Malaysian Journal of Chemistry, 2019, Vol. 21(1), 54-61

The Local Regularity of an Organic Crystal

Alexander Banaru[1*], Yuri L. Slovokhotov[2] and Dmitry Gridin[2]
[1]Moscow State University, Chemistry Faculty, 1/3 Lenin Hills, 119991,
Moscow, Russia
[2]Moscow State University, Faculty of Materials Science, 1/73 Lenin Hills,
119991, Moscow, Russia
*Corresponding author: (e-mail: banaru@phys.chem.msu.ru)

In this paper the calculation of intermolecular interaction energy based on Gavezzotti–Filippini semi-empirical method was performed for the series of α, ω-diols. The initial structural data were retrieved from the Cambridge Structural Database. For each structure, the critical coordination number and the molecular coordination number were calculated. The interrelationship between the coordination numbers and the local regularity of a crystal was also discussed.

Key words: Intermolecular contact; lattice energy; space group; coordination number

Received: October 2018; April 2019

Up to the moment, the most powerful retrieval tool for structural data of organic compounds is Cambridge Structural Database, CSD [1]. Statistics based on CSD has multiple applications: crystal structure prediction, prediction of polymorphs, crystal engineering, a quantitative estimation of intermolecular interactions, justification of phase transitions, an insight into prospective material properties etc. The database was established in 1965 and initially contained published results on structure determination by X-ray and neutron diffraction for compounds having at list one ‘organic’ carbon atom. After that, this ambiguous criterion was moderated, and nowadays CSD contains, for instance, carbonyl complexes $M(CO)_n$ of transition metals. Since 1990th, an unpublished data, if satisfied specific requirements to the reliability, is also acceptable by CSD. According to estimates [2], the capacity of the Cambridge Database may reach a million of crystal structures in a few years. Each item is attributed by a refcode of 6 letters and, optionally, two numbers (for multiple determinations of the same crystal structure and polymorphs). The same crystalline compound studied by different teams and deposited independently may get different refcodes due to a fault of entering checking.

Developing the theory of dense molecular packing, Kitaigorodskii [3] has introduced the concept of molecular coordination number (MCN) defined several neighboring molecules having at least one van der Waals contact with the considered one. In organic crystals without structure-forming specific interactions (like H-bonds), MCN most often accepts value 12 or 14 [4]. However, determination of contacting molecules is based on sums of van der Waals radii that may deviate with actual intermolecular distances within 5% even for equivalent pairs of elements in the same crystal [3]. It induces some ambiguity into molecular coordination shells since various values of radii can lead to multiple coordination numbers.

Long ago it was shown that the periodicity of a crystal is caused by the identical local environment of structural units [5]. For this purpose, a so-called Delone system, or (r, R)-system, was introduced: a system of points, such that the vicinity of radius r (r-vicinity) of any point does not contain other points, and R-vicinity of any point contains at least one other point of the system. The strict criteria of local arrangement resulting in the periodicity are still being developed. It has been proven that the 2-dimensional (r, R)-system has 2D periodicity if all its points have identical environment in the vicinity of radius $4R$ [6]. In the same work, it was supposed that the criterion of a 3D crystal is the identity of environments in the sphere of radius $6R$. Recently, Dolbilin has found out [7] that the limit of correctness for a 3D crystal is most likely equal to $10R$. However, if the system is centrosymmetric (so-called locally antipodal set) and the centres of inversion are located in each point of the system, this boundary is reduced to $2R$.

In molecular crystals, the energy of intermolecular interaction is vital for the characteristics of molecular coordination along with MCN, as the energy of different contacts may differ strongly. In 1972, Zorky and Zefirov suggested [8] that impacts of pairwise non-bonded intermolecular interactions into total energy of a molecular crystal would have a gap at the last close intermolecular contact. Many years later, this proposition was tested by Grineva for series of α,ω-diols [9]. The energetic coordination number (N_E) was calculated as several intermolecular contacts for which the difference in energy with the strongest contact did not exceed 3%. Taking into account symmetric equivalence of some contacts, the N_E value in that series of structures appeared to be 1, 2, 4 or 8.

Recently, Motherwell [10] also addressed an ambiguity of the purely geometrical definition of MCN suggesting a criterion for relative decrease in energy of intermolecular contacts. Using the random set of 1000 crystal structures from CSD, he searched for a relative drop of energy from i-th to $(i+1)$-*th* contact being not less than 40%. The calculated MCN values were 14 in about 45% of structures, 12 in about 13% of structures, and MCN = 10, 13 and 16 were shared by 5-10% structures. However, only in 35% of structures, the first drop ≥40% of energy turned out to be at i = MCN. This hints that not all contacts in MCN but just a fraction of them may define a molecular environment.

In [11] there was introduced the concept of the critical coordination number (CCN), the least number of symmetrically unique intermolecular contacts sufficient for formation of a given crystal. It was also shown that in homomolecular crystals in which molecules occupy one general orbit ($Z' = 1$) min(CCN) equals the cardinality of the minimal set of generators for the space group, the last one accepting values from 2 to 6 inclusive. However, generally CCN is not limited to these values, and in [12] we demonstrated examples of crystal structures with CCN = 1, i.e. generated by repetition of one unique contact. The most surprising examples of these structures belonging to the space group $Fdd2$, Z = 8, are dimethylsulphate (CSD refcode VEGNAX01) generated by equivalent interactions of Me…O=S (H…O 2.4 Å), dichlorofluoromethane (SOPYUR) with a framework of short interactions Cl…F 3.1 Å, and several others. However, the strongest set of intermolecular contacts not necessary corresponds to the set of generators of a space group, i.e., CCN > min(CCN). That is, in the series of structures of substituted prolines [13] one of the strongest intermolecular contacts typically could not be included in any minimal set of generators of the space group.

The objective of the article was to find out whether there is some parsimony principle for CCN to prefer a minimal possible value, and to investigate the interrelationship between MCN, CCN, and R.

METHOD

The series of crystal structures of α, ω-diols, analyzed in [9], is also very easy-to-use for CCN assessment: first, these structures are chemically relative, second, they are homomolecular with $Z' = 1$ (similar molecules occupy the same position of a space group), and third, their intermolecular interactions are reasonably diverse (there are H-bonds together with ordinary van del Waals interactions). From 33 structures of the specified series, we analyzed 28, having excluded identical structures with a greater diffraction R-factor since small differences in geometrical parameters were not essential for CCN. Crystal structures of the same space group in different settings were treated as belonging to the same group. For instance, $P2_1/c$, $P2_1/a$, and $P2_1/n$ were uniformly marked as $P2_1/c$.

The computation of interaction energy was performed in mercury software [14] with the use of Gavezzotti–Filippini 6–exp potential [15]. In this approach, the potential of intermolecular interaction U for a pair of molecules is the sum of all interatomic potentials between these molecules. Intramolecular interactions were neglected. For the chart energies U_1, U_2, … U_{200} of 200 strongest contacts were summarized; the residual contacts commonly share much less than 1% of the lattice energy U_Σ. Thus, $U_1 + U_2 + \ldots + U_{200} \approx U_\Sigma$. For the reason that N_E in these structures did not exceed 3 (not considering symmetrically equivalent contacts), we thoroughly analyzed only the first 6 strongest unique contacts.

Assessment of MCN was also carried out in mercury using *Molecular Shell* utility which allows to visualizing a molecular coordination shell in the confined interval of distances either around a molecule, or around its fragment. As a criterion of inclusion of each following molecule in the calculation of MCN we used the deviation of interatomic distances from the sum of Bondi’s [16] van der Waals radii (ΣR_{vdW}). The lower bound of an appropriate interval of interatomic distances was set to (ΣR_{vdW} – 1,0 Å) in order not to miss any of the shortened interactions, in particular between hydrogen atoms. The upper bound ($\Sigma R_{\mathrm{vdW}} + \delta$), δ = 0; 0.1; 0.2;... Å was set individually for each crystal structure. We found out that if the coordination sphere did not increase for two following values of δ in a row, then it would not also increase for the following values until the new coordination sphere at a significantly greater distance was not reached. This allowed considering triple repetition of coordination number at different δ pointing to the true MCN value.

Each intermolecular contact corresponds to some symmetry operation. Combining these operations, one can obtain a minimal generator set of the space group. This is a minimal set of interactions to form the crystal structure. Although there can be different minimal sets for any space group, we selected the 'best' one in the following manner. Assuming that such minimal set of interactions contained contacts with the highest energy, we selected such a minimal set of interactions that its 1st strong interaction was the strongest among all sets. If it did not lead to a definite set, we selected a set where 2nd strong contact is the strongest among those in other sets, etc.

RESULTS AND DISCUSSION

Each structure was characterized by its min(CCN) value. In particular, space group $P2_12_12_1$ had the minimal set of two generating elements, namely, two mutually perpendicular screw axes 2_1 closest to each other. The third 2_1 axis, perpendicular to the former two, as well as the coordinate translations, were generated. Since this group did not have special positions, the homomolecular crystal with $Z' = 1$ of a given space group could belong only to the structural class with min(CCN) = 2. The space groups $P2_1$, $P2_1/c$ (as well as $P2_1/a$ and $P2_1/n$) and $C2/c$ had three generating element. For instance, the group $P2_1/c$ in β-setting could be generated by translation X together with any pair of the following three elements: 2_1 axis, c plane, and the closest centre of inversion i. The translations Y and Z then were generated. If in a crystal of these groups a single general position was occupied, then its min(CCN) is 3. If in a crystal of the group $P2_1/c$ or $C2/c$ a special position i was occupied, then min(CCN) = 2, because in these cases the inversion centre was included in the generating set imminently. For the $C2/c$ group, there also existed a structural class in which the molecules were located on rotation axes, but this structural class was not represented among the analyzed structures.

To find out whether a critical contact could be detected from the drop of energy for the i-th strong contact, we analyzed various combinations of U_i and U_Σ. Contrary to our anticipations based on the experience of analyzing the structures of other series (in particular, substituted prolines [13]), there was no break in the diagram of U_i vs i for the critical contact (the least strong contact, which was necessary for the formation of a given structure). Also, critical contact energies in the current series of structures had similar values; thus their normalization by U_Σ led to an increase of standard deviation and complicated the analysis. More informative appears to be $\Delta_i = (U_i - U_{i+1})/U_i$ ratio (Figure 1b).

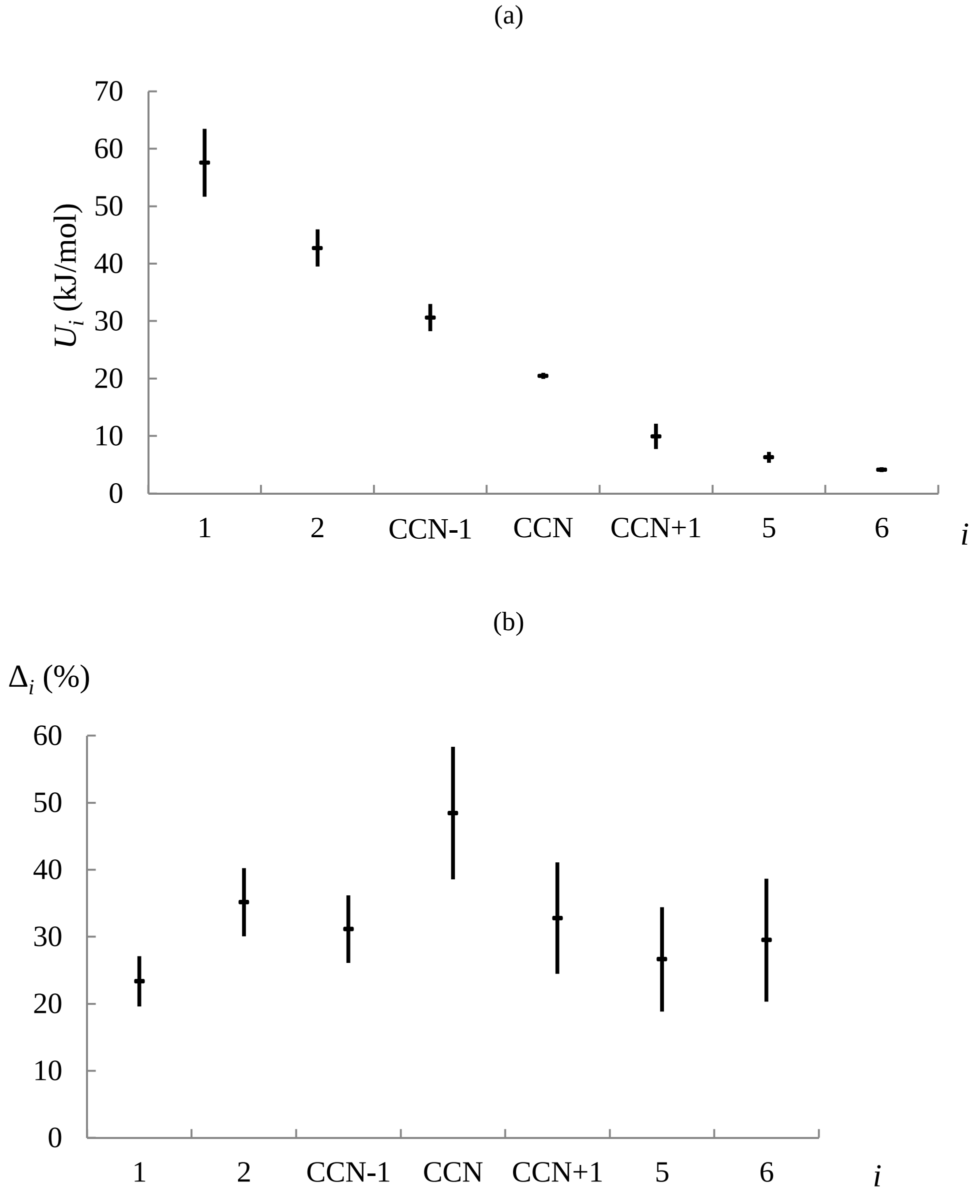


Figure 1. The average absolute (a) and relative (b) decrease of U_i and its confidence interval ($\alpha = 0.05$) for the series of crystal structures of α, ω-diols.

The plot in the Figure 1b shows that even with a very large standard deviation (there were no excluded structures), the confidence interval ($\alpha = 0.05$) of the Δ_i value for i = CCN lays almost entirely higher than the confidence intervals for all other contacts. Unlike N_E estimate [9], our Δ_i difference was much larger than 2-3% was 40-60% for the critical contact and 20-40% for the other contacts. Presently, it could not be stated that the obtained result follows precisely neither from the 'criticalness' of the critical contact, nor from the fact that such a contact mostly corresponds to H-bond To clarify this point, analysis of other series of interrelated crystal structures is required, primarily those without specific interactions (for example, olefins). Surprisingly, the threshold separating the energy of the critical contact from the others is exactly the same as that found by Motherwell for the coordination number [10]. As mentioned above, in more than half of cases such a drop in energy was detected not for MCN, but for some stronger contact, which now is likely to be the critical one.

The MCN values in the considered series were most often equal to 12 or 14 in good accordance to general predictions [3, 4] (Table 1). The only deviation was observed for QATTIO refcode with symmetry $P2_1/c$, $Z = 4$ and MCN = 13. Noteworthy, CCN increased with the molecular size, while a gap between the values of CCN and MCN decreased. The larger were molecules in the coordination sphere, the more possibilities they had to contact with each other varying local arrangement. It led to the emergence of new unique contacts, even if they were not required for initial local symmetry.

Table 1. Critical coordination numbers (CCN) and molecular coordination numbers (MCN) for the crystal structures of α, ω-diols with n(C) carbon atoms. Initial structure data were retrieved from CSD ver. 5.36 + 3 updates.

Space group, Z	min(CCN)	n(C)	CSD refcode	CCN	MCN
$P2_12_12_1$, $Z = 4$	2	2	NOZKES02	2	14
		5	QATTOU	2	12
		7	QATVAI	3	12
		9	WESTET	3	12
		11	HIYHAY	3	12
		13	TIQHAC	4	12
		15	MATLIC	4	12
		17	QIMCUK	4	12
		19	ICAHID	4	12
		21	XAYHOU	4	12
		23	QURRUQ	4	12
$C2/c$, $Z = 4$	2	20	RIWTAT	4	12
		22	RIWTEX	4	12
		24	RIWTIB	4	12

$P2_1/c$, $Z = 2$	2	8	QATVEM	3	14
		10	WESTIX01	3	14
		12	RUVWAG	3	14
		14	CAVDEI	3	14
		16	WERTUI	3	14
		18	GURBUQ	4	14
$P2_1$, $Z = 2$	3	16	WERTUI01	4	12
		18	GURBUQ01	4	12
		20	RIWTAT01	4	12
		22	RIWTEH01	4	12
		24	RIWTIB01	4	12
$P2_1/c$, $Z = 4$	3	3	QATTEK	3	12
		4	QATTIO	3	13
		6	FECCOF01	5	14

From the viewpoint of Dolbilin [7], a homomolecular crystal with $Z' = 1$ can be considered as (r, R)-system, where the points correspond to the centres of mass of molecules, and $R \approx d_{\min}$. Crystal structures in which molecules occupy inversion centres are a so-called locally antipodal set in which the same molecular environment in a sphere of radius $2R$ should lead to 2D periodicity.

Among the analyzed structures, the molecules occupy inversion centres in $C2/c$, $Z = 4$ and $P2_1/c$, $Z = 2$ crystals (Table 2). However, it turned out that the critical contact never fell into the $2R$-vicinity of a molecule. In addition, the number of neighbouring molecules in the $2R$-vicinity, as a rule, did not coincide with the MCN and could be either more or less. The number of molecules within a sphere of radius equal to the length of the critical contact could be so large (>100) that the current possibilities of mercury did not allow to find the exact value of this number. For comparison, in Table 2 we have listed the structures of non-centrosymmetric molecules in the centrosymmetric $P2_1/c$ group, where the critical contact fell into the $2R$-vicinity. However, limitations for the local regularity of such structures was much more rigorous [7].

The crystal structures considered in this article had much in common. All of them contained hydrophobic and hydrophilic alternating regions. Let consider crystal structure of tetracosane-1, 24-diol (RIWTIB). The molecule was centrosymmetric, it had a linear conformation and was packed into hydrophobic slabs with indexes (002). Linking the mass centres of the neighboring molecules, one obtained a plain fragment of the closest sphere packing. Both sides of this fragment form H-bonded tapes combined into plain 'sheets' with indexes (–104). The H-link (Figure 2) in this structure played a role of the critical contact with $U_{\mathrm{CCN}} = 19.9$ kJ/mol and $d_{\mathrm{CCN}} = 29.6$ Å, hydrophobic binding layers, which were very strong. The coverage parameter for the geometric centre of the molecules in the structure was $R = 12.1$ Å.

Table 2. Centro- and non-centrosymmetric molecules in crystal structures of racemic α, ω-diols. $N(<d)$ denotes the number of molecules in d-vicinity of a molecule**.**

Molecule	CSD refcode	d_{min}, Å	$i(d_{\mathrm{min}})$	$N(<2d_{\mathrm{min}})$	d_{CCN}, Å	$N(<d_{\mathrm{CCN}})$
centro-symmetric	QATVEM	4.80	1	18	11.66	32
	WESTIX01	4.94	1	10	13.85	40
	RUVWAG	4.96	1	10	16.08	60
	CAVDEI	4.98	1	10	18.27	72
	WERTUI	4.98	1	10	20.49	86
	GURRUQ	5.00	1	10	22.73	>100
	RIWTAT	5.23	2	14	25.09	>100
	RIWTEX	5.24	2	14	27.33	>100
	RIWTIB	5.24	2	14	29.57	>100
non-centro-symmetric	QATTEK	4.07	1	21	4.65	1
	QATTIO	5.01	6	30	5.30	5
	FECCOF01	4.74	3	22	5.45	3

Figure 2. H-bond in the crystal structure of tetracosane-1, 24-diol (RIWTIB).

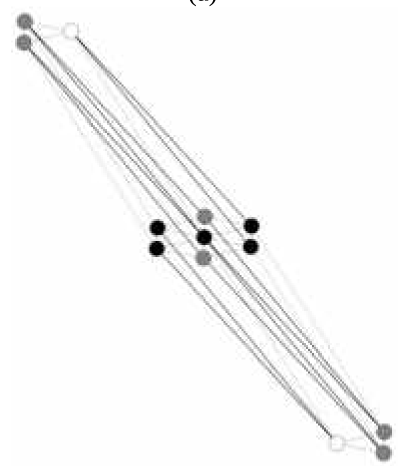
Since the molecule and its environment were centrosymmetric, to form a crystal, it would be enough for molecules to form identical contacts in a sphere of radius $2R = 24.2$ Å. Indeed, the structure had intermolecular contacts that did not exceed this distance in length and were suitable for the role of critical ones. The corresponding molecules are shown in Figure 3b. One of such symmetrically independent contacts had $d = 23.6$ Å, the other $d = 23.6$ Å, but both were not included even in the top 20 of the strongest contacts, and the energy of each $U < 0.2$ kJ/mol.

Moreover, the molecules located at these distances were shielded from the initial, respectively, by 3 and 2 H-linked 'sheets' of molecules, so that direct interatomic contact was out of the question here. The critical contact, as well as the entire first coordination sphere of the molecule in this structure, obviously went beyond $2R$. For a more detailed study of this pattern, a new series of crystal structures was required.

CONCLUSION

The critical coordination number (CCN) calculated in our study, usually exceeded its minimal possible value for a structural class. The heavier the molecule, the larger was CCN–min(CCN) difference due to the increasing diversity of the local molecular environment. On the other hand, CCN was much less than the molecular coordination number (MCN), even if the multiplicity of contact following from its symmetry was taken into account. The least strong contact required for the crystal structure formation (*i.e.,* critical contact) had the greatest average relative decrease of the contact energy. While equality of the local environment within a certain vicinity of a molecule was sufficient for the crystal structure formation, the critical contact was generally outside this vicinity.

(a)



(b)

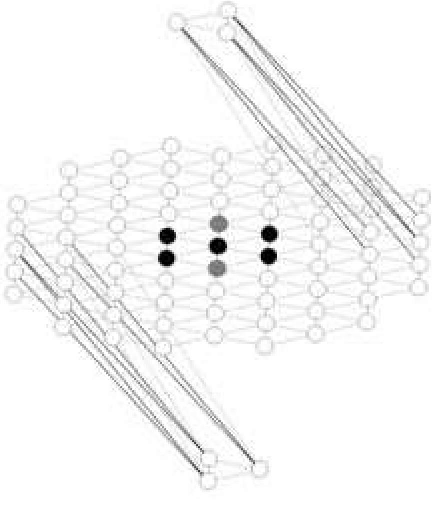


Figure 3. Molecular coordination spheres (concerning the mass centres) in the crystal structure of tetracosane-1, 24-diol (RIWTIB): (a) – the 1st coordination sphere, (b) – the environment in the radius of 24.2 Å. The most strongly bound molecules (E_N) are highlighted in black, the ones included in the CCN calculation are highlighted in gray. H-bonds are shown by a solid line, van-der-Waals contacts are shown by a dashed line.

For a correct comparison with MCN, other types of CN were also more convenient to consider taking into account symmetrically equivalent contacts. Typical values for all CNs were even. E_N and CCN reflected the number of intermolecular contacts, respectively, in the primary (the strongest) and secondary (less strong) agglomerates of the most tightly bound structural units, including molecules. Without taking into account the symmetrically equivalent contacts, CCN was an invariant of the structural class of the crystal. However, taking into account those, CCN was not invariant, because the choice of contacts generating the crystal structure without taking into account the energies was not unambiguous.

As a characteristic of the environment of a molecule in a crystal CCN had certain advantages. First, it took into account the nature of intermolecular interactions, in contrast to the MCN, and not only their topology. Secondly, the MCN had a less wide spread of values than the CCN, taking into account symmetrically equivalent contacts, and therefore was less sensitive to the local features of the crystal structure, as well as to the size and shape of molecules.

REFERENCES

1. Allen, F. H. (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising, *Acta Cryst.*, **B58**, 380-388.

2. Groom, C. R. and Allen, F. H. (2014) The Cambridge Structural Database in retrospect and prospect, *Angew Chem. Int. Ed.*, **53**, 662-671.

3. Kitaigorodskii, A. I. (1961) *Organic Chemical Crystallography*, Consultants Bureau, New York.

4. Peresypkina, E. V. and Blatov, V. A. (2000) Topology of molecular packings in organic crystals, *Acta Cryst.*, **B56**, 1035-1045.

5. Delone, B. N., Dolbilin, N. P., Shtogrin, M. I. and Galiulin, R. V. (1976) A local criterion for regularity of a system of points, *Sov. Math. Dokl.*, **17**, 319-322.

6. Senechal, M. (1986) Geometry and crystal symmetry, *Comp. & Maths. With Appls.*, **12B**, 565-578.

7. Dolbilin, N. (2016) Delone sets with congruent clusters, *Struct. Chem.*, **27**, 1725-1732.

8. Zorkii, P. M. and Zefirov, Y. V. (1972) On molecular coordination number, *Vestnik MGU, Ser. 2*, **13**, 590-593.

9. Grineva, O. V. (2017) Analysis of aliphatic α, ω-diols: crystal structures, application and progress of A. I. Kitaigorodsky ideas, *J. Struct. Chem.*, **58**, 373-383.

10. Motherwell, W. D. S. (2017) Architecture of packing in molecular crystals, *CrystEngComm.*, **19**, 6869-6882.

11. Banaru, A. M. (2009) Critical coordination number in homomolecular crystals, *Moscow Univ. Chem. Bull.*, **64**, 80-82.

12. Lord, E.A. and Banaru, A. M. (2012) Number of generating elements in space group of a crystal, *Moscow Univ. Chem. Bull.*, **67**, 50-58.

13. Banaru, A. and Kochnev, A. (2017) The minimal set of intermolecular interactions in the structures of substituted pralines, *Studia UBB Chemia*, **LXII**, 121-128.

14. Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. and Wood P. A. (2008) New features for the visualization and investigation of crystal structures, *J. Appl. Cryst.*, **41**, 466-470.

15. Gavezzotti, A. and Filippini, G. (1994) Geometry of the intermolecular X-H.cntdot..cntdot.. cntdot.Y (X, Y = N, O) hydrogen bond and the calibration of empirical hydrogen-bond potentials, *J. Phys. Chem.*, **98**, 4831-4837.

16. Bondi, A. (1964) van der Waals volumes and radii, *J. Phys. Chem.*, **68**, 441-451.