

The Structure and Composition Statistics of 6A Binary and Ternary Crystalline Materials

Alon Hever,[†] Corey Oses,^{‡,§} Stefano Curtarolo,^{‡,§,||} Ohad Levy,^{‡,§,⊥} and Amir Natan^{*,†,§,||} 

[†]Department of Physical Electronics, Tel-Aviv University, Tel-Aviv 69978, Israel

[‡]Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, United States

[§]Center for Materials Genomics, Duke University, Durham, North Carolina 27708, United States

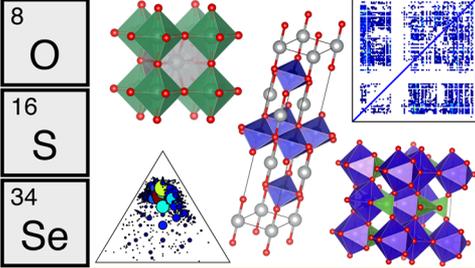
^{||}Fritz-Haber-Institut der Max-Planck-Gesellschaft, 14195 Berlin-Dahlem, Germany

[⊥]Department of Physics, NRCN, P.O. Box 9001, Beer-Sheva 84190, Israel

^{*}The Sackler Center for Computational Molecular and Materials Science, Tel-Aviv University, Tel-Aviv 69978, Israel

Supporting Information

ABSTRACT: The fundamental principles underlying the arrangement of elements into solid compounds with an enormous variety of crystal structures are still largely unknown. This study presents a general overview of the structure types appearing in an important subset of the solid compounds, i.e., binary and ternary compounds of the 6A column oxides, sulfides and selenides. It contains an analysis of these compounds, including the prevalence of various structure types, their symmetry properties, compositions, stoichiometries and unit cell sizes. It is found that these compound families include preferred stoichiometries and structure types that may reflect both their specific chemistry and research bias in the available empirical data. Identification of nonoverlapping gaps and missing stoichiometries in these structure populations may be used as guidance in the search for new materials.



■ INTRODUCTION

The creation of novel materials with optimal properties for diverse applications requires a fundamental understanding of the factors that govern the formation of crystalline solids from various mixtures of elements. Compounds of the nonmetallic elements of column 6A, oxygen, sulfur and selenium, are of particular interest. They serve in a large variety of applications in diverse fields of technology, e.g., chemistry, catalysis, optics, gas sensors, electronics, thermoelectrics, piezoelectrics, topological insulators, spintronics and more.^{1–8} Given the very large number of possibilities, many of the alloy systems of these elements have not been fully investigated, some of them even not at all.

In recent years, high-throughput computational techniques based on ab initio calculations have emerged as a potential route to bridge these experimental gaps and gain understanding of the governing principles of compound formation.⁹ This led to the creation of large databases of computational materials data.^{10,11}

Yet, these computational approaches are practically limited by the number and size of structures that can be thoroughly analyzed, and fundamental issues that limit the applicability of standard semilocal DFT for nonmetallic compounds. The sought-after governing principles are thus still largely unknown.

Nevertheless, the considerable body of experimental data that is already available, although by no means complete, is a useful basis for large-scale data analysis. This experimental data is usually presented in compendiums that lack statistical analysis.

Presenting this data in a structured manner may be conducive for gaining insights into the essential factors that determine structure formation, and may help to provide material scientists with the necessary foundation for rational materials design.

Analyses recently carried out for the intermetallic binaries¹² and ternaries¹³ have uncovered interesting Bravais lattices distributions and an unexpected large prevalence of unique structure types. Here we extend the analysis and discuss trends, as well as special phenomena, across binary and ternary compounds of the 6A nonmetals. This analysis reveals the following interesting observations:

- Considerable overlap exists between the sulfides and selenides: about a third of the total number of structure types are shared among both compound families. In contrast, the overlap between the oxides and the other two families is rather small.
- The prevalence of different compound stoichiometries in the sulfide and selenide families is very similar to each other but different from that of the oxides. Some stoichiometries are abundant in the oxides but are *almost absent* in the sulfides or selenides, and vice versa.
- The number of ternary oxide stoichiometries, $A_xB_yO_z$, decreases when the product of binary oxide stoichiometries, of participating elements, increases. This behavior

Received: September 28, 2017

Published: December 22, 2017

can be explained by general thermodynamic arguments and is discussed in the text.

- Overall, oxide compounds tend to have richer oxygen content than the sulfur and selenium content in their corresponding compounds.
- Across all three compound families, most structure types are represented by only one compound.
- High symmetry lattices, e.g., the orthorhombic face centered, orthorhombic body centered and cubic lattices are relatively rare among these compounds. This reflects the spatial arrangement of the compound forming orbitals of the 6A nonmetals, whose chemistry does not favor these structures.

In the analysis presented here, we adopt the ordering of the elements by Mendeleev numbers as defined by Pettifor,^{14,15} and complement it by investigating the crystallographic properties of the experimentally reported compounds. Pettifor maps constructed for these compound families exhibit similar separation between different structure types as the classical Pettifor maps for binary structure types.^{14,15} For some stoichiometries, the structure types show similar patterns in the maps of the three compound families, suggesting that similar atoms tend to form these stoichiometries with all three elements. Such similarity of patterns is more common between the sulfides and selenides than between either of them and the oxides.

These findings suggest a few possible guiding principles for directed searches of new compounds. Element substitution could be used to examine favorable candidates within the imperfect overlaps of the structure distributions, especially between the sulfides and selenides. Moreover, the missing stoichiometries and structure symmetries mean that data-driven approaches, e.g., machine learning, must use training sets not limited to one compound family, even in studies directed at that specific set of compounds. This hurdle may be avoided by augmenting the known structures with those of the other families. In addition, identified gaps in the Mendeleev maps suggest potential new compounds, both within each family or by correlations of similar structure maps across the different families.

■ DATA METHODOLOGY

The ICSD¹⁶ includes approximately 169 800 entries (as of August 2016). For this study we exclude all entries with partial or random occupation and those that do not have full structure data. The remaining set of structures has been filtered using the AFLOW software,^{17–25} which uses an error checking protocol to ensure the integrity of each entry. AFLOW generates each structure by appropriately propagating the Wyckoff positions of the specified spacegroup. Those structures that produce inconsistencies, e.g., overlapping atoms or a different stoichiometry than the structure label are ignored. If atoms are detected to be too close (≤ 0.6 Å), alternative standard ITC (International Table of Crystallography)²⁶ settings of the spacegroup are attempted. These settings define different choices for the cell's unique axes, possibly causing atoms to overlap if not reported correctly. Overall, these considerations reduce the full set of ICSD entries to a much smaller set of 88 373 “true” compounds. These entries are contained in AFLOW Database.^{27–30} They include the results of the AFLOW generated full symmetry analysis for each structure, i.e., Bravais lattice, space group and point group classifications, and Pearson symbol (the method and tolerances used for this analysis follow the AFLOW standard²⁹). For the analysis presented here we identify all the binary and ternary compounds included in this set, 27 487 binary entries and 37 907 ternary entries. From these, we extract all the entries that contain oxygen, sulfur or selenium as one of the components. Of the binaries, we find 3256

oxides, 1685 sulfides and 1050 selenides. 10 530 oxides, 3190 sulfides and 1786 selenides are found among the ternaries. Duplicate entries representing different experimental reports of the same compound, i.e., the same elements, stoichiometry, space group and Pearson designation, are then eliminated to obtain a list in which every reported compound is represented by its most recent corresponding entry in the ICSD. This reduces our list of binaries to 844 oxides, 495 sulfides and 332 selenides, and the list of ternaries to 5435 oxides, 2041 sulfides and 1256 selenides. These results are summarized in Table 1. Throughout

Table 1. Data Extraction Numerical Summary

	compounds	unique compounds	structure types
total	88373	50294	13324
unary	1752	499	197
binary	27487	10122	1962
binary oxides	3256	844	538
binary sulfides	1685	495	270
binary selenides	1050	332	168
ternary	37907	23398	4409
ternary oxides	10530	5435	2079
ternary sulfides	3190	2041	784
ternary selenides	1786	1256	521
quaternary	15138	11050	3855
5 atoms	4638	3899	2053
6 atoms	1219	1101	682
7 atoms	212	201	154
8 atoms	20	20	12

the rest of the paper, we will refer to these sets of binary and ternary compounds. We choose not to discuss multicomponent structures with four or more elements since their relative scarcity in the database most probably indicates incomplete experimental data rather than fundamental issues of their chemistry. It is also instructive to check the effect of element abundance on the number of compounds. The abundance of oxygen in the earth's crust is $\sim 47\%$ by weight, around 1000 times more than that of sulfur (~ 697 ppm) which is around 5000 more abundant than selenium (120 ppb).³¹ Comparison with the number of elements (O/S/Se) binary compounds, 844/495/332, or ternary compounds, 5435/2041/1256, makes it clear that while a rough correlation exists between the elements abundance and the number of their known compounds, it is by no means a simple proportion.

In the next stage, we identify unique structure types. Structure types are distinguished by stoichiometry, space group, and Pearson designation, without consideration of the specific elemental composition. This implicit definition of structure type is common in the literature,^{32,33} and we use it throughout the manuscript as providing a good balance of clarity and simplicity. However, it should be noted that there are a few rare cases of complex structures where a given structure type under this definition includes a few subtypes (see Figure S1 in the Supporting Information). Examples exist of more complex definitions of structure types, formulated to define similarities between inorganic crystals structures.³⁴

The binary structure type lists contain 538 oxides, 270 sulfides and 168 selenides. The ternary lists contain 2079 oxides, 784 sulfides and 521 selenides. This means that 64% of the binary oxides, 55% of the sulfides and 51% of the selenides are distinct structure types. The corresponding ratios for the ternaries are 38% of the oxides, 38% of the sulfides and 41% of the selenides. All the other entries in the compound lists represent compounds of the same structure types populated by different elements. Differently put, this means that there are on average about 1.6 compounds per structure type in the binary oxides, 1.8 in the binary sulfides and 2 in the binary selenides. Among the ternaries, the corresponding numbers are 2.6 compounds per structure type in the oxides, 2.6 in the sulfides and 2.4 in the selenides. These numbers may be compared to the intermetallics, where there are 20 829 compounds of which 2166, about 10%, are unique structure types.¹² There are about seven compounds per structure types in the binary intermetallics and

about nine in the ternaries. The number for binary intermetallics is considerably larger than for ternary oxides, sulfides or selenides. Together with the higher proportion of unique structure types in the latter, this reflects the limits on materials chemistry imposed by the presence of one of those 6A elements.

It should be noted that this structure selection procedure produces lists that partially overlap, i.e., certain structure types may appear in more than one list, since there might be oxide structure types that are also represented among the sulfide or selenide structures, and vice versa. 11% of the oxide binary structure types also appear in the sulfides binary list and 8% are represented in the selenides binary list. 33% of the sulfide binaries are also represented in the selenides list. The total number of binary oxides, sulfides and selenides structure types is 976, which is reduced by 16%, to 818 structure types, by removing all overlaps. The corresponding overlap ratios for the ternaries are 10% for the oxides and sulfides, 6% for the oxides and selenides and 31% for the sulfides and selenides. The total number of entries in the ternary oxides, sulfides, and selenides structure type lists is 3384, which is reduced to 2797 structure types by removing all overlaps, a 17% reduction. Therefore, the overlaps between these three compound families are similar for the binaries and ternaries. In both, the overlap between the oxides and the other two families is rather small, whereas the overlap between the sulfides and selenides represents about a third of the total number of structure types.

The sequence of Mendeleev numbers includes 103 elements, from hydrogen to lawrencium with numbers 1–6 assigned to the noble gases, 2–16 to the alkali metals and alkaline earths, 17–48 to the rare earths and actinides, 49–92 to the metals and metalloids and 93–103 to the nonmetals. Of these, noble gases are not present in compounds and artificial elements (metals heavier than uranium) have very few known compounds. We are thus left with 86 elements, of which the above compounds are composed. That means there are about ten times more oxide binaries than element-oxygen combinations, about six times more sulfides than element-sulfur combinations and four times more selenides than element-selenium combinations. Oxides are much more common than sulfides and selenides. The corresponding numbers for the ternaries are much lower. There are about 1.6 times more ternary oxides than two-element-oxygen ternary possible systems, about 0.6 times less ternary sulfides and about 0.4 times less ternary selenides than the corresponding two-element combinations. The ternaries are relatively quite rare, more so as we progress from oxides to sulfides and then to selenides. A similar analysis of the intermetallic binaries in ref 12 shows that of the 20 829 intermetallics, 277 are unaries (about three times more than possible metal elements), 6441 are binaries (about two times more than possible metal binary systems), and 13 026 are ternaries (6.5 times less than possible metal ternary systems). This means that unary metal structures are less common among the metallic elements than the oxide, sulfide and selenide binary compounds among their corresponding binary systems. This seems to reflect simply the larger space of stoichiometries available to binaries over unaries. However, on the contrary, the intermetallic binary compounds are more common among the metallic binary systems than the oxide, sulfide and selenide ternary compounds among their corresponding ternary systems. This discrepancy again reflects either the chemical constraints imposed by the presence of a 6A nonmetal on the formation of a stable ternary structure, or simply gaps in the experimental data since many ternary systems have not been thoroughly investigated.

RESULTS AND DISCUSSION

Structure Types. The distribution of the binary and ternary compounds among the corresponding structure types is shown in Figure 1. Detailed data for the most common structure types is presented in Tables S1–S6 in the Supporting Information.

About 84% of the binary oxide structure types represent a single compound, characterizing the tail end of the binary oxide distribution. They include about 53% of the binary oxide compounds. The most common structure type represents 29 compounds, 3.4% of the oxide compounds list. Among the

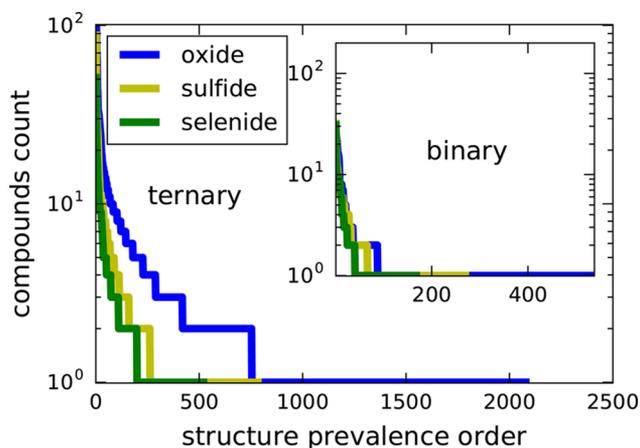


Figure 1. Distributions of the compounds among structure types for binary (inset) and ternary compounds. Oxides are shown in blue, sulfides in yellow and selenides in green. The binary distributions differ mostly by the length of their single-compound prototypes tails, while the ternary distribution of the oxides deviates significantly from those of the sulfides and selenides.

binary sulfides, 76% of the structure types represent a single compound. They include 41% of the binary sulfide compounds. The most common structure type represents 32 compounds, 6.5% of the sulfide compounds list. Among the binary selenides, 76% of the structure types represent a single compound. They include 39% of the binary selenide compounds. The most common structure type represents 31 compounds, 9.3% of the selenide compounds list.

In all three binary lists the most common structure type is rock salt (NaCl). The binary oxide structure type distribution has a much longer tail than the sulfides and selenides, i.e., more oxide compounds have unique structure types. The most common structure type in these three distributions represents a similar number of compounds but a smaller proportion of the corresponding compounds in the oxides. The middle regions of the distributions are very similar (inset Figure 1). This means that the much larger number of binary oxide compounds, compared to the sulfides and selenides, is expressed at the margin of the distribution, in the long tail of unique compounds.

This discrepancy between the three binary distributions is much less apparent among the ternary compounds. 64% of the ternary oxide structure types represent a single compound. They include 24% of the ternary oxide compounds. The two most common structure types, pyrochlore and perovskite, represent 116 and 115 compounds, respectively, about 2% each of the entire compounds list. Among the ternary sulfides, 70% of the structure types represent a single compound. They include 34% of the ternary sulfide compounds. The most common structure type, delafossite, represents 65 compounds, 4% of the entire compounds list. Among the ternary selenides, 62% of the structure types represent a single compound. They include 26% of the ternary selenide compounds. The most common structure type, again delafossite, represents 51 compounds, 4% of the ternary sulfides.

In contrast to the binaries, the larger count of ternary oxides, compared to the sulfides and selenides, is expressed by a thicker middle region of the structure type distribution, whereas the margins have a similar weight in the distributions of the three compound families.

Binary Stoichiometries. The structure types stoichiometry distribution for the binary oxide, sulfide and selenide

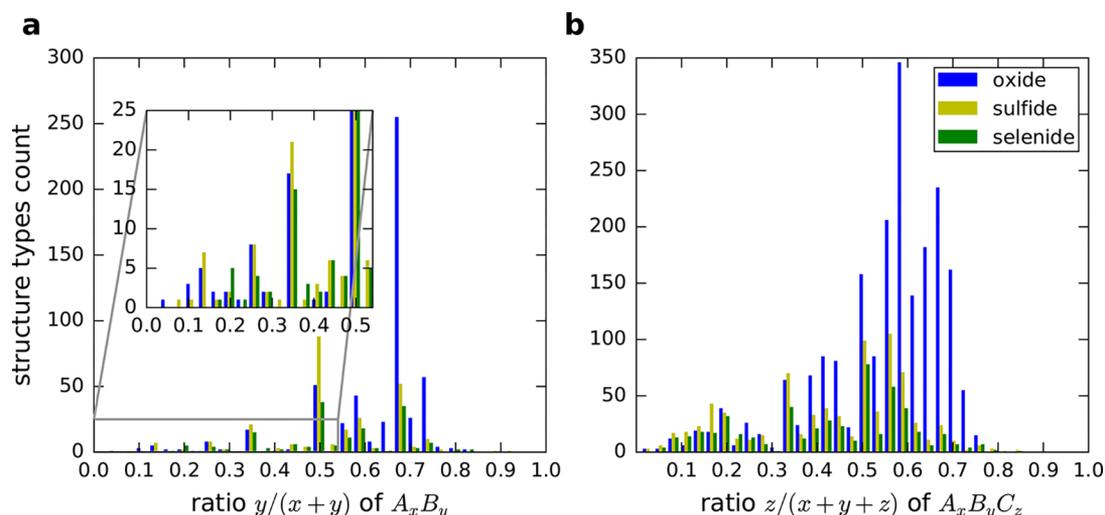


Figure 2. Distribution of (a) binary stoichiometries and (b) ternary stoichiometries. Oxides are shown in blue, sulfides in yellow and selenides in green. The distributions of the selenides and sulfides are quite similar, while those of the oxides deviate significantly, as detailed in the text.

compounds is shown in Figure 2(a). We define the binaries as A_xB_y , where B is O, S or Se, and the number of structure types is shown as a function of $y/(y+x)$. A very clear peak is found for the oxides at the stoichiometry 1:2, AO_2 , while both the sulfides and selenides have a major peak at 1:1, AS and ASe, respectively.

For $y/(y+x) < 0.5$, there are more gaps in the plot (missing stoichiometries) for the oxides compared to the sulfides and selenides, while for $y/(y+x) > 0.6$ there are more gaps in the sulfides and selenides, this behavior is shown in detail in Table S10 in the SI. An important practical conclusion is that augmenting the binary oxide structure types with those of sulfides and selenides will produce a more extensive coverage of possible stoichiometries.

Another interesting property is the number of stoichiometries for each of the elements in the periodic table. The prevalence of binary oxide stoichiometries per element is shown in Figure 3(a). A few interesting trends are evident: the first row of transition metals shows a peak near vanadium (19 stoichiometries) and titanium (14 stoichiometries). Hafnium, which is in the same column of titanium has only a single stoichiometry, HfO_2 . Both the beginning and end of the d -elements exhibit a small amount of stoichiometries—scandium with only one and zinc with only two. The two most abundant elements, silicon and oxygen, form only a single stoichiometry in the ICSD, SiO_2 , with 185 different structure types. Another interesting trend is evident for the alkali metals, where rubidium and cesium have more stoichiometries, perhaps related to the participation of d -electrons in the chemical bonds.

Figures 3(b) and (c) show the binary stoichiometries prevalence per element for sulfur and selenium, respectively. Similar trends are exhibited: there are two “islands” of large number of stoichiometries in the transition metals: one around vanadium and titanium and the other near nickel and copper. Evidently, prime candidates for new compounds should be searched among structures in the vicinity of these high density islands, especially for elements that exhibit a considerably higher density in one family.

Ternary Stoichiometries. Similar to the binaries, the ternary stoichiometries are designated $A_xB_yC_z$, where C is O, S or Se. The distributions of the ternaries are, as might be expected, more complex, with maxima at $z/(x+y+z) = 0.6$ for the oxides, $z/(y+x+z) = 0.55$ for the sulfides and $z/(y+x+z)$

$= 0.5$ for the selenides. The major peaks still appear at integer and half integer values, but with more minor peaks at intermediate values. This behavior is shown in Figure 2(b). The ternary selenide and sulfides distributions are again nearly identical, and there are almost no compounds with ratios larger than 0.75 in the oxides or larger than 0.66 in the sulfides and selenides. However, there are few sulfide and selenide compounds around 0.8 and 0.85 but no oxides.

Another perspective of ternary stoichiometries is demonstrated in Figure 4 which shows the abundance of the most common stoichiometries. The biggest circle in each diagram denotes the prevalence of the most common stoichiometry (number of unique compounds for this stoichiometry), which is 718 ($x = 1, y = 1, z = 3$) for oxides, 242 ($x = 1, y = 1, z = 2$) for sulfides, and 145 ($x = 1, y = 1, z = 2$) for selenides. The smaller circles in each plot are normalized to the corresponding highest prevalence.

These diagrams highlight the similarities as well as important differences between the three families of compounds. In all three cases, the most common stoichiometries appear on the symmetry axis of the diagram, i.e., at equal concentrations of the A and B components, or very close to it. For the oxides, they are concentrated near 0.5–0.6 fraction of oxygen, representing the $A_1B_1O_2$ and $A_1B_1O_3$ stoichiometries, respectively, and form a very dense cluster with many similar reported stoichiometries of lower prevalence. Outside this cluster, the occurrence of reported compositions drops sharply, and other regions of the diagram are very sparsely populated, in particular near the vertices of the B and O components.

The sulfide and selenide diagrams also exhibit prominent clusters on the AB symmetry axes, but they appear at a lower S or Se concentration of about 0.5, i.e., $A_1B_1C_2$ stoichiometry. They are considerably more spread out and include a significant contribution at the ABC stoichiometry. In both sulfides and selenides, an additional minor cluster appears closer to the A vertex (Figure 4). A few members of this cluster are ternary oxides, reflecting the high electronegativity and high Mendeleev number (101) of oxygen. The B and C vertex regions are still sparsely populated, but less so than in the oxides case. Overall, the sulfide and selenide diagrams are very similar to each other and different from that of the oxides. They are more spread out, less AB symmetric than the oxide diagram and less tilted toward

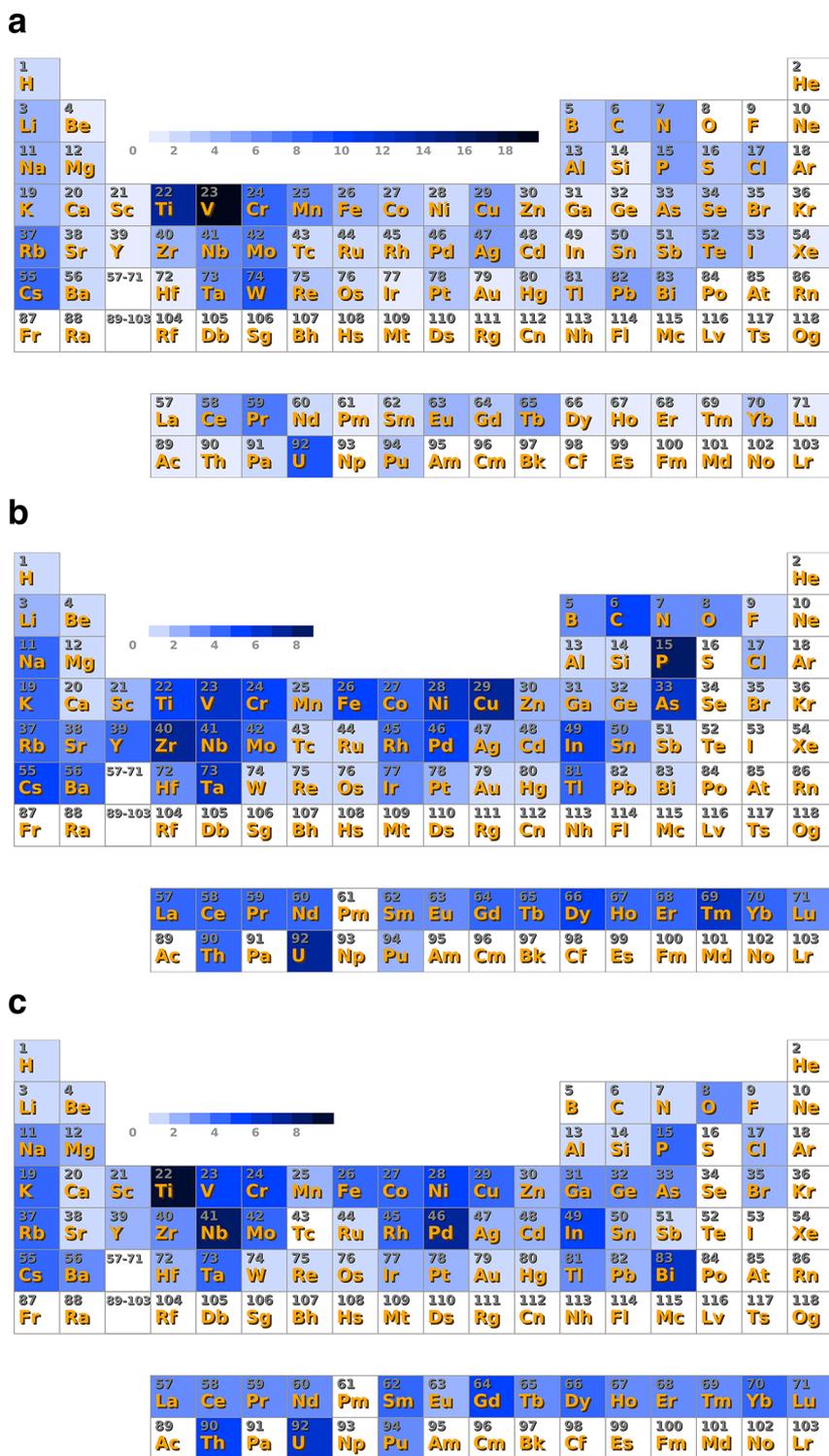


Figure 3. Binary (a) oxide, (b) sulfide, and (c) selenide stoichiometries (number of different stoichiometries that include the respective element). The colors go from no stoichiometries (white) to the maximal number of stoichiometries (dark blue) which is different for each element, 19/8/9 for O/S/Se. Islands of high prevalence appear for the 4B and 5B transition metals and the heavy alkalis in all three compound families. Additional, smaller islands appear in the sulfides and selenides for the 8 and 1B transition metals and the 3A and 5A semimetals.

rich C-component concentration. This discrepancy may reflect some uniqueness of oxygen chemistry compared to sulfur and selenium, or rather simply reflect the oxygen rich environment in which naturally formed compounds are created in the atmosphere. The number of stoichiometries and the differences in the C-component concentration are summarized in Table 2.

Another interesting observation is that while some stoichiometries are abundant in the oxides they are almost absent in the sulfides or the selenides. For example, there are 299 compounds with the $A_2B_2O_7$ stoichiometry (ignoring order between M_A and M_B), but only two $A_2B_2S_7$ compounds and no $A_2B_2Se_7$ compounds. Also, there are 71 $A_1B_3O_9$ compounds but no $A_1B_3S_9$ and $A_1B_3Se_9$ compounds. On the other hand, there

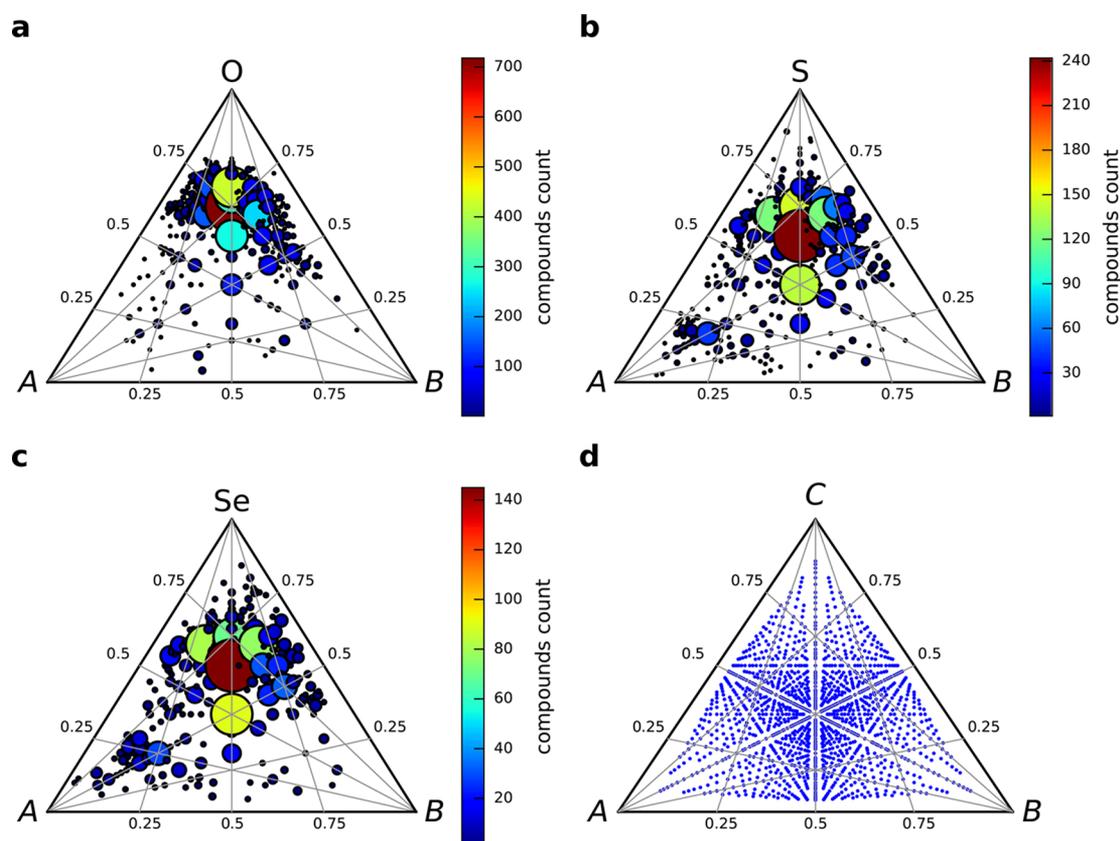


Figure 4. Prevalence of stoichiometries among the ternary (a) oxide, (b) sulfide and (c) selenide compounds. (d) shows, for reference, all the possible stoichiometries with up to 12 atoms of each component per unit cell. In each figure, the smaller circles are normalized to the biggest one, which denotes the highest prevalence, i.e., 718 for oxides, 242 for sulfides, and 145 for selenides. In addition a heat map color scheme is used where blue means low prevalence and red means the highest prevalence for each element. The x and y axes denote the atomic fractions in the ternaries $A_xB_yC_z$, where C is O, S or Se, respectively. A and B are ordered by Mendeleev number where $M_A > M_B$.

Table 2. Ternary Stoichiometry Data: $A_xB_yC_z$ ^a

	oxygen	sulfur	selenium
Number of stoichiometries	585	282	206
C-rich stoichiometries ratio	0.85	0.67	0.66
C-rich compound ratio	0.92	0.77	0.73

^a“C-rich” refers to stoichiometries where $z > x + y$.

are no $A_4B_{11}X_{22}$ oxides, but 20 sulfides and 8 selenides. If we require that $M_A > M_B$, there are no oxides of the $A_3B_2X_2$ stoichiometry, but 25 sulfides and 7 selenides.

Again, an important conclusion is that there are many missing stoichiometries, Figure 4(d) shows all the possible stoichiometries for $A_xB_yC_z$ for $x, y, z \leq 12$, clearly showing rich concentration in the middle, which is not the case for oxides, and also to a lesser degree to sulfides and selenides.

We can repeat the analysis of the binary stoichiometries and ask how many stoichiometries per element are there for the ternaries. This is shown in Figure 5. Here, also, the similarity of sulfides and selenides is clear. In addition, while there are similarities between the distributions of binary stoichiometries per element to the ternary distributions, there are also obvious differences. One might guess that there should be a correlation between the binary and ternary distributions. This is examined in Figure 6(a).

It is evident that the correlation between ternary and binary number of stoichiometries is not strong but the minimal number of ternary stoichiometries tends to grow with the number of

binary stoichiometries. We check this further in Figure 6(b), by comparing the number of ternary stoichiometries of $A_xB_yO_z$ to the product of stoichiometry numbers of A_xO_y and B_xO_y . The general trend obtained is an inverse correlation, i.e., as the product of the numbers of binary stoichiometries increases, the number of ternaries decreases. This trend can be explained by the following argument: when the two binaries are rich with stable compounds, the ternaries need to compete with more possibilities of binary phases, which makes the formation of a stable ternary more difficult. In Figure 6(b), this trend is highlighted for vanadium, the element with the most binary stoichiometries, but this pattern repeats itself for most elements. We analyze this behavior for the sulfides and selenides in the SI, similar trends are found but they are less pronounced due to a smaller number of known compounds.

Composition and Mendeleev Maps. The occurrence of each element in the binary and ternary compound lists has been counted and tabulated. The results are described in Figure 7. For the binary oxides a very prominent peak appears at $M = 85$, the Mendeleev number of silicon. It represents the 185 different silicon oxide structures types (s.t.) reported in the ICSD database for just a *single* stoichiometry, SiO_2 . Smaller peaks appear for $M = 51$ (titanium, 42 s.t., 14 stoichiometries, leading stoichiometry is TiO_2 with 14 s.t.), $M = 54$ (vanadium, 42 s.t., 18 stoichiometries, leading stoichiometry is VO_2 with 10 s.t.), $M = 56$ (tungsten, 24 s.t., 9 stoichiometries, leading stoichiometry is WO_3 with 13 s.t.), and $M = 45$ (uranium, 22 s.t., 9 stoichiometries, leading stoichiometries are UO_2 and U_3O_8

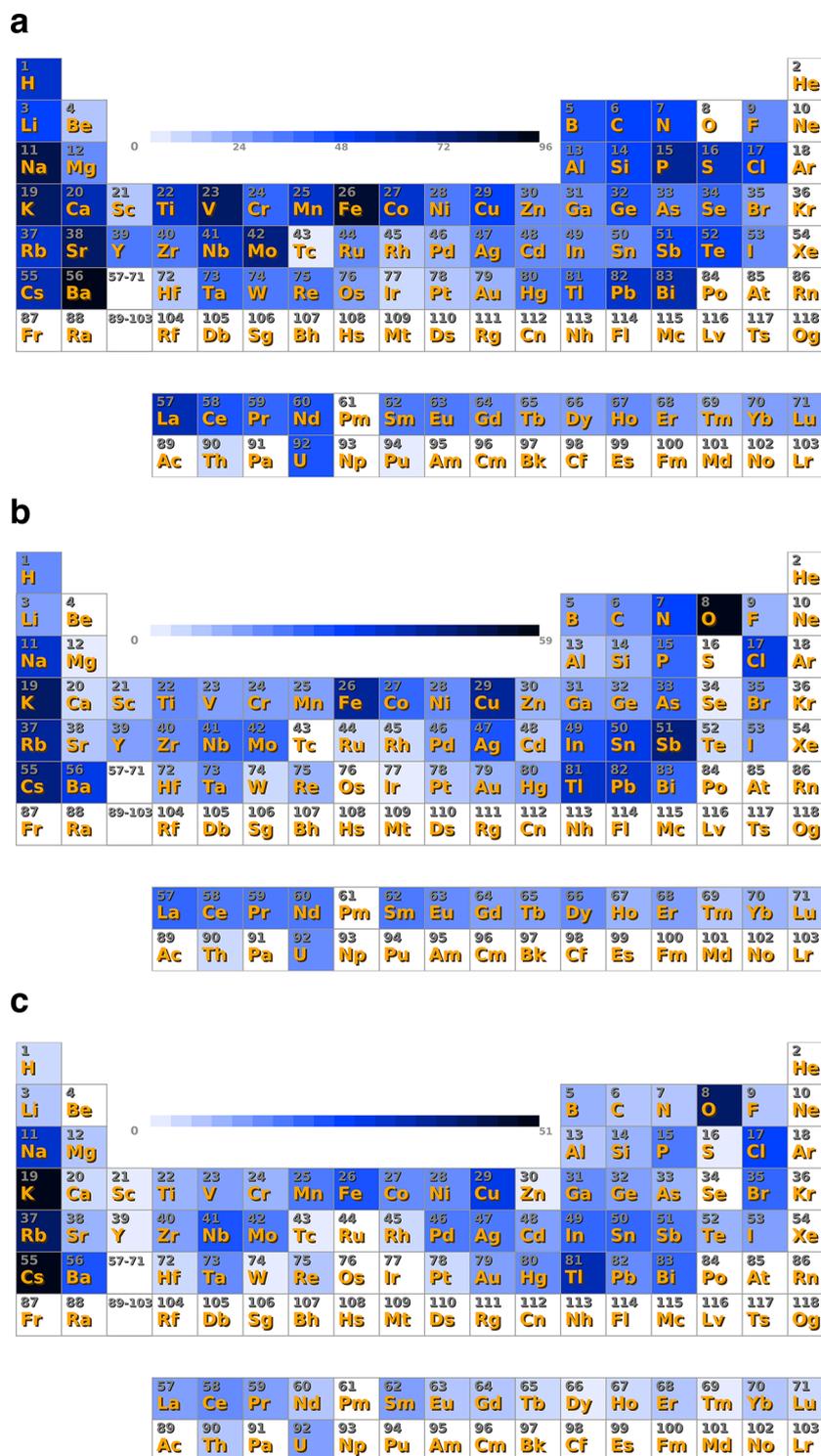


Figure 5. Ternary (a) oxide, (b) sulfide, and (c) selenide stoichiometries (number of different stoichiometries that include the respective element). The colors go from no stoichiometries (white) to the maximal number of stoichiometries (dark blue) which is different for each element, 96/59/51 for O/S/Se. High prevalence appears for the alkali metals in all three compound families. An additional island in the transition metals is much more pronounced in the oxides. The sulfides and selenides distributions are nearly identical, and shown high prevalence of oxygen containing ternaries.

with 6 s.t. each). Unlike the silicon peak which is composed of a single stoichiometry, the other leading peaks evidently include multiple stoichiometries, reflecting the different chemistry of those elements. These differences also carry over into the ternary oxide compounds involving those elements. For example, the stoichiometry distribution of silicon oxide ternaries is more tilted toward the silicon poor compounds compared to the

corresponding distributions of vanadium and titanium ternary oxides, as is shown in Figure S3 in the SI.

The distribution of the sulfides is generally much lower than that of the oxides, due to the much smaller total number of known binaries, but is also more uniformly structured. It has one major peak for $M = 76$ (zinc, 40 s.t., 2 stoichiometries, leading stoichiometry is ZnS with 39 s.t.), and quite a few smaller ones

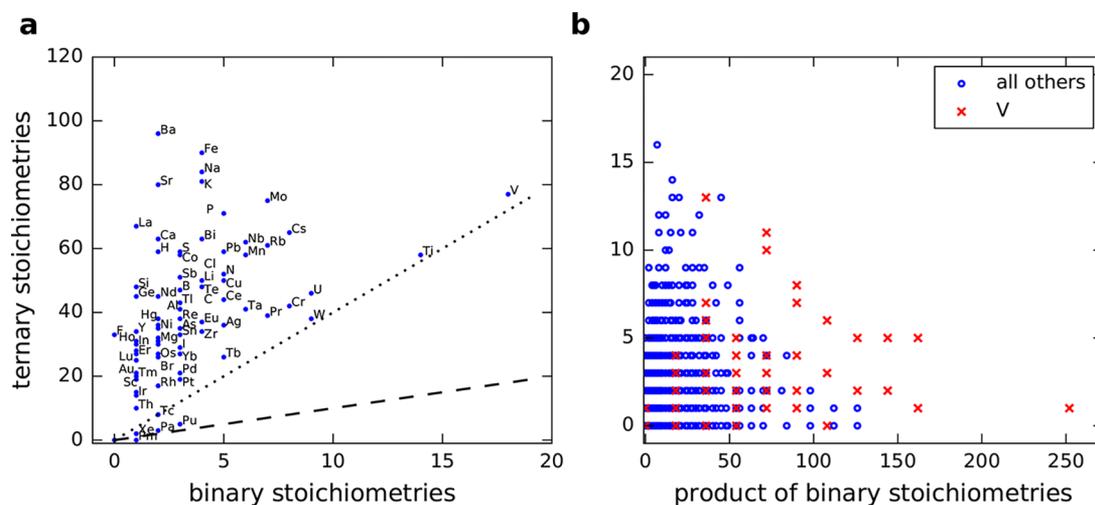


Figure 6. (a) Prevalence of ternary oxide stoichiometries per atom as a function of the prevalence of its binary stoichiometries. The dashed line marks perfect similarity ($y = x$), and the dotted line marks the ratio $y = 4x$. (b) Number of oxide ternary stoichiometries as a function of product of participating elements numbers of oxide binary stoichiometries. The data for vanadium is shown with red crosses, all the rest is shown with blue circles.

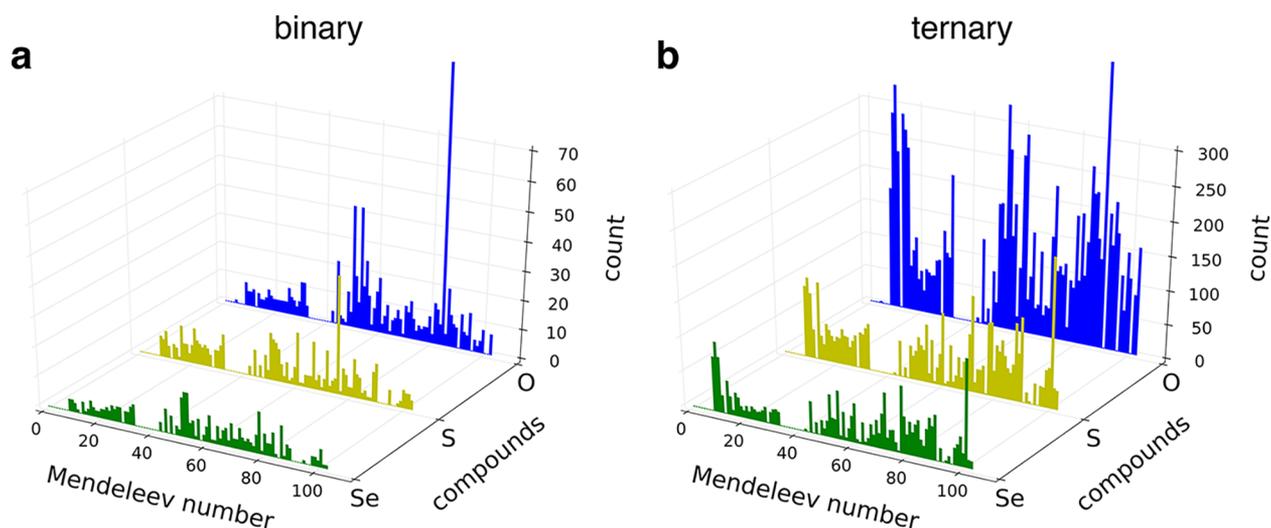


Figure 7. Distribution of the elements in (a) binary and (b) ternary compounds. The binary oxides exhibit a structures distribution with two prominent peaks. The distributions of the binary sulfides and selenides are less structured and more similar to each other. The distributions of the ternary compounds have higher, wider peaks than their binary counterparts. The relative differences between the oxide, sulfide and selenide distributions remain similar to the distributions of the binaries.

such as $M = 51$ (titanium, 16 s.t., 5 stoichiometries, leading stoichiometry is TiS_2 with 9 s.t.), $M = 61$ (iron, 18 s.t., 5 stoichiometries, leading stoichiometry is FeS with 6 s.t.), $M = 67$ (nickel, 16 s.t., 6 stoichiometries, leading stoichiometry is NiS_2 with 8 s.t.), $M = 90$ (phosphorus, 13 s.t., 8 stoichiometries, of which P_2S_7 , P_4S_9 , P_4S_6 , P_4S_5 and P_4S_3 have 2 s.t. each). The $M = 8$ –33 region also exhibits a minor concentration of participating elements. The selenides distribution is yet smaller than that of the sulfides, and even more uniform. Several peaks appear, $M = 51$ (titanium, 13 s.t., 9 stoichiometries, leading stoichiometry is TiSe with 3 s.t.), $M = 52$ (niobium, 15 s.t., 8 stoichiometries, leading stoichiometry is NbSe_2 with 8 s.t.), $M = 53$ (tantalum, 15 s.t., 4 stoichiometries, leading stoichiometry is TaSe_2 with 10 s.t.) and $M = 79$ (indium, 14 s.t., 5 stoichiometries, leading stoichiometry is In_2Se_3 with 6 s.t.). All distributions cover most of the elements except two obvious gaps, one at $M < 9$, which includes the noble gases and the two heaviest alkali metals, cesium and francium, and another at $34 \leq M \leq 42$ which

represents the heavy actinides. Another gap appears in the sulfide and selenide distributions at $91 \leq M \leq 97$, which reflects the rarity of polonium and astatine compounds and shows that the elements of the 6A column, except oxygen, do not coexist, in the known compounds, with each other or with the heavier halogen iodine.

The element occurrence distributions for the oxide, sulfide and selenide ternaries exhibit greater similarity than the corresponding binary distributions. The most apparent difference, however, is the most common component, which is sulfur, $M = 90$, in the oxides, but oxygen itself, $M = 101$, in the sulfides and selenides. The sulfide and selenide distributions are almost the same, except for generally lower numbers in the selenides (due to the smaller total number of compounds) and an apparent lower participation of the lanthanides $M = 17$ –35.

Mendeleev maps for the ternaries are shown in Figures 8–10. Figure 8 shows the cumulated maps for all stoichiometries reported for the respective ternary family. They reflect the same

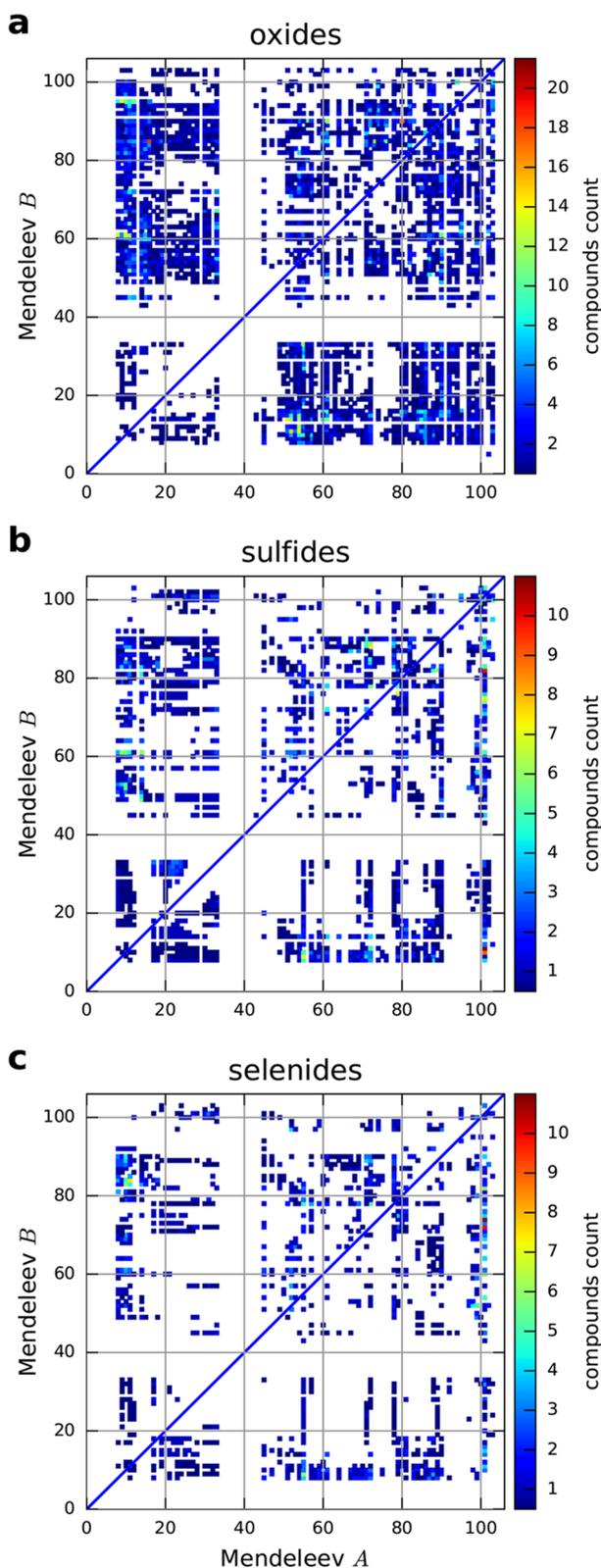


Figure 8. Mendeleev maps of ternary (a) oxide $A_xB_yO_z$, (b) sulfide $A_xB_yS_z$ and (c) selenide $A_xB_ySe_z$ compounds. It is assumed that $x \geq y$ with the x -axis indicating M_A and the y -axis M_B . If the stoichiometry is such that $x = y$, the compound is counted as $0.5 A_xB_yO_z + 0.5 B_xA_yO_z$. A color scheme is used to represent the compound count for each composition, blue means the minimal number (one) and green means the maximal number, which is different for each element.

major gaps as the binary distributions. The maps show that most of the reported compositions are represented by one or two compounds with just a few hotspots that include up to 20 compounds in the oxides and 10 compounds in the sulfides and selenides. The oxides map is obviously denser, reflecting the much richer, currently known, chemistry of the oxides compared to the other two elements. The chemistry becomes more constrained as we proceed down the periodic table column from oxygen to sulfur and then to selenium.

Next, we examine maps of specific stoichiometries. Maps of a few notable oxide stoichiometries and their leading structure types are shown in Figure 9. These maps reflect the dominant features of the full oxide ternaries map (Figure 8), but with significant new additional gaps of absent compounds. These gaps are naturally wider for less prevalent stoichiometries, i.e., the map of the most prevalent stoichiometry, $A_1B_1O_3$, is denser than the three other maps in Figure 9. Different structure types in all stoichiometries tend to accumulate at well-defined regions of the map. The separation between them is not perfect, but is similar to that exhibited by the classical Pettifor maps for binary structure types.^{14,15} A similar picture is obtained for the sulfide and selenide structure types, although more sparse (Figure 10). It is interesting to note that the maps of, e.g., $A_1B_2C_4$ ($C = O, S, Se$), show similar patterns in the map for oxides (Figure 9) and sulfides/selenides (Figure 10)—suggesting that similar elements tend to form this stoichiometry. In the same manner, the 2:1:1 stoichiometry shows very similar patterns in oxides, sulfides and selenides (see also Figure S6 in the SI).

Symmetries. The distribution of the compounds and structure types among the 14 Bravais lattices is presented in Table 3 and Figure 11. It is interesting to note that in all six cases (binary and ternary oxides, sulfides and selenides) the distribution is double peaked, with the majority of the compounds belonging to the monoclinic and orthorhombic primitive lattices, and a smaller local maximum at the hexagonal and tetragonal lattices. All distributions exhibit a local minimum for the orthorhombic face and body centered lattices. The high symmetry cubic lattices are also relatively rare. This reflects the complex spatial arrangement of the compound forming electrons of oxygen, sulfur and selenium, which does not favor the high symmetry cubic structures or the densely packed face and body centered orthorhombic structures.

Figure 12 shows a more detailed distribution of the compounds among the different space groups. The binary compounds show a distinct seesaw structure, with a few local peaks near the highest symmetry groups of each crystal system. The corresponding ternary distributions have three sharp peaks in the triclinic, monoclinic and orthorhombic systems, and much smaller peaks in the hexagonal and cubic groups. It is interesting to note that the three compound families, exhibit distributions of very similar structure. The oxide distributions are the densest, simply due to the existence of more oxide compounds in the database, and become sparser in the sulfide and selenide cases. The compounds of all these families are distributed among a rather limited number of space groups, with most space groups represented by just a single compound or not at all.

Unit Cell Size. The distributions of unit cell sizes (i.e., the number of atoms per unit cell) for the six compound families we discuss are shown in Figure 13. All of these distributions have strong dense peaks at small cell sizes and decay sharply at sizes above a few tens of atoms. However, the details of the distributions differ quite significantly from group to group. Among the binaries, the oxides exhibit the highest and widest

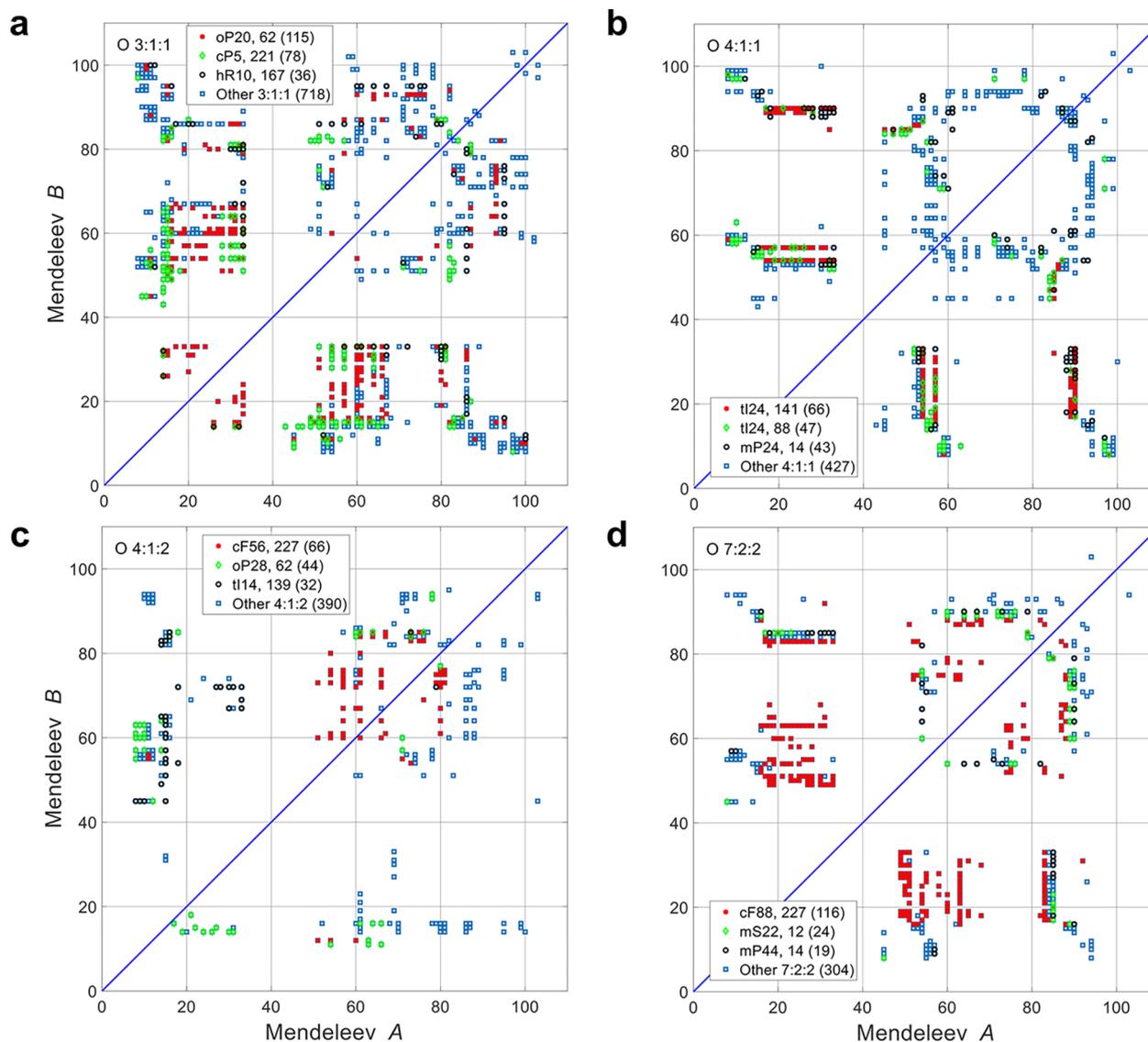


Figure 9. Three leading structure types in each of the four leading stoichiometries in oxide ternaries: (a) $A_1B_1O_3$, (b) $A_1B_1O_4$, (c) $A_1B_2O_4$, and (d) $A_2B_2O_7$. The legend box appears at a region with no data points. The number in parentheses is the number of compounds for this structure type; for “Other”, refers to the total number of compounds with this stoichiometry.

peak with its maximum of 102 oxide binary compounds located at 12 atoms per cell. 90% of the oxide binaries have less than 108 atoms in the unit cell and 50% of them have less than 24 atoms. The sulfides distribution has a lower and narrower peak of 70 compounds at 8 atoms. The distribution of the selenides has a still lower peak of 60 compounds at 8 atoms. The fact that oxygen has a peak at 12 atoms in the unit cell and not at 8 as the sulfides and selenides, is related to the fact that binary oxides prefer the AO_2 stoichiometry over AO , where as both sulfides and selenides prefer the 1:1 stoichiometry over 1:2. This is probably related to the different chemistry of oxygen vs sulfur and selenium. Additional computational analysis would be required to fully understand the effect of the different chemistry on the stoichiometry and number of atoms. Detailed data for these dense parts of the distributions is tabulated in Table S12 the [Supporting Information](#) (SI). The oxides distribution exhibits the longest tail of the binaries, with the largest binary oxide unit cell including 576 atoms. The largest binary sulfide and selenide unit cells include 376 and 160 atoms, respectively.

The distributions of the ternary compounds have higher, wider peaks and longer tails than their binary counterparts. The relative differences between the oxide, sulfide and selenide distributions remain similar to the distributions of the binaries. The oxide ternaries exhibit a high and wide peak. Its maximum of 465 compounds is located at 24 atoms per cell, and 90% of the compounds have less than 92 atoms in the unit cell and 50% of the compounds have less than 32 atoms. As in the binary case, the distribution of the ternary sulfides has a lower and narrower peak than the oxides, where the maximum of 190 compounds at 28 atoms and 90% of the compounds have less than 72 atoms in the unit cell. The distribution of the selenides has a still lower and narrower peak, where the corresponding numbers are 130 compounds at 28 atoms and 90% of the compounds having less than 28 atoms in the unit cell. Detailed data for these dense parts of the distributions is shown in Table S13 of the SI. The ternary oxides distribution exhibits the longest tail of the three types, with the largest oxide ternary unit cell having 1080 atoms. The largest ternary sulfide and selenide unit cells have 736 and 756 atoms, respectively.

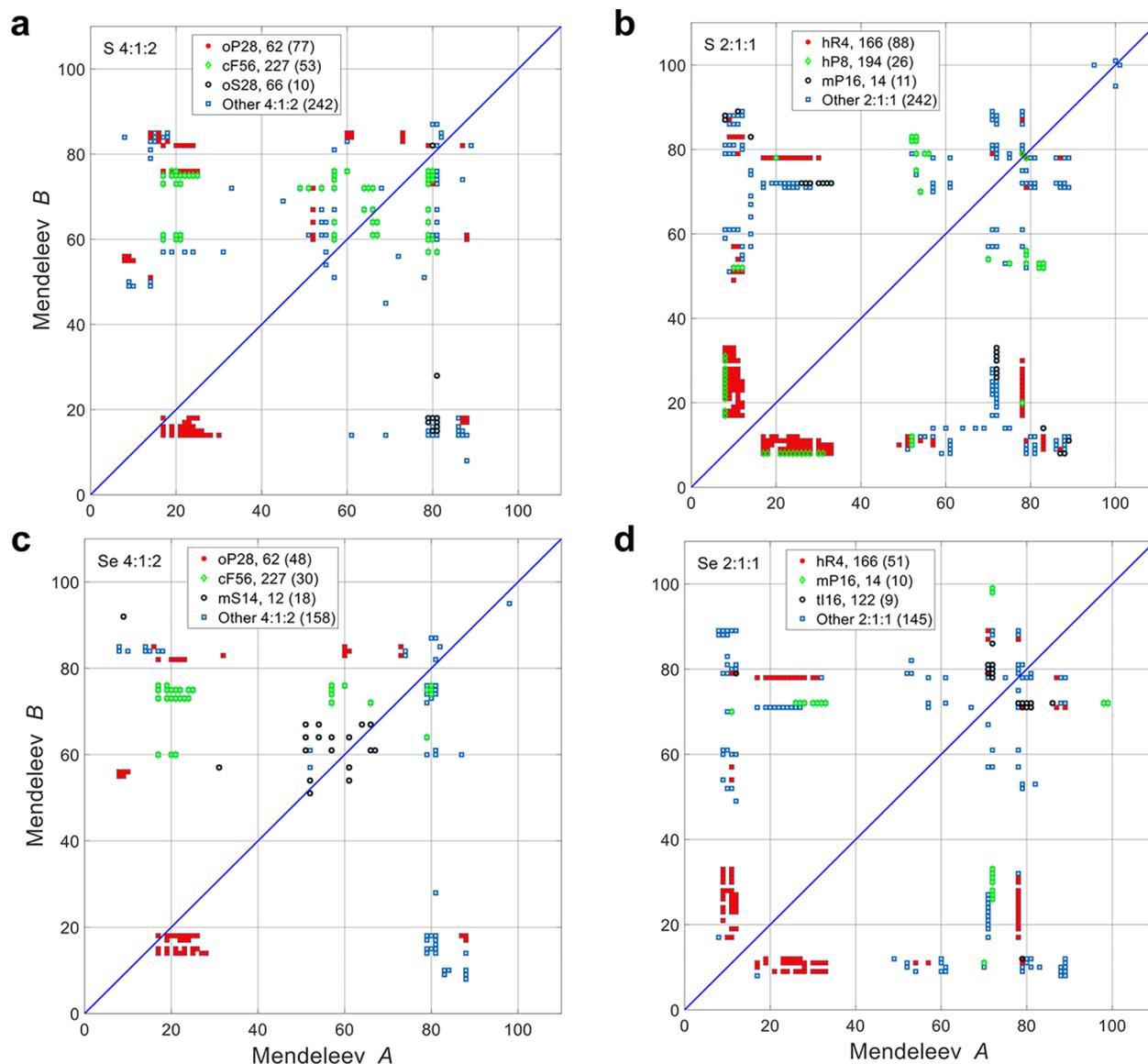


Figure 10. Three leading structure types in each of the two leading stoichiometries in sulfur and selenium ternaries: (a) $A_1B_2S_4$, (b) $A_1B_1S_2$, (c) $A_1B_2Se_4$, and (d) $A_1B_1Se_2$. The number in parentheses is the number of compounds for this structure type; for “Other”, refers to the total number of compounds with this stoichiometry.

Table 3. Distribution of the Oxide, Sulfide and Selenide Compounds and Structure Types among the 14 Bravais Lattices

	binary compounds			binary structure types			binary compounds per structure type			ternary compounds			ternary structure types			ternary compounds per structure type		
	O	S	Se	O	S	Se	O	S	Se	O	S	Se	O	S	Se	O	S	Se
aP	51	13	5	39	12	5	1.3	1.1	1	378	79	60	219	56	39	1.7	1.4	1.5
mP	82	54	31	62	36	20	1.3	1.5	1.6	918	318	198	363	166	109	2.5	1.9	1.8
mS	88	31	22	58	21	15	1.5	1.5	1.5	672	251	170	292	117	77	2.3	2.1	2.2
oP	123	82	48	81	37	30	1.5	2.2	1.6	950	481	266	373	139	105	2.5	3.5	2.5
oS	39	24	11	36	19	9	1.1	1.3	1.2	334	84	60	133	40	25	2.5	2.1	2.4
oF	11	7	11	10	6	4	1.1	1.2	2.8	51	32	23	28	14	8	1.8	2.3	2.9
oI	22	5	2	20	4	2	1.1	1.25	1	89	36	27	39	15	12	2.3	2.4	2.25
tI	41	20	10	31	17	8	1.3	1.2	1.25	418	80	72	101	34	23	4.1	2.4	3.1
tP	78	27	28	48	13	16	1.6	2.1	1.75	239	73	52	107	39	26	2.2	1.9	2.0
hP	94	87	66	62	50	32	1.5	1.7	2.1	435	224	103	198	75	41	2.2	3.0	2.5
hR	40	44	20	30	33	15	1.3	1.3	1.3	420	230	133	123	49	33	3.4	4.7	4.0
cP	42	22	20	21	6	4	2.0	3.7	5.0	187	58	43	45	18	13	4.2	3.2	3.3
cF	75	65	48	19	10	6	3.9	6.5	8.0	251	80	43	27	17	7	9.3	4.7	3.9
cI	58	14	10	21	6	2	2.8	2.3	5.0	92	15	6	30	5	3	3.1	3.0	2.0

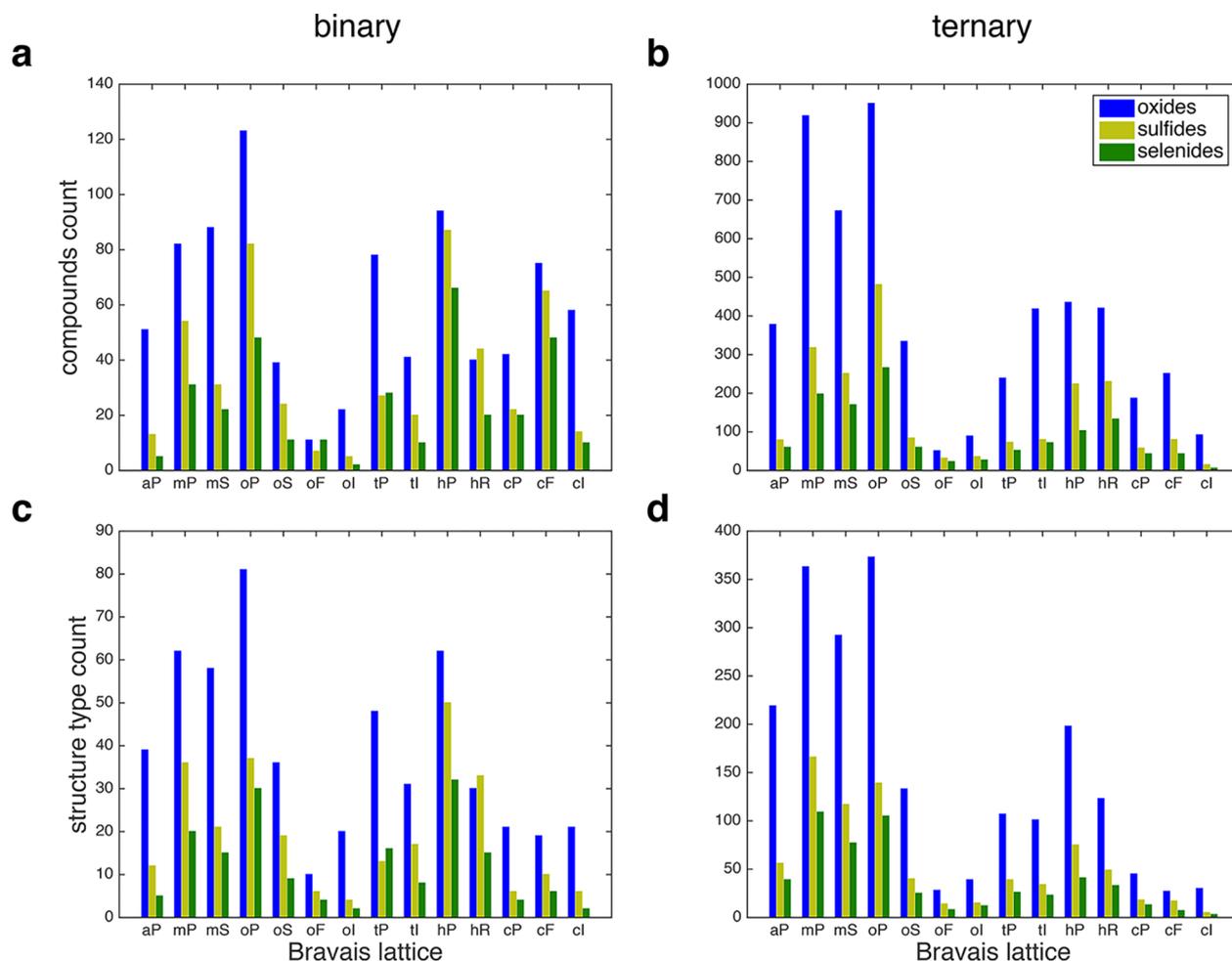


Figure 11. Number of compounds (a and b) and structure types (c and d) for each Bravais lattice. Binaries are on the left (a and c) and ternaries on the right (b and d). Oxides are shown in blue, sulfides in light green and selenides in darker green. All six distributions (binary and ternary oxides, sulfides and selenides) are double peaked with a local minimum for the orthorhombic face and body centered lattices. The high symmetry cubic lattices are also relatively rare. This reflects the complex spatial arrangement of the compound forming electrons of the 6A elements, which does not favor the high symmetry of these structures.

It should be noted that large unit cells, within the tails of all distributions, tend to have very few representatives, with just one compound with a given unit cell size in most cases. Notable exceptions are local peaks near 80 atoms per unit cell in the binary distributions and near 200 atoms per unit cell in the ternary distributions. The oxide distributions exhibit additional peaks, near 300 atoms per unit cell for the binaries and near 600 atoms per unit cell for the ternaries. These minor peaks may indicate preferable arrangements of cluster-based structures.

SUMMARY

We present a comprehensive analysis of the statistics of the binary and ternary compounds of oxygen, sulfur and selenium. This analysis and the visualization tools presented here are valuable to finding trends as well as exceptions and peculiar phenomena.

Oxygen has a higher electronegativity (3.44) than sulfur (2.58) and selenium (2.55), which are similar to each other. Therefore, one can expect that oxygen will form compounds with a stronger ionic character. Oxygen is 1000 times more abundant than sulfur, and more than 10^6 times than selenium;³¹ however, it has less than two times the number of binary compounds compared to sulfur and 2.5 that of selenium. Hence,

the abundance of those elements plays a little role in the relative numbers of their known compounds. These important differences are reflected in our analysis by the significantly larger fraction of oxygen rich compounds compared to those that are sulfur or selenium rich. Structure type classification also shows that there is little overlap between the oxygen structure types to sulfur or selenium structure types, while sulfur and selenium present a much higher overlap. The gaps in these overlaps, especially between the sulfides and selenides, indicate that favorable candidates for new compounds may be obtained by simple element substitution in the corresponding structures. In particular, structures that are significantly more common in one family, such as KrF_2 in the oxides, may be good candidates for new compounds in another. Comparison of these three 6A elements binary and ternary compounds shows significant differences but also some similarities in the symmetry distributions among the various Bravais lattices and their corresponding space groups. In particular, the majority of structure types in all three families have a few or single compound realizations. This prevalence of unique structure types suggests a ripe field for identification of currently unknown compounds, by substitution of elements of similar chemical characteristics. In addition, the analysis of the distribution of

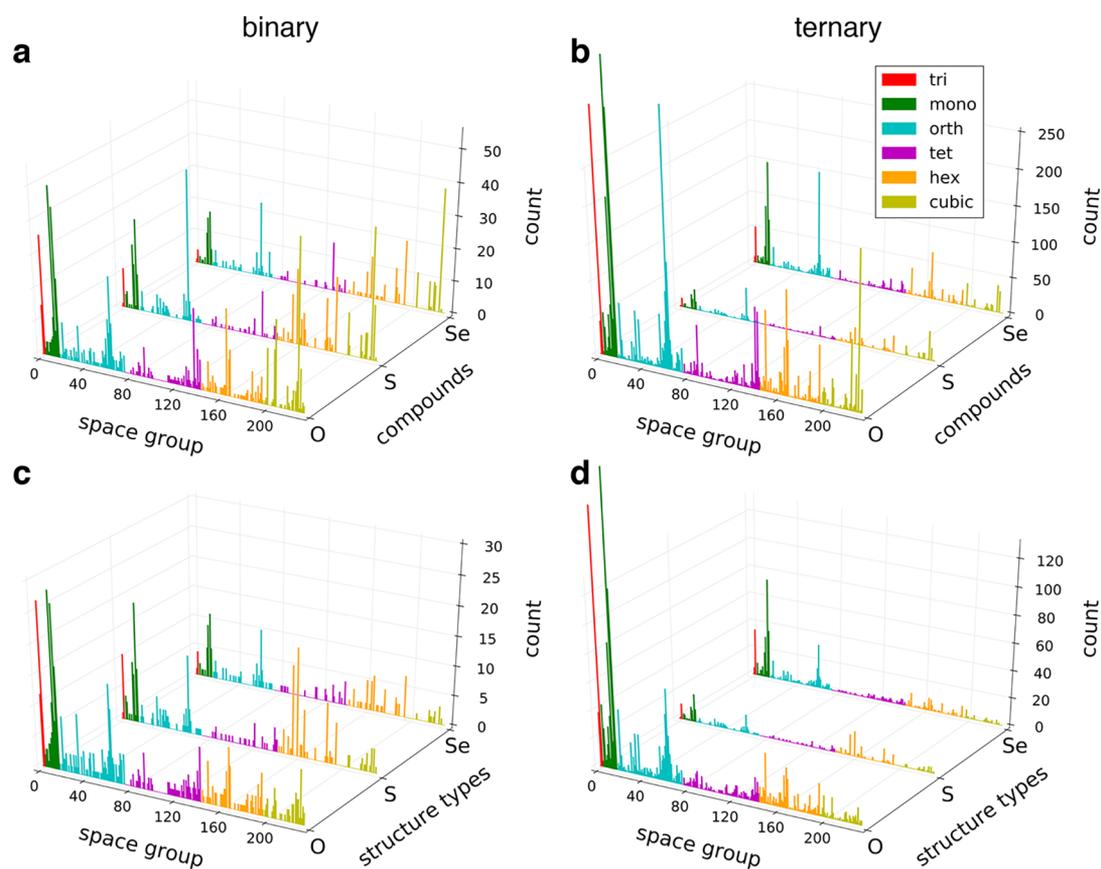


Figure 12. Distribution of compounds and structure types among the 230 space groups. Binaries are on the left (a and c) and ternaries on the right (b and d). Compounds are depicted on the top (a and b) and structure types on the bottom (c and d).

known compounds among symmetry space groups and, in particular, their apparent concentration in specific hot spots of this symmetry space may be serve as a useful insight for searches of potential new compounds.

An important observation is the existence of different gaps (missing stoichiometries) in the stoichiometry distribution of the oxide binary compounds compared to the sulfides and selenides (Figures 2 and 3, and Table S7 in the SI). Stoichiometries such as 5:7 appear in the oxides but are missing in the sulfides and selenides. More rare are nonoverlapping gaps between the selenides and sulfides, e.g., 6:1 and 5:7. These should be prime candidates for new compounds by element substitution between the two families. Future work would be directed at exploiting these discrepancies to search for new compounds within different subsets of those compound families.

Specific elements tend to present very different stoichiometry distributions, for example, silicon forms only one oxide stoichiometry (SiO_2) while transition metals such as titanium and vanadium present 14 and 18 different stoichiometries, respectively. These differences clearly reflect the different chemistry of those elements, while the large number of reported SiO_2 structures might reflect research bias into silicon compounds.

Another important finding is that there is an inverse correlation between the number of ternary stoichiometries to the product of binary stoichiometries of participating elements. This can be caused by the fact that there are too many binary phases and hence it becomes difficult to create a stable ternary that competes with all of them.

A Mendeleev analysis of the common structure types of these families shows accumulation of different structures at well-defined regions of their respective maps, similar to the well-known Pettifor maps of binary structure types. Furthermore, at least for some of the stoichiometries, similarity of the maps for a given stoichiometry is demonstrated across all three elements. These maps should therefore prove useful for predictive purposes regarding the existence of yet unknown compounds of the corresponding structure types. Future work will be directed at exploiting identified nonoverlapping gaps in the Mendeleev maps for a directed search of new compounds in these families. Complementary properties (e.g., partial charges, bond analysis, electronic properties) should be incorporated in the analysis to reveal additional insights of the aforementioned trends among the three elements.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.inorgchem.7b02462.

- (1) Discussion of structures subtypes;
- (2) additional ternary stoichiometry triangles;
- (3) prevalence of structure types among the oxide, sulfide and selenide binary and ternary compounds;
- (4) prevalence of stoichiometries tables;
- (5) correlation between ternary and binary stoichiometries for sulfides and selenides;
- (6) prevalence of unit cell sizes tables; and
- (7) additional Mendeleev plots (PDF)

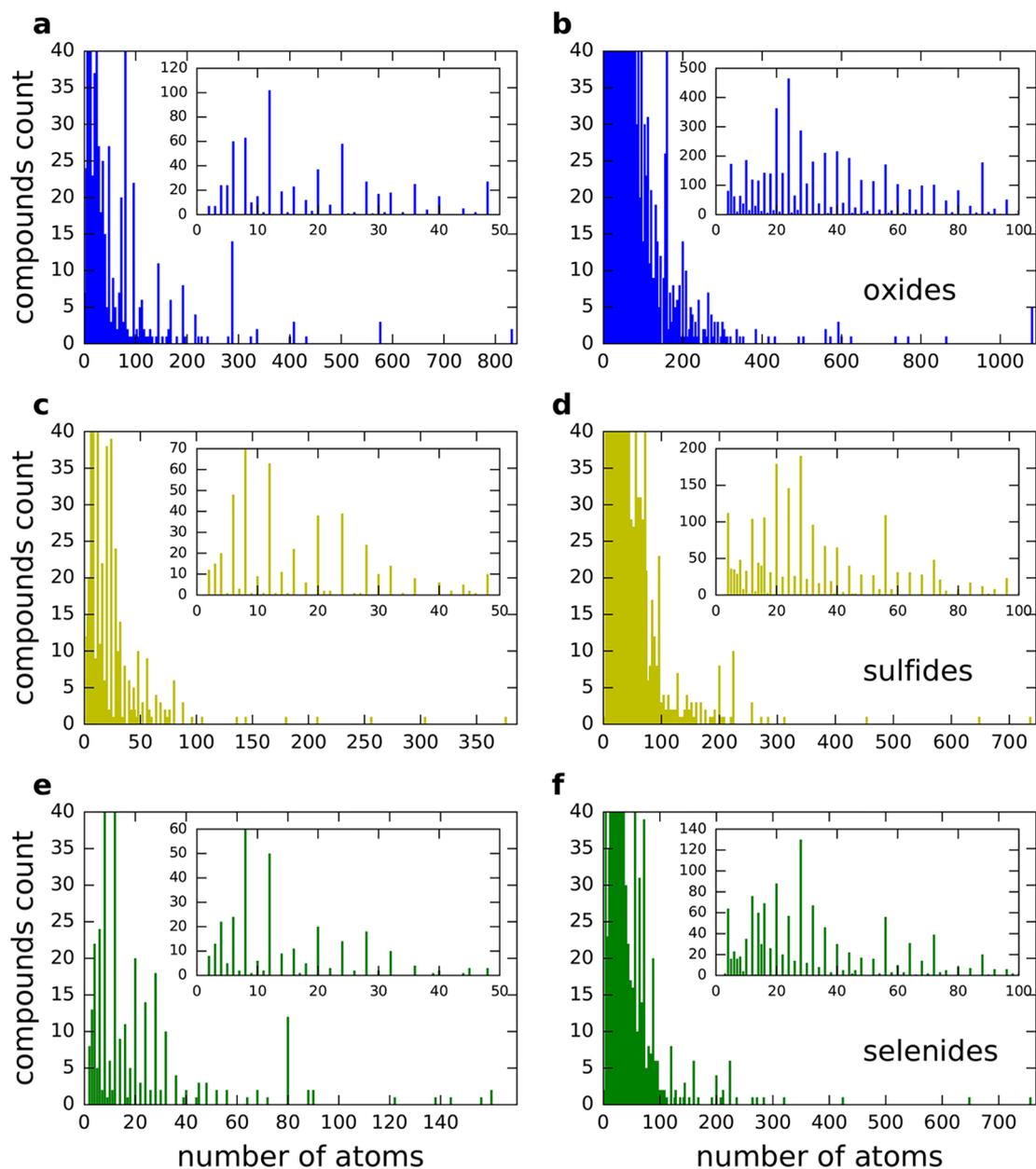


Figure 13. Unit cell size distribution for all compounds. Binaries are on the left (a, c and e) and ternaries on the right (b, d and f). Oxides are at the top (a and b), sulfides in the middle (c and d) and selenides at the bottom (e and f). The insets show the compounds with up to 50 atoms per unit cell in each case. All distributions exhibit long tails of rare very large unit cells which extend much further in the oxides. The dense cores of the distributions reflects the higher prevalence of oxides and are very similar for the sulfides and selenides.

AUTHOR INFORMATION

Corresponding Author

*E-mail: amirmatan@post.tau.ac.il

ORCID

Amir Natan: 0000-0003-4517-5667

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

AN acknowledges financial support from the Israeli National Nanotechnology Initiative (INNI, FTA project). AN and OL acknowledge financial support from the Pazi foundation. CO acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGF1106401.

SC and CO acknowledge support by DOD-ONR (N00014-17-1-2090) and DOE (DE-AC02-05CH11231), specifically the Basic Energy Sciences program under Grant # EDCBEE. SC acknowledges support by the Alexander von Humboldt-Foundation. The authors thank Drs. Cormac Toher, David Hicks, Natalio Mingo, Luca Ghiringhelli, and Jesús Carrete for various technical discussions.

REFERENCES

- (1) Eranna, G.; Joshi, B. C.; Runthala, D. P.; Gupta, R. P. Oxide Materials for Development of Integrated Gas Sensors—A Comprehensive Review. *Crit. Rev. Solid State Mater. Sci.* **2004**, *29*, 111–188.
- (2) Fortunato, E.; Barquinha, P.; Martins, R. Oxide Semiconductor Thin Film Transistors: A Review of Recent Advances. *Adv. Mater.* **2012**, *24*, 2945–2986.

- (3) Tsepis, E. V.; Kharton, V. V. Electrode Materials and Reaction Mechanisms in Solid Oxide Fuel Cells: A Brief Review. *J. Solid State Electrochem.* **2008**, *12*, 1039–1060.
- (4) Jiang, T.; Ozin, G. A. New Directions in Tin Sulfide Materials Chemistry. *J. Mater. Chem.* **1998**, *8*, 1099–1108.
- (5) Panda, P. K. Review: Environmental Friendly Lead-Free Piezoelectric Materials. *J. Mater. Sci.* **2009**, *44*, S049–S062.
- (6) Shi, J.; Cerqueira, T. F. T.; Cui, W.; Nogueira, F.; Botti, S.; Marques, M. A. L. High-Throughput Search of Ternary Chalcogenides for *p*-Type Transparent Electrodes. *Sci. Rep.* **2017**, *7*, 43179.
- (7) Lorenz, M.; et al. The 2016 Oxide Electronic Materials and Oxide Interfaces Roadmap. *J. Phys. D: Appl. Phys.* **2016**, *49*, 433001.
- (8) Rühle, S.; Anderson, A. Y.; Barad, H.-N.; Kupfer, B.; Bouhadana, Y.; Rosh-Hodesh, E.; Zaban, A. All-Oxide Photovoltaics. *J. Phys. Chem. Lett.* **2012**, *3*, 3755–3764.
- (9) Curtarolo, S.; Hart, G. L. W.; Buongiorno Nardelli, M.; Mingo, N.; Sanvito, S.; Levy, O. The High-Throughput Highway to Computational Materials Design. *Nat. Mater.* **2013**, *12*, 191–201.
- (10) Curtarolo, S.; Setyawan, W.; Hart, G. L. W.; Jahnátek, M.; Chepulska, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An Automatic Framework for High-Throughput Materials Discovery. *Comput. Mater. Sci.* **2012**, *58*, 218–226.
- (11) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314–319.
- (12) Dshemuchadse, J.; Steurer, W. Some Statistics on Intermetallic Compounds. *Inorg. Chem.* **2015**, *54*, 1120–1128.
- (13) Dshemuchadse, J.; Steurer, W. More Statistics on Intermetallic Compounds – Ternary Phases. *Acta Crystallogr., Sect. A: Found. Adv.* **2015**, *71*, 335–345.
- (14) Pettifor, D. G. A Chemical Scale for Crystal-Structure Maps. *Solid State Commun.* **1984**, *51*, 31–34.
- (15) Pettifor, D. G. The Structures of Binary Compounds. I. Phenomenological Structure Maps. *J. Phys. C: Solid State Phys.* **1986**, *19*, 285–313.
- (16) FIZ Karlsruhe, Inorganic Crystal Structure Database. <http://icsd.fiz-karlsruhe.de/icsd/>.
- (17) Curtarolo, S.; Setyawan, W.; Hart, G. L. W.; Jahnátek, M.; Chepulska, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An Automatic Framework for High-Throughput Materials Discovery. *Comput. Mater. Sci.* **2012**, *58*, 218–226.
- (18) Setyawan, W.; Curtarolo, S. High-Throughput Electronic Band Structure Calculations: Challenges and Tools. *Comput. Mater. Sci.* **2010**, *49*, 299–312.
- (19) Yang, K.; Oses, C.; Curtarolo, S. Modeling Off-Stoichiometry Materials with a High-Throughput *Ab-Initio* Approach. *Chem. Mater.* **2016**, *28*, 6484–6492.
- (20) Levy, O.; Jahnátek, M.; Chepulska, R. V.; Hart, G. L. W.; Curtarolo, S. Ordered Structures in Rhenium Binary Alloys from First-Principles Calculations. *J. Am. Chem. Soc.* **2011**, *133*, 158–163.
- (21) Levy, O.; Hart, G. L. W.; Curtarolo, S. Structure Maps for hcp Metals from First-Principles Calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *81*, 174106.
- (22) Levy, O.; Hart, G. L. W.; Curtarolo, S. Uncovering Compounds by Synergy of Cluster Expansion and High-Throughput Methods. *J. Am. Chem. Soc.* **2010**, *132*, 4830–4833.
- (23) Hart, G. L. W.; Curtarolo, S.; Massalski, T. B.; Levy, O. 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* **2013**, *3*, 041035.
- (24) Mehl, M. J.; Hicks, D.; Toher, C.; Levy, O.; Hanson, R. M.; Hart, G. L. W.; Curtarolo, S. The AFLOW Library of Crystallographic Prototypes: Part 1. *Comput. Mater. Sci.* **2017**, *136*, S1–S828.
- (25) Supka, A. R.; Lyons, T. E.; Liyanage, L. S. I.; D'Amico, P.; Al Rahal Al Orabi, R.; Mahatara, S.; Gopal, P.; Toher, C.; Ceresoli, D.; Calzolari, A.; Curtarolo, S.; Buongiorno Nardelli, M.; Fornari, M. AFLOW π : A Minimalist Approach to High-Throughput *Ab Initio* Calculations Including the Generation of Tight-Binding Hamiltonians. *Comput. Mater. Sci.* **2017**, *136*, 76–84.
- (26) Hahn, T., Ed.; *International Tables of Crystallography. Vol. A: Space-Group Symmetry*; Kluwer Academic Publishers, International Union of Crystallography: Chester, England, 2002.
- (27) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L. W.; Sanvito, S.; Buongiorno Nardelli, M.; Mingo, N.; Levy, O. AFLOWLIB.ORG: A Distributed Materials Properties Repository from High-Throughput *Ab Initio* Calculations. *Comput. Mater. Sci.* **2012**, *58*, 227–235.
- (28) Taylor, R. H.; Rose, F.; Toher, C.; Levy, O.; Yang, K.; Buongiorno Nardelli, M.; Curtarolo, S. A RESTful API for Exchanging Materials Data in the AFLOWLIB.org Consortium. *Comput. Mater. Sci.* **2014**, *93*, 178–192.
- (29) Calderon, C. E.; Plata, J. J.; Toher, C.; Oses, C.; Levy, O.; Fornari, M.; Natan, A.; Mehl, M. J.; Hart, G. L. W.; Buongiorno Nardelli, M.; Curtarolo, S. The AFLOW Standard for High-Throughput Materials Science Calculations. *Comput. Mater. Sci.* **2015**, *108* (Part A), 233–238.
- (30) Rose, F.; Toher, C.; Gossett, E.; Oses, C.; Buongiorno Nardelli, M.; Fornari, M.; Curtarolo, S. AFLUX: The LUX Materials Search API for the AFLOW Data Repositories. *Comput. Mater. Sci.* **2017**, *137*, 362–370.
- (31) Wedepohl, K. H. The Composition of the Continental Crust. *Geochim. Cosmochim. Acta* **1995**, *59*, 1217–1232.
- (32) Villars, P.; Cenzual, K. *Pearson's Crystal Data—Crystal Structure Database for Inorganic Compounds*; ASM International: Materials Park, OH, 2013.
- (33) Villars, P.; Cenzual, K.; Daams, J. L. C.; Hulliger, F.; Massalski, T. B.; Okamoto, H.; Osaki, K.; Prince, A.; Iwata, S. *Crystal Impact, Pauling File. Inorganic Materials Database and Design System, Binaries ed.*; ASM International: Metal Park, OH, 2003.
- (34) Lima-de-Faria, J.; Hellner, E.; Liebau, F.; Makovicky, E.; Parthé, E. Nomenclature of Inorganic Structure Types. Report of the International Union of Crystallography Commission on Crystallographic Nomenclature Subcommittee on the Nomenclature of Inorganic Structure Types. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1990**, *46*, 1–11.