

Discovering hierarchies among intermetallic crystal structures

Sanjeev Krishna Kolli , Anirudh Raju Natarajan , John C. Thomas, Tresa M. Pollock, and Anton Van der Ven
Materials Department, University of California, Santa Barbara, California 93106, USA



(Received 20 July 2020; revised 16 October 2020; accepted 22 October 2020; published 10 November 2020)

The crystal structure of a compound plays an important role in determining its properties. Here we analyze over 4000 intermetallic compounds, and we identify a hierarchical relationship between their crystal structures. By considering each intermetallic compound as a particular chemical ordering over the sites of a higher symmetry crystal structure, we determine that most intermetallic compounds can be derived from a small number of parent crystal structures. While many compounds are chemical orderings over the sites of simple crystal structures such as body-centered-cubic (bcc) and face-centered-cubic (fcc) structures, the majority map onto more complex parent crystal structures. Surprisingly, many intermetallic compounds map onto parent crystal structures that differ from those of their elemental constituents. We find that the occurrence of several of the more complex parent crystal structures, such as the Laves phases, can be understood in terms of simple descriptors such as the ratio of atomic radii and electronegativity differences.

DOI: [10.1103/PhysRevMaterials.4.113604](https://doi.org/10.1103/PhysRevMaterials.4.113604)

I. INTRODUCTION

A materials designer is often eager to identify the chemical composition space in which a desired crystal structure is stable [1–3]. This requires the calculation of the phase diagrams of many candidate alloy systems by minimizing over the free energies of all crystal structures competing for stability [4]. While the free energy of a phase can be calculated with remarkable accuracy using first-principles statistical mechanics approaches [5–15], these calculations are expensive and the number of candidate crystal structures is vast. A major challenge, therefore, is to identify a small list of likely crystal structures to consider in a first-principles analysis of phase stability.

The task of establishing phase stability in a multicomponent composition space is made easier upon recognizing a natural hierarchy among crystal structures: many compound crystal structures are related to a more manageable number of higher symmetry *parent* crystal structures through a group/subgroup symmetry relation. A solid solution of a face-centered-cubic (fcc) crystal structure, for example, may undergo an ordering transition at low temperatures to form an intermetallic compound whereby different chemical species adopt a periodic arrangement over the sites of the fcc lattice. While an ordered phase is usually treated as a distinct crystal structure in most crystal structure databases, it can also be viewed as a derivative of a higher symmetry parent crystal structure. This has been the general philosophy of alloy theorists [5,16–22], who have traditionally viewed phase stability as a competition between a variety of high symmetry parent crystal structures that each can form a solid solution along with several derivative ordered phases at particular compositions and temperatures.

There are several advantages to analyzing phase stability within a framework that recognizes the hierarchical

relationship between parent crystal structures and their derivative ordered phases. Given a parent crystal structure, there are systematic approaches with which to rapidly enumerate derivative orderings [23]. Furthermore, well-established statistical mechanics schemes based on the cluster expansion approach exist to treat the configurational contributions to the free energy of parent and derivative crystal structures [5,10,15,24]. A recognition that a particular compound may be a derivative ordering of a higher symmetry parent phase also provides insights about the nature of possible phase transformations and the likelihood that a compound may coexist coherently within a disordered matrix phase. In this context, it is often advantageous to measure the degree of symmetry lowering of a compound crystal structure relative to a simpler parent crystal structure with symmetry-adapted order parameters that can be generated algorithmically [25–27] and that can be used to construct generalized free-energy descriptions [8,28].

While the organization of compound crystal structures within a hierarchical framework is appealing for many purposes, the extent with which such hierarchical relationships exist has not yet been established. Furthermore, if such hierarchical relationships do exist among the majority of compounds, it is not evident which parent crystal structures are most common. Among those that are the most common, which ones host the largest number of distinct ordered phases and which ones are truly unique compound crystal structures that host only one ordering? Are there strong correlations between the occurrence of different parent crystal structures within the same alloy system? If so, having prior knowledge about one parent crystal structure can then be used to anticipate the occurrence of other parent crystal structures in a particular alloy system.

In this work, we set out to answer these questions. We organize the crystal structures of multicomponent compounds

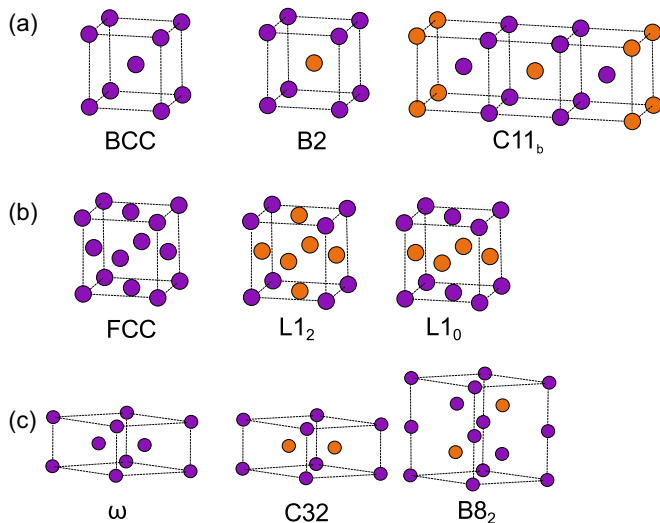


FIG. 1. Illustrations of basic (parent) crystal structures and decorations that occur on them that are commonly seen intermetallics.

into parent crystal structures and derivative ordered structures using a robust mapping algorithm to compare the similarity between crystal structures. We demonstrate the approach for binary intermetallic structures, and we show that many intermetallic compounds map onto simple parent crystal structures such as body-centered-cubic (bcc) and fcc, but that an even larger number map onto more complex parent crystal structures, including those that belong to the family of topologically close-packed phases. A large number of intermetallic compounds are found to be orderings on parent crystal structures that differ from those of the constituent elements in their pure state. We also find strong correlations between the occurrence of pairs of parent crystal structures in the same binary alloy system. The results of this study suggest the utility of a hierarchical cataloguing of crystal structures for first-principles databases [29–32] and for the rapid prediction of phase stability in uncharted composition spaces.

II. ESTABLISHING A HIERARCHY OF CRYSTAL STRUCTURES

Periodic crystals in multicomponent composition spaces can be complex. It is nevertheless possible to organize compound crystal structures hierarchically by viewing each compound as a particular decoration of different chemical species over the sites of a high symmetry *parent* crystal structure. An important example is the B2 crystal structure (as designated within the Strukturbericht classification scheme) adopted by CsCl and NiAl. Figure 1(a) shows that B2 is a simple chemical ordering over the sites of a bcc parent crystal structure. Similarly, the common L1₂ and L1₀ crystal structures formed by Ni₃Al and TiAl, respectively, correspond to periodic orderings of two chemical species over the sites of an fcc parent crystal structure [Fig. 1(b)]. Other intermetallic compounds can be mapped onto more complex parent crystal structures. Figure 1(c), for example, shows that C32 and B8₂ correspond to chemical decorations of an ω parent crystal structure.

The identification of the parent crystal structure of a particular compound is not always trivial. This is especially true when the chemical ordering of the compound has a symmetry that is lower than that of the undecorated parent crystal structure. For example, while the ordering of L1₂ preserves the cubic symmetry of the underlying fcc lattice, the layered ordering of L1₀ does not, and compounds adopting this structure often undergo a slight tetragonal distortion of their unit-cell vectors. Eliminating the distinction between Ti and Al in L1₀ TiAl, for instance, produces an undecorated crystal that is not perfect fcc, but rather a face-centered-tetragonal crystal. Nevertheless, the tetragonal distortion is sufficiently small in most L1₀ forming compounds that they can still be considered as having an fcc parent crystal structure. Other common crystals that can be viewed as a symmetry-breaking ordering of a parent crystal structure include C11_b, which has a lower symmetry than undecorated bcc [Fig. 1(a)], and B8₂, which lowers the symmetry of undecorated ω [Fig. 1(c)].

The examples of L1₀, C11_b, and B8₂ demonstrate the need for a robust method of establishing similarity between any pair of crystal structures when attempting to identify the parent crystal structure of a compound. In comparing two crystal structures, it is necessary to determine the similarity between (i) the lattice vectors and (ii) the atomic positions of both structures. The similarity between two crystal structures can be measured with a mapping score that is zero when the two crystals are identical, and large when the two crystals are qualitatively very different. The Supplemental Material [56] describes a robust approach to measure the dissimilarity of two crystal structures. The approach differs from other structure comparison algorithms that rely on comparing crystal invariants, pair distribution functions, or symmetry features such as space group and Wyckoff positions [33–37]. Instead, the approach is similar in spirit to those that measure geometric similarity by identifying an affine mapping between a pair of crystal structures [38,39]. It measures crystal similarity with a cost function that depends on the symmetry-breaking strains and atomic displacements along different paths that convert one crystal structure into the other. In this sense, care is taken to remove strains and distortions that preserve the symmetry of the reference structure being mapped onto. For example, two hexagonal-close-packed (hcp) crystals may have very different *c/a* ratios and, therefore, require a large strain to convert one into the other. Nevertheless, this strain preserves the symmetry of the hcp crystal and should not contribute to any metric that is used to determine if two crystal structures are qualitatively identical. The mapping of one crystal onto the other relies on a lattice mapping algorithm similar to that described by Trinkle *et al.* [40] and a basis mapping that relies on the Hungarian algorithm [41,42].

III. HIERARCHY AMONG INTERMETALLIC CRYSTAL STRUCTURES

The organization of compound crystal structures into *parent* crystals and derivative *ordered* structures can be performed for any class of compounds, including intermetallics, oxides, sulfides, nitrides, etc. Here we limit the analysis to binary intermetallic compounds. To this end, we collected elemental and intermetallic crystal structures from the Materials

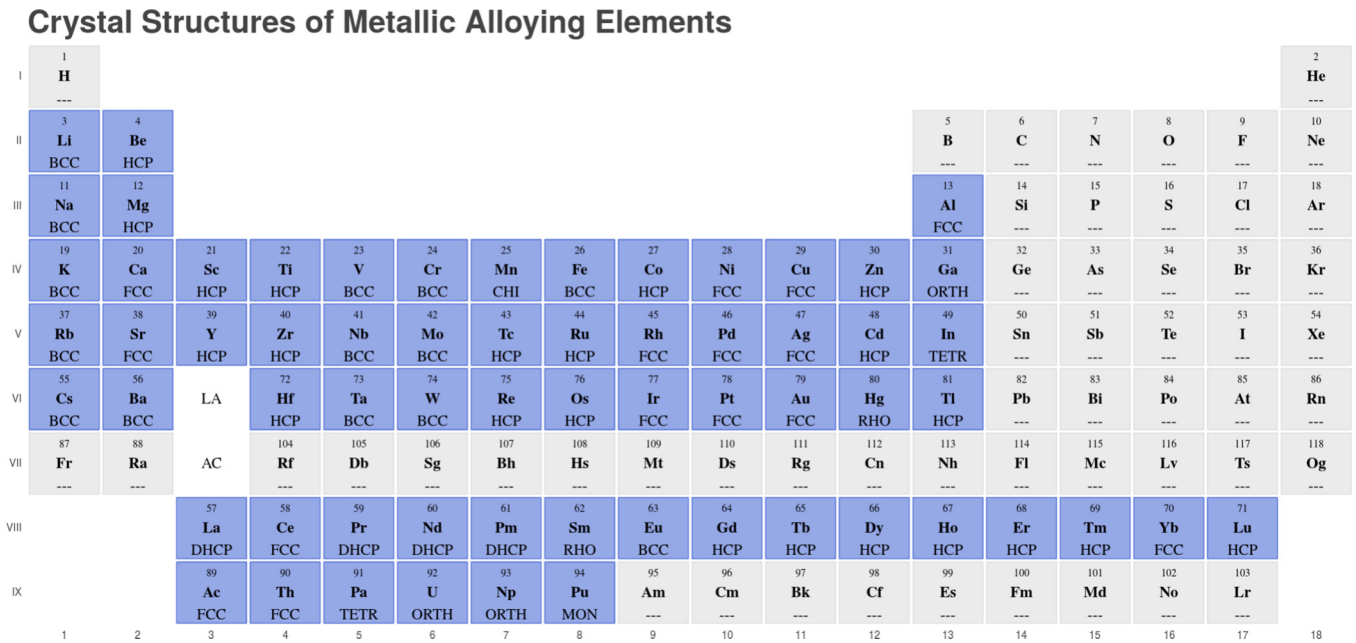


FIG. 2. The crystal structures of elemental compounds. Elements marked with no crystal structure were not considered as a potential alloying element in this study.

Project [31] and the International Crystallography Structure Database (ICSD) [43]. Only single-component and binary crystal structures containing the metallic elements highlighted in blue in Fig. 2 and containing fewer than 60 atoms in their unit cells were considered. The structures extracted from the Materials Project were restricted to those with an ICSD identification number. Structures with partial occupancies and up to two of the elements shown in Fig. 2 were collected from the ICSD. Structures with partial occupancies were converted to stoichiometric compounds by filling each partially occupied site by the majority element of that site. A total of 4335 elements and compounds were collected in this manner. It is important to recognize that crystal structures from the ICSD database are ones that have been observed experimentally. However, not all of these crystal structures correspond to equilibrium phases, and our analysis is therefore not necessarily reflective of trends displayed by thermodynamically stable phases. Other biases are also likely present in this dataset. For example, complex crystal structures with large unit cells tend to be more difficult to refine than simpler ones, and may therefore be underreported. The relative frequencies and crystal structure rankings extracted from this dataset should consequently only be viewed as a qualitative indication of the likelihood of encountering those phases in a particular alloy.

The mapping algorithm described in the Supplemental Material [56] was applied to all 4335 elements and compounds, resulting in the identification of 455 unique crystal structures. This list still included crystal structures that can be viewed as a particular chemical ordering over a higher symmetry parent crystal structure. To identify unique parent crystal structures, each of the 455 unique crystal structures was undecorated and mapped on the remaining set of entries. If a particular crystal structure was able to map onto a simpler structure having the same or higher symmetry with a mapping score below a prede-

fined threshold, it was considered a derivative ordered phase, while the structure it mapped onto was made a candidate for a parent crystal structure. This led to the identification of 267 unique parent crystal structures. All parent crystal structures and their derivative ordered structures have been deposited in an interactive online database [44].

A. The top 20 parent crystal structures

Figure 3 shows a histogram for the top 20 most common parent crystal structures. The horizontal bar for bcc, for example, represents the fraction of binary compounds that are a particular ordering on bcc. The bar also includes all the elements highlighted in blue in Fig. 2 that have been reported as bcc. When assigning a particular compound crystal to a higher symmetry parent crystal, we allowed for the presence of vacancies and simple clusters of atoms such as dumbbells on a subset of sites of the parent crystal. The fraction of compounds that contain vacancies when mapped onto a particular parent crystal structure are shown in purple. Crystallographic information about each of the top 20 parent crystal structures is listed in Table I.

The top 20 parent crystal structures account for almost 75% of the 4335 elements and intermetallic compounds analyzed in this study. Of the remaining 247 parent crystal structures, 158 only have one or two compounds that map onto them. Figure 3 shows that a large fraction of intermetallic compounds are orderings on bcc and fcc. These constitute approximately 30% of all the compounds considered in this study. Perhaps a surprising outcome is the high frequency of complex parent crystal structures such as CaCu_5 , Laves C15, and ω . Each one of these parent crystals appears with a higher frequency than hcp-derived orderings.

TABLE I. The primitive cell size, number of distinct crystallographic sites, layering patterns, and number of chemical orderings of the top 20 most common parent crystal structures. The notation used to describe each two-dimensional layer is illustrated in Figs. S4, S6, and S7 in the Supplemental Material [56].

Parent crystal	No. of atoms in primitive cell	No. of distinct crystallographic sites	Layering pattern	No. of distinct orderings
bcc	1	1		44
fcc	1	1	ABC	32
CaCu ₅	6	3	A α	5
Laves C15	6	2	α [acb] β [bac] γ [cba]	3
ω	3	2	aH _a	8
hcp	2	1	ABAB	14
Laves C14	12	3	α [acb] β [bca]	6
AlTh	4	2	AlTh	1
A15	8	2		2
Be ₃ Nb	12	5	A α [abc] γ A β [bca] α A γ [cab] β	1
CeCu ₂	6	2	CeCu ₂	1
NiY	8	2	NiY	1
Yb ₆ Fe ₂₃	29	29		1
LuBe ₁₃	28	12		1
Cementite	16	3	Cementite	1
Al ₈ Cr ₅	26	9		12
dhcp	4	2	ABAC	7
Al ₂ Cu	6	2	S θ_a S θ_b	1
Co ₂ Si	12	3	Co ₂ Si	4
σ	30	5	$\theta_a\sigma_0\theta_a\sigma_\pi$	4

Table I summarizes crystallographic information about each of the top 20 parent crystals. The second column lists the number of atoms per primitive unit cell. Only bcc and fcc have simple parent crystal structures with one atom in their primitive unit cells. The majority of other parent crystal structures are much more complex and have large unit cells. Nevertheless, 15 of the top 20 parent crystals can be described

as a particular stacking sequence of simpler two-dimensional motifs. The Supplemental Material [56] describes the two-dimensional building blocks and their naming convention. The stacking sequences of the layered parent crystals are listed in the third column of Table I.

While our analysis has yielded 267 distinct parent crystal structures, the fourth column of Table I shows that only a few among the top 20 host multiple chemical orderings. The parent crystal structures that support more than one chemical ordering also tend to occur with a higher frequency. Furthermore, Table I shows that the simpler parent crystal structures such as bcc, fcc, and hcp have the largest number of chemical orderings. Figure 4 breaks down the histogram bars of the top six parent crystal structures of Fig. 3 by the frequency with which different orderings occur. The bar for bcc in Fig. 4 shows that B2 is the most common ordering on bcc. It is in fact the most common derivative ordered structure among all the compounds considered in this work. The L1₂ ordering dominates on fcc, while other well-known fcc orderings such as L1₀ occur much less frequently. The ω parent crystal structure also hosts several different orderings. Two of these have already been discussed and are shown in Fig. 1(c). Overall, most of the ordered compounds occur at simple stoichiometric compositions such as 1/2, 1/3, 1/4, and 1/6.

An interesting result that emerges from our analysis is that some intermetallic compounds can tolerate high concentrations of vacancies. B2 NiAl, for example, accommodates an excess of Al by introducing vacancies on its Ni sublattice [45]. These vacancies, which can reach concentrations as high as 15%, tend to order over the Ni-sublattice sites at low temperatures, thereby forming lower symmetry intermetallic compounds [46]. The D8₈ compound was found to map onto the ω parent crystal structure if one-third of the sites in every triangular layer are vacant. The histograms of Figs. 3 and 4

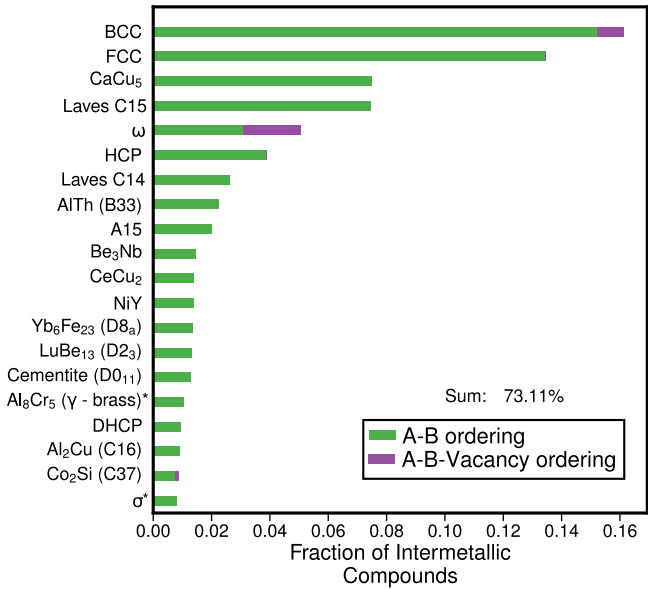


FIG. 3. The distribution of compounds across the top 20 most common parent crystal structures for binary metallic alloys. The purple represents ordered phases that contain vacancies. Parent crystals marked with a star have a large portion of compounds with partial occupancies.

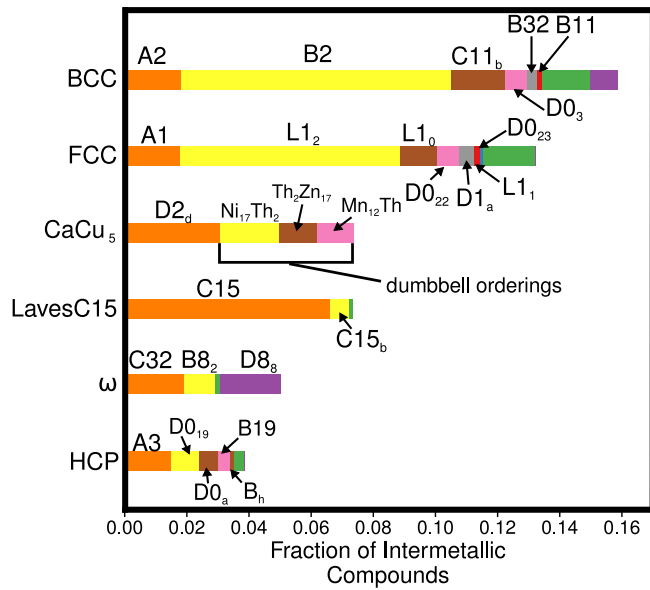


FIG. 4. The breakdown of ordered compounds among the top six most common parent crystal structures for binary metallic alloys. Each ordering is labeled by its Strukturbericht prototype name. Orderings that do not have a Strukturbericht prototype name are pooled together and represented with green blocks. The purple blocks represent orderings that contain vacancies.

account for these vacancy ordered compounds, with the fraction of compounds containing vacancies shown in purple. A striking result is that the vacancy ordered intermetallic compounds are primarily restricted to the bcc and ω parent crystal structures. Furthermore, over half of the bcc intermetallic compounds that contain vacancies also contain Al.

Both the CaCu_5 and Laves C15 parent crystal structures appear prominently, occupying the third and fourth positions, respectively, in the histograms of Figs. 3 and 4. Both are Laves-like phases, consisting of triangular and kagome layers, and they are common among binary compounds in which

the constituents have very different atomic radii [47,48]. The larger constituent occupies a subset of the triangular layers, while the kagome layers are exclusively occupied by the smaller constituent. Figure 4 indicates that Laves C15 does not host many distinct orderings and is dominated by the C15 prototype. The CaCu_5 parent crystal structure, in contrast, has been assigned five prototype orderings in Fig. 4. However, the $\text{Ni}_{17}\text{Th}_2$, $\text{Th}_2\text{Zn}_{17}$, and Mn_{12}Th prototype orderings assigned to CaCu_5 can only be described as superstructure orderings over the sites of CaCu_5 provided that dumbbells occupy a subset of the Ca sites. This is described in more detail in the Supplemental Material [56].

The number of observed ordered compounds on each parent crystal structure (Table I) is only a small fraction of the total number of geometrically possible orderings. The enumeration of symmetrically distinct orderings within supercells of the primitive unit cell of a parent crystal is a solved combinatorics problem [23]. Figure 5 compares the number of observed orderings to the total number of possible symmetrically distinct orderings as a function of their supercell volume (in integer multiples of the primitive unit cell of the parent crystal) [49] for bcc, fcc, and hcp. While the number of possible orderings increases dramatically with increasing supercell size, the number of observed orderings peaks at a supercell volume of four times the primitive unit cell for both the bcc and fcc parent crystal. Even at small supercell volumes, not all possible orderings are represented in nature. For example, there are two symmetrically distinct orderings on bcc in supercells consisting of two primitive bcc unit cells. Only the B2 ordering is observed in nature. The other, made up of alternating (110) planes of pure A and B, is not among the list of bcc orderings found in this study. This is consistent with a ground-state analysis of the bcc crystal, which showed that of the two possible orderings, only B2 can be a ground state for a lattice model with first- and second-nearest-neighbor pair interactions [21]. This suggests that many of the commonly observed orderings are likely stabilized by short-range interactions. It does not, however, mean that more complex orderings are

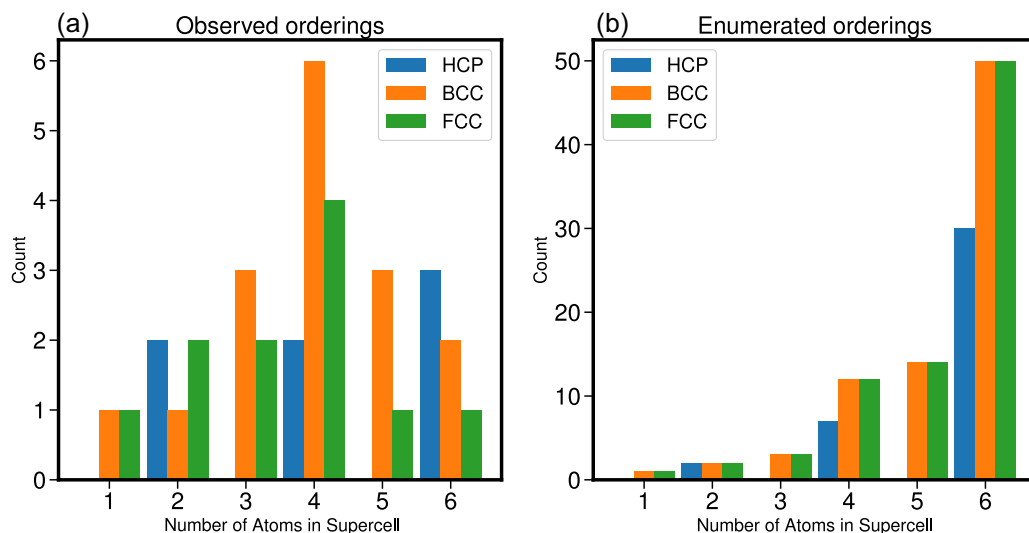


FIG. 5. The number of (a) experimentally observed orderings and (b) geometrically possible (enumerated) orderings as a function of the number of atoms in a supercell for bcc, fcc, and hcp.

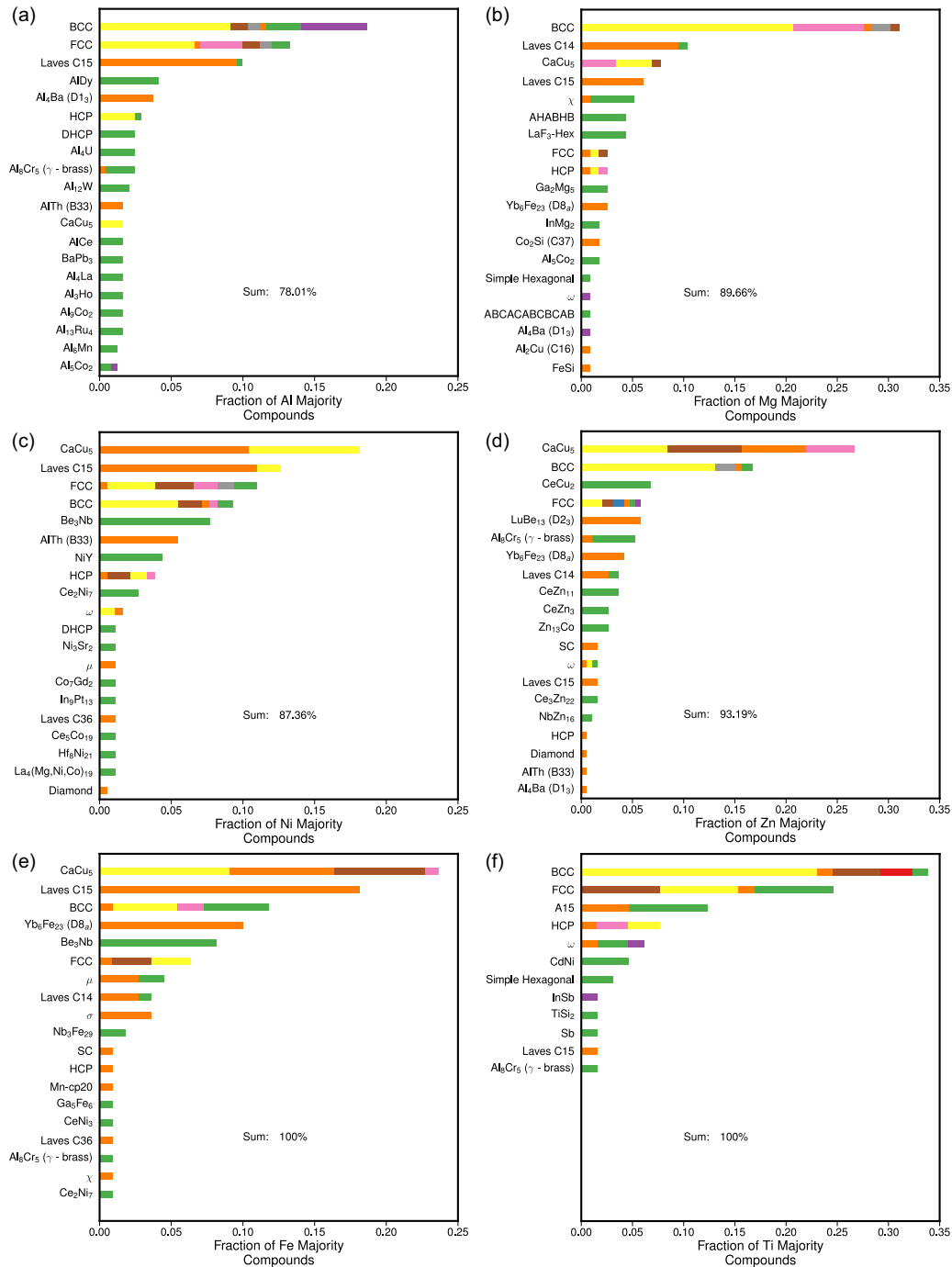


FIG. 6. The distribution of compounds across parent crystal structures for Al-, Mg-, Ni-, Zn-, Fe-, and Ti-rich binary alloys. The colored bars refer to the same orderings as in Fig. 4. For parent crystals not shown in Fig. 4, orange bars indicate an ordering with a Strukturbericht designation, green bars correspond to orderings without any label, and purple indicates orderings that contain vacancies.

necessarily thermodynamically unstable. Ground-state analyses of lattice models show that more complex orderings in large unit cells require long-range interactions (beyond first- and second-nearest-neighbor shells) [20,21]. As long-range interactions tend to be weaker than short-range interactions, the order-disorder transition temperatures of more complex orderings will generally be lower than those of simpler orderings. Hence, they are less likely to be observed as their formation requires equilibration at low temperature. Kinetic

factors may also hamper their formation as it will be easier to quench in local disorder before a complex ordered pattern can be realized through coordinated atomic motion.

B. Crystal hierarchies for important metallic alloys

Parent crystal structures can also be ranked based on their frequency in more restricted alloy classes. Figure 6 shows the distribution of parent crystal structures in Al-, Mg-, Ni-, Zn-,

Fe-, and Ti-rich binary alloys. A comparison of the histograms for each of the six alloy classes shows that there can be a large variability among the top parent crystal structures. For example, the most frequent parent crystal structures when only considering compounds having a majority of either Ni or Fe are CaCu_5 and Laves C15 instead of bcc and fcc. Binary compounds containing a majority of either magnesium, aluminum, or zinc form a large number of bcc orderings even though the elements themselves adopt either the hcp or fcc crystal structure at ambient temperatures and pressures. Aluminum-rich alloys are much more likely than the others to accommodate large vacancy concentrations in bcc based intermetallic compounds, as is evident by the large purple block on its bcc bar. Figure 6 also shows that titanium-rich alloys and iron-rich alloys only have a limited number of parent crystal structures, with 12 for Ti and 19 for Fe. Zn majority compounds frequently show up on the CeCu_2 parent crystal, which is unlike the other five alloy classes.

C. Correlations between the occurrence of crystal structures

We have so far only considered the frequency with which each parent crystal structure occurs in binaries of metallic elements. Also of interest are correlations between pairs of parent crystal structures. That is, for any given pair of parent crystal structures A and B, are they more likely or less likely to appear together in the same binary system than would be expected from independent random sampling? This can be quantified with the following correlation parameter:

$$\kappa = \frac{P(A+B)}{P(A)P(B)} - 1,$$

where $P(A)$ and $P(B)$ are the probabilities with which parent crystal structures A and B occur. For intermetallic phases, we estimate these probabilities by the frequency with which they occur among all the binary systems analyzed in this study. The joint probability $P(A+B)$ represents the probability that the parent crystal structures A and B are both observed in the same binary system.

If the occurrence of parent crystal structures A and B in any binary system is uncorrelated, then $P(A+B) = P(A)P(B)$ and the correlation parameter κ becomes equal to zero. If κ is positive, then the simultaneous presence of A and B in any binary occurs more frequently than that expected from independent random sampling. κ is less than zero when the A and B parent crystal structures are anticorrelated, and κ assumes its minimum value of -1 when A and B never occur together in any binary system [i.e., $P(A+B) = 0$].

Figure 7 displays the matrix of correlation parameters, κ , calculated for all pairs of the top 20 parent crystal structures for the intermetallic compounds considered in this study. The off-diagonal entries contain κ values for pairs of different parent crystal structures. Since several parent crystal structures host multiple derivative ordered structures, it is common that a binary system will exhibit two compounds that are derived from the same parent crystal structure. The diagonal elements of the correlation matrix in Fig. 7 account for these occurrences. Blue signifies a positive correlation, while red indicates a negative correlation.

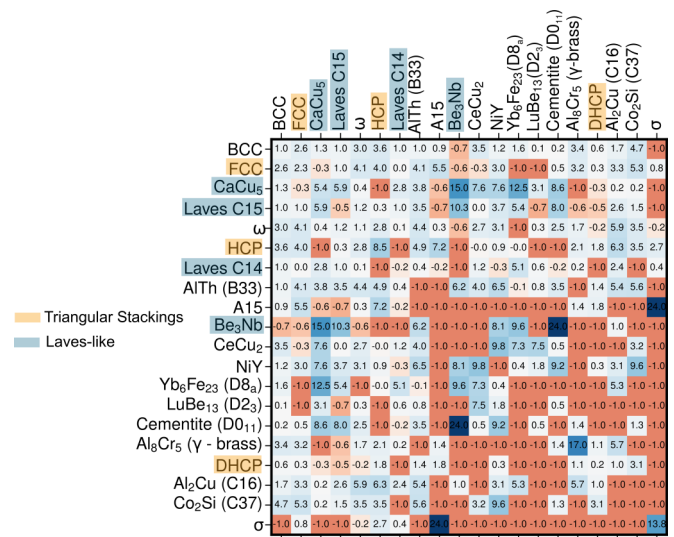


FIG. 7. A matrix that displays the degree with which pairs of parent crystal structures are correlated over binary alloy systems. The numerical values represent the κ correlation parameter. Positive values indicate that derived orderings of pairs of parent crystal structures occur more frequently than independent sampling. Negative quantities indicate that the pair of parent crystal structures are anticorrelated. Positive (negative) values have been colored a degree of blue (red) to reflect their magnitude for easy visualization of the matrix.

Figure 7 shows that the Laves-like phases are highly correlated with each other. The Laves-like CaCu_5 parent crystal, for example, has large κ entries with C15 and Be_3Nb . This is a reflection of the fact that multiple Laves-like phases with different stoichiometries tend to appear in the same binary. It is well known that Laves-like phases form when mixing elements with different atomic radii [47,48]. Figure 7 shows that cementite (D0_{11}) and $\text{Yb}_6\text{Fe}_{23}$ (D8_a) are also highly correlated with Laves phases. The correlation between cementite and Be_3Nb is especially large, with a correlation parameter of $\kappa = 24$. Both cementite and $\text{Yb}_6\text{Fe}_{23}$ also tend to form between elements with large differences in atomic radii, and they can therefore be expected in the same binaries that form Laves-like phases. Laves C14 is unlike the other Laves-like phases in that its correlation parameter, κ , with other parent crystals is generally small, including with other Laves-like phases and cementite, indicating that it does not tend to coexist with other parent crystals in the same binary systems.

Another notable trend in Fig. 7 is that hcp is very strongly anticorrelated with Laves-like phases. The other close-packed parent crystals, including fcc and dhcp, also show negligible to negative correlations with Laves-like phases. This is likely a result of the instability of close-packed phases when mixing constituents with large differences in atomic radii, as was recently shown in a first-principles study [48]. A15 and the σ parent crystals show an especially large number of $\kappa = -1$ entries, meaning that they do not occur with many of the other top 20 parent crystals in binary alloys. However, the two parent crystals are very strongly correlated with each other, with $\kappa = 24$. This coexistence could be due to the fact that both structures have coordination environments in which a central

	BCC	FCC	CaCu ₅	Laves C15	ω	HCP	Laves C14	AlTh (B33)	A15	Be ₃ Nb	CeCu ₂	NiY	Yb ₆ Fe ₂₃ (D8 _a)	LuBe ₁₃ (D2 ₃)	Cementite (D0 ₁₁)	Al ₈ Cr ₅ (Y - brass)	DHCP	Al ₂ Cu (C16)	Co ₂ Si (C37)	σ	other	Number of Compounds	Number of Binaries	Expected Number of Compounds per Binary
HCP HCP	19	7	16	6	6	3	9	1	0	3	1	0	3	4	3	0	0	1	0	0	20	347	484	0.72
FCC HCP	22	13	6	9	4	3	1	3	1	1	1	2	1	1	2	1	2	1	1	0	23	805	330	2.44
BCC HCP	12	4	12	11	1	1	5	0	4	4	2	0	4	5	0	1	0	1	0	6	26	229	308	0.74
FCC FCC	11	12	8	10	4	2	0	3	0	1	3	3	0	0	2	0	0	2	2	0	35	262	225	1.16
BCC FCC	8	22	6	6	2	9	2	1	7	0	1	2	1	0	1	2	2	1	0	3	24	265	210	1.26
BCC BCC	4	0	0	12	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	12	42	24	196	0.12
DHCP HCP	16	6	18	12	5	0	2	0	0	2	2	0	0	5	4	0	10	0	0	0	17	99	88	1.12
DHCP FCC	6	8	10	13	2	2	1	12	0	5	3	6	0	1	2	0	1	0	1	0	27	124	60	2.07

Probabilities shown as %

FIG. 8. The probability (in percent) of the occurrence of each of the top 20 parent crystal structures broken down by the type of crystal structures of the constituents. The “Other” column denotes the percent of all remaining parent crystal structures.

atom is connected to 12 other atoms in its coordination shell. Alternatively, the fact that A15 and σ are both very prominent in refractory containing intermetallics could explain the high κ value. A15 also has large κ values with the close-packed fcc and hcp phases. Both Al₈Cr₅ (γ -brass) and σ have very large diagonal κ values indicating that binaries with these phases have multiple derivatives of these phases.

It is also of interest to inspect the extent to which particular parent crystal structures are correlated with the crystal structures of the pure constituents. This is collected in the matrix of Fig. 8. Each row in Fig. 8 corresponds to a particular combination of pure element crystal structures. The first row, for example, collects all binary systems in which both constituents form hcp as pure elements. The rows are ordered by the number of binaries of each category. For example, the second to last column shows that there are 484 binaries made up of constituents that are both hcp in their pure form. The second most common type of binary is made up of a combination of an fcc and an hcp element; there are 330 binaries of this type.

Figure 8 indicates that a combination of two hcp elements results in a large number of bcc and Laves-like CaCu₅ derived compounds. Surprisingly, the fraction of hcp derived orderings is very low (only 3%) among these binaries, in spite of the fact that both end members are hcp. The next most frequent category of binaries, made of an fcc and an hcp element, also results in a high fraction of bcc derived intermetallic compounds (22%). fcc derived intermetallics are a distant second (13%), and, as with the hcp-hcp binaries, the fraction of hcp intermetallics is very low. The last column of the matrix in Fig. 8 lists the average number of compounds per binary in each category. The combination of an fcc and hcp element leads to an average of 2.44 compounds per binary, which is substantially higher than all the other categories, with

the exception of the dhcp-fcc category, which has a value of approximately 2. The binaries made of two bcc elements show a very high probability of C14 formation and to a lesser extent the σ phase. Also notable about this category is the very low average number of compounds per binary, with a value of 0.12. Furthermore, a high fraction of compounds formed in the bcc-bcc category are categorized among the parent crystal structures below the top 20 discussed in this study (42%).

IV. DISCUSSION

The existence of a hierarchy among intermetallic compounds and simpler parent crystal structures has been recognized in materials science for many decades. Metallic binary systems commonly have intermetallic compounds that undergo order-disorder transitions to a solid solution on the same parent crystal structure at high temperature. A well-known example is CuZn, which forms the B2 ordering at low temperature but disorders through a second-order transition to a bcc solid solution upon heating [50]. Many thermodynamic descriptions of multicomponent crystalline solids exploit the crystallographic relationships that exist between a parent crystal and its derivative orderings [5,17,18,20,22,25,51]. However, these descriptions have primarily been restricted to simpler parent crystal structures such as bcc, fcc, and hcp.

Here we investigated the extent to which intermetallic compounds can more generally be viewed as derivative ordered structures of higher symmetry parent crystal structures. The fact that 3/4 of the intermetallic compounds considered in this work can be mapped onto 20 simpler parent crystal structures suggests that a classification based on this hierarchical relationship is meaningful for a large number of intermetallic compounds. Of particular interest is the large number of parent crystal structures in the top 20 that have a complex

topology and that also host multiple derivative chemical orderings. These include the ω phase, two Laves phases, and Al_8Cr_5 . Furthermore, parent crystal structures such as bcc and ω are found to host chemical orderings with high concentrations of vacancies, a possibility that is often overlooked in high throughput studies of phase stability.

The concept of a parent crystal structure plays a central role in first-principles and empirical thermodynamic descriptions of multicomponent solids. Different orderings on a common parent crystal structure can be distinguished from each other and from a disordered solid solution with the help of long-range order parameters [17,18,25]. This makes it possible to describe the thermodynamic properties of a parent crystal structure and its derivative orderings with a common free-energy surface that is a function not only of concentration, temperature, and pressure, but also of long-range order parameters. Stable chemical orderings on a parent crystal structure then correspond to local minima on this common free-energy surface. First-principles statistical mechanics methods that rely on the cluster expansion approach and Monte Carlo simulations are now routinely used to construct such free-energy surfaces [25,28,52]. Ultimately, global phase stability is determined by a minimization over the free-energy surfaces of multiple competing parent crystal structures.

The calculation of the phase diagram of a new and uncharted alloy system with first-principles statistical mechanics methods requires a manageable, ranked list of candidate parent crystal structures, as the computational cost of calculating a free-energy surface for a single parent crystal can be high [11,46,53]. It is therefore of interest to have the ability to anticipate likely parent crystal structures for a particular composition space. In this context, simple descriptors that depend only on the elemental constituents of the alloy are desirable. Chemical descriptors, such as ionic radii and electronegativity, are commonly invoked to explain and rationalize crystal preferences within an alloy [47,48,54]. They can also be used to assess the likelihood of a particular parent crystal structure within a given composition space. Figure 9, for example, shows the density distribution for each of the top 20 parent crystals as a function of the radius ratio $R_{\text{big}}/R_{\text{small}}$ and electronegativity difference $\chi_{\text{big}} - \chi_{\text{small}}$ of the binary constituents that form derivative orderings on each parent crystal. As is clear in Fig. 9, the density distributions of many of the top 20 parent crystal structures are localized within a restricted region of $R_{\text{big}}/R_{\text{small}}$ and $\chi_{\text{big}} - \chi_{\text{small}}$ space. Exceptions are ω , Laves C14, and Co_2Si (C37), which have more spread out density distributions. The simple parent crystal structures, such as bcc, fcc, and hcp, tend to be favored when combining elements that have similar atomic radii (i.e., $R_{\text{big}}/R_{\text{small}} \approx 1$). Other more complex crystals, such as A15 and σ , also tend to be formed by elements that have similar atomic radii. The Laves-like phases, such as CaCu_5 , C15, and Be_3Nb , in contrast, form in binaries having a large radius ratio $R_{\text{big}}/R_{\text{small}}$ between the constituents. For example, pairs of chemical species that combine to form compounds that are derivative orderings of CaCu_5 have radius ratios between 1.2 and 1.4 and electronegativity differences between 0 and -1.0 .

Overall, the density distributions of Fig. 9 suggest that simple descriptors can be used as a first approximation to assemble a ranked list of likely parent crystal structures. A

further narrowing of likely parent crystal structures in an unexplored composition space can be achieved by exploiting empirically established correlations among parent crystal structures, as, for example, measured by the correlation parameter κ tabulated in Fig. 7. Any prior knowledge about the existence of one parent crystal structure increases the likelihood of other parent crystal structures with which it is highly correlated. For example, the existence of a CaCu_5 derived compound in an alloy system strongly suggests that Be_3Nb , $\text{Yb}_6\text{Fe}_{23}$, cementite (D0_{11}), CeCu_2 , NiY, and Laves C15 may also form in the same binary system, as these parent crystals have a high κ value with respect to CaCu_5 . The high degree of correlation between these particular parent crystal structures is also reflected by the density distributions in $R_{\text{big}}/R_{\text{small}}$ and $\chi_{\text{big}} - \chi_{\text{small}}$ space of Fig. 9. The dashed circles in each density distribution plot of Fig. 9 encompass the domain in $R_{\text{big}}/R_{\text{small}}$ and $\chi_{\text{big}} - \chi_{\text{small}}$ space corresponding to the binary alloys that form CaCu_5 derivatives. Also shown in each density distribution plot are the κ values of Fig. 7 for each parent crystal with respect to CaCu_5 . Figure 9 clearly shows that Be_3Nb , $\text{Yb}_6\text{Fe}_{23}$, cementite (D0_{11}), CeCu_2 , NiY, and Laves C15, which are highly correlated with CaCu_5 in that they commonly appear together in the same binary, have a high density in the same region of $R_{\text{big}}/R_{\text{small}}$ and $\chi_{\text{big}} - \chi_{\text{small}}$ space. This overlap shows that the more complex parent crystal structures such as the Laves-like phases, cementite, CeCu_2 , and NiY tend to form when mixing elements having a large size mismatch. It is also clearly evident in Fig. 9 that parent crystal structures that have a negative correlation κ with respect to CaCu_5 , including HCP, A15, Al_8Cr_5 , dhcp, and σ , have density distributions that fall outside the dashed circle.

While simple descriptors such as radius ratio and electronegativity differences appear to explain correlations between CaCu_5 and other parent crystal structures, it is not always as clear-cut for the other top 20 parent crystal structures. The Supplemental Material [56] shows similar plots to that of Fig. 9 for each of the top 20 parent crystal structures. In many cases, the correlations between pairs of parent crystal structures are consistent with a large overlap of their density distributions in $R_{\text{big}}/R_{\text{small}}$ and $\chi_{\text{big}} - \chi_{\text{small}}$ space. However, there are also many cases in which this is not true, suggesting that further work is required to identify a larger set of descriptors with which the occurrence of each parent crystal structure can be predicted with a high degree of confidence. More data, either experimental or first-principles, will also be required to improve the quantitative reliability of correlations reported here.

Our analysis has so far been restricted to intermetallic compounds. However, similar hierarchies exist among more complex compounds such as oxides, sulfides, nitrides, etc. Corundum Al_2O_3 , for example, can be viewed as a particular Al-vacancy ordering over the interstitial sites of an hcp oxygen sublattice, while anatase TiO_2 is a particular Ti ordering over the octahedral interstitial sites of an fcc oxygen sublattice [55]. More esoteric oxides such as NbO , WO_3 , and MoO_3 can be derived from rocksalt as a particular ordering of vacancies on both its metal and oxygen sublattices. The anions of oxides and sulfides tend to be bigger and more abundant than the metal cations. In many oxides and sulfides, the anions, therefore, form a close-packed sublattice such as fcc or hcp,

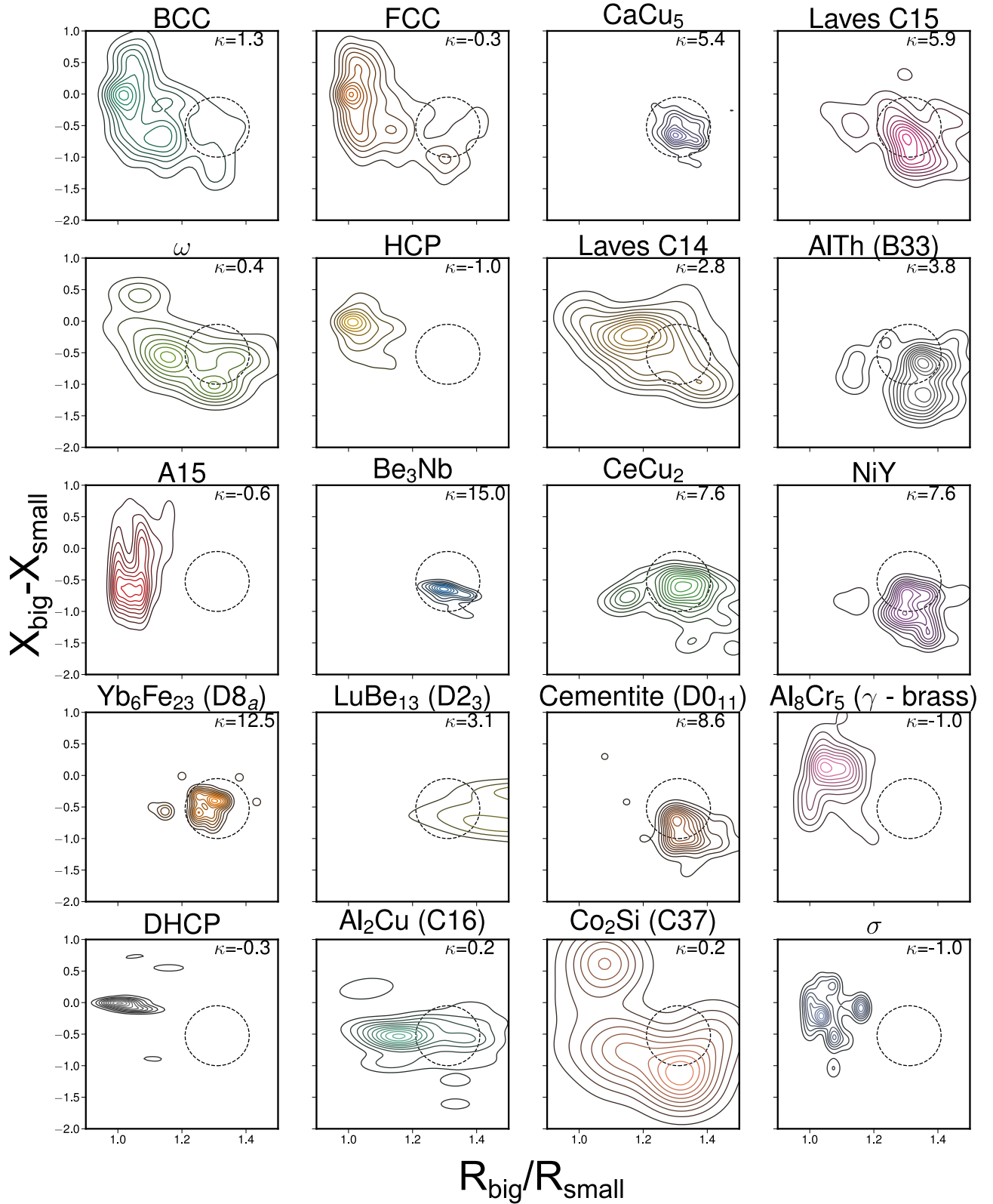


FIG. 9. Contour plots of the kernel density estimates (KDEs) of the distributions of chemistries on the top 20 parent crystals in Fig. 7. The KDEs are over the distribution of radius ratio and electronegativity difference of the elements present in the binary intermetallic. The κ correlation values of each parent crystal with respect to CaCu_5 are also shown in each plot. The majority of CaCu_5 chemistries form with radius ratios between 1.2 and 1.4 and electronegativity differences of -1 to 0 . This region is denoted by the dashed circles.

while the metal cations fill interstitial tetrahedral or octahedral sites. There are, of course, many exceptions to this trend with a vast array of complex oxide and sulfide crystal structures that have more complex anion sublattices. In contrast to oxides and sulfides, the nitrides and carbides commonly have more metal cations than nitrogen or carbon. Many of their crystal structures can therefore be viewed as a close-packed metal sublattice with nitrogen or carbon occupying interstitial sites [53].

V. CONCLUSION

It has long been recognized that common intermetallic crystal structures such as B2 and $L1_2$ are superlattice orderings of simpler parent crystal structures such as bcc and fcc. Common first-principles statistical mechanics approaches to calculate finite-temperature phase stability exploit the hierarchical relationship between ordered compounds and their underlying parent crystal structures. In this work, we have explored the extent to which intermetallic compounds in general can be viewed as derivative orderings on higher symmetry parent crystal structures. Our analysis was enabled by a robust mapping algorithm to measure the similarity between crystal structures. We found that a large fraction of intermetallic compounds can indeed be viewed as chemical orderings over a small number of higher symmetry parent crystal structures.

While many compounds are derivative orderings of bcc and fcc, a larger number are found to be orderings of more complex crystal structures that include ω and Laves phases. A similar approach can be applied to categorize other classes of crystalline materials including oxides, sulfides, carbides, nitrides, etc. The hierarchical organization of crystal structures into high symmetry parent crystals and derived ordered structures should enable the development of a deeper understanding of the relationship between crystal structures, and it should lay the foundation for the efficient calculation of phase diagrams in new alloy systems.

ACKNOWLEDGMENTS

This work was supported by the ONR BRC Program, Grant No. N00014-18-1-2392, the NSF IDEAS center, Grant No. 1934641, and the Schmidt Family Foundation. Use was made of computational facilities purchased with funds from the National Science Foundation (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. We are also grateful for computing resources from the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

-
- [1] T. M. Pollock and A. Van der Ven, *MRS Bull.* **44**, 238 (2019).
 - [2] A. Van De Walle and M. Asta, *MRS Bull.* **44**, 252 (2019).
 - [3] R. Arróyave and D. L. McDowell, *Annu. Rev. Mater. Res.* **49**, 103 (2019).
 - [4] A. Abu-Odeh, E. Galvan, T. Kirk, H. Mao, Q. Chen, P. Mason, R. Malak, and R. Arróyave, *Acta Mater.* **152**, 41 (2018).
 - [5] D. DeFontaine, *Solid State Phys.* **47**, 33 (1994).
 - [6] M. Asta, V. Ozolins, and C. Woodward, *JOM* **53**, 16 (2001).
 - [7] A. Van de Walle and G. Ceder, *Rev. Mod. Phys.* **74**, 11 (2002).
 - [8] A. Van de Walle and M. Asta, *Modell. Simul. Mater. Sci. Eng.* **10**, 521 (2002).
 - [9] D. Alfè, *Comput. Phys. Commun.* **180**, 2622 (2009).
 - [10] A. Van De Walle, *JOM* **65**, 1523 (2013).
 - [11] B. Puchala and A. Van der Ven, *Phys. Rev. B* **88**, 094108 (2013).
 - [12] J. C. Thomas and A. Van der Ven, *Phys. Rev. B* **88**, 214111 (2013).
 - [13] R. K. Rhein, P. C. Dodge, M.-H. Chen, M. S. Titus, T. M. Pollock, and A. Van der Ven, *Phys. Rev. B* **92**, 174117 (2015).
 - [14] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso *et al.*, *Science* **351** (2016).
 - [15] A. Van der Ven, J. C. Thomas, B. Puchala, and A. R. Natarajan, *Annu. Rev. Mater. Res.* **48**, 27 (2018).
 - [16] D. De Fontaine, *Acta Metall.* **23**, 553 (1975).
 - [17] A. Khachaturian, *Progr. Mater. Sci.* **22**, 1 (1978).
 - [18] D. de Fontaine, *Solid State Phys.* **34**, 73 (1979).
 - [19] T. Mohri, J. Sanchez, and D. De Fontaine, *Acta Metall.* **33**, 1171 (1985).
 - [20] F. Ducastelle and F. Ducastelle, North-Holland, Amsterdam, (1991).
 - [21] G. Inden, in *Materials Science and Technology* (Wiley-VCH Verlag GmbH, Weinheim, Germany, 2013).
 - [22] A. G. Khachaturyan, *Theory of Structural Transformations in Solids* (Courier, Mineola, New York, 2013).
 - [23] G. L. W. Hart and R. W. Forcade, *Phys. Rev. B* **80**, 014120 (2009).
 - [24] J. M. Sanchez, F. Ducastelle, and D. Gratias, *Physica A* **128**, 334 (1984).
 - [25] A. R. Natarajan, J. C. Thomas, B. Puchala, and A. Van der Ven, *Phys. Rev. B* **96**, 134204 (2017).
 - [26] J. C. Thomas and A. Van der Ven, *Phys. Rev. B* **96**, 134121 (2017).
 - [27] J. C. Thomas and A. Van der Ven, *J. Mech. Phys. Solids* **107**, 76 (2017).
 - [28] G. H. Teichert, A. Natarajan, A. Van der Ven, and K. Garikipati, *Comput. Methods Appl. Mech. Eng.* **353**, 201 (2019).
 - [29] S. Curtarolo, W. Setyawan, G. L. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy *et al.*, *Comput. Mater. Sci.* **58**, 218 (2012).
 - [30] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, *JOM* **65**, 1501 (2013).
 - [31] 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, *APL Mater.* **1**, 011002 (2013).
 - [32] W. Ye, C. Chen, S. Dwaraknath, A. Jain, S. P. Ong, and K. A. Persson, *MRS Bull.* **43**, 664 (2018).

- [33] L. Gelato and E. Parthé, *J. Appl. Crystallogr.* **20**, 139 (1987).
- [34] H. Burzlaff and Y. Malinovsky, *Acta Crystallogr. Sect. A* **53**, 217 (1997).
- [35] A. Dzyabchenko, *Acta Crystallogr. Sect. B* **50**, 414 (1994).
- [36] G. Flor, D. Orobengoa, E. Tasci, J. M. Perez-Mato, and M. I. Aroyo, *J. Appl. Crystallogr.* **49**, 653 (2016).
- [37] E. Willighagen, R. Wehrens, P. Verwer, R. De Gelder, and L. Buydens, *Acta Crystallogr. Sect. B* **61**, 29 (2005).
- [38] R. Hundt, J. Schön, and M. Jansen, *J. Appl. Crystallogr.* **39**, 6 (2006).
- [39] D. C. Lonie and E. Zurek, *Comput. Phys. Commun.* **183**, 690 (2012).
- [40] D. R. Trinkle, R. G. Hennig, S. G. Srinivasan, D. M. Hatch, M. D. Jones, H. T. Stokes, R. C. Albers, and J. W. Wilkins, *Phys. Rev. Lett.* **91**, 025701 (2003).
- [41] H. W. Kuhn, *Naval Res. Logist. Quart.* **2**, 83 (1955).
- [42] H. W. Kuhn, *Naval Res. Logist. Quart.* **3**, 253 (1956).
- [43] D. Zagorac, H. Müller, S. Ruehl, J. Zagorac, and S. Rehme, *J. Appl. Crystallogr.* **52**, 918 (2019).
- [44] S. K. Kolli, <https://labs.materials.ucsb.edu/vandervan/anton/research/> (2020).
- [45] D. Miracle, *Acta Metall. Mater.* **41**, 649 (1993).
- [46] J. G. Goiri and A. Van der Ven, *Phys. Rev. B* **94**, 094111 (2016).
- [47] M. J. Murray and J. V. Sanders, *Philos. Mag. A* **42**, 721 (1980).
- [48] A. R. Natarajan and A. Van der Ven, *Phys. Rev. Lett.* **121**, 255701 (2018).
- [49] CASM Developers (2020), URL <https://zenodo.org/record/546148#.WbQ2QRdlBhE>.
- [50] T. B. Massalski, H. H. Okamoto, and ASM International, *Binary Alloy Phase Diagrams* (ASM International, Materials Park, Ohio, 1990).
- [51] S. M. Allen and J. W. Cahn, *Acta Metall.* **27**, 1085 (1979).
- [52] G. Teichert, A. Natarajan, A. Van der Ven, and K. Garikipati, *Comput. Methods Appl. Mech. Eng.* **371**, 113281 (2020).
- [53] N. S. Harsha Gunda and A. Van der Ven, *Phys. Rev. Mater.* **2**, 083602 (2018).
- [54] U. Mizutani, *Surface Properties and Engineering of Complex Intermetallics* (World Scientific, Singapore, 2010), pp. 323–399.
- [55] N. S. Harsha Gunda, B. Puchala, and A. Van der Ven, *Phys. Rev. Mater.* **2**, 033604 (2018).
- [56] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevMaterials.4.113604> for a robust approach to measure the dissimilarity of two crystal structures and other additional analysis.