CAMBRIDGE
UNIVERSITY PRESS

Review

Structural and chemical complexity of minerals: an update

Sergey V. Krivovichev[1,2]* , Vladimir G. Krivovichev[2], Robert M. Hazen[3] , Sergey M. Aksenov[1] ,
Margarita S. Avdontceva[2], Alexander M. Banaru[1,4], Liudmila A. Gorelova[2], Rezeda M. Ismagilova[2], Ilya V. Kornyakov[2],
Ivan V. Kuporev[2], Shaunna M. Morrison[3], Taras L. Panikorovskii[1,2] and Galina L. Starova[2]

[1]Kola Science Centre, Russian Academy of Sciences, 14 Fersman Street, Apatity 184200, Russia; [2]Department of Crystallography, St. Petersburg State University, University Emb. 7/9, St. Petersburg, 199034, Russia; [3]Earth and Planets Laboratory, Carnegie Institution for Science, 5251 Broad Branch Road NW, Washington DC 20015, USA.; and [4]Faculty of Chemistry, Moscow State University, Vorobievy Gory, Moscow 119991, Russia

Abstract

The complexities of chemical composition and crystal structure are fundamental characteristics of minerals that have high relevance to the understanding of their stability, occurrence and evolution. This review summarises recent developments in the field of mineral complexity and outlines possible directions for its future elaboration. The database of structural and chemical complexity parameters of minerals is updated by H-correction of structures with unknown H positions and the inclusion of new data. The revised average complexity values (arithmetic means) for all minerals are 3.54(2) bits/atom and 345(10) bits/cell (based upon 4443 structure reports). The distributions of atomic information amounts, $^{chem}I_G$ and $^{str}I_G$, versus the number of mineral species fit the normal modes, whereas the distributions of total complexities, $^{chem}I_{G,total}$ and $^{str}I_{G,total}$, along with numbers of atoms per formula and per unit cell are log normal. The three most complex mineral species known today are ewingite, morrisonite and ilmajokite, all either discovered or structurally characterised within the last five years. The most important complexity-generating mechanisms in minerals are: (1) the presence of isolated large clusters; (2) the presence of large clusters linked together to form three-dimensional frameworks; (3) formation of complex three-dimensional modular frameworks; (4) formation of complex modular layers; (5) high hydration state in salts with complex heteropolyhedral units; and (6) formation of ordered superstructures of relatively simple structure types. The relations between symmetry and complexity are considered. The analysis of temporal dynamics of mineralogical discoveries since 1875 with the step of 25 years show the increasing chemical and structural complexities of human knowledge of the mineral kingdom in the history of mineralogy. In the Earth's history, both diversity and complexity of minerals experience dramatic increases associated with the formation of Earth's continental crust, initiation of plate tectonics and the Great Oxidation event.

Introduction

In modern mineralogy, crystal structure and end-member chemical formula are considered as the most important intrinsic properties of a mineral species (Nickel, 1995; Nickel and Grice, 1998; Hawthorne *et al.*, 2021). The complexity of chemical composition and crystal structure are therefore fundamental characteristics of minerals that have high relevance to the understanding of their stability, occurrence and evolution. The universal quantitative measures of structural and chemical complexity were proposed in Krivovichev (2012, 2013, 2014; Krivovichev *et al.*, 2018) on the basis of Shannon information theory and have been recognised as useful parameters for the analysis of various mineralogical and crystal chemical phenomena, from crystal-structure classifications to metastable crystallisation and mineral evolution

(Tambornino *et al.*, 2015; Christy *et al.*, 2016; Varn and Crutchfield, 2016; Plášil *et al.*, 2017; Plášil, 2018a, 2018b, 2018c; Ovchinnikov *et al.*, 2018; Majzlan *et al.*, 2018; Colmenero *et al.*, 2019; Gurzhiy and Plášil, 2019; Gurzhiy *et al.*, 2019, 2021; Hornfeck, 2020; Bindi *et al.*, 2020; Majzlan, 2020; Mugnaioli *et al.*, 2020; Sabirov and Shepelevich, 2021; Kaußler and Kieslich, 2021, etc.). The aim of this review is to summarise the recent findings in the field of structural and chemical complexity of minerals and to outline possible directions for its further development. This work can be considered as a sequel for the previous review by Krivovichev (2013) that was devoted to the modern introduction of information-based complexity measures into mineralogy (from a historical point of view, the pioneering contributions were made by Petrov (1970), Bulkin (1972a, 1972b) and Yushkin (1977)). Krivovichev (2013) pointed out the importance of quantitative expression of structural and chemical complexity of minerals for mineralogy and the novel insights that may be deduced from complexity considerations. In this review, our analysis will be based upon a comprehensively updated and corrected database of information-based complexity measures of minerals (see Supplementary Information), some of which have

been incorporated into the Global Earth Mineral Inventory (GEMI) database (Prabhu *et al.*, 2021).

The review is organised as follows. First, we discuss basic methodological principles of complexity analysis of minerals and provide an overview of theoretical developments, including extension of complexity measures, quantification and visualisation of different factors that contribute to the overall structural complexity of a mineral, and theoretical relations between complexity and configurational entropy. The second and third parts are based upon the analysis of an updated database of complexity parameters, including average complexities for different sets of minerals, and a review of the most complex minerals based upon the recent mineralogical discoveries. The relations between symmetry and complexity are discussed in the fourth part of the review, along with some interesting (though not completely unexpected) insights into the history of explorative mineralogical research. The applications of the complexity considerations in mineral evolution studies are described in the fifth section, whereas the sixth contains brief discussions of directions for further investigation.

Theory and methodology

Basic principles

In accord with the previous theoretical considerations (Krivovichev, 2012, 2013), the structural complexity is estimated as the amounts of structural Shannon information per atom ($^{str}I_G$) and per unit cell ($^{str}I_{G,total}$) calculated according to the following equations:

$$^{str}I_G = - \sum_{i=1}^{k} p_i \log_2 p_i \text{ (bits/atom)} \quad (1)$$

$$^{str}I_{G,total} = v \times {}^{str}I_G = -v \times \sum_{i=1}^{k} p_i \log_2 p_i \text{ (bits/cell)} \quad (2)$$

where k is the number of different crystallographic sites (crystallographic orbits) in the structure and p_i is the random choice probability for an atom from the ith crystallographic orbit, that is:

$$p_i = m_i/v \quad (3)$$

where m_i is a multiplicity of a crystallographic orbit (i.e. the number of atoms of a specific Wyckoff site in the reduced unit cell), and v is the total number of atoms in the reduced unit cell.

By analogy, the chemical complexity of a mineral can be estimated from its end-member formula extracted from the International Mineralogical Association List of Minerals (Pasero, 2021) and, in some cases, corrected by Krivovichev (2021b). It should be emphasised that, within the context of this study, it is implicitly understood that each mineral has a unique end-member formula (Hawthorne *et al.*, 2021; Hawthorne, 2021). Thus, for the mineral with the end-member formula $E^{(1)}_{c1}E^{(2)}_{c2}...E^{(k)}_{ck}$, where $E^{(i)}$ is an ith chemical element in the formula and c_i is its integer coefficient, the chemical information per atom ($^{chem}I_G$) and per formula unit, f.u. ($^{chem}I_{G,total}$) can be calculated as follows (Siidra *et al.*, 2014; Krivovichev *et al.*, 2018a, 2020a):

$$^{chem}I_G = - \sum_{i=1}^{k} p_i \log_2 p_i \text{ (bits/atom)} \quad (4)$$

$$^{chem}I_{G, total} = e \times {}^{chem}I_G = -e \times \sum_{i=1}^{k} p_i \log_2 p_i \text{ (bits/f.u.)} \quad (5)$$

where k is the number of different elements in the formula and p_i is the random-choice probability for an atom of the ith element, that is:

$$p_i = c_i/e \quad (6)$$

where c_i is the coefficient of the ith element in the formula and e is the total number of atoms in the chemical formula:

$$e = \sum_{i=1}^{k} c_i \quad (7)$$

Krivovichev *et al.* (2018a) demonstrated that, on average, structural and chemical complexities correlate with each other, which agrees well with the intuitive feeling that chemically complex species generally possess complex crystal structures. Though there are exceptions of this empirical rule (Krivovichev, 2013), it is a statistically valid observation.

New developments: Hornfeck extension

In some cases, the information-based measures proposed by Krivovichev (2012, 2013) do not possess enough discriminative power to distinguish between the crystal structures that are thought to be different in terms of their symmetry or thermodynamic stability. For instance, α- and β-quartz have the same structural complexity parameters, 0.918 bits/atom and 8.265 bits/cell, which is counter-intuitive, as the α → β transition is associated with the increase of symmetry from trigonal to hexagonal. In order to overcome these difficulties, Hornfeck (2020) proposed to extend the information-based complexity measures by introducing chemical, combinatorial and coordinational parameters. In the framework of the Hornfeck's approach, the complexity parameters $^{str}I_G$ and $^{str}I_{G,total}$ defined by equations (1) and (2), respectively, are classified as *combinatorial*. The *coordinational* measures are defined as follows (in order to be consistent with the definitions given above, the Hornfeck notations of complexity measures are modified).

Each crystallographic (Wyckoff) site in a crystal structure has a finite number of degrees of freedom depending upon its site symmetry. For instance, the site located on the inversion centre has zero degrees of freedom, whereas one located on the mirror plane has two degrees of freedom. The number of degrees is called by Hornfeck (2020) an arity, by analogy with mathematics, where arity means the number of arguments taken by a function. The arity may be equal to 0, 1, 2, and 3. Let the arity of the ith crystallographic site be denoted as a_i. Then

$$A = \sum_{i=1}^{k} a_i \quad (8)$$

is the sum of arities for a crystal structure consisting of k crystallographic orbits (k different Wyckoff sites). The information amount of the distribution of arities in the crystal structure is defined as (d.f. = degree of freedom)

$$I_A = - \sum_{i=1}^{k} p_i^a \log_2 p_i^a \text{ (bits/d.f)} \quad (9)$$

$$I_{A,total} = A \times I_A = -A \times \sum_{i=1}^{k} p_i^a \log_2 p_i^a \quad \text{(bits/cell)} \quad (10)$$

where

$$p_i^a = a_i/A \quad (11)$$

For instance, for α-quartz (space group $P3_221$), $a_1(\text{Si}) = 1$ (the Si atom is located on the twofold axis) and $a_2(\text{O}) = 3$ (general site). Therefore, $A = 4$, $I_A = 0.811$ bits/d.f., $I_{A,total} = 3.245$ bits/cell. In contrast, for β-quartz (space group $P6_222$), $a_1(\text{Si}) = 0$, $a_2(\text{O}) = 1$ (general site), $A = 1$, $I_A = 0.000$ bits/d.f., $I_{A,total} = 0.000$ bits/cell (Banaru *et al.*, 2021).

The I_A and $I_{A,total}$ parameters are identified by Hornfeck (2020) as coordinational complexity. The *configurational* complexity I_{GA} can be calculated as a combination of combinatorial and coordinational complexities via the formulas:

$$I_{GA} = I(v, A) + (I_{G,total} + I_{A,total})/(v + A) \quad (12)$$

where

$$I(v, A) = \frac{v}{v + A} \log_2 \frac{v + A}{v} + \frac{A}{v + A} \log_2 \frac{v + A}{A} \quad (13)$$

The dimension of the I_{GA} parameters is in bits per degree of freedom per atom (bits/d.f.; each atom is considered as an additional degree of freedom). The total configurational information amount per reduced unit cell is calculated as

$$I_{GA,total} = (v + A) \times I_{GA} \quad (14)$$

The application of the equations (9 to 14) to α- and β-quartz results in different configurational complexities, 1.775 and 1.296 bits/d.f., respectively, and 15.978 and 12.955 bits/cell, respectively. The relative configurational simplicity of β-quartz corresponds to its thermodynamic nature of a high-temperature polymorph (see below).

The analogous calculations for sillimanite and andalusite that, otherwise, have the same structural complexity (Krivovichev, 2013) provide the values of 3.564 bits/d.f. and 160.366 bits/cell, and 3.638 bits/d.f. and 167.350 bits/cell, respectively. The lower complexity of sillimanite agrees well with its nature as a high-temperature Al_2SiO_5 polymorph.

Hornfeck (2020) also provided measures for the chemical complexities of crystalline compounds based upon the atomic numbers of their constituent elements. Combined with combinatorial and coordinational complexities, the Hornfeck chemical complexity parameters provide a measure for a *crystallographic* complexity. The latter may be considered as numerical characteristics of crystalline compounds, and therefore is of use for digital crystal-structure classification purposes. In this review, however, we use the structural and chemical complexity parameters as defined by Krivovichev (2013) and Krivovichev *et al.* (2018a), respectively.

Complexity, configurational entropy, and entropy of mixing

The configurational entropy of crystalline solids has received considerable attention recently due to the development of high-entropy materials (see: Miracle and Senkov (2017), McCormack

and Navrotsky (2021), Dippo and Vecchio (2021), and references therein). In all cases, the authors refer to configurational entropy as to the entropy of mixing, in conformity with geochemical and mineralogical textbooks (Putnis, 1992; Ottonello, 2000). However, it is evident that the configurational entropy as an entropy of an atomic configuration is not restricted to mixing entropy only, as a complex crystal structure with many crystallographic orbits (= many different Wyckoff sites) should obviously have lower configuration entropies compared to the isochemical structure with few crystallographic orbits. Thus, it makes sense to separate configurational entropy (S_{cfg}) into configurational entropy *sensu stricto* ($S_{cfg,ss}$) and configurational entropy of mixing ($S_{cfg,mix}$):

$$S_{cfg} = S_{cfg,ss} + S_{cfg,mix} \quad (15)$$

In the case of ideal mixing and a 'single-lattice model' (i.e. when the solid consists of one lattice or one crystallographic orbit), the value of $S_{cfg,mix}$ is calculated as follows:

$$S_{cfg,mix} = -R \sum_{j=1}^{N} f_j \ln f_j \quad (16)$$

where R is a gas constant ($\sim$8.3145 J·K^{-1}·mol^{-1}), N is the number of different chemical elements occupying the 'lattice' under consideration, and f_j is the concentration of the jth element (note that vacancy can also be considered as an entropic species) so that

$$\sum_{j=1}^{N} f_j = 1 \quad (17)$$

For the 'sublattice-model' (i.e. when the crystal structure consists of several 'sublattices' (= several crystallographic orbits), the formula (16) transforms into (Miracle and Senkov, 2017; Dippo and Vecchio, 2021):

$$S_{cfg,mix} = -R \left(\sum_{x=1}^{x} a^x \sum_{j=1}^{N} f_j \ln f_j \right) / \sum_{x=1}^{x} a^x \quad (18)$$

where x is the number of 'sublattices' (= the number k of crystallographic orbits in our notations above), and a^x is the number of sites in the xth 'sublattice' (= multiplicity m_i of an ith crystallographic orbit relative to reduced unit cell).

Taking into account that $a^x = m_i$ and $x = k$ and applying the equation (3), the formula (18) transforms into:

$$S_{cfg,mix} = -R \left(\sum_{i=1}^{k} p_i \sum_{j=1}^{N} f_j \ln f_j \right) \quad (19)$$

According to the value of $S_{cfg,mix}$, materials with structural disorder are classified into low-entropy ($S_{cfg,mix} < 1R$), medium-entropy ($1R < S_{cfg,mix} < 1.5R$), and high-entropy ($S_{cfg,mix} > 1.5R$) materials (Dippo and Vecchio, 2021).

In order to derive the expression for the estimates of configurational entropy *sensu stricto* (i.e. the entropy associated with structural complexity, assuming that more complex structures have lower entropies), Krivovichev (2016) suggested the following reasoning. The entropy is defined as

$$S = k_B \ln W \quad (20)$$

where k_B is the Boltzmann constant ($\sim$1.38064·10^{23} J·K^{-1}) and W is the number of microstates that realise the given macrostate. For the crystal structure consisting of one crystallographic orbit, the

$S_{\text{cfg,ss}}$ is maximal with atomic entropy equal to

$$S_{\text{cfg,ss}}^{max}/N = k_B\ln(N-1) \qquad (21)$$

where N is the number of atoms that may exchange their positions locally at a given temperature (Krivovichev (2016) assumed that this value is specific for a particular compound; the exchange of positions implies realisation of different microstates and the difference of atoms of the same element and the same crystallographic orbit). For the molar entropy, the equation (21) transforms into

$$S_{\text{cfg,ss}}^{\max} = R\ln(N-1) \qquad (22)$$

Let us now suppose that, due to an unspecified reason, the symmetry of the initial crystal structure is reduced, and the single crystallographic orbit splits into k crystallographic orbits. Obviously, the crystal structure becomes more complex and its configurational entropy *sensu stricto* should decrease. Krivovichev (2016) demonstrated that the amount by which the maximal value $S_{\text{cfg,ss}}^{\max}$ is decreased, $\Delta S_{\text{cfg,ss}}$, is expressed as the following:

$$\Delta S_{\text{cfg,ss}}/N = k_B^{str}I_G\ln2 \qquad (23)$$

Then the configuration entropy *sensu stricto* of the resulting low-symmetry structure, $S_{\text{cfg,ss}}$, can be calculated as

$$S_{\text{cfg,ss}} = S_{\text{cfg,ss}}^{\max} - \Delta S_{\text{cfg,ss}} = R[\ln(N-1) - ^{str}I_G\ln2] \qquad (24)$$

Through the combination of equations (1), (15), (19) and (24), one deduces that

$$S_{\text{cfg,ss}} = R[\ln(N-1) - \ln2\sum_{i=1}^{k}p_i(\log_2p_i - \sum_{j=1}^{N}f_j\log_2f_j)] \qquad (25)$$

or

$$S_{\text{cfg,ss}} = R[ln(N-1) - \ln2^{\text{str}}I_{G,mix}] \qquad (26)$$

where

$$^{str}I_{G,mix} = -\sum_{i=1}^{k}p_i\log_2p_i - \sum_{j=1}^{N}f_j\log_2f_j) \qquad (27)$$

is a new measure of structural complexity that takes into account contributions from complexity *sensu stricto* and structural disorder that by definition decreases structural complexity. Equation (27) provides a procedure to account for the influence of chemical substitutions and vacancies upon the overall structural complexity, and the parameter $^{str}I_{G,mix}$ can therefore be considered as a measure of crystal chemical complexity (or the degree of atomic order) that takes into account not only structural architecture, but also its chemical nature. As many important rock-forming minerals are complex solid solutions and can be considered as low-, medium- or high-entropy materials (McCormack and Navrotsky, 2021), this measure might be useful for mineralogical and geochemical implications.

Equation (27) can be re-written as

$$^{str}I_{G,\text{mix}} = -\sum_{i=1}^{k}p_i\log_2p_i + \sum_{i=1}^{k}p_i\sum_{j=1}^{N}f_j\log_2f_j$$
$$= ^{str}I_G - I_{\text{mix}} \qquad (28)$$

where

$$I_{\text{mix}} = -\sum_{i=1}^{k}p_i\sum_{j=1}^{N}f_i\log_2f_j = S_{\text{cfg,mix}}/(R\ln2) \qquad (29)$$

is a contribution of mixing to the overall crystal-chemical complexity. It is obvious that, as $^{str}I_G \geq 0$ and $I_{\text{mix}} \geq 0$, the case is possible when $^{str}I_G < I_{\text{mix}}$ and $^{str}I_{G,\text{mix}} < 0$. This means that, in contrast to the measure of structural complexity $^{str}I_G$, the parameter of crystal chemical complexity may be negative and provide a positive contribution to the total configurational entropy of a crystalline solid.

Kaußler and Kieslich (2021) recently suggested the following method to calculate the atomic structural complexity parameter (denoted here as $^{KK}I_G$) in order to account for the presence of chemical substitutions (or partially occupied sites) in a crystal structure:

$$^{KK}I_G = -\sum_{j=1}^{N}\sum_{i=1}^{k}f_jp_i\log_2(f_jp_i) \quad \text{(bits/site)} \qquad (30)$$

where f_j is the occupancy of the ith site by jth element.

Taking into account that, for a given site,

$$\sum_{j=1}^{N}f_j = 1 \qquad (31)$$

one may re-group the equation (30) to get

$$^{KK}I_G = -\sum_{i=1}^{k}p_i\log_2p_i$$
$$-\sum_{i=1}^{k}p_i\sum_{j=1}^{N}p_if_j\log_2f_j \text{ (bits/site)} \qquad (32)$$

or

$$^{KK}I_G = ^{str}I_G + I_{\text{mix}} \qquad (33)$$

Through the comparison of equation (33) with equation (28), it is clear that the application of the formula (30) to solid solutions leads to the results exactly opposite to those provided by the formula (27). According to eq. (30), the presence of chemical substitutions (including substitutions by vacancies) results in an *increase* of the overall structural complexity, i.e. the *increase* in the atomic order, which is counter-intuitive, as disorder *decreases* the degree of order (= structural complexity). In order to show this, let us consider the forsterite–fayalite solid solution, Mg_2SiO_4–Fe_2SiO_4. The crystal structure of forsterite–fayalite (space group *Pbnm*) contains two M sites (4a and 4c), one Si site (4c) and three O sites (4c, 4c and 8d). For the sake of convenience, we suppose that, in the Mg_2SiO_4–Fe_2SiO_4 solid solution, Fe distributes equally between two M sites. The application of the formulae (1) and (2) provides $^{str}I_G = 2.522$. The application of the formulae (29) and (28) provides for the composition

$Mg_{2(1-x)}Fe_{2x}SiO_4$:

$$I_{\mathrm{mix}} = -(2/7)[x\log_2 x + (1-x)\log_2(1-x)] \quad (34)$$

$$^{\mathrm{str}}I_{G,\mathrm{mix}} = 2.522 + (2/7)[x\log_2 x + (1-x)\log_2(1-x)] \quad (35)$$

In contrast, equation (30) provides:

$$^{\mathrm{KK}}I_G = 2.522 - (2/7)[x\log_2 x + (1-x)\log_2(1-x)] \quad (36)$$

Graphs corresponding to the behaviour of structural complexity indices versus the content x (the average Fe content in the M sites) are shown in Fig. 1. The value $x = 0.5$ corresponds to the state of maximal disorder (minimal degree of order) and has the minimal value of $^{\mathrm{str}}I_{G,\mathrm{mix}}$, which agrees with general understanding of relations between order and complexity, whereas the $^{\mathrm{KK}}I_G$ value is maximal for $x = 0.5$, which contradicts those relations.

The value of $^{\mathrm{str}}I_{G,\mathrm{mix}}$ allows one to distinguish mineral species that have the same space group, similar unit-cell parameters, and differ in the degree of disorder only. Sanidine (Sa), orthoclase (Or) and microcline are the polymorphs of $KAlSi_3O_8$ with a different degree of Al–Si order at the tetrahedral sites. According to the current nomenclature (see, e.g. Krivovichev, 2020a), in K feldspars, Al preferentially occurs at the T_1 site. If $x_{Al(T1)}$ is the content of Al at the T_1 site, then $x_{Al(T1)} = 0.5$ for orthoclase and 0.25 for sanidine (in microcline, all Al is at the $T_1 0$ site with $x = 1$). The behaviour of $^{\mathrm{str}}I_{G,\mathrm{mix}}$ and entropy of mixing versus $x_{Al(T1)}$ is shown in Fig. 2. The crystal chemical complexity (= the degree of order) decreases with the increasing entropy of mixing.

Ladder diagrams

The complexity of a crystal structure is a sum of contributions of different factors. In order to estimate quantitatively the amount of information generated by a particular structural element, Krivovichev (2018) suggested the use of ladder diagrams for visualising different information amounts arising from specific levels of hierarchical organisation of a crystal structure. For oxysalt minerals in particular, the following levels can be recognised: (1) topological information (TI) reflecting the amount of information with respect to the topological structure of basic structural units (clusters, chains, ribbons, layers and frameworks); (2) structural information (SI) arising from the symmetry reduction of structural units due to their distortions induced by their interactions with their local environments (in some cases, structural information equals topological information); (3) packing information (PI) due to the appearance of several identical structural units in the reduced unit cell (for example, the cell may contain two identical layers that may or may not be symmetrically related); (4) information arising from the interstitial structure (IS) except for H atoms; and (5) information due to the presence of H atoms responsible for the formation of a hydrogen bonding (HB) system. The calculation of different contributions proceeds via sequential calculations starting from the determination of the highest symmetry of structural units present in the structure. This first step is probably the most challenging and corresponds to finding the highest symmetry embedding of the chemical bonding net in 3D Euclidean space. In principle, it can be generated by the *Systre* software program (Delgado-Friedrichs et al.,
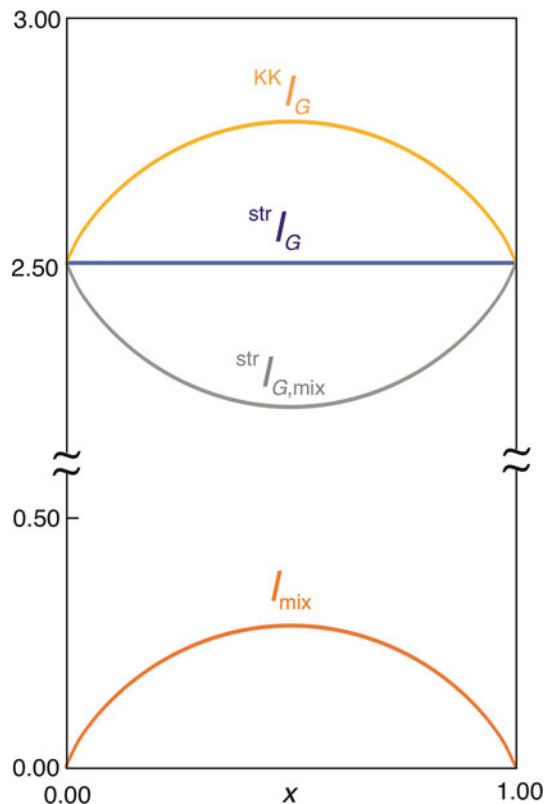

Fig. 1. Atomic structural complexities calculated using different approaches versus x value for the fayalite–forsterite $Mg_{2(1-x)}Fe_{2x}SiO_4$ solid solution.

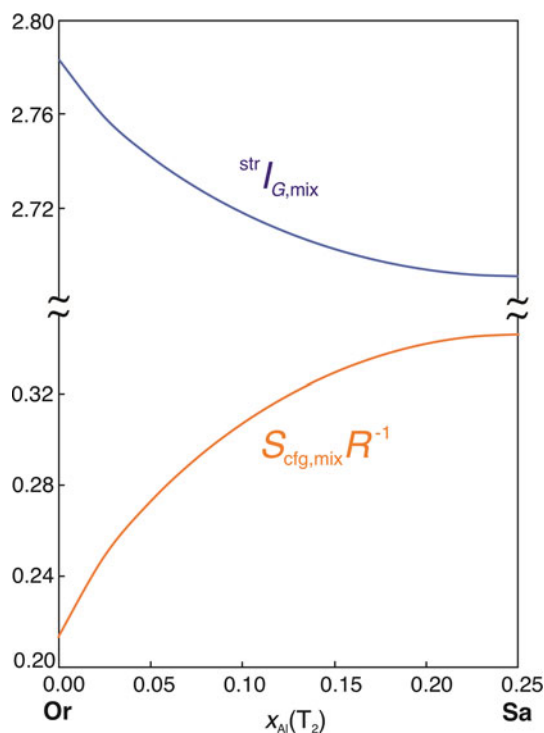**Fig. 2.** The behaviour of $^{\mathrm{str}}I_{G,\mathrm{mix}}$ and entropy of mixing versus the $x_{Al(T1)}$ content of Al at the T_1 site for the orthoclase (Or) – sanidine (Sa) solid solution.

2017). To find the most symmetric embedding of a net, the algorithm of *Systre* software uses barycentric placement in which each vertex of the net is placed into the centre of gravity of its adjacent
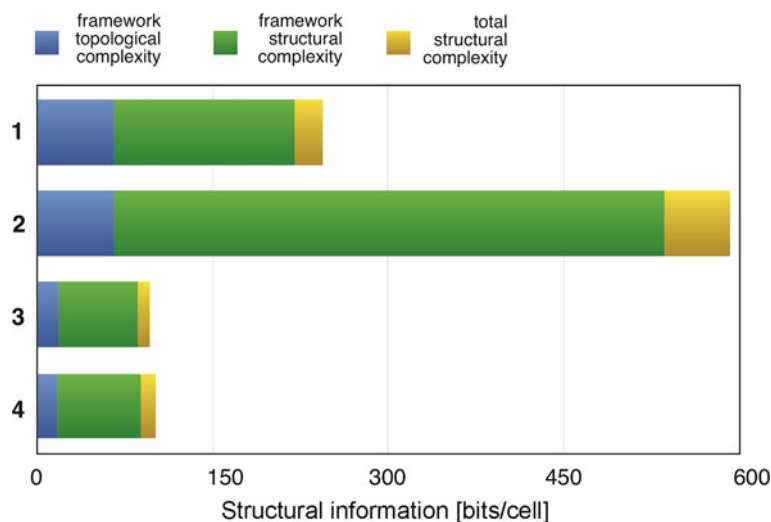
Fig. 3. The ladder diagrams for the Ca[Al$_2$Si$_2$O$_8$] polymorphs with tetrahedron topologies: 1 – anorthite $I\bar{1}$; 2 – anorthite $P\bar{1}$; 3 – dmisteinbergite; 4 – svyatoslavite. After Krivovichev (2020a).

vertices (all vertices have the same weight). The resulting net is called stable if no two vertices collide. Stable nets have the maximal achievable crystallographic symmetry (Delgado-Friedrichs and O'Keeffe, 2003). In many cases, the situation does not need application of *Systre* software, as the highest symmetry embedding may be determined from analysis of the topological structure of structural units by visual inspection. Figure 3 shows ladder diagrams for the four CaAl$_2$Si$_2$O$_8$ polymorphs: anorthite (space group $P\bar{1}$), high-pressure or high-temperature anorthite (space group $I\bar{1}$), dmisteinbergite and svyatoslavite (Krivovichev, 2020a). Both room- and high-temperature anorthites are based upon the feldspar-type tetrahedral framework, which is rather low in complexity (blue colour). The major contribution to the structural complexity is due to the distortion of the tetrahedral framework induced by its interaction to the interstitial Ca^{2+} cations (green colour). The specific contribution of Ca^{2+} cations is quite small (yellow colour). The tetrahedral networks in svyatoslavite and dmisteinbergite are topologically simpler than the feldspar framework, which is typical for metastable polymorphs (see below).

For some interesting applications of ladder diagrams to the analysis of uranyl minerals and inorganic compounds see Gurzhiy and Plášil (2019) and Gurzhiy *et al.* (2019 and 2021).

Database on structural and chemical complexity of minerals

The present review is based upon the updated database of structural and chemical complexity parameters of minerals calculated by means of the equations (1–7). For the crystal structures of hydrated minerals, where the positions of H atoms could not be determined, the H-correction procedure was used as first described by Pankova *et al.* (2018a) through introduction of surrogate H sites with corresponding multiplicities. No positional disorder was taken into account. For the calculation of chemical complexities, the end-member mineral chemical formulae were used as approved by the International Mineralogical Association (IMA; Pasero, 2021). The structural complexity parameters for mineral species have recently been partially incorporated into the Global Earth Mineral Inventory (GEMI) online database (Prabhu *et al.*, 2021; see below).

Complexity statistics

Average complexity for all minerals

Krivovichev (2013) provided the following average structural complexity values (arithmetic means) for the mineral kingdom: 3.23(2) bits per atom and 228(6) bits per reduced unit cell, which corresponds to the average 70.6 atoms per cell (based upon 3949 structure reports). For October 2021, the revised values (arithmetic means) are 3.54(2) bits/atom and 345(10) bits/cell, respectively, corresponding to the average of 97.5 atoms per cell (based upon 4443 structure reports). The difference between the 2013 and 2021 values is due to the H-correction of the 2013 database and a high number of very complex minerals that have been discovered since 2013 (in particular, polyoxometalates; see below). The average chemical complexities (arithmetic means) are 1.63(1) bits per atom and 63(1) bits per formula, which account for the average number of 38.7 atoms per formula (based upon 5455 chemical formulae). The latter value is rather high, keeping in mind that most rock-forming minerals are rather simple in their chemical compositions.

Complexity distributions

The distributions of atomic information amounts, $^{chem}I_G$ and $^{str}I_G$, versus the number of mineral species (N) are shown in Fig. 4a and b, respectively. Both distributions fit the normal modes, with the goodness-of-fit R^2 values of 0.970 and 0.974, respectively. In contrast, the distributions of total complexities, $^{chem}I_{G,total}$ and $^{str}I_{G,total}$, are heavily right-skewed (Fig. 5a and b, respectively) and correspond to log-normal distributions.

The $N-$ $^{chem}I_{G,total}$ and $N-$ $^{str}I_{G,total}$ can be described by the following exponential functions:

$$N = 5285.19 \times \exp(-^{chem}I_{G,total}/52.28)(R^2 = 0.998) \quad (37)$$

$$N = 1988.16 \times exp(-^{str}I_{G,total}/218.62)(R^2 = 0.998) \quad (38)$$

The distributions $N- \log(^{chem}I_{G,total})$ and $N - \log(^{str}I_{G,total})$ for the representative sampling of all minerals are symmetrical and fit
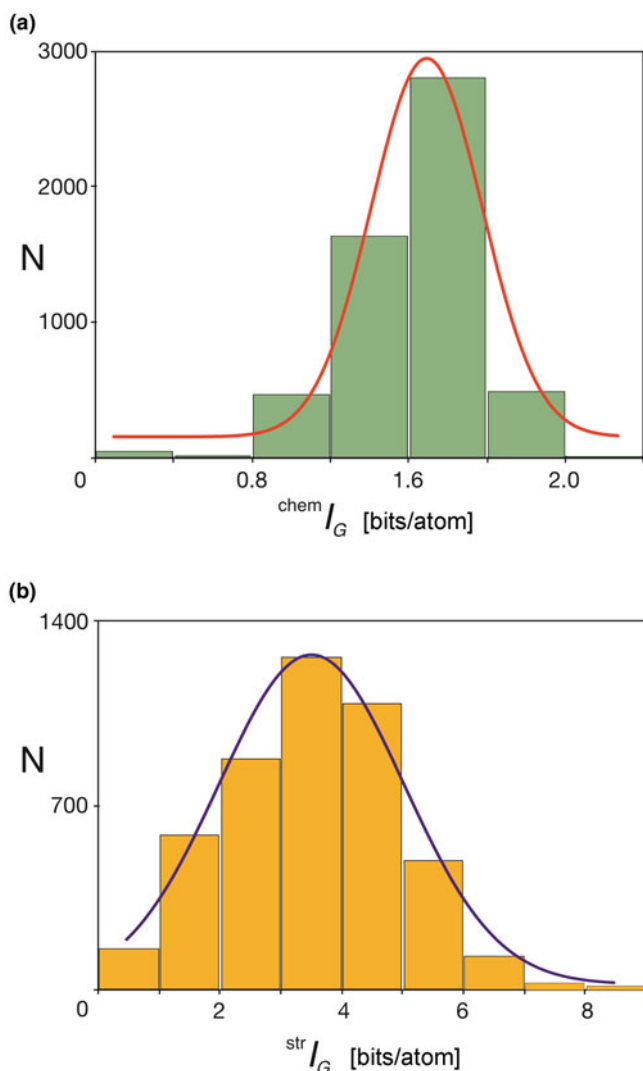
Fig. 4. Distributions of the atomic (a) chemical and (b) structural complexities versus the number of minerals (N).

The mean ($\bar{X}_{\log x_i}$) and empirical standard deviation ($\sigma_{\log x_i}$) of the logarithms of the data and their estimators ($\bar{X}^*$ and s^*) for all minerals are given in Table 1. Note that the geometric means are substantially lower than the arithmetic means given above.

The log-normal character of the total structural and chemical complexity distributions is the result of the log-normal distributions of the numbers of atoms per reduced unit cell and per formula unit (Table 1).

Krivovichev (2012, 2013) mentioned that the atomic structural complexity ($^{str}I_G$) is symmetry-sensitive, whereas the number of atoms per cell (v) is a size-sensitive parameter, while $^{str}I_{G,total}$ combines both symmetry and size as two sides of complexity. In some sense, the $^{str}I_G$ and v quantities can be considered as intensive and extensive parameters, respectively. The normal distribution of $^{str}I_G$ seems to be quite natural and reflects the influence of many mutually unrelated parameters upon the symmetry of crystals. The nature of the log-normal distribution of the v value deserves special attention.

Log-normal distributions are of high importance in many areas of science (Limpert et al., 2001) and, in particular, are used to describe species abundance distributions (SADs) in ecology. According to this model, there are few species with high and low abundances and many with intermediate values. It is, however, unclear how the ecological models can be used to understand the behaviour of mineral systems, where mineral species do not compete for food, territory and other resources. A possible explanation for the log-normal distribution of the number of atoms per formula or per cell is the balance between the need to accommodate different elements in the same cell (most minerals contain from four to five different chemical elements (Krivovichev and Charykova, 2013; Krivovichev et al., 2018b) and the tendency of the crystal structures to be as simple as possible.

Average complexity for mineral classes

Arithmetic means for chemical and structural complexities of different mineral classes are provided in Table 2. It can be seen that the O-free (anoxic) mineral species are both chemically and structurally simpler than the mineral species containing oxygen as a mineral-forming component. Among O-bearing species, oxides are the simplest (taking into account the large number of binary oxides with relative chemical simplicity), whereas sulfates and borates are the most complex. The latter observation is expected, as the majority of sulfates and borates crystallise at low temperatures and are very frequently hydrated. Next in complexity are silicates, arsenates and phosphates that have different modes of origin, including both low- and high-temperature environments.

The higher complexity of O-bearing minerals agrees well with the fact that a large number of them have a secondary origin and form at the expense of primary O-free minerals (e.g. during oxidation of sulfide mineral deposits; Krivovichev et al., 2019b, 2020b). The difference in complexity between anoxic and O-bearing minerals also plays a crucial role in the complexity changes during mineral evolution, i.e. with the formation of a large suite of complex oxysalt minerals as a result of the Great Oxidation Event (Hazen et al., 2008; see below).

Most complex minerals: an update

Twenty most complex minerals: the 2021 list

Krivovichev (2013) provided the following list of twenty most structurally complex minerals known at the time ($I_{G,total}$ in bits

the log-normal distribution with the goodness-of-fit R^2 values of 0.954 and 0.974, respectively.

For a log-normal distribution, the most precise method for estimating the parameters μ^* and σ^* of the statistical population relies on the log transformation. The mean and empirical standard deviation of the logarithms of the data are calculated and then back-transformed (Sachs 1997; Limpert et al., 2001) as follows:

$$\bar{X}^* = \text{antilg}(\bar{X}_{\log x_i}) = \text{antilg}\left(\frac{1}{n}\sum_{i=1}^{n}\log(x_i)\right) = (\Pi_{i=1}^{n} x_i) \quad (39)$$

$$s^* = \text{antilg}(\sigma_{\log x_i}) = \text{antilg}\sqrt{\left(\frac{1}{n-1}\sum_{i=1}^{n}[\log(x_i) - \log(\bar{X}^*)]^2\right)} \quad (40)$$

where $\bar{X}^*$ is the geometric mean of the data; s^* is back-transformed empirical standard deviation of the logarithms of the data.
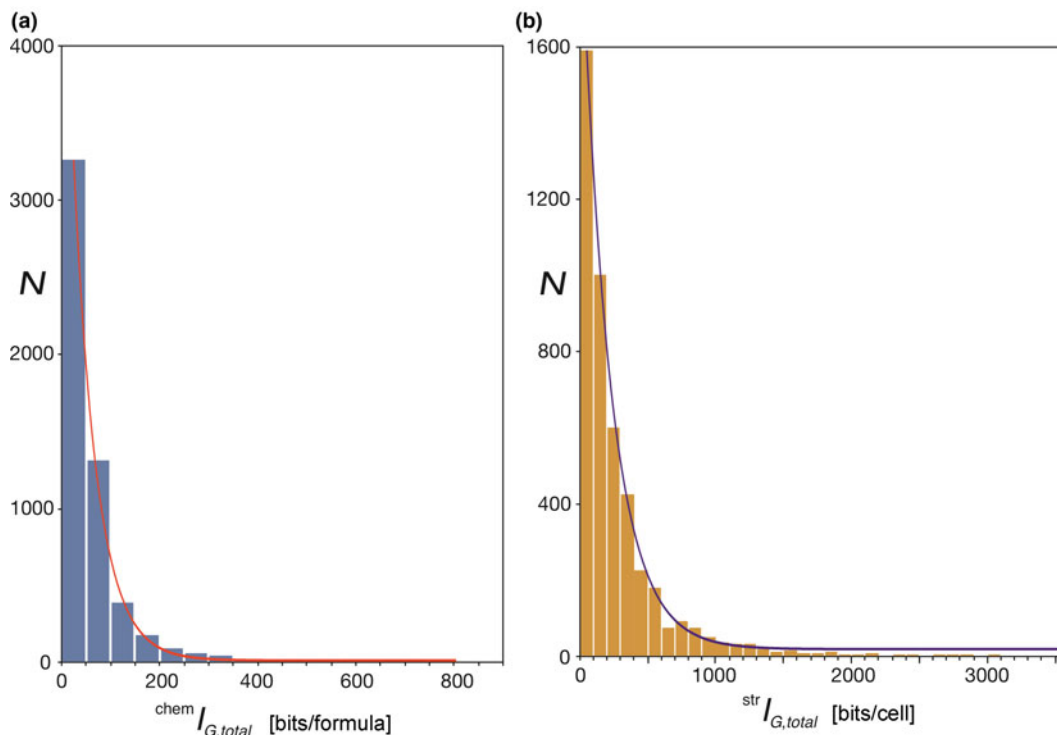
Fig. 5. Distributions of the total (a) chemical and (b) structural complexities versus the number of minerals (N).

per cell): paulingite-(Ca) (6767), fantappièite (5948), sacrofanite (5317), mendeleevite-(Ce) (3399), bouazzerite (3035), megacyclite (2951), vandendriesscheite (2835), giuseppettite (2723), stilpnomelane (2484), stavelotite-(La) (2411), rogermitchellite (2321), parsettensite (2310), apjohnite (2305), antigorite ($m = 17$ polysome) (2250), tounkite (2188), tschörtnerite (2132), farneseite (2094), kircherite (2053), bannisterite (2031) and mutinaite (2025). The updated 2021 list of twenty most structurally complex minerals is given in Table 3 (the difference between the 2013 and 2021 parameters for particular minerals is due to the fact that the 2013 data did not take into account contributions from H atoms). Only one representative of a given structure type is listed: for example, only one member of the voltaite group is given, whereas other mineral species isotypic to voltaite (Majzlan *et al.*, 2013) are omitted; mendeleevite-(Nd) is omitted, etc. Only eight species from the 2013 list remain in the 2021 list. The main reason is the discoveries of seven very complex minerals since 2013 (ewingite, morrisonite, vanarsite, paddlewheelite, gauthierite, rowleyite and meerschautite), the crystal-structure solution of ilmajokite, and the inclusion of the previously neglected or H-corrected species (voltaite, alfredstelznerite, manitobaite, parsettensite and postite).

Ewingite, $Mg_8Ca_8[(UO_2)_{24}(CO_3)_{30}O_4(OH)_{12}(H_2O)_8](H_2O)_{130}$, is the Earth's most complex mineral reported so far with an estimated complexity of 23478 bits/cell. The mineral was found as golden-yellow crystals on a damp wall of the old Plavno mine of the Jáchymov ore district, western Bohemia, Czech Republic (Olds *et al.*, 2017a), where it crystallised from low-temperature uranyl-bearing aqueous solutions. Its crystal structure contains a 54-nuclear (24U + 30C) uranyl carbonate cluster $[(UO_2)_{24}(CO_3)_{30}O_4(OH)_{12}(H_2O)_8]^{32-}$ shown in Fig. 6a. Its skeletal representation (Fig. 6b), where each node corresponds either to U (yellow) or to C (grey) centre emphasises the presence of four U_3 triangles (with U–U distances shorter than 4 Å) that correspond to the trimers of three $((UO_2)O_5)$ pentagonal bipyramids sharing the same equatorial O atom. Two other building units are the $[(UO_2)(CO_3)_3]$ and $[(UO_2)(CO_3)_2(H_2O)_2]$ hexagonal bipyramids. The visual complexity of the cluster architecture can further be reduced by leaving only U atoms and the addition of the U–U links corresponding to the U–U distances in between 4 and 6.2 Å (Fig. 6c). The resulting graph can be considered as consisting of four U_3 trimers (U–U < 4 Å) and six U_4 dihedra of two edge-sharing U_3 triangles (U–U = 4.0–6.2 Å). The

Table 1. Average chemical and structural complexities for minerals.*

Parameter	Units	n	$\bar{X}$	σ	$\sigma_{\bar{X}}$	$\bar{X}_{logx_i}$	σ_{logx_i}	$\bar{X}^*$	s^*
$^{chem}I_G$	bits/atom	5455	1.627	0.339	0.005	–	–	–	–
$^{str}I_G$	bits/atom	4443	3.538	1.391	0.021	–	–	–	–
e	atoms	5455	36.514	49.663	0.676	1.329	0.462	21.33	2.90
v	atoms	4443	70.430	100.50	1.510	1.637	0.447	43.35	2.80
$^{chem}I_{G,total}$	bits/f.u.	5455	63.080	88.092	1.120	1.543	0.504	34.91	3.19
$^{str}I_{G,total}$	bits/cell	4443	320.140	654.040	9.810	2.150	0.617	141.25	4.14

* n – number of minerals taken into account; $\bar{X}$ = arithmetic mean; σ = standard deviation; $\sigma_{\bar{X}}$ = standard error of mean; $\bar{X}_{logx_i}$ = arithmetic mean; σ_{logx_i} = standard deviation; $\sigma_{\bar{X}}$ = standard error of mean; $\bar{X}^* = 10^{\bar{X}_{logx_i}}$; $s^* = 10^{\sigma_{logx_i}}$.

Table 2. Average structural and chemical complexities for different mineral classes.*

Mineral classes	N	$^{chem}I_G$ [bits/atom]		$^{chem}I_{G,total}$ [bits/f.u.]		N	$^{str}I_G$ [bits/atom]		$^{str}I_{G,total}$ [bits/cell]	
		$\bar{X}$	$\sigma_{\bar{X}}$	$\bar{X}$	$\sigma_{\bar{X}}$		$\bar{X}$	$\sigma_{\bar{X}}$	$\bar{X}$	$\sigma_{\bar{X}}$
O-free	1032	1.28	0.01	26.41	1.53	808	2.30	0.05	133.37	10.95
Halides	88	1.30	0.03	14.32	2.75	83	1.95	0.12	62.91	11.57
Sulfides	548	1.48	0.01	39.98	2.63	419	2.88	0.07	197.64	18.98
Selenides	89	1.27	0.04	11.56	1.54	73	1.80	0.10	35.88	5.96
Tellurides	88	1.28	0.04	13.11	1.67	60	2.22	0.16	73.35	14.98
Arsenides	196	1.41	0.03	41.29	5.35	154	2.83	0.13	222.27	35.35
O-based	4434	1.70	0.004	71.75	1.41	3679	3.75	0.02	362.17	11.57
Oxides	275	1.28	0.01	17.87	1.23	252	2.20	0.06	77.64	6.05
Silicates	1556	1.80	0.005	95.01	2.94	1315	3.97	0.03	385.53	17.33
Sulfates	521	1.67	0.01	75.24	3.03	522	4.12	0.05	460.99	25.46
Selenites	38	1.76	0.04	46.01	5.84	32	3.67	0.17	241.16	36.27
Te oxysalts	89	1.62	0.03	45.77	5.36	62	3.22	0.12	191.63	19.83
Arsenates	471	1.73	0.01	73.41	4.98	365	3.93	0.06	377.84	31.43
Phosphates	427	1.72	0.01	64.89	3.22	360	3.93	0.05	370.08	19.14
Carbonates	285	1.74	0.01	61.96	4.18	234	3.63	0.08	343.23	39.65
Borates	144	1.70	0.01	58.97	6.67	133	4.01	0.11	414.99	44.57

* N = number of minerals taken into account; $\bar{X}$ = arithmetic mean; $\sigma_{\bar{X}}$ = standard error of arithmetic mean.

centres of the trimers and the dihedra (denoted by red and blue circles in Fig. 6c) form a tetrahedron and an octahedron, respectively, with six tetrahedral edges in correspondence to six octahedral vertices (Fig. 6d). Such a relation between the tetrahedron and octahedron graphs is known in graph theory as an edge-to-vertex duality.

Morrisonite, $Ca_{11}[As^{3+}V_2^{4+}V_{10}^{5+}As_6^{5+}O_{51}]_2 \cdot 78H_2O$, has 13,558 bits of Shannon information per unit cell and, after ewingite, is the second most complex mineral known so far. The vanarsite-group minerals (including morrisonite itself; Kampf et al., 2016) contain a unique and previously unknown type of the V–As polyoxometalate cluster with the composition $[As^{3+}V_{12}As_6^{5+}O_{51}]$. The cluster consists of twelve (VO_6) octahedra forming a wheel- or corona-shaped unit centred by an As^{3+} cation

in trigonal-pyramidal coordination (due to the stereoactivity of a lone-electron pair) and surrounded by six $(As^{5+}O_4)$ tetrahedra (Fig. 7). The same polyoxometalate cluster occurs in the crystal structure of vanarsite, $NaCa_{12}[As_2^{3+}V_{17}^{5+}V_7^{4+}As_{12}^{5+}O_{102}] \cdot 78H_2O$, but with a different mode of packing of polyoxometalate units. The vanarsite-group minerals form from low-temperature aqueous solutions with V and As derived from the oxidation of primary unoxidised phases.

Zolotarev et al. (2020) reported the crystal structure of ilmajokite, $Na_{11}KBaCe_2[Ti_{12}Si_{37.5}O_{94}(OH)_{31}] \cdot 29H_2O$, the third most complex mineral, which was first reported from a hydrothermal vein in Karnasurt Mountain, Lovozero, Kola Peninsula, Russia (Bussen et al., 1972). The crystal structure is based on a 3D titanosilicate framework consisting of trigonal prismatic titanosilicate

Table 3. Twenty most structurally complex minerals, their space groups and complexity characteristics.

no.	Mineral name	Chemical formula	Space group	$^{chem}I_G$ [bits/atom]	$^{str}I_G$ [bits/atom]	$^{chem}I_{G,total}$ [bits/f.u.]	$^{str}I_{G,total}$ [bits/cell]	Ref.
1	Ewingite	$Mg_4Ca_4(UO_2)_{12}(CO_3)_{15}O_2(OH)_6 \cdot 69H_2O$	$I4_1/acd$	1.576	7.603	512	23,478	1
2	Morrisonite	$Ca_{11}(As^{3+}V_2^{4+}V_{10}^{5+}As_6^{5+}O_{51})_2 \cdot 78H_2O$	$P2_1/c$	1.611	8.622	620	13,588	2
3	Ilmajokite	$Na_{11}KBaCe_2Ti_{12}Si_{37.5}O_{94}(OH)_{31}29H_2O$	$C2/c$	1.843	8.468	567	11,990	3
4	Paulingite-Ca	$Ca_5(Al_{10}Si_{32}O_{84}) \cdot 34H_2O$	$Im\bar{3}m$	1.722	5.830	418	11,591	4
5	Fantappièite	$[Na_{99}Ca_{33}]_{\Sigma132}[Si_{99}Al_{99}O_{396}](SO_4)_{33} \cdot 6H_2O$	$R\bar{3}$	1.932	7.285	1747	6112	5
6	Bouazzerite	$Bi_6Mg_{11}Fe_{14}(AsO_4)_{18}O_{12}(OH)_4 \cdot 86H_2O$	$P2_1/c$	1.648	7.636	658	6063	6
7	Vanarsite	$NaCa_{12}(As_2^{3+}V_{17}^{5+}V_7^{4+}As_{12}^{5+}O_{102}) \cdot 78H_2O$	$P2_1/c$	1.645	7.621	635	5990	2
8	Paddlewheelite	$MgCa_5Cu_2(UO_2)_4(CO_3)_{12}(H_2O)_{33}$	Pc	1.718	8.443	287	5876	7
9	Tschörtnerite	$Ca_4K_3Cu_3(Al_2Si_{12}O_{48})(OH)_8 \cdot 20H_2O$	$Fm\bar{3}m$	2.096	5.081	275	5421	8
10	Vandendriesscheite	$Pb_3(UO_2)_{20}O_{12}(OH)_{22} \cdot 22H_2O$	$Pbca$	1.465	6.539	271	4865	9
11	Mendeleevite-(Ce)	$Cs_6(Ce_{22}Ca_6)_{\Sigma28}(Si_{70}O_{175})(OH)_{14} \cdot 21H_2O$	$Pm\bar{3}$	1.766	5.464	653	4415	10
12	Alfredstelznerite	$Ca_4[B_4O_4(OH)_6]_4 \cdot 19H_2O$	$Pca2_1$	1.549	7.140	218	4027	11
13	Gauthierite	$KPb[(UO_2)_7O_5(OH)_7] \cdot 8H_2O$	$P2_1/c$	1.549	6.943	102	3416	12
14	Rowleyite	$[Na(NH_4)_9Cl_4](V_2^{5+}PO_8)_6 \cdot nH_2O$	$Fd\bar{3}m$	2.108	4.621	251	3382	13
15	Meerschautite	$Ag_{12}Pb_{84}Sb_{90}S_{224}O$	$P2_1$	1.595	7.693	656	3185	14
16	Voltaite	$K_2Fe_5Fe_4Al(SO_4)_{12} \cdot 18H_2O$	$Fd\bar{3}c$	1.750	4.160	221	3178	15
17	Parsettensite	$K(Mn_6Al)_{\Sigma7}(Si_8O_{20})(OH)_8 \cdot 2H_2O$	$C2/m$	1.897	6.903	110	3085	16
18	Manitobaite	$Na_{16}Mn_{25}Al_8(PO_4)_{30}$	Cc	1.706	7.637	340	3039	17
19	Postite	$MgAl_2(OH)_2(V_{10}O_{28}) \cdot 27H_2O$	$Pccn$	1.478	5.993	186	3021	18
20	Rogermitchellite	$Na_6Sr_{12}Ba_2Zr_{13}Si_{39}B_4O_{123}(OH)_6 \cdot 20H_2O$	$P6_3cm$	2.026	5.591	529	3019	19

References: (1) Olds et al. (2017a); (2) Kampf et al. (2016); (3) Zolotarev et al. (2020); (4) Gordon et al. (1966); (5) Cámara et al. (2010); (6) Brugger et al. (2007); (7) Olds et al. (2018); (8) Effenberger et al. (1998); (9) Burns, 1997; (10) Sokolova et al. (2011); (11) Cooper et al. (2010); (12) Olds et al. (2017b); (13) Kampf et al. (2017); (14) Biagioni et al. (2016); (15) Mereiter, 1972; (16) Guggenheim and Eggleton, 1994; (17) Tait et al. (2011); (18) Kampf et al. (2012); and (19) McDonald and Chao (2010).
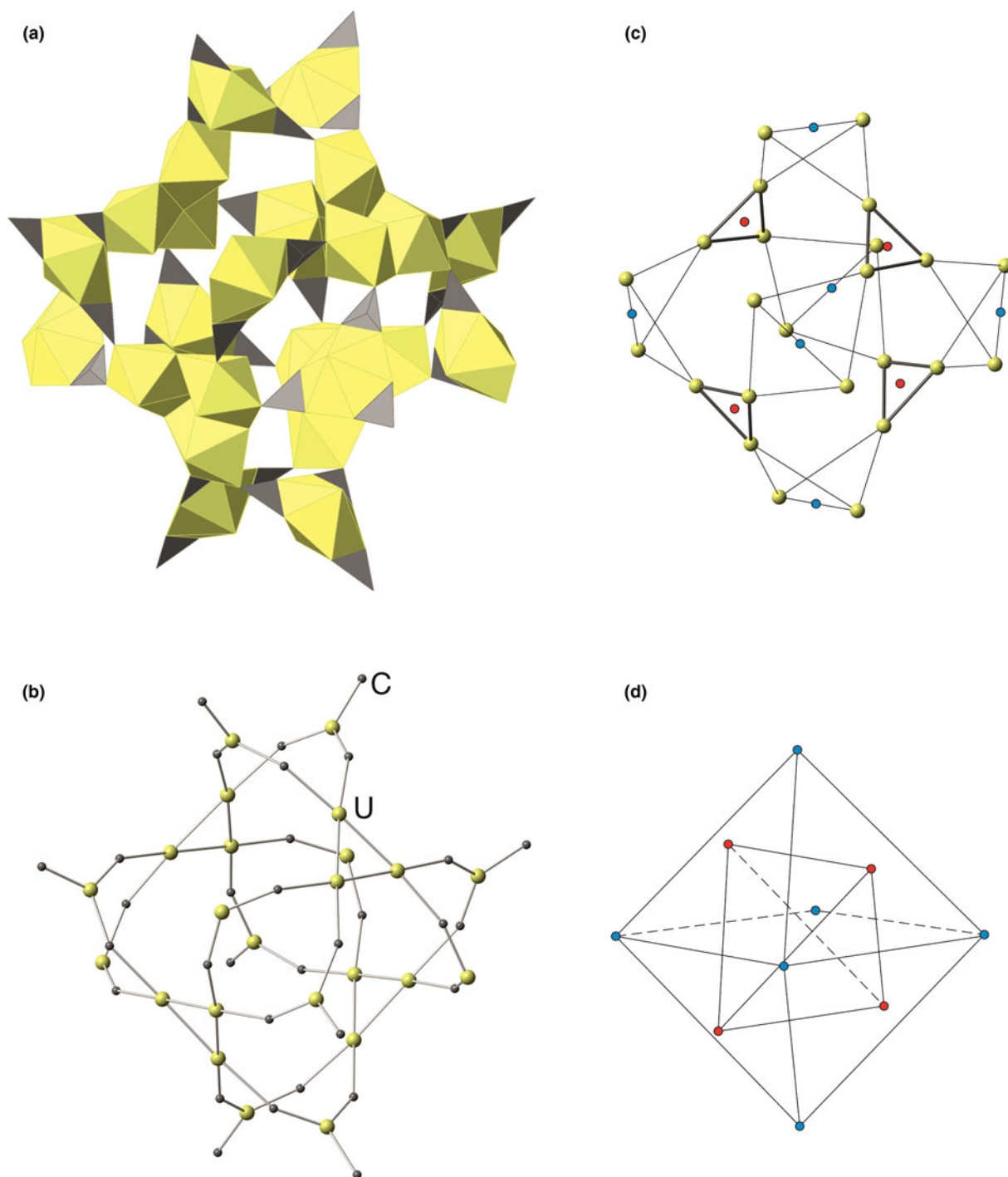
Fig. 6. The $[(UO_2)_{24}(CO_3)_{30}O_4(OH)_{12}(H_2O)_8]$ cluster in: (a) ewingite; (b) its skeletal representation; (c) the U core (the U$\cdots$U contacts shorter than 4 Å are shown as thick black lines; those between 4 and 6.2 Å as thin black lines; red and blue points indicate centres of the U_3 trimers and midpoints of the shared U$\cdots$U edges of the U_4 dihedra, respectively); and (d) the arrangements of the red and blue circles that correspond to the intersection of tetrahedron and octahedron, respectively. After Krivovichev (2020a).

(TPTS) clusters centred by Ce^{3+} in [9]-coordination (Fig. 8). There are two symmetry-independent clusters within the unit cell. In each cluster, three $[Ti_2O_{10}]$ dimers of edge-sharing TiO_6 octahedra in parallel orientation form a trigonal prism centred by Ce^{3+} cations. The triple-dimer titanate structure is surrounded by SiO_4 tetrahedra to form a TPTS cluster. Four adjacent TPTS clusters are linked into four-membered rings within the (010) plane and linked into ribbons parallel to $[\bar{1}01]$. The ribbons are organised into layers parallel to (010) and modulated with a modulation wavelength of 32.91 Å and an amplitude of 13.89 Å. The layers are further linked into a titanosilicate framework via additional SiO_4 tetrahedra. The Na^+, K^+, Ba^{2+} and H_2O groups occur in the framework cavities and have different occupancies and coordination environments.
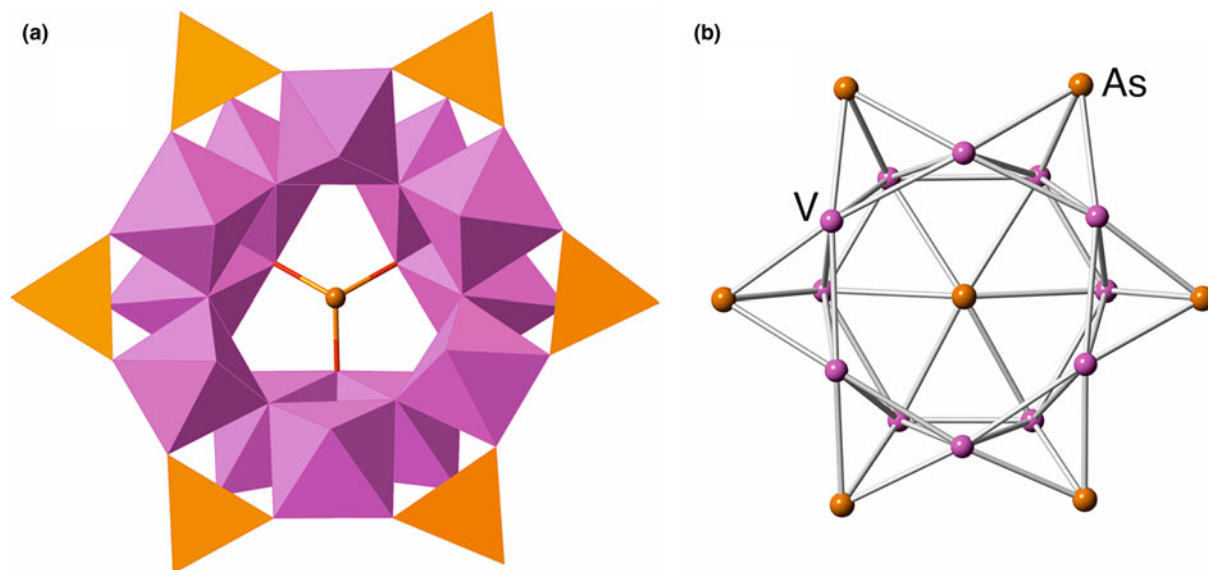
Fig. 7. The $[As^{3+}V_2^{4+}V_{10}^{5+}As_6^{5+}O_{51}]$ cluster in morrisonite in (a) polyhedral and (b) skeletal representations. Adapted from Krivovichev (2020a).

Complexity-generating mechanisms in minerals

The list of twenty most structurally complex minerals known so far given in Table 3 reveals the following most important complexity-generating mechanisms in minerals:

(1) the presence of large (sometimes nanometre-scale) clusters such as polyoxometalates or related finite-cluster structures (Krivovichev, 2020b) isolated from each other; such multi-nuclear atomic units possess a large number of atoms with different topological functions; the examples are ewingite, morrisonite, vanarsite, bouazzerite and postite; as a rule, natural polyoxometalate minerals are also highly hydrated, with the exception of arsmirandite and lehmannite, which are anhydrous polyoxocuprates formed in volcanic fumaroles (Britvin et al., 2020);

(2) the presence of large clusters linked together to form three-dimensional frameworks; the examples are ilmajokite and paddlewheelite (both minerals are also highly hydrated, which contributes greatly to their structural complexities);

(3) the formation of complex three-dimensional modular frameworks formed by cages of different sizes and topologies; this complexity type corresponds to paulingite-group minerals, fantappièite and sacrofanite (members of the sodalite–cancrinite ABC series (Bonaccorsi and Nazzareni, 2015; Chukanov et al., 2021), tschörtnerite, mendeleevite-(Ce) and rowleyite;

(4) the formation of complex layers with different combinations of modules (chains or rings); this type is characteristic for layered uranyl minerals (such as vandendriesscheite and gauthierite) and layered silicates (parsettensite);

(5) a high hydration state in salts with complex heteropolyhedral units (alfredstelznerite and voltaite-group minerals);

(6) the formation of ordered superstructures of relatively simple structure types; the examples are meerschautite (which is the only sulfide species in the list and was described by Biagioni et al. (2016) as an expanded derivative of owyheeite) and manitobaite (that has a fivefold superstructure relative to alluaudite (Tait et al., 2011)); for further discussion on the relations between superstructures and complexity see Krivovichev et al. (2019a, 2021), Kornyakov et al. (2021) and Kornyakov and Krivovichev (2021).

Complexity and symmetry

The Fedorov–Groth statistical law

Fedorov (1913, 1914) and Groth (1921) pointed out that, in general, symmetry of crystalline compounds correlates with their chemical complexity. Less complex compounds (e.g. native elements) tend, on average, to have higher symmetries than compounds consisting of two elements, etc. This empirical observation, known in Russian literature as a Fedorov–Groth law (Shafranovskii, 1973), was recently analysed by Krivovichev and Krivovichev (2020) using modern crystallographic data for minerals and the concepts of structural and chemical complexity of crystals formulated above. They have shown that there is a strong correlation (Fig. 9; confidence level >0.99) between the atomic chemical ($^{chem}I_G$) and structural ($^{str}I_G$) complexities, and the point-group order $|P_G|$ taken as a measure of symmetry ($|P_G|$ may be equal to 1, 2, 3, 4, 6, 8, 12, 16, 24 and 48). The relations between these parameters are described by the following exponential functions:

$$^{chem}I_G = 1.122 + 0.634 \times \exp$$
$$[-|P_G|/30.956], \text{ bits/atom } (R^2 = 0.979) \tag{41}$$

$$^{str}I_G = 1.580 + 3.532 \times \exp$$
$$[-|P_G|/10.718], \text{ bits/atom } (R^2 = 0.987) \tag{42}$$

No good correlations have been obtained for the amounts of chemical and structural information per formula and per reduced unit cell, versus the $|P_G|$ values, which is not surprising, as the point-group symmetry of a system is not a function of its size.
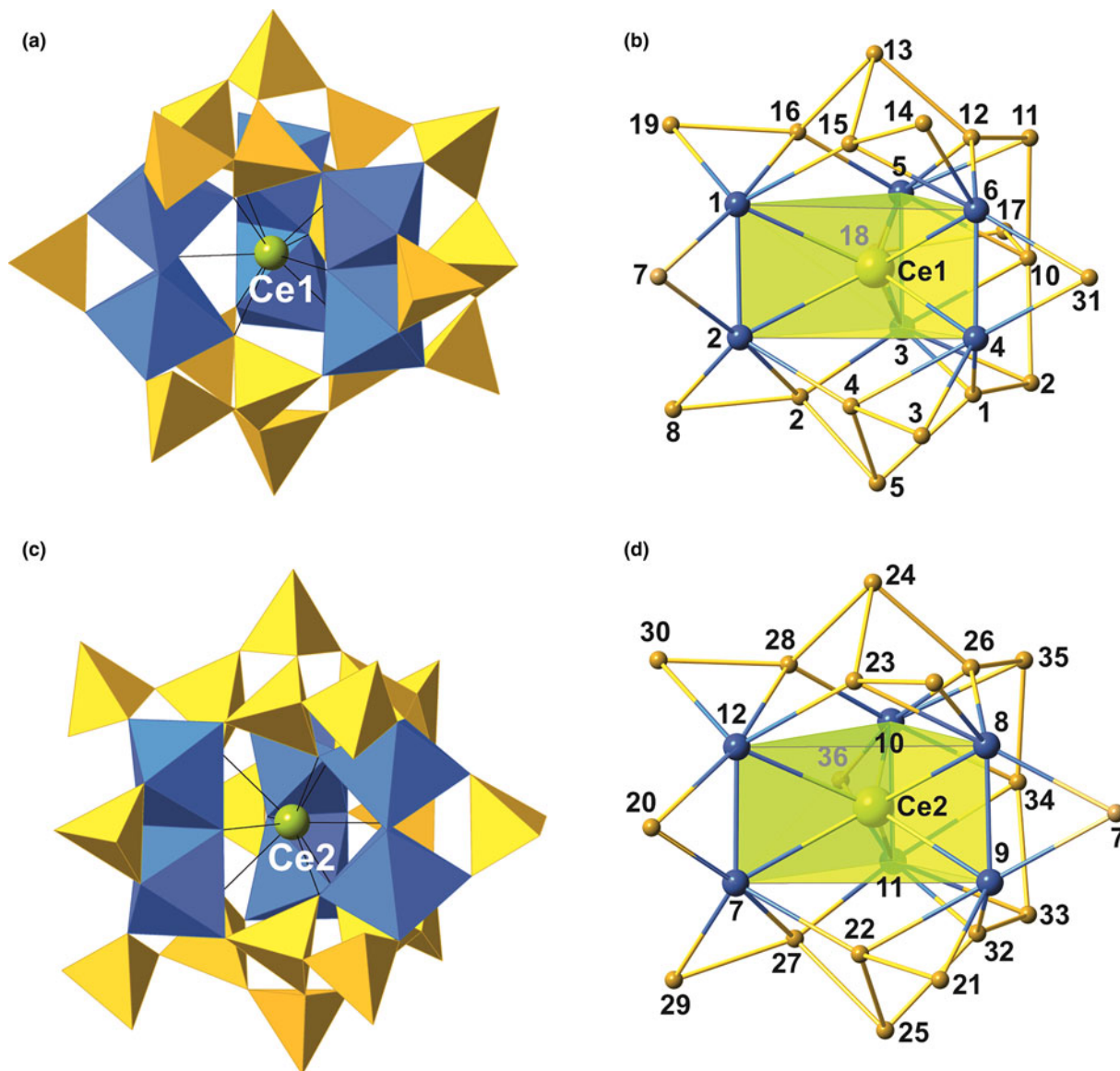
Fig. 8. Trigonal-prismatic titanosilicate (TPTS) clusters in the crystal structure of ilmajokite shown in (a, c) polyhedral and (b, d) nodal representations. The numbering scheme corresponds to the numbering of Si and Ti atoms from the experiment. The Ce-centred Ti_6 trigonal prism is highlighted in yellow. Legend: Si tetrahedra = yellow, Ti octahedra = blue; H_2O molecules, Na, K, Ba and Ce atoms are shown as red, light-blue, green, brown, and orange spheres, respectively (after Zolotarev *et al.*, 2020).

The observed strong correlation between the amount of chemical information per atom, $^{chem}I_G$, and the order of the point group, $|P_G|$, indicates that the Fedorov–Groth law is indeed valid and has a statistical meaning. Chemical simplicity measured as an amount of Shannon information per atom on average corresponds to higher symmetry measured as an order of the point group of a mineral.

Complexity and crystal systems

The distribution of minerals among the crystal systems and classes has been extensively discussed in the past, especially in the Russian mineralogical literature (Povarennykh, 1966; Shafranovskii and Feklichev, 1982; Shafranovskii, 1983a, 1983b, 1985, 1987, 1988; Dolivo-Dobrovolsky, 1987, 1988; Urusov, 2002, 2007; Filatov, 2020). Table 4 provides the modern analysis

of the problem that shows that the monoclinic system is most common (>35% of all minerals), followed by orthorhombic (~20%), trigonal (~11.5%) and triclinic (~10.6%) systems (we distinguish between trigonal and hexagonal systems as having threefold and sixfold rotation or inversion axes, respectively). In their analysis of the 1981 data, Shafranovskii and Feklichev (1982) reported the following sequence of crystal systems: monoclinic (30.8%), orthorhombic (23.4%), cubic (11.2%), hexagonal (9.0%), trigonal (8.7%), tetragonal (8.5%) and triclinic (8.4%). Shafranovskii (1983a, 1983b, 1987, 1988) suggested that the quantitative distribution of minerals among different crystal systems can be considered as a statistical 'law of nature', similar to the Fedorov–Groth law (see above), however Dolivo-Dobrovolsky (1988) challenged this view and showed that the number of cubic and triclinic minerals is decreasing and increasing with the historical time, respectively. Urusov (2002, 2007) supported
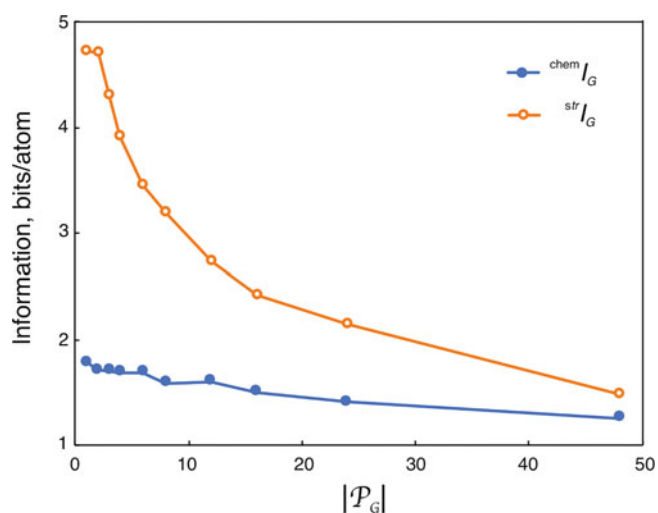
Fig. 9. The plot of average atomic chemical ($^{chem}I_G$) and structural ($^{str}I_G$) complexities of minerals versus the order of the point group, $|P_G|$. After Krivovichev and Krivovichev (2020).

this view by noticing that most of the rare minerals discovered within last few decades are complex and possess low symmetry. Filatov (2020) explained the high abundance of monoclinic minerals by the ability of the monoclinic system to adapt complex atomic arrangements, but did not provide any quantitative metrics to support his conclusion. Moreover, Filatov (2020) based his analysis on old statistical data (e.g. he used 30.5 and 7.4% for the monoclinic and triclinic systems, respectively, which do not correspond to the current situation described in Table 4).

The analysis of atomic chemical and structural complexities for different crystal systems shows general agreement with the Fedorov–Groth law with the exception of the trigonal system, which has relatively high values compared to tetragonal and hexagonal systems. The slightly anomalous behavior of the trigonal system is due to the fact that, surprisingly enough, this system is characteristic for many very complex minerals, including, e.g. the members of the cancrinite–sodalite supergroup.

Complexity and space groups

A list of most common space groups observed in minerals according to data for 2020, along with their average atomic structural complexities is shown in Table 5. As the $^{str}I_G$ parameter is symmetry-sensitive, it is not surprising that it correlates with the degree of symmetry described by corresponding space groups.

Shafranovskii and Feklichev (1982) analysed the relative abundance of different space groups in minerals based upon the 1981 data and reported the following ten space groups taken in the order of their abundance: $P2_1/c$ – $Pnma$ – $P\bar{1}$ – $C2/m$ – $C2/c$ – $R\bar{3}m$ – $P2_1/m$ – $Fd\bar{3}m$ – $Fm\bar{3}m$ – $P6_3/mmc$. Comparison with the data given in Table 5 indicates the number of monoclinic minerals increased significantly since 1981 relative to the numbers of orthorhombic, trigonal and cubic species. Analysis of Table 3 reveals that, among twenty most structurally complex minerals, four crystallise in the $P2_1/c$ space group, whereas all others correspond to unique space groups.

Urusov and Nadezhina (2006) listed 54 space groups that are 'empty' of minerals, i.e. have no mineral species crystallised with those symmetries. Our analysis (Table 6) shows that: (1) several 'empty' space groups were missed in the compilation and (2) several 'empty' space groups have been filled by new mineral species discovered since 2006. For instance, the $Pm\bar{3}$ space group, which was 'empty' in 2006, was 'filled' by the discovery of mendeleevite-(Ce) (Sokolova *et al.*, 2011). It is remarkable that the previous compilation of 'empty' mineralogical space groups given by Povarennykh (1966) listed 112 groups, i.e. slightly less than half of all possible space groups. In 2021, the number of 'empty' space groups (45) is less than 20% of the total number of possible space groups.

Recently Hummer (2021) investigated the distribution of mineral species among 32 point groups and demonstrated that the abundance of minerals belonging to each point group approximately obeys a power law with respect to group order, i.e. corresponds to fractal behaviour.

It is interesting that the order of space-group abundance in inorganic compounds (i.e. both minerals and synthetic compounds) differs from that observed in minerals and corresponds to the following sequence (Urusov and Nadezhina, 2009): $Pnma$ (7.4%), $P2_1/c$ (7.2%), $Fm\bar{3}m$ (5.6%), $Fd\bar{3}m$ (5.1%), $P\bar{1}$ (4.0%), $I4/mmm$ (4.0%), $C2/c$ (3.8%), $P6_3/mmc$ (3.4%), $C2/m$ (3.4%) and $Pm\bar{3}m$ (3.0%). These data show that the total kingdom of inorganic crystal structures is, on average, higher in symmetry relative to the mineral kingdom, which is rather remarkable. One may speculate that this tendency is due to the limited chemical diversity of natural systems compared to synthetic ones, which provides a broad avenue for synthetic chemical exploration of chemically simple, but naturally unstable or geochemically unusual combination of elements. In fact, synthetic chemists rarely study systems of more than three or four elements. Even though they consider combinations of elements not found in Nature, they very rarely attempt to study compounds in a system with seven or eight elements, as are often found in Nature. Therefore, the trend of fewer elements leading to higher average

Table 4. Distribution of mineral species among different crystal systems and their average atomic complexities.

System	O-based		O-free		All		$<^{chem}I_G>$, bits/atom	$<^{str}I_G>$, bits/atom
	n	%	n	%	n	%		
Triclinic	529	11.94	48	4.66	577	10.57	1.721	4.713
Monoclinic	1713	38.66	201	19.51	1914	35.05	1.701	4.036
Orthorhombic	857	19.34	228	22.13	1085	19.86	1.604	3.406
Tetragonal	275	6.21	109	10.58	384	7.03	1.541	2.545
Trigonal	512	11.55	117	11.36	629	11.52	1.643	3.130
Hexagonal	309	6.96	91	8.84	400	7.33	1.526	2.760
Cubic	236	5.34	236	22.92	472	8.64	1.361	1.780
Total	4431	100	1030	100	5461	100		

Table 5. Ten most common space groups for minerals.*

Space group	No.	n_{min}	System	~% from all minerals	$<^{str}I_G>$, bits/atom
$P2_1/c$	14	586	Monoclinic	10.7	4.051
$P\bar{1}$	2	502	Triclinic	9.2	4.711
$C2/m$	12	471	Monoclinic	8.6	3.812
$C2/c$	15	324	Monoclinic	5.9	3.822
$Pnma$	62	255	Orthorhombic	4.7	2.998
$P2_1/m$	11	196	Monoclinic	3.6	4.115
$R\bar{3}m$	166	146	Trigonal	2.7	2.601
$R3m$	160	105	Trigonal	1.9	3.435
$Fd\bar{3}m$	227	97	Cubic	1.8	1.598
$P1$	1	94	Triclinic	1.7	4.727

* n_{min} = number of minerals crystallising in the given space group

symmetry is reflected in synthetic inorganic chemistry. High temperatures of synthesis (compared to all of the low-temperature weathering/alteration minerals) may also play a role. We may hypothesise that the average chemical simplicity of synthetic systems results in their higher symmetry, according to the Fedorov–Groth law, which is statistically valid for large numbers of observations (see above).

For organic compounds, the $P2_1/c$ group is the most common (46%), followed by $P\bar{1}$ (20%), $C2/c$ (7%), $Pbca$ (6%), $P2_12_12_1$ (5%), $P2_1$ (3%), $Pna2_1$ (2%) and $Pca2_1$ (2%) (Rekis, 2020). The space-group distribution pattern here is drastically different from that observed in minerals and inorganic compounds, though the leading role of the $P2_1/c$ group is again remarkable.

As far as we know, there is no comprehensive explanation for the space-group distribution in crystals. Hellner and Sowa (1985) and Dolivo-Dobrovolsky (1988) pointed out that space groups with large percentages of characteristic orbits (i.e. orbits that possess symmetry of the generating space groups and not higher) have higher abundance. However, as was mentioned by Urusov and Nadezhina (2009), this cannot explain the predominance of the $P2_1/c$ space group that contains only one kind of characteristic orbit, whereas all its orbits of other kinds are non-characteristic.

Complexity, symmetry and temporal dynamics of mineralogical research

It was mentioned above that our information on the distributions of minerals among different crystal systems has changed significantly within the last four decades with the relative increase of the number of triclinic minerals. This tendency is related directly to the general tendency of temporal complexification of mineralogical objects in the history of mineralogical research. Kaußler

and Kieslich (2021) provided a temporal analysis of complexity parameters of crystal structures in the history of crystallography, demonstrating that the atomic structural complexities have increased continuously since the birth of X-ray diffraction analysis of crystals in 1914. With time, more and more complex crystal structures are solved and refined, thanks to the enormous advances in diffraction techniques, including recent advancements in detector technologies, X-ray sources and electron diffraction, as well as improved computational methods for the direct solution of complex crystal structures.

We have done a similar analysis of complexity of minerals in the history of mineralogy, starting from 1875, when Weisbach's *Synopsis mineralogica* was published that contained the full list of minerals (631 species) known at the time (Weisbach, 1875). The lists of minerals were then split into quarters of a century, where for each mineral the year of its discovery was assigned as the year of publication of its description according to bibliographic data provided in the International Mineralogical Association (IMA) Database of Mineral Properties (https://rruff.info/ima/). It is important to note that mineralogical knowledge was revised according to the IMA rules in existence at that time. For instance, in plagioclases, only two species were recognised, i.e. anorthite and albite, whereas their varieties such as bytownite, labradorite, andesine, and oligoclase were not considered as separate mineral species, despite their recognition as such in the 19th and 20th centuries. The temporal lists of minerals have been used to calculate their average complexities (arithmetic means). The numerical data are given in Table 7, whereas their visualisation is presented in Fig. 10. Fig. 10a shows the temporal dynamics of mineralogical discoveries since 1875 with steps of 25 years, whereas Figs 10b and c show the increasing chemical and structural complexity of human knowledge of the mineral kingdom in the history of mineralogy. With time, more and more chemically and structurally complex minerals have been discovered, which is exemplified by the family of natural polyoxometalates that belong to the most complex mineral species known so far (see above).

The temporal dynamics of distribution of known minerals among crystal systems is given in Table 8 and visualised in Fig. 11. It can be seen clearly that, in the history of mineralogy, the percentage of triclinic minerals has continuously increased, whereas that of cubic minerals has continuously decreased. It is interesting that the percentage of hexagonal minerals has also increased, though this trend is less pronounced compared to that of triclinic minerals. The most remarkable change is observed for cubic minerals; their percentage has decreased by ~5% since 1875. Our results reported here are in agreement with similar observations made by Dolivo-Dobrovolsky (1988), who also

Table 6. 'Empty' space groups (45) in structural mineralogy.*

System	Space groups
Orthorhombic	$F222$ (22), $I2_12_12_1$ (24), $Pcc2$ (27), $Ccc2$ (37), $Abm2$ (39), $Fmm2$ (42), $Pnnn$** (48), $Pccm$ (49)
Tetragonal	$P4$ (75), $I4_1$ (80), $P\bar{4}$ (81), $P4/m$ (83), $P422$ (89), $P4_222$*** (93), $P4_22_12$** (94), $I4_122$ (98), $P4_2cm$ (101), $P4_2nm$** (102), $P4cc$ (103), $P4nc$ (104), $P4_2mc$ (105), $P4_2bc$ (106), $I4mm$ (107), $I4cm$ (108), $I4_1md$ (109), $P\bar{4};c2$** (116), $P\bar{4}b2$ (117), $P\bar{4}n2$ (118), $I\bar{4}c2$ (120), $P4/nbm$ (125), $P4_2/ncm$ (138)
Hexagonal	$P6$ (168), $P6_2$ (171), $P6_4$ (172), $P6_122$ (178), $P6mm$** (183), $P6cc$ (184)
Cubic	$P23$ (195), $Fm\bar{3}$** (202), $P432$ (207), $P4_232$ (208), $F432$ (209), $F4_132$ (210), $I432$ (211), $Pn\bar{3}n$ (222), $Fm\bar{3}c$ (226)

* The space-group number is given in brackets
** Absent in the list of 'empty' groups by Urusov and Nadezhina (2006)
*** 'empty' group for all inorganic crystal structures

Table 7. Average chemical and structural complexity parameters for minerals in temporal dynamics.

| Year	n	$^{chem}I_G$ [bits/atom]		$^{chem}I_{G,total}$ [bits/f.u.]		n	$^{str}I_G$ [bits/atom]		$^{str}I_{G,total}$ [bits/cell]	
		$\bar{X}$	$\sigma_{\bar{X}}$	$\bar{X}$	$\sigma_{\bar{X}}$		$\bar{X}$	$\sigma_{\bar{X}}$	$\bar{X}$	$\sigma_{\bar{X}}$
2020	5467	1.62	0.004	63.22	1.20	4466	3.49	0.02	320.96	9.78
2000	3762	1.59	0.005	55.91	1.30	3267	3.36	0.02	279.58	8.11
1975	2335	1.56	0.008	48.71	1.31	2118	3.27	0.03	259.74	8.74
1950	1389	1.55	0.010	45.65	1.68	1317	3.23	0.04	249.39	10.48
1925	1103	1.53	0.012	43.85	1.79	1055	3.15	0.04	237.30	10.90
1900	875	1.51	0.014	42.82	2.12	834	3.06	0.05	228.23	11.83
1875	631	1.47	0.017	39.65	2.41	604	2.91	0.06	209.34	13.87
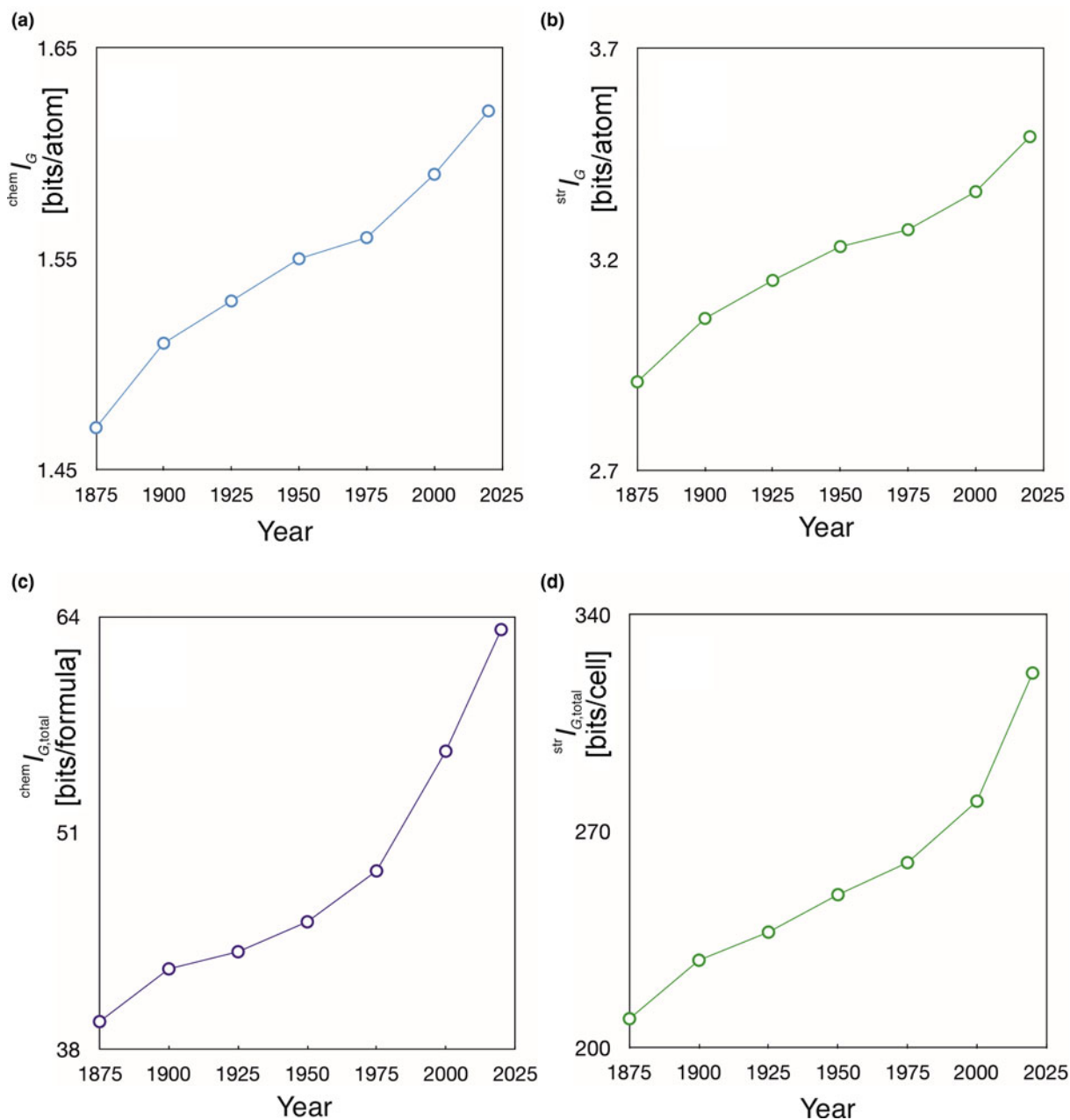
Fig. 10. Temporal dynamics of complexity in mineralogical research: average chemical and structural complexity parameters for the total lists of mineral species in different times.

Table 8. Temporal dynamics of the distribution of minerals among different crystal systems from 1875 to 2020 (in %).

System	Years						
	1875	1900	1925	1950	1975	2000	2020
Triclinic	7.2	7.6	8.1	8.6	9.4	9.8	10.6
Monoclinic	32.8	33.6	34.1	35.3	33.9	34.2	35.1
Orthorhombic	22.3	22.1	22.6	22.2	21.9	21.0	19.9
Tetragonal	7.6	7.6	6.8	6.9	7.6	7.5	7.0
Trigonal	11.1	11.1	11.4	11.0	10.5	11.2	11.5
Hexagonal	5.3	5.5	5.5	5.8	7.0	7.4	7.3
Cubic	13.7	12.6	11.5	10.2	9.7	8.9	8.6

demonstrated the rise of triclinic and the decline of cubic mineral species in the history of mineralogical research.

Complexity, paragenetic modes and mineral evolution

The idea that the mineral kingdom has changed in diversity and distribution with geological time dates back to the 1980s (Zhabin, 1981; Yushkin, 1982), but the most important revolutionary contribution to the field is due to the recent works by Hazen and Morrison, and coworkers (Hazen and Morrison, 2020, 2021a; Morrison and Hazen, 2020, 2021; Hazen *et al.*, 2021, etc.). Krivovichev *et al.* (2018a) used the chronological lists provided by Hazen *et al.* (2008) and Hazen (2013), who subdivided mineral evolution into four partially overlapping eras and ten stages, each of which saw the expansion of mineralogical diversity and/or variation in relative mineral abundances. The starting point of mineral evolution is that of the 'ur-minerals', the twelve earliest mineral phases to appear in the pre-solar nebulae (I). Chondritic meteorites incorporate ∼60 primary mineral phases, which constitute the second phase (II). For the Hadean Eon,

Hazen (2013) estimated 425 mineral species (III), whereas post-Hadean processes were responsible for the appearance of more than 5000 mineral species known today (IV). Figure 12 shows that both chemical and structural complexity are gradually increasing in the course of mineral evolution. However, one should take into account that many very complex minerals (especially surficial ones) have limited stability and are unlikely to survive for very long geological periods of time (Mills and Christy, 2019).

In a recent paper, Hazen and Morrison (2021b) subdivided all known minerals into 57 paragenetic modes, most of which can be assigned to a particular stage of mineral evolution. The average chemical and structural complexities for different paragenetic modes and respective stages of mineral evolution are given in Supplementary table S1. The evolutionary dynamics of mineral diversity and total structural and chemical complexities with the passage of geological time are shown in Figure 13. Both diversity and complexity increase dramatically in association with the formation of Earth's continental crust, the initiation of plate tectonics, and the Great Oxidation event. For paragenetic modes, the highest complexity is associated with secondary uranium mineralisation (P27 and P47f) and basalt-hosted zeolite crystallisation (P10). Both modes are associated with low-temperature hydrous activity, paralleling the conclusion by Hazen and Morrison (2021b) that "by far the most significant factor in enhancing Earth's mineral diversity has been its dynamic hydrological cycle".

Further directions

Structural complexity in Data-Driven Discovery

Hazen (2014) and Hazen *et al.* (2019) advocated Data-Driven Discovery as a new direction in the science of mineralogy that deals with the analysis of large volumes of data on chemical
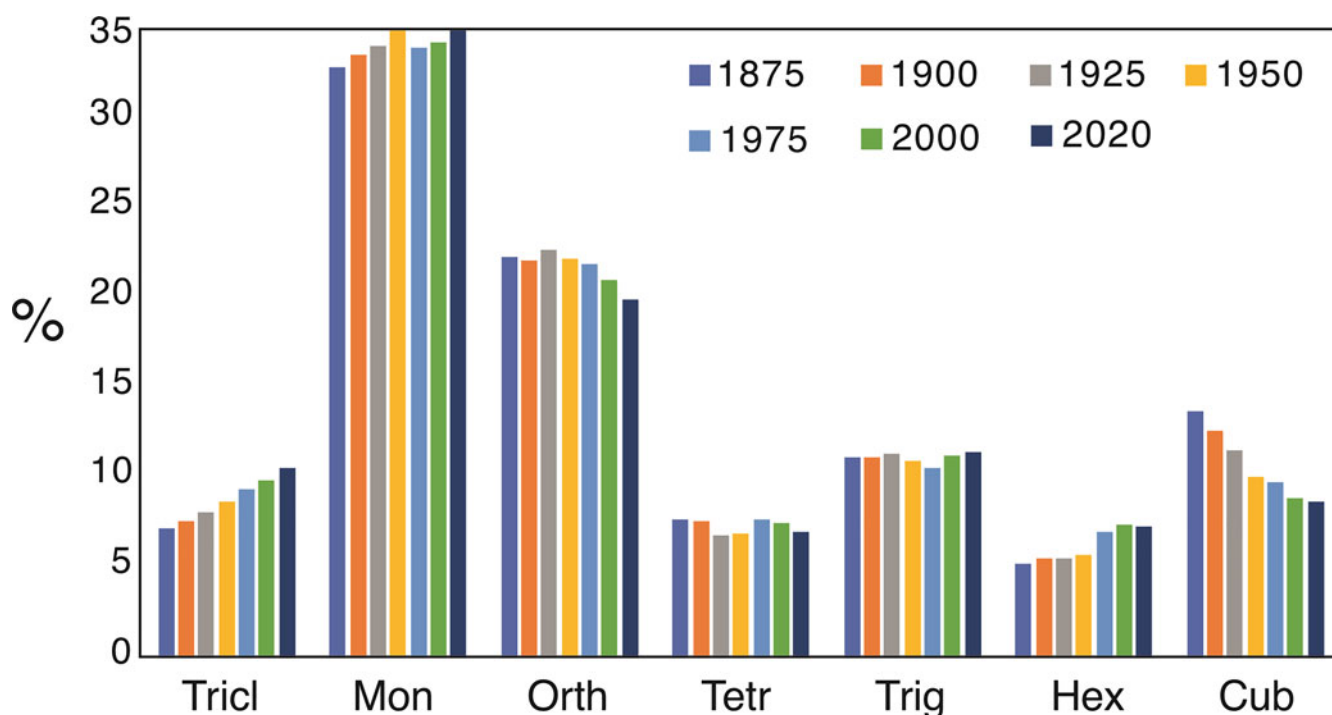

Fig. 11. Relative percentages of crystal systems for minerals known in different time periods of mineralogical research. Legend: Tricl = triclinic; Mon = monoclinic; Orth = orthorhombic; Tetr = tetragonal; Trig = trigonal; Hex = hexagonal; and Cub = cubic.
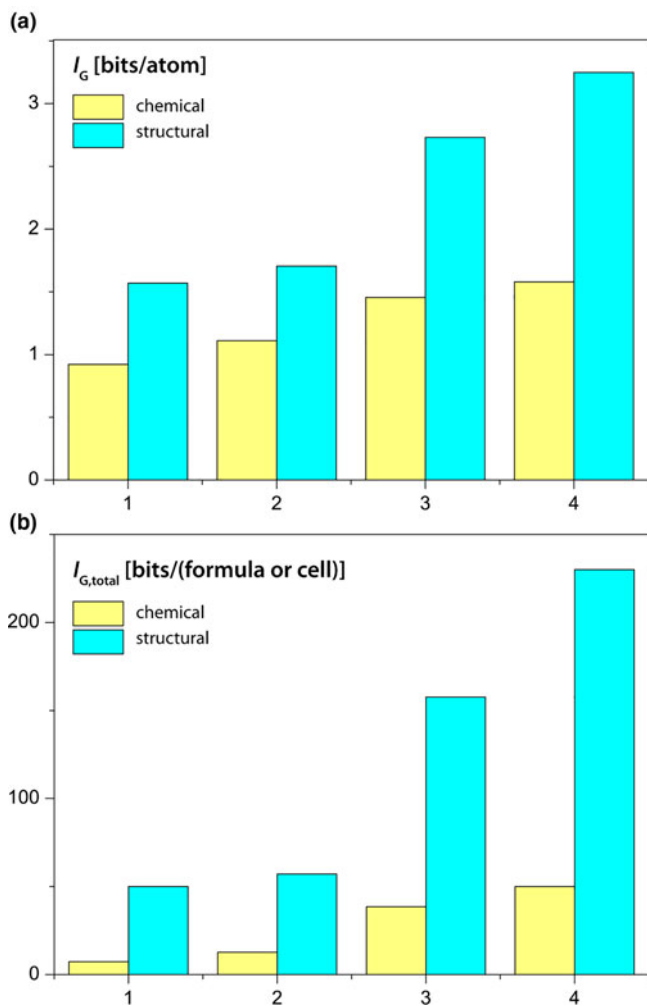
Fig. 12. Mean chemical and structural complexities for minerals occurring in different eras of mineral evolution (1 = 'ur-minerals'; 2 = minerals of chondritic meteorites; 3 = minerals of the Hadean epoch; 4 = minerals of the post-Hadean era): (a) Shannon information per atom and (b) per unit cell or formula unit. After Krivovichev *et al.* (2018a).

Fig. 13. The evolution of mineral diversity (blue), total structural (yellow) and total chemical (green) complexities at different stages of mineral evolution.

composition, crystal structure, physical properties and geological origins of minerals that have been accumulated in scientific literature over several last hundred years. The implementation of various numerical techniques (e.g. network analysis (Morrison *et al.*, 2017, 2020)) would provide us with new insights into distribution and diversity of minerals through space and time. Structural and chemical complexity are important characteristics of minerals reflecting their formation and evolution in geological environments. Incorporation of complexity data into modern mineralogical databases (such as GEMI (Prabhu *et al.*, 2021)) provides additional possibilities to extract novel and interesting regularities, due to the general relations between complexity and configurational entropy of crystalline solids (Krivovichev, 2016).

Structural complexity and thermodynamics

The relations between structural complexity and temperature and pressure have been discussed briefly by Krivovichev (2012, 2013) and is beyond the topic of the current review. However, several general points can be outlined here. As the structural complexity $^{str}I_{G,mix}$, defined by equations (27) and (29) as consisting of the
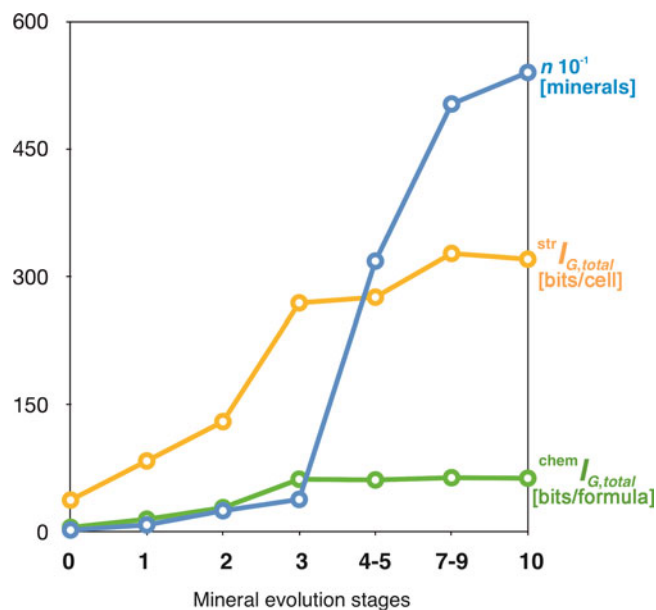
complexity of atomic architecture and positional disorder, is related to configurational entropy, and the latter is related to vibrational entropy, it can be expected that structural complexity decreases with increasing temperature and increases with increasing pressure. The first trend is confirmed by numerous data, including recent results on phase transitions in minerals (e.g. Avdonceva *et al.*, 2015a, 2015b). Increasing temperature also generally leads to an increase in structural disorder. Filatov (2011) emphasised the principle that increasing symmetry is often correlated with an increase of temperature. However, the principle of decreasing complexity with increasing temperature includes symmetry as one of its aspects and better describes the behaviour of solids under heating. For instance, Ismagilova *et al.* (2021) noticed that there are two natural $Cu_3(VO_4)_2$ polymorphs, pseudolyonsite and mcbirneyite, that correspond to high- and low-temperature modifications, respectively, which conforms well with their densities, 4.71 and 4.48 g·cm^{-3}, respectively. However, pseudolyonsite has higher (monoclinic) symmetry than mcbirneyite (triclinic polymorph), which is in obvious contradiction with the rule emphasised by Filatov (2011). In contrast, structural complexity calculations show that triclinic mcbirneyite is structurally simpler (36 bits/cell) than monoclinic pseudolyonsite (72 bits/cell), in agreement with the observation that, generally, complexity decreases upon heating. However, this empirically derived tendency may be violated as, for example, was shown for $BiSeO_3Cl$ (Aliev *et al.*, 2014). The low- and high-temperature modifications, α- and γ- $BiSeO_3Cl$, respectively, have the same complexity, 62 bits/cell. However, the intermediate β-phase is very complex with $^{str}I_{G,total} = 4706$ bits/cell. Aliev *et al.* (2014) showed that the extreme complexity of this phase is due to its transitional character, as its structure can be considered a 'frozen' transitional state between the low- and high-temperature forms. In general, transitional crystalline states existing on the border between stability fields of simple phases, may possess extraordinarily high complexity (Gurzhiy *et al.*, 2016).

Hazen and Navrotsky (1996) analysed the influence of pressure upon order–disorder reactions and showed that, in the

majority of cases, increasing pressure is associated with decreasing positional disorder, though there are exceptions to this rule (Deng *et al.*, 2020). The ordering reactions mean that structural complexity is increasing. Krivovichev (2021a) recently reviewed the crystal chemistry of high-pressure silicates and showed that, in many cases, structural complexity increases with pressure, but only when coordination numbers of atoms remain constant during phase transitions from low- to high-pressure modification. However, even then the situation is not straightforward, as can be seen from the high-pressure phase transitions of coesite, SiO_2. Bykova *et al.* (2018) reported five coesite polymorphs, numbered from I to V with increasing pressure. The coesite-I → coesite-II → coesite-III phase transitions are displacive and do not modify the topology of the framework; Si atoms remain in tetrahedral coordination. Along the pathway, the symmetry is decreasing ($C2/c → P2_1/n → P\bar{1}$), which corresponds to the increasing number of symmetrically independent sites and the increasing structural information per atom ($2.752 → 4.585 → 5.198$ bits/atom). However, the total structural information does not behave similarly ($66 → 440 → 374$ bits/cell). The coesite-IV and V polymorphs contain tetra-, penta- and hexa-coordinated Si and have the same complexity, 4.585 bits/atom and 220 bits/cell. The change in coordination number may modify the complexity behaviour unpredictably. The most probable reason is that the increase in pressure results in the increase of vibrational entropy, which is not a direct function of configurational entropy and structural complexity described by equations (27) and (28).

Structural complexity and crystallization

Krivovichev (2013) provided an overview of relations between structural complexity and metastable crystallisation based upon the Goldsmith's simplexity principle that states that simple structures crystallise more easily. This idea was supported by many subsequent observations, including metastable feldspar polymorphs (Zolotarev *et al.*, 2019; Krivovichev, 2020, and references therein), perovskite-group minerals (Zaitsev *et al.*, 2017), uranyl minerals (Plášil *et al.*, 2018b, 2018c), sulfates (Plášil *et al.*, 2017; Majzlan *et al.*, 2018), phosphates and arsenates (Krivovichev *et al.*, 2016; Krivovichev, 2017; Kolitsch *et al.*, 2020), $Cu_2(OH)_3Cl$ polymorphs (Krivovichev *et al.*, 2017), borates (Grew *et al.*, 2016), borosilicates (Cempírek *et al.*, 2016), oxalates (Huskić *et al.*, 2019), etc. In his review on metastable-mineral formation in oxidation zones and mine wastes, Majzlan (2020) recently noted that "a quantitative link…" between structural complexity and metastability "…is missing". Indeed, no equation could be written that would directly relate structural information and the ease of crystallisation, as complexity is not the only factor that defines the sequence of crystallisation. An example of non-linear behaviour is given by TiO_2 polymorphs. It is usually assumed that, under ambient conditions, anatase and brookite are metastable, whereas rutile is thermodynamically stable (Ranade *et al.*, 2002). The structural complexity parameters of the three phases do not conform to Goldsmith's principle, as anatase and rutile have the same complexity (0.918 bits/atom and 5.5 bits/cell), whereas brookite is the most complex among these TiO_2 polymorphs (1.585 bits/atom and 38 bits/cell). This example shows that the correlation between structural complexity and metastability is not straightforward and outlines an empirical tendency rather than rigorous regularity. The detailed account of the relations between complexity and metastability will be published elsewhere.

Structural complexity and modularity

Modularity is an important structural principle for minerals and inorganic compounds, as many mineral structures can be considered as built up from modules (blocks) derived from simple archetype structures and recombined in a particular way (Ferraris *et al.*, 2004). As one could imagine an infinite number of modular combinations, it could be expected that the number of possible derivative structures is infinite. However, only a few combinations are observed in Nature, and those are usually the simplest possible. Pankova *et al.* (2018b) showed that, in the kurchatovite family, only the two simplest variations are realised in Nature that can be generated according to the given structural principles of modular combinations. The principle of simplicity in modular series and its relation to abundances of structural architectures will be discussed in detail elsewhere.

Summary and outlook

Structural and chemical complexity is an important parameter relevant to the constitution, origin and behaviour of minerals. Quantitative measures of complexity that are applicable to any mineral are therefore useful numerical characteristics that can be used to reveal hidden relations among structure, composition and configuration entropy of crystalline solids, on one hand, and their stability and formation conditions on the other. The key concept behind the complexity parameters is that of information. Minerals do not have a material carrier of information, which is present in living organisms and viruses in the form of a genome, i.e. the sequence of nucleotides that govern the synthesis of proteins. Instead, information is encoded in minerals in various forms, from which an idealised crystal structure and idealised end-member formula are just the first and basic approximations. Further developments may include understanding and quantification of real crystal chemistry that includes chemical admixtures, short-range order, defects and information encoded in microstructural irregularities. The present review further emphasises the importance of the fundamental concept of information in mineralogy, which provides a different perspective to look at such traditional concepts as symmetry. From the viewpoint of an information approach, symmetry is seen as a way to minimise the information content, i.e. as the principle of information minimisation. The advances in experimental techniques provide new and deeper understanding of mineralogical objects, which is evident, in particular, in the modern developments in structural and descriptive mineralogy. For instance, the new discoveries of polyoxometalate minerals in the 21st century witness that the role of polynuclear atomic clusters in geochemical processes may be underestimated and that new models are needed (Rustad, 2010; Friis and Casey, 2018). To conclude, the developments in complexity theory of minerals provide a new way to look at mineralogical objects, which complement the existing theoretical and experimental approaches.

References

Aliev A., Kovrugin V.M., Colmont M., Terryn C., Huvé M., Siidra O.I., Krivovichev S.V. and Mentré O. (2014) Revised bismuth chloroselenite system: evidence of a noncentrosymmmetric structure with a giant unit cell. *Crystal Growth & Design*, **14**, 3026–3034.

Avdontceva M.S., Krzhizhanovskaya M.G., Krivovichev S.V. and Yakovenchuk V.N. (2015a) High-temperature order–disorder phase transition in nacaphite, Na_2CaPO_4F. *Physics and Chemistry of Minerals*, **42**, 671–676.

Avdontceva M.S., Zolotarev A.A. and Krivovichev S.V. (2015b) Order-disorder phase transition in the antiperovskite-type structure of synthetic kogarkoite, Na_3SO_4F. *Journal of Solid State Chemistry*, **231**, 42–46.

Banaru A.M., Aksenov S.M. and Krivovichev S.V. (2021) Complexity parameters for molecular solids. *Symmetry*, **13**, 1399.

Biagioni C., Moëlo Y., Orlandi P. and Stanley C.J. (2016) Lead-antimony sulfosalts from Tuscany (Italy). XVII. Meerschautite, $(Ag,Cu)_{5.5}Pb_{42.4}(Sb, As)_{45.1}S_{112}O_{0.8}$, a new expanded derivative of owyheeite from the Pollone mine, Valdicastello Carducci: occurrence and crystal structure. *Mineralogical Magazine*, **80**, 675–690.

Bindi L., Nespolo M., Krivovichev S.V., Chapuis G. and Biagioni C. (2020) Producing highly complicated materials. Nature does it better. *Reports on Progress in Physics*, **83**, 106501.

Bonaccorsi E. and Nazzareni S. (2015) Crystal chemical models for the cancrinite-sodalite supergroup: the structure of a new 18-layer phase. *Zeitschrift für Kristallographie*, **230**, 345–351.

Britvin S.N., Pekov I.V., Yapaskurt V.O., Koshlyakova N.N., Göttlicher J., Krivovichev S.V., Turchkova A.G. and Sidorov E.G. (2020) Polyoxometalate chemistry at volcanoes: discovery of a novel class of polyoxocuprate nanoclusters in fumarolic minerals. *Scientific Reports*, **10**, 6345.

Brugger J., Meisser N., Krivovichev S., Armbruster T. and Favreau G. (2007) Mineralogy and crystal structure of bouazzerite from Bou Azzer, Anti-Atlas, Morocco: Bi–As–Fe nanoclusters containing Fe^{3+} in trigonal prismatic coordination. *American Mineralogist*, **92**, 1630–1639.

Bulkin G.A. (1972a) *Introduction to Statistical Geochemistry. Application of Information Theory to Geochemistry*. Nedra, Leningrad, 207 pp. [in Russian].

Bulkin G.A. (1972b) Complexity of geochemical systems and their equilibrium. *Doklady Akademii Nauk SSSR*, **204**, 956–959 [in Russian].

Burns P.C. (1997) A new uranyl oxide hydrate sheet in vandendriesscheite: implications for mineral paragenesis and the corrosion of spent nuclear fuel. *American Mineralogist*, **82**, 1176–1186.

Bussen I.V., Gannibal L.F., Goiko E.A., Mer'kov A.N. and Nedorezova A.P. (1972) Ilmajokite, a new mineral from the Lovozero Tundra. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **101**, 75–79.

Bykova E., Bykov M., Černok A., Tidholm J., Simak S.I., Hellman O., Belov M.P., Abrikosov I.A., Liermann H.-P., Hanfland M., Prakapenka V.B., Prescher C., Dubrovinskaia N. and Dubrovinsky L. (2018) Metastable silica high pressure polymorphs as structural proxies of deep Earth silicate melts. *Nature Communications*, **9**, 4789.

Cámara F., Bellatreccia F., Della Ventura G., Mottan A., Bindi L., Gunter M.E. and Sebastiani M. (2010) Fantappièite, a new mineral of the cancrinite-sodalite group with a 33-layer stacking sequence: occurrence and crystal structure. *American Mineralogist*, **95**, 472–480.

Cempírek J., Grew E.S., Kampf A.R., Ma C., Novák M., Gadas P., Škoda R., Vašinová-Galiová M., Pezzotta F., Groat L.A. and Krivovichev S.V. (2016) Vránaite, ideally $Al_{16}B_4Si_4O_{38}$, a new mineral related to boralsilite, $Al_{16}B_6Si_2O_{37}$, from the Manjaka pegmatite, Sahatany Valley, Madagascar. *American Mineralogist*, **101**, 2108–2117.

Christy A.G., Mills S.J., Kampf A.R., Housley R.M., Thorne B. and Marty J. (2016) The relationship between mineral composition, crystal structure and paragenetic sequence: the case of secondary Te mineralization at the Bird Nest drift, Otto Mountain, California, USA. *Mineralogical Magazine*, **80**, 291–310.

Chukanov N.V., Aksenov S.M. and Rastsvetaeva R.K. (2021) Structural chemistry, IR spectroscopy, properties, and genesis of natural and synthetic microporous cancrinite- and sodalite-related materials: A review. *Microporous and Mesoporous Materials*, **323**, 111098.

Colmenero F., Plášil J. and Sejkora J. (2019) The layered uranyl silicate mineral uranophane-β: crystal structure, mechanical properties, Raman spectrum and comparison with the α-polymorph. *Dalton Transactions*, **48**, 16722–16736.

Cooper M.A., Hawthorne F.C., Galliski M.A. and Márquez-Zavalía M.F. (2010) The crystal structure of alfredstelznerite, $Ca_4(H_2O)_4[B_4O_4(OH)_6]_4(H_2O)_{15}$, a complex hydroxy-hydrated calcium borate mineral. *The Canadian Mineralogist*, **48**, 129–138.

Delgado-Friedrichs O. and O'Keeffe M. (2003) Identification of and symmetry computation for crystal nets. *Acta Crystallographica*, **A59**, 351–360.

Delgado-Friedrichs O., Hyde S.T., O'Keeffe M. and Yaghi O.M. (2017) Crystal structures as periodic graphs: the topological genome and graph databases. *Structural Chemistry*, **28**, 39–44.

Deng Z., Kang C.-J., Croft M., Li W., Shen X., Zhao J., Yu R., Jin C., Kotliar G., Liu S., Tyson T.A., Tappero R. and Greenblatt M. (2020) A pressure-induced inverse order–disorder transition in double perovskites. *Angewandte Chemie International Edition*, **59**, 8240–8246.

Dippo O.F. and Vecchio K.S. (2021) A universal configurational entropy metric for high-entropy materials. *Scripta Materialia*, **201**, 113974.

Dolivo-Dobrovolskii V.V. (1987) On the distribution of mineral species according to their symmetry. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **116**, 7–17 [in Russian].

Dolivo-Dobrovolskii V.V. (1988) On the so-called 'laws of statistical mineralogy'. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **117**, 387–393 [in Russian].

Effenberger H., Giester G., Krause W. and Bernhardt H.-J. (1998) Tschörtnerite, a copper-bearing zeolite from the Bellerberg volcano, Eifel, Germany. *American Mineralogist*, **83**, 607–617.

Fedorov E.S. (1913) Crystals of the cubic syngony. *Zapiski Gornogo Instituta*, **4**, 312–320 [in Russian].

Fedorov E.S. (1914) Weitere krystallochemische Behlerungen an der Hand der Tabellen zur krystallochemischen Analyse. *Zeitschrift fuer Krystallographie*, **53**, 337–388.

Ferraris G., Makovicky E. and Merlino S. (2004) *Crystallography of Modular Materials*. IUCr Monographs in Crystallography, Oxford University Press, Oxford, UK, 370 pp.

Filatov S.K. (2011) General concept of increasing crystal symmetry with an increase in temperature. *Crystallography Reports*, **56**, 953–961.

Filatov S.K. (2020) Symmetry statistics of mineral species in various thermodynamic conditions. *Zapiski Rossiyskogo Mineralogicheskogo Obshchetsva*, **148**, 1–13 [in Russian].

Friis H. and Casey W.H. (2018) Niobium is highly mobile as a polyoxometalate ion during natural weathering. *The Canadian Mineralogist*, **56**, 905–912.

Gordon E.K., Samson S. and Kamb W.B. (1966) Crystal structure of the zeolite paulingite. *Science*, **154**, 1004–1007.

Grew E.S., Krivovichev S.V., Hazen R.M. and Hystad G. (2016) Evolution of structural complexity in boron minerals. *The Canadian Mineralogist*, **54**, 125–143.

Groth P. (1921) *Elemente der Physikalischen und Chemischen Krystallographie*. R. Oldenbourg, München, Berlin, pp. 274–278.

Guggenheim S. and Eggleton R.A. (1994) A comparison of the structures and geometric stabilities of stilpnomelane and parsettensite: a distance least-squares (DLS) study. *American Mineralogist*, **79**, 438–442.

Gurzhiy V.V. and Plášil J. (2019) Structural complexity of natural uranyl sulfates. *Acta Crystallographica*, **B75**, 39–48.

Gurzhiy V.V., Tyumentseva O.S., Krivovichev S.V., Krivovichev V.G. and Tananaev I.G. (2016) Mixed uranyl sulfate–selenates: evolution of structural topology and complexity vs chemical composition. *Crystal Growth & Design*, **16**, 4482–4492.

Gurzhiy V.V., Kuporev I.V., Kovrugin V.M., Murashko M.N. Kasatkin A.V. and Plášil J., (2019) Crystal chemistry and structural complexity of natural and synthetic uranyl selenites. *Crystals*, **9**, 639.

Gurzhiy V.V., Kalashnikova S.A., Kuporev I.V. and Plášil J. (2021) Crystal chemistry and structural complexity of the uranyl carbonate minerals and synthetic compounds. *Crystals*, **11**, 704.

Hawthorne F.C. (2021) Proof that a dominant endmember formula can always be written for a mineral or a crystal structure. *The Canadian Mineralogist*, **59**, 159–167.

Hawthorne F.C., Mills S.J., Hatert F. and Rumsey M.S. (2021) Ontology, archetypes and the definition of 'mineral species'. *Mineralogical Magazine*, **85**, 125–131.

Hazen R.M. (2013) Paleomineralogy of the Hadean Eon: A preliminary list. *American Journal of Science*, **313**, 807–843.

Hazen R.M. (2014) Data-driven abductive discovery in mineralogy. *American Mineralogist*, **99**, 2165–2170.

Hazen R.M. and Morrison S.M. (2020) An evolutionary system of mineralogy. Part I: Stellar mineralogy (>13 to 4.6 Ga). *American Mineralogist*, **105**, 627–651.

Hazen R.M. and Morrison S.M. (2021a) An evolutionary system of mineralogy. Part V: Aqueous and thermal alteration of planetesimals (∼4565 to 4550 Ma). *American Mineralogist*, **106**, 1388–1419.

Hazen R.M. and Morrison S.M. (2021b) On the paragenetic modes of minerals: a mineral evolution perspective. *American Mineralogist*, in press, https://doi.org/10.2138/am-2022-8099.

Hazen R.M. and Navrotsky A. (1996) Effects of pressure on order-disorder reactions. *American Mineralogist*, **81**, 1021–1035.

Hazen R.M., Papineau D., Bleeker W., Downs R.T., Ferry J., McCoy T., Sverjensky D. and Yang H. (2008) Mineral evolution. *American Mineralogist*, **93**, 1693–1720.

Hazen R.M., Downs R.T., Eleish A., Fox P., Gagné O.C., Golden J.J., Grew E.S., Hummer D.R., Hystad G., Krivovichev S.V., Li C., Liu C., Ma X., Morrison S.M., Pan F., Pires A.J., Prabhu A., Ralph J., Runyon S.E., Zhong H. (2019) Data-driven discovery in mineralogy: recent advances in data resources, analysis, and visualization. *Engineering*, **5**, 397–405.

Hazen R.M., Morrison S.M. and Prabhu A. (2021) An evolutionary system of mineralogy. Part III: Primary chondrule mineralogy (4566 to 4561 Ma). *American Mineralogist*, **106**, 325–350.

Hellner E. and Sowa H. (1985) The cubic structure types described in their space groups with the aid of frameworks. *Physik Daten – Physics Data*, **16–3**. 194 pp.

Hornfeck W. (2020) On an extension of Krivovichev's complexity measures. *Acta Crystallographica*, **A76**, 534–548.

Hummer D.R. (2021) Fractal distribution of mineral species among the crystallographic point groups. *American Mineralogist*, **106**, 1574–1579.

Huskić I., Novendra N., Lim D.-W., Topić F., Titi H.M., Pekov I.V., Krivovichev S.V., Navrotsky A., Kitagawa H. and Friščić T. (2019) Functionality in metal-organic framework minerals: proton conductivity, stability and potential for polymorphism. *Chemical Science*, **10**, 4923–4929.

Ismagilova R.M., Zhitova E.S., Krivovichev S.V., Sergeeva A.V., Nuzhdaev A.A., Anikin L.P., Krzhizhanovskaya M.G., Nazarova M.A., Kupchinenko A.N. and Zolotarev A.A. (2021) Thermal evolution of volborthite, $Cu_3(V_2O_7)(OH)_2 \cdot 2H_2O$. *Minerals*, **11**, 1312, https://doi.org/10.3390/min11121312

Kampf A.R., Hughes J.M., Marty J. and Nash B. (2012) Postite, $Mg(H_2O)_6Al_2(OH)_2(H_2O)_8(V_{10}O_{28}) \cdot 13H_2O$, a new mineral species from the La Sal mining district, Utah: crystal structure and descriptive mineralogy. *The Canadian Mineralogist*, **50**, 45–53.

Kampf A.R., Hughes J.M., Nash B.P. and Marty J. (2016) Vanarsite, packratite, morrisonite, and gatewayite: four new minerals containing the $[As^{3+}V^{4+,5+}_{12}As^{5+}_6O_{51}]$ heteropolyanion, a novel polyoxometalate cluster. *The Canadian Mineralogist*, **54**, 145–162.

Kampf A.R., Cooper M.A., Nash B.P., Cerling T.E., Marty J., Hummer D.R., Celestian A.J., Rose T.P. and Trebisky T.J. (2017) Rowleyite, $[Na(NH_4,K)_9Cl_4][V^{5+,4+}_2(P,As)O_8]_6 \cdot n[H_2O,Na,NH_4,K,Cl]$, a new mineral with a microporous framework structure. *American Mineralogist*, **102**, 1037–1044.

Kaußler C. and Kieslich G. (2021) CrystIT: Complexity and configurational entropy of crystal structures via information theory. *Journal of Applied Crystallography*, **54**, 306–316.

Kolitsch U., Weil M., Kovrugin V.M. and Krivovichev S.V. (2020) Crystal chemistry of the variscite and metavariscite groups: Crystal structures of synthetic $CrAsO_4 \cdot 2H_2O$, $TlPO_4 \cdot 2H_2O$, $MnSeO_4 \cdot 2H_2O$, $CdSeO_4 \cdot 2H_2O$ and natural bonacinaite, $ScAsO_4 \cdot 2H_2O$. *Mineralogical Magazine*, **84**, 568–583.

Kornyakov I.V. and Krivovichev S.V. (2021) Crystal chemical relations in the shchurovskyite family: Synthesis and crystal structures of $K_2Cu[Cu_3O]_2(PO_4)_4$ and $K_{2.35}Cu_{0.825}[Cu_3O]_2(PO_4)_4$. *Crystals*, **11**, 807.

Kornyakov I.V., Vladimirova V.A., Siidra O.I. and Krivovichev S.V. (2021) Expanding the averievite family, $(MX)Cu_5O_2(T^{5+}O_4)_2$ (T^{5+} = P, V; M = K, Rb, Cs, Cu; X = Cl, Br): Synthesis and single-crystal X-ray diffraction study. *Molecules*, **26**, 1833.

Krivovichev S.V. (2012) Topological complexity of crystal structures: quantitative approach. *Acta Crystallographica*, **A68**, 393–398.

Krivovichev S.V. (2013) Structural complexity of minerals: information storage and processing in the mineral world. *Mineralogical Magazine*, **77**, 275–326.

Krivovichev S.V. (2014) Which inorganic structures are the most complex? *Angewandte Chemie International Edition*, **53**, 654–661.

Krivovichev S.V. (2016) Structural complexity and configurational entropy of crystalline solids. *Acta Crystallographica*, **B72**, 274–276.

Krivovichev S.V. (2017) Hydrogen bonding and structural complexity of the $Cu_3(AsO_4)(OH)_3$ polymorphs (clinoclase, gilmarite): A theoretical study. *Journal of Geosciences*, **62**, 79–85.

Krivovichev S.V. (2018) Ladders of information: What contributes to the structural complexity in inorganic crystals. *Zeitschrift für Kristallographie*, **233**, 155–161.

Krivovichev S.V. (2020a) Feldspar polymorphs: diversity, complexity, stability. *Zapiski Rossiiskogo Mineralogicheskogo Obshchestva*, **149**, 16–66.

Krivovichev S.V. (2020b) Polyoxometalate clusters in minerals: review and complexity analysis. *Acta Crystallographica*, **B76**, 618–629.

Krivovichev S.V. (2021a) High-pressure silicates: crystal chemistry and systematics. *Zapiski Rossiiskogo Mineralogicheskogo Obshchestva*, **150**, 1–78.

Krivovichev V.G. (2021b) *Mineral Species*. St. Petersburg: St. Petersburg University Press, 600 pp. [in Russian].

Krivovichev V.G. and Charykova M.V. (2013) *Classification of Mineral Systems*. St. Petersburg: St.-Petersburg University Press, 196 pp. [in Russian].

Krivovichev S.V. and Krivovichev V.G. (2020) The Fedorov-Groth law revisited: complexity analysis using mineralogical data. *Acta Crystallographica*, **A76**, 429–431.

Krivovichev S.V., Zolotarev A.A. and Popova V.I. (2016) Hydrogen bonding and structural complexity in the $Cu_5(PO_4)_2(OH)_4$ polymorphs (pseudomalachite, ludjibaite, reichenbachite): combined experimental and theoretical study. *Structural Chemistry*, **27**, 1715–1723.

Krivovichev S.V., Hawthorne F.C. and Williams P.A. (2017) Structural complexity and crystallization: the Ostwald sequence of phases in the $Cu_2(OH)_3Cl$ system (botallackite–atacamite–clinoatacamite). *Structural Chemistry*, **28**, 153–159.

Krivovichev S.V., Krivovichev V.G. and Hazen R.M. (2018a) Structural and chemical complexity of minerals: correlations and time evolution. *European Journal of Mineralogy*, **30**, 231–236.

Krivovichev V.G., Charykova M.V. and Krivovichev S.V. (2018b) The concept of mineral systems and its application to the study of mineral diversity and evolution. *European Journal of Mineralogy*, **30**, 219–230.

Krivovichev S.V., Panikorovskii T.L., Zolotarev A.A., Bocharov V.N., Kasatkin A.V. and Škoda R. (2019a) Jahn-Teller distortion and cation ordering: the crystal structure of paratooite-(La), a superstructure of carbocernaite. *Minerals*, **9**, 370.

Krivovichev V.G., Krivovichev S.V. and Charykova M.V. (2019b) Selenium minerals: Structural and chemical diversity and complexity. *Minerals*, **9**, 455.

Krivovichev V.G., Charykova M.V. and Krivovichev S.V. (2020a) Mineral systems based on the number of species-defining chemical elements in minerals: their diversity, complexity, distribution, and the mineral evolution of the Earth's crust: a review. *Geology of Ore Deposits*, **62**, 704–718.

Krivovichev V.G., Krivovichev S.V. and Charykova M.V. (2020b) Tellurium minerals: Structural and chemical diversity and complexity. *Minerals*, **10**, 623.

Krivovichev S.V., Panikorovskii T.L., Yakovenchuk V.N., Selivanova E.A. and Ivanyuk G.Y. (2021) Trigonal variation in the garnet supergroup: the crystal

structure of nikmelnikovite, $Ca_{12}Fe^{2+}Fe_3^{3+}Al_3(SiO_4)_6(OH)_{20}$, from Kovdor massif, Kola Peninsula, Russia. *Mineralogical Magazine*, **85**, 620–626.

Limpert E., Stahel W.A. and Abbt M. (2001) Log-normal distributions across the sciences: keys and clues. *BioScience*, **51**, 341–352.

Majzlan J. (2020) Processes of metastable-mineral formation in oxidation zones and mine waste. *Mineralogical Magazine*, **84**, 367–375.

Majzlan J., Schlicht H., Wierzbicka-Wieczorek M., Giester G., Pöllmann H., Brömme B., Doyle S., Buth G. and Bender Koch C. (2013) A contribution to the crystal chemistry of the voltaite group: solid solutions, Mössbauer and infrared spectra, and anomalous anisotropy. *Mineralogy and Petrology*, **107**, 221–233.

Majzlan J., Dachs E., Benisek A., Plášil J. and Sejkora J. (2018) Thermodynamics, crystal chemistry and structural complexity of the Fe$(SO_4)(OH)(H_2O)_x$ phases: $Fe(SO_4)(OH)$, metahohmannite, butlerite, parabutlerite, amarantite, hohmannite, and fibroferrite. *European Journal of Mineralogy*, **30**, 259–275.

McCormack S.J. and Navrotsky A. (2021) Thermodynamics of high entropy oxides. *Acta Materialia*, **202**, 1–21.

McDonald A.M. and Chao G.-Y. (2010) Rogermitchellite, $Na_{12}(Sr, Na)_{24}Ba_4Zr_{26}Si_{78}(B,Si)_{12}O_{246}(OH)_{24}18(H_2O)$, a new mineral species from Mont Saint-Hilaire, Quebec: description, structure determination and relationship with HFSE-bearing cyclosilicates. *The Canadian Mineralogist*, **48**, 267–278.

Mereiter K. (1972) Die Kristallstruktur des Voltaits, $K_2Fe_5^{2+}Fe_3^{3+}Al[SO_4]_{12}18H_2O$. *Tschermaks Mineralogische und Petrographische Mitteilungen*, **18**, 185–202.

Mills S. and Christy A. (2019) Mineral extinction. *Mineralogical Magazine*, **83**, 621–625.

Miracle D.B. and Senkov O.N. (2017) A critical review of high entropy alloys and related concepts. *Acta Materialia*, **122**, 448–511.

Morrison S.M. and Hazen R.M. (2020) An evolutionary system of mineralogy. Part II: Interstellar and solar nebula primary condensation mineralogy (>4.565 Ga). *American Mineralogist*, **105**, 1508–1535.

Morrison S.M. and Hazen R.M. (2021) An evolutionary system of mineralogy, Part IV: Planetesimal differentiation and impact mineralization (4566 to 4560 Ma). *American Mineralogist*, **106**, 730–761.

Morrison S.M., Liu C., Eleish A., Prabhu A., Li C., Ralph J., Downs R.T., Golden J.J., Fox P., Hummer D.R., Meyer M.B. and Hazen R.M. (2017) Network analysis of mineralogical systems. *American Mineralogist*, **102**, 1588–1596.

Morrison S.M., Buongiorno J., Downs R.T., Eleish A., Fox P., Giovannelli D., Golden J.J., Hummer D.R., Hystad G., Kellogg L.H., Kreylos O., Krivovichev S.V., Liu C., Merdith A., Prabhu A., Ralph J., Runyon S.E., Zahirovic S. and Hazen R.M. (2020) Exploring carbon mineral systems: recent advances in C mineral evolution, mineral ecology, and network analysis. *Frontiers in Earth Science*, **8**, 208.

Mugnaioli E., Bonaccorsi E., Lanza A.E., Elkaim E., Diez-Gómez V., Sobrados I., Gemmi M. and Gregorkiewitz M. (2020) The structure of kaliophilite $KAlSiO_4$, a long-lasting crystallographic problem. *IUCrJ*, **7**, 1070–1083.

Nickel E.H. (1995) The definition of a mineral. *The Canadian Mineralogist*, **33**, 689–690.

Nickel E.H. and Grice J.D. (1998) The IMA commission on new minerals and mineral names: procedures and guidelines on mineral nomenclature. *The Canadian Mineralogist*, **36**, 913–926.

Olds T.A., Plášil J., Kampf A.R., Simonetti A., Sadergaski L.R., Chen Y.-S. and Burns P.C. (2017a) Ewingite: Earth's most complex mineral. *Geology*, **45**, 1007–1010.

Olds T.A., Plášil J., Kampf A.R., Škoda R., Burns P.C., Čejka J., Bourgoin V. and Boulliard J.-C. (2017b) Gauthierite, $KPb[(UO_2)_7O_5(OH)_7] \cdot 8H_2O$, a new uranyl-oxide hydroxy-hydrate mineral from Shinkolobwe with a novel uranyl-anion sheet-topology. *European Journal of Mineralogy*, **29**, 129–141.

Olds T.A., Plášil J., Kampf A.R., Dal Bo F. and Burns P.C. (2018) Paddlewheelite, a new uranyl carbonate from the Jáchymov district, Bohemia, Czech Republic. *Minerals*, **8**, 511.

Ottonello G. (2000) *Principles of Geochemistry*. Columbia University Press, New York, 894 pp.

Ovchinnikov A., Bobnar M., Prots Y., Borrmann H., Sichelschmidt J., Grin Y. and Höhn P. (2018) $Ca_{12}[Mn_{19}N_{23}]$ and $Ca_{133}[Mn_{216}N_{260}]$: structural

complexity by 2D intergrowth. *Angewandte Chemie, International Edition*, **57**, 11579–11583.

Pankova Y.A., Gorelova L.A., Krivovichev S.V. and Pekov I.V. (2018a) The crystal structure of ginorite, $Ca_2[B_{14}O_{20}(OH)_6] \cdot 5H_2O$, and the analysis of dimensional reduction and structural complexity in the $CaO–B_2O_3-H_2O$ system *European Journal of Mineralogy*, **30**, 277–287.

Pankova Y.A., Krivovichev S.V., Pekov I.V., Grew E.S. and Yapaskurt V.O. (2018b) Kurchatovite and clinokurchatovite, ideally $CaMgB_2O_5$: an example of modular polymorphism. *Minerals*, **8**, 332.

Pasero M. (2021) *The New IMA List of Minerals*. International Mineralogical Association. Commission on new minerals, nomenclature and classification (IMA-CNMNC). http://cnmnc.main.jp/.

Petrov T.G. (1970) On the measure of complexity of geochemical systems from the viewpoint of information theory. *Doklady Akademii Nauk SSSR*, **191**, 924–926 [in Russian].

Plášil J. (2018a) Mineralogy, crystallography and structural complexity of natural uranyl silicates. *Minerals*, **8**, 551.

Plášil J. (2018b) Structural complexity of uranophane and uranophane-β: implications for their formation and occurrence. *European Journal of Mineralogy*, **30**, 253–257.

Plášil J. (2018c) Uranyl-oxide hydroxy-hydrate minerals: their structural complexity and evolution trends. *European Journal of Mineralogy*, **30**, 237–251.

Plášil J., Petříček V. and Majzlan J. (2017) A commensurately modulated structure of parabutlerite, $Fe^{III}SO_4(OH) \cdot 2H_2O$. *Acta Crystallographica*, **B73**, 856–862.

Povarennykh A.S. (1966) On the regularities in the distribution of minerals according to syngonies, symmetry classes and space groups. *Mineralogicheskiy Sbornik Lvovskogo Universiteta*, **20**, 341–351 [in Russian].

Prabhu A., Morrison S.M., Eleish A., Zhong H., Huang F., Golden J.J., Perry S.N., Hummer D.R., Ralph J., Runyon S.E., Fontaine K., Krivovichev S., Downs R.T., Hazen R.M. and Fox P. (2021) Global earth mineral inventory: A data legacy. *Geoscience Data Journal*, **8**, 74–89.

Putnis A. (1992) *An Introduction to Mineral Sciences*. Cambridge University Press, Cambridge, UK, 457 pp.

Ranade M.R., Navrotsky A., Zhang H.Z., Banfield J.F., Elder S.H., Zaban A., Borse P.H., Kulkarni S.K., Doran G.S. and Whitfield H.J. (2002) Energetics of nanocrystalline TiO_2. *Proceedings of the National Academy of Sciences USA*, **99**, 6476–6481.

Rustad J.R. (2010) Elementary reactions in polynuclear ions and aqueous–mineral interfaces: a new geology. *Advances in Inorganic Chemistry*, **62**, 391–436.

Sabirov D.S. and Shepelevich I.S. (2021) Information entropy in chemistry: an overview. *Entropy*, **23**, 1240.

Sachs L. (1997) Angewandte Statistik. *Anwendung statistischer Methoden*. Springer, Heidelberg: 881 pp.

Shafranovskii I.I. (1973) Statistical regularities and generalized law in the distribution of minerals according to their symmetry. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **102**, 87–88 [in Russian].

Shafranovskii I.I. (1983a) The Fedorov-Groth statistical law and some related generalizing analogies. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **112**, 177–184 [in Russian].

Shafranovskii I.I. (1983b) Refinement of the concept on the basic law of symmetry statistics of minerals. *Mineralogicheskiy Zhurnal*, **5**, 66–69 [in Russian].

Shafranovskii I.I. (1985) Table of 32 symmetry classes and basic laws of mineralogical crystallography. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **114**, 521–527 [in Russian].

Shafranovskii I.I. (1987) The agreement between symmetry and statistics of minerals. *Mineralogicheskiy Zhurnal*, **9**, 3–7 [in Russian].

Shafranovskii I.I. (1988) Law or not law? *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **117**, 243–245 [in Russian].

Shafranovskii I.I. and Feklichev V.G. (1982) Symmetry statistics of minerals. *Mineralogicheskiy Zhurnal*, **4**, 31–36.

Siidra O.I., Zenko D.S. and Krivovichev S.V. (2014) Structural complexity of lead silicates: Crystal structure of $Pb_{21}[Si_7O_{22}]_2[Si_4O_{13}]$ and its comparison to hyttsjöite. *American Mineralogist*, **99**, 817–823.

Sokolova E., Hawthorne F.C., Pautov L.A., Agakhanov A.A. and Karpenko V.Yu. (2011) The crystal structure and crystal chemistry of mendeleevite-(Ce), $(Cs,\square)_6(\square,Cs)_6(\square,K)_6(REE,Ca,\ \square)_{30}(Si_{70}O_{175})(H_2O,$

OH,F,)$_{35}$, a potential microporous material. *Mineralogical Magazine*, **75**, 2583–2596.

Tait K.T., Ercit S.T., Abdu Y., Černý P. and Hawthorn F.C. (2011) The crystal structure and crystal chemistry of manitobaite, ideally Na$_{16}$Mn$_{25}^{2+}$Al$_{8-}$(PO^4)$_{30}$, from Cross Lake, Manitoba, Canada. *The Canadian Mineralogist*, **49**, 1221–1242.

Tambornino F., Sappl J. and Hoch C. (2015) The Gd$_{14}$Ag$_{51}$ structure type and its relation to some complex amalgam structures. *Journal of Alloys and Compounds*, **618**, 326–335.

Urusov V.S. (2002) The principle of the minimal dissymmetrization and its violation of rare new mineral species. *Doklady Rossiyskoy Akademii Nauk*, **386**, 379–383 [in Russian].

Urusov V.S. (2007) Symmetry statistics of mineral species and the evolutionary dissymmetrization of mineral substance. *Geology of Ore Deposits*, **49**, 497–504.

Urusov V.S. and Nadezhina T.N. (2006) Empty and sparse spatial groups in structural mineralogy. *Vestnik Moskovskogo Gosudarstvennogo Universiteta, Seriya Geologiya*, **6**, 52–59 [in Russian].

Urusov V.S. and Nadezhina T.N. (2009) Frequency distribution and selection of space groups in inorganic crystal chemistry. *Journal of Structural Chemistry*, **50**, 22–37.

Varn D.P. and Crutchfield J.P. (2016) What did Erwin mean? the physics of information from the materials genomics of aperiodic crystals and water to molecular information catalysts and life. *Philosophical Transactions of the Royal Society*, **A374**, 20150067.

Weisbach A. (1875) *Synopsis mineralogica: Systematische Uebersicht des Mineralreiches*. J. G. Engelhardt'sche Buchhdlg, Freiberg, Germany, 78 S [in German].

Yushkin N.P. (1977) *Theory and Methods of Mineralogy*. Nauka, Leningrad, Russia.

Yushkin N.P. (1982) Evolutionary ideas in modern mineralogy. *Zapiski Vsesoyuznogo Mineralogicheskogo Obshchestva*, **116**, 432–442 [in Russian].

Zaitsev A.N., Zhitova E.S., Spratt J., Zolotarev A.A. and Krivovichev S.V. (2017) Isolueshite, NaNbO$_3$, from the Kovdor carbonatite, Kola peninsula, Russia: composition, crystal structure and possible formation scenarios. *Neues Jahrbuch fuer Mineralogie, Abhandlungen*, **194**, 165–173.

Zhabin A.G. (1981) Is there evolution of mineral species on Earth? *Doklady Earth Sciences Section*, **247**, 142–144.

Zolotarev A.A., Krivovichev S.V., Panikorovskii T.L., Gurzhiy V.V., Bocharov V.N. and Rassomakhin M.A. (2019) Dmisteinbergite, CaAl$_2$Si$_2$O$_8$, a metastable polymorph of anorthite: Crystal-structure and Raman spectroscopic study of the holotype specimen. *Minerals*, **9**, 570.

Zolotarev A.A., Krivovichev S.V., Cámara F., Bindi L., Zhitova E.S., Hawthorne F. and Sokolova E. (2020) Extraordinary structural complexity of ilmajokite: A multilevel hierarchical framework structure of natural origin. *IUCrJ*, **7**, 121–128.