CrystEngComm

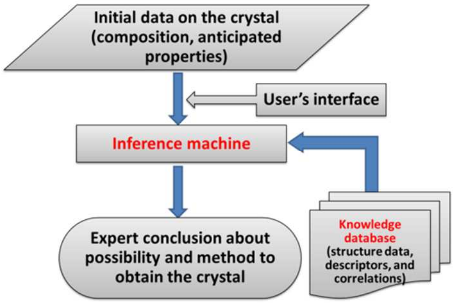
Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/crystengcomm

Table of contents entry

The problem of predicting crystal structures is discussed in the context of artificial intelligence systems.



Journal Name RSCPublishing

ARTICLE

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

New Knowledge and Tools for Crystal Design: Local Coordination *versus* Overall Network Topology and Much More

Eugeny V. Alexandrov,[a] Alexander P. Shevchenko,[a] Abdullah A. Asiri,[b,c] and Vladislav A. Blatov*[a],

The problem of predicting crystal structures is discussed in the context of artificial intelligence systems. The steps of creation of an expert system are considered as applied to crystal design, where the crucial step is invention of new structure descriptors. A number of such descriptors proposed quite recently are listed; most of them characterize local coordination or overall topology of the structure network. An important part of the expert system is the knowledge database that contains correlations between the descriptors; it is used by a computer analyzer, the inference machine, to make a conclusion about possibility of obtaining a particular structure motif. All the steps of developing the expert system are illustrated with the analysis of 811 cyanide complexes and examples of the structure prediction are given.

Introduction

At present, expert systems are the most typical implementation of artificial intelligence; they are used in various areas of science and technology. Banks apply artificial intelligence systems in the insurance business (Actuarial Mathematics), in the games on exchange and management of the property, *etc.*[1] Specialized expert systems are widely used in optical and acoustic recognition (including text and speech), medical diagnostics, spam filters, air defense systems, as well as for solving a number of other problems in national security.[2]

An important feature of expert systems is their ability to self-learning that is associated with the continuous process of knowledge discovery and data mining. That is why the first component of expert system to be created is the *knowledge database*. In general, the process of developing knowledge database consists of the following steps:

- data sampling to select the information that is relevant to the problem;
- data cleaning to remove errors and unreliable information from the sample;
- data transformation to translate the information to a computer language;
- data mining to perform a statistical analysis that includes searching for association rules as well as clustering, classification and regression of the data;
- interpretation and evaluation of the data mining results to extract correlations (knowledge).

At the next step one can try to generate a new knowledge resting upon the correlations and rules from the knowledge database. For this purpose, a data analyzer, the so-called *inference machine*, should be created.

In chemistry, both general knowledge databases and more special datasets as well as full-functioned expert systems have been created in the near past.[3] However, despite long history of developing computer-based knowledge systems in crystallography and crystal chemistry,[4] no full-fledged inference machines have been created that could provide an expert conclusion on the possibility and methods of manufacturing new crystalline materials. At the same time, quantum mechanical packages have emerged, which inherently provide some predictive options of expert systems for solid state design.[5] Obviously, dramatically growing structural information has reached the level, at which its systematization and rationalization become crucial, and expert systems could be an important tool to solve these tasks. In this paper, we outline a possible route to develop expert systems in structural science and provide some examples of how it could work.

Artificial intelligence in crystal design: the problem statement

Huge amount of experimental crystallographic data, most of which are obtained with a high precision, should contain many general laws or more specific regularities that could be used in crystal structure prediction. Following Motherwell,[6] one can say that the crystallographic databases contain a lot of answers, but the problem is to ask correct questions. Expert systems should solve this problem. They should be able to extract the knowledge from the experimental data and apply it for elucidation of the existing crystal structures as well as for prediction of new ones.

When a human expert has been asked what coordination polymer can be obtained from a given set of complexing metal atoms and ligands, or what molecular packing can be formed by a given kind of molecule, he first asks about details, then addresses to his memory and handbooks for the supplementary information and finally to his experience for delivery of the recommendation. He would never say anything definitely, but will use phrases like ‘I suppose…’, ‘most likely…’, *etc*.

The expert system follows the same way (Fig. 1). It is a computer program that asks the user about the experiment details, recognizes structure descriptors in the information, and uses the correlations from the built-in knowledge database to output a conclusion by means of a special tool, the inference machine. The advantages compared to the human are that the number of correlations can be as large as needed, and the conclusion is made with a definite probability. Of course, the expert system cannot replace the human but can essentially help him in making decision.

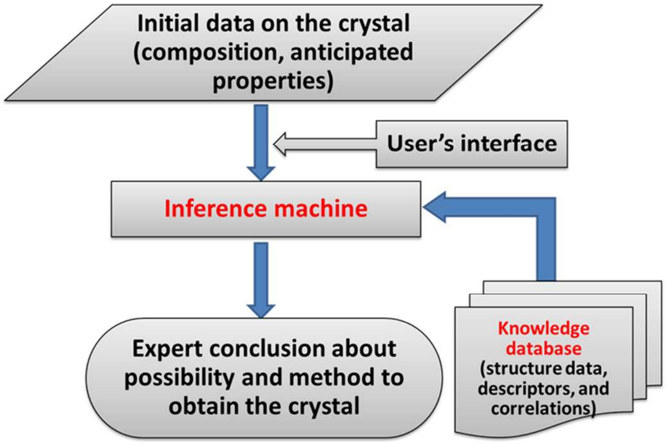


Figure 1. General scheme of the expert system.

Below we consider the main steps on the route to the expert system in crystal design. Some of them have been already passed and hence will be examplified, while others can be merely outlined.

Steps to create an expert system

In this part, we consider how to create an expert system for crystal design. The key steps of this procedure are shown in Fig. 2. The examples considered below mainly concern analysis of topological properties of a crystal structure because it is the field where sufficient progress has been achieved in the last years. Nonetheless the general scheme remains the same for prediction of any type of structural parameters.

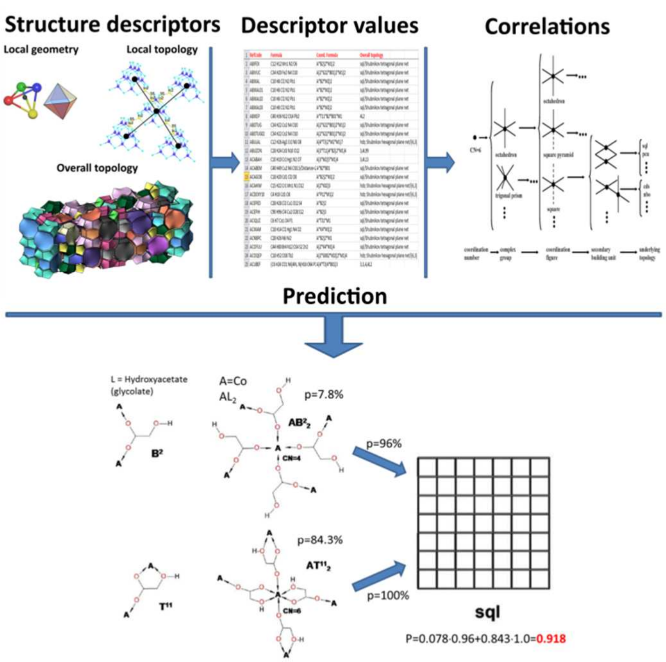


Figure 2. The main steps in creation of an expert system in crystal design.

Inventing new structure descriptors

Any structure correlation exists between some parameters that can either be numerical or in general have a symbolic expression. Any parameter can be considered as a value of some structure descriptor; that is why the more diverse the set of robust descriptors is built, the more possibilities exist to find important correlations between the parameters. By the word ‘robust’ we mean that the descriptor should adequately (*i.e.* acceptably at a given precision level) reflect some structure property and should be adjusted for the computer analysis. Any descriptor can be derived either from experimentally available data or from concepts proved by mathematical modeling. Obviously, the former group of descriptors is the most reliable, large and well known, although the *ab initio* descriptors become more and more popular thanks to the progress of quantum mechanical and other computational methods in the last 20 years.

Probably the first well known and crystallochemically inherent descriptor was *atomic radius*; many efforts in the 20th century were undertaken to find the best system of atomic radii to be applicable to different classes of compounds. Another famous characteristic, *bond strength*, also has almost century-long history. These examples show how difficult is to find a robust universal descriptor.

The first structure descriptors like the mentioned above characterized the local environment of atoms and allowed to find many stereochemical correlations that have formed the basis of modern crystal chemistry. At the same time, local parameters cannot unambiguously characterize the whole structure architecture; to fit the requirements of crystal design one needs a set of global descriptors that are related to overall structure features. It is noteworthy that again a long time passed since the pioneer works,[7] where the overall structure topology had been considered, until the first global descriptors were proposed. Wells introduced into crystal chemistry the notion of periodic net and even used the descriptors[7b] that are now called *point* symbols[8] to characterize the net topology, but he did not propose any method for strict discrimination of nets. Almost in the same time, Kitaigorodskii[9] attracted the attention to the topology of molecular packings, although characterizing them by *molecular coordination number*, *i.e.* by a local descriptor. Packings of atoms as well as ion arrays[10] were explored since the very beginning of the crystallographic era, but again the main classification descriptors were *atomic coordination number* and *coordination polyhedron*. As a result, the nets as well as atomic and molecular packings were ordinarily discriminated by visual analysis until the end of the 20th century. The visual analysis being subjective often resulted in mistakes; moreover, complicated topologies could not be recognized in this way.

In the 2000s, new descriptors[11] were invented to characterize the overall network topology that immediately led to a series of reviews devoted to classification of coordination networks,[11b,12] molecular packings,[13] and ion arrays.[14] An important tendency should be noted: the descriptors become less human-oriented and more adjusted for computer processing. Thus such network topology characteristics as *Systre key*[11a] or *vertex symbol*[8] cannot be used manually for any more or less complicated structure due to their intricacy.

Invention of new local and overall topological descriptors is the crucial problem for the task under consideration. Any novel robust descriptor allows us to find a lot of new correlations. For example, recently,[15] we successfully used the *ligand coordination mode symbol*, *coordination formula*, and *molecular connection type symbol* as efficient descriptors of local connection of ligands or molecules. Another recently proposed local descriptor, *coordination figure*,[12,16] still needs formalization as at the moment it is applied only in the visual analysis. An important further step in this way could be invention of the descriptors that include information on the chemical composition of the entire network as well as of separate structural units.

Creating a database of descriptor values (parameters)

After a set of robust descriptors is determined, the next task is to compute the parameters (*i.e.* values of the descriptors) for all available crystallographic data. The *completeness* of the analysis is important: it means that no possible exceptions will be lost. This is the most time-consuming step: unfortunately, crystallographic databases contain many errors, incompletely determined or highly disordered structures that should be excluded from the analysis or amended in some way. A method for automated extraction of reliable data is required that is directly related to the previous task: robust structure should have all parameters lying within an appropriate range. Such approach is already used in the widespread *checkCIF* procedure,[17] which however does not consider any topological parameters.

The values computed should be stored in a database that has to be flexible enough to deposit any kind of descriptor: number, string, array, graph, *etc*. Design of such universal storage is a separate task that also requires an extension of the crystallographic format (CIF) to include topological parameters into it. In the 2000s, the databases of topological parameters like RCSR[11c] or TOPOS collections[18] emerged. Recently, we have begun to create such databases[15a,c] using the Excel table format and even this simple approach allowed us to solve some tasks on the next level.

Automated search for correlations and forming the knowledge database

The database of structure parameters can be used to search for correlations between the parameters and, hence, to rise to the next level of the data organization. The resulting set of correlations can be considered as a knowledge database; the universal storage mentioned in the previous part could be a technical solution of the problem of developing such database. We have already proposed prototypes of such knowledge databases for 2-periodic coordination polymers[15a] and molecular crystals.[15c] The simplest format of a record (*statement*) in the knowledge database looks like 'value Z1 of descriptor 1 provides value Z2 of descriptor 2 with probability *P*'. A more complicated form can include other descriptors that influence the 1 – 2 correlation: 'value Z1 of descriptor 1 provides value Z2 of descriptor 2 with probability *P* if descriptors 3, 4, 5, … have values Z3, Z4, Z5, …, respectively'. In this case, *P* has the meaning of conditional probability. To be reliable the statements should be based on rather large samples. Two standard criteria can be applied for this purpose: *support* (*S*), which is the part of the bins used for establishing the correlation with respect to the size of the whole sample, and *confidence* (*C*) corresponding to the minimal probability, at which the correlation is worthy to be written to the database. In particular, this means that the statement can usually include no more than 3-4 descriptors since the larger the number of descriptors is, the smaller the S and C values are.

Developing the inference machine

The knowledge database itself cannot provide an effective prediction of possible structures. If it is rather large, the researcher can directly use it only as an electronic manual since he is not able to comprehend all its content. That is why a special interface, the inference machine, is required, which would process the user's request, transform it into the format of the universal storage and use the knowledge accumulated to issue an expert conclusion.

Recently, we presented several examples of how the inference machine can work for coordination polymers[15a] or molecular crystals.[15c] We used the following descriptors: (i) ligand coordination mode symbol or molecular connection type symbol, (ii) coordination formula, (iii) molecular and atomic coordination number, (iv) coordination figure, (v) overall network topology. Significant correlations between the first four local topological descriptors and the last one composed a knowledge database. The test examples, also with specially synthesized ‘random’ compounds showed that the inference machine can successfully predict the topological properties of these classes of compounds resting upon the initial data on their chemical composition.

How it works: more examples

To illustrate the main ideas of developing the expert system approach we extracted from the Cambridge Structural Database[4a] (CSD, version 5.35) and Inorganic Crystal Structure Database[4b] (ICSD, release 2014/1), which had totally contained about 850,000 records, the comprehensive list of crystal structures of coordination compounds with only one type of ligand, namely, cyanide anion CN^-. The choice was caused by a high popularity and large structural diversity of the cyanide complexes, broad variety of their physical properties as well as their importance as potential functional materials.[19] All steps of the analysis were performed with the *ToposPro* program package.[18]

Data sampling and cleaning

We have searched for the relevant information following the step sequence presented in Table 1. The resulting sample contains the crystallographic information on the crystal structures of 811 cyanide-containing complex compounds, including three structures of zinc dicyanoaurates found in the literature.[20]

Table 1. Algorithm for selection of crystallographic information on crystal structure of cyanide-containing complex compounds

Step	Filtering conditions	Number of entries
I	The name of the compound must contain the “cyano” or ”cyanid” fragments	23 110
II	Metal atoms must exist	14 738
III	Metal atoms must coordinate C or N atoms. Coordination of alkali and alkaline earth metals is ignored except simple cyanides	2050
IV	Metal atoms coordinate only cyanide ions. Structures with disordered atoms influencing the network topology (hydrogen atoms are not considered) and duplicated structure determinations are removed.	811

Descriptors and their values

On the next step we study geometrical and topological properties of the 811 crystal structures and for this purpose we have to determine the list of robust descriptors. We analyze the compounds on six levels of their structure organization (Table 2).

Most of the descriptors from Table 2 are either well-known in crystal chemistry or described above. Maximal size of the network ring corresponds to a particular method of selecting secondary building units (‘clusters’) in the crystal structure;[21] this is one of the strict topological criteria for separating polynuclear atomic groups, which are implemented into *ToposPro*. The overall topology of the network was classified with a set of topological invariants in accordance with the reference topological types stored in the *ToposPro TTD* collection.[18] In general, a preliminary simplification of the network is required: the structural groups (ligands, polynuclear complex groups, clusters) are replaced by their centers of mass and the resulting *underlying net*[21] represents the method of connection of the groups in the initial network. The values of all these descriptors for all 811 crystal structures were automatically calculated and stored in a database.

Table 2. Levels and descriptors of the structure organization to be analyzed

Level	Descriptors
Local coordination of metal atom	Chemical sort (*Me*), coordination number (*Cn*), oxidation state (*OxS*), composition (*CCp*) and type (*Cp*) of coordination polyhedron of the metal atom
Mutual coordination of metal atoms and ligands	Coordination mode symbol of the ligand, coordination figure of the metal atom, coordination formula (*CF*)
Topology of secondary building unit	Composition, nuclearity, dimensionality, maximal ring size corresponding to the building unit
Overall topology of coordination network	Point symbol, vertex symbol, coordination sequence for all atoms in the network
Topology of entangled motifs	Types of the entanglement (*Ent*)
Influence of extraframework species	Existence or absence of the extraframework atoms or molecules

Search for correlations between parameters

To extract knowledge from the set of parameters a variety of one-dimensional or multi-dimensional distributions should be analyzed. Below we consider some important correlations, which can be retrieved from the set of parameters computed at the previous step. We should remember that strictly speaking the number of parameters in all correlations is larger than is mentioned: all the distributions are built with the conditions from Table 1.

LOCAL COORDINATION OF METAL ATOM. Table 3 shows the occurrence of metal atoms (*Me*), their coordination numbers (*Cn*), type (*Cp*) and composition (*CCp*) of their coordination polyhedra. These four one-dimensional distributions allow us to explore 12 pairwise correlations, namely, *Me-Cn*, *Cn-Me*, *Me-Cp*, *Cp-Me*, *Me-CCp*, *CCp-Me*, *Cn-Cp*, *Cp-Cn*, *Cn-CCp*, *CCp-Cn*, *Cp-CCp*, *CCp-Cp*. In the correlation sequences, for simplicity we do not specify oxidation state (*OxS*) as a separate descriptor and include it into the *Me* descriptor if needed. The number of triplewise correlations is even larger (24): for each of the four triples *Me-Cn-Cp*, *Me-Cn-CCp*, *Me-Cp-CCp*, and *Cn-Cp-CCp* all six permutations should be considered. At last, there are 4! = 24 quadruple-wise correlations. Here we face with a crucial problem of the process of creating the knowledge database. Which of these correlations should be analyzed and which of them should be stored in the knowledge database? The answers are different for the human expert and the computer.

The human expert will analyze only those correspondences that seem reasonable and most interesting; he rests upon his experience. In particular, he will not consider the correlation

Cp-Cn because it is trivial: coordination polyhedron determines coordination number unambiguously as the number of its vertices. Another trivial correlation is *CCp-Cn* because the number of atoms in coordination polyhedron is equal to coordination number.

Table 3. One-dimensional distributions for metal atoms (*Me*), their coordination numbers (*Cn*), type (*Cp*) and composition (*CCp*) of their coordination polyhedral

Parameter	*N*	*w*, %	Parameter	*N*	*w*, %
Me			*CCp*		
Fe	206	16.72	C_6	354	28.73
Cu	159	12.91	C_4	330	26.79
Cd	122	9.90	C_2	185	15.02
Pt	120	9.74	N_6	116	9.42
Au	118	9.58	N_4	43	3.49
Ag	93	7.55	C_2N	37	3.00
Co	88	7.14	C_8	37	3.00
Ni	71	5.76	CN_2	19	1.54
Mn	48	3.90	C_3	18	1.46
Zn	42	3.41	C_3N	14	1.14
Cr	39	3.17	C_2N_2	13	1.06
Mo	22	1.79	CN_3	11	0.89
Pd	21	1.71	N_3	10	0.81
W	19	1.54	CN	10	0.81
Hg	18	1.46	C_7	8	0.65
Rh	8	0.65	C_3Ni	6	0.49
Ru	8	0.65	C_3N_2	5	0.41
V	4	0.33	C_5	4	0.33
Re	4	0.33	N_2	3	0.24
Os	4	0.33	C_5Co	2	0.16
Li	3	0.24	N_5	2	0.16
Tl	2	0.16	N_2C	1	0.08
Ir	2	0.16	C_2N_4	1	0.08
Nb	2	0.16	C_4Mo	1	0.08
Be	1	0.08	C_4N	1	0.08
Mg	1	0.08	C_5N	1	0.08
Al	1	0.08	*Cn*		
Sn	1	0.08	6	474	38.47
Er	1	0.08	4	417	33.85
In	1	0.08	2	199	16.15
Ga	1	0.08	3	84	6.82
Ti	1	0.08	8	37	3.00
Lu	1	0.08	5	13	1.06
			7	8	0.65
			Cp		
			octahedron	470	38.15
			square	227	18.43
			linear	199	16.15
			tetrahedron	190	15.42
			triangular	84	6.82
			square antiprism	27	2.19
			trigonal dodecahedron	10	0.81
			square pyramid	8	0.65
			pentagonal bipyramid	7	0.57
			trigonal bipyramid	5	0.41
			trigonal prism	4	0.33
			single-capped trigonal prism	1	0.08

Some other pairs could not be considered if they seem chemically or physically unreasonable. Thus, the correlation *Cn-Cp* seems unreasonable if the metal atom (*Me*) is not considered; no chemistry is contained in this pair. The remaining nine pairs are useful for analysis (Table S4); however, the human expert could say that he knows the result in many cases, for example he is aware of typical coordination numbers of metal atoms, *i.e.* of the *Me-Cn* correlation.

The computer (a prototype of the expert system) as usual will work more straightforward and this is both its advantage and its disadvantage. If we consider creation of the knowledge database from scratch, the computer knows nothing about reasonability of the distributions. This means that it will build all possible distributions. However, it will immediately understand which of them are trivial – they will have the 100% correlation, like *Cp-Cn*. Moreover, as opposite to the human, the computer will analyze seemingly uninteresting correlations that nonetheless can implicitly contain important knowledge. All found correlations fitting the support (*S*) and confidence (*C*) parameters will be written to the universal storage and will form the experience base of the expert system.

In our example, we use *S*=0.01 (1% level), *i.e.* any subsample used for searching correlations must be larger than 1% of the whole sample, and *C*=3, *i.e.* the correlations are written in the knowledge database if they are found for at least three bins. Totally there are 1232 non-equivalent metal atoms in the 811 cyanides, but we have considered only 1137 atoms (Table 4) that corresponds to the chosen values of *S* and *C*; the lowest level fitting the support value corresponds to Mo with an occurrence of 1.2%. All subsamples in columns *Cn* and *Cp* have $N \geq 3$ bins that corresponds to the confidence value. Complete data with all exceptions excluded from Tables 4-6 are given in the ESI.

Table 4 shows just two of 60 possible correlations: *Me-Cn* and *Me-Cn-CCp*, but even from this information many rules can be derived. For example, in general, metal atoms prefer a C-end coordination of CN^- ligands with a 90%-probability, but the *Me* oxidation state can strongly influence this correlation. In particular, bivalent manganese with coordination number six has an N-environment, but for atoms of trivalent manganese, C-environment is more typical. This information can be written down to the knowledge database as the following rules:

IF (*Me*='Mn') AND (*OxS*=2) AND (*Cn*=6) THEN *CCp*=’N6’ with *P*=71.9%

IF (*Me*='Mn') AND (*OxS*=3) AND (*Cn*=6) THEN *CCp*=’C6’ with *P*=90.0%

To reveal other correlations a lot of tables like Table 4 should be built and analyzed (see Table S4). Of course, only computer can do such a tedious job!

METAL ATOM AND LIGAND COORDINATION MODES. Now we extend our knowledge beyond the coordination shell of metal atom and consider mutual coordination of metal atoms and ligands. New descriptors are required for this purpose; some of them were mentioned above (see also Table 2). Coordination properties of CN^- ligand (Fig. 3) are restricted with a narrow range of possible modes: M^1 (561 from 811 complexes, 63.2

%), B^2 (303, 34.2 %), B^3 (21, 2.4 %), B^4 (1 compound), and M^2 (1 compound). Combination of most common central atoms *Me* (Fe, Au, Pt, Cu) and CN^- ligand coordination modes (M^1, B^2) predetermine the abundance of coordination formulas AM^1_6 (206 complexes, 25.4 %), AM^1_4 (183, 22.6 %), AM^1_2 (98, 12.1 %), AB^2_3 (87, 10.7 %), and AB^2_2 (51, 6.3 %). These formulas present the most well-studied molecular ($[Me(CN)_6]^{n-}$, $[Me(CN)_4]^{n-}$, $[Me(CN)_2]^{n-}$) and polymer cyanides, among which the Prussian blue analogues[23] and the diamonded-like frameworks[24] are most typical.

Table 4. Correlations between type of the metal atom (*Me*), its oxidation state (*OxS*), coordination number (*Cn*), and composition of its coordination polyhedron (*CCp*) (*S*=0.01, *C*=3)

Me	*OxS*	*Cn*	*CCp*	*N*	*w*, %
Fe	3	6	C_6	128	95.5
			N_6	6	4.5
	2	6	C_6	70	97.2
Pt	2	4	C_4	119	100
Au	1	2	C_2	87	97.8
	3	4	C_4	29	100
Ag	1	2	C_2	80	93.0
			CN	3	3.5
		3	C_2N	3	60.0
Co	3	6	C_6	67	98.5
	2	6	N_6	8	72.7
		4	N_4	3	60.0
Cd	2	4	C_4	66	86.8
			N_4	5	6.6
			C_3N	4	5.3
		6	N_6	40	97.6
		5	C_3N_2	5	100
Ni	2	4	C_4	49	100
		6	N_6	16	100
	1	4	C_3Ni	6	100
Cr	3	6	C_6	35	100
Cu	1	3	C_2N	34	46.6
			CN_2	19	26.0
			C_3	13	17.8
			N_3	7	9.6
Cu	1	4	C_4	17	32.7
			C_2N_2	12	23.1
			C_3N	10	19.2
			CN_3	8	13.5
			N_4	5	9.6
		2	C_2	7	53.8
			CN	6	46.2
	2	6	N_6	14	93.3
Zn	2	4	N_4	25	67.6
			C_4	12	32.4
		6	N_6	3	75.0
Mn	2	6	N_6	23	71.9
			C_6	9	28.1
		4	C_4	3	60.0
	3	6	C_6	9	90.0
Pd	2	4	C_4	20	100
W	4	8	C_8	11	100
	5	8	C_8	8	100
Hg	2	2	C_2	11	100
		4	C_4	7	100
Mo	4	8	C_8	10	100
	5	8	C_8	5	100

Obviously, the local topology of a coordination compound $[Me_x(CN)_y]^{z-}$ depends on coordination modes of the metal atom and the ligand as well as on the ratio *r*= *y*/*x*. Table 5 collects the values of these descriptors for the 209 three-periodic cyanides. There is a number of strong correlations *r*-*CF*; in most cases, the ratio ligand/metal predetermines the local coordination (coordination formula, *CF*) with a high probability.

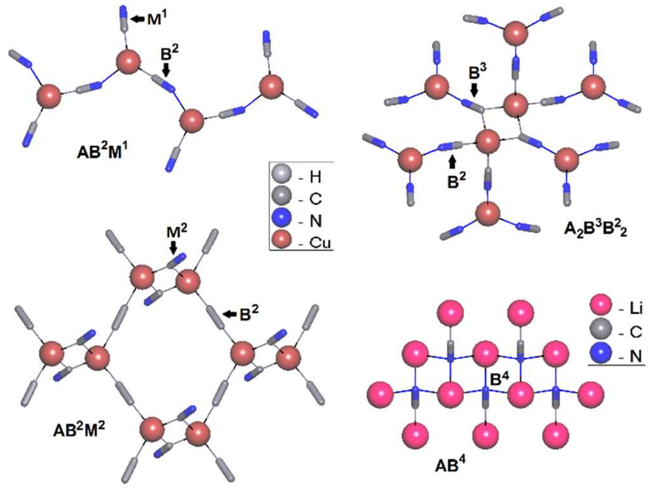


Figure 3. Coordination types of the cyanide group and coordination formulas of the networks: M^1 and B^2 in the chain structure of $[(ptp)Cu(CN)_3]$ (ptp = 2,2'-(1,4-phenylene)(bis-3,4,5,6-tetrahydropyrimidin-1-ium)) (CSD code CEXWAF)[22a] (top left); B^2 and B^3 in the 3D structure of $[Cu^I_2Cu^{II}_2(CN)_6]\cdot(TEA)_2$ (TEA = triethylamine) (CSD code EDAPEG)[22b] (top right); B^2 and M^2 in the layered structure of $[(pami)Cu(CN)_4]$ (pami = (1,4-phenylene)bis(aminomethaniminium)) (CSD code CEXVIM)[22a] (bottom left), B^2 and B^4 in the 3D structure of Li(CN) (ICSD code 77321)[22c] (bottom right).

OVERALL TOPOLOGY OF COORDINATION NETWORK. Further, important correlations can be found on the next level of the structure organization that is the underlying topology (*UT*) of coordination network. Recently we revealed many strong relationships, both direct *CF*-*UT* and reverse *UT*-*CF*, between local and overall topologies in coordination networks[15a] and molecular crystals.[15c] In Tables 5 and 6, these two correlation types are considered for the cyanide sample and the following significant rules can be established (Fig. 4):

IF *CF*='AB^2_3' THEN *UT*='**pcu**' with *P*=97.7%;
IF *UT*='**pcu**' THEN *CF*='AB^2_3' with *P*=91.2%;
IF *CF*='AB^2_2' THEN *UT*='**dia**' with *P*=82.6%;
IF *UT*='**dia**' THEN *CF*='AB^2_2' with *P*=95.0%.
IF *CF*='$A_5B^2_{12}$' THEN *UT*='**cor**' with *P*=58.3%;
IF *UT*='**cor**' THEN *CFs*='$A_5B^2_{12}$' with *P*=100%.

An ‘unusual’ *r*=M:CN ratio 5:12 in the last two rules can be easily explained if one considers the assembling of the corundum (Al_2O_3) **cor**-type topology from octahedral AB^2_3 and tetrahedral AB^2_2 units in the ‘corundum’ (2:3) ratio: $A_5B^2_{12} = 2AB^2_3 + 3AB^2_2$ (Fig. 4).

There are many other rules that can be derived from Tables 5 and 6 but their confidence is much less because they are based on much smaller samples.

The shape of building unit can influence the overall topology even if the structural units have the same coordination numbers, but differ by the coordination figures. A well-known example[12] is the pair of underlying nets **pts** and **dia** with a 4-coordination and an AB_2^2 coordination formula (Table 6). The coordination figures of the **dia** and **pts** highest-symmetry embeddings are tetrahedron and a combination of tetrahedron and square, respectively.

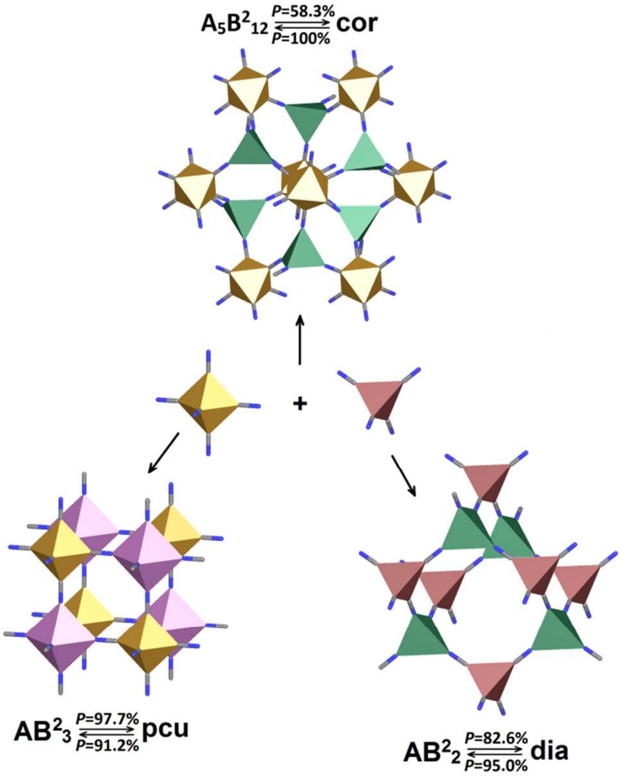


Figure 4. Fragments of the coordination networks with underlying topology **pcu** (left), **dia** (right) and **cor** the corresponding local coordination $AB^2{}_3$ and $AB^2{}_2$ in the crystal structures of $MnCs_2[Fe(CN)_6]$ (ICSD code 151695),[25a] $(NMe_4)[CuZn(CN)_4]$ (CSD code KEGZOL),[25b] and $Zn_3[Fe(CN)_6]_2$ (ICSD code 157850),[25c] respectively.

The coordination and shape of the building unit can be tailored to provide a particular overall topology; this is the subject of reticular design.[12,26] The expert system can separate such units using topological criteria[21] and recognize a particular unit in a network of any complexity[27] to deposit the information about the unit occurrence in the knowledge database. Thus, polynuclear complex groups were determined only in 19 cyanides. In most cases (15), they are dimers $\{Cu_2C_2\}$, which have an octahedral coordination.

TOPOLOGY OF ENTANGLED MOTIFS. The next level of the structure organization of coordination polymers is entanglement of several coordination networks. The entanglement phenomenon being quite rare even in the 1990s, now becomes ordinary.[21,28] There is a clear correlation between the overall topology of the coordination network and its ability to interweave. In particular, such underlying topologies as **dia** or **srs** are suitable for interpenetration, while many others never form entangled arrays.[21] Another common regularity is existence of elongated bridging structural units; in the cyanides, those are linear complex groups like $[Ag(CN)_2]^-$ or $[Au(CN)_2]^-$. Such information can also be stored in the knowledge database.

INFLUENCE OF EXTRAFRAMEWORK SPECIES. A great majority (210 out of 253) of 2-periodic or 3-periodic polymeric cyanides contain extraframework cations, solvate or clathrate molecules. Obviously, they influence the network topology; in particular, they promote unusual underlying topologies that become exceptions in the corresponding distributions. For example, six cyanide structures[29] with r = 3:2 exhibit rare 3-periodic network topologies 3,3,3,3,4T11, 3,3,4T101, 3,3,4T12, 3,3,4T20, 3,3,4T36, and 3,3,4T72 instead of more common for this ratio 2-periodic **hcb** topology (Table S1). All these structures contain large extraframework cations like $(i\text{-Pr})_2NH_2^+$ or $(n\text{-Bt})_4N^+$.

Table 5. Relations between ligand/metal ratio (*r*), coordination formula (*CF*), and underlying topology (*UT*) for the 209 three-periodic cyanides (*S*=0.01, *C*=1). Hereafter the RCSR bold three-letter symbols[11c] and TOPOS *NDk-n* symbols[15b] are used to designate the underlying topology

r	*N*	*w*, %	*CF*	*N*	*w*, %	*UT*	*N*	*w*, %
3	87	41.2	$AB^2{}_3$	86	98.9	**pcu**	84	97.7
						nia	2	2.3
			$AB^2{}_2M$	1	1.15	**pts**	1	100
2	53	25.1	$AB^2{}_2$	46	86.8	**dia**	38	82.6
						pts	6	13.0
						lon	2	4.3
			$AB^3{}_2$	7	13.2	3,6,6T20	7	100
1.5	18	8.5	$A_2B^2{}_3$	10	55.6	**pcu**	7	70.0
						bcs	1	10.0
						noe	1	10.0
						3,3,3,3,3T1	1	10.0
			$A_2B^3B^2{}_2$	5	27.8	3,3,4T101	1	20.0
						3,3,4T12	1	20.0
						3,3,4T20	1	20.0
						3,3,4T36	1	20.0
						3,3,4T72	1	20.0
			$A_4B^3B^2{}_5$	1	5.56	3,3,3,3,4T11	1	100
2.4	12	5.7	$A_5B^2{}_{12}$	12	100	**cor**	7	58.3
						4,5,6T1	3	25.0
						4,5,6T2	2	16.7
2.33	21	100	$A_3B^2{}_7$	21	100	**fsk**	5	23.8
						xat	5	23.8
						4,6T52	3	14.3
						4,4,6T62	3	14.3
						fsh	3	14.3
						fsn	1	4.8
						fsm	1	4.8
1.33	11	5.2	$A_3B^2{}_4$	11	100	**qtz**	4	36.4
						4,4T44	3	27.3
						dia	2	18.2
						mok	1	9.1
						ths	1	9.1
1.75	4	1.9	$A_4B^3{}_2B^2{}_5$	4	100	3,4,4T105	2	50.0
						3,3,4,4,4,4T10	1	25.0
						3,4,4T78	1	25.0
1.67	3	1.4	$A_3B^3{}_2B^2{}_3$	1	33.3	3,3,4,4,4T16	1	100
			$A_3B^3B^2{}_4$	1	33.3	3,3,4,4T80	1	100
			$A_3B^2{}_5$	1	33.3	3,3,3,3,3,3,4,4,4T2	1	100

Voluminous extraframework species preclude entanglement, as a rule. In particular, all 30 cyanides *Me*(CN)$_2$ (*Me* = Cd, Cu, Li, Hg, Zn) that contain ions or molecules like $(CH_3)_4N^+$, CCl_4, C_6H_6 and possess the **dia** underlying topology, are non-interpenetrated. To store such correlations in the knowledge database one has to estimate the size of the molecules that can be done with molecular Voronoi polyhedra in an automated mode.[30]

Table 6. Correlations between underlying topology (*UT*) and coordination formula (*CF*) (*S*=0.01, *C*=3)

UT	*N*	*w*, %	*CF*	*N*	*w*, %
pcu	91	43.1	AB^2_3	84	92.3
			$A_2B^2_3$	7	7.6
dia	40	19.0	AB^2_2	38	95.0
3,6,6T20	7	3.3	AB^3_2	7	100
cor	7	3.3	$A_5B^2_{12}$	7	100
pts	7	3.3	AB^2_2	6	85.7
fsk	5	2.4	$A_3B^2_7$	5	100
xat	5	2.4	$A_3B^2_7$	5	100
qtz	4	1.9	$A_3B^2_4$	4	100
4,4T44	3	1.4	$A_3B^2_4$	3	100
4,6T52	3	1.4	$A_3B^2_7$	3	100
4,5,6T1	3	1.4	$A_5B^2_{12}$	3	100
4,4,6T62	3	1.4	$A_3B^2_7$	3	100
fsh	3	1.4	$A_3B^2_7$	3	100

How the inference machine can work: examples

Having the knowledge database created at the previous steps and completely represented in the ESI we can envisage the work of the inference machine. Although such a tool still has to be developed, the possible results and restrictions can already be outlined.

As was recently discussed,[15a] an important application of the inference machine is prediction of the topology of coordination network. For example, the knowledge database (see the ESI) contains two topological types, **hcb** and **pcu**, as the most suitable choices for the coordination formula $A_2B^2_3$, and three topological types, **bcs**, **noe**, and 3,3,3,3,3T1, as exceptions (Table S1, Fig. 5).

The local coordination in the **hcb** underlying net is characterized by a trigonal coordination figure so to get this topology with a 90.0% probability the inference machine would recommend $[Me(CN)_3]^{n-}$ (*Me*=Cu, Ag) structural units with a trigonal geometry (Fig. 6). To obtain a **pcu** underlying net the best choice is to use the octahedral building units $[Me(CN)_6]^{n-}$ in a combination with 2-coordinated Au^+ or Ag^+ ions, or, *vice versa*, the linear spacers $[Ag(CN)_2]^-$ or $[Au(CN)_2]^-$ in a combination with octahedron-coordinated Me^{n+} ions. This scheme predicts the required result with a probability 87.5%.

Another example demonstrates how the expert system can restrict the set of possible local and overall topologies of the coordination polymer (Fig. 7). If we propose Cd as the complexing atom and CN^- as the ligand, the expert system would conclude that most likely the $[Cd(CN)_r]^{z-}$ complex will be obtained with *r*= 2, 3, 2.33, 2.4, or 4. Depending on the stoichiometry (*r*) and coordination numbers 4, 5, or 6, the corresponding tetrahedral, trigonal bipyramidal, or octahedral coordination polyhedra can be realized, and their combinations are strongly predetermined by *r*. Only a few types of the Cd coordination environment (*CCp*) are characteristic: pure C_4, N_4, N_6, and mixed, usually disordered, C_3N, C_3N_2, and C_2N_3. Some framework types can be assembled only with participation of other metal atoms that provide the C_6-type of environment being unusual for cadmium. Only three ligand coordination modes, M^1, B^2, B^3, occur in cadmium cyanides, and depending on *r* and *Cn* only six coordination modes are possible (Fig. 7). Out of them, four (AB_2, AB^2_3, AB^3_2, and AM^1_4) strongly predetermine the underlying topology (**dia**, **pcu**, 3,6,6T20, and 1,4M5-1, respectively). Two remaining modes ($A_3B^2_7$ and $A_5B^2_{12}$) allow some diversity of the overall topologies depending on the synthesis conditions (counterion type, solvent type, clathrate molecules, thermodynamic parameters). Interpenetration can be realized only in the networks with the **dia** topology if extraframework cations, solvent or clathrate molecules are absent.

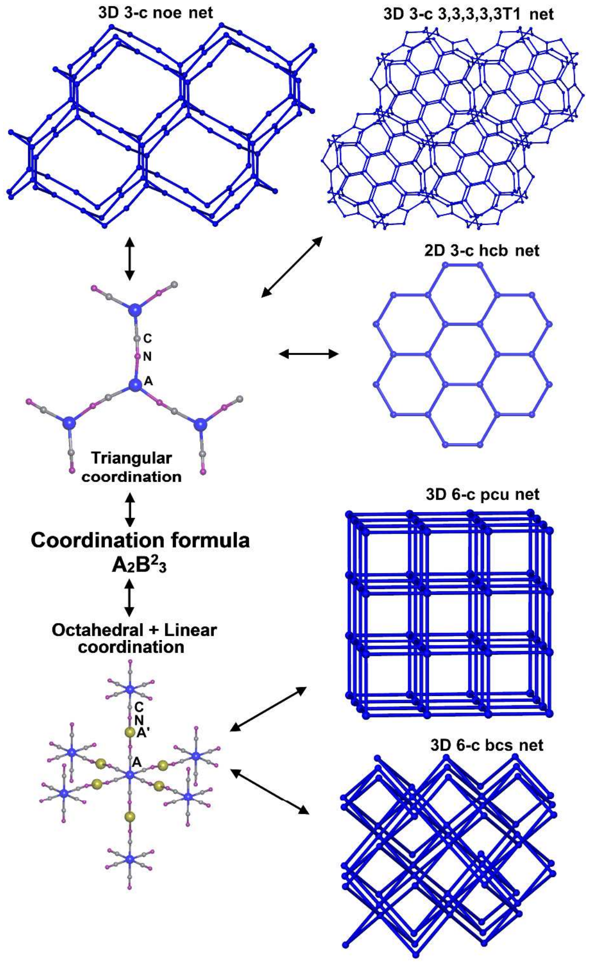


Figure 5. Relations between coordination figure and underlying net topology for cyano-complexes with coordination formula $A_2B^2_3$.

Using the data from Fig. 7 one can estimate the total probability (*P*) of obtaining a particular overall topology as well as other parameters of the cyanide complexes by simple multiplying the corresponding conditional probabilities. For example, the **pcu** topology can be obtained for a $[Cd(CN)_r]^{(r-2)-}$ complex with a probability $P = 0.103\cdot0.857\cdot1\cdot1\cdot1\cdot1 = 0.088$; the **dia** topology appears in two branches of the scheme (Fig. 7) so the total probability consists of two summands: $P = (0.441\cdot0.800\cdot0.958\cdot1\cdot1\cdot0.957) + (0.441\cdot0.100\cdot1\cdot0.667\cdot1\cdot1) = 0.323 + 0.029 = 0.352$. However, if the user specifies additional conditions, except metal atom (Cd) and ligand (CN), the conclusion can be done with a much higher probability. For example, if $r = 2$ is specified, *i.e.* only the $[Cd(CN)_2]$ complexes are considered, then the **dia** topology is expected with $P = (0.800\cdot0.958\cdot1\cdot1\cdot0.957) + (0.100\cdot1\cdot0.667\cdot1\cdot1) = 0.733 + 0.067 = 0.800$.

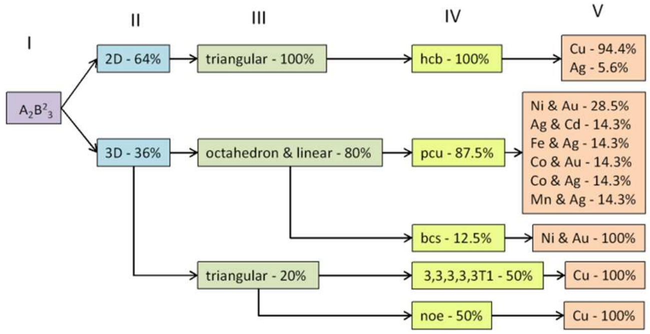


Figure 6. Prediction scheme provided by the inference machine for the cyanide complexes with the coordination formula $A_2B^2_3$ specified by the user at the first (I) step. The subsequent steps contain the prediction results for the dimensionality (II), coordination figure (III), overall topology (IV) and metal coordination centers (V) of possible coordination polymers. The probabilities of realization of the corresponding parameters are given in percentage.

One also can construct a similar scheme for the whole list of the cyanide complexes with any central atom, and propose some topologies for design that have not yet been found for cadmium cyanides. For example, the 4,6-coordinated **cor** topology being abundant for Zn, Fe-cyanides (Table S8), may be suitable also for cadmium complexes with appropriate extraframework templating molecules.

Conclusion remarks

Expert systems inevitably appear at a particular stage of science development, when the experimental information becomes abundant and versatile enough to derive reliable and predictable rules. Huge amount of the information invokes electronic databases like the CSD or ICSD, which should be transformed to knowledge databases with time. Crystal engineering stands at this point and the trends of the past 10-15 years let us hope that first full-fledged knowledge databases will be created soon. The expert system will use correlations stored in the knowledge database in combination with *ab initio* modelling methods to predict possible structure motifs and to facilitate design of new materials. To develop such new tools, novel approaches, models, and methods are required. We believe that formalized descriptors and strict computer algorithms will gain a crucial role in exploration of crystal structures; development of these tools promises to be one of the main trends in crystal chemistry and design of the near future.

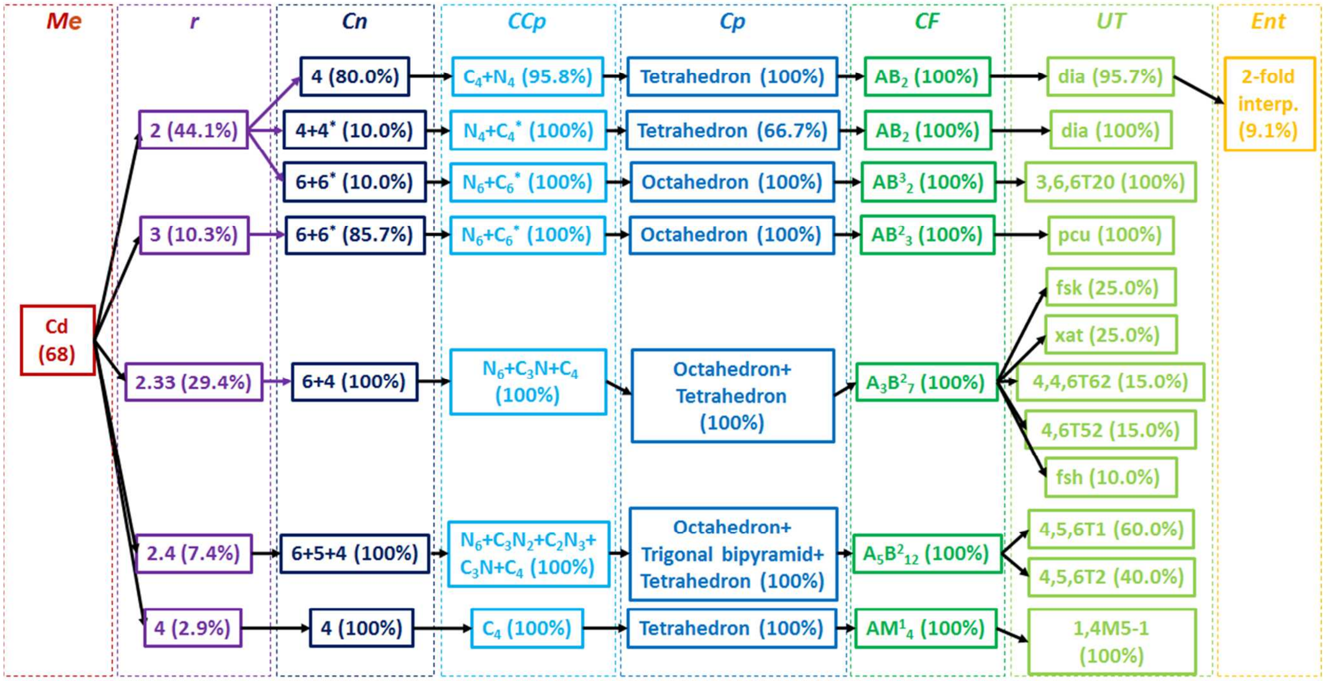


Figure 7. Prediction scheme provided by the inference machine with *S*=0.01 and *C*=2 for 68 compounds containing $[Cd(CN)_r]^{x-}$ complexes with different CN:Cd ratio (*r*), coordination number (*Cn*), coordination polyhedra composition (*CCp*), coordination polyhedra (*Cp*), coordination formula (*CF*), underlying topology (*UT*) and type of entanglement (*Ent*). Additional coordination marked by "*" in the *Cn* and *CCp* columns corresponds to non-cadmium metals (Cu, Fe, Os, Pd, Rh, Ru, Zn). The conditional probabilities are given in parentheses.

Acknowledgements

E.V.A., A.P.S. and V.A.B. thank Russian government (grant No. 14.B25.31.0005) and Russian Foundation for Basic Research (grant No. 13-07-00001) for support as well as Prof. Davide M. Proserpio for helpful discussions.

Notes and references

[a] Samara Center for Theoretical Materials Science (SCTMS), Samara State University, Ac. Pavlov St. 1, Samara 443011, Russia; E-mail: blatov@samsu.ru.

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

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

† Electronic Supplementary Information (ESI) available: The knowledge database as an Excel file containing initial data and descriptor values (parameters) for 811 cyanide crystal structures (Tables S1, S2), one-dimensional distributions of the parameters (Table S3), pairwise correlations between the parameters (Table S4) as well as multiple correlations (Tables S5-S8). See DOI: 10.1039/b000000x/

1 (a) L. F. Pau, J. Motiwalla, Y. H. Pao and H. H. Teh, *Expert Systems in Economics, Banking, and Management*, North-Holland, Amsterdam, 1989; (b) S.-H. Liao, *Expert Syst. Appl.*, 2005, **28**, 93.

2. (a) P. Jackson, *Introduction to Expert Systems*, 3rd ed., Addison-Wesley, 1999; (b) M. Schmitt, H.-N. Teodorescu, Ashlesha Jain, Ajita Jain and S. Jain, *Computational Intelligence Processing in Medical Diagnosis*, Springer-Verlag, 2002.

3. (a) M. C. Hemmer. *Expert Systems in Chemistry Research*, CRC Press; Boca Raton, FL: 2008; (b) P. Judson, *Knowledge-Based Expert Systems in Chemistry: Not Counting on Computers*, RSC Publishing, Cambridge, UK, 2009.

4. (a) F. H. Allen, *Acta Crystallogr.*, 2002, **B58**, 380; (b) A. Belsky, M. Hellenbrandt, V. L. Karen and P. Luksch, *Acta Crystallogr.*, 2002, **B58**, 364; (c) J. Chisholm, E. Pidcock, J. van de Streek, L. Infantes, W. D. S. Motherwell and F. H. Allen, *CrystEngComm*, 2006, **8**, 11; (d) S.P. Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder and K.A. Persson, *Comp. Mat. Sci.*, 2015, **97**, 209; (e) P. T. A. Galek, F. H. Allen, L. Fábián and N. Feeder, *CrystEngComm*, 2009, **11**, 2634; (f) P. A. Wood, N. Feeder, M. Furlow, P. T. A. Galek, C. R. Groom and E. Pidcock, *CrystEngComm*, 2014, **16**, 5839; (g) M. Klintenberg and O. Eriksson, *Comp. Mat. Sci.*, 2013, **67**, 282; (h) J. G. P. Wicker and R. I. Cooper, *CrystEngComm*, 2015, DOI: 10.1039/C4CE01912A; (i) S. Curtarolo, G. L. W. Hart, M. Buongiorno-Nardelli, N. Mingo, S. Sanvito and O. Levy, *Nature Mater.*, 2013, **12**, 191; (j) O. Isayev, D. Fourches, E. N. Muratov, C. Oses, K. Rasch, A. Tropsha and S. Curtarolo, *Chem. Mater.*, 2015, DOI: 10.1021/cm503507h.

5. (a) A. R. Oganov and C. W. Glass, *J. Chem. Phys.*, 2006, **124**, 244704; (b) Q. Zhu, V. Sharma, A. R. Oganov and R. Ramprasad, *J. Chem. Phys.* 2014, **141**, 154102; (c) G. Ceder, D. Morgan, C. Fischer, K. Tibbetts and S. Curtarolo, *MRS Bull.*, 2006, **31**, 981.

6. W. D. S. Motherwell, *Cryst. Rev.*, 2008, **14**, 2, 97.

7. (a) A. F. Wells, *Acta Cryst*., 1954, **7**, 535-544; (b) A. F. Wells, *Three-Dimensional Nets and Polyhedra*, New York, Interscience, 1977; (c) A. F. Wells, *Further Studies of Three-dimensional Nets*, ACA monograph 8, 1979.

8. V. A. Blatov, D. M. Proserpio and M. O'Keeffe, *CrystEngComm*, 2010, **12**, 44.

9. A. I. Kitaigorodskii, *Molecular Crystals and Molecules*, Academic Press, New York, 1973.

10. M. O'Keeffe and B. G. Hyde, *Struct. Bond*. 1985, **61**, 77.

11. (a) O. Delgado-Friedrichs and M. O'Keeffe, *Acta Cryst.*, 2003, **A59**, 351; (b) V. A. Blatov, L. Carlucci, G. Ciani and D. M. Proserpio, *CrystEngComm*, 2004, **6**, 377; (c) M. O'Keeffe, M. A. Peskov, S. J. Ramsden and O. M. Yaghi, *Acc. Chem. Res.*, 2008, **41**, 1782; (d) E. V. Alexandrov, V. A. Blatov and D. M. Proserpio, *Acta Crystallogr.*, 2012, **A68**, 484.

12. N. W. Ockwig, O. Delgado-Friedrichs, M. O'Keeffe and O. M. Yaghi, *Acc. Chem. Res.*, 2005, **38**, 176.

13. I. A. Baburin and V. A. Blatov, *Acta Cryst.*, 2007, **B63**, 791.

14. V. A. Blatov, *Struct. Bond*., 2011, **138**, 31.

15. (a) T. G. Mitina and V. A. Blatov, *Cryst. Growth Des*., 2013, **13**, 1655; (b) F. Aman, A. M. Asiri, W. A. Siddiqui, M. N. Arshad, A. Ashraf, N. S. Zakharov and V. A. Blatov, *CrystEngComm*, 2014, **16**, 1963; (c) P. N. Zolotarev, M. N. Arshad, A. M. Asiri, Z. M. Al-amshany and V. A. Blatov, *Cryst. Growth Des*., 2014, **14**, 1938; (d) O. A. Blatova, A. M. Asiri, Z. M. Al-amshany, M. N. Arshad and V. A. Blatov, *New J. Chem*., 2014, **38**, 4099.

16. (a) O. Delgado-Friedrichs, M. O'Keeffe and O. M. Yaghi, *Phys. Chem. Chem. Phys*., 2007, **9**, 1035; (b) M. O'Keeffe, *Phys. Chem. Chem. Phys*., 2010, **12**, 8580.

17. S. P. Westrip, *J. Appl. Cryst.*, 2010, **43**, 920; http://checkcif.iucr.org/

18. V. A. Blatov, A. P. Shevchenko and D. M. Proserpio, *Cryst. Growth Des.*, 2014, **14**, 3576; http://topospro.com.

19. (a) M. Avila, L. Reguera, J. Rodriguez-Hernandez, J. Balmaseda and E. Reguera, *J. Solid State Chem.*, 2008, **181**, 2899; (b) H. Tokoro and S. Ohkoshi, *J. Chem. Soc.*, *Dalton Trans.,* 2011, **40**, 6825; (c) A. B. Cairns and A. L. Goodwin, *Chem. Soc. Rev.*, 2013, **42**, 4881.

20. M. J. Katz, T. Ramnial, H.-Z. Yu and D. B. Leznoff, *J. Am. Chem. Soc.*, 2008, **130**, 32, 10662.

21. E. V. Alexandrov, V. A. Blatov, A. V. Kochetkov and D. M. Proserpio, *CrystEngComm*, 2011, **13**, 3947.

22. (a) S. Ferlay, P. Dechambenoit, N. Kyritsakas and M. W. Hossein, *J. Chem. Soc.*, *Dalton Trans.*, 2013, **42**, 11661; (b) W.-Y. Jian, W. Li, Q.-Y. Lv, X. Min, Y.-Y. Liu and S.-Z. Zhan, *Synth. React. Inorg., Met.-Org., Nano-Met. Chem.*, 2012, **42**, 1375; (c) J. A. Lely and J. M. Bijvoet, *Rec. Tr. Chim. Pays-Bas Belg*., 1942, **61**, 244.

23. (a) K. R. Dunbar and R. A. Heintz, *Prog. Inorg. Chem.*, 1997, **45**, 283; (b) S. Tanase and J. Reedijk, *Coord. Chem. Rev.*, 2006, **250**, 2501; (c) F. Herren, P. Fischer, A. Ludi and W. Haelg, *Inorg. Chem.*, 1980, **19**, 4, 956.

24. (a) D. Williams, B. Pleune, K. Leinenweber and J. Kouvetakis, *J. Solid State Chem.*, 2001, **159**, 1, 244; (b) R. J. Williams, A. C. Larson and D. T. Cromer, *Acta Crystallogr.*, 1972, **B28**, 858; (c) T. Kitazawa, *J. Mater. Chem.*, 1998, **8**, 3, 671; (d) T. Kitazawa, S.-I. Nishikiori, R. Kuroda and T. Iwamoto, *J. Chem. Soc.*, *Dalton Trans.*, 1994, 1029; (e) B. Xianhui, T. E. Gier and G. D. Stucky, *Acta Crystallogr.*, 1996, **C52**, 14.

25. (a) R. Martinez-Garcia, E. Reguera, J. Rodriguez, J. Balmaseda and J. Roque, *Powder Diffr*., 2004, **19**, 255; (b) B. F. Hoskins and R. Robson, *J. Am. Chem. Soc.*, 1990, **112**, 1546. (c) J. Rodriguez-Hernandez, E. Reguera, E. Lima, J. Balmaseda, R. Martinez-Garcia and H. Yee-Madeira, *J. Phys. Chem. Solids*, 2007, **68**, 1630.

26. O. M. Yaghi, M. O'Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi and J. Kim, *Nature*, 2003, **423**, 705.

27. V. A. Blatov, *Struct. Chem*., 2012, **23**, 955.

28. (a) L. Öhrström and K. Larsson, *Molecule based Materials: The Structural Network Approach*, Elsevier, Amsterdam, 2005; (b) L. Carlucci, G. Ciani, D. M. Proserpio and N. Champness, in *Making Crystals by Design: Methods, Techniques and Applications*, eds. D. Braga and F. Grepioni, Wiley-VCH, Weinheim, 2007, ch. 1.3, ch. 2.4; (c) S. R. Batten, S. M. Neville and D. R. Turner, *Coordination Polymers: Design, Analysis and Application*, Royal Society of Chemistry, Cambridge, 2009.
29. (a) S. Ferlay, P. Dechambenoit, N. Kyritsakas and M. W. Hosseini, *J. Chem. Soc.*, *Dalton Trans.*, 2013, **42**, 11661; (b) W.-Y. Jian, W. Li, Q.-Y. Lv, X. Min, Y.-Y. Liu and S.-Z. Zhan, *Synth. React. Inorg., Met.-Org., Nano-Met. Chem.*, 2012, **42**, 1375; (c) T. Pretsch, I. Brudgam and H. Hartl, *Z. Anorg. Allg. Chem.*, 2004, **630**, 353; (d) Y.-L. Qin, J.-J. Hou, J. Lv and X.-M. Zhang, *Cryst. Growth Des.*, 2011, **11**, 3101.
30. I. A. Baburin and V. A. Blatov, *Acta Cryst.* 2004, **B60**, 447.

Eugeny V. Alexandrov *was born in 1986 in Chelno-Vershini, Samara region, Russia. He graduated from Samara State University in 2009 and defended his PhD thesis in 2014. His research is devoted to topological analysis of coordination polymers and other extended architectures as well as to applications of computer methods in the design of new materials.*

Alexander P. Shevchenko *graduated from Chemistry Department of Samara State University in 1991, and got a PhD degree in inorganic chemistry in 1998. His PhD thesis was devoted to applications of Voronoi polyhedra in crystal chemistry. His research concerned non-bonded interactions in crystals, computer simulation of atomic and molecular systems, as well as metal-organic compounds. Now he participates in the development of software for expert systems in chemistry and materials science.*

Abdullah M. Asiri *received PhD from University of Wales, College of Cardiff, UK in 1995. He is a Professor of Organic Photochemistry, the Head of the Chemistry Department at King Abdulaziz University since October 2009, the founder and the Director of the Center of Excellence for Advanced Materials Research. His research interest covers color chemistry, synthesis of novel photochromic, thermochromic systems, synthesis of novel coloring matters and dyeing of textiles, Materials Chemistry, Nanochemistry and Nanotechnology Polymers and plastics. He is the Editor-in-Chief of King Abdulaziz University Journal of Science and a member of the Editorial Board of Pigments and Resin Technology (UK), Organic Chemistry in Sight (New Zealand), Designed Monomers & Polymers and Journal of Single Molecule Research. He is the Vice-President of Saudi Chemical Society (Western Province Branch). He hold three USA patents, more than 600 Publications in international Journals, four book Chapters, and 10 Books.*

Vladislav A. Blatov *was born in 1965 in Samara (former Kuibyshev), Russia. He graduated from Samara State University (1987) and got degrees of Candidate (1991) and Doctor in Chemistry (1998) from the Institute of General and Inorganic Chemistry (Moscow). He is Full Professor at SSU Chemistry Department since 1998 and Director of Samara Center for Theoretical Materials Science (SCTMS) since 2013. His research and educational interests concern geometrical and topological methods in crystal chemistry and their* *computer implementation in the program package TOPOS being developed since 1989.*