research papers

Methods of crystallochemical analysis of supramolecular complexes by means of Voronoi–Dirichlet polyhedra: a study of cucurbituril host–guest compounds

A. V. Virovets,[a] V. A. Blatov[b]* and A. P. Shevchenko[b]

[a]Institute of Inorganic Chemistry of Siberian Branch of RAS, Ac. Lavrent'ev Pr. 3, Novosibirsk 630090, Russia, and [b]Samara State University, Ac. Pavlov St. 1, Samara 443011, Russia

Correspondence e-mail: blatov@ssu.samara.ru

Crystallochemical analysis methods based on the Voronoi–Dirichlet partition of crystal space are extended to supramolecular complexes of any complexity. The sizes and shapes of receptor cavities and substrate molecules are shown to be successfully estimated as volumes and the second moments of inertia of the corresponding molecular Voronoi–Dirichlet polyhedra. To predict which organic substrates can occupy the receptor cavity a mini-expert system known as *MOLVOL* was created, comprising a database on completely determined crystal structures of almost 60 000 organic molecular compounds. Using the developed methods, volumes and shapes are assessed for cucurbit[n]uril receptors (n = 5–10) and their cavities. A number of organic and inorganic molecules are found which can optimally fit the cavities inside the cucurbit[5]uril and cucurbit[6]uril molecules.

1. Introduction

Supramolecular chemistry is a rapidly developing field of modern chemistry (Lehn, 1995), which deals with compounds based on the weak intermolecular interactions between receptors and substrates. They are supermolecules existing in the liquid phase and solid supramolecular ensembles, which include crystalline molecular aggregates (for instance, crystals of host–guest compounds). The problem of predicting the possible receptor–substrate combinations from structural data is a great challenge for modern crystal chemistry. For this purpose, crystallographic data stored in databases such as that at the Cambridge Structural Database (CSD) could be helpful if the appropriate computer methods were developed to automate crystallochemical analysis.

It is known that the receptor and substrate must be structurally and functionally complementary to each other (Lehn, 1995). Crystallochemical analysis could be useful in this connection to assess the size and shape of possible receptors and substrates. In the case of inclusion compounds it is necessary to estimate the volumes of the substrate molecule and the cavity inside the receptor.

The widespread method of estimating the geometrical parameters of molecules and cavities is based on van der Waals radii (Kitaigorodskii, 1973; Alard & Wodak, 1991; Thomas, 1991; McArdle & Cunningham, 2000). In this approach a molecule is considered to be a spatial region bound by the surfaces of such overlapping spheres. However, the concept of the van der Waals radius has essential disadvantages, *e.g.* the absence of a conventional set of van der

Waals radii (Zefirov, 1997). Therefore, one needs to develop an alternative geometrical molecular-crystal model that requires no van der Waals radii. Such a model is based on the concept of a molecular Voronoi–Dirichlet polyhedron (VDP), *i.e.* the union of Voronoi–Dirichlet polyhedra of the atoms composing a molecule (Fischer & Koch, 1979). A number of studies show the efficiency of this approach to analyze the close environment of molecules (Peresypkina & Blatov, 2000a), molecular packings (Peresypkina & Blatov, 2000b) and steric effects in molecular coordination compounds (Blatova et al., 2001). Recently, Blatov & Shevchenko (2003) extended the VDP method to find voids, cavities and channels in crystal structures and to calculate their geometrical parameters. All the methods are implemented in the computer package *TOPOS* (Blatov et al., 2000), which allows one to process crystal structures of any complexity and large sets of crystallographic database entries. However, this approach has still not been applied in supramolecular chemistry; our investigation is an attempt to overcome this trend.

2. Analysis of supramolecular complexes with Voronoi–Dirichlet polyhedra (VDP)

2.1. Volume and shape of molecules

In a number of studies (Koch & Fischer, 1980; Thomas, 1996; Christensen & Thomas, 1999; Blatov & Serezhkin, 2000) the atomic VDP volume was shown to correctly evaluate not only the relative, but also the absolute atomic size in the crystal structure. Since the molecular VDP is a union of atomic VDPs, one could expect that its volume would correspond to
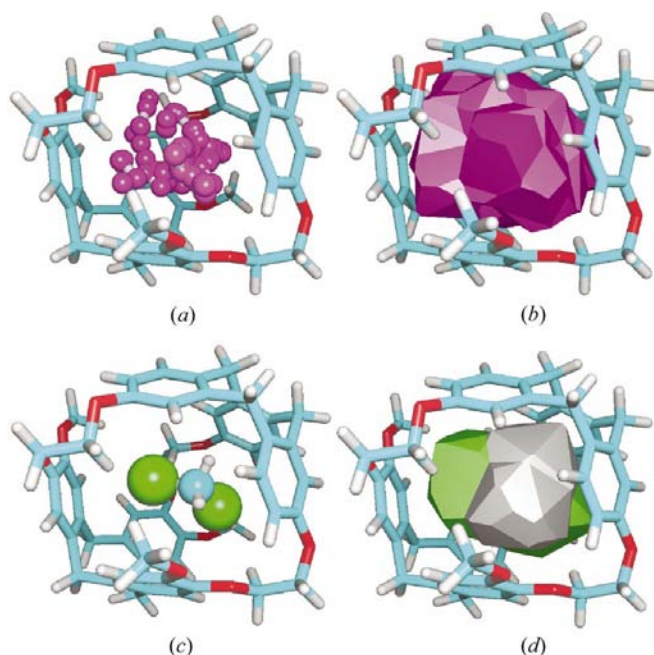
molecular size in a crystal field. Thus, hereinafter we associate molecular VDP volume (V_{VDP}) with the volume of molecule in the crystal. Sometimes the radius of the spherical domain (R_{sd}) is useful to characterize the size of molecule, which is equal to the radius of a sphere of VDP volume.
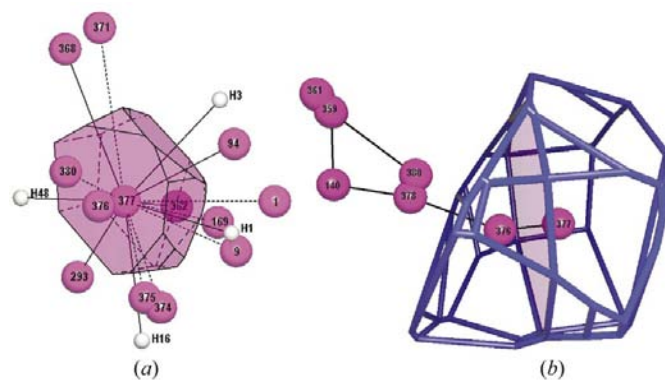
Peresypkina & Blatov (2003) proposed one more geometrical characteristic of molecular VDP to be useful in the analysis of molecular shape. This is the degree of sphericity, which can be evaluated by the VDP normalized second moment of inertia

$$G_3 = \frac{1}{3}\left[\left(\sum_i V_i I_i\right)/V_{VDP}^{5/3}\right], \qquad (1)$$

where V_i and I_i are the volume and the second moment of inertia of an ith simplex in the simplicial dissection of molecular VDP; I_i values are calculated relative to the molecule centroid and the summation is provided over all simplexes in the dissection. The G_3 magnitude is an integral parameter: the greater the G_3 value the smaller the degree of sphericity of a molecule. The smallest value, $G_3 = 0.07697$, corresponds to a sphere (Conway & Sloane, 1988).

2.2. Voids and atomic Voronoi–Dirichlet polyhedra

Blatov & Serezhkin (2000) showed that the topological features of the electron density function in a crystal field determined using the method of Bader (1990) could be associated with the combinatorial properties of atomic VDPs. In particular, VDP vertices indicate void centres and VDP edges correspond to channels between the voids. Furthermore, we will term the crystal space domain embracing a VDP vertex the *elementary void* or simply the *void*. The centre of an elementary void defined in this way does not always correspond to the centre of a real cavity in the crystal structure, which can contain N VDP vertices and, hence, can be a conglomerate of N elementary voids (*N-conglomerate*). Such a situation is ordinal for the crystal structures with bulky cavities decorated with many atoms, in particular, for molecular receptors.

Figure 1
Cryptophane molecules in the crystal structure of cryptophane dichloromethane clathrate (Canceill et al., 1986) containing: an *N*-conglomerate of elementary voids in (a) graph and (b) polyhedral representations; (c) dichloromethane molecule and (d) its molecular VDP. $N = 84$; $V_{VDP} = 130.5$ Å^3; $R_{sd} = 3.15$ Å; $G_3 = 0.0869$.

Figure 2
The crystal structure of cryptophane dichloromethane clathrate (Canceill et al., 1986): (a) VDP and the nearest environment of the void 377; (b) channel between 'direct' neighbours 377 and 376 represented by a VDP face with $\Omega = 36.9\%$.

The two representations of *N*-conglomerates are useful for their visualization and analysis (Blatov & Shevchenko, 2003): as a three-dimensional graph (*graph representation*, Fig. 1*a*) and as a union of *N* VDPs of the conglomerate voids (*polyhedral representation* as a molecular VDP, Fig. 1*b*). The first representation is intended for the topological analysis of *N*-conglomerates; the second is necessary to calculate the geometrical characteristics of the corresponding cavities.

2.3. Determination of *N*-conglomerate topological properties

An automatic determination of the connectivity of elementary voids within an *N*-conglomerate is based on a 'secondary' Voronoi–Dirichlet partition where void centres are the centres of the VDPs. Whilst constructing such VDPs, all atoms are considered together with all the void centres (Fig. 2*a*). Similar to the criterion for the existence of interatomic contacts (Blatov & Serezhkin, 2000), one can assume the existence of a connecting channel between a pair of voids if the following two conditions are satisfied:

(i) the solid angle (Ω) of the separating VDP face between void centres (*i.e.* the face that separates the void centres from each other) is greater than some value which is ordinarily equal to 1.5% of the full solid angle 4π steradian. Actually, this value corresponds to a channel with a minimum allowable section (Fig. 2*b*);

(ii) the voids are 'direct' neighbours (O'Keeffe, 1979), *i.e.* the line passing through the void centres intersects the VDP face (Fig. 2*b*).

The void subspace is often not connected simply (*i.e.* it can contain isolated void conglomerates). Therefore, its dimensionality (zero-, one-, two- and three-dimensional; Blatov & Shevchenko, 2003) should be determined before performing graphical–topological analysis.

2.4. Determination of the geometrical characteristics of void conglomerates

The notion of volume is reasonable only if used with zero-dimensional void conglomerates, which will be considered exclusively in this study. When a conglomerate consists of a single void (1-conglomerate), its volume is equal to the volume of the VDP constructed with all its neighbouring

atoms and its radius is equal to R_{sd}. In this case, V_{VDP} is equal to the volume of an atom, which would have a given atomic environment. This is why the void volumes calculated in this way can be compared with the atomic sizes estimated as the volumes of atomic VDPs in a given environment. Moreover, the problem disappears for positioning VDP faces relative to atoms (this is the main difficulty with using VDPs in crystal chemistry) of both void and atomic VDPs, provided the division coefficient is the same (for simplicity its value can be taken equal to 0.5 that corresponds to the VDP definition).

The V_{VDP} and R_{sd} of a zero-dimensional *N*-conglomerate ($N > 1$) are equal to the corresponding characteristics of the molecular VDP. An estimation of the size of the *N*-conglomerates can be important for solving the problems of supramolecular chemistry, but most of all, the problems of molecular recognition. Thus, the 84-conglomerate in a polyhedral representation (Fig. 1*b*), constructed for a cavity in cryptophane molecules, fits small organic molecules such as dichloromethane and ethanol whose volumes were estimated with X-ray data [DCLMET10 and ETANOL][1] to be 82.1 and 75.5 Å^3, respectively. Indeed, the dichloromethane molecule occupies the cryptophane cavity (Fig. 1*c*) and the volume of its molecular VDP (82.0 Å^3, Fig. 1*d*) nearly coincides with the volume of the CH_2Cl_2 molecule in crystalline dichloromethane.

The degree of sphericity of the cavity can be assessed in the same manner as with a molecule using the G_3 criterion (1).

2.5. Differentiation of voids

After restoring the connectivity the set of VDP vertices is similar to an atomic net in its graph representation. However, in contrast to interatomic distances, the distances between the voids are not limited from below. Besides, voids cannot be *a priori* differentiated by type as atoms can be by chemical type. At the same time the differentiation of voids would be useful when searching for the cavities with given characteristics.

A kind of differentiation can be proposed as follows (Blatov & Shevchenko, 2003). Notice that a VDP vertex is equidistant from a tetrad of non-coplanar atoms. If some of these four atoms are valence-bonded with each other and, hence, are arranged close to each other, the void volume could be small and insufficiently large to fit other atoms. Obviously, if only large cavities are of interest (as in the case of supramolecular complexes) such VDP vertices should be excluded from consideration. Since there are six contacts between four atoms surrounding each VDP vertex, it is useful to introduce the vertex rank, which is equal to the number of non-valence contacts in the vertex atomic environment and which varies in the range 0–6. The vertices of the highest rank are expected to correspond to the cavities of the largest volume. If the vertices of a lower rank are ignored then the *N*-conglomerates corresponding to the small cavities disappear and the void map becomes simpler.
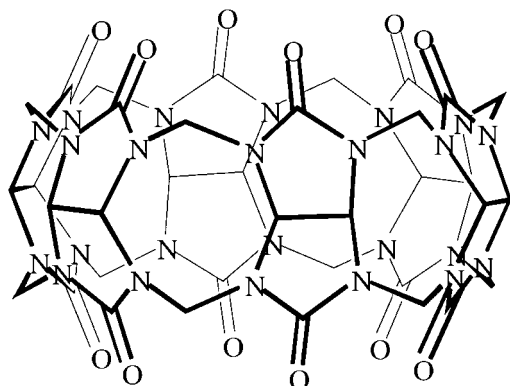
Figure 3
A cucurbit[*6*]uril molecule.

---

[1] Hereinafter, the refcodes for compounds in the CSD (version 5.24, October 2002) are given in square brackets.

Table 1
Some parameters of the cucurbituril molecules and their cavities.

Cucurbituril molecule Qn			Cavity (N-conglomerate)				
n	Volume (Å^3)	G_3	N	Volume (Å^3)	R_{sd} (Å)	G_3	Reference code
5	873.5	0.0849	42	79.0	2.66	0.0878	LIRTEL
	1098.4†	0.0852	47	69.5	2.55	0.0810	IDIWEX
6	1288.4	0.0803	9	101.4	2.89	0.0813	GUMWAM
	1264.1	0.0858	72	121.5	3.07	0.0799	GUMWEQ
	1269.8	0.0964	32	93.7	2.82	0.0838	GUMWIU
	1441.2	0.0942	36	105.6	2.93	0.0816	GUMWOA
	1363.0	0.0905	44	101.9	2.90	0.0807	GUQCEA
	1357.3	0.0922	64	122.1	3.08	0.0784	MAYBET
	1354.2	0.0925	70	122.6	3.08	0.0789	MAYBIX
	1133.2	0.0889	68	118.0	3.04	0.0793	MAYBOD
	1334.0	0.0885	69	119.2	3.05	0.0861	QIMPUX
	1467.7	0.0956	51	110.4	2.98	0.0786	QUQLOD
	1463.4	0.0958	57	112.8	3.00	0.0781	QUQLUJ
	1426.9	0.0843	64	125.2	3.10	0.0806	UCANOB
	1212.3	0.0987	20	100.0	2.88	0.0826	UCAROF
	1287.4	0.0880	66	119.7	3.06	0.0801	VEFWIQ
	1425.6	0.0841	64	118.8	3.05	0.0806	XAVXUN
	1327.2	0.0878	67	122.1	3.08	0.0806	XEMQAH
Average	1338 (94)	0.090 (5)	53 (20)	113 (10)	3.00 (9)	0.081 (2)	
7	1672.8‡	0.0886	78	186.6	3.55	0.0803	LIRTIP
8	1840.6	0.1048	44	219.9	3.74	0.0823	LIRTOV
	2563.8‡	0.0892	86	292.2	4.12	0.0822	LIRTUB
10	4097.6‡	0.0887	123	613.8	5.27	0.0950	IDIWEX

† The $Q5$ molecule is contained inside the $Q10$ molecule.   ‡ The Qn molecule contains other organic molecules.

In this study we use an additional method of void differentiation by taking into consideration only the 'own' voids of the receptor molecules. The term 'own' means that these voids are bounded by only the atoms of the same molecule, not by atoms of other molecules. In other words, the four atoms which are equidistant from the centre of an 'own' void (VDP vertex) should belong to the same receptor molecule. This criterion is found to be very efficient in separating intra- and intermolecular voids.

3. Investigation objects

To extend the VDP method to supramolecular compounds it is important to select the proper species, which are widely investigated and whose geometrical properties vary only slightly from one crystal structure to another. The compounds that seem to fit these criteria are supramolecular complexes based on the organic cucurbituril cavitands. Cucurbit[n]urils (Qn) have the composition $\{C_6H_6N_4O_2\}_n$ and can be obtained by acid-catalyzed condensation of glycoluril and formaldehyde (Fig. 3).

Structural and spectroscopic investigations show that a Qn molecule can include into its cavity various organic molecules or molecular ions, such as mono- and diamines, alkyl- and benzylammonium cations and some organic dyes resulting in guest–host complexes (Mock, 1996). Another important feature of cucurbiturils is the presence of polar carbonyl groups forming two portals ('*occuli*'). The O atoms of portals can interact with H-atom donors to form strong hydrogen bonds as well as acting as ligands for some metal cations

(Mock, 1996; Park *et al.*, 1999; Heo *et al.*, 2000; Buschmann *et al.*, 2001; Gerasko *et al.*, 2002).

From the structural chemistry point of view it is interesting to assess the size and shape of the cavity inside the Qn molecules to predict which inorganic and organic molecules and ions can fit. The only attempt at this assessment was made by Kim *et al.* (2000) for n = 5–8, but the authors gave no information on the methods of estimating Qn cavity volumes.

4. Experimental

At first, only the information on the crystal structures of cucurbit[n]urils and their complexes was extracted from the CSD (Version 5.24, October 2002). Only compounds with free Qn molecules (containing no substrate molecules) or their complexes with small inorganic molecules (like water) were considered. Such compounds were found only for $Q5$, $Q6$ and $Q8$. Some complexes with organic molecules inside Qn molecules were taken into account to compare their characteristics with free Qn molecules. The only $Q10$ complex [IDIWEX] containing a $Q5$ molecule in the $Q10$ void was also taken into consideration. When constructing molecular VDPs all atoms of Qn molecules were considered, while H atoms of other structural fragments were dismissed. Water molecules were also omitted because they were frequently determined only approximately in the cucurbituril crystal structures. All the compounds studied are given in Table 1 together with their main size and shape characteristics.

To investigate cavities in Qn molecules we have improved the program package *TOPOS* (Blatov *et al.*, 2000; Blatov & Shevchenko, 2003). Coordinates of void centres (VDP vertices) and void ranks (0–6) were calculated with the program *Dirichlet*. Only 'own' cucurbituril voids with ranks 5 and 6 were considered when constructing N-conglomerates. Void connectivity was automatically determined with the program *AutoCN*. The automatic recognition of simply connected d-dimensional void conglomerates and the determination of their dimensionality (d = 0–3) were made by means of the program *ADS*. The visualization of voids in graphic/polyhedral representations and their geometrical analysis were realised using the program *IsoCryst*.

To predict which organic molecules can occupy Qn cavities we have created a mini-expert system, *MOLVOL*, comprising a database of completely determined crystal structures from almost 60 000 organic molecular compounds. For every non-equivalent molecule in every compound, the volume and the G_3 parameter of the molecular VDP were calculated and stored in the database. Then the averaged values for each type

of molecule were obtained. *MOLVOL* requires the following data to be input:

   (i) minimum and maximum cavity volume;

   (ii) number of molecules to be placed into the cavity;

(iii) minimum and maximum G_3 of the cavity (if only one molecule is expected to occupy the cavity).

   MOLVOL then turns to the database and looks for the molecules that fit the specified conditions. No 'chemical' but
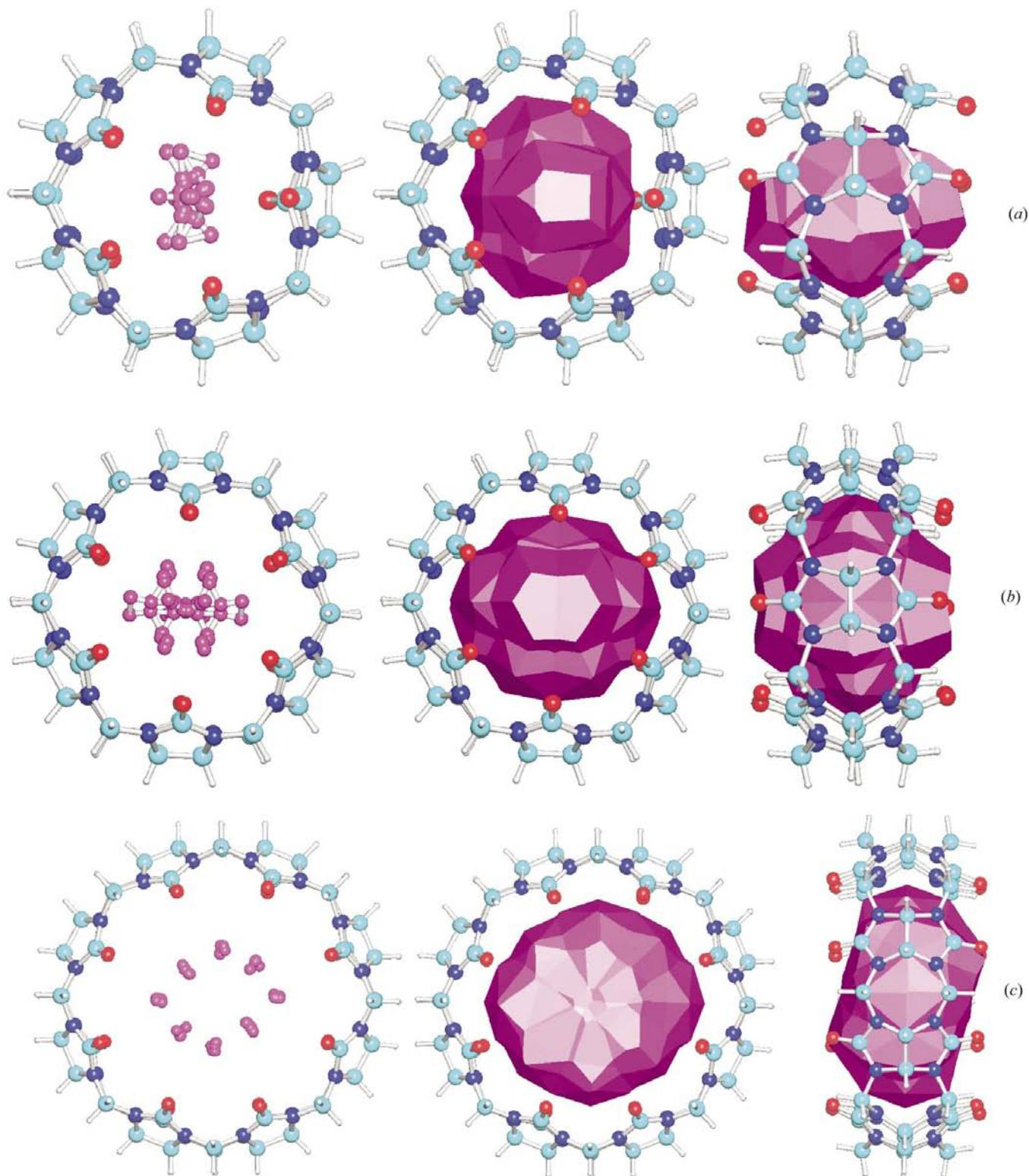

Figure 4
N-Conglomerate and two views of the molecular VDP of a cavity inside a *Qn* molecule: (*a*) $N = 42$, $n = 5$ [LIRTEL]; (*b*) $N = 36$, $n = 6$ [GUMWOA]; (*c*) $N = 44$, $n = 8$ [LIRTOV].

Table 2
Molecules or ions with V_{VDP} = 70–85 Å^3 and G_3 = 0.07695–0.09000, which could optimally fit the cavity inside cucurbit[5]uril.

Molecule	Number of compounds/molecules	V_{VDP} (Å^3)	G_3
$H_3N\cdot BH_3$	10/13	72 (4)	0.086 (2)
SeF_5^-	1/2	72 (2)	0.088 (0)
BrO_4^-	1/1	73.2	0.0865
SiF_6^{2-}	21/26	74 (3)	0.088 (2)
IO_4^-	2/2	74.2 (8)	0.084 (4)
H_2SO_4	1/2	74 (5)	0.085
CHF_2COO^-	1/1	74.4	0.0865
H_3PO_3	1/1	76.2	0.0873
(triazole structure)	2/2	76.3 (4)	0.090 (2)
$H_3N\cdot BF_3$	1/1	76.6	0.0892
$H_3CPH_3^+$	1/1	76.8	0.0876
CH_2Cl_2	335/419	79(5)	0.090(2)
B_2H_6	1/1	79.0	0.0855
H_3PO_4	5/5	79 (2)	0.088 (2)
SiF_5^-	2/2	79 (4)	0.086 (2)
$NH_2C(=O)NH_3^+$	11/13	79 (4)	0.090 (2)
$N_2H_4\cdot BH_3$	1/1	79.4	0.0898
SOF_5^-	1/1	80.8	0.0840
$[H_2N=N(H_2)CH_3]^+$	3/3	81 (6)	0.089 (1)
$CHFCl_2$	1/3	81 (3)	0.087 (2)
TeF_5^-	1/1	82.0	0.0863
(ring structure)	1/1	84.1	0.0879

only 'geometrical' possibilities are taken into account, so the user needs to control functional receptor–substrate compatibility separately.

5. Results and discussion

5.1. Molecular and cavity characteristics

Typical N-conglomerates in cucurbituril molecules and their corresponding molecular VDPs for cavities are given in Figs. 4(a)–(c). The following conclusions can be made when analyzing the shape and geometrical characteristics of Qn molecular VDPs (Table 1):

(i) The sizes of the $Q6$ molecules are practically the same. Indeed, the standard deviation in the VDP volume (7.0%) is close to a typical error for an X-ray experiment. This fact correlates with the effect of constancy in the molecule volume (Blatova et al., 2001).

(ii) The volume for Qn molecules is a linear function of n at n = 5–8: V_{VDP} = 324n–673 with the correlation coefficient ρ = 0.98. Let us emphasize that this volume includes the volume of the Qn cavity because molecular VDPs completely cover cavities for all Qn, except $Q10$. Cavity volumes (V_{cav}) also depend on n linearly: V_{cav} = 50n–173 with ρ = 0.99.

(iii) The sphericity of the Qn molecules regularly decreases with increasing n that is reflected by an increase in G_3 values for their molecular VDPs. At the same time, the sphericity of the Qn cavities varies non-monotonically: it is rather low for

$Q5$, high and practically the same for $Q6$ and $Q7$, and slightly lower for $Q8$.

The Qn molecular VDP characteristics may be correctly estimated only for isolated cucurbituril molecules containing no other organic substrates. The reason for this is that the substrate can occupy not only the Qn cavity, but also can extend outside the cucurbituril portals and can shield the Qn molecule that results in a distortion of its molecular VDP. For instance, in the crystal structure of a p-xylylenediammonium complex with $Q6$ [CISWOQ] (Fig. 5a) the elongated substrate molecules strongly influence the $Q6$ VDP. Actually, the main VDP parameters calculated with or without considering xylylenediammonium ion are: V_{VDP} = 1132.5 and 1378.6 Å^3; G_3 = 0.1041 and 0.0841, respectively. However, substrate molecules very slightly influence the geometrical characteristics of cavities in $Q5$ and $Q6$. Thus, the $Q6$ cavity characteristics for the CISWOQ complex (V_{cav} = 125.3 Å^3; G_3 = 0.0813, Fig. 5b) are typical for free cucurbit[6]uril molecules (Table 1). Apparently, this is due to the high isolation of the cavity (shielding by portal O atoms; see Figs. 4 and 5) and the high rigidity of the cucurbit[6]uril molecules. Starting from $Q8$ (data for $Q7$ are insufficient to make an unambiguous conclusion), the cavity is open and its localization (and hence correct estimation of molecular VDP characteristics) becomes ambiguous. Therefore, our conclusions on the stability of cucurbituril complexes will concern only $Q5$ and $Q6$ receptors.

5.2. Forecasting cucurbituril complexes

Using the MOLVOL system we have obtained a list of molecules which can optimally fit the cavities inside the $Q5$ and $Q6$ molecules (Tables 2 and 3). In fact, the variety of possible guest molecules is much larger. First of all, we selected the molecules whose volumes are close to the volume of the cucurbituril cavity. Smaller molecules, of course, can also fit the cavity, but such host–guest complexes would be less stable. Secondly, our aim was to find guest molecules that would cause no distortion of the cucurbituril hosts, therefore we selected only more or less spherical molecules with rather small G_3 values (less than 0.090). For example, the pyridinium cation, $C_5H_6N^+$, has a suitable average volume of 122.8 Å, but its average G_3 value of 0.092 lies outside the limits. It is known that its inclusion into $Q6$ makes the cucurbituril molecule
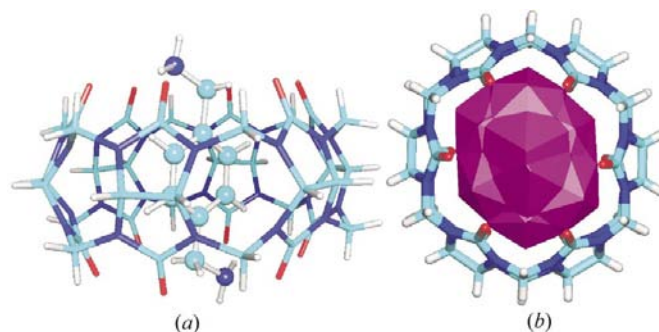
Figure 5
Complex of cucurbit[6]uril with xylylenediammonium [CISWOQ]: (a) molecular structure; (b) cavity view.

Table 3
Molecules or ions with V_{VDP} = 100–125 Å³ and G_3 = 0.07696–0.09000, which could optimally fit the cavity inside cucurbit[6]uril.

Molecule	Number of compounds/molecules	V_{VDP} (Å³)	G_3
$C_2H_5NH_3^+$	9/12	98 (6)	0.086 (1)
$HOCH_2CH_2NH_3^+$	7/9	101 (6)	0.087 (2)
$(CH_3)_2NH_2^+$	27/37	102 (8)	0.087 (2)
[guanidinium, H_2N–C(=NH_2^+)–]	7/7	102 (8)	0.090 (1)
$HOCH_2CH_2CH_2OH$	1/1	104.0	0.0870
$F_2ClCCOOH$	4/4	105 (6)	0.089 (1)
AsF_6^-	80/104	105 (10)	0.083 (1)
C_4H_8 (cyclobutane)	5/5	105 (5)	0.088 (3)
$CF_3SO_3^-$	183/226	106 (6)	0.087 (2)
TeF_7^-	1/1	106.3	0.0839
CF_3I	1/1	106.5	0.0895
[aziridine–N–BH_3]	1/1	107.2	0.0874
$H_2NCH_2CH_2NH_3^+$	2/3	108 (6)	0.088 (1)
CCl_4	24/25	108 (9)	0.086 (2)
$CFCl_3$	1/1	108.6	0.0854
[cyclopropyl–NH_3^+]	1/2	109 (4)	0.085 (0)
$^+NH_3CH_2CH_2NH_3^+$	51/74	110 (6)	0.089 (2)
[S(=O)NH structure]	1/1	110.7	0.0873
$(CH_3)_3N$	2/2	113 (2)	0.0868 (0)
Cl_3CCOO^-	13/14	113 (4)	0.088 (1)
[tetrahydrofuran]	2/2	114 (9)	0.090 (1)
iso-C_3H_7OH	44/62	114 (7)	0.088 (2)
[O O S structure]	1/1	115.9	0.0821
[HN NH S structure]	1/1	116.2	0.0865
$(CH_3)_3NO$	3/4	117 (5)	0.087 (1)
$NCCH_2NH_2CH_3^+$	3/3	118 (2)	0.088 (3)
$CHBr_3$	3/3	118 (3)	0.0892 (0)
Cl_3CCOOH	1/1	119.3	0.0888
$F_3CS(=O)_2OH$	1/1	119.3	0.0839
$HOOCCF_2COOH$	1/1	119.4	0.0892
$(CH_3)_2SH^+$	2/2	119 (3)	0.090 (2)
H_6TeO_6	5/6	120 (6)	0.087 (2)
[pyrazolidine–NH_2^+ structure]	2/3	120 (8)	0.087 (5)
[cyclopropyl–NH_2 structure]	1/1	121.0	0.0857
$F_5SNNO_2^-$	2/2	121 (2)	0.090 (1)
[S N ring structure]	1/1	121.6	0.0876
$(CH_3)_3NH^+$	19/24	122 (10)	0.085 (1)
B_4H_{10}	4/4	122.4 (2)	0.0838 (0)

Table 3 (continued)

Molecule	Number of compounds/molecules	V_{VDP} (Å³)	G_3
$F_2ClCC(=O)NH_2$	1/1	122.7	0.0897
[cyclohexenone]	1/1	123.2	0.0897
[S–Cl trithiadiazine structure]	1/1	124.1	0.0890
$(CF_3)_2NH_2^+$	1/1	124.4	0.0887

ellipsoidally distorted (Samsonenko et al., 2002). Thirdly, in some cases the derivatives of the organic molecules listed in Tables 2 and 3 can also fit the cavity if the additional substituents have the appropriate 'diameters' and can penetrate through the portal of Qn partly lying outside the host molecule (for instance, p-substituted benzenes; Mock, 1996). Finally, so far we have only used the CSD to create the MOLVOL system, therefore, some possible inorganic guests might have been missed.

The possible guest molecules predicted for $Q6$ (Table 3) correspond to the experimental data regarding ligand affinity (Mock, 1996). Thus, according to Mock (1996), the most stable are the complexes with the following ligands:

(i) n-butyl derivatives, such as $Me(CH_2)_3NH_2$;

(ii) cyclobutane and cyclopentane derivatives, such as $(CH_2)_3CHCH_2NH_2$ and $(CH_2)_4CHCH_2NH_2$;

(iii) five-membered aromatic rings (such as furan or thiophene), which fit the cavity better than the six-membered benzenes.

Mock (1996) pointed out that amino substituents protruded into a cucurbituril portal to interact with the portal O atoms, therefore, n-butyl, cyclobutane and cyclopentane substituents, and five-membered aromatic rings optimally filled the cavity. According to Table 3 it is the tetrameric chains (such as $NH_2CH_2CH_2NH_3^+$ or $HOCH_2CH_2NH_3^+$) that have the volume close to V_{cav}. If the atoms in the chain are bonded with fewer H atoms, the number of chain links could reach five ($HOCH_2CH_2CH_2OH$). Note that the volume of the n-butane [DUCKOB04] molecular VDP is 119.6 Å³ and this ligand is not included in Table 3 because of the rather large G_3 value of 0.0910. Cyclobutane can easily fill the cavity (Table 3), but the volume of cyclopentane (131.3 Å³) is slightly larger than the cavity volume. The cucurbit[6]uril molecule probably becomes ellipsoidally distorted when coordinating with cyclopentane derivatives, as in the case of benzenes. Finally, Table 3 contains five-membered aromatic rings (and their hydrated derivatives), not six-membered (except trithiadiazine, $CH_2N_2S_3$, with only two H atoms in the ring).

It should be noted that no chemical considerations were taken into account whilst placing molecules in Tables 2 and 3. Of course, some molecules will not form the host–guest complexes with Qn due to chemical reasons. To plan future syntheses it is necessary to take into account the fact that

cucurbiturils are soluble only in water, not in organic solvents, and only at low pH values, about 2–3, or in the presence of *s*- and *f*-metal cations that form complexes with direct M—$O(Qn)$ coordination (Samsonenko *et al.*, 2002). An additional limitation is that the guest molecule must not react with Qn. For instance, the perbromate anion successfully passes crystallochemical tests (Table 2), but probably will oxidize $Q5$.

6. Conclusions

The data obtained show that molecular VDPs can successfully be used to assess the geometrical fitness of a guest for a receptor's cavity. The principal advantages of the proposed method are its independence of any empirical parameters (such as van der Waals radii) and hence the possibility of being used for ligands and receptors of any chemical composition and geometry. In future this approach can be extended by considering not only the geometrical parameters, but also functional properties. It means that molecular VDP ceases to be 'unicoloured' because its faces will differ from each other according to the chemical nature of the atoms.

References

Alard, P. & Wodak, S. J. (1991). *J. Comput. Chem.* **12**, 918–922.
Bader, R. F. W. (1990). *Atoms in Molecules – a Quantum Theory.* Oxford University Press.
Blatov, V. A. & Serezhkin, V. N. (2000). *Russ. J. Inorg. Chem.* Suppl. S105–S220.
Blatov, V. A. & Shevchenko, A. P. (2003). *Acta Cryst.* A**59**, 34–44.
Blatov, V. A., Shevchenko, A. P. & Serezhkin, V. N. (2000). *J. Appl. Cryst.* **33**, 1193.
Blatova, O. A., Blatov, V. A. & Serezhkin, V. N. (2001). *Acta Cryst.* B**57**, 261–270.
Buschmann, H.-J., Mutihac, L. & Jansen, K. (2001). *J. Incl. Phenom.* **39**, 1–11.
Canceill, J., Cesario, M., Collet, A., Guilhem, J., Riche, C. & Pascard, C. (1986). *J. Chem. Soc. Chem. Commun.* pp. 339–341.
Christensen, S. W. & Thomas, N. W. (1999). *Acta Cryst.* A**55**, 811–820.
Conway, J. H. & Sloane, N. J. A. (1988). *Sphere Packings, Lattices and Groups.* Germany: Springer Verlag.
Fischer, W. & Koch, E. (1979). *Z. Kristallogr.* **150**, 245–260.
Gerasko, O. A., Samsonenko, D. G. & Fedin, V. P. (2002). *Russ. Chem. Bull.* **71**, 840–861.
Heo, J., Kim, S.-Y., Roh, S.-G., Park, K. M., Park, G.-J., Whang, D. & Kim, K. (2000). *Mol. Cryst. Liq. Cryst.* **342**, 29–38.
Kim, J., Jung, I.-S., Kim, S.-Y., Lee, E., Kang, J.-K., Sakamoto, S., Yamaguchi, K. & Kim, K. (2000). *J. Am. Chem. Soc.* **122**, 540–541.
Kitaigorodskii, A. I. (1973). *Molecular Crystals and Molecules.* New York: Academic Press.
Koch, E. & Fischer, W. (1980). *Z. Kristallogr.* **153**, 255–263.
Lehn, J.-M. (1995). *Supramolecular Chemistry: Concepts and Perspectives.* New York: Wiley VCH.
McArdle, P. & Cunningham, D. (2000). *J. Appl. Cryst.* **33**, 993.
Mock, W. L. (1996). *Comprehensive Supramolecular Chemistry*, edited by F. Vögtle, Vol. 2, pp. 477–493. Oxford: Pergamon Press.
O'Keeffe, M. (1979). *Acta Cryst.* A**35**, 772–775.
Park, K. M., Heo, J., Roh, S.-G., Jeon, Y.-M., Whang, D. & Kim. K. (1999). *Mol. Cryst. Liq. Cryst.* **327**, 65–70.
Peresypkina, E. V. & Blatov, V. A. (2000a). *Acta Cryst.* B**56**, 501–511.
Peresypkina, E. V. & Blatov, V. A. (2000b). *Acta Cryst.* B**56**, 1035–1045.
Peresypkina, E. V. & Blatov, V. A. (2003). *Russ. J. Inorg. Chem.* **48**, 237–245.
Samsonenko, D. G, Lipkowski, J., Gerasko, O. A. & Fedin, V. P. (2002). *Zh. Srukt. Khim.* **43**, 715–720.
Thomas, N. W. (1991). *Acta Cryst.* B**47**, 588–597.
Thomas, N. W. (1996). *Acta Cryst.* B**52**, 939–953.
Zefirov, Yu. V. (1997). *Crystallogr. Rep.* **42**, 111–116.