Taylor & Francis
Taylor & Francis Group

Science and Technology of Advanced Materials: Methods

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tstm20

Categorization of inorganic crystal structures by Delaunay tetrahedralization

Yoyo Hinuma & Yukari Katsura

To cite this article: Yoyo Hinuma & Yukari Katsura (2022) Categorization of inorganic crystal structures by Delaunay tetrahedralization, Science and Technology of Advanced Materials: Methods, 2:1, 75-83, DOI: 10.1080/27660400.2022.2059336

To link to this article: https://doi.org/10.1080/27660400.2022.2059336

SCIENCE AND TECHNOLOGY OF ADVANCED MATERIALS: METHODS
2022, VOL. 2, NO. 1, 75–83
https://doi.org/10.1080/27660400.2022.2059336

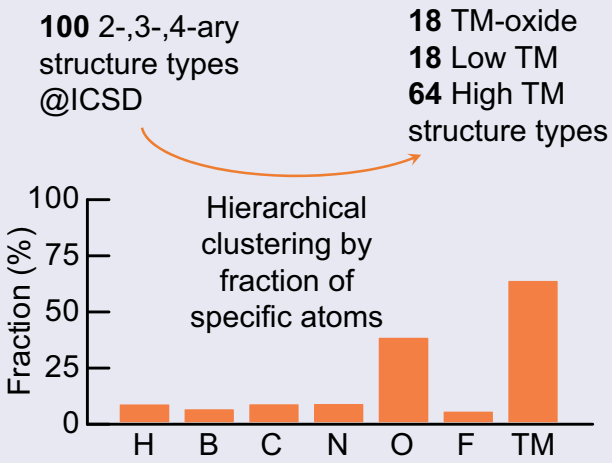
Categorization of inorganic crystal structures by Delaunay tetrahedralization

Yoyo Hinuma [a] and Yukari Katsura [b,c,d]

[a]Research Institute of Electrochemical Energy, Department of Energy and Environment, National Institute of Advanced Industrial Science and Technology (AIST), Ikeda, Japan; [b]Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science (NIMS), Tsukuba, Japan; [c]Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan; [d]Center for Advanced Intelligence Project, RIKEN, Chuo-ku, Japan

ABSTRACT

Crystal structure prediction is demanding in two senses: this is a difficult problem, both experimentally and computationally, yet we want to find new materials with new properties that take unknown crystal structures. Using statistical information on the local environment of already reported stable crystal structures would be helpful when, for instance, finding new crystal structures through visually adding atoms sequentially to a cluster of atoms, which is an intuitive procedure that does not require substantial computational resources. This study conducted hierarchical classification of frequently appearing structure types in the ICSD database by comparing the occurrence ratios of H, B, C, N, O, F, and transition metals (TM) in binary, ternary, and quaternary crystals. The bond angle and solid angle distributions of Delaunay polyhedra with atoms as vertices were considered as descriptors of the local environment for each structure type. Hierarchical clustering of structure types based on bond angle or solid angle distributions revealed that the structure types can be broadly categorized into those favoring TM oxides, having low TM occurrence, and having high TM occurrence. Our results indicate that the anion, or lack of, plays a key role in determining the crystal structure type, and structure types strongly preferred by TM oxides can accommodate a broad range of angles.



ARTICLE HISTORY
Received 19 February 2022
Revised 14 March 2022
Accepted 25 March 2022

KEYWORDS
Bond angle; crystal structure type; Delaunay polyhedra; hierarchical classification; solid angle

1. Introduction

Prediction of crystal structures from the constituent atoms and their ratios is still a big frontier of solid-state inorganic chemistry. The diversity of crystal structures results in a huge uncertainty in the chemists' guesses of the choices of possible structures from information on constituent elements only. There also remains many unreported crystal structures unknown to us.

Oganov et al. conceived a fully-automated method to predict crystal structures named USPEX [1–3], to attack this frontier. This method employs evolutionary algorithms to generate candidate structures and searches the most stable structure by relaxing and calculating their formation energies, for example by the embedded atom model (EAM) or even density functional theory (DFT) calculations. Such efforts are computationally expensive because formation energies of many crystal structures need to be evaluated, which could reach the order of 10^3 to 10^4. Necessary parameters, such as the number of individuals per generation and number of generations, must be obtained from experience. As a result, this technique cannot be casually used to search reasonable crystal

CONTACT Yoyo Hinuma y.hinuma@aist.go.jp Research Institute of Electrochemical Energy, Department of Energy and Environment, National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31 Midorigaoka, Ikeda, Osaka 563-8577, Japan
Supplemental data for this article can be accessed here.

structures for arbitrary materials. As another effort, Wang et al. developed the CALYPSO code that uses particle swarm optimization for determination of energetically stable/metastable crystal structures of materials at given chemical compositions and external conditions such as pressure [4,5]. The CrySPY code can use Bayesian optimization and Look Ahead based on Quadratic Approximation in addition to random search and evolutionary algorithm for crystal structure prediction [6].

We believe that there are needs for alternative tools to guess and identify reasonable inorganic structures that are way less computationally intensive, thus less time-consuming, and easily accessible by a broader group of inorganic chemists. Guidance based on statistical information on the local environment of already reported stable crystal structures would be helpful to avoid absurd structures. Real-time display of such information allows visual addition of atoms, one by one, to build potentially stable crystals while ensuring that all bond lengths and angles are in their reasonable ranges.

Several descriptors have been developed to describe the local environments in inorganic crystal structures. The most classical such descriptor is the element radius. Representative radii values are derived by statistically analyzing the interatomic distances within a selected group of inorganic crystals. Examples are atomic [7,8], ionic [9,10], covalent [11], metallic[12], and van der Waals radii [9,13,14]. These enabled many inorganic chemists to exclude implausible candidates of stable structures where interatomic distances are highly inconsistent with reported element radii. An intensive analysis of local crystal structures has been carried out by Villars et al [15]. They examined the topologies of local coordination polyhedra in ~ 5000 binary compounds and categorized them based on their topologies. They classified 147 classical structure types into 97 coordination types, and expressed their occurrences on a two-dimensional grid map, where the two axes are the constituent elements arranged by the Mendeleev number. However, their analysis of coordination polyhedra was not applicable to ternary compounds, thus they employed a different classification scheme that uses the concentration-dependent sums of the valence-electron numbers $\overline{VE}$, differences in pseudopotential radii $\overline{R}$, and the differences in electronegativity $\overline{X}$.

In this study, we employed Delaunay polyhedra as basic structural units for structural classification of multinary compounds. The Delaunay tetrahedralization technique is a three-dimensional version of Delaunay triangulation [16]. In general, Delaunay tetrahedra can be uniquely determined such that all four vertices of a tetrahedron lie on a sphere where no other vertices are inside. It divides the space into a set of tetrahedra from a given set of vertices such that the resultant set of tetrahedra are most isotropic. One counterexample is a simple cubic structure, where Delaunay polyhedra are cubes and their eight vertices lie on a sphere without any vertices inside. While Delaunay tetrahedralization is a basic technique in three-dimensional computer graphics, a few studies have applied this on inorganic crystal structures. The conjugate technique, Voronoi tessellation, is more commonly used for the analysis of crystal structures, although the unit polygons are quite complex. [17–20] In particular, the area of a Voronoi polyhedron centered at an atom is reported to be proportional to the strength of an atomic interaction of a given type [21,22].

Visualization of Delaunay tetrahedra can be useful in a crystal structure building tool because the tetrahedron is a primitive unit structure that resembles the real process of crystal formation: an atom wanders on the surface of a cluster of atoms and settles in a stable position. However, chemical bonding could cause preference of some forms of Delaunay polyhedra compared to others in a similar manner to preference of some interatomic distances and bond angles based on the chemical bonding. For instance, sp^3 bonding in the diamond structure results in bond angles of 109°, while the bond angles between nearest neighbors are 90° and 180° in the rocksalt structure. Delaunay polyhedra cannot be reasonable building units for crystal structure building, based on constituent elements and their ratios, without knowing the bonding-type dependences of the chemistry.

We attempted a preliminary analysis of this 'Delaunay chemistry' in this study. We carried out Delaunay tetrahedralization on the 100 most common structure prototypes in the Inorganic Crystal Structure Database (ICSD)[23,24]. We then hierarchically classified these prototypes based on the shapes of the constituent Delaunay polyhedra, to investigate if the basic concept of inorganic chemistry appears in the local geometry of the Delaunay polyhedra. The shape of each Delaunay polyhedron was described by the set of bond angles and solid angles at the vertices.

2. Methodology

2.1. Selection of the common crystal structure prototypes

Crystal structure files were downloaded from the ICSD as 242,848 CIF files in June 2021. The entries with the same 'StructureFormula', FormulaUnitsPerCell" and 'HMS' (Hermann-Mauguin symbol of the space group) were removed as duplicates. Note that this procedure does not remove the duplicates with identical space group number but different HMSs, such as *Pnma*

and *Pbnm* in the space group number 62. The HMS and space group number information in the ICSD database is not curated well and contains many mistakes such as non-consistent HMS and space group number relations, hence we refrained from stripping too much information. From the 179,602 non-duplicate structures, we removed structures with partial occupancies or any errors during the conversion into VASP POSCAR format. Out of the remaining 85,403 structure files, we selected the compounds composed of two to four elements and without noble gases (He-Rn) and actinides (Ac-Lr). The resultant 61,256 structure files were classified by the names of the structure types as recorded in the ICSD. The 100 most frequent structure types were selected for analysis.

The occurrence ratio of H, B, C, N, O, F, and transition metals (TM, defined as Sc-Cu, Y-Ag, La, Ce, and Hf-Au) were identified for each structure type. These ratios (0–100%) were used as elements of seven-dimensional feature vectors for hierarchical classification using the Euclidean metric and Ward's method. We focused on occurrence ratios of these elements to detect ionic compounds with popular anions (H, B, C, N, O, F) and to distinguish compounds consisting of main-group elements only. TM metals were included in 63.5% of the 61,256 structure files. O and F were the most and least popular anion with occurrence ratio 38.2% and 5.3%, respectively.

2.2. Bond angles and solid angles of a Delaunay polyhedron

Positions of *all* atoms that lie on a sphere where no other atoms are inside were considered as vertices of a Delaunay polyhedron in this study. Delaunay polyhedra are generally tetrahedra, but may have more than four facets in exceptional cases. We use polyhedra that are not tetrahedra in these degenerate cases; not forcing tetrahedralization avoids making a choice among equally valid options. Bond angles and the solid angle of Delaunay polyhedra of an arbitrary atom A were considered for use as descriptors of a crystal. The solid angle is important because this value reflects the directional relations between four atoms (three bonds) rather than three for a bond angle (two bonds). For example, the solid angle is small when three bonds point close to each other, thus atoms with many bonds tend to have small solid angles. Atoms with distorted coordination environments can have breadth in the distribution of solid angles compared with atoms at highly symmetric sites. Delaunay polyhedron embraces a void, so the size of a face of the Delaunay polyhedron could be considered as a measure of the channel between this void and a neighboring void. The solid angle distribution therefore acts as a fingerprint of a crystal incorporating geometrical meaning.

Angular information describes the crystal from a viewpoint completely different from bond lengths. Angles are, in one sense, normalized by bond lengths. Bond angles of, for example, rocksalt LiF and RbI are identical even though the bond lengths of the latter are almost 2 times that of the former. Normalizing by the shortest bond length may appear as one solution to scale the geometry crystals. However, how to scale is not an easy problem in anisotropic crystals such as hexagonal wurtzite; the c/a ratio varies from 1.60 in ZnO to 1.64 in ZnS. The problem is even more complicated in more complex crystals. This study chose to focuses on angular information as descriptors of a crystal.

The positions of atoms in the crystal other than A are denoted as P_i (=P_1, P_2, ...). The bond angles and solid angles of a Delaunay polyhedron are determined as follows.

- Delaunay polyhedra with vertex A are identified. The qhull code [25] was used in this study. The vertices of a Delaunay polyhedron, other than A, are labeled as Q_j (Q_1, Q_2 ... Q_J) where $\{Q_j\} \subset \{P_i\}$.
- Facets of the Delaunay polyhedron are labeled as F_k (F_1, F_2 ... F_K).
- Vertices of facet F_k are denoted as R_{km} (R_{k1}, R_{k2} ... R_{kM}) where $\{R_{km}\} \subset \{A,Q_k\}$. Facets where A is not included as a vertex are not investigated hereon.
- All angles $\angle R_{km}AR_{kn}$ of facet F_k are enumerated. The largest angle is the bond angle of O for facet F_k.
- The set of vertices of Q_j that share a ridge with O are denoted as S_n (S_1, S_2 ... S_N) where $\{S_n\} \subset \{Q_j\}$, and the projection of S_n on a unit sphere centered at A is labeled as S'_n.
- The convex hull of points A and S_n is obtained. The facets of the convex hull must be triangulated; any choice can be used if there are multiple choices of triangulation.
- The sum of solid angles of apex A, corresponding to triangles where A is not a vertex, is calculated. For a triangle with vertices S_1, S_2, and S_3, the solid angle is [26,27, 28]

$$2\arccos\frac{1+\cos(\angle S_1AS_2)+\cos(\angle S_2AS_3)+\cos(\angle S_3AS_1)}{\sqrt{2(1+\cos(\angle S_1AS_2))(1+\cos(\angle S_2AS_3))(1+\cos(\angle S_3AS_1))}}.$$

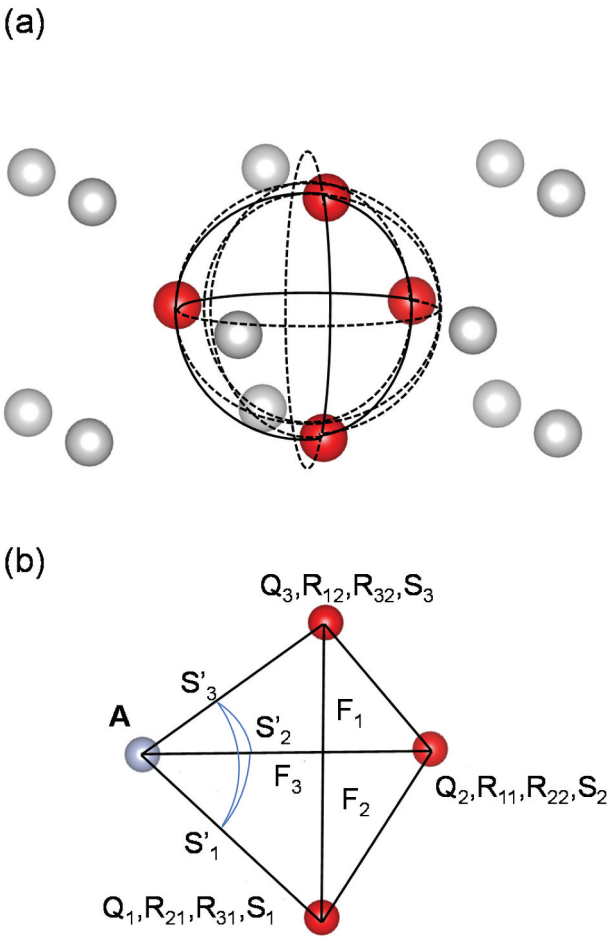


Figure 1. Determination of the bond angle and solid angle around atom O (gray circle). The Delaunay polyhedron is a tetrahedron in this example. Red circles indicate atoms at the remaining vertices of the Delaunay polyhedron (tetrahedron). The unit sphere centered at O is shown with blue curves.

The total of such solid angles is the solid angle of this Delaunay polyhedron for apex A.

Figure 1 shows an example of solid angle determination in a bcc crystal. The tetrahedron with atoms shown in red is discussed here. The above algorithm is almost trivial in tetrahedra; S_1, S_2, and S_3 are simply three atoms other than the apex, and cosines can easily be obtained using the Cartesian coordinates of the atoms.

Things become much more complicated with the Delaunay polyhedron is *not* a tetrahedron. There are multiple, equally reasonable ways to split such a Delaunay non-tetrahedron into tetrahedra, and we want to avoid dealing with this degeneracy issue. The aforementioned algorithm is demonstrated when the Delaunay polyhedron is a cube (Figure 2), which appears in the simple cubic and related structures such as rocksalt, but is valid for all possible polyhedron shapes. The same atom (vertex) is denoted with various symbols depending on the role of the atom in the analysis. There are eight vertices, A and Q_1 to Q_7, and six facets, F_1 to F_6, in Figure 2(a). Facet F_1 is shown in Figure 2(b).

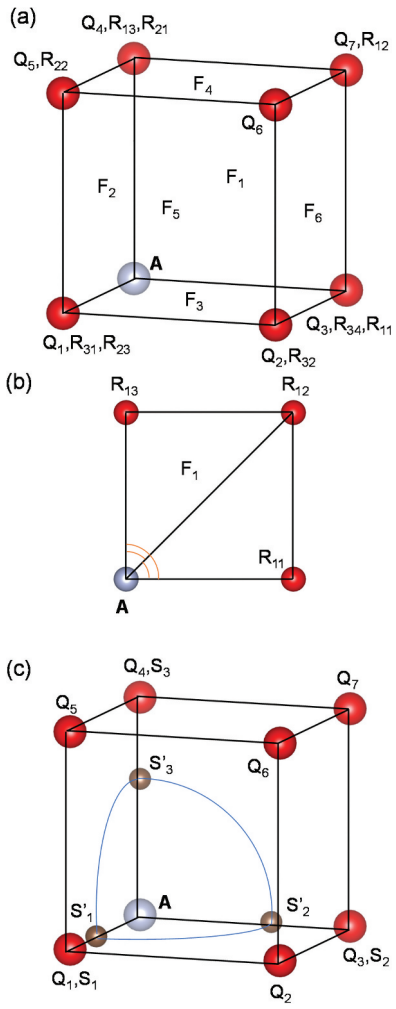


Figure 2. Determination of the bond angle and solid angle around atom O (gray circle). The Delaunay polyhedron is a cube in this example. Red circles indicate atoms at the remaining vertices of the Delaunay polyhedron (cube). Brown balls in (c) represent positions of select atoms (S_1, S_2, and S_3) on a unit sphere centered at O (blue curves).

There are three angles of the form $\angle R_{1m}AR_{1n}$ (orange arcs), and the largest angle, $\angle R_{11}AR_{13} = 90°$, is the bond angle of A for this facet. There are three vertices, S_1 to S_3, that are connected by a ridge of the Delaunay polyhedron to A (Figure 2(c)). These vertices are projected to the unit sphere (projected vertices are S'_1 to S'_3) centered at A, and the area of the spherical triangle with vertices S'_1 to S'_3 (spherical excess) is $\pi/2$. Therefore, the solid angle of A is $\pi/2$ sr, or 1/8 of the solid angle of a sphere (4π sr).

3. Results and discussion

3.1. Hierarchical classification of crystal structure types

Table S1 shows the 100 most frequent structure types together with occurrence ratios of H, B, C, N, O, F, and TMs. The most frequent structure type was Perovskite#CaTiO3 with 747 entries, while the least

Table 1. Summary and description of groups identified from hierarchical classification of structure types. The examples shown are the two most frequently occurring structure type for each subgroup (three in B3 because of a tie in the number of crystals and one in C6 because there is only one structure type).

Group	Count	Description	Subgroup	Count	Description	Examples
A	18	TM oxides	A1	10	Very high O (76–100%) and high to very high TM (64–100%) occurrence	Perovskite#GdFeO3 Pyrochlore#Ca2Nb2O7
			A2	8	Moderate O (41–69%) and moderate to very high TM (46–87%) occurrence	Double_Perovskite#Elpasolite#K2NaAlF6#Ba2CaWO6 Spinel#MgAl2O4
B	18	Low TM	B3	5	Low to moderate O (17–53%) and low TM (18–41%) occurrence	Delafossite#NaFeO2 La2O3#CaAl2Si2 CaV2O4#CaFe2O4
			B4	13	Very low O (0–15%) and very low to moderate TM (10–63%) occurrence	NaCl PbClF
C	64	High TM	C5	2	Heavily borides (B occurrence 69% and 91%)	FeSmB4#CrYB4 Carbide(tau)#Cr23C6
			C6	1	Fluorides (F occurrence 100%)	LiSbF6
			C7	30	Extremely high TM (90–100%) and extremely low H, B, C, N, O, F occurrence (e.g. N ≤ 2%, O ≤ 4%)	ThCr2Si2#CeAl2Ga2#BaAl4 Li2AgSb
			C8	2	Accommodates all of C, N, O (at least 22% each), high TM occurrence (66% and 100%)	Perovskite#CaTiO3 Carbide(eta)#W3Fe3C
			C9	29	Moderate to very high TM (57–90%) and low H, B, C, N, O, F occurrence (e.g. N ≤ 16%, O ≤ 19%)	TiNiSi#MgSrSi Heusler#AlCu2Mn

frequent was MnP#FeAs and PbCl2 with 60 entries each. Figs. S1-S2 show the hierarchical classification result. Structure types were classified into three broad (A-C) groups and nine refined (A1-C9) subgroups, and data in Table S1 is sorted by group. Table 1 shows a summary of these groups (the group of each structure type is also shown in Table S1). Group A, which has a high occurrence ratio of O and TM, branches off first. The remaining structures are then separated into groups B and C, which are distinguished mostly by the TM occurrence ratio. Groups A and B can each be split into two subgroups according to the O occurrence ratio. Boride- and fluoride-favoring cluster types (subgroups C5 and C6) split off first in group C, followed by structures with very high TM and low anion-forming element (H, B, C, N, O, and F) occurrences (subgroup C7). The remaining structure types are separated into two clusters (subgroups C8 and C9) based on the occurrence ratio of C, N, and O. The hierarchical classification of structure types well reflects the nature of bonding, for instance ionic and metallic, as anticipated from the type of constituent elements.

3.2. Relation between structure types and favored bond angles

The relation between structure type and favored bond angle provides a measure of reasonableness on the estimated crystal structure for given constituent elements. For example, one can speculate whether a crystal structure is likely or not likely to be accomplished. The distribution of bond angles can be regarded as a descriptor of the crystal structure. The bond angles in a crystal can be summarized as a histogram with a certain bin size, which is available as a feature vector. The element x_i of the feature vector of the bond angle distribution, $\boldsymbol{x}=(x_0, x_1, x_2, ., x_i, .)$, is the ratio of bond angles between $[ib, (i+1)b)$, where the bin size b was chosen to be 4.5 degrees in this study, and the histogram for a structure type was obtained using all bond angles in all crystals within the structure type. The total number of bond angle records for an entire structure type varied from 9552 in the CdI_2-type structure (hP3) (61 crystals) to 1315200 in the Th_6Mn_{23}-type structure (167 crystals). The sum of all elements of the feature vector $\boldsymbol{x}$ was normalized to unity.

Figures S3-S4 show the hierarchical classification of bond angles, and the feature vector $\boldsymbol{x}$ for each structure type is shown in Figs. S5-S9. The 100 structure types can be categorized into three groups, I to III, and class II is further refined into two subgroups, II2 and II3. Group I (subgroup I1) is the first branch and contains 32 structure types. Subgroups II2 and II3 have 31 and 32 structure types, respectively. The average in each subgroup of the standard deviation, over all bins, of feature vector elements (fraction of bonds in each bin) for constituent structure types is .086, .056, and .048 in subgroups I1, II2, and II3, respectively. The standard deviation is zero in the hypothetical case where the fraction of bond angles in all bins are equal while the standard deviation is maximized when all bond angles are entirely in one bin. This means that the bond angle distribution is relatively diverse in subgroup II3. Group III (subgroup III4) has six structure types, which all have no degrees of freedom in internal coordinates due to symmetry. Possible bond angles there are 45°, 70.5°, and 90°.

The matrix of number of structure types shown with structure type groups A to C and bond angle subgroups I1 to III4 is given in Table 2. The ratio of structure type groups in each bond angle subgroup are visualized in Figure 3. We find a striking difference

Table 2. Matrix of number of structure types based on the structure type and bond angle subgroup.

	Bond angle				
Structure type	I1	II2	II3	III4	Sum
A		1	17		18
B	5	6	5	2	18
C	26	24	10	4	64
Sum	31	31	32	6	100

between subgroup II3 and the others: 17 out of 18 group A structure types are found in subgroup II3 containing 31 out of 100 structure types. This means that structure types favored by TM oxides can accommodate a wider range of bond angles compared to other chemistries.

3.3. Relation between structure types and favored Delaunay solid angles

The bond angle reflects the relation between the atom of interest O and two nearby atoms. In contrast, the solid angle of the Delaunay polyhedron around the atom O is a measure of the spatial distribution of atoms. As with bond angles, the solid angle distributions were obtained and histograms were derived by structure type. The element y_i of the feature vector of the bond angle distribution, $\boldsymbol{y}=(y_0, y_1, y_2, .., y_i, ..)$, is the ratio of bond angles between $[is, (i+1)s)$, where a bin size s of .03π sr was used in this study.

Figures S10-S11 show the hierarchical classifications for solid angles, and the feature vector $\boldsymbol{y}$ for each structure type is shown in Figs. S12-S16. The 100 structure types are categorized into three groups, i to iii, and group iii is further refined into five subgroups, iii3 to iii7. Group i (subgroup i1) is the first branch and contains six structure types with no degrees of freedom in internal coordinates. These are characterized by a large fraction of π/6 sr solid angles. Group ii (ii2) is the second branch with 30 structures. Group iii splits into subgroups iii3 to iii7. The average solid angle among the 100 structure types is the lowest in the four structure types in subgroup iii3. Subgroup iii7 has no degrees of freedom in internal coordinates, and the variety of solid angles are also limited; their fractions of the π/6 sr solid angle is lower than in group i.

The average in each subgroup of the standard deviation, over all solid angle bins, of feature vector elements (fraction of bonds in each of the considered 37 bins) of constituent structure types in subgroups ii2, iii3, iii4, iii5, and iii6 is 0.073, 0.063, 0.065, 0.060, and 0.066, respectively. Therefore, subgroup ii2 (group ii) has a relatively concentrated distribution of solid angles. The distribution of the difference of solid angle fractions, [0.09π, 0.12π) - [0.12π, 0.15π) sr, among subgroups iii4, iii5, iii6 can be expressed using (mean, standard deviation) = (−0.22, 0.13), (−0.03, 0.10), and (0.14, 0.11), respectively. In other words, the fraction of solid angles in the range [0.09π, 0.12π) is smaller, almost the same, and larger than [0.12π, 0.15π) sr in subgroups iii4, iii5, iii6, respectively. Group ii can also be distinguished from subgroups iii4, iii5, and iii6 by comparing the fraction of bonds in the range [0.06π, 0.12π) and [0.15π, 0.21π) sr. The latter is larger than the former by 0.12 or more in all structure types in group ii. On the other hand, this relation is satisfied by only three of the 57 structure types in subgroups iii4, iii5, and iii6 (the maximum among the three is 0.15).

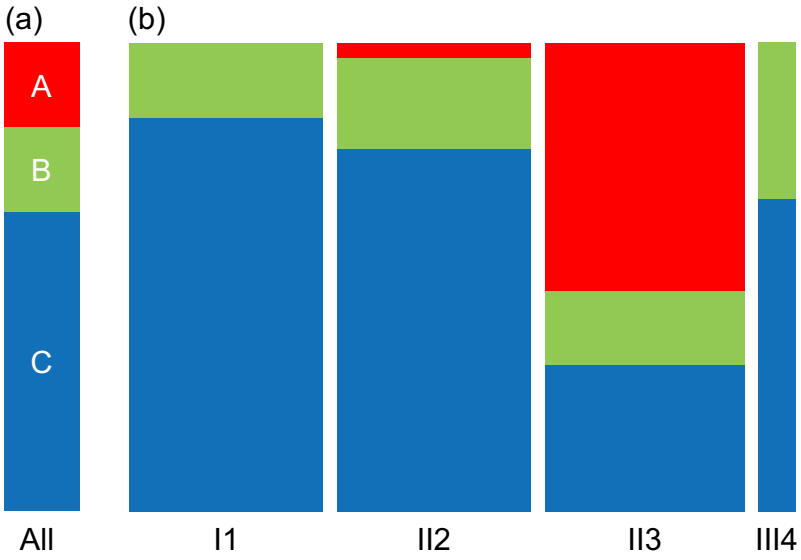


Figure 3. Ratio of structure type groups (A-C) (a) for all structure types and (b) by bond angle subgroups (I1-III4). The bar widths in (b) are proportional to the number of structure types in each bond angle subgroup, and each structure type has the same area in (b).

Table 3. Matrix of number of structure types based on the structure type and solid angle subgroup.

	Solid angle							
Structure type	i1	ii2	iii3	iii4	iii5	iii6	iii7	Sum
A				2	15	1		18
B	1	4	1	3	6	1	2	18
C	5	26	3	10	9	10	1	64
Sum	6	30	4	15	30	12	3	100

A matrix of number of structure types, shown with structure type groups A to C and solid angle subgroups i1 to iii7 as the axes, is given in Table 3, and the ratio of structure type classes in each solid angle subgroup is visualized in Figure 4. Again, group A structure types are concentrated in a solid angle subgroup; 15 out of 18 group A structure types are found in subgroup iii5 with 30 out of 100 structure types. There is a large group, ii2, with 30 structure types that do not contain any group A structure type. The relatively concentrated solid angle distributions as in subgroup ii2 is also not favored in group A TM oxides. These results suggest that, as in bond angles, TM oxides strongly prefer some subgroups of solid angle distributions over others. Relatively concentrated solid angle distributions as in subgroups i1, ii2, and iii7 are not favored in group A TM oxides. The average solid angle in group A structure types is 0.18π sr (~4π/22), which is much larger than what is allowed in subgroup iii3 (average 0.15π sr).

3.4. Comparison between bond angle and solid angle classes

Table 4 shows the matrix of number of structure types shown with bond angle subgroups I1 to III4 and solid angle subgroups i1 to iii7 as axes. The matrix is far from sparse; when subgroups with less than 10 structure types are removed (III4, i1, iii3, and iii7), 10 out of 12 matrix elements are occupied. This means that the bond angle and solid angle subgroups are quite independent. Sizeable matrix elements in Table 4 are I1-ii2 and II3-iii5 with 23 structure types each. The former contains two group B and 21 group C structure types, indicating a high concentration of TM-rich non-oxides. On the other hand, 15 out of 18 groups A structure types are concentrated in the latter 23 out of 100 structure types.

Table 4. Matrix of number of structure types based on the bond angle and solid angle subgroup. Structure type group A is mostly found in subgroups II3 and iii5.

	Solid angle							
Bond angle	i1	ii2	iii3	iii4	iii5	iii6	iii7	Sum
I1	1	23	3	2			2	31
II2		6		8	7	10		31
II3		1	1	5	23	2		32
III4	5						1	6
Sum	6	30	4	15	30	12	3	100

3.5. Additional remarks

One may consider performing crystal structure prediction solely based on bond angles and solid angles and finding new crystals. Unfortunately, this is a very difficult task. Construction of relative atom positions from angular information only is very subtle compared to the inverse problem of obtaining angular information from atomic positions. In one example, angular information does not change when a simple cubic structure is gradually elongated in the <100> direction. Information on the scale (lattice parameter size) is lacking because of normalization. The labels of atoms are removed, thus how to arrange elements at atom positions, even if identified, is not prescribed. Despite these drawbacks, angular information provides an additional viewpoint to

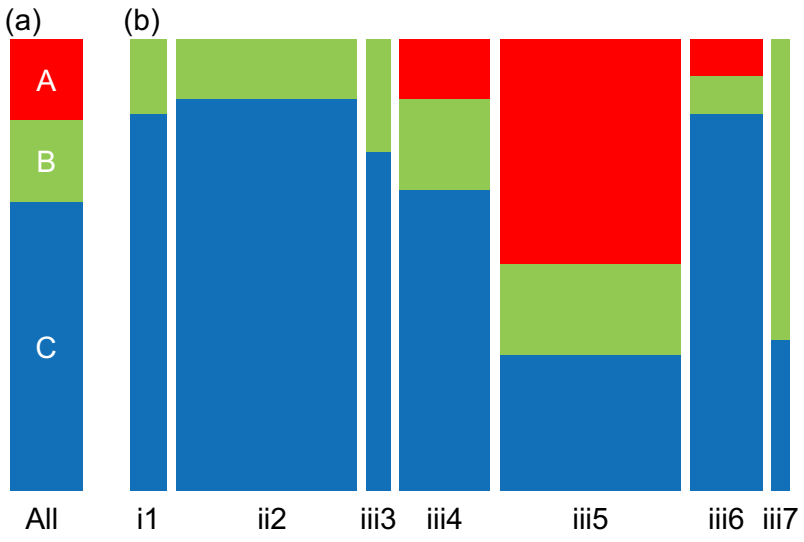


Figure 4. Ratio of structure type groups (A-C) (a) for all structure types and (b) by solid angle subgroup (i1-iii7). The bar widths in (b) are proportional to the number of structure types in each solid angle subgroup, and each structure type has the same area in (b).

discuss crystal structures that complement existing studies. One application is checking whether a predicted crystal structure is reasonable given its stoichiometry.

4. Summary

This study categorized occurrence ratios of H, B, C, N, O, F, and TM in frequently appearing structure types in binary, ternary, and quaternary compounds in the ICSD database. The structure types can be broadly categorized into those favoring TM oxides, low TM occurrence, and high TM occurrence. The bond angle and solid angle distributions of Delaunay polyhedra with atoms as vertices were calculated as descriptors of the local environment for each structure type, and hierarchical categorization was conducted to categorize the distributions. TM oxides preferred a limited number of bond angle and solid angle subgroups. Use of bond angle and solid angle (sub)groups together found a particular sizeable group favored by TM oxide structure types and another by high TM occurrence structure types. Our results indicate that the anion, or lack of, plays a key role in determining the crystal structure, and structures preferred by TM oxides can accommodate a broad range of bond angles and solid angles.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was funded by grants (No. JPMJCR17J3 and JPMJCR19J1) from CREST of the Japan Science and Technology Agency (JST). The VESTA code [28] was used to draw Figs. 1 and 2.

ORCID

Yoyo Hinuma http://orcid.org/0000-0003-2272-1178
Yukari Katsura http://orcid.org/0000-0002-8905-2995

References

[1] Oganov AR, Glass CW. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. J Chem Phys. 2006;124(24):244704.
[2] Oganov AR, Lyakhov AO, Valle M. How evolutionary crystal structure prediction works - and why. Acc Chem Res. 2011;44(3):227–237.
[3] Lyakhov AO, Oganov AR, Stokes HT, et al. New developments in evolutionary structure prediction algorithm USPEX. Comput Phys Commun. 2013;184(4):1172–1182. DOI:10.1016/j.cpc.2012.12.009
[4] Wang Y, Lv J, Zhu L, et al. Crystal structure prediction via particle-swarm optimization. Phys Rev B. 2010;82(9):094116. DOI:10.1103/PhysRevB.82.094116
[5] Zhang Y, Wang H, Wang Y, et al. Computer-assisted inverse design of inorganic electrides. Phys Rev X. 2017;7(1):011017. DOI:10.1103/PhysRevX.7.011017
[6] Yamashita T, Kanehira S, Sato N, et al. CrySPY: a crystal structure prediction tool accelerated by machine learning. Science and Technology of Advanced Materials: Methods. 2021;1(1):87–97. DOI:10.1080/27660400.2021.1943171
[7] Slater JC. Atomic Radii in Crystals. J Chem Phys. 1964;41(10):3199–3204.
[8] Clementi E, Raimondi DL, Reinhardt WP. Atomic screening constants from SCF functions. II. Atoms with 37 to 86 Electrons. J Chem Phys. 1967;47(4):1300–1307.
[9] Pauling L. The nature of the chemical bond. 3rd ed. Ithaca (NY): Cornell University Press. 1960.
[10] Shannon R. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Acta Crystallographica Section A. 1976;32(5):751–767.
[11] Cordero B, Gómez V, Platero-Prats AE, et al. Covalent radii revisited. Dalton Trans. 2008;21(21):2832–2838. DOI:10.1039/b801115j
[12] Greenwood NN, Earnshaw A. Chemistry of the elements. Burlington (MA): Elsevier Butterworth-Heinemann; 1997.
[13] Bondi A. van der Waals Volumes and Radii. J Phys Chem. 1964;68(3):441–451.
[14] Alvarez S. A cartography of the van der Waals territories. Dalton Trans. 2013;42(24):8617–8636.
[15] Hafner J, Hulliger F, Jensen WB, et al The structures of binary compounds (cohesion and structure). Amsterdam: North-Holland. 1990.
[16] Delaunay B. Sur la sphe·re vide. A la me·moire de Georges Voronoi. Bulletin de l'Acade·mie des Sciences de l'URSS. 1934;6:793–800.
[17] Yang L, Dacek S, Ceder G. Proposed definition of crystal substructure and substructural similarity. Phys Rev B. 2014;90(5):054102.
[18] Isayev O, Oses C, Toher C, et al. Universal fragment descriptors for predicting properties of inorganic crystals. Nat Commun. 2017;8(1):15679. DOI:10.1038/ncomms15679
[19] Waroquiers D, Gonze X, Rignanese G-M, et al. Statistical analysis of coordination environments in oxides. Chem Mater. 2017;29(19):8346–8360. DOI:10.1021/acs.chemmater.7b02766
[20] Jalem R, Nakayama M, Noda Y, et al. A general representation scheme for crystalline solids based on Voronoi-tessellation real feature values and atomic property data. Sci Technol Adv Mater. 2018;19(1):231–242. DOI:10.1080/14686996.2018.1439253
[21] Blatov VA. Voronoi–Dirichlet polyhedra in crystal chemistry: theory and applications. Crystallography Reviews. 2004;10(4):249–318.
[22] Blatov VA. A method for hierarchical comparative analysis of crystal structures. Acta Crystallographica Section A. 2006;62(5):356–364.

[23] Zagorac D, Müller H, Ruehl S, et al. Recent developments in the inorganic crystal structure database: theoretical crystal structure data and related features. J Appl Crystallogr. 2019;52(5):918–925. DOI:10.1107/S160057671900997X

[24] Belsky A, Hellenbrandt M, Karen VL, et al. New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design. Acta Cryst B. 2002;58 (3):364–369. DOI:10.1107/S0108768102006948

[25] Barber CB, Dobkin DP, Huhdanpaa H. The quickhull algorithm for convex hulls. ACM Trans Math Softw. 1996;22(4):469–483.

[26] Fillmore DW, Fillmore JP. Whitney Forms for Spherical Triangles I: The Euler, Cagnoli, and Tuynman Area Formulas, Barycentric Coordinates, and Construction with the Exterior Calculus. arXiv. 2014;1404.6592.

[27] Hinuma Y, Kamachi T, Hamamoto N. Algorithm for automatic detection of surface atoms. Transactions of the Materials Research Society of Japan. 2020;45 (4):115–120.

[28] Momma K, Izumi F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. J Appl Crystallogr. 2011;44(6):1272–1276.