], the Materials Project [13], and the Atomic Simulation Environment [14]. However, little work has been done to develop databases of basic elemental or alloy diffusion data using high-throughput *ab initio* methods.

There are a few key advantages to having a large database of diffusion properties formulated from automated *ab initio* calculations. The consistency of the methods used to obtain the results makes for more reliable relative values. Furthermore, using *ab initio* techniques allows one to study systems that are theoretical or do not exist in a stable natural form. Predictions for metastable systems can be useful if these systems are ever synthesized. These predictions are also often necessary when fitting the dependence of diffusion energetics on alloy composition. Finally, having a consistent and consolidated database allows one to establish correlations between diffusion and system properties. These correlations could be used as a materials screening tool that would help facilitate the choice of materials for technological applications. This work will demonstrate the utility of high-throughput calculations for establishing an elemental diffusion database, and this database will be used to explore trends in vacancy formation and migration energetics. While this work is focused on simple pure elements as an initial demonstration, it suggests a path to a more comprehensive effort currently underway that is focused on treating diffusion in alloys with the same high-throughput approach.

2. Computational Methods

2.1. Total Energy Calculations

All calculations were performed using the Vienna Ab Initio Simulation Package (VASP) version 5.2.2, a quantum mechanical code based on plane-wave density functional theory [15-18]. The workflows of the calculations were automated using the MAterials Simulation Toolkit (MAST), a code package developed at the University of Wisconsin-Madison and built on the pymatgen toolset [19]. For all of the studied materials, the Perdew–Burke–Ernzerhof (PBE) exchange–correlation functional with the Generalized Gradient Approximation (GGA) for the exchange–correlation energy was used, and the projector augmented wave method (PAW) was implemented to model the electron-ion interactions within each of these systems [20-23]. All ionic relaxations were performed using a conjugate gradient algorithm. For these, a first-order Methfessel-Paxton smearing method [24] was used with a smearing width of 0.2 eV. For all ground state energy calculations with fixed ions, a tetrahedron method with Blöchl corrections was used with the same smearing width. Spin-polarization was neglected for all systems with the exception of Ni, Co, and Mn in both FCC and HCP. Because Fe is non-magnetic in the FCC structure [25], only HCP Fe was attempted with spin-polarization (although no moment was found). For this subgroup of element-structure combinations, non-spin-polarized and spin-polarized results are both presented.

2.2. Calculation Parameters

The size of the supercell and the density of points in the k-point mesh were held constant for all elements in a given crystal structure. For the FCC systems, a Monkhorst–Pack [26] 4x4x4 k-point mesh and a 3x3x3 supercell of the conventional cell consisting of 108 lattice sites were used. For the HCP systems, a Gamma-centered 9x9x9 k-point mesh and a 3x3x2 supercell of the conventional cell consisting of 36 lattice sites were used. Table 1 summarizes the results of convergence testing for k-point mesh performed on Aluminum in FCC and Magnesium in HCP. Convergence with cell size was tested by using a finite-size scaling approach. Following previous literature we plotted the values of interest (H_{vf} and H_{vm}) as a function of N_{atoms}^{-1} and looked for trends to enable extrapolation to infinite cell size [27-31]. In general

we found very robust fits for a simple linear dependence versus N_{atoms}^{-1} , consistent with the defect interactions being based on strain.

Figure 1 shows the effect of cell size on H_{vf} for Au and Mg in both the FCC and HCP structures. Based on figure 1, the estimated size-effect error for H_{vf} is 20-30 meV for both the FCC and HCP systems. Almost no change in the HCP vacancy formation energy is observed with respect to system size, suggesting that the 3x3x2 supercell is sufficient for HCP calculations. Figure 2 shows the effect of cell size on H_{vm} for Au and Mg in both the FCC and HCP structures. From figure 2, the estimated size-effect error for H_{vm} is again between 20-30 meV for both the FCC and HCP systems. While migration energies were not calculated for the 6x6x4 HCP supercell due to computational limitations, both basal and c-axis migration energies show little dependence with system size and are sufficiently converged. Plane wave cutoff energies were assigned for each element at 1.5 times the maximum energy cutoff in the PAW potential file (denoted ENMAX) for that element. This energy cutoff gives errors in the reference energy of about 5 meV/atom or less. In both the FCC and HCP structures, the same pseudopotential was used for each element.

Table 1. Results of convergence tests performed on FCC aluminum and HCP magnesium.

Element	K-point mesh at constant size ^a	
	H_{vf} (eV)	H_{vm} (eV)
Aluminum (FCC)	0.018	0.021
Magnesium (HCP)	0.009	0.007

^a 108 atom cell for FCC aluminum and 36 atom cell for HCP magnesium. These errors are obtained by comparing Monkhorst-Pack k-point meshes of 4x4x4 to 8x8x8 for aluminum and Gamma-centered k-point meshes of 9x9x9 to 13x13x13 for magnesium.

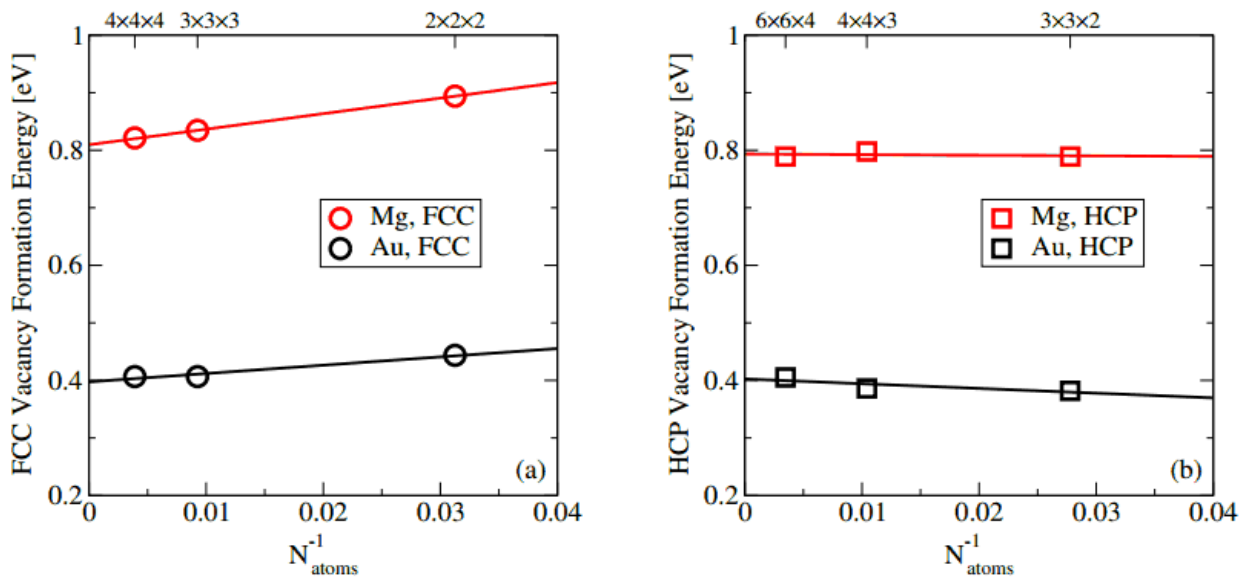


Figure 1. Plots showing the effect of H_{vf} with cell size for two systems in (a) FCC and (b) HCP structures. Based on linear extrapolation to $1/N = 0$ for these test cases, the size-effect error for both the FCC and HCP systems

treated in this paper is around 20-30 meV. Note that, in the case of Mg in FCC and Au in HCP, these are the vacancy formation energies in a metastable host structure.

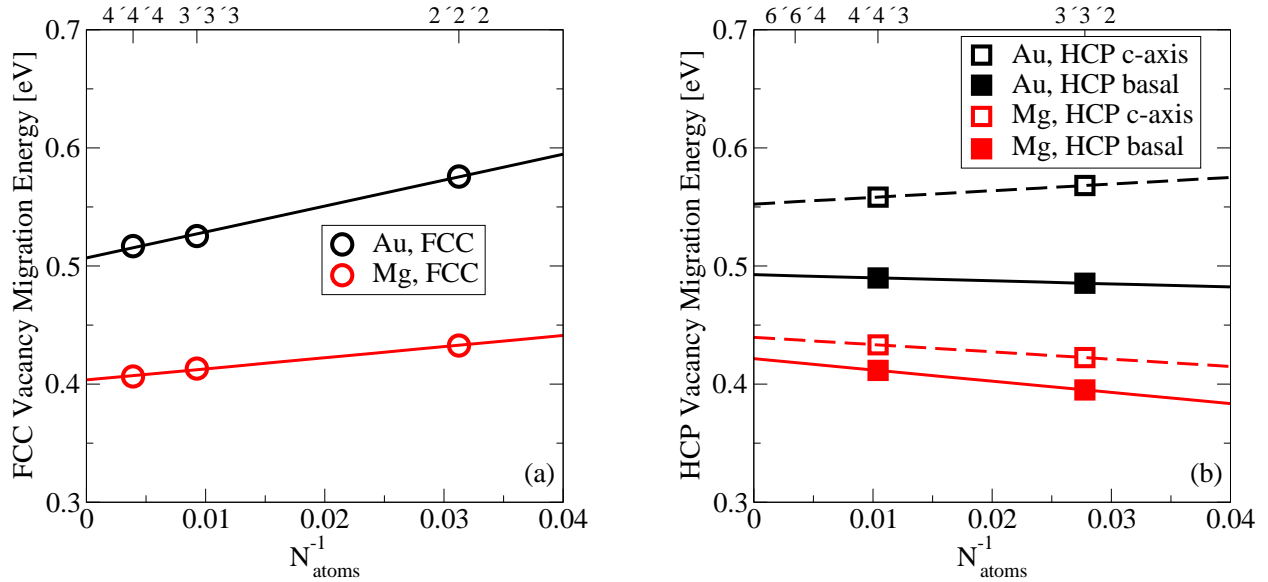


Figure 2. Plots showing the effect of H_{vm} with cell size for two systems in (a) FCC and (b) HCP structures. Based on linear extrapolation to $1/N = 0$ for these test cases, the size-effect error of the migration energy for both the FCC and HCP systems treated in this paper is between 20-30 meV. Note that the migration energies for the 6x6x4 HCP supercell were not calculated.

2.3. Parameters Used in Saddle Point Calculations

The Climbing Image Nudged Elastic Band (CI-NEB) [32] method was used to obtain the vacancy migration enthalpy, H_{vm} , of each system. This procedure involves finding the saddle point structure and the associated minimum energy pathway using a chain of images representing intermediate structures that are linearly interpolated from two endpoint structures. Both endpoint structures consisted of cells containing a monovacancy at one site (107 on-lattice atoms + 1 vacancy or 35 on-lattice atoms + 1 vacancy). In all cases, the vacancies were positioned at nearest-neighbor sites with respect to the migrating atom. For all CI-NEB-based calculations, a $5.0 \text{ eV}/\text{\AA}^2$ spring constant was used to maintain a constant distance between the images. A quasi-Newton algorithm with a force scaling constant of 0.5 was utilized in the relaxation of ions into their respective ground states. Further detail regarding the calculation of migration energies is given in section 2.5.

2.4. Calculation of the Vacancy Formation Enthalpy

The vacancy formation enthalpy, H_{vf} , was determined by comparing the total energy of the cell containing the monovacancy with that of the undefected cell. The undefected cell energy is scaled by a fraction in order to conserve the appropriate number of atoms. This calculation is represented in (1):

$$H_{vf} = H_{N-1} - \left(\frac{N-1}{N}\right)H_N \quad (1)$$

In the above equation, N is the number of lattice sites, H_{N-1} is the enthalpy of the defected cell containing a single vacancy, and H_N is the enthalpy of the undefected cell. The calculation of H_{vf} using VASP involved a series of relaxations of both an undefected and a defected cell. For each element, calculations began with the geometric optimization of the undefected cell, allowing the cell volume and ionic positions

to relax to their minimum energy states. Multiple optimizations were performed until the total energy of the undefected cell was converged to within 1 meV/cell. Once this was achieved, the electronic degrees of freedom of the system were re-optimized in a fixed-ion calculation of the total energy to give H_N in (1). To get H_{N-1} in (1), a new cell was then formed from the relaxed undefected cell by adding a single vacancy of arbitrary placement. A similar series of calculations were performed using the same energy threshold at a fixed volume equal to that of the relaxed undefected cell. From these resulting energies, H_{vf} was calculated. Note that all calculations are performed at approximately zero pressure so the enthalpy and energy (H_{vf} and E_{vf}) are taken to be equal in these calculations. Therefore, the energy is calculated and set equal to the enthalpy.

2.5. Calculation of the Vacancy Migration Enthalpy

Transition state theory [33] was applied to find H_{vm} for each system. H_{vm} was obtained by comparing the total enthalpy of an endpoint containing a monovacancy with the total enthalpy of the system containing an atom at the transition-state position of a migration event. This is quantified in (2):

$$H_{vm} = H_{N-1, TS} - H_{N-1, IS} \quad (2)$$

In the above equation, TS signifies the transition state and IS signifies the initial state of the system. As with formation energies, the vacancy migration enthalpy and energy (H_{vm} and E_{vm}) are taken as equal, as the calculations are performed at approximately zero pressure. For computational efficiency, an exact construction of the second endpoint was made from the first using symmetry in both the FCC and HCP systems. This method involved translating each atom of the first endpoint by a vector equivalent to a nearest-neighbor hop associated with each crystal system.

The first and second endpoint structures were taken and used to linearly interpolate a single transition state image used in the CI-NEB calculation. $H_{N-1, TS}$ in (2) is given by the energy of this intermediate image, while the initial-state enthalpy, $H_{N-1, IS}$, can be obtained from the total energy of either equivalent endpoints. To ensure that the single-image CI-NEB calculations discovered an energy maximum, the migrating atom of the linearly-interpolated middle image was displaced by a fraction of an angstrom in a symmetry-breaking direction. This displacement guaranteed that the climbing image would not be stuck in a local minimum due to the existence of a high-symmetry point. Since all of the atom hops studied in this work are symmetric, this approach will find the maximum of the potential energy surface assuming the migration path has one or two maxima. If the energy surface has three or more maxima then they could be missed by our approach, but no examples of such complex energy surfaces were found in the subset of four element-structure systems checked with additional images.

2.6. Calculation of the Cohesive Energy and Bulk Modulus

As the formation and migration of vacancies can both be considered bond-breaking processes, it is of interest to see how the energy barriers correlate with the cohesive energy, E_C , of each element. Cohesive energies were calculated as the difference in energy per atom between the crystal and the single atom system. The latter reference energy was calculated with a Gamma-centered 1x1x1 k-point mesh in a 12x13x14Å cell at an energy cutoff of 600 eV. In the case of Ni, the asymmetric cell used for the atom calculations could not be converged despite extensive efforts, and a cell with 10x10x10Å dimensions was used. The cohesive energy of Ni therefore runs a risk of errors associated with inadequate symmetry breaking in the ground state of its atom reference state. All isolated atom calculations were performed with spin-polarization.

The bulk modulus B_0 and its pressure derivative B_0' were obtained by fitting energy–volume data of the undefected cell to the third-order Birch–Murnaghan equation of state (EOS), shown in (3) [34]. The fit

Table A1 and table A2 found in the appendix list all of the data calculated in this work for elements in the FCC and HCP structures respectively. Comparisons to selected experimental data are given in figures 4a and 4b. In these figures, as throughout figures in the paper, lines are superimposed as a guide to help the reader see trends in the data. For the FCC systems, calculated vacancy formation enthalpies are generally lower than the corresponding experimental data. It has previously been proposed that, due to an intrinsic error in treating surfaces, GGA DFT can underestimate H_{vf} for certain systems [35]. The intrinsic surface error involved can be removed by adding a correction value that is a function of electron density [35], although this correction is not included in this work. Based on the limited data in figure 4a, this same error is not seen with the HCP systems, as the calculated HCP vacancy formation enthalpies generally overestimate the experimental results. Further work is necessary to understand why the typical errors between FCC and HCP, which are both quite similar close-packed structures, should be so different. The H_{vm} values calculated in this work generally agree well with the experimental data.

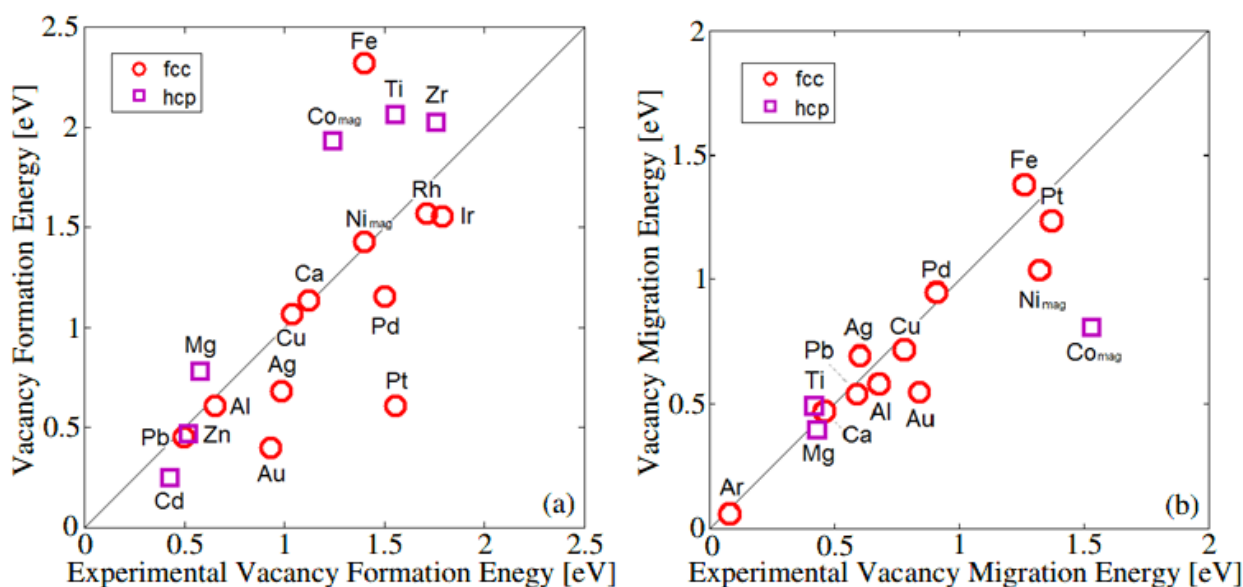
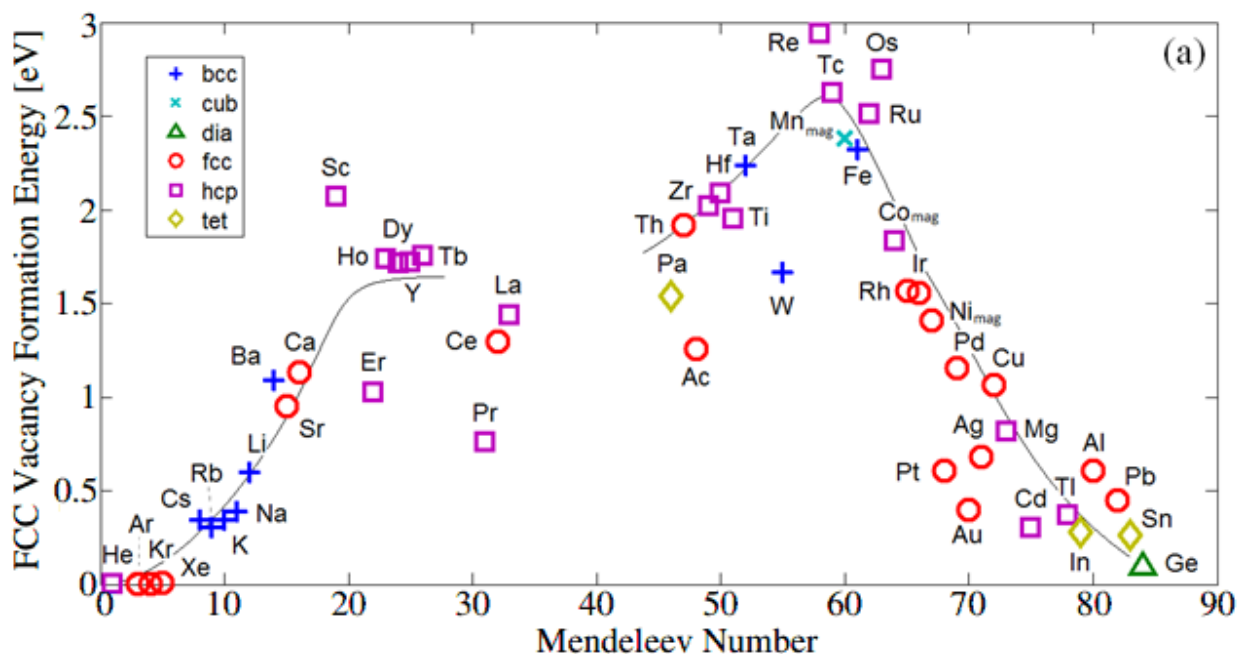


Figure 4. (a) H_{vf} values from this work versus experimental data [2, 36, 37] for the FCC and HCP systems. FCC Fe is non-spin-polarized in this figure. Calculations including spin-polarization are here forth denoted by the _{mag} subscript, where all labels not containing this subscript denote non-spin-polarized calculations. (b) H_{vm} plotted against experimental data [2, 36-39] for both the FCC and HCP systems. Only in-plane HCP migration data are included in this plot. All data presented in both of these plots are for elements in their stable, naturally occurring crystal structures.

It should be noted that previous *ab initio* work has identified an anomalous monovacancy migration pathway for Group IV HCP metals with BCC-HCP phase transitions at higher temperatures. For migration in the basal plane, the migrating atom finds a BCC-like atomic environment along the pathway which is an energy local minimum that exists between the two saddle points [40]. It is very possible that the naturally BCC elements treated in this work follow this tendency when placed in the HCP structure. Because only a single image is used in the CI-NEB calculations, it is possible that the single image falls into this local minimum, thereby underestimating the migration barrier.

In figure 5 and figure 6, the H_{vf} and H_{vm} values respectively are plotted against the elements' Mendeleev numbers. The Mendeleev number is a phenomenological coordinate that captures size, electronegativity, valence, and bond orbitals in a single chemical scale in which each element is assigned one unique number [41]. It is important to note in these figures that the legend corresponds to the natural crystal structure of each element at standard temperature and pressure (STP), not to the structure in which the elements were relaxed. These plots show that the vacancy formation and migration energies are generally smooth and continuous when plotted as a function of chemical species. Notable in both figures is the peak formed by the middle transition metals (Re, Os, Tc, Ru, Mn, and Fe). Also notable are the troughs formed by the weakly interacting elements with essentially zero defect energetics (He, Ar, Kr etc.).



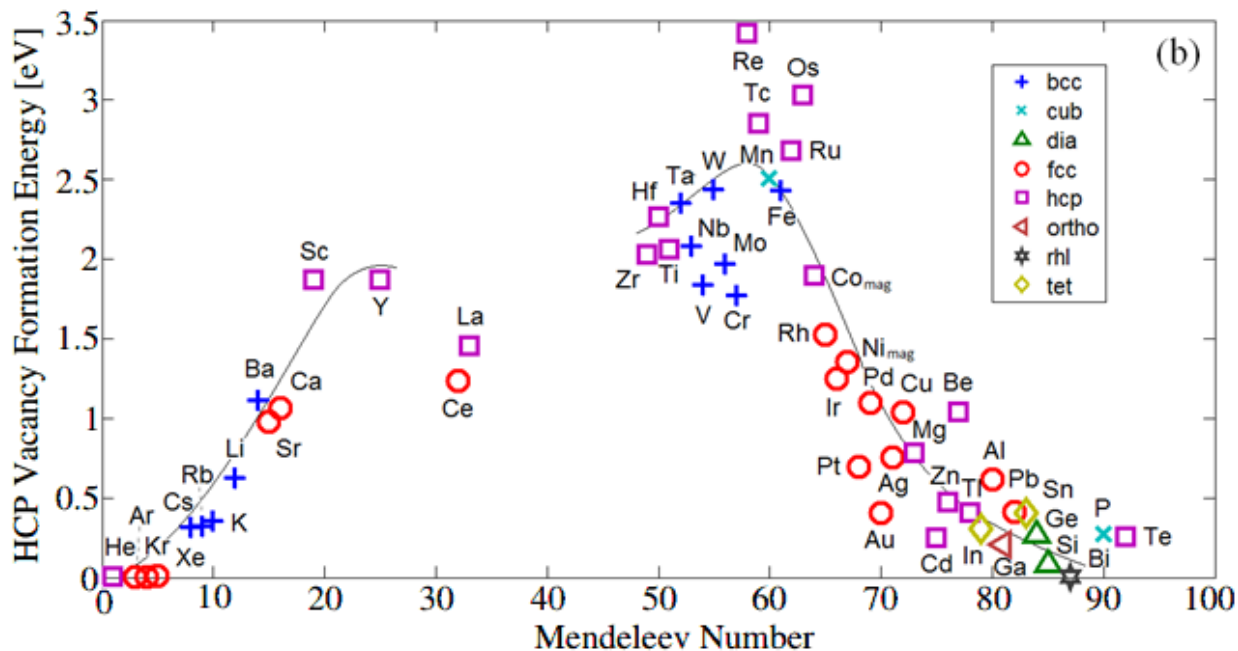
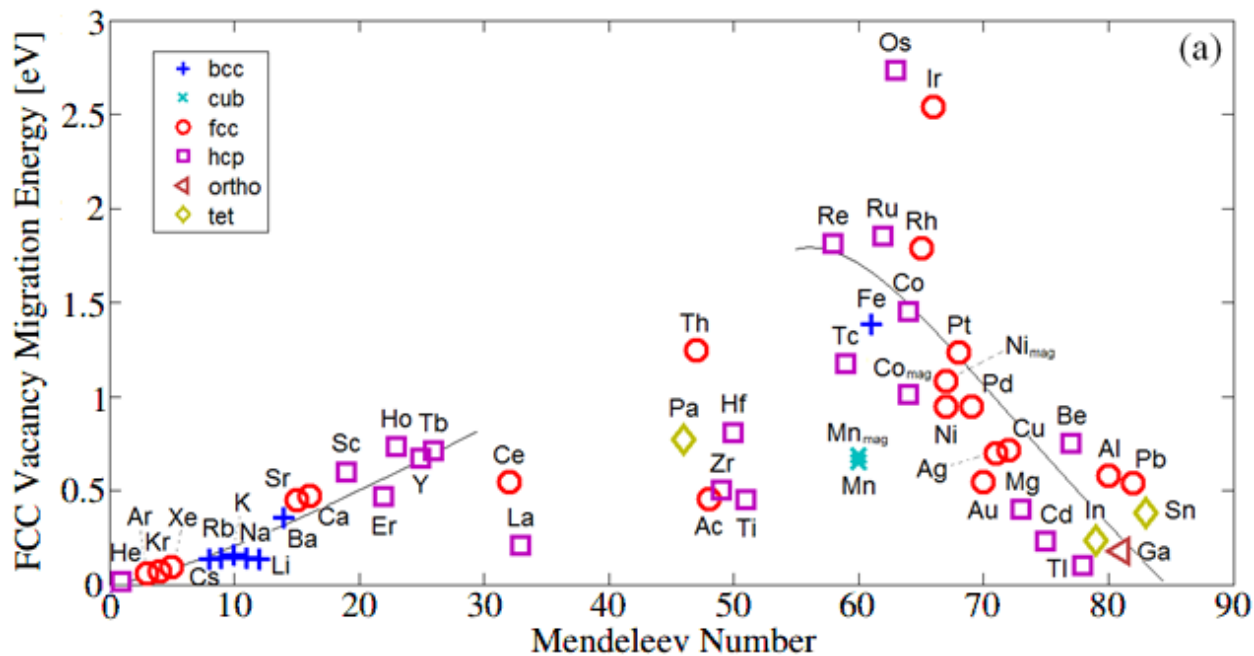


Figure 5. H_{vf} versus Mendeleev number for elements relaxed in the (a) FCC structure and (b) HCP structure. The legend corresponds to the natural crystal structure of each element at standard temperature and pressure (STP), not the structure in which the elements were relaxed. In the case of elements that are gas-phase at room temperature, the legend corresponds to high pressure or low temperature -induced solid phase structures [42]. Gray lines are superimposed as a visual guide for seeing trends. Gaps in the guiding line correspond to portions of the plot that are difficult to interpret due to insufficient or scattered data.



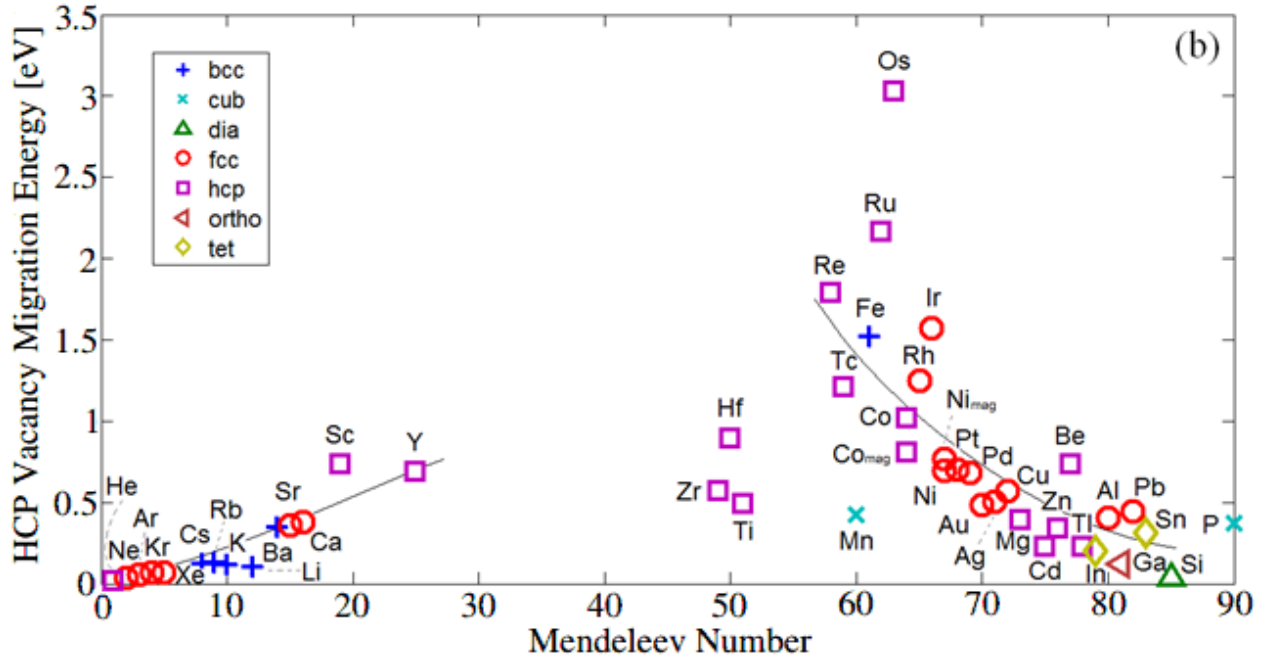


Figure 6. H_{vm} versus Mendeleev number for elements relaxed in the (a) FCC structure and (b) HCP structure. All data in this plot and any future plots containing HCP migration barriers pertain exclusively to the in-plane migration barriers. Gray lines are superimposed as a visual guide for seeing trends. Gaps in the guiding line correspond to portions of the plot that are difficult to interpret due to insufficient or scattered data.

The results in figures 5 and 6 suggest at least some correlation between H_{vf} and H_{vm} and the bonding strength of the systems. To see these trends more clearly, in figure 7 and figure 8, the calculated H_{vf} and H_{vm} values are plotted against the cohesive energy of each element. Figures 7a and 7b show an overall trend of linearly increasing vacancy formation energy with increasing cohesive energy for both structures. A linear regression was performed on both sets of data. For the FCC data, a fitted slope of 0.317 was found with $R^2 = 0.655$ (RMS Error = 0.477 eV). For the HCP data, a slope of 0.323 was found with $R^2 = 0.683$ (RMS Error = 0.501 eV). These results are consistent with the fact that a simple fixed bond-strength model is not sufficient in describing the energetics of vacancy formation. Because the remaining bonds nearby are strengthened when the vacancy is formed, and because the lattice can lower its energy by relaxing into the void space formed by the missing atom, it is common for H_{vf} to fall somewhere between $0.2E_C$ and $0.8E_C$ [38]. Figures 8a and 8b show a continuous trend of H_{vm} with E_C . A fit was performed using data points with cohesive energies less than 6 eV. In this regime, the data appears linear with a slope of 0.191 (RMS Error = 0.243 eV) for the FCC data and 0.159 (RMS Error = 0.209 eV) for the HCP data. The effects of the cohesive energy are especially strong when $E_C > 6$ eV. In this range, there appears to be a non-linear increase of the migration enthalpy with increasing cohesive energy. These trends are consistent with the intuition that the migrating atom's mobility is strongly hindered by an increasingly cohesive lattice. However, the trend of H_{vm} with E_C does not appear to be simply linear, and a change in the migration physics may occur for some of the most stable elements. These simple correlations could provide quick screening for H_{vf} and H_{vm} as the cohesive energy is much faster to calculate than either H_{vf} or H_{vm} .

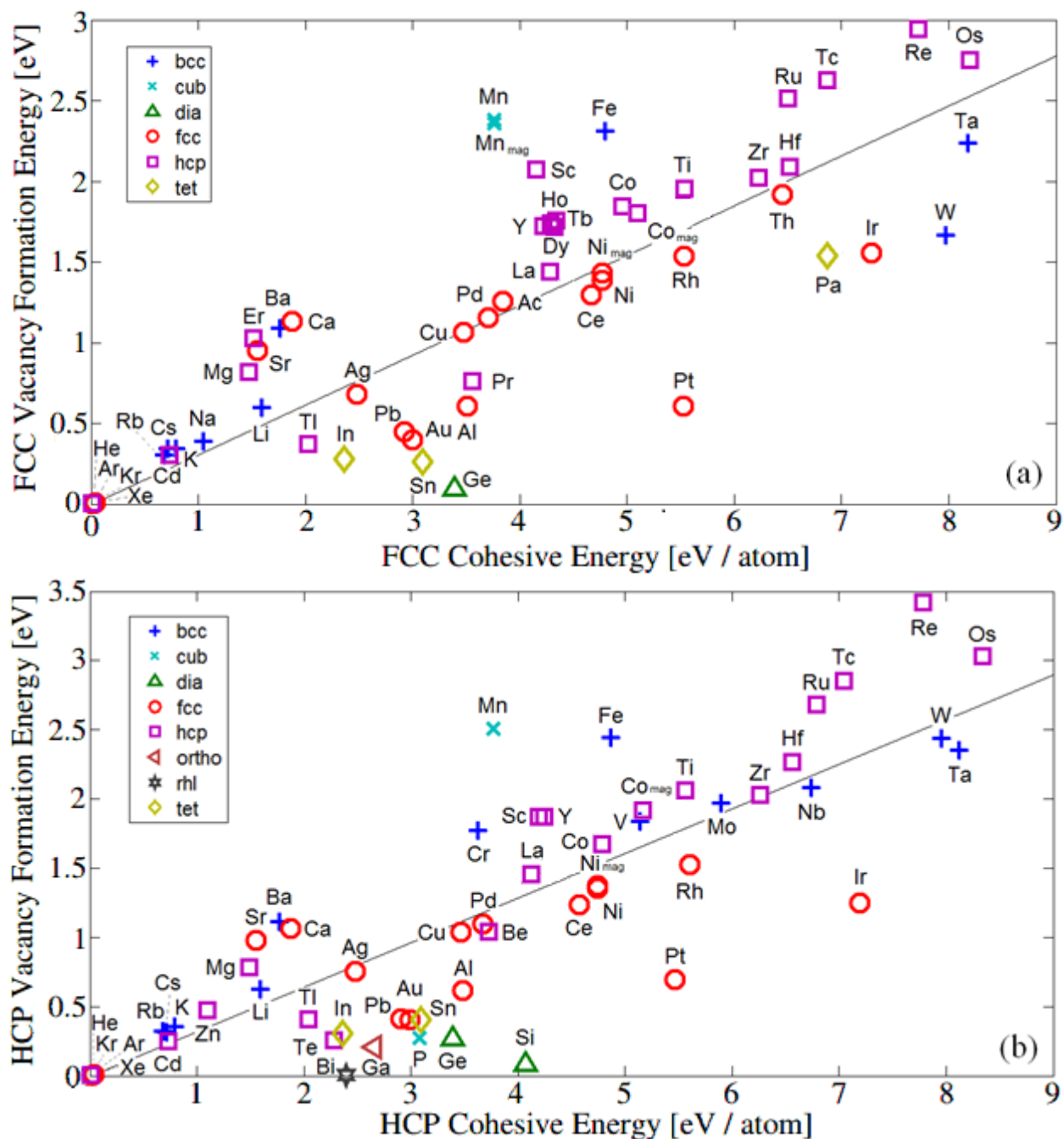
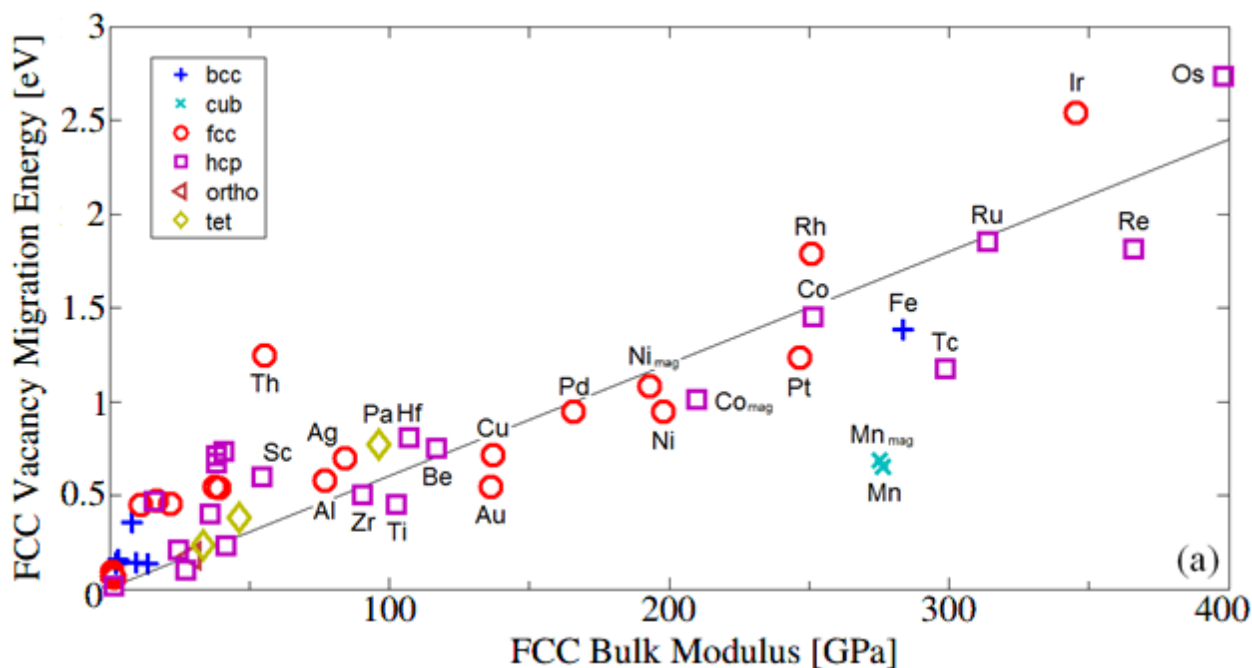


Figure 7. H_v versus E_c for elements relaxed in the (a) FCC structure and (b) HCP structure. The drawn lines are fitted with slopes of 0.317 for the FCC data ($R^2 = 0.655$, RMS Error = 0.477 eV), and 0.323 for the HCP data ($R^2 = 0.683$, RMS Error = 0.501 eV).

regime, H_{vm} increases sharply with cohesive energy, showing that a change in the migration physics may occur for some of the most stable elements. A gray dashed line is added as a visual guide to aid in seeing the increase.

As atomic migration can be considered a bond-breaking phenomenon, it is intuitively reasonable that H_{vm} might scale with cohesive energy. However, the migrating atom must also push its way through a constrained space, and therefore H_{vm} might additionally be expected to scale with elastic properties. This is explored further in figures 9a and 9b, in which the calculated H_{vm} values are plotted against the calculated bulk moduli of the corresponding elements in the FCC or HCP structures. The plots for both structures show a linearly increasing trend with bulk modulus with slopes of 5.5×10^{-3} eV/GPa ($R^2 = 0.70$, RMS Error = 0.329 eV) for FCC (figure 9a) and 5.1×10^{-3} eV/GPa ($R^2 = 0.79$, RMS Error = 0.273 eV) for HCP (figure 9b).



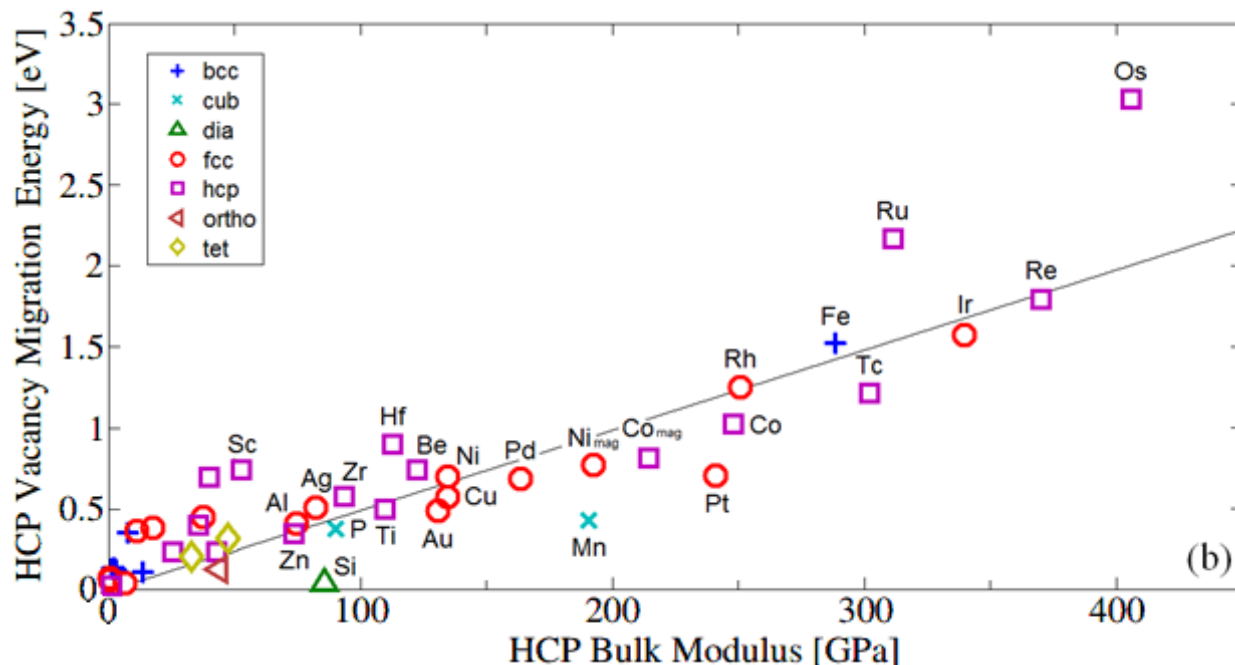


Figure 9. H_{vm} versus bulk modulus for elements relaxed in the (a) FCC structure and (b) HCP structure. The drawn lines are fitted with slopes of 5.5×10^{-3} eV/GPa ($R^2 = 0.70$, RMS Error = 0.329 eV) for the FCC data, and 5.1×10^{-3} eV/GPa ($R^2 = 0.79$, RMS Error = 0.273 eV) for the HCP data. The bulk moduli are obtained by calculations done in this work and pertain to the FCC or HCP structure.

The linear trends seen in figures 9a and 9b suggest a simple relationship between the migration energy and the bulk modulus. Previous work has derived expressions relating H_{vm} to elastic properties in the form $H_{vm} = f a^3 \langle c \rangle$, where f is a structure factor, a is the lattice constant, and $\langle c \rangle$ is an average elastic constant formed from c_{11} , c_{12} , and c_{44} (see Ehrhart p. 98) [2, 43]. Using the bulk modulus as the elastic constant and performing a linear fit ($R^2 = 0.84$, RMS Error = 0.240 eV) on the H_{vm} data versus $a^3 B_0$ yields the expression given in (4):

$$H_{vm} = 0.016 a^3 B_0 \quad (4)$$

This equation applies only to FCC systems due to the relatively more complex symmetries in the HCP structure. Attempts were made to find a similar expression for the H_{vm} of HCP systems by substituting the volume of four HCP atoms for a^3 in (4) and performing a linear fit. However, the resulting equation provided no better representation of H_{vm} than the regression against the bulk modulus itself shown in figure 9b. The validity of (4) is depicted in figure 10 in which H_{vm} values given by (4) are plotted against the *ab initio* H_{vm} values obtained in this work.

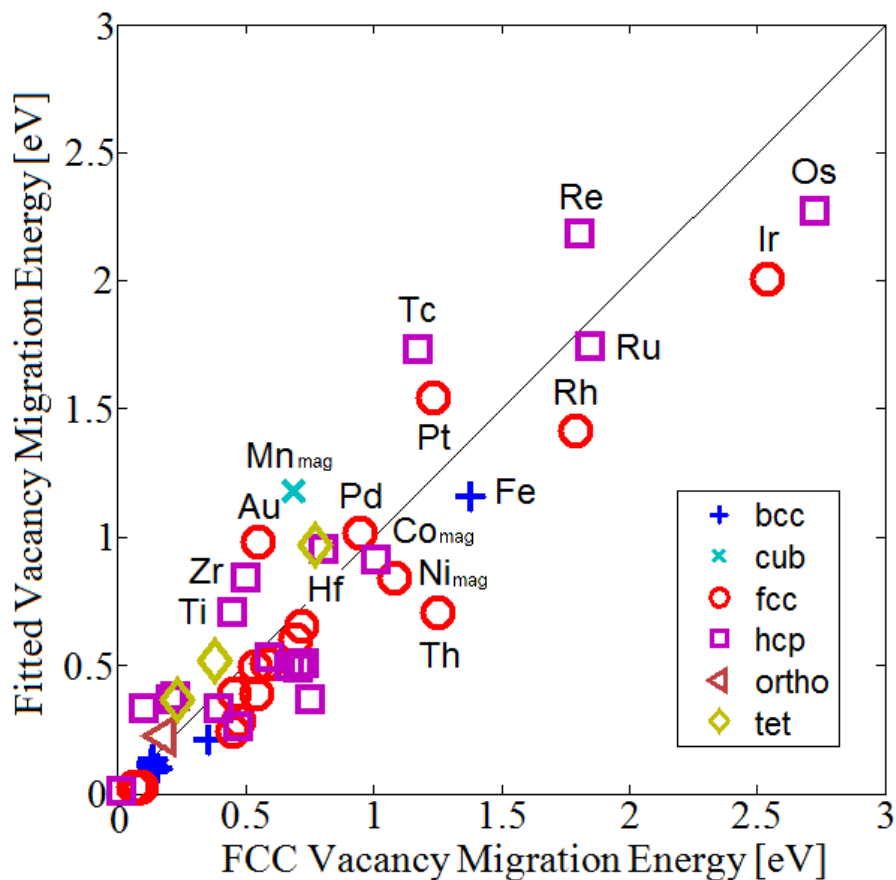


Figure 10. *Ab initio* H_{vm} values plotted against values generated using (4). These data correspond only to the natural and metastable FCC structure-element combinations. The RMS error is 0.24 eV with a maximum error of 0.56 eV.

4. Conclusion

Using a high-throughput approach, a large elemental diffusion energetics database was created that contains data for both metastable and naturally occurring systems. Having such a database is advantageous in that the methods used to calculate the results are rigorously consistent, accurate, and able to provide data for theoretical systems on which it is potentially impossible to perform experiments. The inclusion of data for metastable systems is useful if these systems are ever synthesized or if metastable predictions are required as end-member data inputs for trend fitting software, such as the thermodynamic simulations performed in CALPHAD.

By plotting the data against various chemical, physical, and mechanical properties, several important controlling factors for diffusion in the elemental systems have been discovered. These correlations are fundamental to understanding diffusion in various systems. The vacancy formation enthalpy was shown to depend linearly on the cohesive energy, and the vacancy migration enthalpy plotted against cohesive energy showed an initial linear region followed by a rapidly increasing region for the most stable systems. As vacancy formation and migration can be interpreted as bond-breaking phenomena, cohesive energy as a controlling factor makes intuitive sense. The relationship shown by equation (4) demonstrates that there is also a strong correlation between the bulk modulus and the migration enthalpy. As the migrating atom must push through a constrained space, incompressibility as a controlling factor also makes intuitive

sense. It is interesting to observe that, in general, the metastable data fit into the same trends as those formed by data from the naturally occurring systems.

The above correlations could all be used as screening tools for more complicated systems, such as alloys, although the validity of the correlations must be demonstrated for alloy systems. Properties such as the cohesive energy and the bulk modulus are less computationally intensive to compute, and therefore knowledge of these properties as controlling factors for migration could be useful in the screening process for specific migration energetics. This work has by no means explored all of the potential ways that the database can be used to discover trends of diffusion energetics with other properties of solids. Additional high-throughput diffusion work will increasingly generate more data and allow new correlations to be discovered.

Acknowledgements

Financial support for Thomas Angsten provided by the US Department of Energy (DOE) Nuclear Engineering University Program (NEUP) program under grant No. 10-888. Financial support for Tam Mayeshiba provided by the National Science Foundation (NSF) Graduate Fellowship Program under grant No. DGE-0718123. Financial support for Dane Morgan, Henry Wu and travel, materials and supplies provided by NSF Software Infrastructure for Sustained Innovation (SI²), award No. 1148011. MAST was developed at the University of Wisconsin-Madison by Tam Mayeshiba, Glen Jenness, Thomas Angsten, Henry Wu, and Parker Sear under NSF award No. 1148011. This research used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. NERSC computational resources were provided through the Center for Nanophase Materials Sciences, which is sponsored at Oak Ridge National Laboratory by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The authors gratefully acknowledge helpful discussions with M. Asta and K. Ding.

Appendix

Table A1. (FCC) Listing of the cohesive energy, E_{coh} ; bulk modulus, K ; pressure derivative of the bulk modulus, K' ; equilibrium atomic volume of the undefected cell, Ω ; and diffusion energetics, H_{vf} and H_{vm} , for elements relaxed in the FCC structure. See section 2 (2.6 for columns 2-4, 2.4 and 2.5 for columns 6 and 7) to see how these parameters were defined and calculated.

Symbol	E_{coh} (eV/atom)	K (GPa)	K'	Ω ($\text{\AA}^3/\text{atom}$)	H_{vf} (eV)	H_{vm} (eV)
Ac	3.70	21.63	2.03	45.37	1.26	0.45
Ag	2.49	84.11	5.15	17.87	0.68	0.70
Al	3.43	76.90	4.23	16.47	0.61	0.58
Ar	0.02	1.60	3.45	45.00	0.01	0.06
Au	2.99	136.19	7.45	18.07	0.40	0.53
Ba	1.87	8.01	2.81	64.13	1.09	0.36
Be	3.64	116.96	3.90	7.88	-0.06 ^a	0.75
Ca	1.91	16.82	1.70	42.17	1.13	0.47
Cd	0.74	41.66	5.71	22.60	0.30	0.23
Ce	4.58	37.65	4.10	26.10	1.30	0.54
Co	4.92	251.54	5.17	10.30	1.84	1.45
Co _{mag}	5.11	209.89	5.07	10.90	1.79	1.01
Cs	0.70	2.40	2.36	116.46	0.34	0.13
Cu	3.48	136.99	5.86	11.97	1.07	0.72
Dy ^b	4.23	-	-	31.37	1.72	-
Er	1.55	16.02	2.66	40.96	1.02	0.47
Fe	4.69	283.59	4.89	10.22	2.32	1.38
Ga	2.61	29.64	6.02	18.91	-0.04 ^a	0.18
Ge ^b	3.40	-	-	19.51	0.10	-
He	0.01	1.60	3.36	16.54	0.00	0.02
Hf	6.41	107.35	3.59	22.26	2.09	0.81
Ho	4.20	40.86	4.04	30.88	1.74	0.73
In	2.31	33.65	4.53	27.33	0.28	0.23
Ir	7.23	345.27	5.44	14.53	1.55	2.54
K	0.86	3.20	6.19	73.60	0.34	0.16
Kr	0.02	0.80	7.90	57.44	0.00	0.07
La	4.22	24.83	3.36	37.10	1.44	0.21
Li	1.61	13.62	3.75	20.22	0.60	0.13
Mg	1.49	36.05	4.04	23.03	0.82	0.41
Mn ^c	3.76	276.38	5.42	10.73	2.38	0.65
Mn _{mag} ^c	3.76	275.57	5.58	10.72	2.36	0.69
Na	1.11	9.61	3.08	35.08	0.38	0.14
Ni	4.77	197.87	5.29	10.85	1.39	0.94
Ni _{mag}	4.83	193.06	5.45	10.90	1.43	1.08
Os	8.17	398.14	4.99	14.29	2.75	2.73
Pa	6.86	96.13	4.16	25.21	1.54	0.77
Pb	2.94	39.25	4.07	31.69	0.45	0.54
Pd	3.71	165.83	6.62	15.37	1.16	0.95

Pr ^b	3.58	-	-	23.79	0.76	-
Pt	5.45	246.74	6.10	15.67	0.61	1.24
Rb	0.76	2.40	7.13	90.81	0.30	0.14
Re	7.73	366.10	4.73	14.94	2.94	1.81
Rh	5.63	250.74	5.86	14.13	1.57	1.79
Ru	6.56	314.03	4.81	13.88	2.51	1.85
Sc	4.09	54.47	4.07	24.43	2.07	0.60
Sn	3.11	46.46	6.10	27.82	0.26	0.38
Sr	1.61	11.22	3.72	54.74	0.95	0.45
Ta ^b	8.09	198.67	3.86	18.63	2.23	-
Tb	4.26	38.45	4.55	31.88	1.75	0.71
Tc	6.84	298.81	4.91	14.50	2.62	1.17
Th	6.34	55.28	3.50	32.07	1.92	1.25
Ti	5.43	102.54	2.89	17.25	1.95	0.45
Tl	1.99	27.24	5.55	30.50	0.37	0.10
W ^b	7.98	-	-	16.25	1.67	-
Xe	0.03	0.80	2.22	80.21	0.01	0.09
Y	4.14	38.45	4.10	32.35	1.72	0.67
Zr	6.21	90.52	4.42	23.26	2.02	0.50

^a These slightly negative values signify a lack of stability for these element-structure combinations.

^b The instability of these elements in the FCC structure precluded the gathering of certain data.

^c Due to convergence issues, FCC Mn and Mn_{mag} properties were calculated with a different PAW-PBE pseudopotential than that used to calculate HCP Mn properties.

Table A2. (HCP) Listing of the cohesive energy, E_{coh} ; bulk modulus, K ; pressure derivative of the bulk modulus, K' ; equilibrium atomic volume of the undefected cell, Ω , and diffusion energetics, H_{vf} , $H_{vm \perp}$ (in-plane), and $H_{vm \parallel}$ (out-of-plane) for elements relaxed in the HCP structure. See section 2 (2.6 for columns 2-4, 2.4 and 2.5 for columns 6 through 8) to see how these parameters were defined and calculated.

Symbol	E_{coh} (eV/atom)	K (GPa)	K'	Ω ($\text{\AA}^3/\text{atom}$)	H_{vf} (eV)	$H_{vm \perp}$ (eV)	$H_{vm \parallel}$ (eV)
Ag	2.49	82.51	5.15	17.93	0.76	0.50	0.60

Al	3.40	74.50	4.70	16.63	0.62	0.41	0.46
Ar ^b	0.02	0.88	7.43	46.11	0.01	0.06	-
Au	2.98	130.58	7.58	18.10	0.40	0.49	0.57
Ba	1.87	8.01	3.38	63.55	1.11	0.34	0.37
Be	3.72	122.57	3.57	7.92	1.04	0.74	0.87
Bi ^b	2.38	-	-	31.44	0.00	-	-
Ca	1.91	17.62	3.44	42.39	1.06	0.38	0.42
Cd	0.74	43.26	6.04	22.60	0.25	0.23	0.16
Ce ^b	4.49	-	-	26.43	1.23	-	-
Co	4.90	248.34	3.29	10.33	1.67	1.02	1.19
Co _{mag}	5.12	214.69	4.91	10.85	1.93	0.81	0.78
Cr ^b	3.63	-	-	11.77	1.77	-	-
Cs	0.70	1.60	7.60	116.42	0.31	0.12	0.13
Cu	3.47	134.58	5.61	12.00	1.04	0.57	0.69
Fe	4.77	288.39	5.60	10.16	2.43	1.52	1.51
Ga	2.60	44.06	4.37	18.92	0.21	0.12	0.16
Ge ^b	3.40	-	-	19.25	0.26	-	-
He	0.01	1.60	3.46	16.54	0.00	0.02	0.01
Hf	6.48	112.95	3.59	22.26	2.26	0.89	1.00
In	2.31	32.84	4.68	27.40	0.31	0.20	0.21
Ir	7.16	339.66	3.32	14.60	1.25	1.57	1.95
K	0.86	3.20	5.95	73.72	0.35	0.12	0.14
Kr	0.02	0.80	6.73	58.10	0.00	0.07	0.07
La ^b	4.20	-	-	37.42	1.45	-	-
Li	1.61	13.62	3.82	20.24	0.63	0.10	0.12
Mg	1.50	36.05	4.31	22.85	0.78	0.40	0.42
Mn	3.82	190.66	11.04	10.63	2.51	0.43	0.11
Mo ^b	5.91	-	-	15.95	1.96	-	-
Nb ^b	6.71	-	-	18.61	2.08	-	-
Ne	0.01	6.41	1.65	19.15	-0.01 ^a	0.04	0.04
Ni	4.74	134.58	12.07	10.87	1.35	0.69	0.79
Ni _{mag}	4.81	192.26	5.34	10.94	1.37	0.77	0.89
Os	8.31	406.15	5.04	14.23	3.03	3.03	3.22
P ^b	3.09	90.52	3.96	15.12	0.28	0.38	-
Pb	2.92	37.65	4.49	31.51	0.41	0.44	0.41
Pd	3.67	163.42	6.50	15.45	1.10	0.68	0.73
Pt ^b	5.39	241.13	6.22	15.77	0.70	0.70	-
Rb	0.76	2.40	6.45	91.19	0.32	0.13	0.12
Re	7.79	370.10	4.67	14.89	3.42	1.79	1.17
Rh	5.59	250.74	5.76	14.18	1.53	1.25	1.47

Ru	6.68	311.62	5.16	13.81	2.68	2.17	2.18
Sc	4.14	52.87	4.82	24.48	1.87	0.74	0.69
Si	4.04	85.72	4.88	14.35	0.08	0.04	0.26
Sn	3.11	47.26	5.92	27.55	0.40	0.31	0.33
Sr	1.61	11.22	3.77	55.16	0.98	0.36	0.40
Ta ^b	8.05	-	-	18.55	2.35	-	-
Tc	6.91	302.01	4.90	14.44	2.85	1.21	0.63
Te ^b	2.23	48.87	5.03	31.47	0.25	-	-
Ti	5.49	109.75	5.07	17.29	2.06	0.49	0.59
Tl	2.01	25.63	5.47	30.97	0.40	0.23	0.16
V ^b	5.15	-	-	13.75	1.84	-	-
W ^b	7.96	-	-	16.32	2.44	-	-
Xe	0.03	0.80	2.32	79.60	0.01	0.07	0.08
Y	4.16	40.05	3.72	32.82	1.87	0.69	0.67
Zn	1.10	73.70	6.02	15.23	0.47	0.34	0.20
Zr	6.25	93.73	4.34	23.44	2.03	0.57	0.70

^a This slightly negative value signifies a lack of stability for this element-structure combination.

^b The instability of these elements in the HCP structure precluded the gathering of certain data.

References

- [1] Shewmon P 1989 *Diffusion in Solids* (Warrendale, Pa: TMS)
- [2] Ehrhart P 1991 *Properties and Interactions of Atomic Defects in Metals and Alloys* (Berlin: Springer)
- [3] Kornblit L, Pelleg J and Rabinovitch A 1977 Self-diffusion calculation for fcc metals *Phys. Rev. B* **16** 1164-7
- [4] Choudhury S, Barnard L, Tucker J D, Allen T R, Wirth B D, Asta M and Morgan D 2011 Ab-initio based modeling of diffusion in dilute bcc Fe–Ni and Fe–Cr alloys and implications for radiation induced segregation *Journal of Nuclear Materials* **411** 1-14
- [5] Mantina M, Wang Y, Arroyave R, Chen L, Liu Z and Wolverton C 2008 First-Principles Calculation of Self-Diffusion Coefficients *Physical Review Letters* **100** 215901
- [6] Ganeshan S, Hector Jr. L G and Liu Z-K 2010 First-principles study of self-diffusion in hcp Mg and Zn *Computational Materials Science* **50** 301-7

- [7] Morgan D, Ceder G and Curtarolo S 2005 High-throughput and data mining with ab initio methods *Meas. Sci. Technol.* **16** 296-301
- [8] Vitos L, Ruban A V, Skriver H L and Kollar J 1998 The surface energy of metals *Surface Science* **411** 186-202
- [9] Curtarolo S, Morgan D and Ceder G 2005 Accuracy of ab initio methods in predicting the crystal structures of metals: A review of 80 binary alloys *Calphad* **29** 163-211
- [10] Curtarolo S, Setyawan W, Wang S, Xue J, Yang K, Taylor R H, Nelson L J, Hart G L W, Sanvito S, Buongiorno-Nardelli M, Mingo N and Levy O 2012 AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations *Computational Materials Science* **58** 227-35
- [11] Greeley J, Stephens I E L, Bondarenko A S, Johansson T P, Hansen H a, Jaramillo T F, Rossmeisl J, Chorkendorff I and Nørskov J K 2009 Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. *Nature chemistry* **1** 552-6
- [12] Jain A, Hautier G, Moore C, Kang B, Lee J, Chen H, Twu N and Ceder G 2012 A Computational Investigation of $\text{Li}_9\text{M}_3(\text{P}_2\text{O}_7)_3(\text{PO}_4)_2$ ($\text{M} = \text{V}, \text{Mo}$) as Cathodes for Li Ion Batteries *Journal of The Electrochemical Society* **159** A622
- [13] Materials Project <http://www.materialsproject.org/>
- [14] Bahn S R and Jacobsen K W 2002 An Object-Oriented Scripting Interface to a Legacy Electronic Structure Code *Computing in Science & Engineering* **4** 56-66
- [15] Kresse G and Hafner J 1993 Ab initio molecular dynamics for liquid metals *Physical Review B* **47** 558-61
- [16] Kresse G and Hafner J 1994 Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium *Physical Review B* **49** 14251
- [17] Kresse G and Furthmüller J 1996 Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set *Computational Materials Science* **6** 15-50
- [18] Kresse G and Furthmüller J 1996 Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set *Physical Review B* **54** 11169
- [19] Ong S P, Richards W D, Jain A, Hautier G, Kocher M, Cholia S, Gunter D, Chevrier V L, Persson K A and Ceder G 2013 Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis *Computational Materials Science* **68** 314-319
- [20] Perdew J P, Burke K and Ernzerhof M 1996 Generalized Gradient Approximation Made Simple *Physical Review Letters* **77** 3865-8
- [21] Perdew J, Burke K and Ernzerhof M 1996 Errata: Generalized Gradient Approximation Made Simple. *Physical review letters* **77** 3865-8
- [22] Blochl P E 1994 Projector augmented-wave method *Physical Review B* **50** 17953
- [23] Kresse G and Joubert J 1999 From ultrasoft pseudopotentials to the projector augmented wave method. *Physical Review B* **59** 1758
- [24] Methfessel M and Paxton A T 1989 High-precision sampling for Brillouin-zone integration in metals *Physical Review B* **40** 3616-21
- [25] Reed-Hill R E and Abbaschian R 1994 *Physical Metallurgy Principles* (Boston: PWS)
- [26] Monkhorst H J and Pack J D 1976 Special points for Brillouin-zone integrations *Physical Review B* **13** 5188-92
- [27] Castleton, C.W.M., A. Hoglund, and S. Mirbt, *Density functional theory calculations of defect energies using supercells*. Modelling and Simulation in Materials Science and Engineering, 2009. **17**(8).
- [28] Lany, S. and A. Zunger, *Accurate prediction of defect properties in density functional supercell calculations*. Modelling and Simulation in Materials Science and Engineering, 2009. **17**(8).
- [29] Castleton, C.W.M., A. Hoglund, and S. Mirbt, *Managing the supercell approximation for charged defects in semiconductors: Finite-size scaling, charge correction factors, the band-gap problem, and the ab initio dielectric constant*. *Physical Review B*, 2006. **73**(3).
- [30] Castleton, C.W.M. and S. Mirbt, *Finite-size scaling as a cure for supercell approximation errors in calculations of neutral native defects in InP*. *Physical Review B*, 2004. **70**(19).
- [31] S.-k. Lin, C.-k. Yeh, B. Puchala, Y.-L. Lee, and D. Morgan, Ab initio energetics of charge compensating point defects: A case study on MgO, *Computational Materials Science* **73**, p. 41-55 (2013).
- [32] Henkelman G, Uberuaga B P and Jonsson H 2000 A climbing image nudged elastic band method for finding saddle points and minimum energy paths *J Chem Phys* **113** 9901-4
- [33] Vineyard G H 1957 Frequency factors and isotope effects in solid state rate processes *J. Phys. Chem. Solids* **3** 121-7
- [34] Birch F 1947 Finite Elastic Strain of Cubic Crystals *Physical Review* **71** 809-824

- [35] Mattsson T and Mattsson A 2002 Calculating the vacancy formation energy in metals: Pt, Pd, and Mo *Physical Review B* **66** 214110
- Nazarov R Hickel T and Neugebauer J 2012 Vacancy formation energies in fcc metals: Influence of exchange-correlation functionals and correction schemes *Physical Review B* **85** 144118
- [36] Miedema A R 1979 The Formation Enthalpy of Monovacancies in Metals and Intermetallic Compounds *Zeitschrift Fur Metallkunde* **6** 345-353
- [37] Sheng H, Kramer M, Cadien a, Fujita T and Chen M 2011 Highly optimized embedded-atom-method potentials for fourteen fcc metals *Physical Review B* **83** 134118
- [38] Flynn C P 1972 *Point Defects and Diffusion* (Oxford: Clarendon)
- [39] Smithells C J and Brandes E A 1983 *Smithells Metals Reference Book* (London: Butterworths)
- [40] Shang S L, Hector L G, Wang Y, and Liu Z K 2011 Anomalous energy pathway of vacancy migration and self-diffusion in hcp Ti *Physical Review B* **83** 224104
- [41] Pettifor D G 1984 A Chemical Scale for Crystal-Structure Maps *Solid State Communications* **51** 31-34
- [42] Donohue J 1974 *The Structures of the Elements* (New York: John Wiley & Sons)
- [43] Flynn C P 1968 Atomic Migration in Monatomic Crystals *Physical Review* **171** 682-698