

Topological complexity of crystal structures: quantitative approach

Sergey Krivovichev

Received 14 February 2012

Accepted 20 March 2012

Crystallography, St Petersburg State University, University Emb. 7/9, St Petersburg, 199034, Russian Federation. Correspondence e-mail: skrivovi@mail.ru

The topological complexity of a crystal structure can be quantitatively evaluated using complexity measures of its quotient graph, which is defined as a projection of a periodic network of atoms and bonds onto a finite graph. The Shannon information-based measures of complexity such as topological information content, I_G , and information content of the vertex-degree distribution of a quotient graph, I_{vd} , are shown to be efficient for comparison of the topological complexity of polymorphs and chemically related structures. The I_G measure is sensitive to the symmetry of the structure, whereas the I_{vd} measure better describes the complexity of the bonding network.

© 2012 International Union of Crystallography
Printed in Singapore – all rights reserved

1. Introduction

Which crystal structure is complex? Despite the obvious and widely accepted use of the adjective ‘complex’ in the characterization of crystal structures, there is still no strict definition of crystal-structure complexity that can be applied on an everyday basis in scientific research. Various approaches to structure complexity have been outlined (Baur *et al.*, 1983; Burdett *et al.*, 1994; Mackay, 2001; Estevez-Rams & González-Férez, 2009; Steurer, 2011) but none of them provided simple and unambiguous quantitative measures. At the same time, advances in complex network and chemical graph theories (Barabási, 2002; Dorogovtsev & Mendes, 2003) allowed the formulation of a whole series of complexity measures for chemical, biological, technological and social networks (Dehmer, 2011). The most interesting parameters are based upon application of the Shannon information concept. Since Shannon information is closely related to the Boltzmann statistical entropy (Haken, 2000), the correct estimation of structure complexity may provide important insights into the entropy-based part of the crystal free energy and thus can be used for investigation of the evolution of information during processes involving the formation and transformation of crystalline phases. On the other hand, it is useful to have a set of quantitative parameters describing the structural complexity of crystalline solids. It would provide important tools for the comparison and classification of related structures according to their complexity. One may also identify chemical and physical parameters that may have an important influence upon structural complexity in the course of various physical and chemical processes, both of natural and artificial origin.

In this paper, we propose to use information-based topological complexity measures developed for graphs and networks for characterization of the topological complexity of crystal structures. In order to be of use for crystallographers

and crystal chemists, we tried to avoid complex complexity measures and to focus upon those that are simple, easy to calculate and agreeable with intuitive feelings of structural complexity.

2. Crystal structure as a complex network

From the abstract point of view, crystal structure may be viewed as an infinite network (graph) consisting of atoms as vertices (nodes) and chemical bonds as edges (links). The assumption is usually in order that one bond links precisely two atoms. The graph describing the bond topology of a crystal has some important properties: (i) it is *simple*, *i.e.* it contains no loops or multiple edges; (ii) it is 3-periodic, *i.e.* there are three linearly independent translation vectors that map the graph onto itself (Klee, 2004; Eon, 2005; Blatov & Proserpio, 2011). 3-Periodic simple nets have been used extensively for the description of complex crystal structures such as those observed in zeolites (Smith, 1988, 2000; Baerlocher *et al.*, 2001) and coordination polymers (Yaghi *et al.*, 2003; Blatov *et al.*, 2004; Ockwig *et al.*, 2005). A number of comprehensive databases of 3-periodic simple nets have been compiled such as the Reticular Chemistry Structure Resource (RCSR; O’Keeffe *et al.*, 2008) and the Database of Hypothetical Zeolite Structures (Foster & Treacy, 2010). A 3-periodic simple net can be described using finite graphs containing translationally non-equivalent vertices and edges of the net only. From a mathematical point of view, transition from a net to a finite graph corresponds to the operation of projection, when all translationally equivalent vertices of the net project onto one vertex of the finite graph. This graph is called the *quotient graph* of a 3-periodic net (Chung *et al.*, 1984; Klee, 2004). Quotient graphs are not necessarily simple and may contain loops and multiple edges. Figs. 1(a) and 1(b) show two periodic nets with quotient graphs depicted in Figs. 1(c) and 1(d), respectively. As can be seen, the 2–4 edge of the quotient

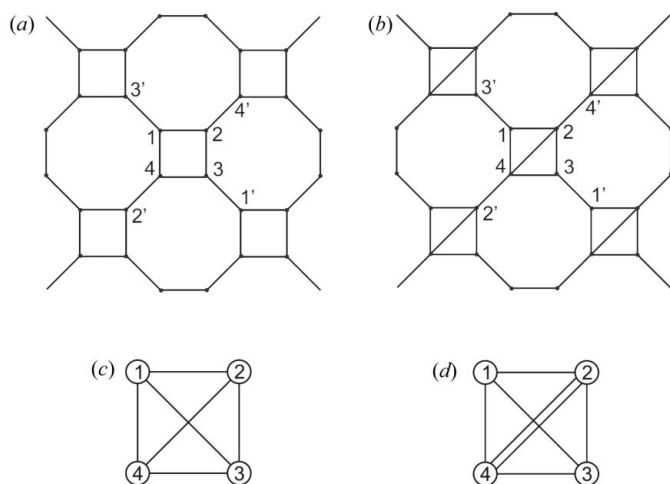


Figure 1
Two planar 2-periodic nets (a), (b) and their quotient graphs (c), (d), respectively.

graph in Fig. 1(d) is a double edge, which indicates that the vertices 2 and 4 are twice connected to each other in the respective parent net.

In general, quotient graphs are directed labelled graphs, *i.e.* their edges have directions and labels. A direction of an edge linking vertices *A* and *B* indicates that the respective chemical bond is directed from the respective atom *A* to the atom *B*, whereas a label identifies a vector in three-dimensional space, usually expressed in terms of crystallographic indices [*pqr*]. A quotient graph with directions and labels unambiguously determines the *topology* of a structure network consisting of atoms and chemical bonds. In some cases, simplifications may be made in order to describe the structure topology in a more straightforward and transparent way (*e.g.*, for tetrahedral frameworks, coordination polymers *etc.*).

Since a quotient graph completely determines crystal structure topology, its complexity measures can be used as complexity measures of the whole structure. As the first step, we suggest considering the topological properties of quotient graphs without directions and labels.

3. Complexity measures

Many measures of the complexity of graphs have been proposed recently in the chemical, physical and mathematical literature (Bonchev, 2003, 2005; Kim & Wilhelm, 2008; Dehmer, 2011). The simplest measure is the number of vertices *v* of a graph *G*(*v*; *e*) (*e* is the number of edges). The larger the graph is, the more complex it is. A useful complexity measure based on the topological properties of graphs is the average edge complexity, *E_a*, defined as (Bonchev, 2005)

$$E_a = 2e/v = \langle a_i \rangle \quad (1)$$

where $\langle a_i \rangle$ is the average vertex degree (vertex degree is the number of edges incident upon a vertex).

Another class of complexity measures is based upon the information content (entropy) of graphs. Rashevsky (1955),

Trucco (1956) and Mowshowitz (1968*a,b,c,d*) were the first to apply Shannon's information measure to derive an entropy of a graph characterizing its topology [see Dehmer & Mowshowitz (2011) for the history of graph entropy measures]. In the framework of Shannon's information theory (Shannon & Weaver, 1949), the entropy of information *H* encoded in a message of *N* symbols, which are subdivided into *k* classes consisting of *N*₁, *N*₂, ..., *N*_{*k*} symbols (the symbols inside each group are thought to be equivalent), is calculated according to the following formula:

$$H = - \sum_{i=1}^k p_i \log_2 p_i \quad (\text{bits symbol}^{-1}), \quad (2)$$

where the ratio $p_i = N_i/N$ is the probability of occurrence of the symbols of the *i*th class.

Application of Shannon's theory to graphs requires the formulation of an equivalence criterion that allows one to subdivide the vertices of a graph into classes. One of the possible criteria is to classify vertices according to their symmetry relationships (graph automorphisms). Two vertices of a graph *G* are equivalent if their permutation does not result in a graph topologically different from (= non-isomorphic to) *G*. For instance, simultaneous permutations 1 ↔ 3 and 2 ↔ 4 of vertices of the graph shown in Fig. 1(d) do not change its topology. Its vertices can be classified into two equivalence classes: *C*₁ = {1, 3} and *C*₂ = {2, 4}. The probabilities for a randomly chosen vertex *i* in the graph to belong to the classes *C*₁ and *C*₂ are $p_1 = v_1/v = 2/4 = 0.5$ and $p_2 = v_2/v = 2/4 = 0.5$, respectively, where *v*₁ and *v*₂ are the numbers of vertices in *C*₁ and *C*₂, respectively, and *v* is the total number of vertices ($v = v_1 + v_2$).

According to Rashevsky (1955), Trucco (1956) and Mowshowitz (1968*a,b,c,d*), the topological information content *I_G* of a graph *G* is defined as

$$I_G = - \sum_{i=1}^k p_i \log_2 p_i \quad (\text{bits vertex}^{-1}), \quad (3)$$

where p_i is the probability of occurrence of the vertex of the *i*th class.

The topological information content of the graph shown in Fig. 1(d) is

$$I_G = -(0.5 \log_2 1/2 + 0.5 \log_2 1/2) = 1.000 \quad (\text{bits vertex}^{-1}). \quad (4)$$

Let us consider another graph with four vertices shown in Fig. 1(c). All its vertices are equivalent and it is obvious that its structural information content is zero.

The maximal topological information content *I_{G,max}* for a graph *G* with *v* vertices is achieved when all vertices are non-equivalent according to the automorphism group of *G*. It can be calculated according to the following formula:

$$I_{G,\max} = \log_2 v \quad (\text{bits vertex}^{-1}). \quad (5)$$

In order for *I_G* to be independent from *v*, we may define the normalized topological information content, *I_{G,n}*, as

$$I_{G,n} = I_G/I_{G,\max} = -\left(\sum_{i=1}^k p_i \log_2 p_i\right) / \log_2 v. \quad (6)$$

It should be noted that proper calculation of the topological information content requires proper determination of the automorphism group of the quotient graph. In order to simplify the calculation procedure, vertices of the quotient graph can be classified into equivalence groups according to the Wyckoff sites which corresponding atoms occupy in the crystal structure. As an example let us consider the structure of $\text{K}_2\text{Cr}_2\text{O}_7\text{-IIb}$ (Krivovichev *et al.*, 2000) (Fig. 2*a*) which has the space group $C2/c$. Its reduced unit cell, which defines the content of the quotient graph, contains four symmetrically equivalent K, four Cr, four O1, four O2, two O3 and four O4 sites. As a consequence, the quotient graph (Fig. 2*c*) contains 22 vertices that can be grouped into six equivalence classes consisting of 4 (K), 4 (Cr), 4 (O1), 4 (O2), 2 (O3) and 4 (O4) vertices. The corresponding finite probability scheme can be written as follows: 4/22, 4/22, 4/22, 4/22, 2/22 and 4/22, respectively. The topological information content is calculated as

$$I_G = -[5 \times (4/22) \log_2 4/22 + (2/22) \log_2 2/22] = 2.550. \quad (7)$$

According to this approach, it is obvious that the I_G information measure is very sensitive to the crystallographic symmetry of the structure.

Another criterion of equivalence and the corresponding measure of the complexity of a graph were proposed by Bonchev (2005). They are based upon the subdivision of vertices into classes according to their degrees. The corresponding measure of complexity is the information content of the vertex-degree distribution of a graph, I_{vd} :

$$I_{\text{vd}} = \sum_{i=1}^v a_i \log_2 a_i, \quad (8)$$

where a_i is a degree of the i th vertex and summation proceeds along all v vertices. According to this definition, the graphs

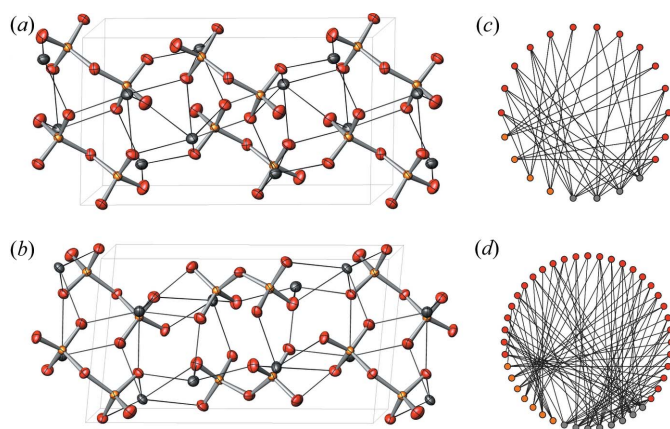


Figure 2
Crystal structures of $\text{K}_2\text{Cr}_2\text{O}_7\text{-IIb}$ (a) and $\text{K}_2\text{Cr}_2\text{O}_7\text{-II}$ (b) and their quotient graphs (c), (d), respectively. Legend: K atoms = dark grey; Cr atoms = dark yellow; O atoms = red. Quotient graphs are drawn using *Pajek* software (Batagelj & Mrvar, 2003).

shown in Figs. 1(c) and 1(d) have their I_{vd} measures equal to 19.020 and 25.510, respectively. Thus, in terms of both structural and vertex-degree-distribution information content, the graph in Fig. 1(b) appears to be *more complex* than that in Fig. 1(a) (the more information the graph contains, the more complex it is). The I_{vd} value is a convenient complexity measure for the comparison of two graphs with the same number of vertices. Its normalized version in our case is hard to define, since the quotient graph may contain loops and multiple edges, which make the maximal value of I_{vd} indefinite.

Unfortunately, there is no universal measure of graph complexity that works well for any type of chemical structure. In the case of a crystal-structure network and its quotient graph, the most important complexity parameters may be listed as follows: (i) number of vertices, v , of the quotient graph = number of atoms in a reduced unit cell (or its part in the case of distorted structures; see below); (ii) the topological information content, I_G , or its normalized version, $I_{G,n}$, which reflects the space-group symmetry of the structure; (iii) parameters characterizing the number of edges and their distribution over the graph in terms of vertex degrees (I_{vd} , E_g , E_a , E_n). Complexity measures of groups (i) and (ii) specify the number, chemical nature and crystallographic (in some cases, topological) equivalence of atoms, whereas those of group (iii) characterize the complexity of bond networks. A particular choice of these or other measures (*e.g.*, partition-independent information measures: see Dehmer & Emmert-Streib, 2008; Dehmer, 2008; Dehmer *et al.*, 2009; Dehmer & Mowshowitz, 2011) depends upon the particular class of crystal structures under consideration.

4. Applications

The most obvious application of topological complexity measures is to compare the complexity of polymorphs of the same compound. Fig. 2 shows the structures of $\text{K}_2\text{Cr}_2\text{O}_7\text{-IIb}$ (Krivovichev *et al.*, 2000) (Fig. 2*a*) and $\text{K}_2\text{Cr}_2\text{O}_7\text{-II}$ (Kuz'Min *et al.*, 1967; Brandon & Brown, 1968; Brunton, 1973; Weakley *et al.*, 2004) (Fig. 2*b*), and their quotient graphs [Figs. 2*c*, 2*d*, respectively; only K—O bonds shorter than 3.25 Å have been taken into account in order for the bond valences to be higher than 0.05 valence units (Brown, 1981)]. Values of the complexity parameters for these structures are provided in Table 1. It can be seen that, according to all the information-based measures (I_G , $I_{G,n}$, I_{vd}), the structure of $\text{K}_2\text{Cr}_2\text{O}_7\text{-IIb}$ (a metastable modification) is much simpler than that of $\text{K}_2\text{Cr}_2\text{O}_7\text{-II}$. This seems to be in agreement with the 'simplicity principle' proposed by Goldsmith (1953) which states that, in the case of rapid crystallization, the metastable phase that forms first is structurally simple compared to the stable phase. Indeed, the metastable polymorph $\text{K}_2\text{Cr}_2\text{O}_7\text{-IIb}$ forms as a result of spontaneous crystallization from a supersaturated aqueous solution of potassium bichromate. As proposed by Goldsmith (1953), preferential crystallization of the metastable phase with higher simplicity (= lower complexity) is due to the smaller size of its critical nuclei in

Table 1
Values of topological complexity measures for selected crystal structures.

Compound/name	v	e	I_G	$I_{G,n}$	I_{vd}	E_a	References
K ₂ Cr ₂ O ₇ polymorphs							
K ₂ Cr ₂ O ₇ -II	44	96	4.459	0.817	429.421	4.4	Weakley <i>et al.</i> (2004)
K ₂ Cr ₂ O ₇ -IIb	22	48	2.550	0.572	215.020	4.4	Krivovichev <i>et al.</i> (2000)
CaCO ₃ polymorphs							
Vaterite	10	18	1.371	0.413	69.059	3.6	Demichelis <i>et al.</i> (2012)
Calcite I	10	18	1.371	0.413	69.059	3.6	Chessin <i>et al.</i> (1965)
Calcite II	10	18	1.371	0.413	69.059	3.6	Merrill & Bassett (1975)
Calcite III	15	27	3.107	0.795	103.588	3.6	Smyth & Ahrens (1997)
Aragonite	20	48	1.922	0.445	229.137	4.4	Dal Negro & Ungaretti (1971)
Post-aragonite	10	30	1.922	0.579	165.207	6.0	Ono <i>et al.</i> (2005)

comparison with the stable phase. It is obvious that the proposed complexity parameters, at least partially, account for this peculiarity: the complex phase has a larger reduced unit cell and, as a result, a larger quotient graph, compared with the simple phase. In general, the evaluation of the topological complexity of phases in the course of a crystallization sequence governed by the Ostwald's step rule (Ostwald, 1897) is an interesting problem [one, however, must be aware that the structure of at least the first phase in the sequence might be seriously affected by the structure of complexes present in the crystallization liquid (Threlfall, 2003)].

Another interesting problem is investigating changes in the topological complexity of crystal structures along phase transition pathways. Fig. 3 depicts quotient graph diagrams for different CaCO₃ polymorphs: vaterite (Kamhi, 1963; Meyer, 1969; Medeiros *et al.*, 2007; Demichelis *et al.*, 2012), calcite I

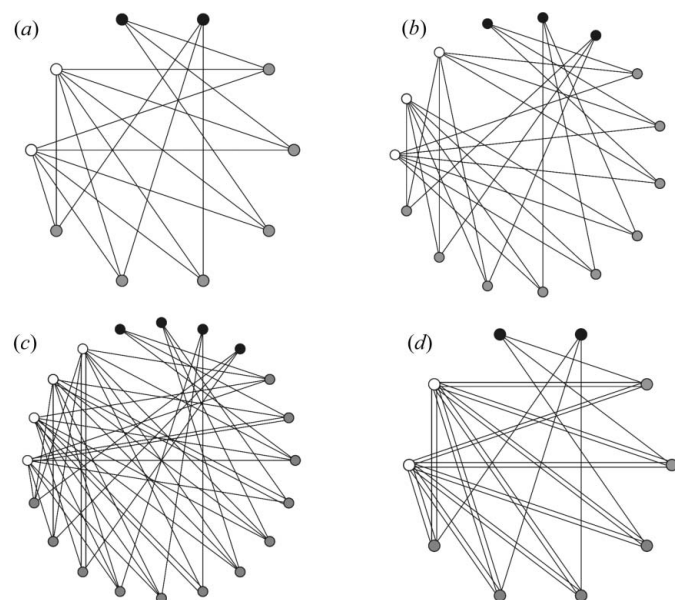


Figure 3
Quotient graphs of the crystal structures of CaCO₃ polymorphs: vaterite, calcite I and calcite II (a), calcite III (b), aragonite (c) and post-aragonite (d). Legend: white circles = Ca; black circles = C; grey circles = O. Quotient graphs are drawn using Pajek software (Batagelj & Mrvar, 2003).

(Chessin *et al.*, 1965), calcite II (Merrill & Bassett, 1975), calcite III (Smyth & Ahrens, 1997), aragonite (Dal Negro & Ungaretti, 1971) and post-aragonite (Ono *et al.*, 2005). Table 1 provides values of the complexity measures. The structure of vaterite has the same quotient graph as that of calcite I. However, it is disordered from the viewpoint of orientation of carbonate triangular ions in adjacent unit cells and therefore has additional degrees of freedom (Demichelis *et al.*, 2012). As a disordered phase, vaterite has lower structural complexity than calcite I,

which may explain its metastable appearance during crystallization of CaCO₃ from aqueous solutions (Gebauer *et al.*, 2008). The high-pressure treatment of CaCO₃ results in the following sequence of phases: calcite I → calcite II → calcite III → aragonite → post-aragonite. The structure of calcite II (Merrill & Bassett, 1975) is a distorted version of the structure of calcite I. Despite the fact that its reduced unit cell contains twice as many atoms as that of calcite I, their topologies are identical and the quotient graph for calcite II might be reduced to that shown in Fig. 3(a). Inspection of the data given in Table 1 shows that the increase of pressure along the line calcite I → calcite II → calcite III → aragonite corresponds to the increase of the I_{vd} complexity measure. However, the aragonite → post-aragonite transition is accompanied by a decrease in the I_{vd} parameter, whereas the E_a value reaches its maximum (6.0). The latter observation accounts for the increase of the coordination number of Ca²⁺ ions, which implies complexification of the bonding network with increasing pressure. Similar tendencies of high-pressure phase transitions (initial increase and subsequent decrease of complexity) have also been observed for SiO₂ and TiO₂ polymorphs, which will be reported elsewhere.

In general, it seems that the I_{vd} complexity measure provides the most adequate description of topological complexity of a crystal structure. However, it is not as sensitive to crystal symmetry as I_G and does not reflect an increasing number of bonds in the network as well as E_a . In addition, it depends upon the specific choice of interatomic interactions considered as bonds, which is not always obvious, especially in the case of large low-valent cations (*e.g.*, alkali metals). In many cases, it is the I_G information measure that provides the most adequate measure of structural complexity. This is the case for zeolite frameworks. For uninterrupted zeolite frameworks TO₂ ($T = \text{Si, Al, P, Zn etc.}$), the coordination numbers of the T and O atoms are 4 and 2, respectively. Therefore, quotient graphs of zeolite frameworks consist of $v = 3t$ vertices and $e = 4t$ edges where t is the number of T atoms in the reduced unit cell of the structure. The I_{vd} information measure is equal to

$$I_{vd} = 4t \log_2 4 + 4t \log_2 2 = 8t + 4t = 12t = 4v. \quad (9)$$

Thus, for zeolite frameworks, the I_{vd} information measure is a simple function of ν and does not discriminate between frameworks with the same ν . In such a situation, the use of topological information content measures appears to be more reliable for complexity characterization. For instance, the zeolite frameworks SOD and NPO (Baerlocher *et al.*, 2001) have the same number of vertices of quotient graphs ($\nu = 18$), the same E_a value (2.667) and the same I_{vd} value (72.000). Standard representation of the SOD framework is in the $Im\bar{3}m$ space group with the T and O atoms in the $12d$ and $24h$ sites, respectively. The reduced cell contains 6 T and 12 O sites ($t = 6$; $\nu = 18$). The quotient graph consists of two equivalence classes consisting of 6 and 12 vertices, respectively.

Therefore,

$$I_G = 6/18(\log_2 18 - \log_2 6) + 12/18(\log_2 18 - \log_2 12) = 0.918; \\ I_{G,n} = 0.918 / \log 218 = 0.220. \quad (10)$$

Standard representation of the NPO framework is in the space group $P\bar{6}2c$. The T atoms are in the $6h$ site, whereas O atoms occupy the $6g$ and $6h$ sites. Vertices of the quotient graph are classified into three equivalence classes consisting of 6 vertices each.

Therefore,

$$I_G = 3[6/18(\log_2 18 - \log_2 6)] = 1.585; \\ I_{G,n} = 1.585 / \log 218 = 0.380. \quad (11)$$

Comparison of the I_G and $I_{G,n}$ measures provides the conclusion that the NPO framework is more complex than the SOD framework, which is in agreement with general intuition.

5. Conclusions

In this work, information-based measures of quotient graphs (topological information content I_G , normalized topological information content $I_{G,n}$ and the information content of the vertex-degree distribution, I_{vd}) are proposed as parameters characterizing the complexity of a crystal structure. The quotient graphs represent structure content within a reduced crystallographic cell of the structure. Thus, the topological information content can easily be calculated using Wyckoff sets of crystallographic sites, whereas the information content of the vertex-degree distribution can be calculated on the basis of coordination numbers assigned to specific atoms in the structure. Both types of information parameters (*i.e.* those based on symmetry and those based upon vertex-degree distribution) should be used in combination depending upon the problem under consideration.

The proposed parameters can be used for estimating the complexity of polymorphs of the same compound, for comparison of the complexities of different related structures, for investigation of accumulation and dissipation of information in various processes involving crystalline phases *etc.* Another point of potential interest is the use of crystal-structure topological complexity measures in the study of quantitative structure–property relationships in crystalline materials. However, it is unclear whether the concept of

complexity can be of any value in this field of materials research.

I am grateful to two anonymous referees for constructive comments on the manuscript. This work was supported through the Russian Federal Grant-in-Aid Program (state contract No. 16.740.11.0490) and the Russian Academy of Sciences Basic Research Program.

References

- Baerlocher, Ch., Meier, W. N. & Olson, D. H. (2001). *Atlas of Zeolite Framework Types*. Amsterdam: Elsevier.
- Barabási, A.-L. (2002). *Linked. The New Science of Networks*. Cambridge, MA: Perseus Publishing.
- Batagelj, V. & Mrvar, A. (2003). *Graph Drawing Software*, edited by M. Jünger & P. Mutzel, pp. 77–103. Springer: Berlin.
- Baur, W. H., Tillmanns, E. & Hofmeister, W. (1983). *Acta Cryst.* **B39**, 669–674.
- Blatov, V. A., Carlucci, L., Ciani, G. & Proserpio, D. M. (2004). *CrystEngComm*, **6**, 378–395.
- Blatov, V. A. & Proserpio, D. M. (2011). *Modern Methods of Crystal Structure Prediction*, edited by A. R. Oganov, pp. 1–28. Weinheim: Wiley–VCH.
- Bonchev, D. (2003). *SAR QSAR Environ. Res.* **14**, 199–214.
- Bonchev, D. (2005). *Complexity in Chemistry, Biology and Ecology*, edited by D. Bonchev & D. H. Rouvray, pp. 191–235. New York: Springer.
- Brandon, J. K. & Brown, I. D. (1968). *Can. J. Chem.* **46**, 933–941.
- Brown, I. D. (1981). *Structure and Bonding in Crystals*, edited by M. O'Keefe & A. Navrotsky, Vol. 2, pp. 49–72. New York: Academic Press.
- Brunton, G. (1973). *Mater. Res. Bull.* **8**, 271–274.
- Burdett, J. K., Mariani, C. & Mitchell, J. F. (1994). *Inorg. Chem.* **33**, 1848–1856.
- Chessin, H., Hamilton, W. C. & Post, B. (1965). *Acta Cryst.* **18**, 689–693.
- Chung, S. J., Hahn, Th. & Klee, W. E. (1984). *Acta Cryst.* **A40**, 42–50.
- Dal Negro, A. & Ungaretti, L. (1971). *Am. Mineral.* **56**, 768–772.
- Dehmer, M. (2008). *Appl. Math. Comput.* **201**, 82–94.
- Dehmer, M. (2011). *Symmetry*, **3**, 767–779.
- Dehmer, M. & Emmert-Streib, F. (2008). *Comput. Biol. Chem.* **32**, 131–138.
- Dehmer, M. & Mowshowitz, A. (2011). *Inf. Sci.* **181**, 57–78.
- Dehmer, M., Varmuza, K., Borgert, S. & Emmert-Streib, F. (2009). *J. Chem. Inf. Model.* **49**, 1655–1663.
- Demichelis, R., Raiteri, P., Gale, J. D. & Dovesi, R. (2012). *CrystEngComm*, **14**, 44–47.
- Dorogovtsev, S. N. & Mendes, J. F. F. (2003). *Evolution of Networks: From Biological Networks to the Internet and WWW*. Oxford University Press.
- Eon, J.-G. (2005). *Acta Cryst.* **A61**, 501–511.
- Estevez-Rams, E. & González-Férez, R. (2009). *Z. Kristallogr.* **224**, 179–184.
- Foster, M. D. & Treacy, M. M. J. (2010). *A Database of Hypothetical Zeolite Structures*, <http://www.hypotheticalzeolites.net>.
- Gebauer, D., Völkel, A. & Cölfen, H. (2008). *Science*, **322**, 1819–1822.
- Goldsmith, J. R. (1953). *J. Geol.* **61**, 439–451.
- Haken, H. (2000). *Information and Self-Organization*. Berlin, Heidelberg: Springer-Verlag.
- Kamhi, S. R. (1963). *Acta Cryst.* **16**, 770–772.
- Kim, J. & Wilhelm, T. (2008). *Physica A*, **387**, 2637–2652.
- Klee, W. E. (2004). *Cryst. Res. Technol.* **39**, 959–968.
- Krivovichev, S. V., Kir'yanova, E. V., Filatov, S. K. & Burns, P. C. (2000). *Acta Cryst.* **C56**, 629–630.

- Kuz'Min, E. A., Plyukhin, V. V. & Belov, N. V. (1967). *Dokl. Akad. Nauk SSSR*, **173**, 1068–1071.
- Mackay, A. L. (2001). *Crystallogr. Rep.* **46**, 524–526.
- Medeiros, S. K., Albuquerque, E. L., Maia, F. F. Jr, Caetano, E. W. S. & Freire, V. N. (2007). *Chem. Phys. Lett.* **435**, 59–64.
- Merrill, L. & Bassett, W. A. (1975). *Acta Cryst.* **B31**, 343–349.
- Meyer, H. J. (1969). *Z. Kristallogr.* **128**, 183–212.
- Mowshowitz, A. (1968a). *Bull. Math. Biophys.* **30**, 175–204.
- Mowshowitz, A. (1968b). *Bull. Math. Biophys.* **30**, 225–240.
- Mowshowitz, A. (1968c). *Bull. Math. Biophys.* **30**, 387–414.
- Mowshowitz, A. (1968d). *Bull. Math. Biophys.* **30**, 533–546.
- Ockwig, N. W., Delgado-Friedrichs, O., O'Keeffe, M. & Yaghi, O. M. (2005). *Acc. Chem. Res.* **38**, 176–182.
- O'Keeffe, M., Peskov, M. A., Ramsden, S. J. & Yaghi, O. M. (2008). *Acc. Chem. Res.* **41**, 1782–1789.
- Ono, S., Kikegawa, T., Ohishi, Y. & Tsuchiya, J. (2005). *Am. Mineral.* **90**, 667–671.
- Ostwald, W. (1897). *Z. Phys. Chem.* **22**, 289–330.
- Rashevsky, N. (1955). *Bull. Math. Biophys.* **17**, 229–235.
- Shannon, E. & Weaver, W. (1949). *The Mathematical Theory of Communications*. Urbana: University of Illinois Press.
- Smith, J. V. (1988). *Chem. Rev.* **88**, 149–182.
- Smith, J. V. (2000). *Tetrahedral Frameworks of Zeolites, Clathrates and Related Materials*. Landolt–Börnstein New Series IV/14, Subvol. A. Heidelberg: Springer.
- Smyth, J. R. & Ahrens, T. J. (1997). *Geophys. Res. Lett.* **24**, 1595–1598.
- Steurer, W. (2011). *Acta Cryst.* **A67**, C184.
- Threlfall, T. (2003). *Org. Proc. Res. Devel.* **7**, 1017–1027.
- Trucco, E. (1956). *Bull. Math. Biophys.* **18**, 129–135.
- Weakley, T. J. R., Ylvisaker, E. R., Yager, R. J., Stephens, J. E., Wiegel, R. D., Mengis, M., Bauer, M. R., Wu, P., Photinos, P. & Abrahams, S. C. (2004). *Acta Cryst.* **B60**, 705–715.
- Yaghi, O. M., O'Keeffe, M., Ockwig, N. W., Chae, H. K., Eddaoudi, M. & Kim, J. (2003). *Nature (London)*, **423**, 705–714.