

A Possible Route toward Expert Systems in Supramolecular Chemistry: 2-Periodic H-Bond Patterns in Molecular Crystals

Pavel N. Zolotarev,[†] Muhammad Nadeem Arshad,^{‡,§} Abdullah M. Asiri,^{‡,§} Zahra M. Al-amshany,[‡] and Vladislav A. Blatov*,^{†,‡}

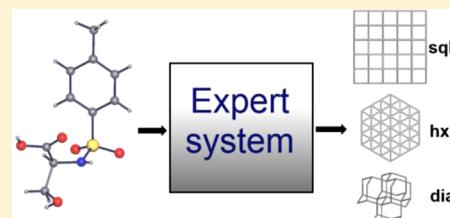
[†]Samara Center for Theoretical Materials Science (SCTMS), Samara State University, Ac. Pavlov St. 1, Samara 443011, Russia

[‡]Chemistry Department, Faculty of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

[§]Centre of Excellence for Advanced Materials Research (CEAMR), King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

Supporting Information

ABSTRACT: A novel approach to prediction of supramolecular motifs was applied to more than 6000 monomolecular structures containing 2-periodic H-bond patterns. It is shown that a number of topological descriptors allow one to rationalize supramolecular motifs, find the regularities in their structure, and store the information in a knowledge database. The knowledge database can then be used in an expert system to mimic the work of a human expert and to forecast the method of assembling molecules into supramolecular ensembles and into extended (periodic) architectures. The crystals of *N*-[(4-methylbenzene)sulfonyl]serine were synthesized, and the principles of the expert system were used to successfully predict the 2-periodic square-lattice H-bond pattern in this compound.



1. INTRODUCTION

In the past ten years, the methods for predicting crystal architectures have reached an essential progress. First of all, it concerns quantum-mechanical methods and other simulation approaches^{1,2} that were successfully used to model crystal structures of quite different classes of compounds: intermetallic, ionic inorganic, metal–organic, molecular. As a result, one can quantitatively evaluate a number of physical properties of a particular chemical substance.

Another way to predict crystal structures is to some extent controversial to the modeling methods and rests upon comprehensive geometrical and topological analysis of large samples of data on crystal structures usually taken from electronic crystallographic databases.^{3,4} An important advantage of the geometrical–topological methods is that they allow one to process the whole set of experimental data on crystal structures, i.e., more than 700 000 records collected in the CSD,⁵ ICSD,⁶ and Pearson's Crystal Data⁷ to find general regularities that can be intrinsic to large groups of compounds. The topological description of substances in terms of graphs and nets is habitual for chemists and widely used for various classes of chemical compounds^{8–10} including H-bonded molecular packings.^{11–13} Being not always directly applicable for determination of physical properties, the geometrical–topological methods help to formulate problems that can encourage experimental or quantum-mechanical investigations.

Automated processing of the databases requires computing a number of parameters that characterize topology of crystal structures.^{1,14} It is the correlations between these parameters that allow one to discover solid-state laws in a huge amount of

experimental data. Thus, invention of new parameters is very important to find new correlations and regularities in crystallographic data.

The regularities can then be used to predict crystal structures. Recently,⁹ we have discussed how to create and use an expert system that can let the user forecast overall topological motifs in coordination polymers resting upon minimal data on the chemical composition of the building units, from which the polymer is assembled. In this study, we will show how new topological descriptors can help to develop expert systems in supramolecular chemistry to predict the methods of assembling molecules into extended architectures. Special attention will be paid to 2-periodic motifs of H-bonded organic molecules because the experimental data on this class of compounds are very rich while the correlations between the topological parameters are clear enough for demonstration.

2. EXPERIMENTAL SECTION

2.1. Topological Description of Molecular Ensembles. To automate the analysis of H-bonded molecular ensembles, we have used our recently proposed topological approach.¹⁵ This approach includes several steps to formalize the crystallographic data on the molecular crystal and render them processable in an automated mode.

Starting from the crystallographic information, we determine all intra- and intermolecular interactions. For the compounds under consideration, we deal only with conventional H bonds as intermolecular contacts. By the term “conventional”¹⁶ we mean the

Received: January 13, 2014

Revised: February 16, 2014

Published: February 21, 2014



interatomic contacts that obey the geometrical criteria for H bonding.^{13,16,17} We consider the contact $H\cdots B$ in a fragment $A-H\cdots B$ ($A = N, O$; $B = N, O, F, S, Cl$) as an H bond if the conditions $d(H\cdots B) \leq 2.5 \text{ \AA}$ for $B = N, O, F$ or $d(H\cdots B) \leq 2.7 \text{ \AA}$ for $B = S, Cl$; $d(A\cdots B) \leq 3.5 \text{ \AA}$ for $B = N, O, F$ or $d(A\cdots B) \leq 3.7 \text{ \AA}$ for $B = S, Cl$; $\angle A-H\cdots B \geq 120^\circ$ hold. Note that, if required, the H bonds can be determined with other methods such as empirical force field, quantum-mechanical, or experimental analysis of electron density; the next steps of our algorithm will be the same in all cases. For simplicity reasons, we have not considered other intermolecular interactions such as van der Waals, $\pi-\pi$, or halogen–halogen, but being determined they can be easily included into our scheme in the same way as H bonds.

In the next step we generate the underlying net of the structure with an appropriate simplification procedure. It allows us to obtain the information about crystal structure topology as a network whose nodes and edges correspond to molecules and chosen intermolecular bonds. The simplification procedure includes the representation of a molecule by its center of mass, keeping the connectivity of the molecules with each other by means of H bonds; all H bonds between a given pair of molecules transform to the same edge between the molecular centers of mass in the simplified net. Such a representation of the structure is called *standard*. On this level of the structure representation, the local topology of molecular associates is described by a *coordination figure*, which just shows the adjacency and space arrangement of molecular centers of mass (Figure 1). This simplification procedure is

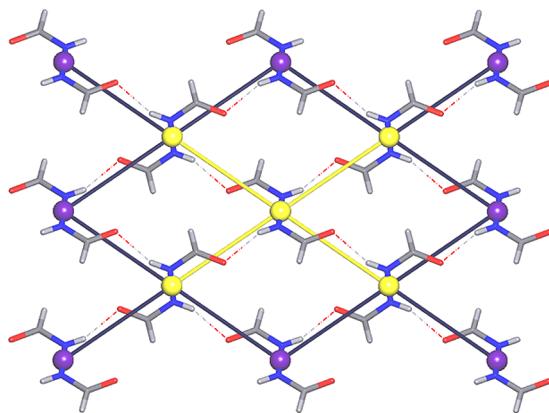


Figure 1. Simplification procedure for H-bonded molecules in diformylhydrazine (FOMHAZ13).¹⁹ The underlying net is formed by centers of mass of the molecules (balls) and edges corresponding to H bonds between the molecules. The coordination figure of a node of the underlying net is highlighted by yellow.

implemented in the program package TOPOS¹⁸ as a standard for molecular crystals; all topological analyses in this study were performed with TOPOS.

In the next step we determine the topological type of the underlying net. In this work, we use three-letter symbols of the Reticular Chemistry Structure Resource (RCSR) notation²⁰ or Fischer and Koch's symbols for 1- or 2-periodic sphere packings.²¹ Those nets that are absent in RCSR are designated with the TOPOS ND_n nomenclature,³ where N is a sequence of coordination numbers of

all nonequivalent nodes of the net, D is periodicity of the net ($D = M$ (molecular), C (chain), L (layer), T (three-periodic) for 0-, 1-, 2-, 3-periodic nets), and n is the ordinal number of the net in the set of all nonisomorphic nets with the given ND sequence.

To describe the local mutual arrangement of molecules, we proposed¹⁵ to generalize the notation that was initially designed for coordination compounds.²² Each molecule (L) is designated by letters $M, B, T, K, P, G, H, O, N, D$ (see the Supporting Information for encoding of the letters) depending on the number $n = 1-10$ of its atoms (both donors and acceptors) involved in the formation of intermolecular H bonds (we call them *active centers*). The atoms that form intramolecular H bonds are not counted, because these atoms do not affect the connectivity of molecules and, hence, the overall topology of the structure. When the molecule has more than 10 active centers, the sign $X[n]$ is used. The total number of molecules connected to a given one is listed as the upper index in the form of the line $mbtkphond...$, where each integer m, b, t, k, \dots is equal to the number of molecules connected by one, two, three, four, ... H bonds. The sum $MCN = m+b+t+k+\dots$ is equal to the number of molecules connected by H bonds with the central one, i.e., to its molecular coordination number (MCN). Moreover, the number of H bonds formed by a given molecule can be computed as $NB = m \cdot 1 + b \cdot 2 + t \cdot 3 + k \cdot 4 + \dots$. Finally, the molecular connection type symbol (MCTS) is written as $L^{mbtkphond}$. For example, the MCTS notation K^{02} for fumaric acid^{23,24} (Figure 2) means that the molecule has four active centers ($L = K$): two H bond donors (atoms H1) and two H bond acceptors (atoms O1), and it is bonded to two molecules by two H bonds in each case ($m = 0, b = 2$); $NB = b \cdot 2 = 4$.

The proposed notation allows one to characterize the local arrangement of molecules in the cases of any complexity. For example, 5-bis(hydroxyimino)-3-phenyl-1,2-oxazole molecule (OBOBOW)²⁵ has five active centers (atoms H1, H2, O1, O2, and N1); therefore, $L = P$ (Figure 3). Molecule 1 is bonded to two molecules 3 and 4 each by one H bond ($m = 2$) and to molecule 2 by four H bonds ($k = 1$); $NB = 2 \cdot 1 + 0 \cdot 2 + 0 \cdot 3 + 1 \cdot 4 = 6$. As a result, the MCTS of this molecule is P^{2001} .

MCTS is an example of a new topological parameter that can be used to find correlations between local topological features of molecules, topology of their supramolecular ensembles, and the overall topology of a molecular crystal as a whole. Indeed, MCTS is a more detailed level of description of intermolecular bonding compared to a molecular coordination figure; it directly relates to the chemical structure of the molecule through the notion of the active center. As a result, the sequence of relations “chemical structure of molecule—number of active centers—molecular connection type—coordination figure—overall topology” can serve as a predictable algorithm for creating an expert system that deals with possible types of molecular packings and connection motifs. The general scheme of such a system is similar to that we proposed for coordination networks.⁴ It rests upon the knowledge database that includes the following electronic collections:

- collection of topological types of molecules (TTM collection) that contains the information on molecular composition, structure, coordination figure, connection type;
- collection of the corresponding overall topologies (TTD collection);

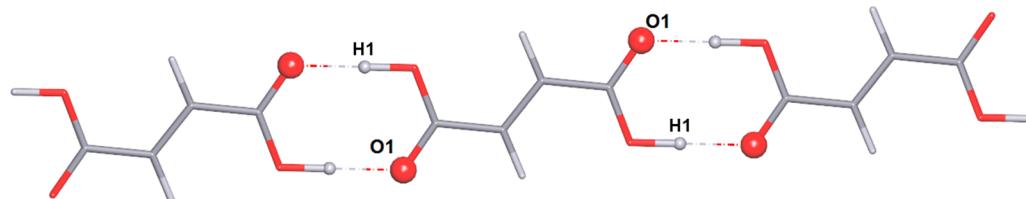


Figure 2. Connection type of a molecule of fumaric acid (K^{02}) both in monoclinic (FUMAAC) and triclinic (FUMAAC01) polymorphs. Hereafter the CSD reference codes of the compounds are given in parentheses.

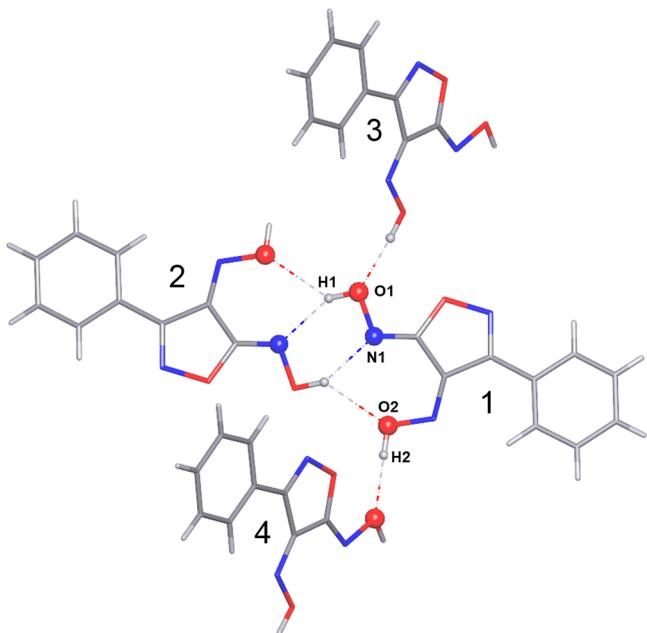


Figure 3. Connection type P^{2001} of a 4,5-bis(hydroxyimino)-3-phenyl-1,2-oxazole molecule (OBOWBO).

- collection of occurrences of the molecules and overall topologies in molecular crystals (TTO collection);
- collection of relations between different overall topologies (TTR collection) that shows how one topology can be transformed to another;
- collection of statements that show the probability of fulfillment of the relations “chemical structure of molecule—number of active centers—molecular connection type—coordination figure—overall topology”; this collection is obtained by an automated analysis of other collections.

Let us consider how the knowledge database is being filled. First, with the TOPOS software, the TTM, TTD, TTO, and TTR collections are updated with the last update of the crystallographic databases accumulating the experimental data on molecular crystals obtained all over the world. At present, the most comprehensive set of data is gathered in the CSD. TOPOS allows one to get the necessary topological information in a proper time; thousands of structures can be processed in several hours with an ordinary desktop computer. Two TOPOS applied programs are used to extract the topological information from crystallographic data. With the program AutoCN, both intramolecular (valence and nonvalence) and intermolecular (H-bond, specific halogen–halogen, and van der Waals) contacts are determined. The program ADS separates molecules, determines MCTS, simplifies the structure by squeezing the molecules into their centers of mass, constructs the underlying net, and determines its topology by assigning it to a topological type.

Second, the statistical data on local and overall topological parameters are collected. The distributions on MCTSs, molecular crystal formula, molecular coordination number, coordination figure, and type of underlying net are constructed and relations between them are established.

The relations found are then stored in the knowledge database and can be used by the inference machine of the expert system to predict possible motifs. Below we will see how this scheme works for H-bonded molecular crystals.

2.2. Objects. We have analyzed homomolecular organic crystals, i.e., containing molecular units of the same kind, crystallographic data for which were taken from the CSD (release 5.34, November 2012). All interatomic bonds, including H bonds, were identified with the TOPOS AutoCN method of intersecting sectors²⁶ improved for determination of intermolecular interactions.²⁷ Entries containing incomplete, erroneous data and disordered structures as well as those

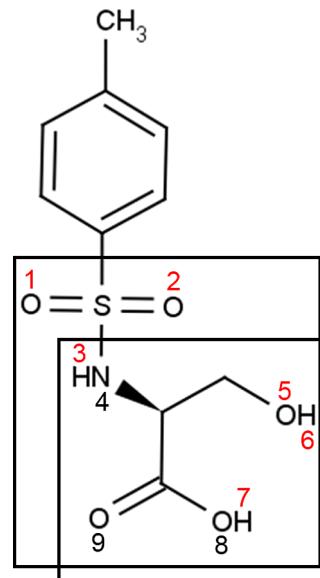
with $R_f > 10\%$ were excluded. For all the structures, we have applied the standard simplification described above and obtained underlying nets, whose nodes and edges represent molecules and H bonding between them. The periodicity of the resulting networks was determined with TOPOS; then we separately analyzed 0- (molecular), 1- (chain), 2- (layer), and 3-periodic nets (frameworks), paying special attention to the 2-periodic ones. For all molecules, their MCTSs were determined by means of a specially tailored TOPOS procedure. Total numbers of the structures studied depending on their periodicity are given in Table 1; the whole list of structures is presented in the Supporting Information.

Table 1. Total Numbers of Structures Studied

periodicity of H-bond patterns	0-periodic	1-periodic	2-periodic	3-periodic
structures with crystallographically equivalent molecules	8478	20235	5194	2657
structures with crystallographically nonequivalent molecules	2339	3267	858	386
total number of structures	10817	23502	6052	3043

2.3. Synthesis and Structure Determination of *N*-(4-Methylbenzene)sulfonyl]serine. To check the approach proposed in this paper, we have synthesized molecular crystals of *N*-(4-methylbenzene)sulfonyl]serine (3-hydroxy-2-(toluene-4-sulfonylamino)propionic acid) (**I**) that has at most nine active centers being able to participate in H bonds (Scheme 1). The chemicals used

Scheme 1. Molecule of I and Its Nine Possible Active Centers^a



^aThe red numbers enumerate the active centers that participate in H bonds in the crystal structure of **I**. Large and small rectangles show *N*-sulfonylserine and L-serine residues, respectively.

were purchased from internationally reputable suppliers and used without further purification. L-Leucine (9.5 mmol, 1 g) was dissolved in water using 1 M solution of Na₂CO₃, and the pH was adjusted to 8–9. 4-Toluenesulfonyl chloride (9.5 mmol, 1.81 g) was added to the above solution. Reaction progress was observed by consumption of suspended 4-toluenesulfonyl chloride to a clear solution. Then pH was adjusted to 2–3 using dilute HCl. The precipitate observed was filtered, washed, and recrystallized in methanol.

A suitable crystal of **I** was selected and mounted, using a glass fiber fixed to a copper pin held on a magnetic base. An Agilent SuperNova (Dual source) Agilent Technologies diffractometer, equipped with a

graphite-monochromatic Cu/Mo K α radiation was used to collect the data. The data collection was performed using CrysAlisPro software²⁷ at 296 K using Cu K α radiation. The structure solution was performed by direct methods using SHELXS-97²⁸ and refined by full-matrix least-squares methods on F² using SHELXL-97,²⁸ operating under the X-Seed.²⁹ All non-hydrogen atoms were refined anisotropically by full-matrix least-squares methods.²⁸ All the C–H hydrogen atoms were positioned geometrically and treated as riding atoms with C–H = 0.93 Å, 0.96 Å, 0.97 Å, and 0.98 Å for aromatic, methyl, chiral carbon, and methylene H atoms, respectively. They were refined using a riding model with U_{iso}(H) = 1.5 U_{eq}(C) for methyl, and U_{iso}(H) = 1.2 U_{eq}(C) for all other carbon atoms. The N–H = 0.76 Å and O–H = 0.82 Å hydrogen atoms were located with difference Fourier maps and refined using the riding model with U_{iso}(H) = 1.2 U_{eq}(N) and U_{iso}(H) = 1.5 U_{eq}(O).

3. RESULTS AND DISCUSSION

3.1. Distributions of Topological Parameters. The distributions of topological parameters and the correlations between these parameters play the key role in the creation of the knowledge database, and they are the primary subject for discussion. Below we present the occurrences for the most abundant cases in the 2-periodic H-bond patterns; all the distributions are listed in the Supporting Information. In this part, the structures with all symmetry-equivalent molecules in 5194 structures (Table 1) are considered; the peculiarities of the packings of nonequivalent molecules in 858 structures (Table 1) will be discussed in a separate part. In Table 2 and Figure 4, the occurrences of connection types of molecules and MCN are given, respectively.

Table 2. Distribution of the Most Abundant Connection Types in 2-Periodic H-Bonded Motifs

MCTS	number of structures	percentage, %
K ⁴	2046	39.4
K ²¹	1118	21.5
G ²²	286	5.5
T ²¹	244	4.7
T ⁴	162	3.1
p ²²	162	3.1
G ⁶	142	2.7
P ⁶	92	1.8
p ⁰³	82	1.6
p ⁴¹	82	1.6
O ⁴²	72	1.4
G ⁴¹	70	1.4
O ⁰⁴	70	1.4
G ⁰³	53	1.0
G ⁰⁴	50	1.0
Others	463	8.9

The distribution shows that the connection types K⁴ and K²¹ (Figure 5) are the most common for 2-periodic H-bond patterns. In general, the most abundant are those connection types, where the number of active centers coincides with the number of H bonds formed by the molecule that conforms with other statistical data on H bonds.³⁰ This is true for K⁴, K²¹, G²², G⁶, and other types from Table 2, for which L is determined by NB, i.e., two-center H bonds are most common.

The distribution of overall topologies of the H-bonded motifs (Table 3) shows that the leaders (square plane net **sql** and honeycomb net **hcb**) are the same as in 2-periodic coordination polymers,⁴ while the third place is occupied by

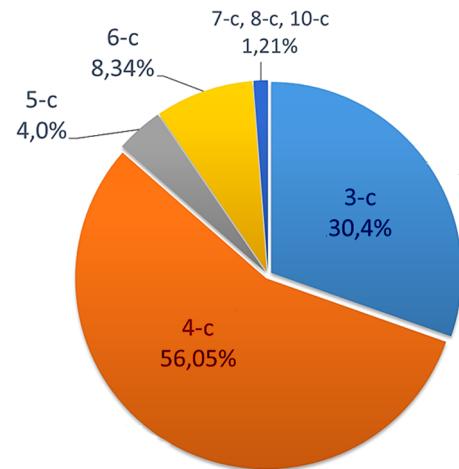


Figure 4. Distribution of coordination numbers of molecules.

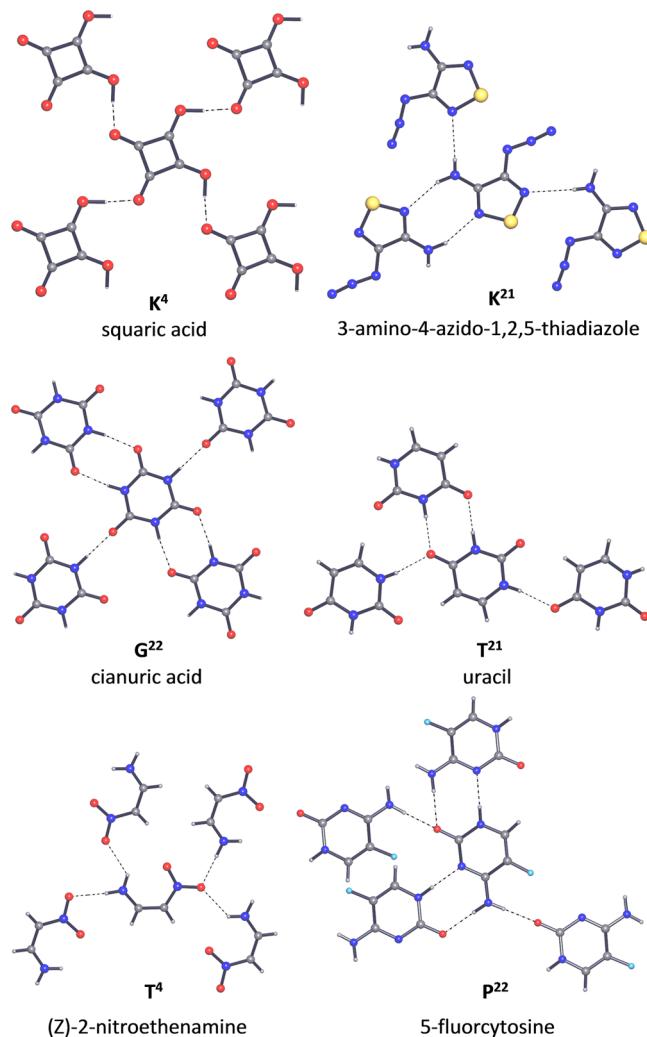


Figure 5. Examples of molecules with the most widespread connection types.

hexagonal close packing **hxl**, which is rather rare in coordination polymers. This occurrence/distribution is due to the tendency of the molecules to form close packing in the layer, where each molecule adjoins six other molecules, but MCN is usually less than six because of the lack of active centers or their improper arrangement in the molecule. The

Table 3. Distribution of Overall Topologies of 2-Periodic H-Bonded Motifs

overall topology	number of structures	percentage, %
sql	2814	54.2
hcb	1556	29.9
hxl	244	4.7
4 ⁴ Ila	87	1.7
4 ⁴ Ia	87	1.7
4L2	59	1.1
4 ⁴ Ilb	52	1.0
cem	49	0.9
tts	47	0.9
others	199	3.8

most frequent MCNs are four and three that lead to the **sql** and **hcb** topologies (Table 3, Figure 6) as derived from **hxl** after breaking some contacts (Figure 7).

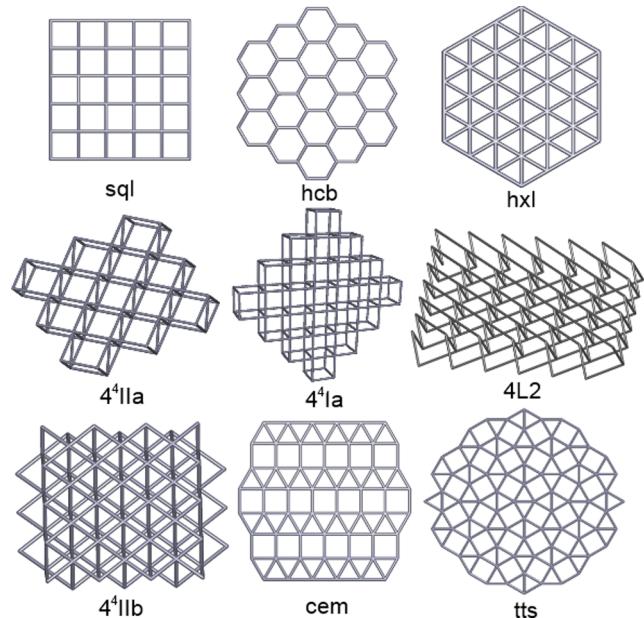


Figure 6. Most frequent H-bonded motifs.

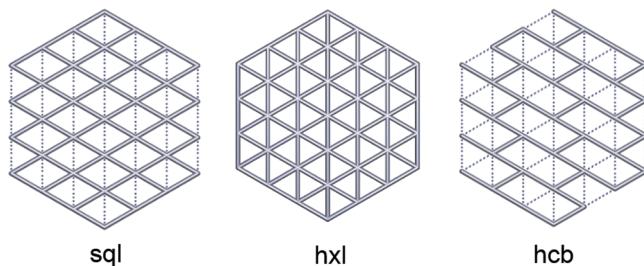


Figure 7. Relation between **sql**, **hcb**, and **hxl**. The links to be broken in **hxl** are shown by dotted lines.

3.2. Correlations between Local and Overall Topological Parameters of H-Bond Patterns. The next step for creation of the knowledge database is to find the correlations between different topological parameters, first of all between parameters characterizing the local environment of molecules (MCN, MCTS, coordination figure) and those that describe the whole system of H bonds, in particular, dimensionality and topological type of the underlying net. As we will see in this

part, such correlations are the strongest and easiest to reveal because the corresponding parameters are well-defined and formalized.

Knowing MCTS, we can predict periodicity of the underlying net with high probability. For instance, connection types K²¹ and T²¹ are ordinary for chain or layer H-bonded networks and very rarely occur in molecular or 3-periodic motifs (Table 4). Instead, many other connection types, such as G⁶ or P⁶, can lead to 2- or 3-periodic patterns but not to molecular or chain ensembles.

Table 4. Distribution of MCTS in H-Bonded Motifs with Different Periodicity

MCTS	number of structures	percentage (%) for periodicity			
		0	1	2	3
K ⁴	3134	0	16.2	65.3	18.5
K ²¹	1821	0.05	36	61.4	2.5
G ²²	439	0	12.3	65.1	22.6
T ²¹	402	0	38.3	60.7	1
T ⁴	279	0.36	20.8	58.1	20.8
P ²²	268	0	14.9	60.4	24.6
G ⁶	417	0	0.2	34.1	65.7
P ⁶	230	0	0.4	40	59.6
P ⁰³	115	0	27	71.3	1.7
P ⁴¹	138	0	0	59.4	40.6
O ⁴²	194	0	0	37.1	62.9
G ⁴¹	125	0	0	56	44
O ⁰⁴	93	0	3.2	75.3	21.5
G ⁰³	65	0	13.8	81.5	4.6
G ⁰⁴	66	0	1.5	75.8	22.7

For the nets of the same dimensionality, MCTS allows one to predict overall topology, in some cases with a 100% probability. For example, if a molecule has four active centers and forms bonds with four other molecules, i.e., it has connection type K⁴, the resulting underlying net will have the **sql** topology with a 96.5% probability (Table 5).

Table 6 contains the information on the reverse correlation between the underlying net type and MCTS at a given MCN. For the most typical MCN = 4, the number of existing overall topologies is essentially limited; in most cases, a **sql** net is realized. One of the reasons for such a high occurrence is that the **sql** topology is compatible with a large number (34) of connection types, where the number of active centers varies in a wide range from 3 to 24 (Table S2, Supporting Information). The next topology in the list, 4L2, fits only to four different MCTSs. However, in all cases, there is one preferable connection type such as K⁴ for **sql** that is evidently the most suitable for a particular overall topology.

At the same time, Tables 5 and 6 show that even if a correlation between topological parameters is strong, there can exist exceptions, such as 4L7 or 4L12 topologies for connection type K⁴, or there can be several equiprobable choices such as 4⁴Ilb and **hxl** for P⁶ (Table 5), or many MCTSs with a single occurrence (Table S3, Supporting Information). Obviously, there are some other factors that influence the topological parameters. Thus, the next step in creation of the knowledge database is to find the correlations between topological parameters and other factors, first of all, chemical properties of the molecules.

3.3. Correlations between Topological Parameters and Other Properties of Molecules. The following

Table 5. Possible Topologies of Underlying Nets at Some MCTSs of H-Bonded Molecules in 2-Periodic Motifs

MCTS	topological type	number of structures	percentage, %
K ⁴	sql	1975	96.5
	4L2	48	2.4
	4L1	19	0.9
	kgm	2	0.1
	4L12	1	0.05
	4L7	1	0.05
K ²¹	hcb	1099	98.3
	KIa	17	1.5
	fes	2	0.2
G ²²	sql	277	96.9
	4L2	4	1.4
	4L1	3	1.0
	6 ³ Ib	2	0.7
P ²²	sql	154	95.0
	4L2	3	1.9
	6 ³ Id	3	1.9
	4L1	2	1.2
G ⁶	hxI	99	69.7
	4 ⁴ Ila	26	18.3
	others (7)	17	12.0
	4 ⁴ Iib	33	35.8
P ⁶	hxI	28	30.4
	4 ⁴ Ila	12	13.0
	6L7	10	10.8
	Others (4)	9	9.6

Table 6. Possible Connection Types of H-Bonded Molecules at Some Topologies of Underlying Nets in 2-Periodic Motifs

topological type	MCTS	number of structures	percentage, %
sql	K ⁴	1975	70.2
	G ²²	277	9.8
	T ⁴	156	5.5
	P ²²	154	5.5
	O ⁰⁴	69	2.5
	G ⁰⁴	50	1.8
	K ²²	31	1.1
	others (27)	102	3.6
4L2	K ⁴	48	81.4
	G ²²	4	6.8
	T ⁴	4	6.8
	P ²²	3	5.1
4L1	K ⁴	19	70.4
	G ²²	3	11.1
	P ²²	2	7.4
	T ⁴	2	7.4
	H ⁰⁴	1	3.7
6 ³ Ib	G ²²	2	66.7
	H ³⁰⁰¹	1	33.3
6 ³ Id	P ²²	3	100
kgm	K ⁴	2	100
4L12	K ⁴	1	100
4L13	O ⁰⁴	1	100
4L7	K ⁴	1	100

additional molecular properties influence the overall topology of an H-bonded motif at a given connection type of the molecules:

(i) Deviation in mutual arrangement of active centers in a homologous series can lead to a regular change of coordination

figure and hence overall topology of the H-bonded motif. A striking example is the homologous series of α,ω -diols with the common formula HO-(CH₂)_n-OH, where n varies from 4 to 24. Even members of the series (with even n) form a 2-periodic sql net, while odd members form a 3-periodic diamondoid (dia) net. In odd diols, one of hydroxyl groups goes out of plane and adopts a gauche position. The authors³¹ mentioned that even diols have flat molecules, but starting from $n = 16$, polymorphic modifications with nonplanar molecules emerge. Molecular packing in the structures of odd diols and even diols with nonplanar molecules is the same. The difference in topological types arises from alternation in the orientation of the hydroxyl group that results in differences of coordination figures for flat and nonplanar molecules or to differences in mutual orientation of coordination figures for nonplanar molecules of even and odd diols (Figure 8).

(ii) Steric hindrances in molecular packing caused by voluminous atomic groups of the molecule can give rise to very rare topologies for the structures under consideration, such as the kagome (kgm) pattern in the structure of 2,5-di-*tert*-butylhydroquinone (HESKOF01) or 2,5-bis(trimethylsilyl)-hydroquinone³³ (NECFUW; Figure 9), where voluminous *tert*-butyl or trimethylsilyl groups meet in the hexagonal holes of the pattern. At the same time, many other substituted hydroquinones with smaller substituents possess a sql topology with smaller square holes.

(iii) Environment and physical conditions during the nucleation and growth of a crystal. Minor changes in the composition of a solution lead to different dimensionalities of structures, though the connection type remains the same as observed in two polymorphs of 5-fluorocytosine. The monoclinic form (MEBQEQQ) was discovered by crystallization from 5% aqueous 2-propanol, while the tetragonal form (MEBQEQQ01) crystallizes when the method of vapor diffusion of ether into solution in 2-propanol was applied.³⁴ The difference in topologies arises from different chain stacking modes (Figure 10).

The main factors influencing the connection type of a molecule are as follows:

(i) Number of active centers in the molecule. This parameter directly determines the letter L in the MCTS notation. At the same time, the molecule cannot use all active centers and hence can have different connection types even in the same structure, although our analysis shows that molecules tend to use all active centers in H bonding, if their number is not too large (<4–5). For example, the nicotinamide molecule in most cases involves all four possible active centers in H-bonding (Figure 11). Also the ratio between H bond donor and acceptor active centers should be taken into account; in most cases, this ratio is equal to 1:1.

(ii) Arrangement of active centers in the molecule. In general, the closer the active centers are disposed, the more probable are MCTSs with zero first numbers (m, b) and a low-periodic H-bonded motif. Two derivatives of uracil with different positions for the methyl group^{35,36} (METURA04, WUDVVAS) elucidate such a correlation (Figure 12).

(iii) Positions and size of the molecular parts that do not participate in H bonding. Each such part plays a role of spacer; it can shield the active centers and prevent formation of H bonds. A series of derivatives of urea give examples of such a behavior. Starting from *N*-methylurea³⁷ (MEUREA) with a 3-periodic dia underlying net, one can come through an hcb net

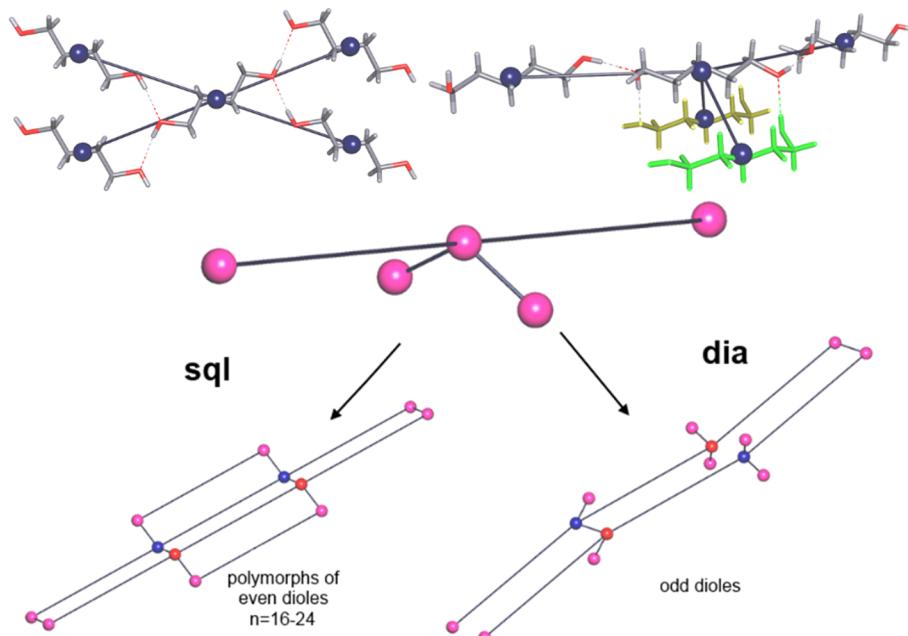


Figure 8. Coordination figures in H-bond patterns $\text{HO}-(\text{CH}_2)_4-\text{OH}$ (QATTIO,³² top left) and $\text{HO}-(\text{CH}_2)_5-\text{OH}$ (QATTOU,³² top right) and different orientation of the coordination figures in even and odd diols (bottom).

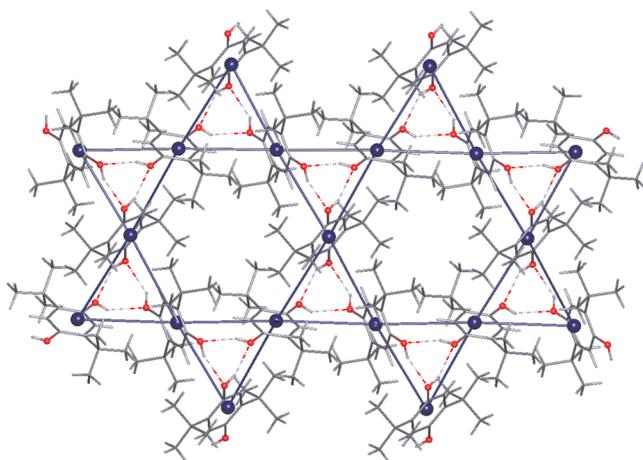


Figure 9. The kagome pattern in the structure of 2,5-di-*tert*-butylhydroquinone (HESKOF01).

in *N*-ethylurea³⁸ (YAQLAE) to a **sql** net in *N*-butylurea³⁸ (YAQLEI) (Figure 13).

(iv) Distribution of the electron density and electrostatic potential over the molecule. The influence of the distribution of electron density in a molecule can be shown in the case of two related molecules: cyanuric acid³⁹ (CYURAC03) and its sulfur analogue, trithiacyanuric acid⁴⁰ (CEHQEM). In both cases, each molecule is connected to two other molecules by two H bonds via amide and thioamide synthons. The difference appears in the mode of connection to other molecules. Oxygen in cyanuric acid is able to form a H bond with an angle C–O...H close to 180° , because of presence of negative charge on its pole. Instead, sulfur in trithiacyanuric acid has toroidal distribution of negative charge that leads to quite different geometry of H bonds. As a result, cyanuric and trithiacyanuric acids have different connection types, G²² and P⁰³, respectively (Figure 14), that gives rise to different underlying nets **sql** and **hcb** (see below).

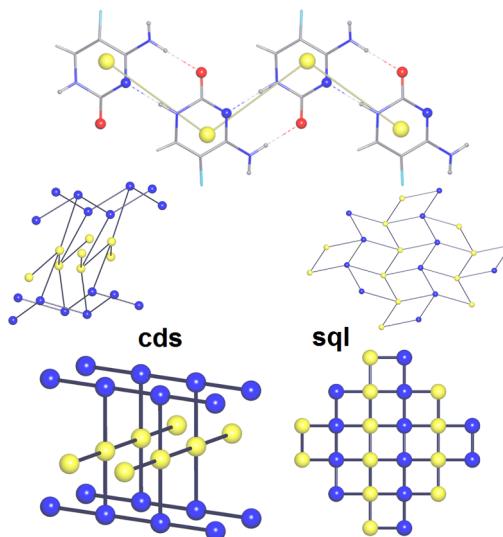


Figure 10. A chain of 5-fluorocytosine molecules (top; active centers of molecules are shown as balls) and different stacking modes of the chains (center) in the polymorph structures resulting in different overall topologies: 3-periodic **cds** and 2-periodic **sql** for tetragonal and monoclinic polymorphs, respectively. Idealized **cds** and **sql** nets are shown on the bottom.

(v) Physical conditions and environment. In particular, solvent can influence the formation of a particular connection type. For example, 5-fluoururacil has two polymorphs⁴¹ with different connection types, K²¹ or K⁰², for triclinic (FURACL01) and monoclinic (FURACL03) modification, respectively (Figure 15). The former type leads to a 2-periodic **fes** topology, while the latter one causes a 1-periodic chain. The triclinic form can be obtained from a variety of solvents, while the monoclinic one was obtained only from dry nitromethane, which is a solvent with very low donor activity. Likely, this property of the solvent results in different methods of

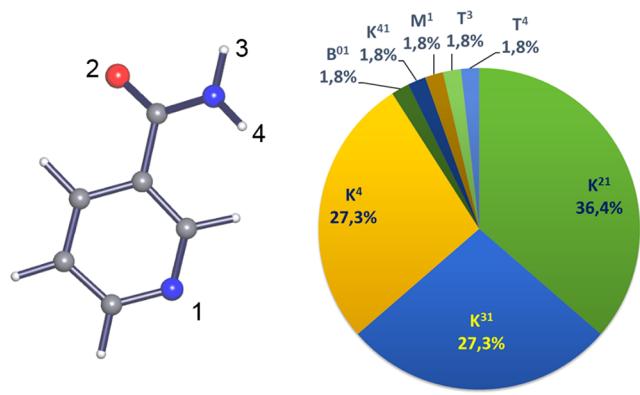


Figure 11. (Left) The nicotinamide molecule and its four possible active centers; (right) distribution of connection types of the nicotinamide molecule in its polymorphs, hydrates, or cocrystals. Totally, 49 structures containing 55 crystallographically different nicotinamide molecules were retrieved from the CSD.

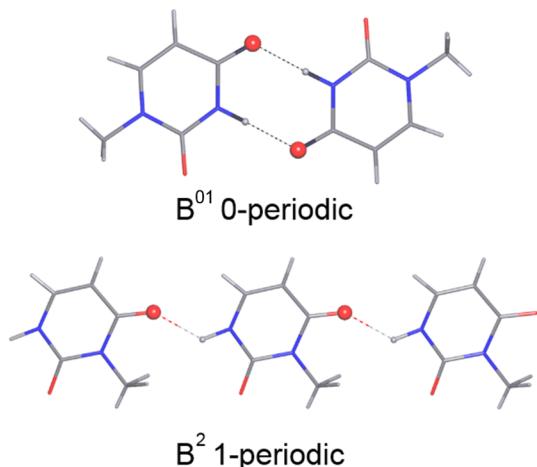


Figure 12. Molecules of 1-methyluracil (METURA04, top) and 3-methyluracil (WUDVAS, bottom) with MCTSs B^{01} and B^2 , respectively.

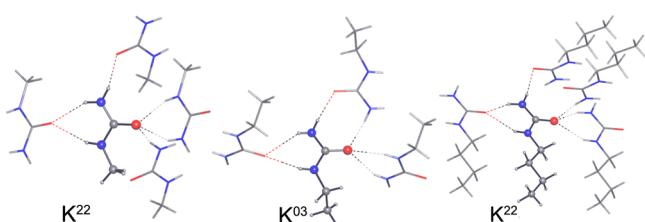


Figure 13. Influence of the volume of alkyl groups on the connection type of molecules: *N*-methylurea (MEUREA, left), *N*-ethylurea (YAQLAE, middle), *N*-butylurea (YALQLEI, right). Number of active centers involved in formation of H-bonds remains the same ($L = K$).

connection of dimers, which can be considered as building units of both H-bond patterns.

3.4. Topological Properties of Supramolecular Ensembles. When one can separate several molecules united in a supramolecular ensemble in the crystal structure, there are different ways to describe the local and overall topology. MCTS can help to find such ensembles and to select the way of simplification and description of the whole structure. For example, connection types such as $L^{mb..1}$ with the last unity number can be considered corresponding to molecular dimers.

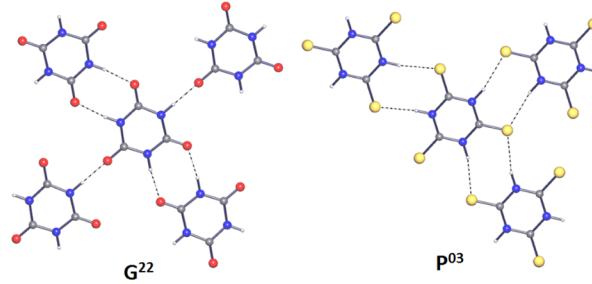


Figure 14. Connection types of cyanuric (CYURAC03; G^{22}) and trithiacyanuric (CEHQEM; P^{03}) acids.

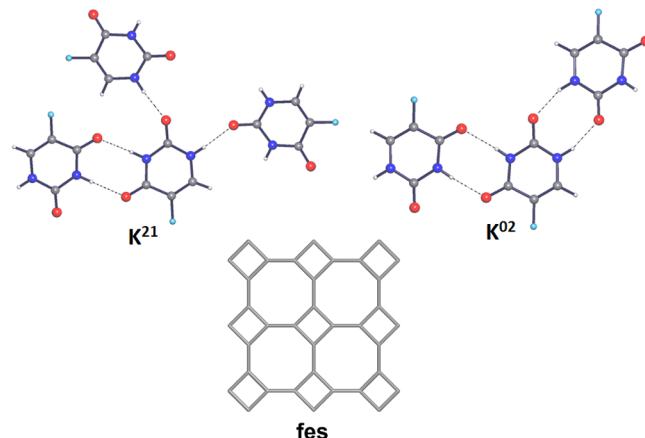


Figure 15. (Top) Connection types of molecules of 5-fluorouracil in its triclinic (FURACL01; K^{21}) and monoclinic (FURACL03; K^{02}) modifications; (bottom) a network with the **fes** topology.

The topologies of H-bond patterns as well as connection types of separate molecules and those of the corresponding supramolecular ensembles are strictly interrelated. For example, if the connection type of a molecule looks like $L^{200...01}$, when only first and last numbers are not zero, one may select dimeric ensembles that have connection type K^4 . Simultaneously, the **hcb** underlying net of molecules transforms to the **sql** underlying net of dimers. Thus, formamide has two polymorphs^{42,43} (Figure 16); one of them (FORMAM02) contains T^{21} -type connected molecules, i.e., belonging to the $L^{mb..1}$ kind; hence, the structure can be considered as an **sql** net of dimers. The other polymorph (FORMAM03) is assembled of molecules connected by types T^{21} and T^4 . The H-bond pattern of molecules has an unusual 3,4L28 topology, but if we separate dimers of the T^{21} -type connected molecules and consider the structures as composed of molecules T^4 and the dimers with MCTS K^4 , we get the **kgm** topology (Figure 16). Another striking example is the structure of one of the polymorphs of 2-thiobarbituric acid⁴⁴ (THBARB03), which contains differently organized H-bonded motifs. In this case, there are alternating layers of K^{21} and K^4 -type connected molecules (Figure 17) that can be considered as **sql** motifs of dimers and separate molecules, respectively. In our opinion, these examples indicate that dimers can build the structure together with separate molecules.

3.5. Structures with Crystallographically Nonequivalent Molecules. The question why some monomolecular structures contain nonequivalent molecules remains a puzzle. Simplified nets of such structures may contain nodes of different topological type, i.e., their underlying nets can be

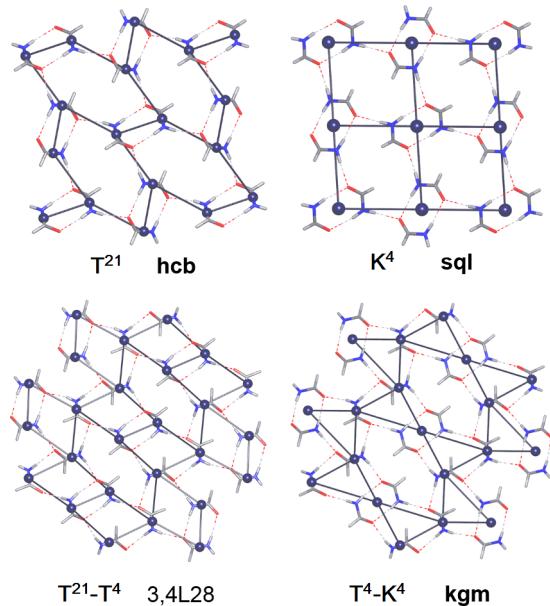


Figure 16. H-bond patterns in two formamide polymorphs FORMAM02 (top) and FORMAM03 (bottom) and different representations of their structures considering separate molecules and dimers. MCTSs refer to the building units under consideration.

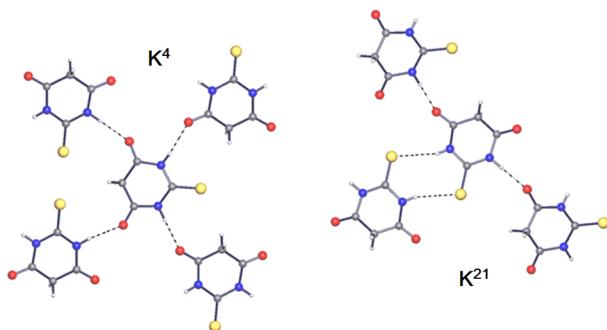


Figure 17. MCTSs K^4 and K^{21} in molecular layers of 2-thiobarbituric acid (THBARB03) with **sql** and **hcb** topologies, respectively.

polynodal. Nonetheless, uninodal nets prevail over polynodal ones and cover 76.2% of the structures under consideration (Table 7). In 10.3% of the structures, there are molecules with bridging connection types such as B^2 or T^{11} (Table 8), which

Table 7. Distribution of Overall Topologies of Underlying Nets in the Structures with Nonequivalent Molecules

overall topology	number of structures	percentage, %
hcb	291	33.7
sql	267	30.9
fes	32	3.7
tts	15	1.7
hxl	14	1.6
KIa	14	1.6
SL7	13	1.5
cem	12	1.4
4L1	12	1.4
4⁴Ia	10	1.2
4⁴Ila	9	1.0
4,4L15	9	1.0
others	166	19.2

Table 8. Distribution of the Most Abundant Connection Types in 2-Periodic H-Bonded Motifs with Several Nonequivalent Molecules

MCTS	number of structures	percentage, %
K^4	279	22.1
K^{21}	243	19.2
T^3	72	5.7
T^{21}	56	4.4
G^{22}	41	3.2
B^2	39	3.1
P^{31}	37	2.9
T^{11}	34	2.7
P^{22}	29	2.3
P^{41}	22	1.7
G^{41}	21	1.7
G^6	18	1.4
others	373	29.5

extend the net, i.e., they play a role of spacers and do not influence the topology of an underlying net. The first two leaders remain the same but in an inverse order. The third place is unexpectedly occupied by the fes net, which was found only in two structures with equivalent molecules. The most abundant binodal net 4,4L15 was found only in 1% of compounds; thus, molecules despite their geometrical non-equivalence tend to play the same topological role in the structure. Only six compounds with two different 2-periodic underlying nets, i.e., two topologically nonequivalent layers of H-bonded molecules coexisting in the same structure, were found.

3.6. A Route to an Expert System in Supramolecular Chemistry. We have seen above that many correlations between local and overall topological parameters are very strong and provide predictable statements. For example, one can state “if a molecule has four active centers, the most probable connection type will be K^4 that in 96.5% of cases results in an **sql** overall topology” or “with the assumption that a molecule forms two H-bonds with every neighbor molecule, which results in, for example, P^{03} or G^{04} connection types, the overall topology can be determined unambiguously with 100% probability and the resulting topologies will be **hcb** and **sql**, respectively”. Such statements can form the foundation of a knowledge database that can be used to produce expert conclusions about the possibility of obtaining a particular topological supramolecular motif. Note that expert systems for H-bond prediction are already being created for separate classes of organic crystals.⁴⁵

Within our approach, a prototype of the knowledge database is the set of distributions of topological parameters and the set of relations between these distributions gathered in Tables 1–8. A special computer program, the inference machine, will ask the user for initial data that at least should include the structure of a molecule and then, after analysis of the knowledge database, will form an expert conclusion. The sequence of steps of the inference machine from input of the initial data to the conclusion about the underlying topology of the supramolecular ensemble is shown on Figure 18. Note that this sequence is appropriate for the task under consideration; for other tasks of supramolecular design, some other steps could be required.

To illustrate the steps of the inference machine, let us apply it to structure I that can be considered as a random example. First

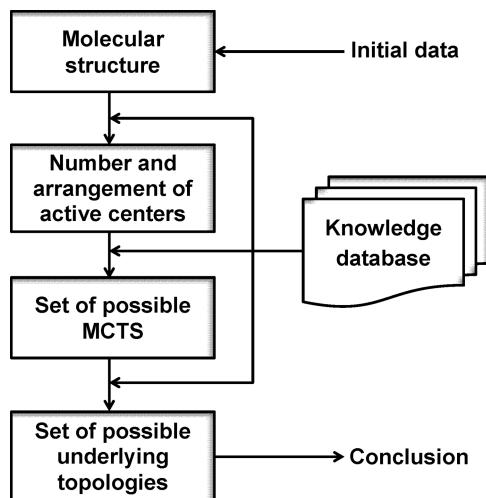


Figure 18. The sequence of steps of the inference machine from initial data provided by the user to the resulting conclusion about the topological motif of the supramolecular ensemble.

we have to get the information about the number of active centers and their arrangement in the molecule. Totally, in the *N*-sulfonylserine residue of **I**, there are nine atoms that can be involved in H bonding (Scheme 1) but obviously not all of them play the role of active centers in H-bond patterns. To enrich the knowledge database with the typical maps of the active centers of the *N*-sulfonylserine residue, we have screened the CSD for the molecules containing this residue. There were no such molecules, but we have retrieved 11 structures containing a fragment of L-serine (Scheme 1, Table 9). In

Table 9. Active Centers of the L-Serine Fragment Participating in H Bonding

molecule	CSD reference code	numbers of active centers (see Scheme 1)
L-serinium	EYOQOY	3 5 6 7
N-(6-amino-3,4-dihydro-3-methyl-5-nitro-4-oxopyrimidin-2-yl)serine	FUQKUX	6 7 9
L-serinium	JEJVAW	3 5 6 7 9
D-serinium	JEJVEA	3 5 6 7 9
L-serinium	KERMID	3 6 7 9
D-serinium, L-serinium	MIFQUN	3 5 6 7
D-serinium, D-serine	NELRAX	3 5 6 7
N-dodecanoyl-L-serine	RESHOM	3 5 6 7
D-serinium, L-serinium	RIWKUD01	3 5 6 7
L-serinium	TIXGOW	3 5 6 7
L-serinium	XATVUK	3 6 7 9

eight cases out of 11, only four out of seven active centers of the fragment were involved in H bonding. To expand this information to the *N*-sulfonylserine residue, we assumed that both oxygen atoms of the sulfonyl group participate in H bonds. This assumption is supported by the fact that the numbers of donor and acceptor active centers should be equal in a monomolecular structure (see part 3.3); in this case, the ratio is 3:3 = 1:1. Hence, there are totally six active centers and the molecule should have MCTS of a G type; this is the first step of the inference machine. In the next step, the analysis of the knowledge database (Table 2) shows that the G^{22} connection type is the most abundant among all G types.

The last step of the inference machine is the analysis of relations between MCTS and underlying topologies (Table 5) that predicts the **sql** supramolecular motif for the G^{22} connection type with a 97% probability. Indeed, we have the G^{22} connection type and the **sql** undulated underlying net in the crystal structure of **I** (Figure 19).

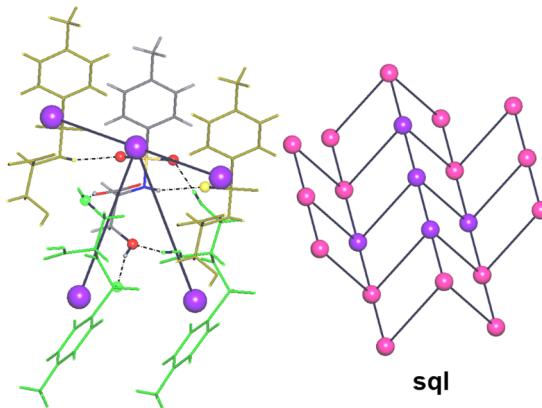


Figure 19. The G^{22} connection type and coordination figure of molecules as well as a **sql** underlying net in the crystal structure of **I**. The molecules connected by one or two H bonds to the central one are colored in yellow and green, respectively.

A rather weak point in the reasoning outlined above is the assumption of a G type of molecular connection. Although it was supported by the analysis of similar L-serine fragments and the requirement of an equal numbers of donor and acceptor active centers, strictly speaking, we need more experimental data to enrich the knowledge database and to make the conclusion with a higher probability. In some cases, this knowledge gap can be filled in with quantum-mechanical or other modeling methods. However, it is important that the final conclusion does not depend on this assumption. Indeed, if we consider other possible connection types, where only one or none of the sulfonyl oxygen atoms participate in H bonds, we get the most probable MCTSs P^{22} or K^4 , respectively (Table 2); both of them with a high probability lead to the same **sql** supramolecular motif (Table 5).

The next point is that from the very beginning of our reasoning we implicitly assumed a 2-periodic symmetry of the supramolecular motif. Indeed, most of the information in our knowledge database is so far restricted to 2-periodic extended structures. However, Table 4 shows that although the G^{22} connection type occurs in 1-, 2-, and 3-periodic H-bonded supramolecular architectures, the 2-periodic ones are most typical. In general, the same algorithm (Figure 18) should be applied to 1- and 3-periodic structures and some 1- and 3-periodic motifs should be added to the list of possible underlying nets. We have done it for the G connection types (see below, Tables 10 and 11).

The last point concerns other possible G connection types. Except for G^{22} , there can be realized G^6 , G^{41} , G^{03} , or G^{04} connection types, which with a high probability lead to **hxl**, **4⁴a**, **hcb**, or **sql** supramolecular motifs, respectively (Table 5, and Table S2, Supporting Information). One can compute the probability (P) of occurrence of a particular motif in H-bond patterns as a product of probabilities taken from the knowledge database:

Table 10. Probabilities of Realization of a Particular G Connection Type in H-Bonded Motifs of Different Periodicity

MCTS	number of structures	$p(\text{MCTS})$	p (periodicity)			
			0 (molecular)	1 (chain)	2 (layer)	3 (framework)
G^{22}	439	0.3948	0	0.1230	0.6515	0.2255
G^6	417	0.3750	0	0.0024	0.3405	0.6571
G^{41}	125	0.1124	0	0	0.5600	0.4400
G^{04}	66	0.0594	0	0.0152	0.7576	0.2273
G^{03}	65	0.0585	0	0.1385	0.8154	0.0462

Table 11. Probabilities of Realization of Underlying Topologies for a Particular G Connection Type

MCTS	underlying topologies their periodicities (after slash) and probabilities (P)					
G^{22}	$\text{sql}/2$	$\text{dia}/3$	$3^6(1,2)/1$	$\text{neb}/3$	$4\text{L}2/2$	
	0.2491	0.0665	0.0477	0.0045	0.0036	
G^6	$\text{hxl}/2$	$\text{pcu}/3$	$\text{sxd}/3$	$\text{lcy}/3$	$\text{acs}/3$	$4^4\text{Ia}/2$
	0.0890	0.0674	0.0468	0.0459	0.0297	0.0234
G^{41}	$4^4\text{Ia}/2$	$\text{sqp}/3$	$\text{tts}/2$	$\text{bnn}/3$	$\text{cem}/2$	$\text{nov}/3$
	0.0288	0.0171	0.0171	0.0153	0.0110	0.0090
G^{04}	$\text{sql}/2$	$\text{dia}/3$				
	0.0450	0.0090				
G^{03}	$\text{hcb}/2$	$4^4(0,2)/1$	$\text{lig}/3$			
	0.0477	0.0081	0.0018			

$$P = p(\text{MCTS})p(\text{periodicity})p(\text{underlying topology}) \quad (1)$$

where $p(\text{MCTS})$, $p(\text{periodicity})$, and $p(\text{underlying topology})$ are probabilities of occurrence of a particular G-type MCTS, a particular periodicity at a given MCTS, and a particular underlying topology at given MCTS and periodicity. The first two probabilities are given in Table 10; the last one can be taken from Table 5 or from similar data for other periodicities. For example, the probability of occurrence of a **sql** motif for the G^{22} connection type can be calculated according to eq 1 as $P = 0.3948 \times 0.6515 \times 0.9685 = 0.2491$. The P values for various underlying nets are given in Table 11, from which one can see that the **sql** motif is realized for G^{22} and G^{04} connection types with the total probability of $0.2491 + 0.0450 = 0.2941$ to be the highest among the probabilities for other overall topologies. In particular, the next possible underlying motifs are **hxl** and **dia** with $P = 0.089$ and 0.0755 , respectively.

The resulting list of the possible underlying topologies is not so long because in many cases different connection types (local topologies) give rise to the same topological type of the whole motif. The **sql** underlying net predominates in this list (it has the highest probability) that agrees with the experimental H-bond pattern in I.

4. CONCLUSIONS

In the exploration and prediction of supramolecular architectures, precise quantum-mechanical methods and qualitative topological approaches play complementary and equally important roles. At the same time, if the role of the modeling methods is now recognized by almost all chemists, the topological approaches are still not so widespread. One could ask why do we need such kinds of prediction as was presented above, if we make conclusions just with some probability? Why not simply apply molecular modeling to get a stricter answer? Such questions ignore the fact that those two groups of methods have different destinations. Working with all available experimental information that can include also the results of structure modeling, the topological methods can solve the following tasks:

- to find general regularities in the experimental data and hence to formulate new chemical laws;
- to generate the questions that require a closer inspection by other methods, both experimental and theoretical;
- to provide a quick preliminary conclusion about possible supramolecular architectures that could bypass expensive quantum-mechanical or experimental treatment.

The expert system built in accordance with the principles described above in fact mimics the work of a human expert. Indeed, when the expert is asked "what is a possible connection type of this molecule?" or "what is a possible overall motif of the supramolecular ensemble assembled from these molecules?", he just answers "my experience says that we have these possible choices and most likely that one will be realized". Here the expert experience and logic correspond to the knowledge database and inference machine of the electronic expert system, but unlike the human expert, the knowledge database accumulates all experimental information and can make a much more precise conclusion. Being unable to replace the human completely, such a system could become his indispensable theoretical tool along with the modeling methods.

ASSOCIATED CONTENT

S Supporting Information

A PDF file with experimental data for compound I (CCDC reference number 975469). An Excel file with Tables S1–S11 that contain the distributions of connection types of molecules and overall topologies of all studied 2-periodic H-bond patterns (Table S1), the correlations between the connection types of molecules and the corresponding overall topologies (Tables S2 and S3), the underlying topologies for all studied 0-, 1-, 2-, and 3-periodic H-bonded motifs in crystal structures with one nonequivalent molecule (Tables S4, S6, S8, S10) and with several nonequivalent molecules (Tables S5, S7, S9, S11). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Phone: +7-8463345445. Fax: +7-8463345417. E-mail: blatov@samsu.ru.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant No. (12-130-35-HiCi). The authors, therefore, acknowledge with thanks DSR technical and financial support. P.N.Z. thanks Russian government (grant No. 14.B25.31.0005) and Russian Foundation for Basic Research (grant No. 13-07-00001) for support. P.N.Z. and V.A.B. are grateful to Prof. Davide M. Proserpio for fruitful discussion of the work.

REFERENCES

- (1) *Modern Methods of Crystal Structure Prediction*; Oganov, A. R., Ed.; Wiley-VCH: Weinheim, 2011.
- (2) Gavezzotti, A. *Top Curr. Chem.* **2012**, *315*, 1.
- (3) Alexandrov, E. V.; Blatov, V. A.; Kochetkov, A. V.; Proserpio, D. M. *CrystEngComm* **2011**, *13*, 3947.
- (4) Mitina, T. G.; Blatov, V. A. *Cryst. Growth Des.* **2013**, *13*, 1655.
- (5) Allen, F. H. *Acta Crystallogr.* **2002**, *B58*, 380.
- (6) Allmann, R.; Hinek, R. *Acta Crystallogr.* **2007**, *A63*, 412.
- (7) Villars, P.; Cenzual, K. *Pearson's Crystal Data: Crystal Structure Database for Inorganic Compounds*; ASM International: Materials Park, OH, 2013.
- (8) Öhrström, L.; Larsson, K. *Molecule-Based Materials: The Structural Network Approach*; Elsevier: Amsterdam, 2005.
- (9) Carlucci, L.; Ciani, G.; Proserpio, D. M.; Champness, N. In *Making Crystals by Design: Methods, Techniques and Applications*; Braga, D.; Grepioni, F., Eds.; Wiley-VCH: Weinheim, 2007; Ch. 1.3 and Ch. 2.4.
- (10) Batten, S. R.; Neville, S. M.; Turner, D. R. *Coordination Polymers: Design, Analysis and Application*; Royal Society of Chemistry: Cambridge, 2009.
- (11) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120.
- (12) Grell, J.; Bernstein, J.; Tinhofen, G. *Cryst. Rev.* **2002**, *8*, 1.
- (13) Baburin, I. A.; Blatov, V. A. *Acta Crystallogr.* **2007**, *B63*, 791.
- (14) Blatov, V. A.; O'Keeffe, M.; Proserpio, D. M. *CrystEngComm* **2010**, *12*, 44.
- (15) Aman, F.; Asiri, A. M.; Siddiqui, W. A.; Arshad, M. N.; Ashraf, A.; Zakharov, N. S.; Blatov, V. A. *CrystEngComm* **2014**, *16*, 1963.
- (16) Gilli, G.; Gilli, P. *The Nature of the Hydrogen Bond: Outline of a Comprehensive Hydrogen Bond Theory*; Oxford University Press: New York, 2009.
- (17) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48.
- (18) Blatov, V. A. *Struct. Chem.* **2012**, *23*, 955 <http://www.topos.samsu.ru>.
- (19) Hope, H.; Ottersen, T. *Acta Crystallogr.* **1978**, *B34*, 3623.
- (20) O'Keeffe, M.; Peskov, M. A.; Ramsden, S. J.; Yaghi, O. M. *Acc. Chem. Res.* **2008**, *41*, 1782.
- (21) Koch, E.; Fischer, W. Z. *Kristallogr.* **1978**, *148*, 107.
- (22) Serezhkin, V. N.; Vologzhanina, A. V.; Serezhkina, L. B.; Smirnova, E. S.; Grachova, E. V.; Ostrova, P. V.; Antipin, M. Y. *Acta Crystallogr.* **2009**, *B65*, 45.
- (23) Brown, C. J. *Acta Crystallogr.* **1966**, *21*, 1.
- (24) Bednowitz, A. L.; Post, B. *Acta Crystallogr.* **1966**, *21*, 566.
- (25) Tselinsky, I. V.; Melnikova, S. F.; Pirogov, S. V.; Shaposhnikov, S. D.; Nather, C.; Traulsen, T.; Friedrichsen, W. *Heterocycl. Commun.* **2000**, *6*, 35.
- (26) Peresypkina, E. V.; Blatov, V. A. *Acta Crystallogr.* **2000**, *B56*, 1035.
- (27) Agilent. CrysAlis PRO. Agilent Technologies, Yarnton, England, 2012.
- (28) Sheldrick, G. M. *Acta Crystallogr.* **2008**, *A64*, 112.
- (29) Barbour, L. J. *J. Supramol. Chem.* **2001**, *1*, 189.
- (30) Infantes, L.; Motherwell, W. D. S. Z. *Kristallogr.* **2005**, *220*, 333.
- (31) Kenjiro, U.; Yoshihiro, O.; Naotake, N. *Cryst. Growth Des.* **2008**, *8*, 592.
- (32) Thalladi, V. R.; Boese, R.; Weiss, H.-C. *Angew. Chem., Int. Ed.* **2000**, *39*, 918.
- (33) Bock, H.; Nick, S.; Nather, C.; Bats, J. W. *J. Prakt. Chem.-Chem. Ztg.* **1996**, *338*, 363.
- (34) Hulme, A. T.; Tocher, D. A. *Cryst. Growth Des.* **2006**, *6*, 481.
- (35) Micklitz, W.; Lippert, B.; Schollhorn, H.; Thewalt, U. J. *Heterocycl. Chem.* **1989**, *26*, 1499.
- (36) Portalone, G.; Ballirano, P.; Maras, A. *J. Mol. Struct.* **2002**, *608*, 35.
- (37) Huiszoon, C.; Tiemessen, G. W. M. *Acta Crystallogr.* **1976**, *B32*, 1604.
- (38) Hashimoto, M.; Tajima, F.; Eda, K.; Yamamura, K.; Okazaki, T. *J. Mol. Struct.* **2005**, *734*, 23.
- (39) Verschoor, G. C.; Keulen, E. *Acta Crystallogr.* **1971**, *B27*, 134.
- (40) Guo, F.; Cheung, E. Y.; Harris, K. D. M.; Pedireddi, V. R. *Cryst. Growth Des.* **2006**, *6*, 846.
- (41) Hulme, A. T.; Price, S. L.; Tocher, D. A. *J. Am. Chem. Soc.* **2005**, *127*, 1116.
- (42) Stevens, E. D. *Acta Crystallogr.* **1978**, *B34*, 544.
- (43) Gajda, R.; Katrusiak, A. *Cryst. Growth Des.* **2011**, *11*, 4768.
- (44) Chierotti, M. R.; Ferrero, L.; Garino, N.; Gobetto, R.; Pellegrino, L.; Braga, D.; Grepioni, F.; Maini, L. *Chem.—Eur. J.* **2010**, *16*, 4347.
- (45) Delori, A.; Galek, P. T. A.; Pidcock, E.; Patniac, M.; Jones, W. *CrystEngComm* **2013**, *15*, 2916.