SAND2016-9936R

DENSITY FUNCTIONAL THEORY APPLIED TO TRANSITION METAL ELEMENTS AND BINARIES: DEVELOPMENT, APPLICATION, AND RESULTS OF THE V-DM/16 TEST SET

ELIZABETH R. DECOLVENAERE* AND ANN E. MATTSSON†

Abstract. Density functional theory (DFT) is undergoing a shift from a descriptive to a predictive tool in the field of solid state physics, heralded by a spike in "high-throughput" studies. However, methods to rigorously evaluate the validity and accuracy of these studies is lacking, raising serious questions when simulation and experiment disagree. In response, we have developed the V-DM/16 test set, designed to evaluate the experimental accuracy of DFT's various implementations for periodic transition metal solids. Our test set evaluates 26 transition metal elements and 80 transition metal alloys across three physical observables: lattice constants, elastic coefficients, and formation energy of alloys. Whether or not a functional can accurately evaluate the formation energy offers key insights into whether the relevant physics are being captured in a simulation, an especially important question in transition metals where active *d*-electrons can thwart the accuracy of an otherwise well-performing functional. Our test set captures a wide variety of cases where the unique physics present in transition metal binaries can undermine the effectiveness of "traditional" functionals. By application of the V/DM-16 test set, we aim to better characterize the performance of existing functionals on transition metals, and to offer a new tool to rigorously evaluate the performance of new functionals in the future.

1. Introduction. In the last decade, density functional theory (DFT) has shifted in use from a primarily *descriptive* to an increasingly *predictive* tool in the field of solid state physics[42]. However, compared to the field of quantum chemistry, solid-state DFT has only the most rudimentary standards of verification, validation, and reproducibility[11] – vital aspects in the design and use of a predictive method. The shift in focus, from descriptive to predictive, has been heralded by a spike in high-throughput studies[39, 65, 5, 9, 12, 21, 38], which leverage increasingly powerful and inexpensive computing resources to churn through tens or hundreds of thousands of structures in search of novel materials, properties, or behavior. Unfortunately, without rigorous methods to evaluate the accuracy and precision of DFT-derived results, the reliability of any of these studies is questionable when experiments and simulation (often) disagree[15, 48].

While several verification and validation studies have been performed on periodic systems[36, 37, 26, 71, 41, 24], these simulations all have been to compare various DFT codes against each other, and/or focused on lattice constants/volume, and/or limited to a small (and not necessarily representative) set of pure or binary metals and oxides. The most popular high-throughput databases have been excellent at providing input files capable of reproducing published outputs[25, 60, 13], but tend to concentrate only on one functional in their studies, usually PBE[49, 50]. Individual studies in other publications, while sometimes encompassing a greater spread of functionals, often provide insufficient information on input parameters such that results are difficult to compare and impossible to reproduce. Finally, in both cases, precious little attention is given to situations where DFT and experiment have strong, fundamental disagreements.

To address these issues, we have developed a test set designed to evaluate the experimental accuracy of DFT's implementations for periodic transition metal (TM) solids. Our test set, V-DM/16, evaluates 26 TM elements and 80 TM alloys across

*University of California, Santa Barbara, elizabeth@umail.ucsb.edu
†Sandia National Laboratories, aematts@sandia.gov

three physical observables: lattice constants, elastic coefficients, and formation energy of alloys. Our enumeration of structures for the binary compounds takes a hybrid approach, blindly enumerating configurations containing up to 6 atoms in the symmetrically-distinct supercell, and also containing experimentally observed phases (when known) and similar supercell decorations. Overall, our full test set contains roughly 12,000 structures; we additionally aim to provide mini and micro test sub-sets for fast and/or computationally cheap initial evaluations.

The V-DM/16 test set is designed to capture the wide variety of unique physics that arise in TM alloys, a consequence of the interaction of chemically active *d*-electrons[43]. These interactions are often ill-captured by present functionals[22], as most functionals are designed to focus on (and accurately solve) one "model" electron system or situation, e.g., the local density approximation[28] (LDA) functional models the uniform electron gas. Some functionals, such as PBESol[54] and AM05[3, 40], are modeled to solve systems with more "mixed" character, but these functionals still fail to capture the *simultaneous* highly localized and delocalized electron behavior that can arise in systems with active *d*- and *s*- electrons. Additionally, multiple elements and many alloys among the TMs have strong and sometimes complex magnetic characteristics[61]; our test set is designed to include representative cases of potential magnetic behaviors. Finally, late-period TMs have important chemical contributions from spin-orbit coupling[59, 66], a phenomena very poorly modeled by most common formulations of DFT[47]. Our test set considers multiple alloy systems where the presence (or lack) of spin-orbit coupling is critical in correctly predicting material properties.

2. Simulation Methods. All of our simulations were carried out using the Vienna Ab initio simulation package[29, 30, 32, 31] (VASP) with projector-augmented wave (PAW) pseudopotentials[6, 33]. Enumeration of superstructures, as well as management of high-throughput simulations, was handled by the CASM[57, 64, 58, 16] software package. All calculations were performed at zero temperature and pressure, with spin polarization, using the interpolation formula of Vosko, Wilk, and Nusair[69], and with automatically generated gamma-centered k-point meshes[45]. Electronic self-consistency loops were converged to energy changes below 10^{-3} meV, while ionic relaxations were converged to energy changes below 1 meV. The general simulation scheme consisted of two steps: first, repeated relaxation runs were performed until the simulation converged within 3 or fewer ionic steps, then, a final static run was performed with no changes allowed in cell geometry or ionic positions. This scheme was chosen to avoid problems introduced by possible non-isotropic changes in the basis set over a volume change during a single run. A quasi-Newton ionic relaxation scheme was used, RMM-DIIS[32], as implemented in VASP. During relaxation runs the method of Methfessel-Paxton[44] was used for electronic smearing; in most cases, the final static run switched to the tetrahedron method with Blöchl corrections[7] for smearing. Energy cutoffs, as well as k-point meshes, were chosen for each element and binary using the techniques described in Section 2.1. Specific functionals used, as well as special simulation schemes used with those functions, are described in Section 2.2.

2.1. Convergence Criteria. As formation energy forms such a large component of our test set metrics, ensuring numerical convergence is especially critical. Rather than aim for a particular k-point mesh density or energy cutoff in some global sense, each element was treated individually and converged, with respect to k-point mesh and energy cutoff, to within 1meV. This was achieved in two stages:

First, the k-point mesh was optimized, using the fully-automatic k-point gener-

ation scheme provided by VASP. For each element, a series of k-point lengths[1] were enumerated and calculations were performed with energy cutoffs of 150% of the VASP default value. Each cell was fully relaxed, and then a static run was performed. When the energy difference between two subsequent k-point lengths was less than 1 meV and *remained* less than 1 meV at increasing k-point densities, this value was taken as the converged k-point length. Step sizes were of 1 angstrom, and the starting k-point length was 30 angstroms.

Next, using the k-point length previously determined, a series of energy cutoffs were enumerated starting from 150% of the VASP default cutoff for that element. Again, the starting geometry was that of the experimental primitive cell, and each energy cutoff was allowed to independently relax, and a final static run was used to collect data. When energy differences between two subsequent energy cutoffs was less than 1 meV and *remained* less than 1 meV at increasing cutoffs, this value was taken as the converged energy cutoff. Step sizes were of 5 eV.

In all cases, starting structures were taken from the Materials Project[25] database, using primitive cells corresponding to the lowest-temperature experimentally observed structure. For binary compounds, the higher of the two components' k-point length and energy cutoff was used. Both k-point and energy cutoff convergences were independently performed for each functional explored. While not used as a *convergence* criteria, the convergence of cell volume and lattice constant *at* the converged k-point and energy cutoff values was collected.

2.2. Functionals Explored. To demonstrate the wide variety of behaviors captured by our test set, we ran simulations across eight different functionals, detailed in table 2.1. Our purpose here was two-fold: our test set characterizes the behavior of these functionals in the case of transition metals, and the differences in behaviors across these functionals demonstrates the necessity of our test set. The functionals chosen represent a wide variety of techniques used in DFT in materials science and different rungs on the "Jacob's ladder"[55] of functional accuracy. Specifically, we aimed to analyze the most commonly used functionals (PBE and PBESol), in addition to a representative of the local density approximation, the (historically) most commonly used generalized gradient approximation functionals, a subsystem functional, and meta-generalized gradient approximation functionals with historically good performance for transition metals.

All functionals were fully and self-consistently converged with regards to both electronic and ionic degrees of freedom (lattice constants and atom positions). For PZ, PW91, BLYP, PBE, and PBESol, simulations were carried out in a straightforward manner using the techniques described at the start of this section. For AM05, rTPSS, and M06-L, each configuration first underwent a single relaxation run using PBE, then, the resulting wavefunctions were used as the starting condition for the chosen functional[2]. For AM05 and M06-L, the chosen algorithms (All/Damped) were incompatible with tetrahedral smearing, and so the final static run was also performed using the smearing method of Methfessel-Paxton. The GW-labeled PAW-PBE potentials from the v.54 set were used in all simulations except for PZ, where the GW-labeled PAW-LDA potentials were used instead.

[1] A k-point "length" generates a k-point mesh along reciprocal direction $\hat{\imath}$ with $\max(1, l_k|\vec{b}_i| + 0.5)$ spacings, where l_k is the k-point length, and $\vec{b}_i$ is the reciprocal lattice vector along $\hat{\imath}$.

[2] Specifically, a single run was performed using the PBE functional, and then the WAVECAR (containing Kohn-Sham wave functions) copied over for use as a starting condition for the next run, this one using the listed functional

Name	Type	Algo
PZ[56]	Local Density Approximation	Normal
PW91[51, 52]	Generalized Gradient Approximation	Normal
BLYP[4, 35]	Generalized Gradient Approximation	Normal
PBE[49, 50]	Generalized Gradient Approximation	Normal
PBESol[54]	Generalized Gradient Approximation	Normal
AM05[3, 40]	Sub-System Functional	All/Damped
rTPSS[63, 53]	Meta-Generalized Gradient Approximation	Normal
M06-L[70]	Meta-Generalized Gradient Approximation	All/Damped

TABLE 2.1

Table of functionals used to test the V-DM/16 test set, indicating the functional's common name (usually an acronym), the general category of the functional, and the electronic convergence algorithm (according to VASP's tags) used with that functional.

We have specifically chosen not to analyze any DFT+U[2] techniques. The DFT+U functionals, which employ a Hubbard-U correction to the on-site coulomb interaction term, add an additional one or two degrees of freedom to our elements and binaries, respectively. Additionally, the U's are semi or totally empirical, usually chosen to match up some experimental parameter (often the band gap) to a simulation result. While some self-consistent methods exist to determine the U *in-situ*[34, 10, 1], these approaches are computationally involved and therefore beyond the scope of our analysis.

We have also neglected to include any hybrid-type functionals, owing primarily to the massively increased computational cost associated with hybrids. Not only can hybrid functionals be thousands of times slower and have significant memory requirements[62, 8, 17], but additionally, results for transition metals in the past have been poor or a-physical[26, 17, 22]. Finally, for both DFT+U and hybrid functionals, these techniques result in energies that are no longer purely *functionals* of the electron density, and hence, these techniques are not "true" Kohn-Sham DFT.

While the GW PAWs include scalar-relativistic corrections, we have not made any attempts to include additional relativistic terms in our analyses. While all calculations were performed using spin-polarization, the magnetic moments calculated were both collinear and decoupled from the crystal structure. Although for some TM binaries non-collinear magnetic effects (e.g., in Ni-Fe[61]) or spin-orbit coupling contributions (e.g., in Co-Pt[46]) can be important, the computational difficulty of performing these calculations in VASP puts them beyond the reach of our current high-throughput infrastructure.

3. The V-DM/16 Test Set. The V-DM/16 test set consists of 26 single-element TM systems and 80 TM binaries, and contains approximately 12,000 structures. Our choice of systems represents roughly one quarter of the potential space spanned by all possible TM binaries, offering a compromise between completeness and computational efficiency. While use of the entire structure space is recommended, we will in the future provide guidelines for "mini" and "micro" versions of the test set that consider decreasingly sized subsets of increasing significance. While significantly larger than the test sets traditionally found in quantum chemistry, the efficiency of periodic codes for "simple" bulk systems (i.e., small unit cells), as well as the increasing availability of cheap but powerful computational resources, ensures our test set remains accessible and justifiable in use.

3.1. Single-Component Systems. Of the potential space of transition metals, we have chosen to include elements in periods 4–6 and groups 3–11, excluding Lutetium. Lu is often considered a lanthanide instead of a transition metal, additionally, the f-electrons in Lu actively participate in determining the electronic structure[18], putting it outside the scope of this test-set. Group 12, containing zinc, cadmium, and mercury, was also excluded as these elements contain completely filled d-shells and therefore behave significantly differently than other metals in their respective periods. For the 26 remaining elements, the experimentally observed lowest-temperature crystal structure was used to select the primitive cell, with geometry inputs acquired from the Materials Project database. These primitive cells were used for all convergence tests, as well as inputs for enumeration for binary mixtures.

3.2. Binary Systems. Combinatorially, $\binom{26}{2} = 325$ binary mixtures of our selected TMs exist; from this we selected a subset of 80 according to several criteria. We wished our test-set to explore as many different electronic structure phenomena as possible while utilizing the minimum number of required calculations (both counting by number of binaries, and total number of structures). Our constraints, in order of priority, were then:

1. Each element is present in 6–7 binaries
2. Each element shares a binary with 2–3 elements from each period
3. No element shares a binary with two elements in the same period and in adjacent groups
4. Each element participates in at least one binary where it is known to form an ordered mixed phase, and one where it is known to form no ordered mixed phases
5. Each element shares a binary with at least one element from each space group

Constraints 1, 2, and 3 were strictly adhered to; constraint 4 was adhered to as a fortunate consequence of constraints 1–3, and constraint 5 was adhered to when possible (for elements like Mn with a unique space group for the primitive cell, due to constraints 1 and 2, only some elements could share a binary with Mn). The resulting binaries chosen, as well as a qualitative view of *all* possible binaries with experimental mixing behavior indicated, is shown in Figure 3.1.

For enumerating structures, if both elements in the binary shared a space group, the element with the lower atomic number was used to prototype the enumeration. If both elements did not share a space group, two enumerations were performed, one using each primitive cell. Enumeration was carried out by constructing all symmetrically-distinct supercells containing up to 6 atoms, and decorating those supercells with all possible symmetrically-distinct configurations. This scheme resulted in 137 unique configurations for all structures enumerated from one-atom primitive cells, and 63 unique configurations for all structures enumerated from two-atom supercells. If one or more known ordered phases formed in a binary, that ordered phase was also included in the enumeration, as well as any alternate decorations of the underlying supercell if the supercell contained 10 or fewer atoms (otherwise, only the exact, experimentally-observed structure was enumerated). If the ordered phase did not share a space group with either of the elemental primitive cells, this process was carried out separately as a third enumeration. In the case of Mn, owing to the relatively large size of the primitive cell (32 atoms), a body-centered-cubic primitive cell was used for enumeration of Mn binaries (though the true primitive cell was used as a chemical reference).

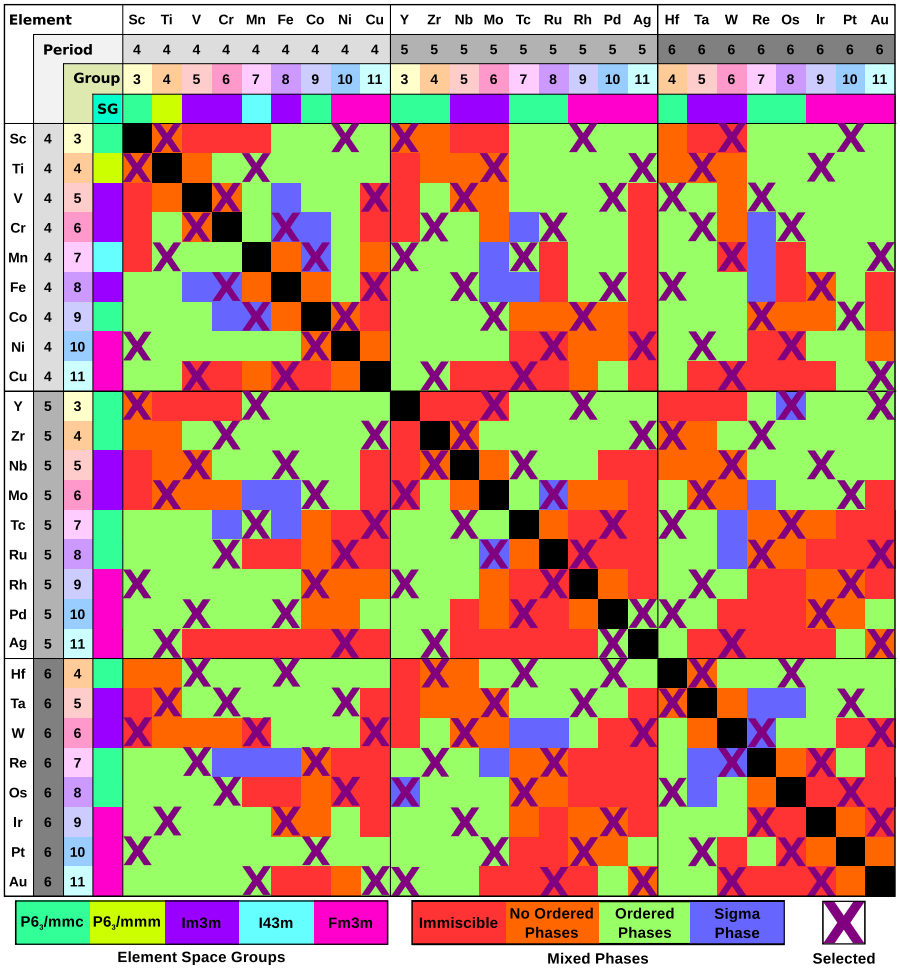

FIG. 3.1. *Table of transition element binaries, ordered by period then group, with the space group (SG) of the lowest-temperature stable solid listed. Binaries selected for inclusion in the V-DM/16 test set are indicated by purple X's. In each cell the experimentally-observed mixing behavior of the two alloys is given by color: red indicates a miscibility gap, orange indicates a random solid solution, green indicates the presence of ordered phases, while blue indicates that only a sigma (quasi-ordered) phase is formed.*

3.3. Experimental Data. Experimental data was derived from a variety of resources, though the authors have attempted to use as few *different* sources as possible to ensure regularity of the results. Owing to the large nature of this study, several databases were used, including Materials Project[25], the ASM Alloy Phase Diagram Database[68], and the Springer Materials[67] search tool. When multiple sources existed, the most modern experimental source was used; this choice is not a statement of distrust in older results, but only is meant to serve as an objective and simple selection criteria.

A summary of the experimental techniques utilized for each parameter is given in Table 3.1. As the individual measurements used in this study were taken from a wide variety of sources, a general guideline as to that measurement's precision is given, rather than a specific error estimate. For lattice constants and elastic coefficients, the

Parameter	Technique	Precision
Lattice Constant	Powder Diffraction	$< 0.1\%$[23]
Elastic Coeff.	Continuous Wave Ultrasonic Inferometry	$< 0.1\%$[14]
Formation Enthalpy	High Temperature Direct Reaction Calorimetry	± 10 meV[27]

TABLE 3.1

Table of experimental techniques used to acquire the various parameters used as a standard for the test-set. Precision estimates are taken from general analysis of the technique or of the instruments used, and not from individual experiments.

precision is expressed as $\frac{\Delta i}{i} \times 100\%$ where i is some property, where as for formation enthalpy a more direct estimate of Δi can be made.

Some consideration must be given in comparing experimental measurements taken at *finite* temperatures, versus DFT simulations performed at zero kelvin. Thermal expansion effects, as well as zero-point phonon energies, can modify lattice constants, elastic coefficients, and formation enthalpies. In the case of lattice constants, these corrections are smaller than 1% of the value and strictly decrease the lattice constant[37, 20], and so have been neglected. Determining elastic coefficient corrections requires pressure derivatives of the coefficients, which are lacking in experimental information[37], and so have also been neglected. Finally, the order of the corrections to the formation enthalpies are on the order of 10 meV[37, 26], which is within the precision of our experimental technique, and so can also be neglected.

4. High-Throughput Results. At the present moment, 33674 single-element convergence calculations and 53294 production binary calculations have been performed across 8 functionals. These calculations represent full convergence calculations for 23 of the 26 individual elements (Mn, Co, and Cu are still in-progress for several functionals), and approximately a quarter of the production runs for the binaries. Completion of all calculations, numbering approximately 100,000 individual VASP runs, is projected for December of 2016. Failures that could not be automatically recovered from have occurred in less than 1% of simulations, however, this still represents hundreds of calculations that must be inspected and restarted by-hand.

4.1. Single-Component Systems. Convergence statistics for the completed elements can be found in Tables 4.1 and 4.2, showing k-point grids and energy cutoffs resulting in energy differences ≤ 1 meV, respectively. Convergence simulations were primarily straightforward, however, in several cases (e.g., Cu, Zr, Co, Fe), the number of bands used had to be manually increased after several warnings and errors were encountered. Additionally, for Y, the GW potentials resulted in non-converging electronic self-consistency loops, and so were subsequently replaced by the non-GW versions of those potentials. In the majority of cases, the smallest k-point mesh density tested (corresponding to a length of 30 angstroms in the fully automatic scheme) already achieved 1 meV convergence, however, in a few cases (especially in periods 9 and 10), a significantly greater number of k-points had to be used. For energy cutoffs, 1.5 times the default value provided by VASP usually proved sufficient. Unsurprisingly, AM05, rTPSS, and M06-L required both the densest k-point meshes and the highest energy cutoffs; these convergence needs are explained by the finer-grained information required to correctly determine subsystem parameters and kinetic energy densities, respectively.

Element	PZ	PW91	BLYP	PBE	PBESol	AM05	rTPSS	M06-L
Ag	20×20×20	13×13×13	13×13×13	13×13×13	20×20×20	30×30×30	20×20×20	25×25×25
Au	13×13×13	13×13×13	13×13×13	13×13×13	13×13×13	33×33×33	13×13×13	25×25×25
Co	20×20×11	19×19×10	19×19×10	19×19×10	20×20×11	16×16×9	26×26×14	22×22×12
Cr	16×16×16	15×15×15	15×15×15	15×15×15	16×16×16	29×29×29	16×16×16	28×28×28
Hf	11×11×6	12×12×7	12×12×7	12×12×7	12×12×7	14×14×8	12×12×7	11×11×6
Ir	24×24×24	24×24×24	24×24×24	24×24×24	24×24×24	30×30×30	24×24×24	33×33×33
Mo	16×16×16	15×15×15	15×15×15	16×16×16	16×16×16	27×27×27	16×16×16	22×22×22
Nb	14×14×14	14×14×14	14×14×14	14×14×14	14×14×14	25×25×25	14×14×14	22×22×22
Ni	21×21×21	21×21×21	24×24×24	21×21×21	21×21×21	32×32×32	24×24×24	24×24×24
Os	15×15×8	15×15×8	15×15×8	16×16×9	16×16×9	29×29×16	16×16×9	24×24×13
Pd	23×23×23	23×23×23	19×19×19	23×23×23	23×23×23	30×30×30	23×23×23	28×28×28
Pt	21×21×21	20×20×20	20×20×20	20×20×20	21×21×21	22×22×22	21×21×21	31×31×31
Re	16×16×9	18×18×10	16×16×9	18×18×10	15×15×8	20×20×11	15×15×8	17×17×9
Rh	19×19×19	22×22×22	25×25×25	19×19×19	19×19×19	18×18×18	22×22×22	23×23×23
Ru	16×16×9	16×16×9	18×18×10	16×16×9	16×16×9	25×25×14	16×16×9	29×29×16
Sc	11×11×6	11×11×6	11×11×6	11×11×6	11×11×6	11×11×6	11×11×6	11×11×6
Ta	13×13×13	13×13×13	13×13×13	13×13×13	13×13×13	23×23×23	13×13×13	16×16×16
Tc	17×17×9	17×17×9	17×17×9	13×13×7	17×17×9	22×22×12	13×13×7	19×19×10
Ti	13×13×7	12×12×7	12×12×7	12×12×7	12×12×7	12×12×7	12×12×7	12×12×7
V	16×16×16	15×15×15	16×16×16	15×15×15	16×16×16	22×22×22	15×15×15	20×20×20
W	15×15×15	15×15×15	15×15×15	15×15×15	15×15×15	21×21×21	15×15×15	21×21×21
Y	14×14×8	14×14×8	14×14×8	14×14×8	14×14×8	14×14×8	14×14×8	14×14×8
Zr	12×12×7	11×11×6	11×11×6	11×11×6	11×11×6	14×14×8	11×11×6	14×14×8

TABLE 4.1
Converged values of the k-point mesh for 23 elements, across 8 functionals.

4.2. Binary Systems. For binary systems, in addition to the difficulties encountered in single elements, a new challenge was introduced in the form of highly unstable structures. Since the enumeration technique we employed relied solely on symmetry (to produce unique structures), some generated structures proved to be unstable. These structures would, through relaxation, reconfigure into either a new structure or another existing structure. Relaxation of this sort result in DFT energies that no longer represent the original structure; this was handled by discarding any offending configurations. Specifically, by measuring deformations both of the unit cell and of the basis sites from the start to the end of a relaxation, a "mapping score" was generated. All structures in a binary family were mapped onto all other structures, and if a configuration mapped better (scored lower) onto any structure other than its starting structure, the configuration was removed. These relaxations tended to occur only in certain binaries, though even in these situations no more than 10% of a binary's enumerated configurations were discarded.

Although much of the calculation work remains to be finished, some interesting trends have already begun to emerge. In binaries where no mixed phases were experimentally observed, but the two components were mutually soluble, the PBE functional usually predicted a series of shallow (< -10 meV formation enthalpy) stable ordered phases. However, in cases where not only were there a lack of ordered phases, but the two elements were experimentally immiscible, no stable ordered were predicted (> 0 meV formation enthalpy). In binaries known to form one or more ordered phases, most of the configurations calculated were negative in enthalpy, though usually not as negative as the experimentally observed ordered phases. Additionally, alloys with two elements with different space groups had a higher percentage of structures that relaxed onto different structures, than alloys that shared a space group (again, in the PBE functional).

Element	PZ	PW91	BLYP	PBE	PBESol	AM05	rTPSS	M06-L
Ag	535	535	535	590	535	535	580	640
Au	470	505	515	495	465	465	515	525
Co	635	635	635	635	635	635	685	700
Cr	575	575	575	575	575	575	655	635
Hf	425	425	425	425	425	425	425	425
Ir	535	535	535	535	535	535	535	535
Mo	525	525	525	525	525	525	525	525
Nb	535	535	535	535	535	535	535	600
Ni	585	585	585	585	595	585	665	710
Os	535	535	535	510	535	535	535	535
Pd	535	570	535	545	535	535	580	605
Pt	515	505	505	560	505	505	560	505
Re	505	505	505	505	475	950	505	585
Rh	525	525	525	525	525	525	580	525
Ru	525	525	525	525	525	525	525	525
Sc	565	565	565	565	565	565	565	565
Ta	435	435	435	435	435	435	435	435
Tc	525	525	525	525	525	525	525	530
Ti	575	575	575	575	575	575	575	575
V	575	575	575	575	575	575	595	605
W	485	485	485	485	485	485	485	485
Y	515	585	515	515	595	515	515	640
Zr	515	515	515	515	515	560	645	625

TABLE 4.2
Converged values of the energy cutoff, in eV, for 23 elements, across 8 functionals.

5. Discussion of Design Choices. Our study places a special emphasis on our third experimental comparison — the formation energy. This parameter plays a critical role in determining whether a given approximation of DFT (i.e., a specific functional) successfully captures the relevant physics present in a material. While present functionals (e.g., PBESol and AM05) can provide excellent results for lattice constants, in materials with active d-electrons (e.g., TMs) the predicted formation energies can be off by a factor of two or more. These results are surprising, but can be understood considering the differences in how formation energy and lattice constant are calculated. DFT, as presently implemented, provides only a *direct* measure of formation energies, specifically, the only density functional we "know" how to evaluate is one of energy. All other properties, including lattice constant, are second (or higher) order extrapolations from the energy, usually found via minimizing the energy with respect to some perturbation in said property. Therefore, for correct evaluation of the lattice constant and other secondary properties (given proper conference settings[42]), the functional only needs to give a good representation of the derivative of the energy near equilibrium.

A small number of previous studies have taken inspiration from quantum chemistry and utilized the cohesive energy as an experimental measure of the accuracy of a functional[26, 37]. While attractive due to the ease of calculation as well as the availability of experimental data for TMs, this approach overlooks important differences between how quantum chemistry and solid-state physics handle DFT. Most

importantly, solid-state, *periodic* codes are designed to function best in bulk materials with a well-defined unit cell. When calculating surfaces, defects, and isolated atoms, so-called "image problem" appears if a sufficiently large simulation cell size is not employed, as the isolated feature can begin to see copies of itself across a periodic boundary[19]. Additionally, most functionals developed for periodic materials are based on some solution of the *uniform* electron gas (PBESol and AM05 being two obvious exceptions), whereas an isolated atom more closely represents a localized system of mostly vacuum. Finally, in systems with higher electron densities (i.e., in a solid, where neighboring atoms can contribute electron density), the fraction of the coulomb energy term made up by self-interaction energy is smaller than in a geometry mostly constructed of vacuum. All of these factors feed into the larger problem that periodic codes are not *designed* for isolated atoms, and while accurate results can be obtained on such systems, such a unique application cannot be considered to be representative of the overall predictive power and accuracy of periodic codes when applied to true periodic systems.

6. Conclusions. While DFT has become a powerful, expansive, and insightful tool in the past decade, caution must be taken in the interpretation of results. Specifically, in the case of *d*-electron materials, DFT is known to suffer several difficulties owing to the mixed localized-and-delocalized electronic behavior present. In light of several high-profile failures of DFT to describe mixed TM binaries, we have developed a detailed test-set suited to probe a technique's accuracy in predicting TM mixture behaviors. Our test set utilizes binary formation enthalpies, an often-neglected characteristic, to provide a more direct measure of the quality of approximations made by a given functional. We have applied our test set to eight of the most commonly used functionals, and compared their performance. No one functional provides an accurate description of all members of our test set, and the differences in their failures highlight the wide variety of unique properties present in TM alloys. The authors urge that forward progress using DFT as a *predictive* rather than a *descriptive* tool must be approached with care and rigor.

REFERENCES

[1] L. A. Agapito, S. Curtarolo, and M. Buongiorno Nardelli, *Reformulation of* DFT + *u as a pseudohybrid hubbard density functional for accelerated materials discovery*, Phys. Rev. X, 5 (2015), p. 011006.
[2] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein, *First-principles calculations of the electronic structure and spectra of strongly correlated systems: the lda + u method*, Journal of Physics: Condensed Matter, 9 (1997), p. 767.
[3] R. Armiento and A. E. Mattsson, *Functional designed to include surface effects in self-consistent density functional theory*, Phys. Rev. B, 72 (2005), p. 085108.
[4] A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A, 38 (1988), pp. 3098–3100.
[5] S. Bhattacharya and G. K. H. Madsen, *High-throughput exploration of alloying as design strategy for thermoelectrics*, Phys. Rev. B, 92 (2015), p. 085205.
[6] P. E. Blöchl, *Projector augmented-wave method*, Phys. Rev. B, 50 (1994), pp. 17953–17979.
[7] P. E. Blöchl, O. Jepsen, and O. K. Andersen, *Improved tetrahedron method for brillouin-zone integrations*, Phys. Rev. B, 49 (1994), pp. 16223–16233.
[8] E. Bylaska, K. Tsemekhman, N. Govind, and M. Valiev, *Large-Scale Plane-Wave-Based Density Functional Theory: Formalism, Parallelization, and Applications*, John Wiley & Sons, Inc., 2011, pp. 77–116.
[9] J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, *Finding unprecedentedly low-thermal-conductivity half-heusler semiconductors via high-throughput materials modeling*, Phys. Rev. X, 4 (2014), p. 011019.

[10] M. Cococcioni and S. de Gironcoli, *Linear response approach to the calculation of the effective interaction parameters in the* LDA+U *method*, Phys. Rev. B, 71 (2005), p. 035105.
[11] C. J. Cramer and D. G. Truhlar, *Density functional theory for transition metals and transition metal chemistry*, Phys. Chem. Chem. Phys., 11 (2009), pp. 10757–10816.
[12] S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, *The high-throughput highway to computational materials design*, Nature Materials, 12 (2013), pp. 191–201.
[13] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, *Aflowlib.org: A distributed materials properties repository from high-throughput ab initio calculations*, Computational Materials Science, 58 (2012), pp. 227 – 235.
[14] J. P. Day, P. S. Ho, and A. L. Ruoff, *A cw interferometer for elastic constant measurements*, Review of Scientific Instruments, 44 (1973), pp. 478–481.
[15] E. Decolvenaere, M. J. Gordon, and A. Van der Ven, *Testing predictions from density functional theory at finite temperatures: β_2-like ground states in co-pt*, Phys. Rev. B, 92 (2015), p. 085119.
[16] A. V. der Ven, J. Thomas, Q. Xu, and J. Bhattacharya, *Linking the electronic structure of solids to their thermodynamic and kinetic properties*, Mathematics and Computers in Simulation, 80 (2010), pp. 1393 – 1410.
[17] F. Furche and J. P. Perdew, *The performance of semilocal and hybrid density functionals in 3d transition-metal chemistry*, The Journal of Chemical Physics, 124 (2006).
[18] E. Furet, K. Costuas, P. Rabiller, , and O. Maury, *On the sensitivity of f electrons to their chemical environment*, Journal of the American Chemical Society, 130 (2008), pp. 2180–2183. PMID: 18225894.
[19] M. J. Gillan, *Calculation of the vacancy formation energy in aluminium*, Journal of Physics: Condensed Matter, 1 (1989), p. 689.
[20] P. Haas, F. Tran, and P. Blaha, *Calculation of the lattice constant of solids with semilocal functionals*, Phys. Rev. B, 79 (2009), p. 085104.
[21] G. L. W. Hart, S. Curtarolo, T. B. Massalski, and O. Levy, *Comprehensive search for new phases and compounds in binary alloy systems based on platinum-group metals, using a computational first-principles approach*, Phys. Rev. X, 3 (2013), p. 041035.
[22] J. N. Harvey, *On the accuracy of density functional theory in transition metal chemistry*, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem., 102 (2006), pp. 203–226.
[23] F. H. Herbstein, *How precise are measurements of unit-cell dimensions from single crystals?*, Acta Crystallographica Section B, 56 (2000), pp. 547–557.
[24] J. Heyd and G. E. Scuseria, *Efficient hybrid density functional calculations in solids: Assessment of the heydscuseriaernzerhof screened coulomb hybrid functional*, The Journal of Chemical Physics, 121 (2004), pp. 1187–1192.
[25] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. a. Persson, *The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation*, APL Materials, 1 (2013), p. 011002.
[26] P. Janthon, S. A. Luo, S. M. Kozlov, F. Vies, J. Limtrakul, D. G. Truhlar, and F. Illas, *Bulk properties of transition metals: A challenge for the design of universal density functionals*, Journal of Chemical Theory and Computation, 10 (2014), pp. 3832–3839. PMID: 26588528.
[27] O. Kleppa and L. Topor, *A new calorimeter for temperatures above 1400 k*, Thermochimica Acta, 139 (1989), pp. 291 – 297.
[28] W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev., 140 (1965), pp. A1133–A1138.
[29] G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, Phys. Rev. B, 54 (1996), pp. 11169–11186.
[30] G. Kresse and J. Furthmller, *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set*, Computational Materials Science, 6 (1996), pp. 15 – 50.
[31] G. Kresse and J. Hafner, *Ab initio molecular dynamics for liquid metals*, Phys. Rev. B, 47 (1993), pp. 558–561.
[32] G. Kresse and J. Hafner, *Ab initio molecular-dynamics simulation of the liquid-metal˘amorphous-semiconductor transition in germanium*, Phys. Rev. B, 49 (1994), pp. 14251–14269.
[33] G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*, Phys. Rev. B, 59 (1999), pp. 1758–1775.

[34] H. J. Kulik, M. Cococcioni, D. A. Scherlis, and N. Marzari, *Density functional theory in transition-metal chemistry: A self-consistent hubbard u approach*, Phys. Rev. Lett., 97 (2006), p. 103001.

[35] C. Lee, W. Yang, and R. G. Parr, *Development of the colle-salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, 37 (1988), pp. 785–789.

[36] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. Di Marco, C. Draxl, M. Dułak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Grånäs, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Hamann, P. J. Hasnip, N. A. W. Holzwarth, D. Iuşan, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Küçükbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordström, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. Van Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang, and S. Cottenier, *Reproducibility in density functional theory calculations of solids*, Science, 351 (2016).

[37] K. Lejaeghere, V. V. Speybroeck, G. V. Oost, and S. Cottenier, *Error estimates for solid-state density-functional theory predictions: An overview by means of the ground-state elemental crystals*, Critical Reviews in Solid State and Materials Sciences, 39 (2014), pp. 1–24.

[38] O. Levy, G. L. W. Hart, and S. Curtarolo, *Structure maps for hcp metals from first-principles calculations*, Phys. Rev. B, 81 (2010), p. 174106.

[39] S. B. Maisel, M. Höfler, and S. Müller, *Configurationally exhaustive first-principles study of a quaternary superalloy with a vast configuration space*, Phys. Rev. B, 94 (2016), p. 014116.

[40] A. E. Mattsson and R. Armiento, *Implementing and testing the am05 spin density functional*, Phys. Rev. B, 79 (2009), p. 155101.

[41] A. E. Mattsson, R. Armiento, J. Paier, G. Kresse, J. M. Wills, and T. R. Mattsson, *The am05 density functional applied to solids*, The Journal of Chemical Physics, 128 (2008).

[42] A. E. Mattsson, P. A. Schultz, M. P. Desjarlais, T. R. Mattsson, and K. Leung, *Designing meaningful density functional theory calculations in materials sciencea primer*, Modelling and Simulation in Materials Science and Engineering, 13 (2005), p. R1.

[43] A. E. Mattsson and J. M. Wills, *Density functional theory for d- and f-electron materials and compounds*, International Journal of Quantum Chemistry, 116 (2016), pp. 834–846.

[44] M. Methfessel and A. T. Paxton, *High-precision sampling for brillouin-zone integration in metals*, Phys. Rev. B, 40 (1989), pp. 3616–3621.

[45] H. J. Monkhorst and J. D. Pack, *Special points for brillouin-zone integrations*, Phys. Rev. B, 13 (1976), pp. 5188–5192.

[46] N. Nakajima, T. Koide, T. Shidara, H. Miyauchi, H. Fukutani, A. Fujimori, K. Iio, T. Katayama, M. Nývlt, and Y. Suzuki, *Perpendicular magnetic anisotropy caused by interfacial hybridization via enhanced orbital moment in* Co/Pt *multilayers: Magnetic circular x-ray dichroism study*, Phys. Rev. Lett., 81 (1998), pp. 5229–5232.

[47] L. Nordström, J. M. Wills, P. H. Andersson, P. Söderlind, and O. Eriksson, *Spin-orbit coupling in the actinide elements: A critical evaluation of theoretical equilibrium volumes*, Phys. Rev. B, 63 (2000), p. 035103.

[48] V. Ozoliņš, C. Wolverton, and A. Zunger, *Cu-au, ag-au, cu-ag, and ni-au intermetallics: First-principles study of temperature-composition phase diagrams and structures*, Phys. Rev. B, 57 (1998), pp. 6427–6443.

[49] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*, Phys. Rev. Lett., 77 (1996), pp. 3865–3868.

[50] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple [phys. rev. lett. 77, 3865 (1996)]*, Phys. Rev. Lett., 78 (1997), pp. 1396–1396.

[51] J. P. Perdew, J. A. Chevary, V. S. H., K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation*, Phys. Rev. B, 46 (1992), pp. 6671–6687.

[52] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Erratum: Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation*, Phys. Rev. B, 48 (1993), pp. 4978–4978.

[53] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, and J. Sun, *Workhorse semilocal density functional for condensed matter physics and quantum chemistry*, Phys. Rev. Lett., 103 (2009), p. 026403.

[54] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, *Restoring the density-gradient expansion for exchange in solids and surfaces*, Phys. Rev. Lett., 100 (2008), p. 136406.

[55] J. P. Perdew and K. Schmidt, *Jacobs ladder of density functional approximations for the exchange-correlation energy*, AIP Conference Proceedings, 577 (2001), pp. 1–20.

[56] J. P. Perdew and A. Zunger, *Self-interaction correction to density-functional approximations for many-electron systems*, Phys. Rev. B, 23 (1981), pp. 5048–5079.

[57] B. Puchala, J. C. Thomas, J. Goiri, M. Radin, N. S. H. GUNDA, A. R. Natarajan, and L. Decolvenaere, *Casmcode: v0.2.0*, Aug. 2016.

[58] B. Puchala and A. Van der Ven, *Thermodynamics of the zr-o system from first-principles calculations*, Phys. Rev. B, 88 (2013), p. 094108.

[59] P. Pyykko, *Relativistic effects in structural chemistry*, Chemical Reviews, 88 (1988), pp. 563–594.

[60] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, *Materials design and discovery with high-throughput density functional theory: The open quantum materials database (oqmd)*, JOM, 65 (2013), pp. 1501–1509.

[61] S. K. Sidorov and A. V. Doroshenko, *On the magnetic structure of some alloys of transition metals*, physica status solidi (b), 16 (1966), pp. 737–744.

[62] A. Sorouri, W. M. C. Foulkes, and N. D. M. Hine, *Accurate and efficient method for the treatment of exchange in a plane-wave basis*, The Journal of Chemical Physics, 124 (2006).

[63] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Climbing the density functional ladder: Nonempirical meta˘generalized gradient approximation designed for molecules and solids*, Phys. Rev. Lett., 91 (2003), p. 146401.

[64] J. C. Thomas and A. V. d. Ven, *Finite-temperature properties of strongly anharmonic and mechanically unstable crystal phases from first principles*, Phys. Rev. B, 88 (2013), p. 214111.

[65] M. C. Troparevsky, J. R. Morris, P. R. C. Kent, A. R. Lupini, and G. M. Stocks, *Criteria for predicting the formation of single-phase high-entropy alloys*, Phys. Rev. X, 5 (2015), p. 011041.

[66] J. H. van Vleck, *On the anisotropy of cubic ferromagnetic crystals*, Phys. Rev., 52 (1937), pp. 1178–1198.

[67] P. Villars, M. Berndt, K. Brandenburg, K. Cenzual, J. Daams, F. Hulliger, T. Massalski, H. Okamoto, K. Osaki, A. Prince, H. Putz, and S. Iwata, *The pauling file, binaries edition*, Journal of Alloys and Compounds, 367 (2004), pp. 293 – 297.

[68] P. Villars, H. Okamoto, and K. Cenzual, *Asm alloy phase diagram database*.

[69] S. H. Vosko, L. Wilk, and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Canadian Journal of Physics, 58 (1980), pp. 1200–1211.

[70] Y. Zhao and D. G. Truhlar, *A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions*, The Journal of Chemical Physics, 125 (2006).

[71] Y. Zhao and D. G. Truhlar, *The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four m06-class functionals and 12 other functionals*, Theoretical Chemistry Accounts, 120 (2008), pp. 215–241.